Developing a Knowledge Management Approach to Support Managing Credit Risk in Jordanian Banks

Abdallah Al-Shawabkeh

A thesis submitted in partial fulfilment of the requirements of
University of Greenwich for the degree of
Doctor of Philosophy

School of Computing Science and Mathematics

Department of Information Systems and Digital Media

2010

Declaration:

I certify that this work has not been accepted in substance for any degree, and is not concurrently being submitted for any degree other than that of PhD in Information Systems and Digital Media being studied at the University of Greenwich. I also declare that this work is the result of my own investigations except where otherwise identified by references and that I have not plagiarised another's work.

Students: Abdallah Al-Shawabkeh

Signature:

Date :

First Supervisor: Dr Miltos Petridis

Signature:

Date :

Second Supervisor: Dr Aleksandar Stojanovic

Signature:

Date :

Table of Contents

List of Tables

List of Figures

Abstract

It is becoming increasing clear that; in banks; the sharing of knowledge amongst the senior executives has not been as effective as it should have been. The lack of knowledge amongst senior executives about the level of risks taken in sub-prime lending, the resulting 'toxic assets' and the global nature of the instruments used to spread risks is said to be the main contributing reason for the current worldwide crisis in banks. Banks in Jordan, the focus of this study, are not immune from the exposure to the risks.

The banking sector in Jordan is the most important in the Jordanian national economy and has effectively contributed to improving economic development through its important role in mobilising savings and channelling them into different fields of investment

Therefore, the overall aim of this research is to propose that developing a knowledge management (KM) approach to support managing credit risk will help banks in general, and Jordanian banks in particular, in improving the process of managing credit risk.

To reach the aim, several objectives have been constructed:

1. Reviewing current status of KM and its relationship with CRM
2. Developing a scale to measure KM behaviour and practices
3. Determining current KM status in Jordanian banks
4. Building a CR decision support system using internal implicit knowledge to reduce the rate of defaults.

As a result of this research, a KM approach has been developed to support managing credit risk. The approach contains the following steps: identify, measure, analyse, improve, and evaluate.

Using the new KM approach, the main conclusion of this research suggest that considering credit risk management and KM together gives a much stronger basis for banks to manage credit risk.

Keywords: Credit risk management, Knowledge management, Knowledge management measurement models, Credit scoring models, Logistic Regression Analysis and Data mining.

Acronyms

JOD	Jordanian Dinar
CBJ	Central Bank of Jordan
FMI	Financial Market International
MCA	Multi Criteria Analysis
CSF	Critical Success Factors
CRM	Credit Risk Management
KM	Knowledge Management
IT	Information Technology
CR	Credit Risk
CRK	Credit Risk Knowledge
CBJ	Central Bank of Jordan
DM	Data Mining
AI	Artificial Intelligent
LDA	Linear Discriminant Analysis
KMS	Knowledge Management System
IC	Intellectual Capital
HR	Human Resources
PR	Public Relations
SECI	Socialisation, Externalisation, Combination, and Internalisation
BSC	Balanced Score Card
CAS	Customer-banks-age
AGE	Age
OCC	Occupation
INC	Income
LAM	Loan amount
EDU	Education
NN	Neural Networks
PCC	Percentage Correctly Classified
PIC	Percentage Incorrectly classified
SPSS	Statistical Package for Social Sciences
DSS	Decision Support Systems
PCL	Percentage of bad loans to gross loans
KDD	Knowledge Discovery in Database
DW	Data Warehousing

Acknowledgement

I am greatly grateful to my supervisor, Dr Antony Tambyrajah, for his excellent guidance and critical comments, which were crucial in completing this thesis. In particular, he reviewed many previous drafts of this work with regular encouragement and long patience. He has set an example as a supervisor, a teacher, an academic and a person whose character I will aspire to follow. He has provided a continual source of intellectual stimulation and motivation, which will extend beyond this study.

I would like to express my deep gratitude to my supervisor, Dr. Miltos Petridis for his guidance, patience and generosity in giving considerable time, effort and attention throughout the progress of this thesis.

I am deeply indebted to my supervisor, Dr. Aleksandar Stojanovic for his contribution and academic help in completing the PhD.

Also I wish to thank Professor Keith Rennols for his support, assistance and encouragement and for his kind efforts and his great assistance in explaining some technical aspects of my research during the study.

My deep, sincere and great gratitude to my wife, Maram, for her support, encouragement, understanding, patience and the great responsibilities which she handled in looking after our children (Mesk and Ahmed) during my study in the university , without which I would never have come this far.

Special thanks for my father and mother. Without their encouragement and emotional support a lonely endeavour such as doctoral research would have been an impossible task.

Finally, I would like to thank my sisters and brothers, friends and colleagues for their support and encouragement.

May Allah bless us

Abdallah Al-Shawabkeh

Chapter 1: Introduction

1.1 Background

Developing countries have historically been more in control of the activities of their banks than industrialised countries (Gruben and McComb, 1997). However, the banking sector in many developing countries has been experiencing low levels of efficiency and subsequently poor performance, mainly due to heavy government intervention (Barth *et al*., 2004; Pasiouras *et al.,* 2007). An example of a banking system heavily subjected to government intervention and restrictions imposed by regulatory authorities is that of Jordan. Before the 1990s, the Jordanian banking industry served as an agent of the government, channelling investment funds to selected sectors under the country's economic development policy, while imposing many requirements and restrictions on banks' activities. For example, the Central Bank of Jordan (CBJ) determined lending limits for banks, set a ceiling on interest rates for loans and deposits, restricted entry to the Jordanian banking market and imposed a high reserve requirement ratio on banks and tight restrictions on foreign exchange transactions

In 1989, the Jordanian economy faced a major crisis when the Jordanian Dinar (JOD) suffered a major devaluation (i.e. the JOD lost 51% of its value against the US Dollar), total government debt reached 197% of gross domestic product (GDP), foreign currency reserves declined sharply, the inflation rate reached 25.6% and the budget deficit, excluding grants, reached 24% of GDP (Isik et al., 2004; Alissa, 2007). This crisis negatively affected the banking sector in Jordan and led to many problems such as an increase in the ratio of non-performing loans, which subsequently led to serious consequences for the banking industry.

Currently, banks in Jordan face difficulties in gaining competitive opportunities from their vast information and knowledge resources in managing credit risk (CR) and making decisions. Most of the problems, in fact, are internal to the banks themselves. In their report, Financial Market International (FMI) (2006) found that Jordanian banks are not effective lenders. The banks lend on the basis of collaterals and relationships and, besides, credit officers in Jordanian banks evaluate CR using their personal judgements (mop, 2007). While the international average for default rates is lower than 5% (CBJ, 2005), on the basis of the researcher's interviews the banking sector in Jordan estimated industry-wide non-performing loans at 11%, which indicates the critical problem that

Jordanian banks face and the need for finding effective solutions for managing credit risks.

1.2 Aim and Objectives

The growing need for an accurate assessment of credit risk inspires academicians and practitioners to introduce theoretical models for CR (Shirly *et al.*, 2009). In this context, the Basel Committee on Banking Supervision has decided to introduce a new capital adequacy framework which encourages the active involvement of banks in measuring the likelihood of defaults (Shirley *et al.*, 2009).

Therefore, the overall aim of this research is to propose that developing a knowledge management (KM) approach provides CR department in banks with new opportunities in the decision making process that support managing credit risk. In addition, considering CR and KM together gives a much stronger basis for banks to manage CR. This research accentuates the opportunity, which is to augment decision making using internal knowledge and advanced performance support systems.

The potential is a KM approach which makes managing CR and decision making processes more rational, instead of being dependent on experience and trial and error logic which carry great risks and may affect the adequacy of these processes.

To accomplish the aim, the following objectives have been constructed:

5. Reviewing current status of KM and its relationship with CRM
   - 1.1 Research and review into KM and CRM implementations in banks and its relationship with issues such as human resources, information technology, culture, benchmarking, knowledge structure, and organisational structure.
   - 1.2 Explore the impact of KM on the performance of organisations.
   - 1.3 Current problems in KM practice and methodologies will be identified.

6. Developing a scale to measure KM behaviour and practices
   - 2.1 Propose a preliminary (2-dimensional) matrix has been developed.
   - 2.2 This model may be used as benchmarks to evaluate the degree to which KM practices have been effectively implemented.
   - 2.3 To develop a number of indicators which relate to each criterion, and which may be used to monitor and model changes in KM adoption in Jordanian banks.

7. Determining current KM status in Jordanian banks

    3.1 Design and conduct a questionnaire survey of staff in Jordanian banks.

8. Building a CR decision support system using internal implicit knowledge to reduce the rate of defaults.

    Preliminary interviews with banks managers suggest that one of the most important problems relates to risk management in credit-loans. According to Shirley *et al.* (2009), the ability to measure CR clearly has the potential to greatly improve banks' risk management capabilities. Therefore, the question of whether knowledge-based systems in the particular area of credit-loans can lead to improvements in decision making will be investigated using 2755 cases provided by banks in Jordan.

Despite the vast literature on KM roles in improving business performance in the banking industry of the United States and Europe, and the rising empirical research in the context of developing countries (for example, see Berger and Humphrey, 1997; Goddard et al., 2001), very few studies have been made on KM efficiency in the banking industry in Jordan. One study tested KM status in Jordanian banks (Al-Omary, 2004), but it should however be noted that this study does not link KM to credit risk management (CRM). The study identifies problems in managing KM in Jordanian banks with no suggestions to overcome them. Therefore this current research is important for several reasons:

First, this study will be the first

(1) To build a KM approach to support CRM in Jordanian banks,

(2) To study the status of KM activities in CRM departments in Jordanian banks,

(3) To build a KM measurement models using KM critical success factors,

(4) To identify gaps in managing knowledge in CRM departments in Jordanian banks,

(5) To examine the variables that influence the risk of loan default in Jordanian banks, and,

(6) To build a credit classification system using internal implicit knowledge from Jordanian banks.

Secondly, by developing a KM approach to support managing CR, an area on in which there is limited empirical evidence, this study will complement the existing KM international banking literature, which is currently significantly skewed towards developed countries.

Thirdly, one of the main objectives of the CBJ is to ensure the safety and soundness of the banking system. This study will contribute to the CBJ efforts to improve the overall performance of the banking sector and identify the causes of inefficiencies in managing CR.

Finally, the study will highlight some of the positive determinants of efficiency in managing knowledge in CR that could benefit bank owners and managers in improving the level of efficiency in their organisations.

This research seeks to answer the following questions:

1. Are Jordanian banks efficient in managing credit risk?
2. Are Jordanian banks efficient in managing knowledge in CRM departments?
    2.1 What is the contribution of the 'People & Culture' factor in supporting KM in banks in Jordan?
    2.2 What is the contribution of the 'Processes' factor in supporting KM in banks in Jordan?
    2.3 What is the contribution of the 'IT' factor in supporting KM in banks in Jordan?
3. What are the enablers of KM in Jordanian banks?
4. What is the relationship between KM and CRM?
5. What are the variables that influence the risk of loan default in Jordanian banks?
6. What is the impact of using transformation for variables on the accuracy of the produced model?
7. What are the implications of the suggested model using internal implicit knowledge on managing CR in Jordanian banks?

Based on the stated purpose and the questions mentioned above, the following hypotheses are formulated (corresponding null hypotheses are the converse):

H1. There is a significant impact of the factors 'Processes and IT' on the factor 'People and Culture'

H2. There is a significant difference between the Jordanian commercial banks and Islamic banks in the performance of KM practices

H3. There is a significant relationship between KM and CRM in Jordanian banks

1.3 Research Methods

1.3.1 Overview

As outlined in Chapter 2, various KM approaches and scale models have been used in the empirical literature to implement and measure KM efficiency. Following the studies reviewed in Chapter 2; using Jordan as a case study; the following Chapters (2-8) have implemented the suggested KM approach as a research process.

As one of the objectives of this study is to investigate the efficiency of Jordanian banks in managing knowledge in CRM departments, it employs the questionnaire method as a quantitative approach. Using the suggested KM scale model (see Chapter 4), the researcher decided to obtain the view of credit officers and managers in Jordanian banks (see Chapter 5).

All Jordanian banks were included in the survey setting. The study concentrates on studying the Jordanian banks only. According to the CBJ (2007) report, in Jordan there are fourteen local commercial and two Islamic banks. The sixteen banks were approached by the researcher and twelve agreed to be part of this study, which represents 75% of the banking community. The sample is comprised of 242 respondents from twelve banks within CRM departments in Jordanian banks. The participants answered a survey questionnaire structured in Likert format. Data gathered from this research instrument were then computed for analysis and interpretation. Along with primary data, secondary sources; in the form of published articles and literatures to support the survey results have been used (see Chapter five).

Second, in order to answer research questions 5, 6 and 7, CRISP-DM methodology has been used to analyse historical data of previous retail loans (2755 cases) (see Chapter 6 and 7).

1.3.2 Data

Two types of data have been collected:

1. The author handed 320 questionnaires to the various head offices and then the branch managers forwarded the questionnaires to targeted officers. The questionnaires were distributed to officers involved in CRM, namely branch managers, senior CRM officers and credit officers. The number of questionnaires to be distributed in each bank was determined according to the banks' wishes after meeting the authorised people in the HR department. From the 300 returned responses, 68 questionnaires were excluded from the study because of missing data. The remaining 242 responses represent an effective response rate of 75.6 % (see Chapter 5).
2. As reviewed in Chapters 6, building a credit classification system requires historical data on previous customers in Jordanian banks. Historical data of previous retail loans (2755 cases) were collected from three Jordanian banks which dominate 55% of the total assets of all banks in Jordan (CBJ, 2006): the Arab Bank, the Housing Bank and the Jordan Bank.

1.4 Thesis Structure

The thesis is structured as follows:

Chapter 1: Introduction

This Chapter starts with the research background. Then a discussion of the significance of the study, aims and objectives, and benefits of the research will be presented.

Chapter 2: Literature Review

This Chapter introduces the concepts of CRM and its status in Jordanian banks. Then, an overview of topics needed to understand KM: different perspectives on knowledge, its processes, KM approaches, and the reasons why an organisation needs KM. Finally, a discussion of the focus of this research which is about the ways knowledge could be managed in CRM departments in Jordanian banks.

Chapter 3: Research Methodology

This Chapter gives a review of research methods used and the rationale behind the choice of the case study to reach the research aim and objectives. Selection methods and approaches are based upon the research problem and the presented research questions. Reasons for choosing these methodological choices will be provided.

Chapter 4: Performance Model of Knowledge Management Activities

This Chapter illustrates the process of building a scale model to measure the status of KM activities. It is shown here that the techniques of multi-criteria analysis (MCA) can be used to develop performance indicators for activities in KM.

The framework developed incorporates both tacit and explicit knowledge as part of its measurement process. The well established critical success factors (CSFs) are categorised appropriately so that a measurement process can be established. The techniques of MCA are used to develop indicators where the measurements can relate to both quantitative and qualitative features of KM activities.

The overall objective of the produced measurement scale is to provide a method for the measurement of the performance of KM activities which should lead to identification of areas where action needs to be taken. Comparisons of organisations by sector or within a sector can also be achieved by suitable extensions.

This Chapter is divided as follows: (1) Determining of CSFs (2) Establishing a KM framework (3) Determining indicators (4) Determining KM performance indicators using MCA.

Chapter 5: Knowledge Management Practices in Credit Risk Management Departments of Banks in Jordan

This Chapter reports on a survey of the current performance of KM practices in CRM departments in Jordanian banks. The secondary objective is to compare the performance of KM practices between commercial and Islamic banks.

Data have been collected using a survey questionnaire, administered to staff at all levels in the twelve banks in Jordan. The construction of the questionnaire has been based on the concepts and framework of dynamic Multi-Criteria Analysis - Critical Success Factors (KM-CSF). The analytical approach is based on MCA, with the objective of producing vertical and horizontal hierarchies of Bank-KM performance indicators/indices, as well as an overall summary Bank-KM performance indicator. The various indicators for the twelve banks provide a multi-level picture of the KM-performance of the banks, which can be used either as a basis for a year on year monitoring of Bank-KM performance or to allow a bank to compare itself with its competitors.

Chapter 6: Data Analysis of Loan Defaults in Jordanian Banks, Design of a Credit Classification System

This Chapter discusses the process of building a CR classification system which will help credit officers in CRM departments in Jordanian banks in improving the decision performance in granting retail loans. The system will be built using logistic regression analysis. The dataset has been collected from three banks in Jordan (The Arab Bank, the Housing Bank and the Jordan Bank).

This Chapter further examines the variables that influence the risk of loan default in Jordanian banks, the impact of using transformation for potential risk variables on the accuracy of the produced model and, finally, the implications of the suggested model using internal implicit knowledge on managing CR in Jordanian banks.

Chapter 7: Using Data Mining to Support Business Intelligence in Managing Credit Risk

The objective of the Chapter is to classify and compare the predictive accuracy of the improved logistic regression model that has been created (see Chapter 6) with the effectiveness of DM techniques. Moreover, it will provide the concepts and theories that should be reviewed and considered during the selection of the predictor variables for the development of any type of credit scoring system.

Chapter 8: Conclusions and Future Research

This Chapter summarises and concludes the research. The main findings of the present study are outlined. Some implications are discussed. The limitations of the study and future research are also highlighted.

Chapter 2: Literature Review

2.0 Chapter Structure

The second Chapter will introduce the concepts of CRM and its status in Jordanian banks. Then, an overview of topics needed to understand KM: Different perspectives on knowledge, its processes, KM current frameworks, and the reasons why banks need KM. The penultimate Sections discuss the focus of this research which is about the ways knowledge could be managed in CRM departments in Jordanian banks and KM applications in managing CR in banks. Finally, current models of measuring the status of KM will be analysed.

2.1 Brief Overview of Jordanian Economy

Jordan is a small country of 5 million people comprising a highly skilled labour force of 1.2 million. The Jordanian economy suffers from limited financial and natural resources and is dominated by trade and services-related activities which account for more than two-thirds of GDP. Jordan depends heavily on capital inflows of international financial assistance and workers to maintain its economy (Alissa, 2007).

The current structure of the Jordanian economy demonstrates the same traditional characters of a small open economy, which make it extremely susceptible to regional as well as international pressures. Due to the geographical position of Jordan at the centre of regional transformations, the current state of the Jordanian economy can be apparent as a reasonable reflection of developments in the economies of other countries within the region (Reichel, 2003).

2.2 Overview of banking system in Jordan

The banking and financial system accounted for an average 17 per cent of Jordan's GDP. This gives an indication of the vital role that the banking system is playing in the economic development of Jordan.

At the end of 2006, there were 23 banks operating in Jordan: 13 national commercial banks, 8 foreign and 2 Islamic (Table 2.1 shows a profile of Jordanian banks). The banking system in Jordan has JOD17.8 billion in assets and a vast network of branches covering about 11,900 persons per branch. However, the three largest banks account for

55% of the total assets: the Arab Bank dominating the sector with 29% of all assets, the Housing Bank as the second largest with the most extensive branch network and the Jordan National Bank is the third (CBJ, 2006).

Bank	Year of Foundation	Financial Capital (MJOD)	*Ownership Structure	Branches Local	Branches Abroad
Bank of Jordan	1960	86	J= 78% F= 28%	44	7
Jordan National Bank	1956	110	J= 76% F= 28%	42	14
HSBC	1949	14	F=100%	2	-
Rafdien	1957	10	F=100%	1	-
Arab Bank	1930	356	J=43% F= 57%	82	96
Arab Egyptian Bank	1951	20	F= 100%	8	-
Cairo Amman Bank	1960	67.5	J= 77% F= 23%	36	16
Standard Chartered	1969	13	F= 100%	8	-
City Bank	1974	23.5	F= 100%	2	-
Jordan Kuwait Bank	1977	75	J= 42.2% F= 58%	34	3
Jordan Commercial Bank (Gulf Bank)	1978	57.5	J=74% F= 26%	24	3
Arab Investment Bank	1978	44	J=76% F= 22%	8	1
Jordan Islamic Bank	1979	64.1	J=38% F= 62%	54	0
Housing Bank	1974	250	J=22% F= 76%	96	7
Jordan Bank for Investment and Finance	1989	33	J= 89 % F=11%	7	0
ABC	1989	44.8	J= 11 % F= 89%	12	0
Union Bank	1991	55	J= 66 % F= 34%	12	1
Societe General	1993	27	J= 42% F= 54%	16	-
Capital Bank	1996	116	J=80% F= 20%	5	-
International Arabic Islamic Bank	1998	40	J=100%	11	0
National Bank of Kuwait	2004	50	F= 100%	1	-
Audi Bank	2004	20	F= 100%	7	-
Blom Bank	2004	20	F= 100%	3	-

* J = Jordanian, F = Foreigners

Table 2.1: Profiles of Jordanian BanksSource: CBJ (2007)

Jordan has made solid progress in the financial sector by implementing a comprehensive economic adjustment and reform programme. The reforms have resulted in a well-developed financial sector, especially in banks. Despite this relatively high level of development for the region, the need for more financial development is still obvious throughout the Jordanian economy (Reichel, 2003). Jordan's banking system is fully privately owned, well-developed, profiTable and efficient and its banks are advanced in comparison with the other banks in the region (Siam, 2007).

The banking sector in Jordan is the most important in the national economy and has effectively contributed to improving economic development through its important role in mobilising savings and channelling them into different fields of investment (CBJ, 2007). The funds in Jordanian banks were used to:

1. Grant loans for the private and public sectors.
2. Deposit with national and international banks, and
3. Invest in organisations' stocks and governments bonds.

2.2.1 Credit Risk Management in Jordanian Banks

Commercial banks in Jordan granted JOD 6003 million by the year 2004 from JOD 4348 million in 2000. Out of this Figure, 92.4% was granted to the private sector and 7.3% of the total credit went to the public (CBJ, 2006). Notwithstanding the benefits of lending money for borrowers and banks, retail lending is among the riskiest of all banks' functions (Siam, 2007).

Although Jordanian banks are operating efficiently as well as using modern banking methods; such as the Arab Bank which is in the process of electronising all its operations and services; there are still gaps that need to be narrowed or eliminated and these gaps or faults come from the banks themselves(Siam, 2007). In this context, Bani Hamdan (2002) demonstrates that in Jordanian banks there is a shortage of formalised and organised knowledge relating to HR management concerning the attraction, training and development of staff, and the development of knowledge-based expert and decision support systems. Jordanian banks are healthy organisations which operate profitably but need to manage risk more efficiently (Siam, 2007).

According to Siam (2007), banks in Jordan face these problems in lending:

1. Loans have become uncollectible due to mismanagement
2. Loans have become uncollectable due to inadequate lending policy or
3. Loans have become illegal manipulation.

Most of the problems, in fact, are internal to the banks themselves. In its report, the Financial Market International (FMI) (2007) found that Jordanian banks are not effective lenders. Its study found that Jordanian banks lend on the basis of collaterals and relationships. Furthermore, credit officers in Jordanian banks evaluate CR using their personal judgement (mop, 2007).

While the international average for default rates is lower than 5% (CBJ, 2005), in Jordan the non-performing loans, based on the researcher’s interviews, are estimated to be in the region of 11%. This indicates the critical problem that Jordanian banks face and their need to find effective solutions using their internal knowledge to reduce the rate of defaults. In order to achieve this, Jordanian banks need first to improve their performance in managing knowledge. In 2006, Figure 2.1 shows that most Jordanian banks were not effective lenders (percentage of bad loans to gross loans > 0.05).

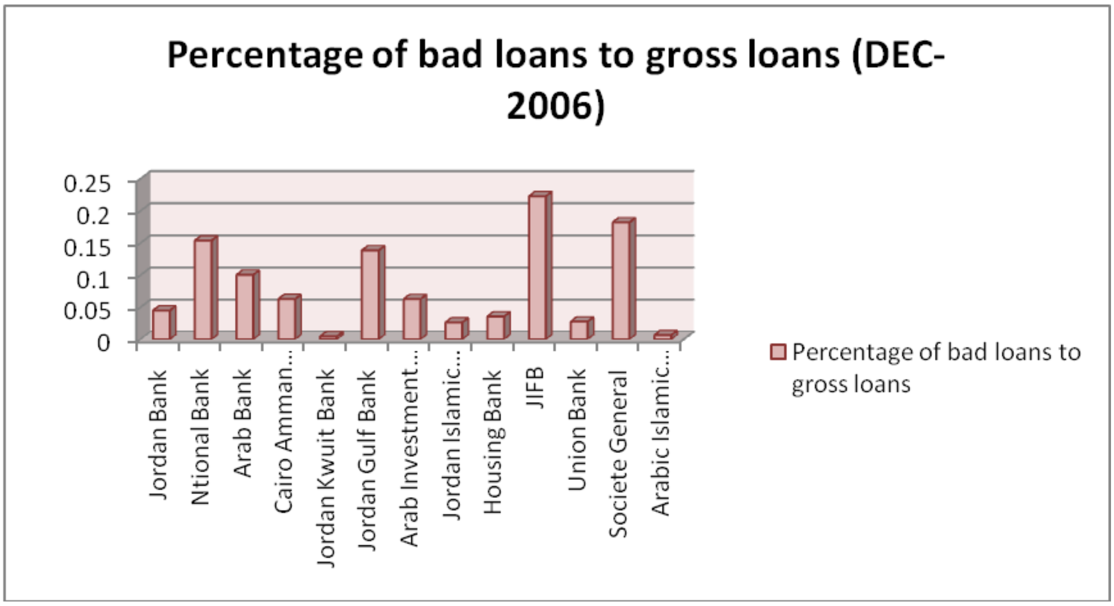


Figure 2.1: Percentage of bad loans to gross loans (December, 2006)
Source: CBJ (2007)

2.2.2 Knowledge Management in Jordanian Banks

To date, there is no sufficient literature review covering KM status in Jordanian banks. The only research is by El-Omary (2004) who reached the following conclusions:

1. In Jordanian banks, use of Decision Support Systems is lacking. This is because they do not know their advantages and such systems are not trusted.
2. Banks in Jordan care about data accuracy deal with the data in a secure way and give their employees easy access to data.
3. Jordanian banks do not apply their internal knowledge efficiently in different banking operations.

4. The banks focus more on developing their internal knowledge by training their employees rather than by using research and success stories.
5. Banks in Jordan do not support appropriately the building of an environment for sharing knowledge.

2.3 Knowledge and Knowledge Management

2.3.1 Knowledge

The world is experiencing an era which has been termed the 'Knowledge Economy'. In this new world, knowledge is the primary key to success in the economy (Sunassee, 2003). This is illustrated by Leibowitz (2000), who states that "Most managers believe that the most important factor that separates themselves from their competitors is knowledge".

The meaning of knowledge is subject to a number of different interpretations. According to Nonaka(1994), knowledge is a complex factor which affects the nature of knowledge creation and sharing. In defining knowledge there are two approaches, the first of which focuses on a non-human view of knowledge (Hussein, 2007). As an example, Brooking (1999) defines knowledge as an information in context with understanding to applying that knowledge. In the same way, Turban (2006) defines knowledge as information that is contextual, relevant and actionable. It is also seen as a collection of facts, values, skills and rules (Pemberton and Stonehouse, 1999). In this respect, knowledge is what gives a 'meaning', therefore the lack of meaning leads to non-useful information (Bhatt, 2000).

The second approach focuses on the structural analysis of the knowledge concepts, with knowledge classified in accordance with its human feature. As an example, Wiig (1999) defines knowledge as "the insights, understandings and practical know-how, that we all possess". Drucker (1993a), for his part, defines knowledge as an information that changes something or somebody either by making grounds for an action or by making an individual or an organisation capable of different and more effective actions. Debowski (2006) sees

knowledge as the process of translating information and past experience into a meaningful set of relationships which are understood and applied by an individual.

Davenport and Prusak (1998) have combined the two approaches in their definition of knowledge (which is one of the most referenced definitions in the literature of knowledge): "*Knowledge is a fluid mix of framed experience, values, contextual information, expert insight and grounded intuition that provides an environment and framework for evaluating and incorporating new experiences and information. It originates and is applied in the minds of the knower. In organisations, it often becomes embedded not only in documents or repositories but also in organisational routines, processes, practices, and norms.*" (Davenport and Prusak, 1998a).

2.3.1.1 Types of Knowledge

Additionally, it is important to recognise the different types of knowledge in order to reveal its potential contribution to the performance of the organisation (Pemberton and Stonehouse, 2000). The differentiation between tacit and explicit knowledge is important since their management is different and needs different means to transfer or share it. Researchers suggest the existence of several forms of knowledge; tacit, explicit, implicit and systemic knowledge at the individual and organisational levels (Polanyi, 1958; Nonaka and Takeuchi, 1995; Inkpen, 1996; Davenport and Prusak, 1998b; and Dixon, 2002).

Polany (1966) first divided knowledge into explicit and tacit knowledge. According to Duffy (1999), explicit knowledge is easier to be documented, articulated and transformed in different formats (structured knowledge). Explicit knowledge is knowledge that can be shared with others, it can be documented, categorised, transmitted as information and illustrated to others through demonstrations, explanations and other forms of sharing (Debowski, 2006). Explicit knowledge is a formal knowledge that has been captured by the organisational knowledge-based systems. It defines the intellectual assets of an organisational independently of its employees. Thus, it is structural knowledge (Stewart, 1999).

In contrast, tacit knowledge is linked to personal perspectives, experiences, habits, culture, believes, skills and know-how. Storing and communicating tacit knowledge is a complex task since it is both social and contextual (Davenport and Prusak, 1998a). Tacit knowledge is a practical knowledge that produces an action (unstructured knowledge), it is the key to getting things done. It needs extensive personal trust in order to transfer it to others. Although tacit knowledge is hardly documented, codified and shared, organisations depend on it to ensure good-quality choices and judgements (Debowski, 2006).

According to this, organisational knowledge could be defined as a fluid mix of framed, experience, documents, repositories, routines, processes, practices, strategies, products, employee's innovations and organisational information that could be used to deal with complex situations.

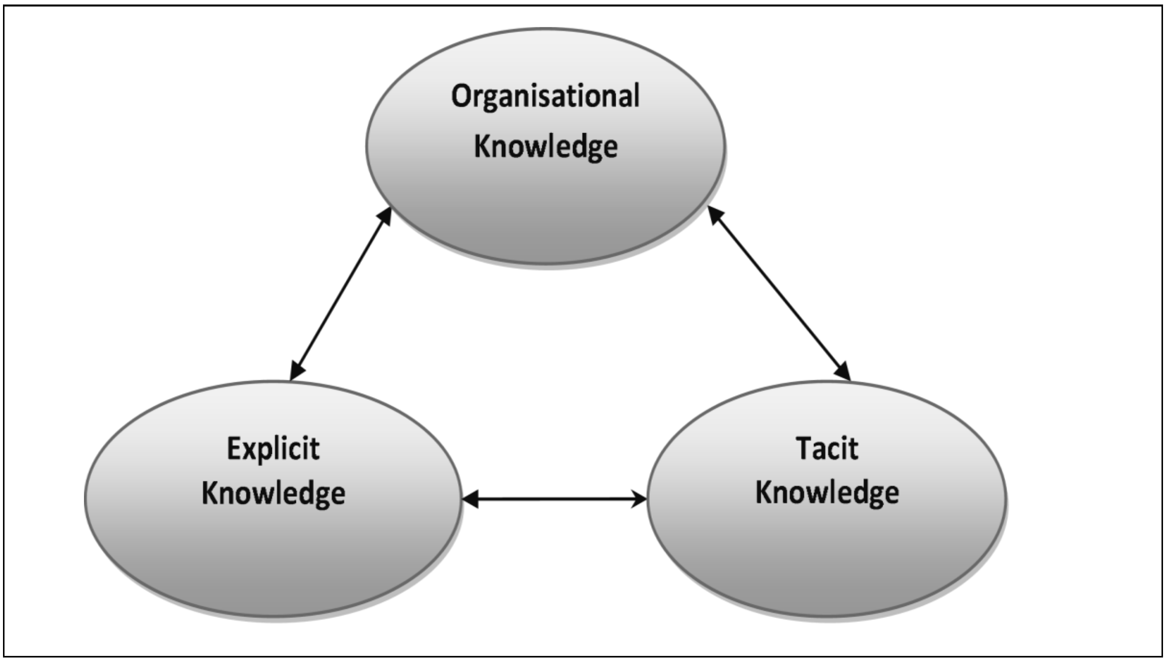


Figure 2.2: Flow of knowledge within an organisation

Figure 2.2 represents the flow of knowledge within an organisation. Firstly, expertise and documented or categorised knowledge that could be transmitted as information emerges ('explicit knowledge') (Debowski, 2006) that constitutes the knowledge representations available to employees. Secondly, it is the collective of the action originated by knowers that, over space and time, appears as a fluid mix of framed experience, documents, repositories, routines, processes, practices, strategies, products, employee's innovations and organisational information ('organisational knowledge'). Thirdly, knowledge that originates in a particular context through the framed experiences and principles of knowers is the emergent ('tacit') knowledge (Polanyi, 1966) that captures the practical knowledge that produces an action.

2.3.1.2 Knowledge Reuse

How organisations create, transfer and reuse knowledge has been a subject of interest to organisations in recent years (Argote, 1999; Levine & Moreland, 2000; and Markus, 2001). Under the environment of open global competition, organisations need an effective reuse of knowledge in order to function effectively (Drucker, 1991). Knowledge reuse is the process of locating and using an existing knowledge to produce a new knowledge. It focuses on knowledge integration through which others' knowledge is integrated into an existing knowledge in order to accomplish an innovative task (Cheung *et al.*, 2008).

According to Markus (2001), knowledge reuse which increases an organisation's performance is an area poorly addressed by many organisations. In the previous Section, two types of knowledge have been defined (tacit and explicit). However, there is middle type between tacit and explicit knowledge which with some focused efforts can be transferred from tacit into explicit. This type of knowledge is called implicit knowledge. Once an organisation has captured this type of knowledge, it needs to identify and document the processes needed to transfer it from tacit to explicit. The process of capturing implicit knowledge and making it available for employees in an organisation is what is called *knowledge harvesting.* Reusing knowledge could be enhanced by transferring implicit knowledge into explicit (Cheung *et al.*, 2008).

This study is examining this argument in the context of CRM to explore how credit officers can use a credit classification system by reusing internal knowledge to tackle managing CR.

2.3.2 Knowledge Management

Much more important than having the data is the firm's ability to combine a high value form of it with experience, context, interpretation and reflection and then apply them to decisions and actions. Managing knowledge will increase its value (Davenport, 1998; Zack, 1999; Sveiby, 2000). Recently, this managerial activity has been known as Knowledge Management (KM) (Sunassee, 2003).

Penrose, acknowledged as one of the first scholars to recognise the role of knowledge in business organisations, saw acquiring knowledge as a social learning process. This increase in knowledge not only causes the productive opportunity of a firm to change in ways unrelated to changes in the environment but also contributes to the 'uniqueness' of the opportunity of each individual firm (Penrose, 1959).

Nonaka (1991) who was first to structure the concept of 'personal tacit' knowledge; originally proposed by Polanyi (1967); as a spiral model for knowledge creation and transfer, afterwards designated the SECI (Socialisation, Externalisation, Combination, Internalisation) model. This model, along with its original assumption that tacit knowledge can be transformed into explicit knowledge, all set in a corporate context, is likely the most broadly adopted KM concept in the first generation. (Nonaka, 1994; Nonaka and Takeuchi, 1995).

Knowledge management can be viewed from different perspectives which then produce different interpretations (Hussein, 2007). In the literature, there are two different approaches in defining KM: IT-driven and culture-driven. Researchers who define knowledge as information in a context, define KM as an advanced level which uses IT to facilitate create, codify, store, transfer and apply knowledge within an organisation (Davenport, 1993; Boisot, 1995, 1998; Davenport and Prusak, 1998; Stewart, 1999).

In this context, Davenport (1998) states that "Knowledge management systems (KMS) are tools to effect the management of knowledge and are manifested in a variety of implementations, including document repositories, expertise databases, discussion lists, and context-specific retrieval systems incorporating collaborative filtering technologies. In the same context, Debowski (2006) defines KM as "the process of identifying, capturing, organising and disseminating the intellectual assets that are critical to the organisation's long-term performance". Also, Turban (2006) defines KM as "a process that helps organisations identify, select, organise, and transfer important information and expertise that are part of the organisation's memory and that typically reside within the organisation in an unstructured manner".

The culture-driven approach claims that KM is about facilitating the environment that supports creating/sharing knowledge within an organisation by applying social networks or communities of practice (Nonaka and Takeuchi, 1995; Wenger, 1998; Alavi and Leinder, 1999). In this context, Nonaka and Takeuchi (1995) stress the crucial importance of individual intuition, hunches and (deliberately) rather vague open-ended vision as essential components in the knowledge creation and innovation processes. They point out that theories of KM in western companies have been largely based on, and limited to, the western epistemological traditions of rationality and empiricism, and little or no account has been taken of the human creative processes that need to be involved.

Sveiby (2000b) argues that knowledge is not something that can be managed; KM is about managing the environment in which knowledge is created. Alavi and Leinder (1999) define KM as "the systematic and organisationally specified process for acquiring, organising, and communicating knowledge of employees so that other employees may make use of it to be more effective and productive in their work ",

As it has been defined two types of knowledge, tacit and explicit, this implies that managing knowledge is about managing both of them, which combines the IT-driven approach which facilitates more managing of explicit knowledge and the culture-driven approach that facilitates more managing of tacit knowledge. In a business context, KM is about managing the activities of knowledge workers, which is achieved through facilitating, motivating, leading and supporting knowledge workers and providing an environment which facilitates KM processes.

2.3.1.2 KM Frameworks

Despite the fact that knowledge management has been considered as an ultimate factor in the success of any organisation in the current competitive market, many KM initiatives have ended up in failure due to their scarcities of covering the different basic perspectives for managing knowledge, this includes the culture, the technological infrastructure, and the different types of knowledge (Storey and Barnett, 2000). According to Wong and Aspinwall (2004), implementation of KM may fail without a proper guidance about what to do, how to do, and where to start. Based on that, there is a great necessity for creating and adapting systematic practices of management knowledge that must be built on a significant KM approach (Drucker, 1993).

As shown in the literature, different KM frameworks that have been developed were concentrated on different sides of KM for different purposes. One of the well known frameworks provided by Nonaka et al (1995) is the knowledge creation framework which focuses on knowledge evolution and knowledge conversion between explicit and tacit knowledge. As it can be noticed from its name, Nonaka's framework is not a comprehensive as it concentrates on one part of KM that is knowledge focus.

Other frameworks that have been introduced to answer the question "What is KM?" are focusing on describing knowledge cycle activities (Rubinstein-Montano et al, 2001). The main concentration of those frameworks emphasises on knowledge flow from creation to application, but not guidance on how to implement KM.

Another type of KM frameworks has been introduced to investigate and analyse the KM initiatives that were performed in particular industries. A well known framework of this category developed by Apostolou and Mentzas (1998), and Lai and Chu (2002).

Wong and Aspinwall (2004) have stated the need of KM frameworks that provide more clear guidance on the implementation phase, which require not only answering "What is KM?" but also describing and advising how to implement KM.

2.3.1.2.1 Why a KM Implementation Approach is needed

Implementing a knowledge management approach that only concentrate on one particular aspect such as information technology leaving out other aspects, would not bring out the desired outcome for the organisation as KM is a complementary process (Arora, 2002). For example, managing explicit knowledge without giving attention of tacit knowledge may end up the whole process with failure. Therefore, an appropriate KM approach must consider all different aspects, and must be aligned with organisational strategy for accomplishing the successful implementation (Arora, 2002). According to Wong and Aspinwall (2004), the main reasons for developing a comprehensive KM framework are:

- To provide the clarification of KM domain increasing the awareness, helping to understand the processes and activities concerned.
- To help understand the KM in a broader perspective offering a holistic view of KM.
- To enable communication of KM in the organisation providing the common language for people. Make easier the communication between managers and employees.
- To outline the elements, processes and influences, showing the scope and phases of the tasks that need to be done.

- To help the managers in auditing the process of KM implementation as it provides a detailed checklist to be followed.

- To help the managers to control the whole process of KM implementation, ensuring that all organisational efforts are coordinated in systematically.

In the same context, Wong has proposed four main elements to have to be remembered when developing a KM framework, these are:

1. The Structure.

2. Knowledge types or knowledge resources.

3. KM processes or activities.

4. KM influences or factors.

Clear framework structure should include all milestones and tasks in a logical order that maintain a successful implementation of KM. in order to accomplish that, a necessity to organise and divide the activities into separate segments and different stages is needed (Wiig et al., 1997; and Rubenstein-Montano et al., 2001). The framework should make clear identification of different presented types of knowledge in the organisation giving the opportunity to align different management strategies for explicit and tacit knowledge. Managing explicit knowledge which is created by individual and can be obtained from outer can be done by different tools such as channels a knowledge map; lessons learnt databases, groupware and electronic data interchange. On the other hand, managing tacit knowledge can be done through listing of the experts possessing this tacit knowledge, group meeting, face to face conversations and practice forums (Sanghani, 2009).

The next logical step in KM framework is to encompass the process and activities that effectively handle knowledge assets, as these processes are considered the main activities in knowledge cycle (Rubenstein-Montano et al., 2001). Holsapple and Joshi (2002) assert that one of the main tasks of KM is to manage organizational knowledge resources to boost organizational effectiveness. Both organizational knowledge-related efficacy and return from knowledge assets depends on KM since it manages effective knowledge processes (Wiig, 1997). Additionally, the awareness about KM enablers and disablers is vital for successful KM implementation (Sanghani, 2009). Obviously, the acknowledgement of those factors will help to make the optimal decisions, which is about the measures and future course of actions directed to use the enablers for the organisational advantage and to decrease the effect of the disablers (Wong and Aspinwall, 2004).

Organisational culture has been considered as the crucial aspect that influence the process of KM implementation (Jarrar, 2002; Apostolou and Mentzas, 1998; Liebowitz, 2000) through either enabling it to be successful or causing it to fail. This has been concluded since the creation, sharing and distribution of knowledge is in direct relation with the organisational culture.

Technological support plays a significant role in KM implementation (Jarrar, 2002). Based on that, KM framework should support the balance between a technological and a social approach to ensure successful KM implementation (Gao et al., 2002;and Offsey, 1997). According to Carter and Scarbrough (2001), most KM initiatives that were focused on the technological side only ignoring the social aspect have failed. Concurrently, the humans cannot manage the knowledge in an adequate manner as they could do with using the technologies.

2.3.2.1 Knowledge Management Processes in Context of Credit Risk Management

According to Boisot (1995), Davenport (1993), Davenport and Prusak (1998) and Stewart (1997), KM contains the following processes:

1. Knowledge creation
2. Knowledge sharing
3. Knowledge storing
4. Knowledge application

1. Knowledge Creation

In banks, a new CR implies a new knowledge that needs to be identified to measure and identify the new CR. For a bank, there are different resources to create knowledge:

- Internal resources like corporate memory, communication between communities of practice, sharing expertise, attending internal meeting and conferences.

- External resources like attending external conferences, and marketing and academic researches facilitated through using technologies such as Internet and teleconferencing (Davice, 1998; Lim, 1999; Fielden, 2001).

  According to Mertins (2001), Beveren (2002) and Despres and Chauvel (2003), knowledge is created through the cooperation between group work and communities of practice to solve problems innovatively, which will add value for an organisation.

  Nonaka and Takeuchi (1995), however, represent knowledge creation through their innovative SECI model which represents the knowledge-life cycle:

- Socialisation: creating CR knowledge through sharing tacit knowledge between individuals
- Externalisation: creating CR knowledge through coding the tacit knowledge to become explicit.
- Combination: creating new explicit from existing explicit CR knowledge through coding it using different applications.
- Internalisation: transferring explicit CR knowledge into tacit through using explicit knowledge in practice or as a routine.

2. Knowledge Sharing

According to Snowden (2000), many organisations have started to understand the importance of sharing knowledge between employees through using intranet or by spreading the success and failure stories which will improve their business performance and increase their competitive advantages. Also, Descouza (2003) asserts that, an organisation could not add a value from knowledge unless it spreads knowledge and encourages employees to share it. To spread and encourage sharing CR knowledge, banks can use formal and non-formal channels like training, Internet, intranet and teleconferencing.

3. Knowledge Storing

Duffy (2000) asserts the importance of storing knowledge which will enable others to retrieve and use it. According to Alavi and Leidner (2001), it is an organisational responsibility to store knowledge in different formats using different applications that support both explicit and tacit knowledge like knowledge-based and expert systems. As well, Miller (2002) states, “A knowledge-based system where an organisation stores its knowledge forms the centre that enables it to survive in a hard competitive market.”

4. Knowledge Application

According to Alavi and Liender (2001), managing knowledge is useless unless an organisation applies it. Creating and sorting knowledge will not improve an organisation’s performance unless knowledge is applied in strategic process, production and decision making (Fielden, 2001). Application of KM in managing CR will be analysed in more detail in Section 2.7.

2.4 Why banks need Knowledge Management

Many researchers (Grant, 1996; Boisot, 1998) suggest that knowledge is the key resource which is capable of achieving a sustainable competitive advantage. Organisations have shifted their focus from the resource-based to the knowledge-based view (Armisted, 1999). Studies have identified employees and an organisation’s knowledge to be the most important factors of competence in today’s competitive market (e.g. Lubit, 2001; and Rogers, 2001). All of them assert the importance of creating, transferring and applying knowledge, which support an organisation’s existence. This implies that the knowledge-based view has the premise that managing knowledge plays an essential role in the existence of an organisation. Thus, the important message we can conclude is that the emergence of the knowledge-based view requires a new synthesis of training, education and other forms of learning and communication skills.

Researchers and KM experts (Nonaka, 1991; Allee, 1997; and Davenport and Prusak, 2000) have proposed the advantages of managing knowledge in an organisation, which in turn will be used as a core resource to achieve a competitive advantage. Actually, an organisation manages its knowledge by reason of two factors: internal and external.

2.4.1 Internal Factors

Managing knowledge needs changing people’s mindsets. Previously, organisations did not encourage dissemination and sharing of knowledge among their employees. Later on, this managing of knowledge has become more crucial for organisational existence. This is because many organisations’ activities are knowledge driven (Sunassee, 2003).

According to Armisted (1999), it is in the activities that the role of knowledge in organisational effectiveness will be demonstrated. Davenport (1999) relates KM activities with activities that affect an organisation's profitability such as organisational performance measurements, indicators of the employees' capacities to carry out tasks related to knowledge and generating of ideas and innovation.

According to Karkoulian (2008), by managing knowledge, employees are empowered to resolve problems effectively, take decisions, respond to customers' queries and create new products and services tailored to clients' needs.

Knowledge is the primary resource which has an economic value. It is an important resource to increase an organisation's performance, increase productivity and provide it with a stability factor in an suitable environment (Ermin, 2000). In this context, Bhatt (2002) states that, knowledge plays an important role as a crucial and strategic resource in an organisation. According to him, knowledge is a key component in building and sustaining an organisation's core competencies.

To survive in this hard competitive environment, an organisation needs to support sharing knowledge among its employees and create new knowledge to anticipate the future (Abecker, 1998). As Armisted (1999) states, "For organisations, by having greater access to their employees' knowledge, they make better decisions, reduce work, increase innovation which will improve overall an organisation's performance.

2.4.2 External Factors

In the second half of the 1990s, the knowledge economy became recognised as the cornerstones of the global economy and , the balance between knowledge and resources has shifted so far towards the former that knowledge has become the most important factor determining the standard of living, more than land, tools and labour. By the mid-1990s, KM initiatives were flourishing and, importantly, a number of European and Japanese firms had instituted focused KM systems.

As the speed of markets is growing and their dimensions tend towards globalisation, response time is decreasing and aggressive pressure is increasing. Because of that, knowledge plays an important role in an organisation's survival. In the last decade, business organisations started dealing with knowledge as their most valuable strategic asset (Zack, 1999a).

Knowledge of various forms is recognised as a crucial business asset, to be used for the development of new products and services, leading to competitive advantage. Knowledge is the main factor in an organisation's success. As Nonaka(1998) puts it, "Because of high uncertainty, the one sure source of competitive advantage is knowledge". Knowledge has been recognised as an important asset for sustaining advantage. Lebowitz (2000) states that "Most managers believe that the most valuable asset to help in competing against their competitors is knowledge and more specifically tacit knowledge which resides in their employees' minds". Because of that, managing knowledge is highly critical for organisations to achieve a competitive advantage.

2.5 Risk Management

Risk management is commonly perceived to mean minimising risk but actually risk management is the goal of optimising risk/reward trade-offs (Kanwar, 1999). Risk management is all about risk control, designation, mensuration and risk monitoring; also it complies with business strategy within the boundaries of targets and goals ordered by the Board of Directors. The main idea is to anticipate the profit from the risk taken; making sure that the organisation (the bank) has adequate means to reduce the impact of risk in the case if something goes wrong.

Banks are in the business of taking risk which means risk management is a cornerstone of prudent banking practice (Al-Tamimi, 2007). Where there are uncertainty and complexity there is risk. Therefore, risk management is about reducing uncertainty and analysing complexity (Mondale, 2006).

2.5.1 Credit Risk Management in Banks

Simply put, the management of risk associated with credit is what we term 'CRM'. The goal of CRM is to maximise a bank's risk-adjusted rate of return by maintaining credit risk exposure within acceptable parameters (Basel report, 1999).

The very first purpose of a bank's credit strategy is to determine how much risk the bank is willing to carry. Once this is determined, the bank can develop a plan to optimise return while minimizing the CR and ensuring it falls within predetermined limits. The bank's CR strategy should thus spell out the institutes plan to grant credit based on various client segments and products, economic sectors, geographical location, currency and maturity.

As stated by the informative paper supplied by the Basel Committee on Banking Supervision (1999), CR commonly originates from a situation where a debtor fails to pay his/her debt and it has an impact on the economic state of the bank. Commercial banks are mainly faced with CR and retail loans are the largest and most obvious source of this type of risk (Al-Tamimi, 2007).

Therefore, over the last decade, a number of international largest banks have developed sophisticated credit risk systems in an attempt to model credit risk. The main aim of such models is to aid decision makers in banks in quantifying, and managing risk. The outputs of these models also play increasingly important roles in banks' risk management and performance measurement processes, such as customer profitability analysis, and risk-based pricing (Crouhy et al., 2000).

According to Basel report (2003), banks develop credit risk models for the following potential benefits:

- The use of credit risk models offers banks a framework for examining credit risk in a timely manner, and analysing factors contributing to risk. These benefits contribute to banks' overall ability to identify, measure, and manage risk.
- Credit risk models provide an approximation of unexpected loss which reflects individual portfolio composition. This provides a better reflection of concentration risk compared to non-portfolio approaches.
- The expected default frequency or EDF– the probability of a particular credit facility defaulting during the time horizon – is a critical model input. Within most credit risk modelling systems, a customer's internal credit risk rating (as determined by a bank's credit staff) is a key – if not the sole – criterion for determining the EDFs applicable to the various credit facilities associated with that customer; generally all the customer's facilities are presumed to default concurrently, or not at all.
- Additionally, credit risk models may offer: (a) the motivation to improve systems and data collection efforts; and (b) more accurate risk- and performance-based pricing, which improve decision-making process.

2.5.1.1 Conditional and Unconditional Credit Risk Approaches

Practitioners distinguish between conditional models that attempt to incorporate information on the state of the economy, such as levels and trends in domestic and international employment, inflation, stock prices and interest rates, and even indicators of the financial health of particular sectors, and unconditional models that reflect relatively limited borrower or facility-specific information (Chabane et al., 2004).

Approach	Examples	Description
Unconditional Approach	unexpected losses (UL) approach, CreditMetricsTM and CreditRisk+TM	These models base EDFs and derived correlation effects on relationships between historical defaults and borrower-specific information such as internal risk ratings.
Conditional Approach	McKinsey and Company's CreditPortfolioViewTM	The rating transition matrices are functionally related to the situation of the economy, as the matrices are modified to give an increased likelihood of an upgrade (and decreased likelihood of a downgrade) during an upswing (downswing) in a credit cycle.

Table 2.2: Conditional and Unconditional Approaches
Source: Basel report (2003)

According to Rodriguez and Edwards (2009), risk management practice does not use KM to improve and develop new answers to the threats. A main reason is that it is not obvious how to break down the 'organisational silos' view of risk management (RM) that is regularly taken. As a result, there has been relatively little work on finding the relationships between RM and KM.

2.6 Relation between Knowledge Management and Credit Risk Management in Banks

As earlier stated, financial institutions like banks have their main business in loans. As a consequence, a level of risk is associated with it and the management of this kind of risk is an essential foundation for a successful business.

CR arises when a bank as a business practice does not demand cash up-front as payment for products or services. When a transaction is first completed and the customer is billed later, the bank carries a risk between the delivery of the service or product and the customer's ability for payment. The availability of CR is influenced by different variables, in most of which knowledge is an underlying factor, hence being able to efficiently evaluate and measure CR is of the utmost importance. In order to achieve this, organisations can apply credit scorecards to assess and rank customers, after which the appropriate strategies are implemented, e.g. high or low interest rates dependent on the ranking.

RM is a knowledge domain used at the centre of each fiscal organisation and which covers all operations with an impact on its risk profile. It is of the great importance that people who are involved with risk distinctly comprehend it. According to Marshal and Prusak (1996), *"Risk Management is frequently not a problem of a lack of information, but rather a lack of knowledge with which to interpret its meaning*".

CRM in banks uses a combination of key variables with knowledge at the centre. As a result, the quality of information and knowledge must be of a very high standard and knowledge has to be effectively managed both formally and informally within the banks. Though not shown as direct financial implications on the asset register, this has a major financial impact on the bank, which in turn will affect effective risk management. Knowledge is a risk asset which if not appropriately managed can become costly to the organisation. As information has a finite life span, it is necessary to have access to real-time information and up-to-date knowledge regarding issues, products, customer status and even the reputation of a company and its brand. This in turn affects the availability of credit and the ranking of the customer on the risk scale. As Fourie and Shilawa (2005) stated, when a new risk is identified a new knowledge is required.

Knowledge as we have come to understand is either articulated or not, but experiences and non-articulated knowledge are also classified as knowledge, which provides an extra edge in handling risk and dealing with credit transactions and creates an avenue for innovation and uncovering of new information which could help in giving the business a leading edge.

Since we have illustrated that knowledge is essential for the management of CR, knowledge needed to manage CR can therefore be called 'CRK' (CRK) and this can be divided into two main forms:

1) Explicit CRK: Articulated or documented knowledge about customers, their information, behaviour, credit history, and the application of data mining over historical data of previous customers who borrowed money from the bank to uncover new knowledge which could be used to predict behavior and patterns of customers. This in turn will help to minimise the risk associated with the decision of lending to a new customer.
2) Tacit CRK: Experiential knowledge gained by credit officers from dealing with past customer behaviour, needs, motivators, patterns and portfolios, which is not articulated, is of great advantage in being able to predict and give insight into the probability that a risk is not worth taking or can be positively used, especially when documentation status is uncertain or sits on the border.

2.7 Knowledge Management Applications in Managing Credit Risk

As it has been illustrated in Section 2.3.2.1, managing knowledge by organisations is useless unless knowledge is applied in banking operations. In banks, different applications for knowledge in managing CR are available which include statistical and data mining (DM) applications. The following Sections represent a comparison between different methods that could be used by banks in managing CR.

2.7.1 Credit Risk Management and Decision Support Systems

In the late 1970s, a number of organisations developed information systems (IS) that used data and models to help decision makers analyse semi-structured problems. These systems were called Decision Support Systems (DSS). DSS can be designed to support decision-makers at any level in an organisation in operations, financial management, and strategic decision-making process.

Turban (1995) defines DSS as "an interactive, flexible, and adaptable computer-based information system, especially developed for supporting the solution of a non-structured management problem for improved decision making.

According to Power (2009), DSS could be divided into following types:

1. Communication-driven DSS: Most of the managers use this type of DSS in order to communicate within the organization and come to a decision or solution to the approach of organization, like it can be seen or on the web.
2. Document-driven DSS: Organizations which give importance to security like banks, they use document DSS. The requirement for Document DSS are the Spread sheets, text documents or data base records, which help them to take decision and use the information to something more effective.
3. Data-driven DSS: This uses the already present internal and external data in a manipulative way in order to fit the requirements of the decision maker. Quite widely used in Companies, Ware house and stores and for Geographic Information System (GIS).
4. Model-driven DSS: The decision maker already provides the data and the constraints in has to fit in the particular multi-criteria model in order to aid in making decisions. This type of DSS is highly useful in predicting the effects early on bases of changes the past data. It works like a tailor made model for organization.
5. Knowledge-driven DSS: Knowledge –driven DSS suggests the actions to the user. Large amount of data is sift, the behaviour of the data is supervised, and the knowledge which is bought forward on bases of sift is used to identify the suitable decision.

   In the context of CRM, Joao (2008) states, credit risk management and scoring is one of the most important issues related to banking and lending organisations. The credit score is used by banks to evaluate the probability as to whether the loan applicant is viable based on his characteristics noted through various attributes associated with his financial obligations. This information is crucial to banks as empowered with is information banks tend to increase volume of granted credits at the same time keeping the likelihood of defaulted loans low (Joao, 2008).

2.7.2 Methods of Predicting Defaulting Borrowers

In a well-built financial system, crisis management is on the downstream and risk prediction is on the upstream (Yeh and Lien, 2009). According to Fensterstock (2005), most CR evaluation systems currently in use are based on some form of judgemental-based system which makes quantifying a risk a big challenge. According to him, this problem can be solved through a statistical or artificial intelligent (AI).

In the banking industry, customers regularly apply for a credit to make purchases. The risk for banks depends on how well they differentiate between the good applicants from the bad applicants. One widely adopted technique for solving this problem is "Credit Scoring" (Altman *et al.*, 2006).

Banks commonly makes two types of decisions: first, whether to grant credit to a new applicant or not, and second, how to deal with existing customers, including whether to increase their credit limits or not. In both cases, to build a credit scoring model, it is critical that there is a large sample of previous customers with their application details, behavioural patterns, and subsequent credit history available (Jarrow, 2001).

Credit scoring is a set of decision models and their underlying techniques that aid decision makers in banks in the granting loans. These techniques decide who will get credit, how much credit they should get, and what operational strategies will improve the profitability of the borrowers to the banks (Altman *et al.*, 2006).

The main task of estimating the risk of default has been enhanced by credit scoring models to include other aspects of credit risk management: at the pre-application stage (identification of potential applicants), at the application stage (identification of acceptable applicants), and at the performance stage (identification of possible behaviour of current customers) (Bakshi, 2001). According to Dudoit (2002) scoring models with different objectives have been developed:

1. Forecast the future behaviour of a new customer by predicting the risk of loan-default or poor repayment behaviours at the time the credit is granted.
2. Select best possible collections policies in order to minimize the cost of administering collections or maximizing the amount recovered from a delinquent's account.

There are several methods suitable for credit scoring in the banking segment:

1. Neural networks: A neural network (NNW) is a mathematical representation inspired by the functioning of the human brain. A typical network is composed of a series of interconnected nodes and the corresponding weights between them. It aims at simulating the complex mapping between the input and output (Hui-Chung, 2007). Many types of NNW have been specified in the literature.

   NNWs have higher credit scoring capability than other statistical methods (e.g. LDA and logistic regression) (Hui-Chung, 2007). Their major drawback is their lack of explanatory capability. While they can achieve a high prediction accuracy rate, the reasoning behind why and how the decision was reached is not available. For example, in a case of not accepting a loan, it is impossible to determine which characteristic(s) was (were) exactly the key one(s) to prompt rejection of the application. Accordingly, it is very difficult to explain the decision results to managers (Baesens, 2003; Lee, 2002; West, 2000).

2. Linear Discriminant Analysis (LDA): The aim of LDA is to classify a heterogeneous population into homogeneous subsets and further the decision process on these subsets by assuming the prior probabilities of the analysis target are equal (Hui-Chung, 2007). As stated by Hand (1997), LDA is the first proposed technique for building credit scoring models.

   The advantages of the LDA method are that it is simple, it can be very easily calculated, indeed it works very well and it is often used by banks for credit-scoring purposes. The disadvantage is that LDA requires as a rule distributed data but the credit data are often non-normal (and categorised) (Eisenbeis, 1977; Dudoit, 2002; and Webb, 2002).

3. K-nearest Neighbour Classifier: This is a standard technique in pattern recognition. It serves as an example of the non-parametric statistical approach. K-nearest neighbour classifier is one of the simplest and most straightforward classifiers (Kilian, 2005).This technique assesses the similarities between the pattern identified in the training set and the input pattern (Henley, 1996).

   However, it has been shown that when the points are not uniformly distributed, predetermining the value of K becomes very difficult (Domeniconi, 2002). Holmes and Adams (2002) conceived that there is a shortage of a formal framework for choosing the k and that the method can only make discrete predictions by accounting the relative frequencies which have no probabilistic interpretation. They tried to overcome these difficulties by presenting the Bayesian approach as a solution, which integrates over the choice of k. Such an approach draws the conclusion that marginal predictions are given as proper probabilities.

4. Logistic regression: This is used for modelling the binary outcome variable which accepts continuous and categorical predictors.

   Logistic regression is a variation of ordinary regression used when the dependent variable is a binary variable (i.e., it takes only two values, which usually represent the occurrence or non-occurrence of some outcome event) and the independent (input) variables are continuous, categorical, or both. Unlike ordinary linear regression, logistic regression does not assume that the relationship between the independent variables and the dependent variable is a linear one. Nor does it assume that the dependent variable or the error terms are distributed normally.

   Logistic regression has been widely used in credit scoring applications due to its simplicity and explainability. Recently, Charitou (2004) found that the logistic regression analysis is superior to other methods in predicting defaults. Desai (1996) examined neural networks, logistic regression and linear discriminant analysis for scoring credit systems. He concluded that NNWs outperform linear discriminant analysis in classifying loan applicants into good and bad applicants, and logistic regression is good as NNWs.

In the previous Sections, the importance of managing knowledge in banks internally and externally has been discussed. This implies that any gap in implementing KM will affect a bank performance negatively in accomplishing internal activities such as making decisions, innovation and its ability to reach a competitive advantage. These implications assert the importance of using measurement models (tools) which help organisations in defining different gaps in implementing KM, and suggest solutions to fill these gaps.

2.8 Defining Gaps in Knowledge Management Implementation

Several studies proposed the concept of ‘knowledge management gap’ to describe the difference between the enterprise’s current capability and the capabilities required for KM.

Beyond the KM gaps, there exist different perceptions of KM activities and implementation amongst employees of differing levels and positions. The inability to identify and resolve any gaps will greatly impact on the organisation's performance. Thus, it would be beneficial for the organisation to build a scale model that would analyse the organisation KM status, evaluate the implementation activities of KMS and identify any gaps to success.

In KM measurement models so far there are two broad directions in which developments have taken place; one treats knowledge as intellectual capital (IC) and the other is concerned with the performance of activities critical to the creation and sharing of knowledge. The latter clearly feeds in to intellectual capital. Methods of IC measurement reported in the literature include Skandia navigator (Edvinsson and Malon, 1997), IC Index (Roos *et al.*, 1997) and IC audit model (Lynn, 1998). One could argue that the measurement of knowledge wealth is a sufficient indicator of the performance of KM. The problem however is that IC approaches in themselves do not give the 'drill down' capability required for identifying problem areas of KM environment.

Methods currently used to measure performance of KM activities include the extended use of the Balanced Score Card (BSC) (Kaplan and Norton, 1996). The BSC method can be used to identify training needs and staff development required to improve internal processes in relation to any initiatives to create or satisfy customer demands, not necessarily in the KM context. This has been extended (Gooijer, 2000) to measure KM performance incorporating a KM framework. The framework does not however include tacit knowledge. To date, there are no methods available to give performance indicators incorporating both tacit and explicit knowledge in an established framework.

Another method, the KMMM-KM Maturity Model (Gallagher and Hazlett, 1999; Kochikar, 2000; Klimko, 2001), identifies the maturity states for effective KM activities to take place. KMMM defines five KM maturity levels: Default, Reactive, Aware, Convinced,and Sharing. Each maturity level is characterised by certain capabilities relating to People, Process and Technology; scores can be given to the capabilities using questionnaires. Table 2.3 presents different examples of KM measurement models.

Category	Model	Researchers	Characteristics	Limitations
Knowledge Asset	Balanced Scorecard (BSC)	Kaplan& Norton(1996),	Clear correlation between indicators and performance.	No external comparison possible. No focus on KM activities
Knowledge Asset	Intellectual capital (IC)	Roos *et al.* (1997)	Focus on monitoring dynamics of IC. Single index of several indicators based on correlating changes in IC with market changes	Very context specific and limited in universality.
Qualitative	KM Maturity Model	Langen (2000)	Systematic development of KM structures	Not based directly on processes and activities in KM.
Metric based	Knowledge Work Measurement	Smits and Moor(2004)	Determines impacts and interrelationships of influence factors	Operative limitations, since numerous different metrics necessary

Table 2.3: Examples of existing models for measuring KM

The problems with the above such models are that most of them tend to focus on a specific aspect of the KM system, such as technology for example. This is too narrow a focus for KM measurement purposes. It does not give the complete picture of the interacting factors required for the KM system. Because of that, one of the objectives of this research is to develop a scale to measure the status of KM behaviour and practices. The overall purpose of developing this measurement scale is to provide a method for the measurement of the performance of activities in KM which should lead to identification of gaps in managing and reusing knowledge where action needs to be taken.

Chapter 3: Research Methodology

3.1 Research Design

This research was conducted to determine whether developing a KM approach to CR makes the decision-making process and CRM more effective. The research seeks to answer the following questions:

1. Are Jordanian banks efficient in terms of managing credit risk?
2. Are Jordanian banks efficient in managing knowledge in CRM departments?
   2.1 What is the contribution of the 'People & Culture' factor in supporting KM in banks in Jordan?
   2.2 What is the contribution of the 'Processes' factor in supporting KM in banks in Jordan?
   2.3 What is the contribution of the 'IT' factor in supporting KM in Banks in Jordan?
3. What are the enablers of KM in Jordanian banks?
4. What is the relationship between KM and CRM?
5. What are the variables that influence the risk of loan default in Jordanian banks?
6. What is the impact of using transformation for variables on the accuracy of the produced model?
7. What are the implications of the suggested model using internal implicit knowledge on managing CR in Jordanian banks?

Based on the stated purpose and the questions mentioned above, the following hypotheses are formulated (corresponding null hypotheses are the converse):

H1. There is a significant impact of the factors 'Processes and IT' on the factor 'People and Culture'

H2. There is a significant difference between the Jordanian commercial banks and Islamic banks in the performance of KM practices

H3. There is a significant relationship between KM and CRM

First, in order to answer research questions 1, 2, 3, and 4, the views of credit officers and managers in Jordanian banks have been obtained. All Jordanian banks were included in the survey setting. The study concentrates on studying the Jordanian banks only. According to the CBJ (2007) report, in Jordan there are fourteen local commercial and two Islamic banks. The sixteen banks were approached by the researcher and twelve agreed to be part of this study, which represents 75% of the banking community. The sample is comprised of 242 respondents from twelve banks within CRM departments in Jordanian banks. The participants answered a survey questionnaire structured in Likert format. Data gathered from this research instrument were then computed for analysis and interpretation. Along with primary data, secondary sources in the form of published articles and literature to support the survey results have been used (see Chapter five).

Second, in order to answer research questions 5, 6 and 7, CRISP-DM methodology has been used to analyse historical data of previous retail loans (2755 cases) (see Chapter 6 and 7).

3.1.1 Case Study

Case study research is an applied technique in the study of business markets which is particularly suitable to investigate complex and ambiguous contexts (Flyvbjerg, 2006).

Yin (2003) defines case study research as:

"It is an empirical inquiry that investigates a contemporary phenomenon within its real-life context; when the boundaries between phenomenon and context are not clearly evident; and in which multiple sources of evidence are used".

Yin (2003) argues that there are major benefits to be derived from utilising this method in research which include:

- Exposure to a variable number of real life cases and situations from which the researcher is able to get a clear and contextual picture of the situation being researched.
- Being able to cope with the technically distinctive situation in which there will be many more variables of interest than data points
- Use of multiple sources of evidence, with data needing to converge in a triangulating fashion
- Taking advantage of earlier developed theoretical propositions in guiding data collection and analysis.

A case study is particularly good for examining “what” and “why” as well as “how” questions (among question series: “who”, “what”, “where”, “how” and “why”), which are enquiries about a contemporary set of events over which the researcher has little or no control (Yin, 2003, Saunders et al., 2007). Especially, the “what” and “how” questions are suitable for a case study because these questions deal with investigating the reasons and the links between different factors in the study (Yin, 2003).

Based on previous researches (Stake, 1995; Yin, 2003), six main steps have been suggested for carrying out a successful case study research. And the method can be summarised as in Figure 3.1:

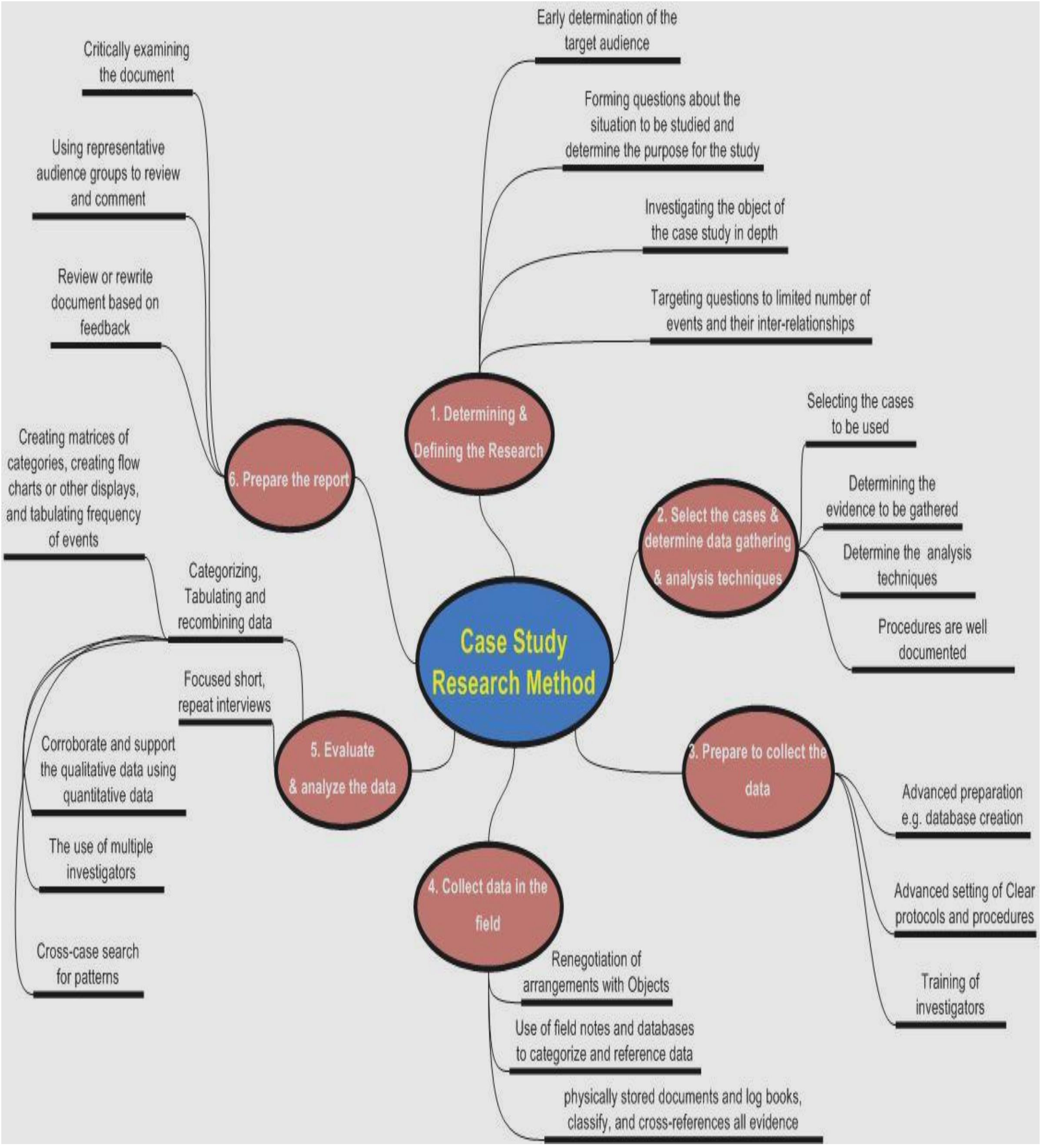


Figure 3.1: Summary of Case Research Method

1. Determine and define the research questions

   As been stated before, the overall aim of this research is to propose that developing a knowledge management (KM) approach provides banks with new opportunities in the decision-making process that supports managing CR. To achieve the aim, several questions have been determined (see 3.1).

2. Select case study

Jordanian banks have been selected as a case study for this research.

3. Prepare to collect data

   Due to the potential complex nature of this research, the multiplicity of the sources and volumes of data, advanced preparation is needed. Clear protocols, procedures have been put in place as well as the preparation of databases for the collection, categorising, sorting, storing and retrieval of data for analysis (see Chapter 5 and 6).

4. Collect the data in the field

   The data that has been collected were entered into a database and others are physically stored but the main documents and log books, classify, and cross-reference all evidence so that it can be efficiently and easily accessible for sorting and examination over the course of the study. Collected data have been stored in formats that were easily accessible and reference-able

5. Evaluate and analyse data

   Data collected have been analysed to discover patterns and relationships in reference to the original research questions. In order to achieve this, a number of processes have been outlined for use. Key techniques involved with data analysis in this research include:

   a) MCA techniques to find out the relative importance of the factors and criteria in this study, a rating algorithm has been used. This method has been used because of its simplicity; high confidence and applicability for our research (see Chapter five).
   b) Frequencies and percentages to describe the selected characteristics of the study participants (see Chapter 5).
   c) Means and standard deviations to describe the status of the factors of KM in the selected Jordanian banks (see Chapter 5).
   d) One-way and Two-way ANOVA models to test the research hypotheses due to their ability to test multiple variables at the same time instead of running two separate tests, but the main reason was their ability to determine whether one variable affects the others (see Chapter 5).

e) Log-linear model to find out the most significant enablers in managing knowledge, because of its ability to analyse and highlight the interactions and interrelationships underlying categorical data (see Chapter 5).

f) Logistic regression analysis to build a CR classification system which will help credit officers in CRM departments in Jordanian banks in improving the decision performance in granting retail loans (see Chapter 6).

g) Different DM techniques have been used (CHAID, QUEST, C&R Tree, C5.0, Bayesian Net, and Neural Networks) to analyse the Jordanian banks' historical data (see Chapter 7).

6. Report

Case studies usually have reports that are easily understood by the public and applicable to their own real-life situations. As a result, boundaries of the case are highlighted as well as conflicting propositions, also a lot of emphasis is placed on displaying enough evidence that suggests that the research has explored all available angles and therefore confidence is inspired in the readers.

However, Barkley (2001) listed a number of advantages and disadvantages of using a case study which are summarised below:

Advantages	Exploratory/Descriptive Analysis, and Hypothesis Testing.
Disadvantages	Research Design, and Biased Responses.

Table 3.1: Advantages and disadvantages of case study approach

- *Advantages of the Case Study Methodology for Research*

1. Exploratory/Descriptive Analysis.

   This is useful in dealing with very new programs/organisations; as in this research (the relation between KM and CRM); where little or no information exists (outside the organisation) regarding its workings and impacts.

2. Hypothesis Testing.

Hypothesis testing is usually carried out to test relationships between the research factors; the use of the case study research method in this instance can show such relationships.

- *Disadvantages of the Case Study Methodology for Research*

    a. Research Design.

    Potential shortcoming sometimes shows up as a result of inadequate review of the relevant literature or not enough attention to the appropriate theoretical model. In this research, an extensive review of the relevant literature has been done to cover all the related areas (see Chapter 2).

    b. Biased Responses.

    This is usually possible when information is gathered through personal interviews which can produce responses that do not properly reflect the situation. To overcome this disadvantage, a questionnaire has been distributed to collect the required data which has been statistically analysed.

3.1.2 Quantitative and Qualitative Approaches

There are two major approaches in research that produce two different types of data:

1. Quantitative research approach – produces Quantitative data
2. Qualitative research approach - produces Qualitative data

Different schools of thought have argued about the superiority of one form of research method over the other but a well balanced research is one that can systematically take advantage of both qualitative and quantitative research methods and data to get the full picture of the object being investigated (Harvey, 2002). Mixed methods research as the third research paradigm can also help bridge the approach between quantitative and qualitative research (Onwuegbuzie and Leech, 2004a).

Using the survey questionnaire, and published literature, this research combines both quantitative and qualitative approaches of research. Using this combined 'mixed' approach, the researcher was able to get the advantages of both and overcome their limitations.

3.2 CRM-KM Approach

Many Knowledge Management (KM) systems have failed to give the results or outputs expected and many of them simply failed. The reason for this is lack of careful understanding of knowledge, KM and its processes (Storey and Barnett, 2000). Some executives are hoping that KM may somehow solve all existing and future problems without defining those problems' causes. Starting a KM system with no aim, regrettably, will most certainly end up in failure. To avoid that, a CRM-KM approach is proposed which contains steps to ensure a successful KM implementation to support managing CR. Using the suggested KM approach will support developing internal models to better quantify their credit risks and improve the quality of information. Additionally, using this approach, beyond the standard CR Models as applied in large banks, will help where there is not enough historic or reliable quantitative data as needed for the CR models. Figure 3.2 describes how to apply this approach.

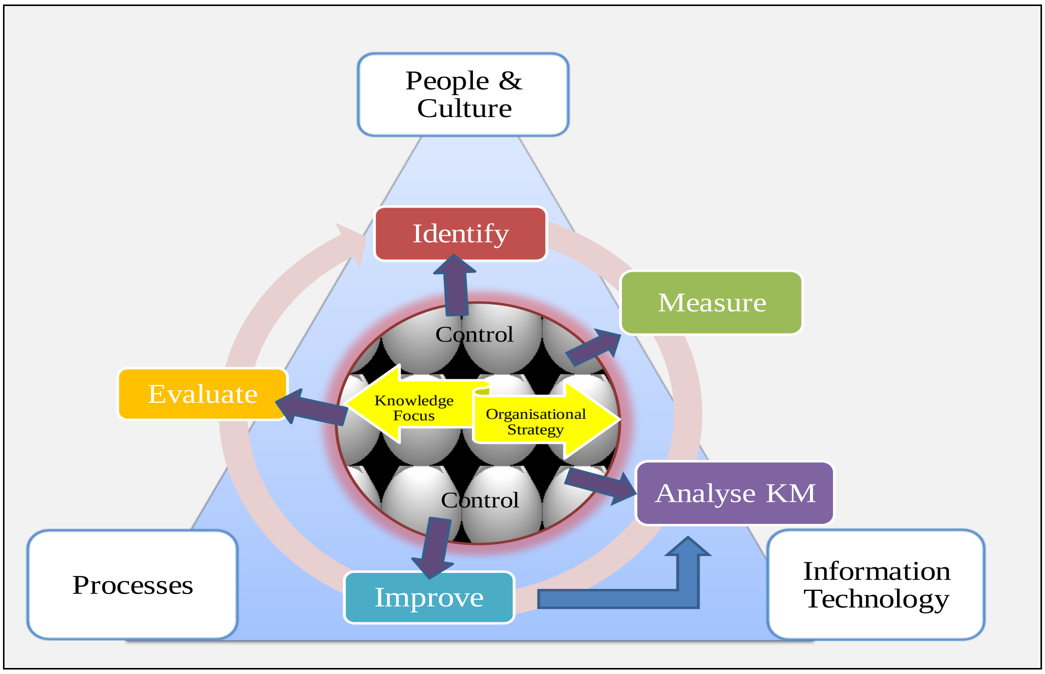


Figure 3.2: CR-KM Approach Cycle

Step 1: Identify

a) Identify a problem: addressing a recognised problem is important as it provides the incentive for employees to implement the KM solution and make it work.

b) Brainstorm and understand the impact of the problem

c) Define a process to investigate the problem

- Choose a proper research methodology (i.e case study, empirical research...)
- Use suitable research methods (i.e surveys and/or questionnaires, interviews...)

Step 2: Measure

- Collect related data
- Measure the current performance: It precisely pinpoints the area causing the problems. It forms the basis of the problem-solving. Subsequently such problems are analysed statistically.

Step 3: Analysis

- In this step, when and where the defect occurs is investigated. Problems are statistically analysed from KM point of view. This includes:

1. Identify the problem's root causes
2. Find out current KM status

- Analyse reasons of the problem

1. Define KM gaps
2. Identify KM barriers and enablers

Step 4: Improve

- Improvements for potential causes, identified in 'Analysis' step 3 are carried out in this step through proposing a KM solution, This includes:

1. Recommend and establish a KM Solution: improvements for potential problem causes identified in step 3 are carried out in this step. The recommended KM Solution needs to be put in the context of the recognised problem
2. Build a culture that accepts the proposed solution, this contains building an environment that facilitates:
    - Knowledge creation
    - Knowledge dissemination
    - Knowledge sharing, and
    - Knowledge implementation

Step 5: Evaluate

1. Collect related data
2. Measure performance against developed deliverables
3. Monitor continuously of whether the improved process is well maintained.

- Control

Control should exist at all steps before, during, and after the KM implementation. This includes activities involved in ensuring a process is predictable, stable, and consistently operating at the target level of performance with only normal variation.

- Knowledge Focus

Knowledge Focus includes managing different types of knowledge. Since explicit knowledge is created by individuals and can be attained from outer channels a knowledge map, lessons-learnt databases, groupware and electronic data interchange can be the appropriate tools for managing this type of knowledge. On the other hand, when it comes to deal with the tacit knowledge then a listing of the experts possessing a tacit knowledge, group meetings, face-to-face conversations and practice forums can be the way of managing tacit knowledge. (Wong et al. 2008)

- Organisational Strategy Focus

All KM processes must be aligned with business strategy and goals. At each step of the framework there should be a possibility to refer to the organisational strategy and goals and ensure that they are integrated in the core of all decisions.

As a guide to implement this CRM-KM approach, Table 3.2 describes its contents:

	Identify	Measure	Analyse	Improve	Evaluate
Objectives	Identify a problem	Define the current performanc e	Identify root causes of the problem	Recommend and establish a KM Solution	Evaluate current KM solution and outputs
Tasks	1. Brainstorm and understand the impact of the problem. 2.Define a process to investigate	Collect the related data	1.Analyse the data and the process map for improvement opportunities	1. Brainstorm KM Solutions to the problem. 2. Involve the users. 3..Build a culture that supports recommended solution, and 4.Reward Contributions.	1.KM Auditing 2. Measure the performance. 3.Collect the related data, and 4. Compare the results before and after.
Techniques	1.Using surveys and/or questionnaires 2. Literature review	Use a scale model	1.Use SPSS (Statistical Software) 2. DM Techniques	1.Hypothesis Testing: test/brainstorm potential solutions. 2.Cost/benefit analysis of proposed solution	1.Use a scale model

Table 3.3: CRM-KM approach contents

3.3 KM-CSF Scale Model

To measure the performance of activities in KM systems and to point out areas in which actions will need to be taken, a KM scale model has been developed (see Chapter four). The developed scale model incorporates both tacit and explicit knowledge as part of its measurement process. The well established CSFs are categorised appropriately so that a measurement process can be established. The techniques of Multi Criteria Analysis (MCA) are used in developing as an effective method for producing weighted factors and criteria. Thus, the overall objective is to provide a method for the measurement of the performance of activities in KM which should lead to identification of areas where action needs to be taken. The scale model has been transformed into a questionnaire (see Chapter five).

3.4 Data

Applying the rules of theoretical sampling, sixteen banks were approached by the researcher. Twelve banks agreed to be part of this study, which represents 75% of the banking community (see Chapter five). Three of those banks have agreed to provide the researcher with historical data of previous retail loans (2755 cases). Those banks dominate55% of the total assets of all banks in Jordan (CBJ, 2006): the Arab Bank, the Housing Bank and the Jordan Bank (see Chapter six and seven).

Chapter 4: Performance Model of Knowledge Management Activities

4.0 Chapter Structure

KM continues to play an important part in an organisation's capability to respond to not only day to day decision-making problems but also in its ability to deal with challenges in a competitive environment. The question whether the measures can be developed to gauge the performance of KM activities is the focus of this Chapter. The assertion that organisational knowledge should be subject to the same degree of control and management as any other organisational function is not in dispute. However, as KM is in a continuous process of defining itself, the techniques for measuring KM performance will therefore need to evolve. In other management areas, critical success factors (CSFs) have been used successfully to monitor performance. With this in mind, several authors have suggested and identified CSFs for KM. Some of the CSFs are qualitative in nature and others can be quantified. It is shown here that the techniques of multi-criteria analysis (MCA) can be used to develop performance indicators for activities in KM.

To date there are no methods available that are able to give performance indicators incorporating both tacit and explicit knowledge in an established framework (see Section 2.7). Thus, one of the contributions of this Chapter is, the developed framework incorporates both tacit and explicit knowledge as part of its measurement process. The well established CSFs are categorised appropriately so that a measurement process can be established. The techniques of Multi Criteria Analysis (MCA) are used in developing as an effective method for producing weighted factors and criteria.

The overall objective is to provide a method for the measurement of the performance of activities in KM which should lead to identification of areas where action needs to be taken. Comparisons of organisations by sector or within a sector can also be achieved by suitable extensions.

This Chapter is divided into three Sections. The first Section is about analysing the current KM measurement models. The second Section illustrates the process followed to build the scale mode, and finally, discussion of the findings is presented.

4.1 KM measurement models

In KM measures to date, there are two broad directions in which developments have taken place: one treats knowledge as intellectual capital (IC) and the other is concerned with the performance of activities critical to the creation and sharing of knowledge. The latter clearly feeds in to intellectual capital.

Category	Model	Researchers	Characteristics	Limitations
Knowledge Asset	Balanced scorecard	Kaplan& Norton(1996),	Clear correlation between indicators and performance.	No external comparison possible. no focus on KM activities
Knowledge Asset	Intellectual capital	Roos *et al.* (1997)	Focus is on monitoring the dynamics of IC. It provides a single index of several indicators based on correlating changes in IC with market changes	Very context specific and limited in universality.
Qualitative	KM Maturity Model	Langen (2000)	Systematic development of KM structures	Not based directly on processes and activities in KM.
Metric based	Knowledge Work Measurement	Smits and Moor(2004)	- Determines the impacts and interrelationships of influence factors	Operative limitations, since numerous different metrics are necessary

Table 4.1: Examples of existing models for measuring KM

4.2 KM-CSF Model

In this Section our method for developing the performance measures is discussed in the following stages (1) Determination of CSFs (2) Establishing a KM framework (3) Determination of indicators (4) Determining KM performance indicators using MCA.

4.2.1 Critical Success Factors

For any business, there are a limited number of areas in which satisfactory results ensure successful competitive performance and those areas are called CSFs. The identification of CSF for KM has been considered many researchers (Davenport et al, 1998; Ryan and Prubutok, 2001; Moffat et al., 2003). Oakland (1995) defined CSFs as what the organisation must have to achieve the mission by examination and categorisation of the impacts. He adds that they are the minimum key factors that the organisation must have and which together will achieve the mission. Digman (1990) defined CSFs as the areas where things must go right for the business to survive.

The most comprehensive list to date from Chong and Choi (2005) includes 11 CSFs. However, in terms of research carried out on CSFs, the following have been identified by various authors. Literature review shows that many researchers have proposed varying CSFs for KM implementation. Belbaly (2005) proposed five factors:

- Organisation culture
- Organisation structure
- People
- Information Technology
- Processes.

Glendon and Kundtz (2000) propose just

- Human.
- Culture and environment.
- Technology.

Helfk *et al.* (2003) have also proposed five and obtained a score for their relative importance as perceived by CEOs:

- Corporate culture: 47.1%
- Structures and processes: 29.8%
- IT: 27.9%
- Skills and motivation: 27.9%
- Management support: 26.9%

For CSFs to be effective in management, they must be few in number. It is suggested that anything from 3 to 6 is optimum. Again, it is argued that the particular choice of CSFs is dependent on the focus and style of the manager in question. From the research analysis, it is proposed that the reported CSFs can be grouped into the following categories (The justification for this is discussed in the next Section):

- People and Culture
- Processes

- Information Technology.

Most of the above CSFs were identified through qualitative research with their importance established through structured analysis. Future research needs to consolidate these factors into a single CSF model (Jennex and Zakharova, 2005). Those authors added that the established CSF model needs to be quantitatively validated against a diversity of organisations which will improve its validity and general application. Chong (2006) states further, "*It is timely that these factors be tested in terms of its level of implementation in organisations. To what extent the KM success factors are implemented in organisations is still elusive. It is equally important to also assess the level of implementation of these factors so that gaps, if any, can be identified. Subsequently, strategies can be proposed on how the gaps can be minimised.*"

4.2.1.1 People and Culture

It is people who create and share knowledge, therefore managing people who are willing to create and share knowledge is important (Belbaly, 2005). Davenport *et al.* (1998) state that companies should ensure that their KM initiatives fit into their organisational culture. It is believed that the direct impact of organisational culture on KM is the willingness for knowledge sharing activities. This includes two dimensions, namely knowledge contributing and knowledge transferring.

Belbaly (2005) identifies three enablers for culture:

1. Collaboration, 2. Learning and 3. Trust

For people, the two enablers are:

1. Team leader characteristics and 2. Team skills

Team leader characteristics have to focus on a democratic and participation style of leadership that encourages the establishment of trust and collaboration (which are indeed culture enablers). Actually, culture is about people and people who formulate the culture. The performance of organisational culture can be measured (in the context of KM) through knowledge contributing and knowledge transferring.

Therefore, it is reasonable and meaningful to put people and culture in one category. To measure the performance of people and culture in implementing KM, the following criteria have been established:

- *Employee training*

Training and education is an important consideration for successful KM. In a basic sense, organisational members need to be aware of the need to manage knowledge and to recognise it as a key resource for the viability of a company (Wong, 2005). Employees must also be provided with training and education of quality which support KM processes and concepts. Yahya and Goh (2002) showed that training related to creativity, team building, documentation skills and problem solving had a positive impact on the overall KM process (Wong, 2005). Therefore, for effective KM, skills development should occur in the following areas: communication, soft networking, peer learning, team building, collaboration and creative thinking (Horak, 2001). Additionally, training for individuals to recognise their new roles for performing knowledge-oriented tasks might be needed. Equally important is to provide them with the skills they need to foster creativity, innovation and knowledge sharing (Wong, 2005).

IT is an important success factor but has no values if employees do not have the proper skills to use it effectively. Therefore, employees should be trained and educated in using the KM system and other technological tools for managing knowledge. This helps to make certain that they can make use of the full capabilities offered by these tools (Davenport *et al.*, 1998).

- *Trustworthy teamwork and employee involvement*

It is suggested as one of the most important CSFs depending on the role that it plays in building the suitable environment and culture for knowledge sharing. Choy and Suk (2005) assert that organisations with team-oriented employees who trust one another are more successful at sharing knowledge than those who are merely technologically superior. When their relationship is high in trust, people are more willing to participate in knowledge exchange. The collaboration strengthens the relationships between the users and establishes an environment of high trust between people that enhances their willingness to participate in the knowledge exchange, and knowledge creation. In many organisations, to ensure participation by employees in knowledge sharing, collaboration and re-use, this requires changing traditional mindsets and organisational culture from knowledge hoarding to knowledge sharing and creating an atmosphere of trust (Hariharan, 2002).

- *Top-management leadership support*:

Researchers view top-management leadership support as one of the most important success factors for a KM system, particularly in knowledge creating and culture sharing activities (Belbaly, 2005). Management leadership plays a key role in supporting the success of KM (Horak, 2001).

According to Belbaly (2005), leadership is responsible for:

1. Creating the knowledge vision of the organisation
2. Communicating the vision
3. Building a culture that regards knowledge as a vital company resource.

4.2.1.2 Processes

KM involves distinct but interdependent processes of knowledge creation, knowledge storage and retrieval, knowledge transfer and knowledge application (Alavi and Liender, 2001). Without question, many of the established and emerging technologies available today hold great potential to facilitate and drive effective KM initiatives. But before an organisation can reap the benefits from employing these tools, there must first be a sound framework developed. One place to begin with this framework development is with process analysis (Glendon and Kundtz, 2000). This includes standard processes for knowledge contribution and content management (Hariharan, 2002).

Besides, these processes must be assessed by workflow diagrams which represent how these processes work and should be implemented within the organisation. Another form of assessment is organisation networks analysis, which can be very useful in understanding the formal and informal processes and interactions in a given organisation. Because the tacit to tacit knowledge transfer process is a social process, its performance can be measured by knowing the density of social networks and the number of social interactions, which is directly correlated with the degree of trust and commitment to the network (Anantatmula and Kanugo, 2005). The two criteria that will be used to map KM processes are

1. Standard processes are established for knowledge contribution and content management, and
2. The organisation has the ability to structure, categorise and access the content of knowledge.

The implementation of KM processes lies at the heart of creating a successful knowledge-based organisation. Thus, it is important that organisations adopt a process-based view of KM (Wong, 2005) and he added that an appropriate interventions and mechanisms need to be in place in order to guarantee that KM processes are addressed in a systematic and structured manner.

4.2.1.3 Information Technology

Beyond the 'soft issues' of process, people and culture, there remains the 'how' of KM. This is where technology begins to show its worth, with appropriate and well-developed processes in place and steps having been taken to develop a sharing, open culture to facilitate KM implementation. IT is the final step that facilitates and speeds the processes that have been built (Glendon and Kundtz, 2000). According to Alavi and Leidner (2001), IT can play a diversity of roles in supporting an organisation's KM processes.

KM technology solutions provide functionality to support knowledge sharing, collaboration, document management, etc. across the enterprise and beyond into the extended enterprise (Hariharan, 2002). IT is widely employed to connect people with reusable codified knowledge and it facilitates conversations to create new knowledge (Belbaly, 2005). According to a survey by Covin (1997), top executives of both Canadian Financial Post 300 firms and US Fortune 500 firms view IT as one of the most critical success factors for KM success (Chong and Choi,2005).

There is a broad collection of IT supports for KM which can be used and integrated into an organisation's technological platform which could be categorised into one or more of the following: business intelligence, knowledge base, collaboration, content and document management, portals, customer relationship management, data mining, workflow, search and e-learning (Wong, 2005). Hariharan (2002) asserts that it is very important to ensure that a robust and user-friendly technology has been put in place to ensure good implementation of KM.

4.3 KM Measurement framework

In this Section the developing of KM measurement framework will be presented which is based on the SECI (Nonaka and Takeuchi, 1995) model combined with the CSFs. The SECI expresses the interplay between tacit and explicit knowledge through a cyclic process of Socialisation, Externalisation, Combination and Internalisation.

4.3.1 Knowledge

In spite of the SECI model, KM research has tended to focus on improving and measuring explicit knowledge. Knowledge creation and application require far more than well-structured knowledge artefacts. A recognition of the importance of employee tacit knowledge is based on the assumption that successful performance improvement may not only depend on how work is organised and the skills of the worker, but on the willingness of employees to convert tacit knowledge into a continuous process of improvement and innovation (Crauise, 1995b); the organisation in KM must therefore also focus on its employees, the real and only repositories of tacit knowledge. According to Borghoff and Pareschi (1997), the early spate of business process re-engineering (BPR) initiatives, where cost reduction was generally identified with the laying-off of people, damaged the tacit knowledge of many organisations.

However, regardless of whether knowledge is tacit or explicit, its contribution must be measurable not only by traditional measures but also by other performance measurements. Lang (2001) asserts that successful KM initiatives rely more on interpersonal interaction and social relationships than the technology itself. It is in this way that tacit knowledge becomes expressed, shared and documented.

The best model which expresses these ideas is the well-known SECI (Socialisation, Externalisation, Combination and Internalisation) model of cyclical knowledge creation of Nonaka and Takeuchi (1995), the focus of which is on the interplay between tacit and explicit knowledge in the organisation, leading to the process of knowledge conversion, expansion and innovation (see Figure 4.1). Under this view, KM can be explained as management of the environment that makes knowledge flow through all the different phases of its life-cycle (Borghoff and Pareschi, 1997).

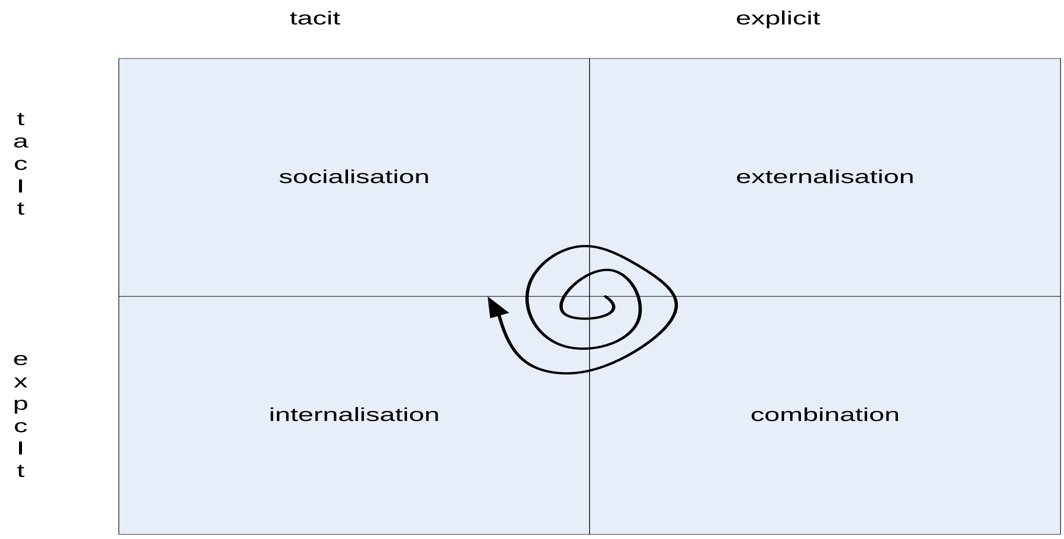


Figure 4.1: SECI spiral model of Nonaka and TakeuchiSource: Nonaka and Takeuchi (1995)

4.3.2 KM Performance New Model

On the basis of the above discussion, the framework model takes the following schematic form (Table 4.2):

CSFs / Knowledge	People and Culture	Processes	Technology
Internalisation Tacit Socialisation	Indicators	Indicators	Indicators
Externalisation Explicit Combination	Indicators	Indicators	Indicators

Table 4.2: SECI-CSF Performance model

The initial set of indicators to be included is based on literature review .Three Tables have been created to show the expanded version of the schematic diagram including

CSFs and the initial set of indicators (see appendix F). The following Section is a discussion of Multi Criteria Analysis (MCA) as a useful technique for implementing our performance model (see Appendix D for full details).

4.4 Multi Criteria Analysis

Multi Criteria Analysis (MCA) is a decision-making tool (for guidelines see Mendoza et al, 1999) developed for complex decision-making problems that include qualitative or quantitative factors. A key characteristic of MCA is its stress on the judgement of the decision-making team in establishing objectives and criteria, estimating relative importance weights and, to some extent, in judging the contribution of each option to each performance criterion (Communities, 2009). The implementation of MCA can be a top-down or a bottom- up process.

However, at the initial stage of the implementation it can only be a top-down process. As MCA continues to be used within the organisation modifications can be incorporated in the performance model from the evidence gathered at ground level. In effect a combination of top-down and bottom-up process begins to take place.

In order to set up a composite measure of performance across all the factors and criteria selected, they must be weighted according to how important each is regarded in relation to the others. To do that, there are different methods available:

1. Ranking: The simplest method for estimating the relative importance of weights for factors and criteria under consideration. In this method, each decision maker allocates a rank for each factor or criterion depending on his/her preference. Because of its simplicity, the ranking method is very attractive. One of the disadvantages is its inappropriateness when the number of factors or criteria is large (Mendoza et al., 1999; Kleindorfer, 1993).
2. Rating: Here, the decision maker allocates weights ranging from 0 to 100 for each factor/criterion. 0 indicates low importance while 100 is assigned to the most important factor/criterion. The sum of weights must equal 100. One of the disadvantages of this method is the difficulty of justification for the assigned weights (Mendoza et al., 1999; Kleindorfer, 1993).
3. Pairwise comparison: This is another method for determining the weights for the factors and criteria under consideration and can convert the relative importance into a linear set of weights (Saaty, 1999; Heywood et al., 1993). It involves the

comparison of the factors and criteria and allows the comparison of only two criteria or factors at once.

Features	Rating	Ranking	Pairwise Comparison
Response Scale	Interval	Ordinal	Ratio
Hierarchical	Yes	Yes	Yes
Confidence	High	Low	High

Table 4.3: Comparison between methods of estimating weights

4.4.1 Mathematical Specification of Multi-criteria Analysis Model, structured by a Criteria and Indicators' matrix

Table 4.4 represents the principles used in the KM-CSF Model, they are three:

1. People and Culture
2. Processes
3. Information Technology (IT)

Also, the Table shows the eight criteria used in the model:

1. Employee training
2. Employee involvement and teamwork
3. Top management leadership & commitment
4. Employee awareness about knowledge inside the organisation
5. Standard processes for knowledge contribution and content management
6. Organisation ability to structure, categorise and access the content of knowledge
7. Robust and user friendly technology
8. Tools for managing knowledge cycles activities

Principle	Criteria

1. People & Culture	1.1 Employee training 1.2 Trustworthy teamwork and employee involvement 1.3 Top-management leadership support 1.4 Employee awareness about knowledge inside the organisation
2. Processes	2.1 Standard processes established for knowledge contribution and content management 2.2 Organisation has ability to access, structure, and categorise content of knowledge
3. IT	3.1 Robust and user-friendly technology available to employees put in place 3.2 Tools establishedfor managing knowledge cycles activities

Table 4.4: SECI-CSF Performance model Factors and Criteria

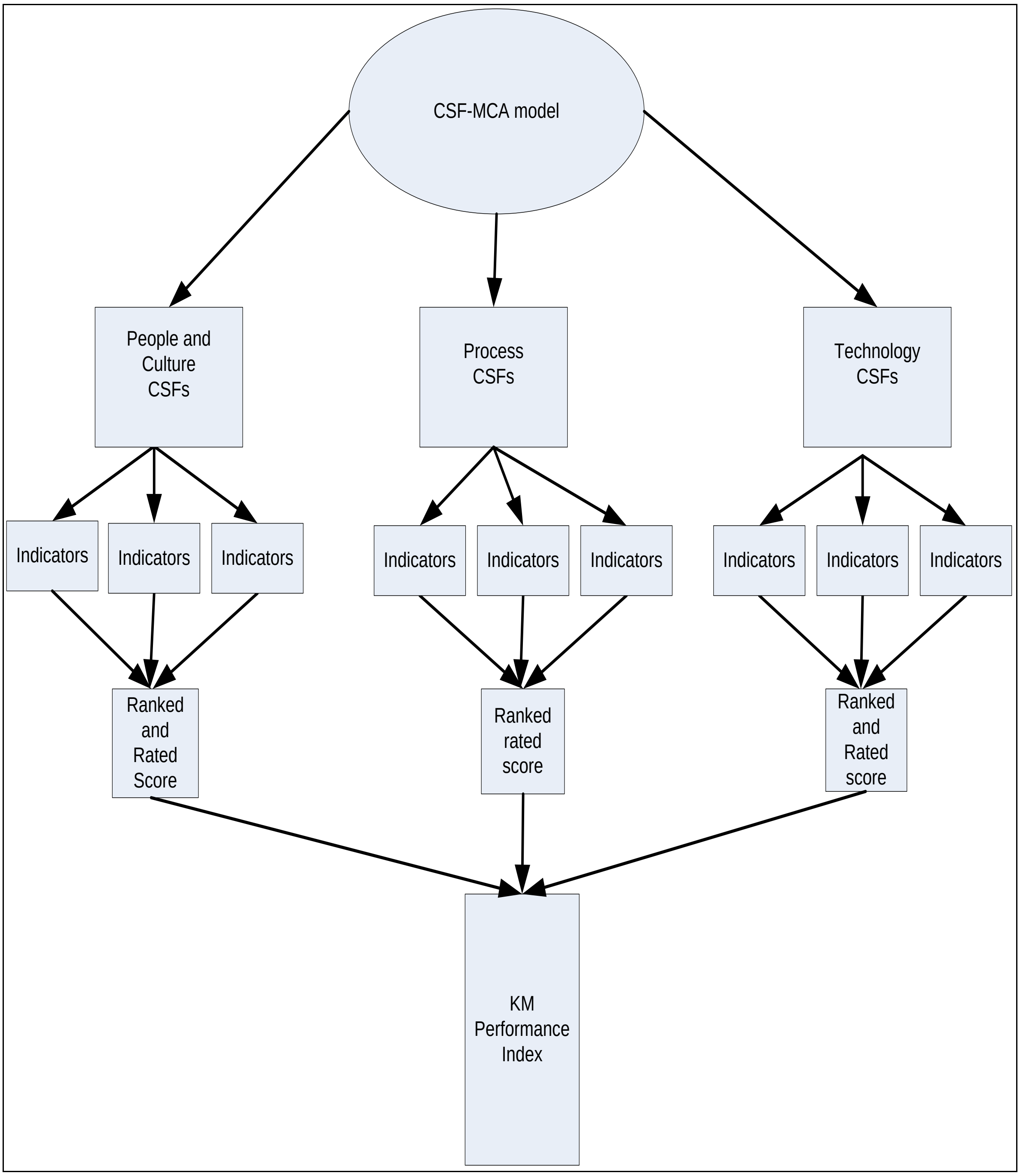


Figure 4.2: KM-CSF Implementation model

The steps involved in generating the KM performance index are as follows

1. The process begins with the initial set of factors and criteria being ranked and rated by the first team of experts. The score obtained from this forms the basis for weights attached to the factors and criteria.

2. In the second step, scores are given for each of the indicators by the employees of the particular business unit under consideration.

The KM Performance index is obtained by normalising the score, using the maximum possible to produce a suitable index range. The process is cyclic and adjustments to weights are necessarily occurring based on the experience at ground level. The implementation of the above model can be achieved in its simplest form using a spread sheet such as MS-Excel with the feature of displaying the result in graphical form.

4.5 Conclusion

The possibility that a performance measurement model for KM consistent with both known CSFs and our current understanding of the role of knowledge in organisations can be developed in an evolutionary way has been demonstrated in this work. The use of MCA technique as an effective method for producing weighted factors and criteria has been explored and shown to be very useful.

The model produced is a dynamic model in the sense that it can be adapted to the changing situation within the organisation as well as being transportable to different organisational environments.

The model is very dependent on the CSFs chosen. The fact that these are subjective need not be regarded as a negative feature of the model. As explained in the text, MCA is both a top-down as well as a bottom-up process. This means that experience at ground level will change not only the choice of CSFs but also features such as the ranking and rating of factors used.

The objective of the proposed model is to measure the performance of activities in KM systems and to point out areas in which actions will need to be taken. This model however gives only a snapshot at a given time of the performance of KM activities over a specified period for the particular strategic business unit under consideration. The extension of the model's availability to all organisations within a sector would require the establishment of acceptable core indicators. This can only be achieved by implementation practice over a period of time.

Chapter 5: Knowledge Management Practices in Credit Risk Management Departments of Banks in Jordan

5.0 Chapter Structure

Arguably now, even more than in normal times, it is crucial for banks to leverage their knowledge resources so that they are able to respond to deal with the undoubted major strategic challenges that exist. This Chapter reports on a survey of the current performance of KM practices in CRM departments in Jordanian banks, and determine the relationship between KM and CRM. The secondary objective is to compare the performance of KM practices between commercial and Islamic banks.

Data have been collected using a survey questionnaire administered to staff at all levels in twelve banks in Jordan. The construction of the questionnaire has been based on the concepts and framework of dynamic Knowledge Management/Critical Success Factors (KM-CSF) created in Chapter 4. The analytical approach is based on Multi-Criteria Analysis (MCA), with the objective of producing vertical and horizontal hierarchies of Bank-KM performance indicators/indices, as well as an overall summary Bank-KM performance indicator. The various indicators for the twelve banks provide a multi-level picture of the KM-performance of the banks, which can be used either as a basis for a year on year monitoring of Bank-KM performance or to allow a bank to compare itself with its competitors.

This Chapter is divided into six Sections. The first Section is an introduction. The following Section presents (KM-CSF) Model (see Chapter four) that will be used in this Chapter as a questionnaire. Section 5.3 discusses the research methodology used in this Chapter. This is followed by Section 5.4 which presents the two questionnaires that have been used to get the results. The penultimate Section deals with the results and finally the discussion of the results.

5.1 Introduction

Risk Management is a knowledge domain which is used at the centre of each fiscal organisation and which covers all operations having an impact on its risk profile. CRM in banks uses a combination of key factors. Knowledge is a key to this and as a result the quality of information and knowledge must be at very high standards.

It is becoming increasing clear, however, that in banks the sharing of knowledge amongst the senior executives has not been as effective as it should have been. The lack of knowledge amongst senior executives about the level of risks taken in sub-prime lending, the resulting ‘toxic assets’ and the global nature of the instruments used to spread risks is said to be the main contributing reason for the current worldwide crisis in banks. Banks in Jordan, the focus of this study, are not immune from the exposure to the risks. The question that naturally follows is how well are banks in Jordan doing with respect to sharing of knowledge, especially amongst their senior executives?

5.2 Multi-Criteria Analysis-Critical Success Factors (KM-CSF) Model

The objective of this model (see Chapter 4) is to provide a method for the measurement of the performance of KM practices which should lead to identification of areas where action needs to be taken. The first step in the MCA-CSA model is to ask respondents, usually key staff, to assign weights for each factor and criterion in the model. The next step is to build a scoring system completed by all respondents that reflects the performance of the organisation in implementing its KM practices. Table 5.11 lists the factors and criteria used in this study.

Factor	Criteria
1. People & Culture (F1)	1.1 Employee Training (C1) 1.2 Trustworthy teamwork & employee involvement (C2) 1.3 Top-management leadership support (C3) 1.4 Employee awareness about knowledge inside the organisation (C4)
2. Processes (F2)	2.1 Standard processes have been established for knowledge contribution and content management (C5) 2.2 Organisation has ability to access, structure and categorise content of knowledge (C6)

3. Information Technology (F3)	3.1 Robust and user friendly technology available to employees put in place (C7) 3.2 Tools for managing knowledge cycles activities established (C8)

Table 5.1: KM-CSF Model Factors and Criteria

5.3 Research Methodology

5.3.1 Research questions and hypotheses

This study attempts to answer the following questions:

RQ1: What is the overall performance of KM practices in Banks in Jordan?

RQ1.1: What is the contribution of the 'People & Culture' factor in supporting KM in banks in Jordan?

RQ1.2: What is the contribution of the 'Processes' factor in supporting KM in banks in Jordan?

RQ1.3: What is the contribution of the 'IT' factor in supporting KM in banks in Jordan?

RQ2: What are the enablers of KM in banks in Jordan?

Based on the stated purpose and the questions mentioned above, the following hypotheses are formulated (corresponding null hypotheses are the converse):

H1. There is a significant impact of the factors 'Processes and IT' on the factor 'People and Culture'

H2. There is a significant difference between the Jordanian commercial banks and Islamic banks in the performance of KM practices

H3. There is a significant relationship between KM and CRM in Jordanian banks

5.4 Questionnaires

There were two questionnaires in this survey. First, a weighting questionnaire was distributed to assess the weights for each of the three factors and the eight criteria. A second questionnaire was then distributed to obtain scores for the listed KM practices. The questionnaires had a good response, largely because of the close involvement of the managers at the various banks.

5.4.1 Relative-Importance Questionnaire Data Analysis

To find out the relative importance of the factors and criteria in this study, a rating algorithm has been used. This method has been used because of its simplicity, high confidence and applicability for our research. MCA techniques usually apply numerical analysis to a performance matrix in two stages:

1. Weighting: Numerical weights are assigned to identify, for each criterion, the relative valuations of a shift between the top and bottom of the chosen scale (Communities, 2009).
2. Scoring: Expected consequences of each option are assigned a numerical score on strength of preference scale for each option for each factor or criterion. The scores for all elements being compared under each factor or criterion must add up to100. More favoured options score higher on the scale and less favoured score lower. In practice, scales extending from 0 to 100 are often used, where 0 represents a real or hypothetical least favoured option and 100 is associated with a real or hypothetical most favoured option. All options measured in the MCA would then fall between 0 and 100.

5.4.1.1 Results

Table 4.2 shows the weights of the perceived importance of the various KM factors and criteria and that the highest weight was allocated to *People and culture* with 41.7%. The perceived importance of criteria measuring this factor is as follows:

The highest perceived importance (weight) was allocated to *Employee involvement & teamwork in managing knowledge* (C2=31.4%), *Top-management leadership support* (C3=27.3%), *Employee training* (C1=24.2%), and *Employees awareness about knowledge inside the organisation* with the lowest weight (C4=17.1%).

The second important factor is *Information technology* (F3=29.74%). *Organisation has the ability* comes first (C6=50.15%) and in second place *Standard processes*

(C5=49.85%). The third important factors were weighted similarly to IT (F2=28.61%). In this group, the *Robust & user-friendly technology* criterion comes first (C7=53.6%), and *Tools for managing knowledge cycles activities* comes next (C8=46.4%).

Factor	Perceived Importance	Criterion	Perceived Importance
1. F1	41.64%	C 1.1 C 1.2 C 1.3 C 1.4	24.2% 31.4% 27.3% 17.1%
2.F2	28.61%	C 2.1 C 2.2	49.85% 50.15%
3. F3	29.74%	C 3.1 C 3.2	53.6% 46.4%

Table 5.2: Factors and criteria perceived importance (weights)

5.4.2 KM-CSF Questionnaire

The purpose of this study is to measure the overall performance of KM practices in the various banks in Jordan and also validate the KM-CSF model (see Chapter 4). The second questionnaire was used to obtain scores for the various factors and criteria in the KM-CSF model and each of the questions was directly related to its indicators.

5.4.2.1 Sampling and data collection

All Jordanian banks were included in the survey setting. The study concentrates on studying the Jordanian banks only. According to the CBJ (2007) report, in Jordan there are fourteen local commercial and two Islamic banks. The sixteen banks were approached by the researcher and twelve agreed to be part of this study, which represents 75% of the banking community. The targeted respondents of this questionnaire were employees in CRM departments in Jordanian banks. They were selected because the study is about developing a KM approach to support managing CR.

The banking sector in Jordan is considered as one of the biggest in Jordan and CBJ (2007) estimated that it represents 31% of the total investment in the country. The GDP in May 2007 was 11,225.3 millions JOD, the calculated banking services were 360.4

millions JOD and a vast network of branches covered about 11,900 persons per branch. The three largest banks account for 55% of the total assets. The Arab Bank dominates the sector with 29% of all assets and the Housing Bank is the second largest, with the most extensive branch network.

The authors handed 320 questionnaires to the various head offices and then the branch managers forwarded the questionnaires to targeted officers. The questionnaires were distributed to officers involved in CRM, namely branch managers, senior CRM officers and credit officers. The number of questionnaire to be distributed in each bank was determined according to the banks' wishes after meeting the authorised people either in the HR or the PR department. From the 300 responses, 68 questionnaires were excluded from the study because of missing data. The remaining 242 responses represent an effective response rate of 75.6 % (see Table 5.3).

No	Bank Name	Number of Distributed Questionnaires	Number of respondents	Response Percentage
1	Arab Bank	37	31	83.8%
2	Cairo Amman Bank	30	22	73.3%
3	Jordan Ahli Bank	25	19	76%
4	Jordan Bank	25	19	76%
5	Jordan Commercial Bank (Gulf Bank)	25	16	64%
6	Jordan Islamic Bank	30	24	80%
7	Jordan Kuwait Bank	25	18	72%
8	Housing Bank for Trade and Finance	30	22	73.3
9	Industrial Development Bank	20	17	85%
10	Arab Islamic Bank	20	15	75%
11	Societe General Bank-Jordan	25	19	76%
12	Union Bank for Saving and Investments	28	22	78.6%
	Total	320	242	75.6%

Table 5.3: Percentage Response of Observed Banks

5.4.2.2 Data Analysis

The scoring system in the questionnaire was based on a 5-point Likert-scale with the values (1) strongly disagree, (2) disagree, (3) neutral, (4) agree and (5) strongly agree. The English version of the questionnaire was translated into Arabic to ensure that all of the participants understood the items perfectly.

The draft questionnaire was pilot tested with a selected sample of staff of banks in Jordan. Feedback from the pilot round resulted in minor modifications to the survey questionnaire.

The Statistical Package for Social Sciences (SPSS) version 16 (2007) was used to analyse the collected data because of its included features which were very important to achieve the main objectives of this research. The following statistics were used:

1. Frequencies and percentages to describe the selected characteristics of the study participants.
2. Means and standard deviations to describe the status of the factors of KM in the selected Jordanian banks.
3. One-way and Two-way ANOVA models to test the research hypotheses due to their ability to test multiple variables at the same time instead of running two separate tests, but the main reason was their ability to determine whether one variable affects the others.
4. Log-Linear model to find out the most significant enablers in managing knowledge, because of its ability to analyse and highlight the interactions and interrelationships underlying categorical data.

5.4.2.3 Reliability of measures

The main goal of the reliability test is to ensure that the instrument and its questions will achieve the objectives of the study. In addition, the reliability test is considered as one of the basic criteria used for evaluating the accuracy and precision of the research work.

Cronbach's alpha allows measurement of the reliability of different variables. This coefficient varies from 0 to 1 and a value of 0.6 or less generally indicates poor internal consistency reliability (Malhotra, 2003, p.268). The questionnaire proved to be average (0.83).

5.5 Results

5.5.1 Population Analysis

Table 5.4 summarises the sample characteristics. Since the sample was fairly extensive in the banks studied, it is taken to be representative of the population of bank employees included in the study.

Characteristics		Frequency	Percentage
Age (years)	20-29	111	45.9%
	30-39	76	31.4%
	40-49	49	20.2%
	50-59	5	2.1%
	60 or over	1	0.4%
Gender	Male	143	59.1%
	Female	99	40.9%
Educational level	High School	7	2.9%
	Diploma	54	22.3%
	Bachelor	166	68.6%
	Master	2	0.8%
	PhD	13	5.4%
Position	Credit Officers	130	53.7%
	Senior Managers	104	43%
	Branch Managers	8	3.3%
Experience (years)	1-4	106	43.8%
	5-9	26	10.7%
	10-15	57	23.6%
	15-19	30	12.4%
	20 or more	23	9.5%
Pre-Experience (years)	Less than 5 years	208	85.95%
	5 years or more	34	14.05%

Table 5.4: Characteristics of sample (n=242)

Table 5.4 shows that, the majority of respondents were males (59.1%), while 40.9% were females. This indicates that women working in the banking industry in Jordan are filling banking positions more than in many other developing countries (Siam, 2007). The majority of the employees in banks in Jordan are young (under 39). This is because CRM is a new activity and is mostly implemented in a separate department.

The qualification attribute in Table 5.4 shows that 68.6% of respondents have a bachelor's degree. As far as position is concerned, 53.7% of respondents are credit officers, 43% are senior managers and 3.3% are branch managers.

5.5.2 Results of Testing Research Questions and Hypotheses

As Shaw (2005) states, KM is a discipline that can contribute positively to risk management, especially in risk analysis and risk knowledge sharing. In order to answer the first research question (What is the overall performance of KM practices in banks in Jordan?), means and standard deviations were calculated for the questionnaire items. The means were then classified into three levels:

1. High: mean equal to or more than 3.5
2. Average: mean 2.50-3.49
3. Poor: mean less than 2.50

RQ1: What is the overall performance of KM practices in banks in Jordan?

Table 5.5 shows the weighted average scores for KM and its factors. In the Table, we can note that the factor F2 (Processes) scores highest (3.546), while factor F1 (People & Culture) scores lowest (3.37). However, the weighted average for all factors was relatively close. Finally, the overall performance of KM practices is average (3.446). This suggests that banks in Jordan have considerable room for improvement. Clearly the importance of leveraging knowledge in the current climate has already been emphasised. In later Sections, the areas where efforts need to be addressed will be identified. This will be done by analysing different criteria related to each factor.

Factor	Weighted Average (max 5.0)	Weighted Average (%)
People & Culture	3.37	67.4%
Employee training	3.282	65.64%
Employee involvement	3.286	65.72%
Employee aware about knowledge	3.507	70.14%
Visible top management	3.424	68.48%
Processes	3.546	70.92%
Standard processes	3.5	70%
Organisation's ability to access knowledge	3.591	71.82%
Information Technology	3.46	69.2%
Robust & friendly technology	3.647	72.94%
Tools for managing knowledge	3.246	64.92%
KM Performance	3.446	68.92%

Table 5.5: Weighted average for KM and its factors

RQ1.1: What is the contribution of 'People & Culture' in supporting KM in Banks in Jordan?

The 'People & Culture' factor contains four criteria. To find out the contribution of this factor in supporting KM and identifying the gaps, each criterion has been analysed separately.

Criterion 1: Employee Training

Table 5.6 shows the results of frequency analysis for the *employee training* that supports tacit knowledge. The weighted average for respondent answers was approximately 68%, indicating that the contribution of the criterion is acceptable, but not very high. However, this can be explained by the fact that 40.5% and 27% of respondents answered "strongly disagree" on the first item, "The specialised training offered before starting my current job was very useful", and the second, " I am satisfied with the IT training given for my job requirements ". Accordingly, we can conclude that the specialised training courses before starting the job are not useful and should be restructured to increase their efficiency. In addition, more attention should be given to IT training with emphasis on job-related training.

No.	Item	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
8	The specialised training offered before starting my current job was very useful	13.5%	39.6%	4.5%	1.8%	40.5%
10	I am satisfied with the IT training given for my job requirements	9.9%	45.9%	15.3%	1.8%	27.0%
12	I am satisfied with the IT training given to me when new software is installed	12.6%	55.0%	13.5%	0.9%	18.0%
13	I am satisfied with my training plan as provided by the Human Resource Department	11.7%	52.3%	21.6%	13.5%	0.9%
15	My management supports and encourages me to attend seminars and conferences that relate to my work	18.9%	46.8%	23.4%	9.9%	0.9%
16	The bank provides me well structured training to recognise knowledge that is valuable to my work.	13.5%	51.4%	29.7%	5.4%	0.0%
Average		13.35%	48.50%	18.00%	5.55%	14.55%
Weighted average (max. 5.0) *		3.404				
Weighted average (%)		68.08%				

Table 5.6: Employee Training Supporting Tacit Knowledge

* Weighted Average calculated by multiplying average for each answer category by weight of each category, and then summing. Categories given following weights: (strongly agree = 5, agree = 4, neutral = 3, disagree = 2, strongly disagree = 1).

On the other hand, Table 5.7 shows the results of frequency analysis for the *employee training* criterion that supports explicit knowledge. Based on the weighted average for responses (68%), very close to the overall weighted average, we can conclude that the contribution of employee training that supports explicit knowledge in the Jordanian banks is acceptable, but not very high. Since 28.8% were not satisfied with the relevancy of the presentations given to them on their work, this explains the need for more presentations and workshops relevant to work in the banking sector.

No.	Question	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
17	The presentations given to me by the bank on global case studies are very relevant to my work	6.3%	24.3%	32.4%	28.8%	8.1%
18	I am satisfied with my training on how to extract knowledge from my bank's knowledge base	8.1%	42.3%	27.0%	21.6%	0.9%
Average		7.20%	33.30%	29.70%	25.20%	4.50%
Weighted average (max. 5.0)		3.132				
Weighted average (%)		62.64%				

Table 5.7: Employee Training Supporting Explicit Knowledge

Criterion 2: Employee Involvement

Table 5.8 shows the results of frequency analysis for the *employee involvement* criterion that supports tacit knowledge. The weighted average for responses was 66% approximately, which is acceptable, but not very high. The reason is that a significant percentage of respondents (about 32% on average) answered "neutral" on the relevant items. This weakness can be overcome by encouraging team work and presentation attendance, motivating social activities, attracting external experts and increasing the level of employee involvement in meetings and decision making.

No.	Question	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
19	I am satisfied with teamwork inside the bank	12.6%	55.9%	22.5%	6.3%	2.7%
20	I am satisfied with the attendance in group's presentations	5.4%	58.6%	25.2%	9.9%	0.9%
21	I am satisfied with the social activities in the bank	4.5%	21.6%	33.3%	28.8%	11.7%
22	I am satisfied with the amount of hours I spend with external experts per month	5.4%	27.0%	38.7%	17.1%	11.7%
23	I am satisfied with regular meetings between departments to discuss market trends and developments	5.4%	22. 5%	39.6%	22.5%	9.9%
36	The bank involve the employees in a decision making process	8.1%	34.2%	29.7%	23.4%	4.5%
Average		6.90%	39.46%	31.50%	18.00%	6.90%
Weighted Average out of five *		3.297				
Weighted average as a percentage		65.95%				

Table 5.8: Employee Involvement to Support Tacit Knowledge

With respect to the *employee involvement* criterion that supports explicit knowledge, the results in Table 5.9 show that the weighted average was approximately 65%. This is close to the weighted average for the employee involvement tacit knowledge criterion. The weakness in this factor is that Jordanian banks do not support their employees in their research activities. According to Lee (2000), the number of published patents reflects the quality of tacit knowledge of employees in an organisation and encourages them to develop their skills, which will be reflected positively in their ability to accomplish tasks. In the CRM context, this will be reflected positively in credit officers' abilities to estimate the CR and find effective solutions to manage it successfully.

No.	Question	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
24	I am satisfied with regular meetings within my department	9.9%	55.0%	19.8%	13.5%	1.8%
25	I am satisfied with the number of shared reports produced	9.9%	42.3%	27.9%	16.2%	3.6%
26	I am satisfied with the number of patents that we published	5.4%	13.5%	42.3%	28.8%	9.9%
Average		8.40%	36.93%	30.00%	19.50%	5.10%
Weighted average (max. 5.0) *		3.238				
Weighted average (%)		64.77%				

Table 5.9: Employee Involvement to Support Explicit Knowledge

Criterion 3: Visible Top Management

Table 5.10 shows the results of frequency analysis for the *visible management support for activities* criterion that supports tacit knowledge. The weighted average for respondents' answers was 67% approximately, which is acceptable, but not very high. The reason is that top management do not encourage employees to share their knowledge. This gap could be narrowed through:

1. Knowledge communities. The Jordanian banks should establish an environment providing a friendly and effective communication channel and emphasising more the sharing of knowledge and explicitly, so that the employees will be more willing to share and apply new knowledge with each other (Glendon and Kundtz, 2000).
2. Reward system. This reward system will encourage employees to form a sharing culture (Goh, 2002). The banks should take steps to promote a trust culture by establishing an incentive system for sharing knowledge between employees (Kim, 2004).

No.	Question	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
27	Managers encourage sharing knowledge between employees	12.6%	39.6%	28.8%	14.4%	4.5%
28	Top management gives appropriate rewards to motivate employees to share knowledge	9.0%	27.0%	38.7%	18.9%	6.3%
29	I am satisfied with the work space that I have been allocated in the bank	13.5%	45.0%	18.0%	18.9%	4.5%
63	IT managers have built strong interpersonal relationship with employees	9.0%	50.5%	24.3%	14.4%	1.8%
Average		11.03%	40.53%	27.45%	16.65%	4.28%
Weighted average (max. 5) *		3.372				
Weighted average (%)		67.43%				

Table 5.10: Visible Top Management to Support Managing Tacit Knowledge

With respect to the *visible management* criterion that supports explicit knowledge, the results in Table 5.11 show that the weighted average was approximately 62% lower than the weighted average for *tacit visible management*. The weakness in this factor is the lack of presentations and workshops relevant to work and lack of support for research which affects negatively the creating of explicit knowledge needed for managing CR.

No.	Item	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
17	The presentations given to me by the bank on global case studies are very relevant to my work	6.3%	24.3%	32.4%	28.8%	8.1%
30	Top-management supports research and development well	9.9%	32.4%	35.1%	18.9%	3.6%
Average		8.10%	28.35 %	33.75%	23.85%	5.85%
Weighted average (max. 5.0) *		3.087				
Weighted average (%)		61.74%				

Table 5.11: Visible Top Management to Support Managing Explicit Knowledge

Criterion 4: Employee awareness about knowledge

As shown in Table 5.12, the weighted average for responses on the *employee awareness* criterion that supports tacit knowledge was relatively high (about 73%). The major weakness here can be summarised as the lack of references and availability of knowledge related to the work. Improving these two aspects will increase the credit officers' abilities in finding the knowledge they need and then implementing it in a required task.

No.	Item	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
14	I am satisfied with the reference documents available for me in relation to my work	9.9%	47.7%	25.2%	16.2%	0.9%
31	I am satisfied with knowledge availability	6.3%	54.1%	24.3%	15.3%	0.0%
34	I know exactly who to ask when I need information for a specific task	25.2%	54.1%	14.4%	4.5%	1.8%
Average		13.80%	51.97%	21.30%	12.00%	0.90%
Weighted average (max. 5.0) *		3.657				
Weighted average (%)		73.13%				

Table 5.12: Employee awareness about knowledge (tacit)

However, the results for the employees' *awareness about knowledge* criterion that supports explicit knowledge, as shown in Table 5.13, were lower than those for tacit awareness, with an average of 66%. The contribution of this criterion could be improved by improving the banks' knowledge base and making access to explicit knowledge easier. This will encourage credit officers in banks in Jordan to use the knowledge-base more frequently (implementing knowledge about customers in the decision-making process in granting loans).

No.	Item	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
32	I use the knowledge-base frequently	13.5%	45.9%	24.3%	12.6%	3.6%
33	I call the help-desk frequently	4.5%	18.0%	30.6%	35.1%	11.7%
35	I use knowledge about customers to do my tasks	11.7%	57.7%	21.6%	6.3%	2.7%
Average		9.90%	40.53%	25.50%	18.00%	6.00%
Weighted average (max. 5.0) *		3.301				
Weighted average (%)		66.03%				

Table 5.13: Employee Awareness about Knowledge (Explicit)

In general, the contribution of 'People & Culture' in supporting KM in banks is found to be average (M=3.3, SD=0.57), but this factor is important in KM (see Section 5.4.1). Regarding the contributing criteria, Table 5.14 indicates that training of employees in Jordanian banks is average (M=3.28, SD=0.82). Sharing of knowledge on risk is only just acceptable (M=3.29, SD=0.64). To overcome this problem, the Jordanian banks need to facilitate knowledge sharing by building an environment that supports employees to share their knowledge explicitly. This would suggest the use of Communities of Practice and Social Networks as a solution for this problem.

	N	Min	Max	Mean	Std. Deviation
C1: Employee training	242	1.33	5.00	3.282	.82719
C2: Employee involvement	242	1.67	5.00	3.286	.66720
C3: Top-management support	242	2.50	4.33	3.424	.49530
C4: Employee awareness	242	1.50	5.00	3.501	.78004
F1: People & Culture	242	2.16	4.71	3.370	.56822
Valid N	242				

Table 5.14: Performance of People & Culture

RQ1.2: What is the contribution of the 'Processes' factor in supporting KM in banks in Jordan?

KM processes act as a means to provide new knowledge and give employees the opportunity to apply existing CR knowledge. The 'Processes' factor contains two criteria. To find out the contribution of this factor in supporting KM and identifying the gaps, each criterion has been analysed separately

Criterion 1: Standard Processes

Table 5.15 shows responses on the *standard processes* criterion that supports tacit knowledge. The weighted average for answers was acceptable (about 67%), but not very high. This may be due mainly to the lack of work breaks or because there are no breaks at all. According to Moor and Smits (2002), more than 25% of the shared knowledge between employees occurs within work breaks.

No.	Item	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
37	By following routine business processes I gain knowledge about the business.	14.4%	68.5%	11.7%	5.4%	0.0%
38	Peer evaluation/review and collaboration are an established practice in the bank	18.0%	47.7%	22.5%	8.1%	3.6%
39	The bank provides enough time for work breaks	5.4%	18.9%	21.6%	29.7%	24.3%
Average		12.60%	45.03%	18.60%	14.40%	9.30%
Weighted average (max. 5.0) *		3.370				
Weighted average (%)		67.41%				

Table 5.15: Standard Processes Supporting Tacit Knowledge

Table 5.16 shows the results of frequency analysis for the *standard processes* criterion that supports explicit knowledge. The weighted average for responses (70%) was higher than that for tacit standard processes. However, the explicit standard processes have a main shortfall in documenting the successful and unsuccessful services, besides a gap in creating clear standard processes for retrieving knowledge that could be used in a mathematical modelling.

No.	Question	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
40	Standard processes for retrieving knowledge (used in mathematical modelling) are defined clearly	9.0%	50.5%	31.5%	8.1%	0.9%
41	Organisational knowledge is updated regularly	9.9%	51.4%	32.4%	6.3%	0.0%
42	Case notes on successful and unsuccessful services are documented and archived	5.4%	36.0%	41.4%	14.4%	2.7%
Average		8.10%	45.97%	35.10%	9.60%	1.20%
Weighted average (max. 5.0) *		3.501				
Weighted average (%)		70.01%				

Table 5.16: Standard Processes to Support Explicit Knowledge

Criterion 2: Organisation's ability to access knowledge

Table 5.17 indicates that the weighted average for Jordanian banks' ability to *access tacit knowledge* criterion is acceptable. However, there is a clear need for providing employees in CRM departments in Jordanian banks with the main requirements for sharing knowledge, such as work breaks. They should establish an environment providing a friendly and effective communication channel and emphasise more the sharing of knowledge and explicitly, as stated before.

No.	Item	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
43	I know who to ask when I need specific knowledge	14.4%	61.3%	18.0%	5.4%	0.9%
46	Requirements for sharing tacit knowledge have been provided	4.5%	36.9%	27.9%	22.5%	8.1%
Average		9.45%	49.10%	22.95%	13.95%	4.50%
Weighted average (max. 5.0) *		3.449				
Weighted average (%)		68.98%				

Table 5.17: Jordanian Banks' Ability to Access Tacit Knowledge

Also, as shown in Table 5.18, the weighted average for Jordanian banks' *ability to access explicit knowledge* criterion is the highest among other factors, with an average of 72.5%. However, it is clear that knowledge about customers is not managed systematically (30 % of respondents answered 'neutral' on item 44). This low performance could affect negatively the banks' abilities to manage CR. Additionally, Jordanian banks need to put a clear formatting to document best practices and case studies which will increase the organisational learning and decrease the possibility of failures in managing CR.

No.	Item	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
44	Professional knowledge such as customer information is managed systematically	10.8%	46.8%	30.6%	10.8%	0.9%
45	Processes and workflows are well documented	27.0%	52.3%	18.0%	2.7%	0.0%
47	Internal best-practices are well recorded	9.0%	45.0%	28.8%	15.3%	1.8%
48	Sources of explicit knowledge are well defined	11.7%	51.4%	26.1%	10.8%	0.0%
49	Formats to document best practice and case studies are clearly defined	10.8%	40.5%	36.9%	7.2%	4.5%
Average		13.86%	47.20%	28.08%	9.36%	1.44%
Weighted average (max. 5.0) *		3.625				
Weighted average (%)		72.50%				

Table 5.18: Jordanian Banks' Ability to Access Explicit Knowledge

In general, the contribution of KM Processes in supporting KM was found to be average (M=3.49, SD=0.57). The performance of C5 is high (M=3.43, SD=0.61). This means that banks in Jordan need to identify KM processes that support CRM more clearly (especially knowledge in mathematical modelling). This will reduce poor application of knowledge. The interpretation of rules in, for example, granting loans will become more effective. While the performance of C6 is high (M=3.59, SD=0.64), which means that banks in Jordan have the organisational abilities to create a clear structure for credit risk KM processes, they are however still not able to transfer CR knowledge effectively to their employees. Table 5.19 shows the overall performance of Processes' criteria

	N	Min	Max	Mean	Std. Deviation
C5: Standard Processes	242	2.17	5.00	3.500	.61425
C6: Organisational Abilities	242	1.90	4.90	3.591	.63474
F2: Processes	242	2.03	4.72	3.546	.57491
Valid N	242				

Table 5.19 Performance of Processes

RQ1.3: What is the contribution of IT in supporting KM in Banks in Jordan?

IT has the impact on CRM processes as a means to provide the required analysis and solutions to solve CR problems. Using IT will enhance the quality of CR knowledge that is generated, which can be used in other applications. Additionally, CRM needs good IT support for the following KM processes for CR: knowledge creation, knowledge storage and retrieval, knowledge sharing and knowledge application. The ‘IT’ factor is divided into two criteria. To find out the contribution of the IT factor, each criterion has been analysed separately.

Criterion1: Robust and friendly technology

As shown in Table 5.20, the weighted average for the *robust and friendly technology* criterion that supports tacit knowledge is high and acceptable. The contribution of IT in managing KM could be enhanced by encouraging employees to download the explicit knowledge residing in the banks’ knowledge-base which will be transformed into tacit knowledge when employees hold it. This in turn will increase the performance of managing CR. In the same context, banks could increase the contribution of IT by making access to knowledge faster. This could be done through enhancing the hardware, software and networks within banks.

No.	Item	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
50	I use the knowledge base easily	12.6%	53.2%	22.5%	9.0%	2.7%
52	I download knowledge frequently from the knowledge-based	9.9%	49.5%	26.1%	11.7%	2.7%
53	Data access is fast	12.6%	46.8%	27.9%	8.1%	4.5%
55	I can use the Information Systems very easily.	11.7%	61.3%	18.9%	4.5%	3.6%
Average		11.70%	52.70%	23.85%	8.33%	3.38%
Weighted average (max. 5.0) *		3.609				
Weighted average (%)		72.18%				

Table 5.20: Robust and Friendly Technology that Supports Tacit Knowledge

Table 5.21 shows the results for the *robust and friendly technology* criterion that supports explicit knowledge. The weighted average for responses was 70% approximately. The major weakness here can be summarised as insufficient help-instructions in the banks’ database. This in turn could affect the credit officers’ abilities in managing explicit CR knowledge.

No.	Question	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
51	Help-instructions in the bank's database are sufficient	8.1%	47.7%	28.8%	11.7%	3.6%
54	The Information Systems process the data very well.	12.6%	42.3%	30.6%	11.7%	2.7%
Average		10.35%	45.00%	29.70%	11.70%	3.15%
Weighted average (max. 5.0) *		3.474				
Weighted average (%)		69.48%				

Table 5.21: Robust & Friendly Technology that Supports Explicit Knowledge

Criterion 2: Tools for Managing Knowledge

Table 5.22 indicates that the weighted average for the *tools for managing tacit knowledge* is the lowest among other factors (about 60%). It seems that banks in Jordan dismiss the importance of supporting tacit knowledge. This gap could be overcome by providing employees with the proper IT tools to support sharing and creating tacit knowledge such as teleconferencing and groupware.

No.	Item	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
56	We use teleconferencing frequently in the bank	3.6%	15.3%	39.6%	31.5%	9.9%
57	The bank makes good use of groupware (such as Lotus notes) to share documents on services and notes	7.2%	39.6%	31.5%	18.0%	3.6%
Average		5.40%	27.45%	35.55%	24.75%	6.75%
Weighted average (max. 5.0) *		2.997				
Weighted average (%)		59.94%				

Table 5.22: Tools for Managing Tacit Knowledge

Finally, as shown in Table 5.23, the weighted average for the *tools for managing explicit knowledge* is acceptable but there is a major gap in using decision support systems and expert systems, which indicates the lack of reusing internal knowledge efficiently within Jordanian banks.

No.	Item	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
58	The bank uses Decision Support Systems extensively.	5.4%	36.9%	38.7%	14.4%	4.5%
59	I am satisfied with the availability of knowledge in Information Systems.	6.3%	48.6%	27.0%	13.5%	4.5%
60	The bank uses the expert systems extensively.	9.0%	43.2%	30.6%	13.5%	3.6%
61	The bank uses the Internet in executing transactions	9.0%	56.8%	20.7%	9.0%	4.5%
62	The bank uses non-IT systems to save knowledge	2.7%	21.6%	28.8%	39.6%	7.2%
Average		6.90%	42.93%	28.35%	17.40%	4.35%
Weighted average (max. 5.0) *		3.304				
Weighted average (%)		66.09%				

Table 5.23: Tools for Managing Explicit Knowledge

From this evaluation of IT in banks in Jordan (Table 5.24), we can conclude that the contribution of IT to KM practices in those banks is average (M=3.46, SD=0.6). On the one hand, they have built a robust and user-friendly technology (M=3.64, SD=0.7) which contributes positively to their performance from the knowledge that resides in IT systems but, on the other hand, the performance of C8 is average (M=3.25, SD=0.61) (the lowest among all the factors). This means that sharing knowledge and transferring knowledge from tacit into explicit is not effectively supported by IT. This indicates that the organisational abilities of banks to protect their knowledge are at risk. Lack of managing CR knowledge cycle practices can be a serious problem. For example, this can affect the transfer of the CR knowledge. This will result in poor assessment of customers, poor understanding of the present situation of Jordanian's credit market and reduce the ability to forecast through risk knowledge measurements.

	N	Min	Max	Mean	Std. Dev.
C7: Robust and user-friendly technology	242	1.12	5.00	3.647	.73327
C8: KM tools	242	1.10	4.70	3.246	.66018
F3: IT	242	1.11	4.73	3.460	.64159
Valid N	242				

Table 5.24: Performance of Information Technology

RQ.3: What are the enablers of KM in banks in Jordan?

The complexity and interaction of the factors associated with KM demand that bank managers use appropriate statistical procedures that will permit isolating the most significant factors influencing KM. With log linear analysis, the interactions and interrelationships underlying categorical data can be highlighted. The major emphasis of log linear analysis is to obtain a log linear model that is linear in the logarithms of expected frequencies of a contingency Table that adequately describes or fits the associations and interactions existing in the original frequency Table (Wrigley, 1985).

However, in this research, three factors have been identified that improve and shape KM. These factors are 'People & Culture' (F1), 'Processes' (F2) and 'IT' (F3). As the combination of these factors can lead to more effective KM, it is important to examine the interactions among these components, leading to more efficient management.

- *Application of Log-Linear Model*

The data from the questionnaires (63 items) were summarised into three broad categories by estimating the average of answers concerning each variable. After that, the average, ranging from 1 to 5 according to the Likert scale, is transformed into five categories;

< 1.5	=> very poor performance (recoded 1)
1.5 to 2.49	=> poor performance (recoded 2)
2.5 to 3.49	=> average performance (recoded 3)
3.5 to 4.25	=> high performance (recoded 4)
> 4.25	=> very high performance (recoded 5)

The log-linear model is used here to look at the effects of F1, F2, F3 and KM upon each other. This is necessary to understand the interactions that might exist among the KM factors. According to the log-linear model, all possible interactions among variables are included which can be summarised by (F1*F2*F3*KM).

- *Simultaneous test of all K-factor interactions*

A first step towards understanding the degree of complexity of the Table is to review the Table of simultaneous tests for all K-factor interactions and the tests of all marginal and partial association models.

In this research, K factors are used to introduce the nature of the log-linear models to be tested. In Table 5.25, K factors have been estimated using the Log-linear Module in SPSS 16.0. The K factor relates to the number of interactions in the frequency Table.

K-Factor	df	Max. Likert. Chi-squared.	Probability. P	Pearson Chi-squared.	Probability P
1	16	571.8336	0.000000	427.4755	0.000000
2	96	142.2620	0.001526	124.6915	0.026114
3	256	35.1736	1.000000	32.9057	1.000000
4	256	15.9046	1.000000	15.4165	1.000000

Table 5.25: Results of fitting all K-factor interactions (Simultaneous tests that all K-factor interactions are simultaneously zero)

In Table 5.26, we note that only one-way and two-way interactions are significant, while three-way and four-way interactions seem to be insignificant.

Effect	df	Partial Association. Chi-squared	Partial Association P	Marginal. Association Chi-squared.	Marginal Association P
1 [F1]	4	0.4815	0.975276	0.4815	0.975276
2 [F2]	4	2.0137	0.733239	2.0137	0.733239
3 [F3]	4	29.8654	0.000005	29.8654	0.000005
4 [KM]	4	539.4706	0.000000	539.4706	0.000000
12 [F1,F2]	16	1.7858	0.999996	1.7910	0.999995
13 [F1,F3]	16	0.6023	1.000000	0.5883	1.000000
14 [F1,KM]	16	1.0766	1.000000	1.0626	1.000000
23 [F2,F3]	16	4.6361	0.997287	4.9986	0.995760
24 [F2,KM]	16	6.5146	0.981523	6.8800	0.975505
34 [F3,KM]	16	126.9438	0.000000	127.2921	0.000000
123 [F1,F2,F3]	64	5.4496	1.000000	5.4196	1.000000
124 [F1,F2,KM]	64	3.8805	1.000000	3.8361	1.000000
134 [F1,F3,KM]	64	2.5829	1.000000	2.5847	1.000000
234 [F2,F3,KM]	64	23.4659	0.999999	23.2943	0.999999

Table 5.26: Tests of Marginal and Partial Association between F1, F2, F3 and KM

The procedures of marginal and partial association are used to choose and examine a subset of models from the four-dimensional Table. Table 5.26 provides a summary of results from the marginal and partial association tests. Briefly, the latter are done using the full four-variable frequency Table. To test the four main effects, F1, F2, F3 and KM, the base log-linear model of [F1][F2][F3][KM] is used. All four main effects are included in the base model, which is fitted to the original frequency Table and its level of significance is noted. Models are then fitted that omit each main effect in succession.

Using a significance level of 0.05, it can be observed from Table 5.26 that F3 and KM are the only significant main effects. In addition, the only significant two-way interaction is that between F3 and KM.

- *Best Model*

Table 5.27 reveals the best model selected, which is the log-linear model [F3,KM]. Therefore, using the procedure of backward elimination and partial and marginal association tests, the final choice of log-linear model is [F3, KM] with Chi-squared equal to 68.53, df equal to 600and a P value of 1.000.

Model	Chi-Squared	df	P
34 [F3, KM]	68.543	600	1.0000

Table 5.27: Best Model Selection

- *Discussion of Results*

The results of the log-linear modelling permit an assessment of the KM enablers. Based on the model [F3, KM] that best describes the four-dimensional Table, it is clear that the main enabler for KM in Jordanian banks is IT (F3). This is shown in the log-linear model by the term [F3, KM], which indicates an interaction between the factors 'Knowledge Management' [KM] and 'IT' (F3).

The finding that 'IT' is the most significant factor in enabling KM can be used to increase KM performance in Jordanian banks. As an example, enhancing the performance of employees by training them how to manage knowledge using IT, supporting them from the top management by providing them with suitable IT, and facilitating the culture to accept the new IT. These procedures are crucial to increase the contribution of IT to KM in general and managing CR knowledge in particular.

H1. There is a significant impact of the factors Processes and IT on the factor People & Culture

In order to test the first hypothesis, Two-Way ANOVA model analysis has been applied. Two-Way ANOVA is a parametric statistical test that requires that the normality assumptions are met. This ANOVA variant is used when there are two independent variables, each with two or more treatment types but only a single dependent variable. This test looks for interactions between the two independent variables, as well as how much influence each of them has had on its own. Table 5.28 investigates if there is a significant impact of the factors 'Processes' and 'IT' on 'People & Culture' in banks in Jordan at a significance level ($\alpha \leq 0.05$). With Sig. < .05, H1 has been accepted.

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	23.910	2	11.955	111.244	.000
Residual	11.606	108	.107		
Total	32.116	110			

Table 5.28: Impact of Factors Processes and IT on People & Culture (2-way ANOVA)

By the coefficients for each factor, Table 5.29 indicates that the impact of the 'Processes' factor on the contribution of 'People & Culture' is greater than 'IT.' These results assert the priority to establish standard processes for knowledge contribution and content management. This will increase the Jordanian banks' abilities to structure, categorise and access the content of knowledge.

Model	Unstandardised Coefficients		Standardised Coefficients	t	Sig.
	B	Std. Error	Beta		
Constant	.426	.195		2.181	.031
F2: Processes	.617	.081	.624	7.655	.000
F3: IT	.215	.072	.243	2.980	.004

Table 15.29: Coefficients for 'Processes' and 'IT' Factors

In the same context, the positive impact of 'IT' on 'People & Culture' indicates the important role that 'IT' plays in enhancing the contribution of KM practices in general and 'People & Culture' in particular. Enhancing the performance of employees by training them in how to manage knowledge using IT, supporting them from the top management by providing them with suitable IT and facilitating the culture to accept the new IT are crucial to increase the contribution of 'People & Culture' and 'IT' to managing CR knowledge.

H2: There is a significant difference between the Jordanian commercial banks and Islamic banks in the performance of KM practices

In order to test the second hypothesis, One-way ANOVA analysis has been used to investigate and compare the KM situation in the banks in Jordan at a significance level ($\alpha \leq 0.05$), with results shown in Table 5.30.

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	5.573	1	5.573	23.180	.000
Within Groups	57.696	240	.240		
Total	63.269	241			

Table 5.30: KM practices between bank types in Jordan (One-way ANOVA)

For the overall performance of KM activities, t-test results show a significant difference between Jordanian commercial and Islamic banks (see Table 5.31). It is clear that the Jordanian commercial banks implement KM activities better than Islamic banks. This result is consistent with previous researches (Al-Osaimy and Bamakhramah, 2004; Al-Ajlouni, 2008).

	Bank type	N	Mean	Std. dev.	Std. Error Mean
KM	Commercial	203	3.5294	.47662	.05553
	Islamic	39	3.1020	.54812	.08777

Table 5.31: KM performance of Commercial and Islamic banks

Table 5.32 shows that Islamic banks do not perform well in their support for both tacit and explicit knowledge. This clearly affects all operations in the CRM department. It might seem surprising that the performance of KM practices supporting tacit knowledge in Islamic banks is low. More analysis was done to understand this. The main contributing criterion to this result is experience. It is found that employees with less than five years' experience do not manage tacit knowledge (mean=3) well, while the others whose experience is more than five years perform well (mean=3.5). Further investigation shows that the Islamic banking industry in Jordan is suffering a shortage in well qualified staff who understands the regulations of Islamic bank operations. As around half the staffs in Islamic banks in Jordan have less than five years' experience, training is needed in delivering the different Islamic banking operations. This explains the low performance of KM practices in those banks.

	Bank type	N	Mean Tacit	Mean Explicit
KM	Commercial	72	3.5	3.5
	Islamic	39	3.1	3.1

Table 5.32: KM performance level of Commercial and Islamic banks in tacit and explicit knowledge

Also, the contributions of the three factors have been tested and these are illustrated in Table 5.33. Results show that the main difference between commercial and Islamic banks is the contribution of IT. While its contribution to KM that supports tacit knowledge is high (3.5) in commercial banks, we find that in Islamic banks it is average (2.99).

Factors	Knowledge Type	Commercial Banks	Islamic Banks
F1: People & Culture	Tacit	3.4741	3.1651
	Explicit	3.3709	2.9859
F2: Processes	Tacit	3.5279	3.1967
	Explicit	3.6835	3.3447
F3: Information Technology	Tacit	3.5071	2.9949
	Explicit	3.6130	2.9490

Table 5.33: Factors and contributions of knowledge types by banks

Results show that the main difference between commercial and Islamic banks is the contribution of IT. While its contribution to KM that supports tacit knowledge is high (3.5), in commercial banks, we find that in Islamic banks it is average (2.99).

In KM practices that support explicit knowledge, People & Culture and IT contributions in Islamic banks are average. To understand the low performance of People & Culture and IT factors, the data for the two factors have been analysed. For the People & Culture factor, the results in Table 5.34 indicates that employees in Islamic banks are not happy with the training they have received (C1 < 3.0). In the same context, results in Table 5.34 show that employees in Islamic banks are not fully aware of the availability of explicit knowledge where they work (C4 for explicit knowledge = 2.82).

Criteria	Knowledge Type	Commercial Banks	Islamic Banks
C1: Employee Training	Tacit	3.41	2.97
	Explicit	3.27	2.88
C2: Employee Involvement	Tacit	3.28	3.00
	Explicit	3.39	2.94
C3: Top-management Support	Tacit	3.72	3.53
	Explicit	3.34	3.22
C4: Employee Awareness	Tacit	3.50	3.12
	Explicit	3.50	2.82

Table 5.34: People & Culture Comparisons between Commercial and Islamic banks

Regarding the IT factor, the results in Table 5.35 indicate that there is a clear gap in providing employees in Islamic banks with IT tools that support managing both tacit and explicit knowledge (C8 < 3.0). This means that Islamic banks are not able to connect their employees with reusable codified knowledge and facilitate conversations to create new knowledge.

Criteria	Knowledge Type	Commercial Banks	Islamic Banks
C7: Robust & User- friendly Technology	Tacit	3.77	3.30
	Explicit	3.70	3.05
C8: KM Tools	Tacit	3.19	2.64
	Explicit	3.50	2.83

Table 5.35: IT criteria Comparisons between Commercial and Islamic banks

H3. There is a significant relationship between KM and CRM

In order to test the third hypothesis, Linear Regression (LR) has been applied. LR is used to specify the nature of the relation between two variables.

According to Basel report (1999), CR commonly originates from a situation where a debtor fails to pay his/her debt and it has an impact on the economic state of the bank. Therefore, in this hypothesis, the independent variable is KM, while the dependent variable CRM has been represented by the percentage of bad loans to gross loans (PCL) variable for Jordanian banks in 2007 (see appendix E).

Model	Unstandardised Coefficients		Standardised Coefficients	t	Sig.
	B	Std. Error	Beta		
Constant	.126	.012		10.366	.000
KM	-.019	.003	.-337	-5.550	.000

Table 5.36: Linear regression output between KM and PCL

Table 5.36 gives the values needed in order to write the regression equation. The regression equation will take the form:

Predicted variable (dependent variable) = slope * independent variable + intercept

The slope (B) represents how steep the line regression is. As it has been shown in Table 5.36, B= -.019 which indicates the negative relationship between KM and PCL. Therefore, as the contribution of KM increases the PCL will decreases. This indicates the importance of managing knowledge in supporting CRM.

5.6 Conclusion

The KM-CSF model has been shown to be an effective framework for studying the performance of KM activities and defining gaps in organisations, banks in Jordan in this case. Research questions and hypotheses were processed by descriptive statistical analysis, using means, standard deviation, frequencies and inferential statistics: Pearson correlation, One-way ANOVA, Two-way ANOVA, linear regression, and log-linear analysis.

Broadly speaking, the overall performance indicator for CRM departments in banks in Jordan shows the health of the KM activities. Consideration of performance for the KM factors and criteria in banks in Jordan shows that banks in Jordan need to do more to overcome gaps in managing KM:

1. For the ‘People & Culture’ factor, Jordanian banks need to identify the core CR knowledge required to maintain their competitive advantages. Additionally, Jordanian banks need to restructure their training plans on how to manage knowledge. For effective KM, skills development should occur in the following areas: communication, soft networking, peer learning, team building, collaboration and creative thinking (Horak, 2001).

   The banks should also establish an environment emphasising knowledge sharing and encouraging employees to form such a culture through creating communities of practice and reward systems and set up comforTable work breaks.

2. For the ‘Processes’ factor, results indicate that standard processes have not been established well for knowledge contribution and content management. This suggests that there is room for improvement. Jordanian banks need to identify KM processes that

support CRM more clearly (especially knowledge of mathematical modelling). This will reduce poor application of knowledge. The interpretation of rules in, for example, granting loans will become more effective. However, it is found that the CRM departments in Jordanian banks have the abilities to access, structure and categorise the content of credit risk knowledge which will help their employees to transfer CRK from tacit to explicit and vice versa.

3. For the 'IT' factor, its contribution is found to be average. Banks in Jordan have built a robust and user-friendly technology, which will affect positively their performance in benefiting from the knowledge that resides in IT, but they have failed to establish effectively tools that manage knowledge cycle practices. To overcome this, there is a broad collection of IT supports for KM which can be used and integrated into Jordanian banks' technological platform which could be categorised into one or more of the following: business intelligence, knowledge base, collaboration, content and document management, portals, customer relationship management, data mining, workflow, search and e-learning (Wong, 2005).

Testing the first hypothesis, results have shown that there is a significant impact of the factors 'Processes' and 'IT' on 'People & Culture' in banks in Jordan. Additionally, results indicate that the 'Processes' factor's impact on the contribution of 'People & Culture' is greater than 'IT'. These results assert the priority to establish standard processes for knowledge contribution and content management. This will increase the Jordanian banks' abilities to structure, categorise and access the content of knowledge. In the same context, the positive impact of 'IT' on 'People & Culture' indicates the important role that 'IT' plays in enhancing the contribution of KM practices in general and 'People & Culture' in particular.

In consideration of the second hypothesis, results indicate that commercial banks are managing CRK better than Islamic banks. Results further show that 'People & Culture' and 'IT' factors are the main causes for the low performance of KM practices in Islamic banks. The contributions of these two factors are average in terms of our classification. In looking at the presence of KM practices that support tacit knowledge compared with those which support explicit knowledge, the following has been found:

Bank Type	Tacit Knowledge	Explicit Knowledge
Commercial	High	High
Islamic	Average	Average

Table 2Table 5.37: KM practices, types of knowledge and kinds of bank

Regarding to the enablers of KM in Jordanian banks, results have shown that the main enabler for KM in Jordanian banks is IT. This indicates the importance of the 'IT' factor in managing knowledge and supporting managing CR. Enhancing the performance of employees by training them in how to manage knowledge using IT, supporting them from the top management by providing them with suitable IT and facilitating the culture to accept the new IT are crucial to increase the contribution of 'IT' to KM in general and managing CRK in particular.

Finally, regarding the third hypothesis *"There is a significant relationship between KM and CRM"*, results have shown (see Table 5.36) the negative relationship between KM and PCL. This result indicates the importance of managing knowledge in supporting CRM in Jordanian banks.

Chapter 6: Data Analysis of Loan Defaults in Jordanian Banks, Design of Credit Classification System

6.0 Chapter Structure

The previous Chapter analysed the current performance of KM practices and behaviours in CRM departments in Jordanian banks. The following gaps have been identified, in Jordanian banks:

First, there is a need to identify KM processes that support CRM more clearly (especially knowledge in mathematical modelling) which will reduce poor application of knowledge in granting retail loans. Second, there is a lack in using DSS and expert systems. This indicates the lack of using internal knowledge efficiently.

Therefore, Chapter 6 is to propose and discuss the process of building a CR classification system which will help credit officers in CRM departments in Jordanian banks in improving the decision performance in granting retail loans. The system will be built using logistic regression analysis. The dataset has been collected from three banks in Jordan (The Arab Bank, the Housing Bank and the Jordan Bank.

This Chapter further examines the variables that influence the risk of loan default in Jordanian banks, the impact of using transformation for potential risk variables on the accuracy of the produced model and, finally, the implications of the suggested model using internal implicit knowledge on managing CR in Jordanian banks. By using the knowledge gained from the model, the banks will be able to make better decisions when new customers apply to them for a loan. Banks will direct their marketing schemes to attract people with low risk profiles.

6.1 Introduction

Banks' competition for obtaining more market share and profit has become more and more aggressive in recent years, some banks taking more risks to reach a competitive advantage in the market. Banks have faced difficulties over the years for a multitude of reasons but the major cause of serious banking problems can be traced back to poor CRM. As a result, many banks suffered a huge loss from a steady increase of defaults and bad loans.

The lack of knowledge amongst senior executives about the level of risks taken in sub-prime lending, the resulting 'toxic assets' and the global nature of the instruments used to spread risks is said to be the main contributing reason for the current worldwide crisis in banks. Banks in Jordan, the focus of this research, are not immune from the exposure to the risks.

Alongside the findings from the literature review (Chapter 2), the analysis of the questionnaire (Chapter 5) has shown the need for CRM departments in Jordanian banks to build a system that use implicit knowledge to convert it into explicit which will improve the performance of granting retail loans.

In the literature review (Chapter 2), it has been found that CRM departments in Jordanian banks need a solution to improve their performance in retail loans. In order to achieve this, Jordanian banks need to build a credit classification system which helps them in differentiating between the profiles of customers that they might consider as high or low risks. This could be done by analysing patterns of behaviour of former customers so as to be able to identify distinguishing characteristics that help them to develop the most profiTable lending strategy.

The research questions examined in this Chapter are:

In Jordan,

1. Do demographic variables such as educational level, occupation, marital status, age, and gender affect the risk of default and to what extent if they do?
2. Do non-demographic variables such as loan amount, bank's location and customer-bank-age affect the risk of default and to what extent if they do?
3. What is the impact of using transformations for the potential risk variables on the performance of the produced model?
4. What are the implications of building a CR classification model using internal implicit knowledge on managing CR in Jordanian banks?

This Chapter is divided into ten Sections. The first Section is about the business background for the research problem. The following Sections illustrate the process of building the risk model for credit defaults using logistic regression. The penultimate Section deals with the discussion of the results and the final Section with the implications and conclusion.

6.2 Business Context

Jordanian banks are healthy organisations, they operate profitably, but they need to manage risk more efficiently (Siam, 2007). While the international average for default rates for retail loans is lower than 5% (CBJ, 2005), in Jordan non-performing retail loans, based on the researcher's interviews, are estimated to be in the region of 11%. This indicates the critical problem that Jordanian banks face and their need for finding effective solutions to reduce the rate of defaults.

In the next Section, improvements to risk management in retail lending in Jordanian banks will be considered through analysing the historical data.

6.3 Data Sources

Historical data of previous retail loans (2755 cases) have been collected from three Jordanian banks which dominate55% of the total assets of all banks in Jordan (CBJ, 2006): the Arab Bank, the Housing Bank and the Jordan Bank.

After cleaning the data, the dataset that will be used for analysis and modelling contains 2755 customer records in total, with 9 potential risk variables (predictors). The primary target variable (response variable) for analysis and modelling is 'Perform' (PERF), which indicates a satisfactory record of repayments on a loan. The target is a binary variable: 1 indicates 'perform' and 0 indicates 'non-perform'. The potential risk variables are divided into two groups:

1. Demographic variables: educational level, occupation, marital status, age and gender.
2. Non-demographic variables: loan amount, bank's location and customer-bank-age.

One of the missing data in this dataset (this variable has not been provided by Jordanian banks) is the time to default. Knowing this variable helps in building a survival analysis. Survival analysis can be applied to approximate the time to default or to early repayment. It is possible that if the time to default is long, the acquired interest will pay off or even exceed losses resulting from default (Stepanova and Thomas, 2002). Table 6.1 presents the variables used in this study (complete variables' codes are shown in Appendix B).

6.4 Preliminary Data Descriptions and Analysis

As the first step, descriptive statistical analysis has been carried on the 2775 customer data set. The results of the analysis are shown in Table 6.1.

Variable	Abbreviation	Type	Code	Frequency	Percent
Perform	PERF	Flag	0	364	13.2%
			1	2391	86.8%
Loan-Amount	LAM	Set	1	461	16.7%
			2	670	24.3%
			3	1222	44.4%
			4	305	11.1%
			5	77	2.8%
			6	15	0.5%
			7	3	0.1%
			8	2	O.1%
Customer-Bank-Age	CAS	Set	1	1772	64.3%
			2	223	8.1%
			3	65	2.4%
			4	47	1.7%
			5	286	10.4%
			6	362	13.1%
Gender	GND	Flag	0	401	14.6%
			1	2354	85.4%
Income	INC	Set	1	236	8.6%
			2	757	27.5%
			3	744	27.0%
			4	455	16.5%
			5	563	20.5%
Marital Status	MAR	Flag	0	874	31.7%
			1	1881	68.3%
Educational Level	EDU	Set	1	8	0.3%
			2	112	4.1%
			3	1056	38.3%
			4	1452	52.7%
			5	127	4.6%
Occupation	OCC	Set	0	7	0.3%
			1	213	7.7%
			2	93	3.4%
			3	268	9.7%
			4	279	10.1%
			5	384	13.9
			6	59	2.1%
			7	480	17.4%
			8	242	8.8%
			9	730	26.5%
Age	AGE	Set	1	1	0.0%
			2	300	10.9%
			3	1584	57.5%
			4	530	19.2%
			5	277	10.1%
			6	63	2.3%

Table 6.1: Descriptive statistics for simple subset of Jordanian banks' default data

The dataset contains 2755 entries of previous customers. This is a rather large sample size and will have the consequence that even fairly weak predictive relationships will be detected to a high degree of statistical significance.

Table 6.1 gives the descriptive statistics for the variables used in the study. Values of CR variable Perform (PERF) shows that 86.8% of clients in this study did not default.

While, the loan amount (LAM) is measured as an interval variable ranging from 1 (for amounts less than or equal to 5000 J.D) to 8 (350000 J.D and over). Results show that 44% of customers in this dataset took loans range 5000-30000 J.D.

Also customer-bank-age (CAS) is an interval variable ranging from 1 (less than or equal one year) to 6 (more than five years). Results show that 64.3% of borrowers were new customers when they took loans (CAS ≤ 1 year).

Gender (GND) is a dummy variable showing the borrower's gender as either male (1) or female (0). Table 6.1 shows that the majority borrowers of the data set (85.4%) were males while 14.6% were females.

Educational level (EDU) is a ordinal variable ranging from 1(un-educated) to 5 (postgraduate). Results show that 52.7% of borrowers had a bachelor degree.

Income (INC) is a scale variable ranging from 1 (200-400) J.D to 5 (greater than 1000 J.D). Table 6.1 shows that 52.4% of borrowers' income is in the range 401-800 J.D.

Marital status (MAR) is a dummy variable (2 if married and 1 otherwise). Results show that most of borrowers (68.3%) are married.

The age (AGE) of borrower is measured as a scale variable ranging from 1 (19 years or less) to 6 (60 years and over). Results show that 57.5% of borrowers' ages range from (30 – 39) years.

Finally, Occupation (OCC) is a categorical variable. 0 represents 'un known' jobs while 9 represents 'Business, Management and administration' jobs. The biggest cluster in the data set is 'Business, Management and administration' (26.5%), while the smallest cluster is 'un-known' jobs (0.3%).

The intention in this Chapter is to see whether some of the variables that reflect specific circumstances in Jordan appear in the model or not. If they do, this would mean they are important in customer behaviour and worth including into the modelling. Depending on previous studies, (Vandal and Thibodean, 1985; Zorn and Lea, 1989; and Quercia and Stegman, 1992), some of the variables we expect to see in the produced model are:

- Occupation: we expect that clients without a job are more likely to default,
- Income: clients with higher income should have lower probability of default,
- Marital Status: married clients should have lower probability of default,
- Education: educated clients should have lower probability of default,
- Age: we expect that young clients should have higher probability of default,
- Loan amount: we expect high loan amount to have lower probability of risk, and
- Customer-Bank-Age: we expect new customer to have lower probability of risk.

To find out the impact of the independent variables on the output variable (PERF), a correlation analysis has been carried out. Table 6.2 gives the correlations between the predictor variables and the output variable PERF.

Variable	Abbreviation	Perform
Perform	PERF	1
Loan-Amount	LAM	0.289**
Customer-bank-age	CAS	-.576**
Gender	GND	0.009
Income	INC	.389**
Marital Status	MAR	.255**
Educational Level	EDU	.254**
Occupation	OCC	.115**
Age	AGE	-.374**
Location	LOC	0.003

Table 6.2: Correlation analysis
**. Correlation significant at 0.01 level (2-tailed).

Depending on the results, variables are divided into two groups:

1. *Significant variables*

1.1 Variables that influence non-defaulting: INC, MAR, EDU and OCC

The results of the present study in this group are in conformity with earlier studies (Vandal and Thibodean, 1985; Zorn and Lea, 1989; Quercia and Stegman, 1992) which have shown that unmarried or divorced borrowers are more likely to default than married ones. Hayashi (1987) has shown that borrowers with low income are more likely to default. Peter (2006) stated that uneducated and un-skilled borrowers are at a high level of risk.

1.2 Variables that influence the risk of default: CAS, and AGE

CAS and LAM variables are confirmatory of another study, on Spanish banks (Gabriel, 2002), showing that "when there is an exclusive or very close relationship between the bank and its borrowers, the bank is more willing to grant higher risk loans; also that large loans are at a lower risk, probably because the loan operation has been studied in greater detail".

As to the AGE variable; in contrast to previous studies that have shown that young people are high risk borrowers (Peter, 2006); this study shows that in Jordan young customers are less likely to default than older ones. To find out the reason behind this difference, further analysis has been needed.

2. *Non-significant variables*: Gender (GND) and Location (LOC) have no significant impact on risk of defaulting.

At this stage, we re-state the aim of this risk analysis exercise. We wish to be able to use covariate data normally available to the bank prior to the loan being granted to make a prediction of the probability of default of an applicant. This is the model-based form of loan analysis which has the counterpart of Classification analysis, in which a set of classifier variables will be used to classify a person as either a defaulter or non-defaulter, before a loan is actually made. The problem with classification techniques is that it can be the case that all applicants have a fairly low probability of defaulting and the probability threshold for a classification of a defaulter has to be set very low. This is not normally regarded as acceptable. In consequence, it is usually the case that classification diagrams based on a logistic model will be useless, since they use by default the prob. value of 0.5 as the classifier boundary. Hence, the usual result is that nobody is classed as a defaulter.

The use of the logistic regression risk model allows a comparative analysis of risk, even though it is uniformly low.

6.5 Methods of Predicting Defaulting Borrowers

According to Fensterstock (2005), most CR evaluation systems currently in use are based on some form of judgemental-based system which makes quantifying a risk a big challenge. According to him, this problem can be solved by using a statistical analysis.

As was explained in the previous Section, the target outcome is a binary variable which takes values 0 and 1, while the predictor variables are two types: continuous and categorical.

There are several methods which are suitable for credit scoring in the banking segment, such as:

1. Neural networks: A neural network (NNW) is a mathematical representation inspired by the functioning the human brain. A typical network is composed of a series of interconnected nodes and the corresponding weights between them. It aims at simulating the complex mapping between the input and output (Hui-Chung, 2007). Many different types of NNW have been specified in the literature.

The neural networks have higher credit scoring capability than other statistical methods (e.g. LDA and logistic regression) (Hui-Chung, 2007). The major drawback of NNWs is their lack of explanatory capability. While they can achieve a high prediction accuracy rate, the reasoning behind why and how the decision was reached is not available. For example, in a case of not accepting a loan it is impossible to determine which characteristic(s) was (were) exactly the key one(s) to prompt rejection of the application. Accordingly, it is very difficult to explain the decision results to managers (Baesens, 2003; Lee, 2002; West, 2000).

2. Linear Discriminant Analysis: The aim of Linear Discriminant Analysis (hereafter LDA) is to classify a heterogeneous population into homogeneous subsets and further the decision process on these subsets assuming the prior probabilities of analysis target is equal (Hui-Chung, 2007).

As stated by (Hand, 1997), LDA is the first proposed technique for building credit scoring models. The advantages of the LDA method are that it is simple, it can be very easily calculated and indeed it works very well; it is often used by banks for credit-scoring purposes. The disadvantage is that LDA requires as a rule distributed data but the credit data are often non-normal (and categorised) (Dudoit, 2002; and Webb, 2002).

3. K-nearest Neighbour Classifier: The k-nearest neighbour classifier is a standard technique in pattern recognition. It serves as an example of the non-parametric statistical approach and is one of the simplest and most straightforward classifiers (Kilian, 2005). This technique assesses the similarities between the pattern identified in the training set and the input pattern (Henley, 1996).

However, it has been shown that when the points are not uniformly distributed, predetermining the value of k becomes very difficult (Domeniconi, 2002). Holmes and Adams (2002) conceived that there is a shortage of a formal framework for choosing the k and that the method can only make discrete predictions by accounting the relative frequencies which have no probabilistic interpretation. They tried to overcome these difficulties by presenting the Bayesian approach as a solution which integrates over the choice of k. Such an approach draws the conclusion that marginal predictions are given as proper probabilities.

4. Logistic regression: This is used for modelling the binary outcome variable which accepts continuous and categorical predictors.

Logistic regression is a variation of ordinary regression used when the dependent variable is a binary variable (i. e., it takes only two values, which usually represent the occurrence or non-occurrence of some outcome event) and the independent (input) variables are continuous, categorical or both. Unlike ordinary linear regression, logistic regression does not assume that the relationship between the independent variables and the dependent variable is a linear one. Nor does it assume that the dependent variable or the error terms are distributed normally.

Logistic regression has been widely used in credit scoring applications due to its simplicity and explainability. Recently, Charitou (2004) found that the logistic regression analysis is superior to other methods in predicting defaults. Desai (1996) examined neural networks, logistic regression, and linear discriminant analysis for scoring credit systems. He concluded that neural networks outperform linear discriminant analysis in classifying loan applicants into good and bad applicants and logistic regression is as good as neural networks.

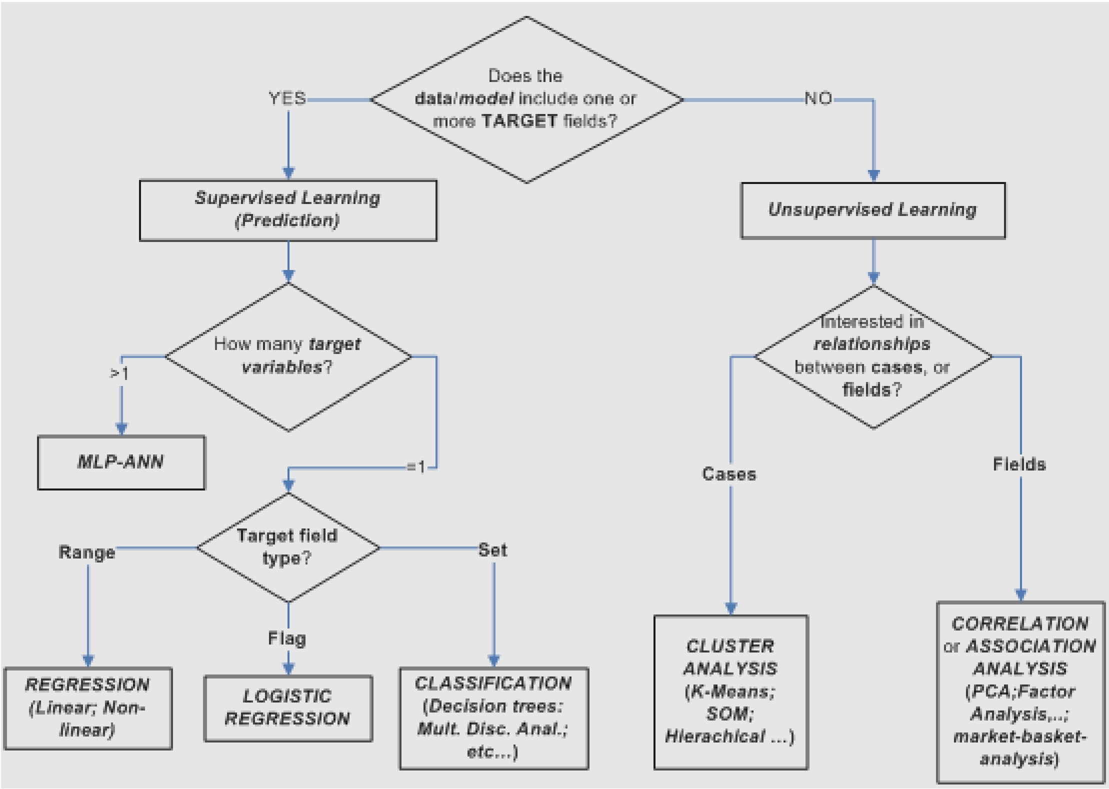


Figure 6.1: Simple Flow-chart for Model Selection

Since the target field is a flag variable (Perform), it is therefore quite acceptable that logistic regression analysis can be used to predict the probability of risk in this study.

$$P(PERF = 1) = \frac{1}{1 + e^{-(c + \sum_{k=1}^{n} B_k X_k)}} \qquad (6.1)$$

Where (P) is the probability that PERF=1 and x_1, x_2,..., x_n are the independent variables (predictors). *C* is the constant, $B_1, \ldots, B_n$ are known as the regression coefficients, which have to be estimated from the data, and n represents number of independent variables. Logistic regression estimates the probability of a certain event occurring.

6.6 Predicting Defaulting Borrowers using Logistic Regression Analysis

As a first step, variables have been entered to the SPSS 16.0 separately (with no mapping transformations). The accuracy of predicting non-defaulting of the produced model (see Equation C1 in Appendix C) is 79.1%. To increase the performance, the model transformations for variables have been done. Besides, all the potential interactions between variables have been entered to the SPSS 16.0 to get the best model using logistic regression (see Table A3 in Appendix A).

The sequence of covariate inclusion in the dataset was determined by forward step wise selection methods. However, for reporting purposes, it is prefered to build up to the best model in a stepwise manner, giving a detailed analysis of the performance and the improvements at each stage. Performing forward stepwise logistic regression can improve a model since it excludes any insignificant variables. The model starts with the strongest significant variable and then the next significant variable is entered step by step.

To assess the appropriateness and usefulness of the model, the overall goodness of fit of the model is tested through using measures such as Cox and Snell R^2, Nagelkkere R^2 and Hosmer & Lemshow. Additionally, the ability of the model to discriminate between the two groups defined by the response variable is evaluated. Finally, Wald statistics is used to assess the importance of each predictor by carrying out statistical tests of the significance of the coefficients.

In linear regression R^2 represents the percentage of total variation explained by the model. While in logistic regression an equivalent statistic to R^2 does not exist. However, to evaluate the goodness-of-fit of logistic models, several R^2 have been developed.

Cox and Snell R^2 and Nagelkkere R^2 indicate how much of the variability in our data is successfully explained away by our model. Large values of these R^2 (Cox and Snell and Nagelkkere) have a minimum value of zero and a maximum of 1, indicating that a model captures more of the data variability.

In logistic regression, the Cox and Snell R^2 (R^2_{CS}) method is used to test the goodness of fit of the model, where

$$R^2_{cs} = 1 - \left\{ \frac{L\ (M_{Intercept})}{L\ (M_{Full})} \right\}^{\frac{2}{N}} \tag{6.2}$$

M $_{Full}$ = Model with predictors

M $_{Intercept}$ = Model without predictors

N= Number of independent variables

L = Estimated likelihood, where

$$LL = \sum_{i=1}^{n}\{Y_i \ln(P(Y_i)) + (1-Y_i)\ln[1-P(Y_i)]\} \qquad (6.3)$$

Where P is the probability of the event Y to occur for a customer I, and n represents number of independent variables.

The problem in using Cox and Snell R² (R^2_{CS}) is that the maximum value it reaches is 0.75. To correct this problem, another test is used (Nagelkkere R² (R^2_N)). The form of this test is,

$$R^2_N = \frac{R^2_{cs}}{1 - L(M_{Intercept})^{\frac{2}{N}}} \qquad (6.4)$$

From the goodness of fit in Figure 6.1, it is found that Nagelkere R² = .887, which means that 88.7% of the variation in the dependent variable, default risk, has been explained by the independent variables, which indicates the power of the model in predicting defaulting customers.

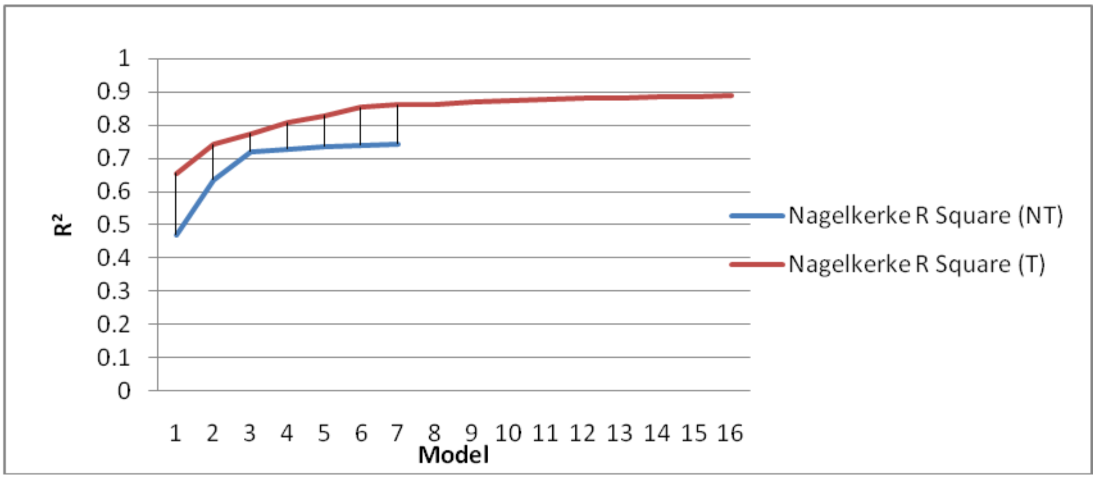


Figure 6.2: Goodness of fit statistics performance's comparison before and after using transformation as complexity increased by forward stepwise logistic regression
******. NT (without using transformation), T (using transformation)

From Figure 6.2, comparing the performance of predictability shows the improvements made to the model using transformation on variables.

The Hosmer-Lemeshow goodness of fit index is useful for assessing the overall model fit, particularly where there are many predictor variables, which may be continuous or categorical. It proposes a Pearson x^2 - statistic for logistic regression based on a grouping of the estimated probabilities (Hosmer and Lemeshow, 2000).

In Table 6.3, significant values for models from step 2 to step 10 indicate that there is lack of fit (sig. > 0.05), while subsequent steps are significant, so apparently any lack of fit has been corrected.

Step	Chi-squared	df	Sig.
1	139.207	6	.000
2	15.041	8	.058
3	12.983	8	.112
4	9.761	8	.282
5	5.125	8	.744
6	9.447	8	.306
7	14.016	8	.081
8	17.196	8	.028
9	5.132	8	.743
10	6.274	8	.617
11	22.581	8	.004
12	63.183	8	.000
13	51.450	8	.000
14	121.358	8	.000
15	108.967	8	.000
16	78.142	8	.000

Figure 6.3: Hosmer and Lemshow goodness of fit index

To select the best model, different methods are available (such as Akaike Information Criterion (AIC) and F-Test). The F-test seems to be more robust towards higher statistical error (0.15 and 0.2), while the AIC is better in selecting the correct model for lower (0.05 and 0.1) statistical error (Kletting and Glatting, 2009). Since the acceptable error in this study is ≤0.05, the AIC test has been used (see Table 6.4).

$$AIC= -2 \text{ log-likelihood} + 2K \tag{6.5}$$

where K is the number of estimated parameters included in the model (to find number of parameters see Table A3 in Appendix A).

Step	K	Log-Likelihood	AIC
1	1	945.703	947.703
2	2	736.379	740.379
3	3	652.705	658.705
4	4	571.523	579.523
5	5	512.276	522.276
6	6	443.914	455.914
7	7	425.786	439.786
8	8	415.64	431.64
9	9	405.643	423.643
10	10	392.845	412.845
11	11	374.533	396.533
12	10	362.965	382.965
13	11	363.227	385.227
14	12	350.341	374.341
15	10	350.712	370.712
16	13	345.406	371.406

Table 6.4: AIC results of candidate models

According to Table 6.4, step 15 is the best model (lowest AIC). This means that we will depend on the 15th model in this study (for complete steps, see Table A2 in Appendix A). In order to obtain an overall picture of the prediction capability of this model, we perform Classification Accuracy and results are given in Table 6.5.The classification Table shows the overall percentage of correct prediction in step 15 (the most accurate) is 97.9%. The accuracy of the model in predicting non-perform customers is 89.3%. Comparing this result with that of the previously produced model which was 79.1% (see EquationC1 in Appendix C), shows the improvements made to the model using transformation for variables. This means that we will depend on the 15th model.

	Observed		Perform		
			Actual	Predicted	Accuracy Results
Step 15	Perform	No= 0	364	325	89.3
		Yes= 1	2391	2372	99.2
	Overall Percentage		2755	2697	97.9

Table 6.5: Classification Accuracy Results for Produced CR Model

6.6.1 Over-fitting Test

Another important consideration is the performance metrics. The performance is assessed by testing the model against a testing and possibly a validation sample. In many cases there are no validation samples and the testing sample is used to measure the model performance. Some algorithms use the testing sample to stop the training of the model. Since this may introduce a bias (called "over-fitting") it is important to measure the performance on a validation sample. Therefore, another dataset has been used. Since some of the variables are missing, using the mean value of the missing variables has been used as a strategy to overcome this problem.

The models were tested for their effectiveness on several measures - percentage correctly classified (PCC), and Percentage Incorrectly classified (PIC). The results of classification indicate that there is no bias in the produced model (see Table 6.6).

Dataset	Observed	PCC	PIC
one	defaulters	89.3%	10.7%
	Non defaulters	99.2%	0.8%
two	Defaulters	80.3%	19.7%
	Non defaulters	97.1%	2.9%

Table 6.6: Over fitting- test Results

6.7 Building Risk Model

Table 6.7 presents the output of step 15 from the logistic regression analysis on the 2755 borrower dataset (for complete steps, see Table A3 in Appendix A).

It is important to note that the variables in the dataset have correlations between them and therefore they have a combined effect on the creditworthiness of the borrower. For example, the 'age' factor in isolation does not have a direct impact on the risk. But when taking into consideration interaction variables, for example age with loan amount (LAM) and the customer bank age (CAS), we find that the age factor (AGE) increases the predictability of our model (increases model accuracy). That is why a multivariate analysis of all these variables have been done (results are shown in Table 6.7).

Variable	B	S.E.	Wald	df	Sig.	Exp(B)*
OCC			59.607	9	.000	
OCC(2)	-5.388	1.617	11.100	1	.001	.005
OCC(3)	-5.340	2.059	6.730	1	.009	.005
OCC(4)	19.869	3.616	30.199	1	.000	4.255E8
OCC(5)	-7.742	1.739	19.827	1	.000	.000
CAS^2	.196	.040	24.234	1	.000	1.217
INC^2	2.137	.342	39.127	1	.000	8.474
AGE^0.5	11.573	1.820	40.448	1	.000	106213.819
Log(EDU)	23.846	5.395	19.539	1	.000	2.271E10
INC by LAM	-.220	.074	8.838	1	.003	.803
CAS by AGE	-.687	.088	60.618	1	.000	.503
EDU by INC	-1.177	.339	12.095	1	.001	.308
INC * OCC			66.109	9	.000	
INC by OCC(2)	4.487	1.067	17.681	1	.000	88.812
INC by OCC(3)	3.219	1.240	6.735	1	.009	25.011
INC by OCC(4)	-6.360	1.166	29.762	1	.000	.002
INC by OCC(5)	3.210	.770	17.394	1	.000	24.777
INC by OCC(8)	2.496	.956	6.820	1	.009	12.134
INC by AGE	-1.186	.216	30.122	1	.000	.305
AGE by MAR	-.349	.100	12.115	1	.001	.705
Constant	-19.926	3.184	39.157	1	.000	.000

Table 6.7: Estimation results of Logistic Regression Analysis

The first column represents the coefficients associated with each variable along with their sign. The probabilities of significance are reported in the penultimate column in Table 6.7. If Sig. is less than 0.05, the independent variables are pointed as significant at 5% confidence level.

The last column in Table 6.7, Exp (B), which is the exponential function (e = 2.718) raised to the value of each regression coefficient, indicates the value by which the odds of the event change when the i^{th} independent variable increases by one unit. If the value is greater than 1, the odds are increased; if the value is less than 1, the odds are decreased. A value of 1 leaves the odds unchanged. Based on the above, we can conclude that OCC (2), OCC (3), OCC (5), INC*LAM, CAS*AGE, EDU*INC, INC* OCC (4), INC*AGE, and AGE*MAR affect the probability of not defaulting negatively, while the other variables have a positive impact on the probability of not defaulting. Based on the results in Table 6.7, the following equation has been built:

$$P(PERF = 1) = \frac{1}{1 + e^{-(C + \sum_{i=1, j=1}^{n} b_i f(x_i) * g(x_j))}} \qquad (6.6)$$

where,

C= -19.926

$$\sum_{i=1, j=1}^{n} b_i\ f(x_i) * g(x_j) = \alpha * OCC + 0.196\, CAS^2 + 2.137\, INC^2 + 11.573\, AGE \wedge 0.5$$

$$+ 23.846\, Log(EDU) - 0.22\, INC * LAM - 0.687\, CAS * AGE$$

$$- 1.177\, EDU * INC + \beta * INC * OCC - 1.186\, AGE * MAR$$

where α and β are as given in Table 6.8.

	OCC(1)	OCC(2)	OCC(3)	OCC(4)	OCC(5)	OCC(6)	OCC(7)	OCC(8)	OCC(9)
α	0	-5.388	-5.340	19.869	-7.742	0	0	0	0
β	0	4.487	3.219	-6.360	3.210	0	0	2.496	0

Table 6.8: α and β Values

Using the above equation one would get the risk probability value for a loan if information on all the 10 parameters specified above is available. If one of the parameters is missing, the risk probability value would not give the exact picture about the creditworthiness of the borrower. Therefore, to ensure more accurate prediction about the borrower's ability to pay back the loan, it is advised that information is collected about all 10 parameters.

From the order in which the variables have been entered in the forward stepwise process for building the risk model, Table 6.9 shows the priority of significance of each variable.

Priority	Variable
1	CAS * AGE
2	CAS^2
3	OCC
4	INC*OCC
5	Log(EDU)
6	AGE*MAR
7	INC * LAM
8	AGE^0.5
9	INC^2
10	EDU*INC

Table 6.9: Order of variables according to importance

6.8 Building risk classification system

Based on our analysis, a general system for grouping risk classes has been formulated by adopting the following steps:

Step One: Estimating the new independent parameters from our data set (this includes the following: CAS^2, INC^2, AGE^0.5, Log (EDU), INC*LAM, CAS*AGE, EDU*INC, INC*OCC and AGE*MAR. Finally, we divided OCC into five variables (OCC2, OCC 3, OCC4, OCC5 and OCC8) which were coded by 0 and 1 according to the occupation of client.

Step Two: Applying the previous equation to the data set (the independent variables from step one), which included 2755 cases, and solving for Pr(Y=1), obtaining the probability of not defaulting.

Step Three: Sorting the file according to Pr(Y=1) in ascending order (from the lowest to the highest value).

Step Four: Splitting the file into three categories according to the following criteria:

- Part 1: if Pr(y=1) less than or equal to 0.250 (high risk)
- Part 2: if Pr(y=1) higher than 0.250 and less than 0.750 (moderate)
- Part 3: if Pr(y=1) higher than or equal to 0.750 (low risk)

(It is an example of how to divide the risk into levels. There is no consensus in the literature about number of levels).

Step Five: Running frequency analysis for each category and determining the characteristic that distinguishes each one. The results of this analysis have been confirmed in comparing them with the cross-tabulation for each parameter.

6.8.1 Eliciting Typical Characteristics of Each Risk Level

As it has been stated in the previous Section, by running frequency analysis some typical characteristics of each risk level can be elicited. Table 6.9 reveals the results of frequency analysis and crosstabs.

As shown before in the correlation test (Table 6.2) some of the results in Table 6.10 are not consistent with previous studies with international data. This indicates that CR systems which are built using international data are not useful in the Jordanian context. Jordanian banks need to adopt a system that is built using internal knowledge (as recommended by the Basel report (1999)).

Risk Level / Variable	High Risk	Moderate Risk	Low Risk
Customer-Bank-Age	Greater than or equal to 6 years	4 years – 5 years	Less than or equal to 3 years
Age	Greater than or equal to 60 years	40 years – 59 years	Less than or equal to 39 years
Education	Uneducated High school	Diploma	Bachelor Postgraduate
Income	JOD 200 – JOD 400	JOD 401 – JOD600	Greater than or equal to JOD 601
Occupation	Education, Training and Media Unknown	Vocational Marketing, Sales and Finance, Customer services	Other clusters

Table 6.10: Some Typical Jordanian Borrowers' Risk Characteristics

From the point of view of banking supervisors concerned with the quality of loans, it is of interest to obtain an evaluation of the marginal impact of each characteristic associated with a loan. In this way, for example, supervisors can better guide their CR monitoring. Alert systems could be designed to detect if a bank approves an increasing number of loans whose characteristics signal greater CR. Because of that, results in Table 6.9 need more careful study to obtain the hidden interactions between the different variables.

According to the order of their importance (Table 6.9), the parameters in the logistic regression model (Equation 6.6) have been analysed further using cross-tabulation analysis and the results were then compared with the output of the regression model (Equation 6.6). The most important parameter has been found to be CAS*AGE.

6.8.1.1 Customer-banks-age and Age variables (CAS * AGE)

From the correlation analysis shown in Table 5.2, CAS (customer bank age) variable is found to be the main risk variable determining defaults in Jordanian banks. The mean for this variable is (2.25), which suggests that the borrowers on average have been customers with the bank for 2.25 years. The correlation between CAS and PERF is (-0.576), which means that the possibility of defaulting for old customers (CAS≥ 3years) is more than for new customers (CAS< 3) (From Table 8A we can see that 86.5% of defaulters had been customers for more than three years, i.e. > average). Clearly, Jordanian banks have failed in applying the knowledge they have about their customers (i.e. previous loans, his/her payable abilities) in managing CR and making decisions.

In the logistic regression model (Equation 6.6), CAS variable has been combined with another important variable (AGE) to be the first parameter entered into the model. AGE also has a negative correlation with PERF (correlation= -0.374) which means that young borrowers are less likely to default than old ones. It has been found (see Table 9A in Appendix A) that 73% of defaulters have been customers for more than 3 years and their ages were greater than or equal to 40 years. This indicates the importance of analysing this parameter.

Using crosstab analysis, further study has been done for both variables to obtain some revealing information. It has been found that 78% of borrowers who were customers for more than three years are low income (less than or equal to JOD 600) and low income has been shown to be one of the main drivers for defaulting (89.3% of defaulters are in the low income cluster.

Regarding the AGE variable, it has been found that 93% of defaulters whose age is greater than or equal to 50 are from the low income cluster (less than or equal to JOD 600). From the data, it has been found that in Jordan old people (older than or equal to 50) earn less than young ones (see Figure 6.3 and 6.4).

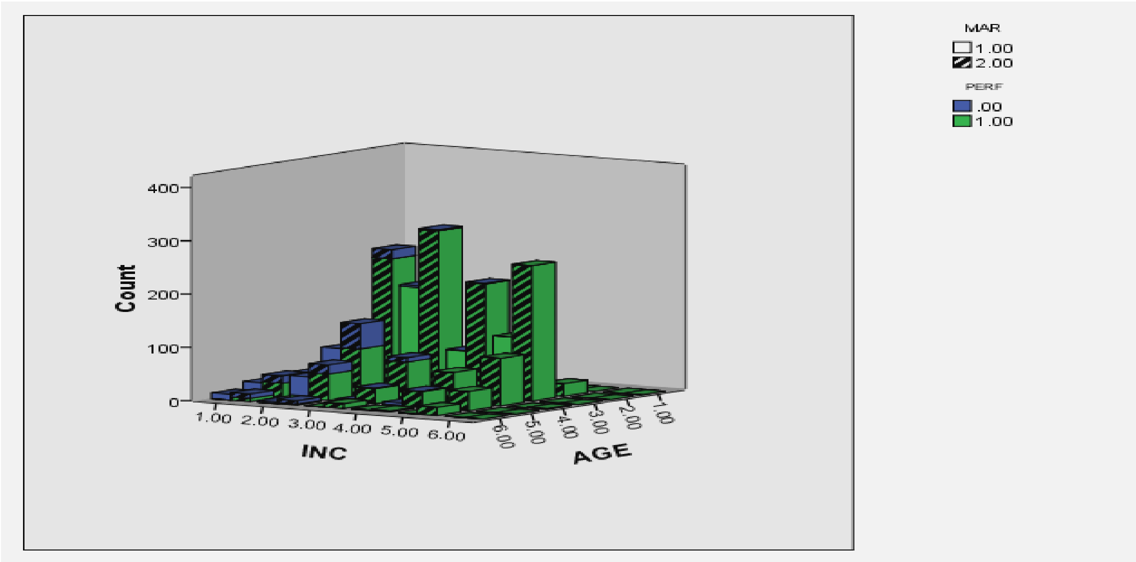


Figure 6.3: Relation between age, income and marital status

Another important factor contributing to older (age ≥50) customers' defaulting is their marital status (MAR). 62.7% of defaulters aged more than or equal to 50 were not married or were separated (see Figure 6.3).

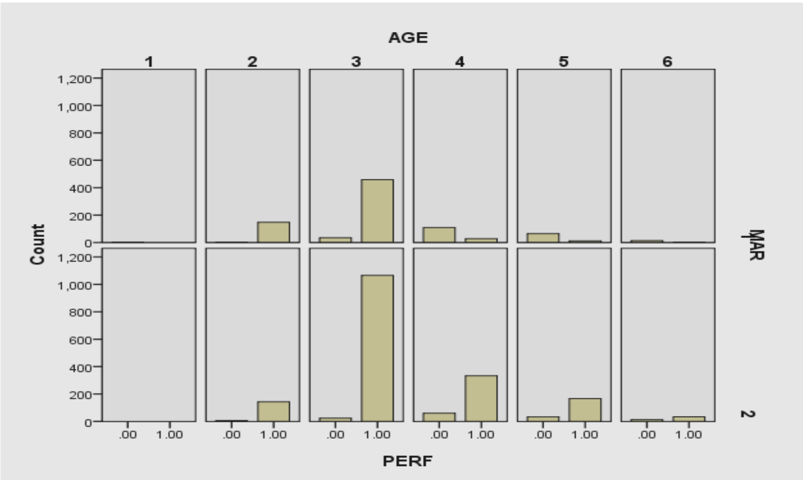


Figure 6.4: Impact of marital status variable on age behaviour

From the above analysis, we can conclude that the factors that influence the behaviour of CAS and AGE are income and marital status.

6.8.1.2 Occupation (OCC) and Income (INC)

The second parameter in order of importance is occupation (OCC). From the correlation analysis (see Table 6.2), we see that there is a positive relationship between occupation and PERF. Since the occupation variable is categorical (the order of occupation codes is meaningless) this output is not useful as it is. Because of that, more analysis has been done on this variable to understand its behaviour.

- The biggest cluster of borrowers is Business, Management and administration with 26.4%.
- The most risky clusters are Job unknown (100% defaulted), Education, Training and Media (61% defaulted), followed by Marketing, Sales and Finance and Customer services (47% defaulted).
- Marketing, Sales and Finance and Customer services is a risky cluster because 36% of defaulters were from this group.
- 100% of borrowers from the Job unknown cluster are from low income (less than or equal to JOD600).
- 72% of borrowers from Education and Training and Media are from the low income cluster (less than or equal to JOD600), 86% of those defaulted.
- 49.8% of borrowers from Marketing, Sales and Finance and Customer services are from the low income cluster (less than or equal to JOD 600), all defaulters from this group are low income borrowers.

As a result of the above analysis, it is reasonable to conclude that the income variable (INC) is the main factor that influences the probability of defaulting.

6.8.1.3 Income and Loan amount variables (INC*LAM)

According to Table 6.9, one of the main parameters to determine the risk level is LAM*INC. The correlation between LAM and PERF is 0.289. This means that high loan amounts were noticeably less risky than small. To know the reason behind this, cross-tabulations were produced (Table A. 6 in Appendix A).

First of all, the correlation between LAM and INC is 0.21. This means that large loan amounts were granted to high income applicants and vice versa. As revealed before, 89.3% of defaulters are from low income (less than or equal to JOD 600), while 72% of borrowers who did not default were from high income (≥ JOD 601).

It is the income that is the critical factor not the loan amount. Using Table A6 in Appendix A, it has been found that 89.7% of defaulters from the small amounts' cluster (less than or equal to JOD30000) are from the low income cluster (less than or equal to JOD 600). While for 77% of borrowers who were granted loans greater than JOD30000, their income was greater than JOD 600. This explains why the correlation between LAM and PERF is positive.

It is clear that Jordanian banks' policy became more risky with their old customers (CAS ≥ 3 years) as high loan amounts were given to low income borrowers which increased the percentage of defaulting loans (estimated default rate in Jordan is 11%).

6.8.1.4 Education & Income Variables (EDU*INC)

Both variables have positive correlations with PERF (Corr (EDU) = 0.254 and Corr (INC) = 0.389). Although on average Jordanian borrowers in this dataset were educated (mean= 3.57), 67% of defaulters from all income groups were low educated. It has also been found that 66% of defaulters were uneducated and from low income category.

Another important factor affecting the EDU variable is marital status (MAR). Analysis has shown that 92% of applicants who defaulted were low educated (uneducated or high school) and unmarried or divorced. Along with that, 55% of defaulters were unmarried or divorced and from the low income cluster (less than or equal to JOD 600).

From the above arguments, it has been shown that the main factors affecting CR are (in order): Income, Customer-Bank-Age and Marital status, which is consistent with the correlation analysis (see Table 6.2).

6.9 Risk Determined Loan Policy

Taking into account the above analysis, Jordanian banks need to adopt a credit policy using their internal knowledge. They need to carefully consider an applicant:

1. whose income is less than or equal to JOD 600, or
2. who is not married or divorced, and whose age is greater than or equal to 40, or
3. who is a customer for more than or equal to 3 years.

6.9.1 Simple Risk Loan Policy

Table 6.10 could be used to obtain a simple risk loan policy which uses risk variables individually. Suppose the bank wishes to lend only to applicants with low risk (Pr (PERF) ≥ 0.75). Then they should

1. Lend to applicants whose income is greater than or equal to JOD 601 and who are married, or
2. Lend to applicants whose education level is bachelor or postgraduate and who have been customers of the bank for less than or equal to 3 years.

The above policy is an example of the types of analysis that could be done in Jordanian banks simply by running a frequency analysis. In spite of its simplicity, this policy is not efficient since it takes variables independently without consideration of the interactions between them. Because of that, credit officers cannot rely on it.

6.9.2 Second Level Loan Policy

Through using crosstabs in Section 6.8, it has been shown that variables have interactions between them which affect the probability of defaulting. Therefore, results from Section 6.8 were used to build a higher level of loan policy which uses interactions between two or three variables at maximum:

1. Low Risk Level (Pr (PERF) ≥ 0.75):
   a) If the applicant's income is less than or equal to JOD 400 (high risk applicant) and his/her education level is postgraduate then the risk becomes low.
   b) If the applicant is uneducated (high risk) and his/her income greater than JOD 600 then the risk becomes low.
   c) If the applicant is unmarried or divorced (high risk) and his/her educational level is Bachelor then the risk becomes low.
   d) If the loan amount is less than or equal to JOD 5000 (high risk) and he/she is married then the risk becomes low.
   e) If the applicant is aged 40 - 49 (moderate risk applicant) and his/her income is greater than JOD 600 then the risk becomes low.

2. Moderate Risk Level (0.75>Pr (PERF>0.25):
   a) If the applicant's income is less than or equal to JOD 400 (high risk applicant) and his/her education level is Diploma then the risk becomes moderate.

b) If the applicant's education level is high school (high risk applicant) and he/she is married then the risk becomes moderate.
c) If the applicant is unmarried or divorced (high risk) and his/her educational level is Diploma then the risk becomes moderate.
d) If the applicant's educational level is high school (high risk applicant), and his/her income is JOD 400 – JOD 600 then the risk becomes moderate.
e) If the applicant's income is less than JOD 600 (high risk applicant), and he/she is married then the risk becomes moderate.

3. High Risk Level (Pr (PERF) $\leq$ 0.25):
   a) If the applicant has been a customer for 2 years or fewer (low risk applicant) and his/her job is from cluster 6 (Education, Training and Media) then the risk is high.
   b) If the applicant has been customer of the bank for 4 years (moderate risk applicant) and his/her age is $\geq$ 40 years then the risk is high.
   c) If the applicant is aged 40 – 49 (moderate risk applicant) and he/she is not married or divorced then the risk becomes high.
   d) If the applicant is aged 40 - 49 (moderate risk applicant) and his/her income is less than or equal to JOD 600 then the risk becomes high.

It is clear that this suggested CR policy is better than the previous one. Although it takes into account the interactions between different variables, it fails to use all the potential interactions between variables. This could be solved by using the produced model (Equation 6.6) which overcomes the problems in the previously suggested CR policies.

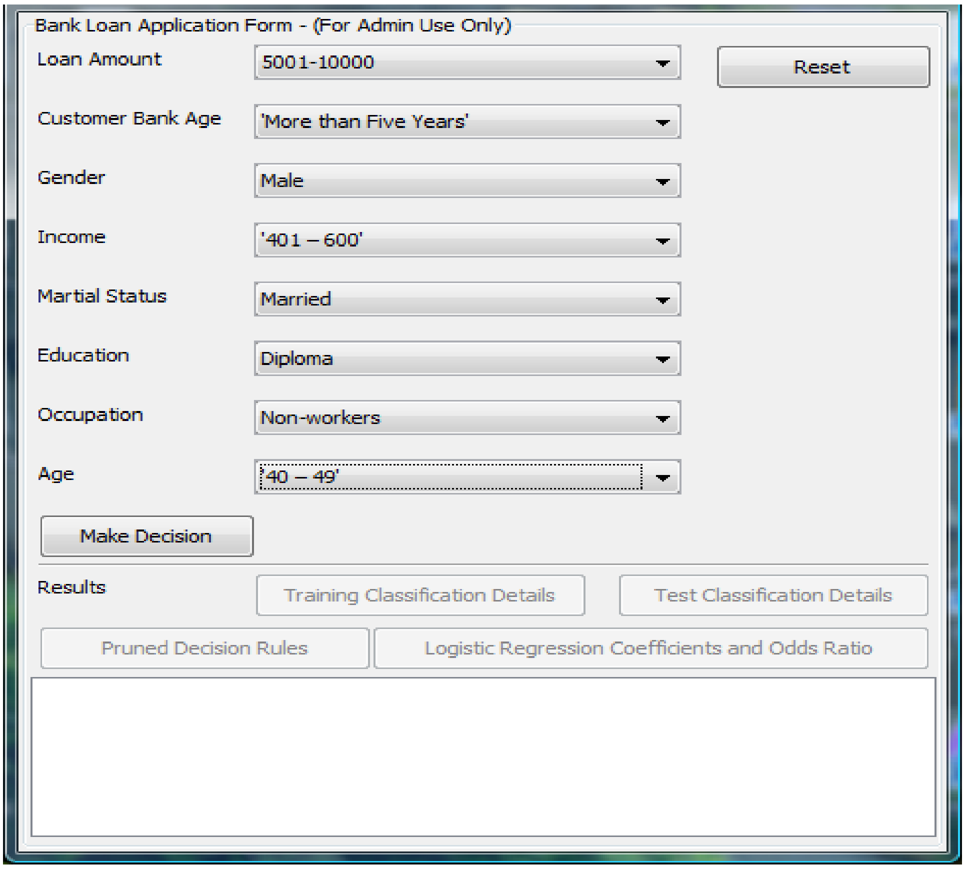
6.10 Automated Retail Loan Policy System

In the developed software application, the probability of perform and non performing for loan applicant can be predicted based on the coefficient values - substituted in the logistic regression model that has been created (see Section 6.7).

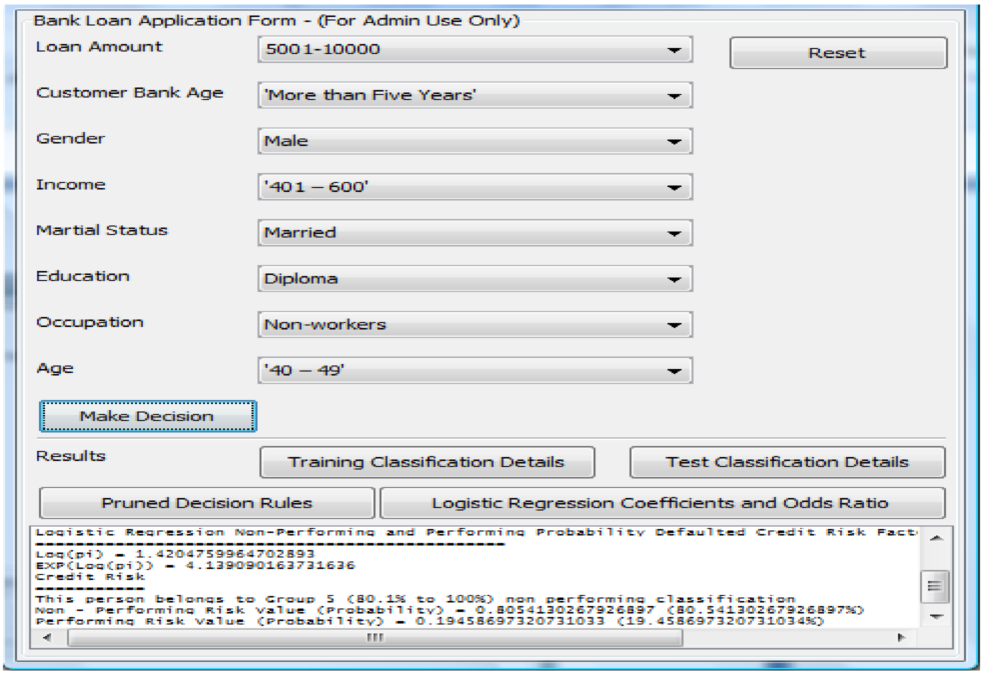
The following snapshot shows the predicted perform/non perform probability in the software application for a loan applicant. For testing purpose, random entries from the training data have been taken.

- Steps to perform this task:

1. Enter the parameter values by selecting values from drop boxes. For example as shown in Figure below.



2. Click make decision button to submit the input/values for predictor variables.
3. The results will be shown in text window on result Section of application window.



4. The non-performing probability shown here is 80.542% for this loan applicant and the group value of 5 – ranges from 80.1 to 100%
5. Considering that the threshold is provided, the loan approval officer can easily decide upon approval.

The CR classification system that has been developed and tested is a useful tool for credit officers in Jordanian banks to assess the risk of defaulting and develop appropriate strategies to minimise risk. The model that has been developed would help in the following ways:

- Improve the decision making process, making it more quantitative and dependable instead of using a judgemental-based system (based on previous experience). A judgemental system makes quantifying a risk a big challenge (Fensterstock, 2005) and combines many disadvantages; for example if for any reason the system does not work well, it becomes very difficult to determine which variables and weights need to be adjusted. The same problem can be solved easily using logistic regression.
- Provide a process to quantify risk which makes evaluating risk explicit, systematic and consistent. This will improve the speed of loan processing. As a consequence, customer service will be improved and more loans will be granted.

6.11 Conclusion

Policy makers at all levels are faced with new challenges of promoting maximum individual borrowers with minimum CR.

This Chapter started with a business understanding for the research problem. As a result, it has been found that the international average for default rates is lower than 5%, while the banking sector in Jordan estimated industry-wide non-performing loans at 11%. This indicates the critical problem that Jordanian banks face and the urgent need to address it.

In 2005, the CBJ asked Jordanian banks to adopt an internal CR policy. The interviews with managers in the banks have shown that most banks do not follow their credit policies in granting loans. Instead, credit officers and managers use their own judgements.

In Jordanian banks, analysis has shown the lack of analysing the internal implicit knowledge to use it in managing CR for retail loans and their need to build a CR policy using their internal knowledge to improve the process of managing CR (see Chapter five). They need also to improve their risk assessment systems and parameters.

This problem illustrates the need for Jordanian banks to build a CR system to help them to overcome this obstacle and increase their performance. The suggested solution was in building a CR classification system through logistic regression analysis.

Before building the risk classification system, a descriptive statistical analysis has been made of the 2775 customer data set. Several variables have been used including loan amount, customer bank age and other demographic variables such as age, income, educational level, occupation and marital status of borrowers.

Several CR classification models (see Appendix B) have been created and tested before reaching the most accurate model (Equation 6.6). Transformation for all variables has increased the ability of the final system to predict defaulting borrowers Using the credit classification system produced, the research questions have been answered:

1. Demographic variables such as income, age marital status, occupation and education have been shown to have an impact on the risk of default. While gender has not, income and marital status have been shown to play important roles in the probability of defaulting. The riskiest factor for defaulting in Jordanian banks is low income (less than JOD 600).
2. Non-demographic variables such as loan amount, bank location,and customer-bank-age also affect the risk of default. In this group, customer-bank-age (CAS) has been found as the main driver for risk of defaulting. Although knowledge about previous customers was implicitly present in Jordanian banks, they were not able to make use of it in managing CR.
3. The last question was about the implications of building a risk model on managing CR in Jordanian banks. The results have shown that, behaviour of Jordanian customers differs from international ones. Therefore the need for adopting a system using internal knowledge to manage CR is very important.

The CR model that has been developed and tested is a useful tool for credit officers in Jordanian banks to assess the risks of default and develop appropriate strategies to minimise them. The model that has been developed would help in the following way:

- Improve the decision making process, making it more quantitative and dependable instead of using a judgemental-based system (based on previous experience). A judgemental system makes quantifying a risk a big challenge (Fensterstock, 2005) and combines many disadvantages. For example, if for any reason the system does not work well, it becomes very difficult to determine which variables and weights need to be adjusted. The same problem can be solved easily using logistic regression.

- Provide a process to quantify risk which makes evaluating risk explicit, systematic and consistent. This will improve the speed of loan processing. As consequence, customer service will be improved and more loans will be granted.

The risk classification model thus emerges as a key contribution of this research. The produced system has shown the importance of using internal knowledge in managing CR and making decisions. Alongside with that, as an output of the research, typical Jordanian borrowers' risk characteristics have been determined.

Furthermore, the resulting credit classification system needs to be upgraded to include new data on default probabilities.

To conclude, empirical studies on credit loans default have shown that the major drivers of default risk are OCC (2), OCC(3), OCC(5), INC*LAM, CAS*AGE, EDU*INC, INC* OCC(4), INC*AGE, and AGE*MAR, while the other variables have a positive impact on the probability of not defaulting.

Chapter Seven: Using Data Mining to Support Business Intelligence in Managing Credit Risk

7.0Chapter Structure

Correctness, transparency and effectiveness are the principal attributes of knowledge derived from databases using data mining tools. In the current data mining research there is a focus on efficiency of algorithms for knowledge discovery derived from database (Daniels and Dissel, 2002).

As the aim of this research is to propose that developing a knowledge management (KM) approach to support managing credit risk will help banks in general, and Jordanian banks in particular, in improving the process of managing credit risk. DM as a KM technique (which is also known as Knowledge Discover in Database), will be used to reach following objectives:

1. To evaluate the proposed solution (see Chapter six),
2. To show the impact of using DM techniques in supporting CRM in Jordanian banks.
3. To classify and compare the predictive accuracy of the improved logistic regression model that has been created (see Chapter six) with the effectiveness of DM techniques, and
4. To provide the concepts and theories that should be reviewed and considered during the selection of the predictor variables for the development of any type of credit scoring system.

This Chapter is divided into five Sections. The first Section is an introduction to DM and its characteristics that differentiate it from typical statistical exploratory data analysis. The second Section illustrates the concepts of DM and its relationship with KDD. The third Section is a description for the CRISP-DM methodology that will be used in this Chapter. The fourth Section illustrates the process of building the credit risk model using DM techniques and finally, discussion of the findings is presented.

7.1 Introduction

Database usage is extensive, ranging from the mundane (such as Credit/Debit Card usage, Billing Systems, and Stock Control Systems) to less familiar uses (such as

Criminal Records, Satellite Navigation, and Scientific Research). According to Hand *et al.* (2001) “Progress in digital data acquisition and storage technology has resulted in the growth of huge databases”. Banks accumulate a large amount of data during their day to day activities. This data is useful for creating business strategy and hence to steer towards the long term objectives of the organisation. However, it is not essential that this data is always used or rather translated/mined in such a way that it can be used as knowledge set. With economic climate fiercely against the common person, the requirement of additional loan amount is at all time high, however, banks are turning down the loan applications that may adversely affect their profit margins and income from gains generated from interest charged

Unsurprisingly then, owners of databases are continually searching for increasingly sophisticated methods to extract valuable information from data stored within them. This process of extracting value from a database is referred to as *Data Mining*. Furthermore the process of deriving knowledge from extracted data is termed *knowledge discovery*.

Mining the data collected by banks and finding decision rules that can serve as guidelines for future transactions is very active research topic. Banks specifically can take advantage from an application that can accurately provide decision based on loan applicant’s profile and financial obligations.

Data mining is differentiated from typical statistical exploratory data analysis by two core attributes. Firstly, data mining tends to facilitate the use of large data sets. For this reason various constraints are placed upon data miners. These constraints include; storage and access to data, time constraints, determining the representativeness of the data, and validity of patterns or relationships identified (Hand et al., 2001). The second major distinction indicates that data mining techniques are generally applied to data sets which have been collected to answer specific questions independent of questions being researched by the data mining strategist. Interestingly, the distinction between data mining and statistical data analysis is often blurred and many individuals dispute the individual boundaries of each field.

7.2 DM & KDD

Recent research undertaken by Wang and Wang (2008) stated that “While data mining has been perceived to be a potentially powerful tool, the real benefit of data mining for business intelligence has not been fully recognized". The Information Technology community further recognizes that data mining can be damaging to business intelligence if the context of the data and application of the data is not understood by all individuals involved (Merkel, 2004). It is therefore imperative that individuals involved in the application of business intelligence recognise the core phases and principles involved in the manipulation and discovery of knowledge as described below.

In a business context, two typically unique "actors" are integrated into the data mining process. These roles are commonly termed as "Business Insider/Expert" and "Data Miner" (IBM, 2008). A business insider characteristically has grand knowledge of the business and often provides the business hypotheses or scenarios to which the data miner investigates. The business insider aims to apply the resulting knowledge (or business intelligence) to optimize or increase business performance (Wang and Wang, 2008). The role of a data miner therefore focuses on the recognition of patterns within datasets in addition to identifying how patterns relate to the business specific scenarios. Crucially then a data miner must understand the nature of a business or problem (Tabladillo, 2009). Typically, in a small business both roles may be fulfilled by one individual.

Creative and innovative knowledge creation and discovery, and knowledge deployment, must involve intuitive processes, both in the individual and in social interactions involving knowledge, if they are to be fruitful. Any hands-on data analyst will accept that perception and intuition are essential components of data exploration, necessary in addition to the routine and formal processes of training of logical/mathematical models.

The data mining (DM), data warehousing (DW) and knowledge discovery in databases (KDD) movements have had high profiles for the two last two decades years (for example (Fayyad, et al., 1996; and Mitchell, 1999). DM has tended to focus on the toolsets (models and fitting algorithms) that are available for drawing inductive inferences from datasets, and how such results can be deployed: a jointly empirical and rationalist approach. The CRISP-DM life cycle has become the commonly accepted gold standard of good-practice in DM-methodology in a knowledge and information management setting. This gold-standard of DM-methodology is largely based upon the sequential project-management paradigm of Software Engineering, in the Information Systems (IS) framework. Machine learning (M-L) techniques from AI were included in the under the KDD umbrella as means of achieving "intelligent" automaticity into KDD, as well as providing a range of specific algorithmic techniques to fill the DM tool-box. The KDD software facilities provided over the last few decades by the main KDD software producers (e.g. Oracle, IBM, Microsoft, SPSS, etc…) have, as a result, tended to data warehouses, with DM tool-boxes bolted on. Such an *ad hoc* KDD software support framework does not really provide an adequate conceptual, methodological or philosophical infrastructure to effectively enable the creative processes of knowledge discovery and exploration, or the communication and sharing of the resulting discovered knowledge. It is possibly a consequence of such inadequacies that many DW/DM/KDD projects have floundered. Efforts to enhance to relevance and applicability of DW and DM to business processes have led to the "Business Intelligence" (BI) movement.

The BI approach has been advocated as a means of enhancing the quality and timeliness of business decision making, and has grown largely from the aim to fully utilize data-warehouses, based on the underpinning IS techniques and technologies (e.g. microsoft, 2009). The knowledge products from DM have been seen as adding to the knowledge capital of an enterprise, and as an information resource which will strengthen the corporate decision making process.

7.2.1 DM Comparisons

In the literature describing business applications of data mining, Desai et al. (1996) explored the ability of neural networks and traditional methods, such as discriminant analysis and logistic regression, in building credit scoring models in the financial institutions. They studied datasets containing 18 variables collected from three credit institutions and showed that neural networks were particularly useful in detecting bad loans, whereas logistic regression outperformed neural networks in the overall (bad and good loans) classification accuracy.

In the same context, Barney et al. (1999) analyzed the performance of neural networks and regression analyses in differentiating between farmers defaulting on farmers Home Administration Loans. Using an unbalanced data, Barney found that neural networks outperform logistic regression in correctly classifying performed farmers.

7.3 CRISP-DM Methodology

There are different DM methodologies such as CRISP-DM, SEMMA, KDD, and Domain-specific methodology. CRISP-DM methodology makes large data mining projects faster, cheaper, more reliable and more manageable (see www.crisp-dm.org).

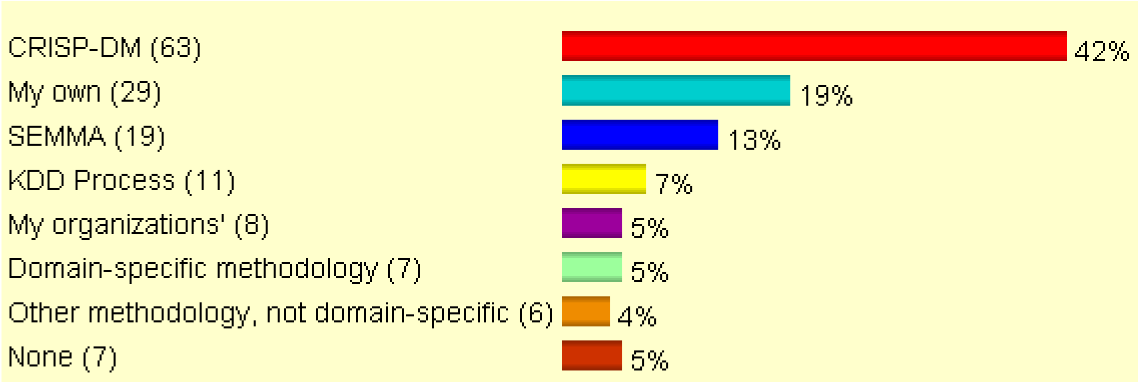


Figure 7.1 Methodologies used for data mining
Source: www.kdnuggets.com, (2009)

CRISP-DM proposes the following methodology for data mining (see fig. 7.2): (1) business understanding, (2) data understanding and data preparation, (3) modelling, (4) evaluation, and (5) deployment. Business understanding is critical as it identifies the business objectives and hence the success criteria of data mining projects. Further, as the term "data mining" implies, data is a crucial component, that is, no data means no mining. Hence, CRISP-DM includes data understanding and data preparation (for example, sampling and data transformation) as an essential antecedent for modelling.

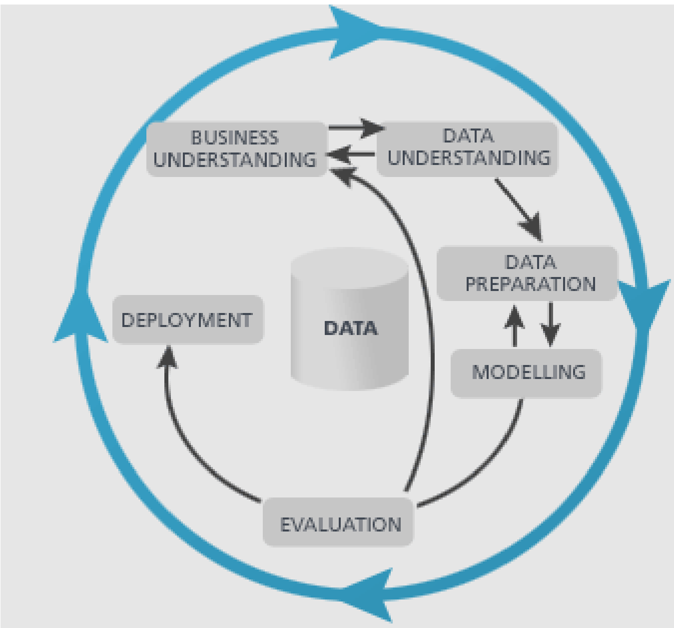


Figure 7.2: CRISP-DM Methodology
Source: www.crisp-dm.org, (2010)

The modelling stage is the actual data analysis. Most data mining software include OLAP (on-line analytical processing), traditional statistical methods (for example, cluster analysis, discriminant analysis and regression analysis) as well as non-traditional statistical analysis (such as neural networks, decision trees, link analysis and association analysis). This extensive range of techniques is not surprising given that data mining has been viewed as the offspring of three different disciplines, namely database management, statistics and computer science.

The evaluation stage allows the comparison of models and results from any data mining model by using a common yardstick (for example, lift charts, profit charts or diagnostic classification charts). Finally, deployment relates to the actual implementation and operationalisation of the DM models.

7.4 Building Credit Risk Model Using DM Techniques

In managing credit, a bank is interested in learning whether a potential costumer will pay back its credit. The aim of using DM methods in this Chapter is to model the probability with which an applicant can be categorized as a default or non default customer. If Jordanian banks can discriminate between these two groups, they can then use the predictive model to classify or predict new cases where they have the below information (Table 7.1) but do not know the person's credit standing. This will help decision makers, for example, to decide whether to qualify a person for a loan.

Data mining techniques can be broadly classified based on what they can do, namely: (1) description and visualisation; (2) association and clustering; and (3) classification and regression (that is, prediction). Using SPSS PASW Modeler, these techniques will be used to identify the factors or predictors that differentiate "risky" customers from others (based on patterns pertaining to previous customers), identify predictive techniques that perform well on test data, and later deploy those models to predict new risky customers (the first two steps of CRISP DM methodology have been described in Chapter five (see 5.2 and 5.3)). The following Table 6.1 provides the variables and their definitions (detailed discrete values for these variables are provided in appendix B).

Variable	Definition
Occupation	The current occupation of the loan applicant
Loan Amount	The loan amount that loan applicant is applying for
Customer Bank Age	The amount of time (years) the loan applicant has been with the bank
Gender	The gender of loan applicant
Age	The age (years) of the loan applicant
Education	The highest qualification achieved by the loan applicant
Income	The current income of the loan applicant in Jordanian currency
Marital Status	The marital status of loan applicant
Perform/Non Perform	The target attribute informing the loan applicants viability

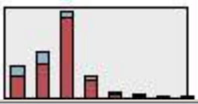
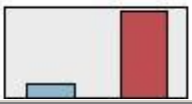
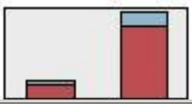
Table 7.1: Data Definitions

7.4.1 Data Preparation

With *PASW Modeler Data Miner*, it is straightforward to apply powerful modelling tools to data and judge the value of resulting models based on their predictive or descriptive value. This does not diminish the role of careful attention to data preparation efforts. Data is the main resource for data mining – therefore it should be prepared properly before applying any data-mining tool. Thus, it is important to pre-process the data and improve the accuracy of the model so that one can make the best possible decision.

The following steps summarize the data preparation during this stage:

1. Exploring the data: Descriptive statistics (by looking at distributions, means, minimum and maximum values, etc).
2. There are no outliers in the data, and
3. Dealing with missing values.

	Field	Sample Graph	Type	Min	Max	Unique	Valid
1	Loan-Amount		Set	1.00	8.00	8	2755
2	Customer-bank-age		Set	1.00	6.00	6	2755
3	performing-non-performing		Flag	0.00	1.00	2	2755
4	Gender		Flag	0.00	1.00	2	2755

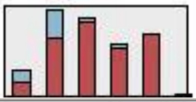
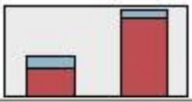
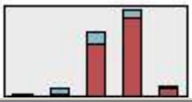
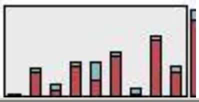
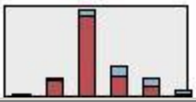
5	Ind Income		Set	1.00	6.00	6	2755
6	married (Y/N)		Flag	0.00	1.00	2	2755
7	Education		Set	1.00	5.00	5	2755
8	Occupation		Set	0.00	9.00	10	2755
9	Age		Set	1.00	6.00	6	2755

Table 7.2 Dataset audit

7.4.2 Modelling

Next stage is the modelling in which will present the analysis results to classify and compare the predictive accuracy of the improved logistic regression model that has been created (see Chapter six) with the effectiveness of DM techniques. To select the model, a simple flow-chart for model selection will be used (see Figure 7.4). Our models includes on target field which is PERF (flag variable). This guides us to use Logistic regression analysis.

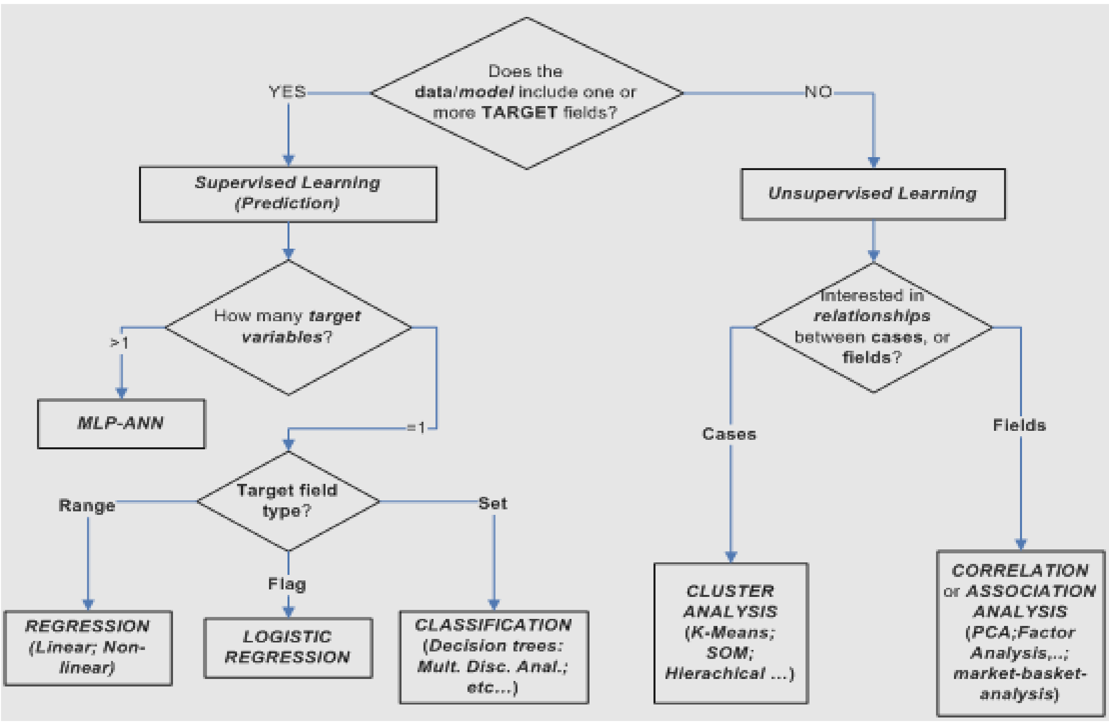


Figure 7.3: simple flow-chart for Model Selection

The following Section represents the Data Mining modelling process. Different intelligent methods and algorithms will be applied to generate models and patterns. Data Miner needs to analyze these patterns to get useful information. The following are the applied algorithms and my interpretations to the results.

7.4.2.1 Decision Trees

Decision trees are powerful and popular tools for classification and prediction. The fact that decision trees can readily be summarized graphically makes them particularly easy to interpret. A decision tree of rule based classifier is a predictive model where the transactions are mapped to conclude about that itemsets target value. Each node is the attribute and each edge presents the value for that attribute leading to a child. The target/predicted values is represented by the leaf. Therefore, that path from root to the leaf represents the itemset. One of the most important advantages of decision trees is the fact that knowledge can be extracted and represented in the form of classification (if-then) rules (Zurada and Lonial, 2005).

Decision trees can be less cognitive; however, in case of high risk scenarios such as banking credit scoring, it is but natural to allow for probabilities of classification (perform/ non-perform) to decide the credit approval. Four algorithms have been used namely, CHAID, QUEST, C&R Tree, and C5.0 (for full results see Appendix G).

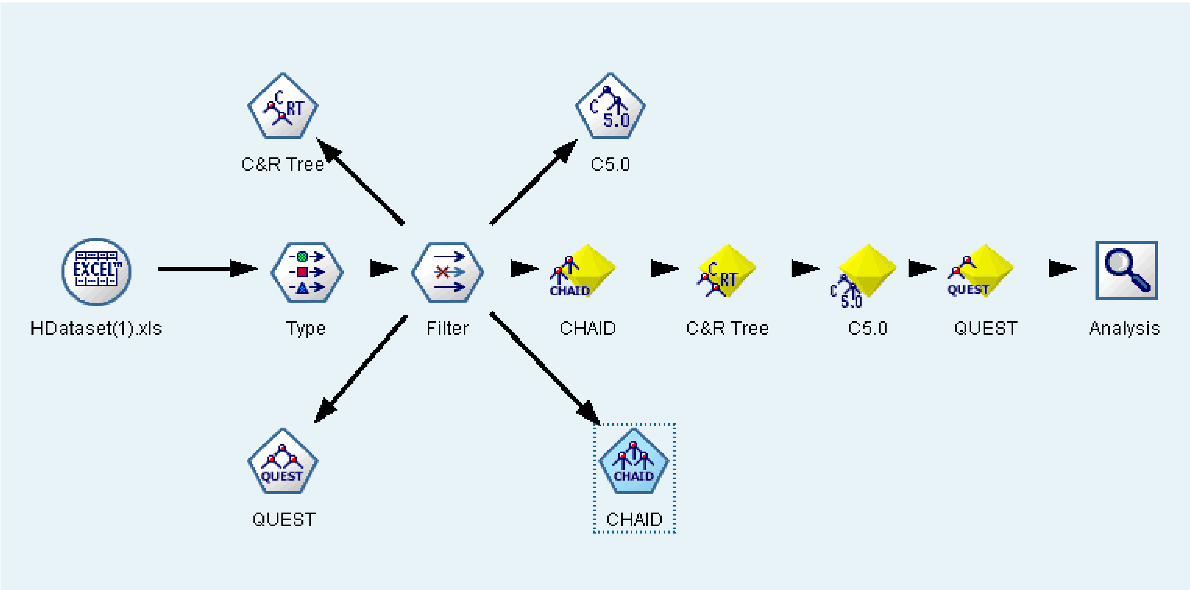


Figure 7.4: Decision trees stream

- *C&R Tree*

One of the interesting techniques is the Classification and Regression Tree node (C&R Tree) (Satchidananda & Simha, 2006). It is a tree-based classification and prediction method that uses recursive partitioning to split the training records into segments with similar output field values.

First, a C&R Tree node is connected to the Type node. The launch the Interactive Tree window will allow the data miner to grow and edit the tree before generating the model. After that, the Prune tree and Use standard error rule. Set the minimum change in impurity value to 0.003. Increasing this value tends to result in a simpler tree by preventing splits that result in only a very small improvement.

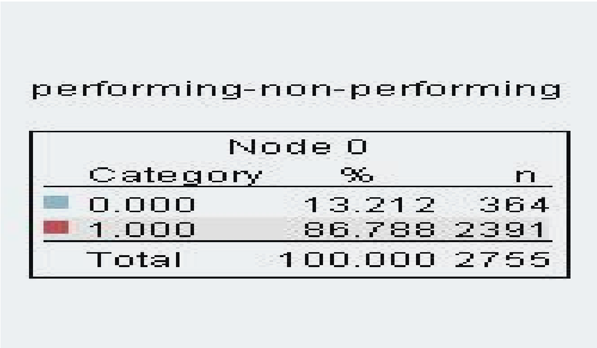


Figure 7.5: C&R Tree root node

Initially, only the root node is displayed (Figure 7.6). The statistics indicate that the training data has 2755 records. Since the tree has yet to split, all of them (100%) fall into this node. From the entire sample, there are 2391 perform customers and 364 non-perform customers. Next, the tree can be improved by targeting the subgroups most likely to perform or non-perform.

The resulting tree has six levels and nine terminal nodes (see Appendix G). If the pruning option hadn't been chosen, the tree would have been significantly more complex. Pruning is based on a cost-complexity algorithm that adjusts the risk estimate based on the number of terminal nodes.

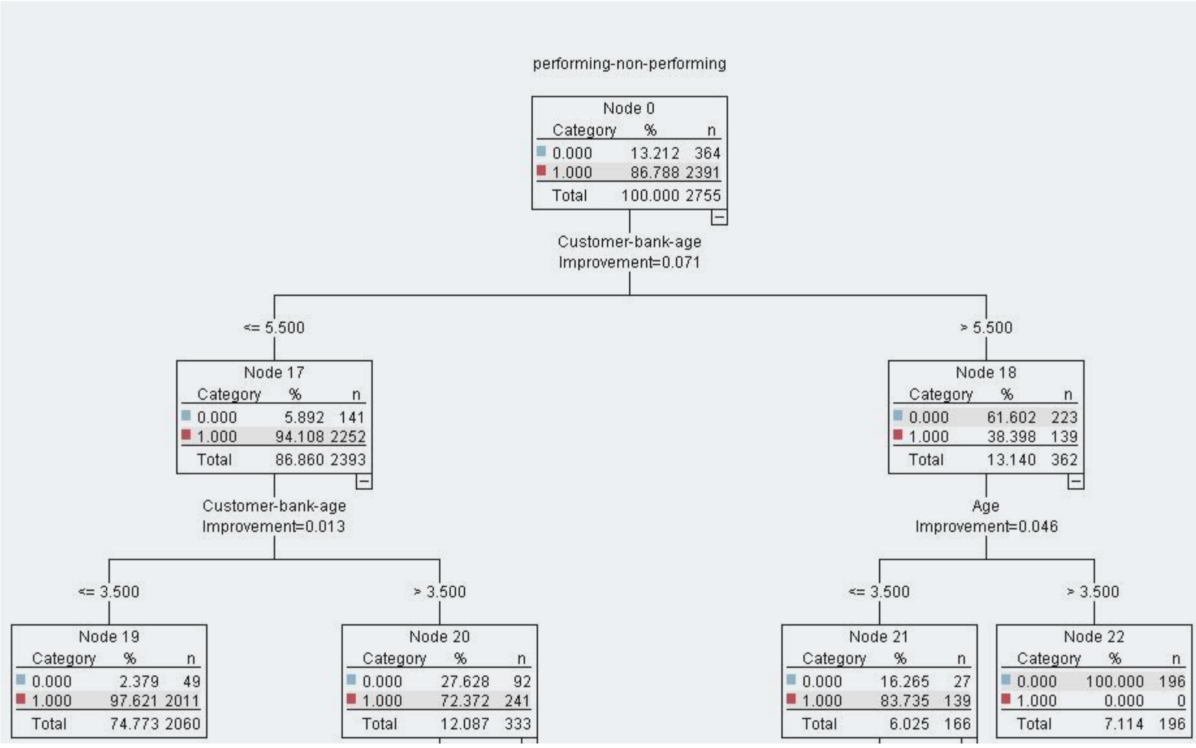


Figure 7.6: C&R tree sub-tree

Next, the gains Table for this tree was examined. The gain index percentage tells how much greater the proportion of a given target category at each node differs from the overall proportion.

	Nodes	Node: n	Node (%)	Gain: n	Gain (%)	Response (%)	Index (%)
1	22	196.0	7.11	196.0	53.85	100.0	756.87
2	24	70.00	2.54	70.00	19.23	100.0	756.87
3	31	13.00	0.47	13.00	3.57	100.0	756.87
4	30	31.00	1.13	23.00	6.32	74.19	561.55
5	32	57.00	2.07	8.00	2.20	14.04	106.23
6	29	24.00	0.87	2.00	0.55	8.33	63.07

Table 7.3: Interactive Tree of C&R Tree (target category 0)

Nodes with index values greater than 100% indicate that you would have a better chance of getting respondents who will not default by selecting records from these nodes instead of randomly selecting from the entire sample. Looking at the index values in Table 7.1, nodes 22, 24, and 31 have the highest possible rate for the entire dataset, with a value of nearly 756%. This indicates that Jordanian banks are almost 7.56 times as likely to get a non-perform customer with these records as by using random selection. As an example, the rule for node 31 is: If the customer has been a bank's customer for more than 3.5 years and his/er age is less than or equal 35 years and income is less than or equal J.D 600 and occupation form groups (2, 4, and 8) then the customer is more likely to default.

	Nodes	Node: n	Node (%)	Gain: n	Gain (%)	Response (%)	Index (%)
1	28	193.00	7.01	192.00	8.03	99.48	114.63
2	26	111.00	4.03	109.00	4.56	98.20	113.15
3	19	2060.00	74.77	2011.00	84.11	97.62	112.48
4	29	24.00	0.87	22.00	0.92	91.67	105.62
5	32	57.00	2.07	49.00	2.05	85.96	99.05
6	30	31.00	1.13	8.00	0.33	25.81	29.74
7	22	196.00	7.11	0.00	0.00	0.00	0.00
8	24	70.00	2.54	0.00	0.00	0.00	0.00
9	31	13.00	0.47	0.00	0.00	0.00	0.00

Table 7.4: Interactive Tree of C&R Tree (target category 1)

Looking at the index values in this category, nodes 28 has the highest possible rate for the entire dataset, with a value of nearly 114%. This indicates that Jordanian banks are almost 1.14 times as likely to get a perform customer with these records as by using random selection.

7.4.2.2 Neural Networks

According to Hornick *et al.* (2007) a NN is created by set of interconnected simulated neurons that can be used as computational models. The neurons are mathematical functions that an input and produce an output. The input is subjected to weight to produce an output which is in turn subjected to activation function to determine the threshold for classification of target variable as truly positive or truly negative (o or 1).

A better scoring provider used recently is a NN because of the predictive accuracy of the NN based model for decision making is high. Contrasting to statistical methods, neural networks do not depend on the assumptions regarding the independence and distribution of residuals or collinearity of input variables. On the other hand, a large volume of data is required for training, and the neural networks parameters (weights) offer little insight into the physics of the process. This is a disadvantage because weights cannot be easily converted to if-then rules which one can understand (Zurada and Lonial, 2005).

To build a NN model, using the Neural Network node, I constructed, trained, and validated a multiplayer feed-forward network with error back-propagation. An expert choice has been used with one input layer (17 neurons), two hidden layers (20 neurons and 15 neurons in order), and one output layer (1 neurons) (see the appendix G).

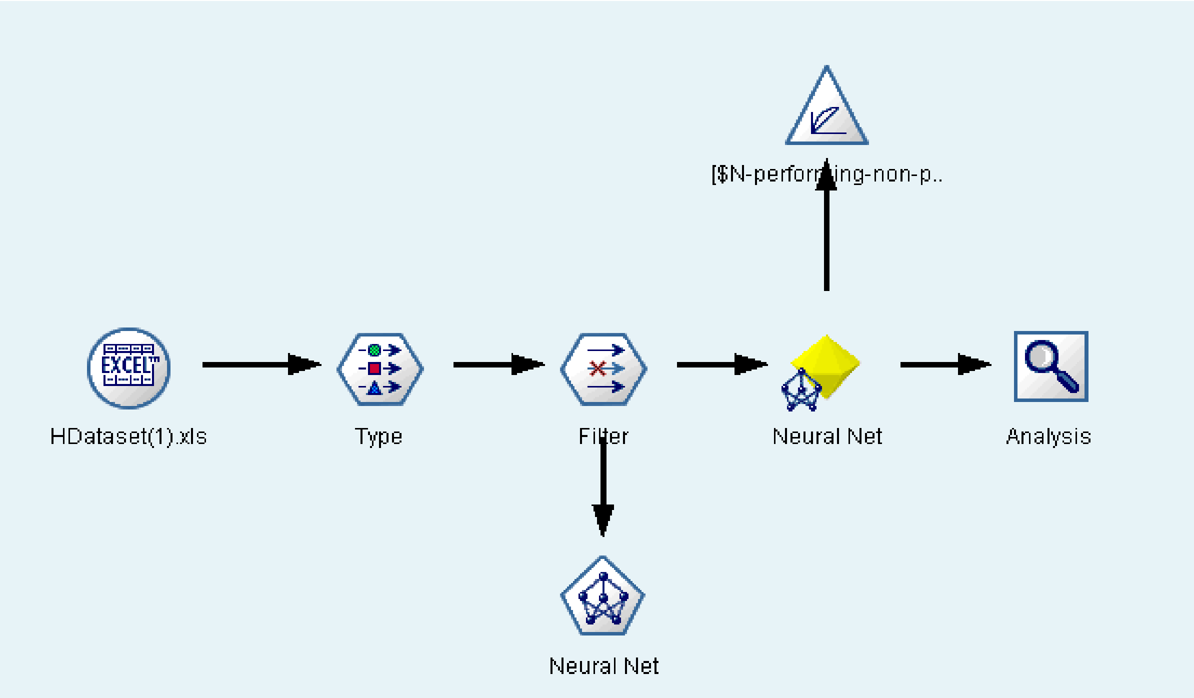


Figure 7.7: Neural Networks Stream

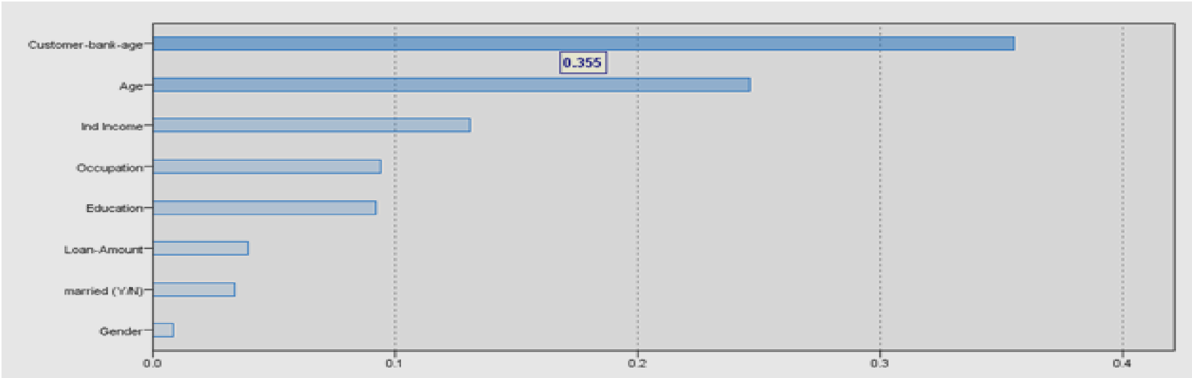


Figure 7.8: Neural Networks Variable Importance

One of the interesting features of DM techniques is calculating the variable importance (see Figure 7.9). According to the results, the most important variable that banks needs to take into account is CAS variable (0.355) which indicates how long a customer has been registered with the bank, the customer's age comes second (0.246), and third, Income variable (0.131). These results confirm with the results in the previous Chapter.

7.4.3 Evaluation

The models were evaluated for their effectiveness on several measures - percentage correctly classified (PCC), and Percentage Incorrectly classified (PIC), and lift charts (Satchidananda & Simha, 2006). The results of classification are shown in following Table.

Model	Observed	PCC	PIC	Overall accuracy
Logistic Regression	defaulters	89.3%	10.7%	94.25%
	Non defaulters	99.2%	0.8%	
CHAID	Defaulters	79.1%	20.9%	89.4%
	Non defaulters	99.7%	0.3%	
QUEST	Defaulters	82.7%	17.3%	90.55%
	Non defaulters	98.4%	1.6%	
C&R Tree	Defaulters	86.5%	13.55	92.5%
	Non defaulters	98.5%	1.5%	
C 5.0	Defaulters	81%	19%	90.4%
	Non defaulters	99.8%	0.2%	
Bayesian Net	Defaulters	87.9%	12.1%	93.25%
	Non defaulters	98.6%	1.4%	
Neural Networks	Defaulters	87.5%	12.5%	93.3%
	Non defaulters	99.1%	0.9%	

Table 7.5: Comparisons results between the improved logistic regression model and DM models

Table 7.5 compares the test set classification accuracy of logistic regression, CHAID, QUEST, C&R Tree, C 5.0, Bayesian Net, and NN. Among DM methods, NN outperformed the others (93.3%) which confirms with other researches (Desai *et al*., 1996; Barney *et al.*, 1999; and Zurada and Lonial, 2005). However, it is observed that the improved logistic regression model has outperformed DM methods on most parameters in case on Non Performing and by very less margin is similar to DM methods in performing classification. As we are working with bank and credit risk factor is crucial the risk of non-performing has to have a larger weightage and therefore logistic regression model is better fitting this data. More importantly, I attribute the outperformance of logistic regression model to the fact that binary classification (optimisation of linear hyper surface values are either high or low) has been used in data set as opposed to the other DM methods that works well for axis oblique hyper surface (where values can be continuous).

7.5 Conclusion

The main objective of this Chapter was to classify and compare the predictive accuracy of the improved logistic regression model that has been created with the effectiveness of DM techniques. Dislike previous researches (Desai *et al*., 1996; Barney *et al.,* 1999; and Zurada and Lonial, 2005) results have shown that the improved logistic regression model has outperformed DM methods on most parameters in case on Non Performing and by very less margin is similar to DM methods in performing classification.

Also, analysis has shown that DM can perform the following tasks:

1. Identify risk factors that predict defaulters and non-defaulters customers.

   One critical question in credit risk management is the following: “What are the risk factors or variables that are important for predicting the likelihood of claims and the size of a risk?” although many risk factors that affect rates are obvious, subtle and non-intuitive relationships can exist among variables that are difficult if not impossible to identify without applying more sophisticated analysis. DM techniques such as decision trees and NN can more accurately predict credit risk than actuarial models, therefore Jordanian banks can set more rates more accurately, which in turn can result a better competitive position.

2. Customer level analysis

   Successfully retaining customers requires analysing data at the most appropriate level, the customer level, instead of cross aggregated collections of customers. Using the

DM techniques, Jordanian banks can more accurately select which policies need to be established. With DM techniques Jordanian banks can:

- Segment the customer dataset to create a customer profiles.
- Dataset segmentation and more advanced techniques enable decision makers to more accurately choose whom to target.

The study has also introduced and described various characteristics, concepts and frameworks for data mining and knowledge discovery in support of business intelligence. The following recommendations summarize the core findings of the study:

1. Throughout the data mining process, the phases in CRISP-DM Methodology should be applied. Each phase of the framework must be completed sequentially. In particular it is recommended that attention must be given to the data preparation phase as this is often neglected. Neglecting this stage is liable to provide inconclusive or misleading results.
2. Before beginning the process of knowledge discovery it is integral that data miners and data analysts fully recognise the business context or scenario about which the project is being undertaken. This recommendation is particularly significant if the data mining process is outsourced. Failure to action this is likely to minimise the effectiveness of projects, with vital trends often missed, ignored or hypotheses misunderstood.
3. Throughout the "data mining" cycle communication between business insiders and data miners is critical. It is therefore advisable to encourage a number of communication channels to establish a platform for feedback on both current and future projects.
4. During the data visualisation phase, it is vital the context and end users are considered. Visualisations must be fit for purpose or it is probable that information will be mistranslated. It is also recommended therefore that if the end user is unknown, traditional techniques (such as graphs, Tables, simple diagrams) and the "keep it simple" philosophy should be applied.
5. Knowledge derived from data mining techniques must be fully scrutinised before being integrated into decision making processes and business intelligence. This process will minimise potential risk of misleading or incorrect data corrupting decision making procedures. This course of action should be integrated in addition to common risk management polices and functions applicable in general business scenarios.

Finally, regarding CRISP-DM, one of the main limitations is that it is essentially sequential and linear. Even though feedback loops are represented, the sequential nature of the representation suggests an ordering of the knowledge space, and its exploration, which does not appropriately characterize the hierarchical and interactive network features of corporate knowledge space, or the dynamics of knowledge discovery. For example, data understanding and data preparation are given a place between business understanding, and modelling. But, the representation of business understanding is essentially a prior model of the knowledge space and needs modelling techniques for its adequate representation. Similarly, data understanding depends on a suitable representation of available data and an evaluation of its suitability for business aims and DM modelling. Such evaluative procedures can really only be effectively conducted within the context of a pervasive model-based framework. Hence a more interactive and networked representation of the corporate knowledge space is required than is provided by the life-cycle paradigm.

To overcome the limitations of CRISP-DM methodology, a hierarchy of processes operating on the constructs at each level of the DIK hierarchy may also be defined, and is illustrated in Figure 7.9. Corresponding to the data-level of the DIK hierarchy, operations concerned with manipulations and storage of basic data resources we term “technical”. Information is derived by operations performed on data using the tools and techniques of DM, most of which can be regarded as model-based: we refer to the processes at this level as “technological” in Figure 7.9. Finally, derived and discovered information becomes knowledge only when it is subjected to processes which put the information in the context of the whole enterprise, and its aims. Hence the processes which deal such “understanding” of information/knowledge are termed “strategic” in Figure 7.9.

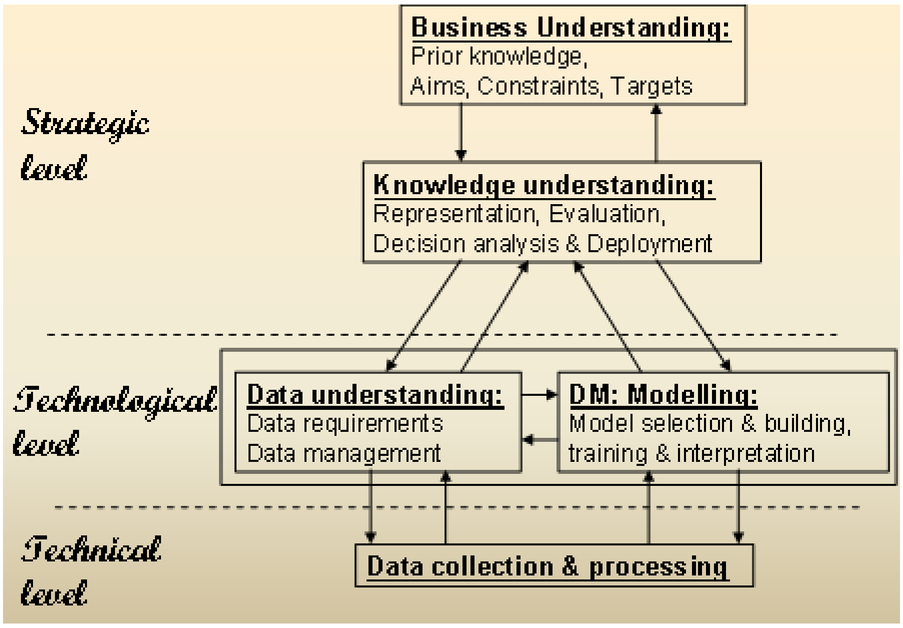


Figure 7.9: The DIK-process hierarchy

Source: Rennols and Al-Shawabkeh, (2008)

At the strategic level, the business and knowledge understanding processes are both complementary and overlapping. Appropriate and consistent representations of both prior business knowledge and discovered knowledge is needed if aims are to feed effectively into the knowledge discovery process, and discovered knowledge is to be deployed appropriately. Such communicable knowledge representations must rest on both data-understanding and model-based representations of the DM modelling process. Strategic level knowledge evaluation, decision analysis and optimization can only be effected in terms of target criteria defined on what must be essentially a model-based knowledge representation.

At the technological level of Figure 7.9, data-understanding processes (including data management and data requirements) are constrained by, and influence considerations at the strategic level. Understanding data requirements follows from aims coming from the strategic level and the models that might be used to represent and capture information and knowledge from the available or required data resources. Hence data understanding and modelling are overlapping concerns and should not be separated into distinct processes; the technological level should be regarded as a holistic whole, while at the same time being represented as the two component processes in Figure 7.9. The data collection and processing of the technical level in Figure 7.9 are concerned with mechanisms of data capture, and their storage in appropriate forms in databases and warehouses, informed by considerations of data understanding and modelling from the technological level.

This hierarchy of DIK processes (represented in Figure 7.9) contrasts with what has become the traditional and default views of DM, DW and BI in industrial and business practice. Within this traditional perspective, data collection, processing, storage and management are captured in operational databases by technologies (DBMS) of the IS discipline. DW may be seen within the IS paradigm as the primary response to targeting corporate information and knowledge needs. By the inclusion of data mining, modelling, and visualization (now often termed "analytics") within this paradigm, to extend the traditional SQL-based analytical tools for tabulation, "Business Intelligence" is reached.

This traditional perspective, in relation to DM and modelling, has the unfortunate consequence that the central importance of model-based knowledge representations to the information and knowledge flows, and to model-based strategic level functions, are not fully acknowledged.

Chapter 8: Conclusions and Future Research

8.1 Summary of Findings and Implications

By the end of 1989 the Jordanian economy faced a major crisis when the Jordanian Dinar (JOD) suffered a large devaluation (i.e. the JOD lost 51% of its value against the US Dollar), total government debt reached 197 per cent of GDP and the country's reserves of foreign currency declined sharply (Isik et al., 2004). Consequently the Jordanian banking sector suffered from problems such as an increase in the ratio of non-performing loans and this led to serious consequences for the banking industry and the collapse of many Jordanian Banks (e.g. Petra Bank, Islamic National Bank, Amman Bank) (Al-Jarrah, 2002). While the international average for default rates is lower than 5% (CBJ, 2005), based on the interviews done by the researcher, the banking sector in Jordan estimated industry- wide non-performing loans at 11%, which indicates the critical problem that Jordanian banks face and their need for finding effective solutions.

Therefore, the overall aim of this research is to propose that developing a knowledge management (KM) approach provides credit risk department in banks with new opportunities in decision making process that support managing credit risk.

Several scholars and studies have investigated the importance of knowledge resources in the banking industries (Kridan and Goulding, 2006; Salmador and Bueno, 2007; Karkoulian et al, 2008). However, no research has been conducted on the KM capabilities in supporting CRM in the banking industry. Therefore, this study can be considered a pioneer in this area. As a result of this research, the following have been achieved:

First, a framework has been developed which incorporates both tacit and explicit knowledge as part of its measurement process. The well established CSFs are categorised appropriately so that a measurement process can be established. The techniques of Multi-Criteria Analysis (MCA) have been used to develop indicators where the measurements can relate to both quantitative and qualitative features of KM activities. According to different researchers (Jennex and Zakharova, 2005, Chong, 2006) most of KM-CSFs were identified through qualitative research with their importance established through structured analysis. According to them, future research is needed to consolidate these factors into a single CSF model.

The overall objective of the developed framework is to provide a method for the measurement of the performance of activities in KM which should lead to identification of areas where action needs to be taken. Comparisons of organisations by sector or within a sector can also be achieved by suitable extensions.

Second, this framework has been transformed into a questionnaire to answer research Q2 (Chapter 1) (Are Jordanian banks efficient in terms of managing knowledge in CRM departments?), all Jordanian banks were included in the survey setting. The study concentrates on studying the Jordanian banks only. According to the CBJ (2007) report, in Jordan there are fourteen local commercial and two Islamic banks. The sixteen banks were approached by the researcher and twelve agreed to be part of this study, which represents 75% of the banking community. The sample is comprised of 242 respondents from twelve banks within CRM departments in Jordanian banks. The participants answered a survey questionnaire structured in Likert format.Data gathered from this research instrument were then computed for interpretation. Along with primary data, the researcher also made use of secondary resources in the form of published articles and literature to support the survey results.

Broadly speaking, the overall performance indicator for CRM departments in banks in Jordan shows the health of the KM activities. Consideration of the performance of the KM factors and criteria in banks in Jordan shows that they need to make more effort to overcome gaps in managing KM as follows:

1. For the ‘People & Culture’ factor, Jordanian banks need to identify the core CRK required to maintain their competitive advantages. Additionally, Jordanian banks need to restructure their training plans on how to manage knowledge. For effective KM, skills development should occur in the following areas: communication, soft networking, peer learning, team building, collaboration and creative thinking (Horak, 2001).

The banks should also establish an environment emphasising knowledge sharing and encouraging employees to form such a culture through creating communities of practice and reward systems and set up comforTable work breaks. Uzzi and Lancaster (2003) stated that internal relationships have an effect on the knowledge sharing in an organisation. Furthermore, enhancing collaboration and communication increase the performance of managing risk (Waldvogel and Whelan, 2008).

2. For the ‘Processes’ factor, results indicate that standard processes have not been established well for knowledge contribution and content management. This suggests that there is room for improvement. Jordanian banks need to identify KM processes that support CRM more clearly (especially knowledge of mathematical modelling). This will reduce poor application of knowledge. The interpretation of rules in, for example, granting loans will become more effective. However, it is found that the CRM departments in Jordanian banks have the abilities to access, structure and categorise the content of credit risk KM which will help their employees to transfer CRK from tacit to explicit and vice versa.

3. For the ‘IT’ factor, its contribution is found to be average. Banks in Jordan have built a robust and user-friendly technology which will affect positively their performance in benefiting from the knowledge that resides in IT but they have failed to establish effective tools that manage knowledge cycle practices. To overcome this, there is a broad collection of IT supports for KM which can be used and integrated into Jordanian banks’ technological platform which could be categorised into one or more of the following: business intelligence, knowledge base, collaboration, content and document management, portals, customer relationship management, data mining, workflow, search and e-learning (Wong, 2005).

Regarding the enablers of KM in Jordanian banks (Q3), results have shown that the main enabler in Jordanian banks is the ‘IT’ factor. This indicates its importance in managing knowledge and supporting managing CR. Enhancing the performance of employees by training them in how to manage knowledge using IT, supporting them from the top management by providing them with suitable information technologies and facilitating the culture to accept the new IT are crucial to increase the contribution of ‘IT’ to KM in general and managing CRK in particular.

Third, to answer research questions 5, 6 and 7, historical data of previous retail loans (2755 cases) have been analysed using logistic regression and DM techniques. In Jordanian banks, analysis has shown the lack of analysing the internal implicit knowledge to use it in managing CR for retail loans and their need to build a CR policy using their internal knowledge. The lack of analysing the internal implicit knowledge illustrates the need for Jordanian banks to build a CR system to help them to overcome this obstacle and increase their performance. The suggested solution was in building a CR classification system using logistic regression analysis.

Using the produced credit classification system, Q5 (What are the variables that influence the risk of loan default in Jordanian banks?) has been answered (see Section 6.4):

1. Demographic variables such as income, age, marital status, occupation, and educational have been shown to have an impact on the risk of default, while gender has not.

   Income and marital status have been shown to play important roles in the probability of defaulting. The riskiest factor for defaulting in Jordanian banks is low income (less than JOD 600).

2. Non-demographic variables such as loan amount, and customer-bank-age also affect the risk of default. In this group, customer-bank-age (CAS) has been found as the main driver for risk of defaulting. Although knowledge about previous customers was implicitly present in Jordanian banks, they were not able to make use of it in managing CR.

Regarding research question 6, results have shown that using transformation for potential risk variables has increased the ability of the final system adopted to predict defaulting borrowers. An important question for both banks and their regulators is the accuracy of a model's forecasts of credit losses (Shu-Ling *et al,* 2009). In this research, the accuracy of the model in predicting non-performing customers is 89.3%. Comparing this result with the previous produced model (see Equation C1 in Appendix C) (was 79.1%). It shows the improvements made on the model using transformation of variables (see Section 6.6).

The last question (Q7) was about the implications of building a risk model on managing CR in Jordanian banks. The results have shown (see Section 6.9.2) that behaviour of Jordanian customers differs from international ones. Therefore the need for adopting a system using internal knowledge to manage CR is very important. The CR model developed and tested is a useful tool for credit officers in Jordanian banks to assess the risk of default and develop appropriate strategies to minimise them. The model developed would help in the following way:

- Improve the decision making process, making it more quantitative and dependable instead of using a judgemental-based system (based on previous experience). According to Fensterstock (2005) a judgemental system makes quantifying a risk a big challenge and combines many disadvantages, for example if for any reason the system does not work well, it becomes very difficult to determine which variables and weights need to be adjusted. The same problem can be solved easily using logistic regression.
- Provide a standard KM technology-based solution to support knowledge workers in credit risk departments in banks.

- Provide a process to quantify risk which makes evaluating risk explicit, systematic and consistent. This will improve the speed of loan processing. As a consequence, the customer service will be improved and more loans will be granted (Shirley *et al.*, 2009).

With regard to the additional research hypotheses formulated in Chapter 5, this research also found:

- Testing the first hypothesis

H2: "*There is a significant impact of the factors 'Processes and IT' on the factor 'People and Culture'.*"

Results have shown that there is a significant impact for the factors 'Processes' and 'IT' on 'People & Culture' in the Jordanian banks. Additionally, results indicate that Processes factor's impact on the contribution of 'People & Culture' is greater than 'IT'. These results assert on the priority to establish standard processes for knowledge contribution and content management. This will increase the Jordanian banks' abilities to structure, categorise, and access the content of knowledge. In the same context, the positive impact of 'IT' on 'People & Culture' indicates the important role that 'IT' plays in enhancing the contribution of KM practices in general and 'People & Culture' in particular.

- For the consideration of the second hypothesis:

H2: "*There is a significant difference between the Jordanian commercial banks and Islamic banks in the performance of KM practices."*

Results indicate that commercial banks are managing CRK better than Islamic banks. Results further show that 'People & Culture' and IT factors are the main causes for the low performance of KM practices in Islamic banks. The contributions of these two factors are average in terms of our classification.

- Finally, question 4 has been transformed into a hypothesis *"There is a significant relationship between KM and CRM"*, results have shown (see Table 5.36) the negative relationship between KM and PCL. This result indicates the importance of managing knowledge in supporting CRM in Jordanian banks.

8.2 Policy Implications and Recommendations

In the light of the main aim of this research, which was to develop a KM approach to support managing credit risk to help banks in supporting CRM, it could be argued that since one of the main objectives of the CBJ is to ensure the safety and soundness of the banking system. This study may therefore help the CBJ in its efforts to improve the overall performance of the banking sector and identify the causes of inefficiencies in managing CR. As an implication, CBJ needs to encourage Jordanian banks to build a KM department to support different banking operations in general and CRM in particular. In the same context, the lack of sharing risk knowledge can create issues in the CRM processes and the controls may not be enough. Therefore, banks need to create the culture that everyone is responsible for managing credit risk. Weak means for transferring knowledge can provide inadequate knowledge of the operation, poor assessments of the lessons learned and poor understanding of the present and forecasts through risk knowledge (Shu-Ling *et al*, 2009).

Moreover, the Islamic banks in Jordan are just two banks and these banks are suffering substantial KM inefficiencies, this might encourage the CBJ to increase the number of Islamic banks to improve competition within Islamic banks sector.

Another implication comes in light of using the KM scale model that has been developed; this could help CBJ in adopting efficiency scores for Jordanian banks according to their efficiency in managing KM in credit risk department.

A final implication, the study has highlighted some positive determinants of efficiency in managing knowledge (such as IT) to support managing CR that could benefit bank owners and managers in improving the level of efficiency in their organisations.

8.3 Contribution of the Research

Banks have faced difficulties over the years for a multitude of reasons but the major cause of serious banking problems can be traced back to poor CRM. As a result, many banks suffered a huge loss from a steady increase of defaults and bad loans. The lack of knowledge amongst senior executives about the level of risks taken in sub-prime lending, the resulting 'toxic assets' and the global nature of the instruments used to spread risks is said to be the main contributing reason for the current worldwide crisis in banks. Banks in Jordan, the focus of this research, are not immune from the exposure to the risks.

Therefore, the aim of this research is to propose that developing a knowledge management (KM) approach to support managing credit risk will help banks in general, and Jordanian banks in particular, in improving the process of managing credit risk.

8.3.1 Contribution to KM Implementation Approaches

Several scholars and studies have investigated the importance of knowledge resources in the banking industries (Kridan and Goulding, 2006; Salmador and Bueno, 2007; Karkoulian et al, 2008). However, so far no research has been conducted on the KM capabilities in supporting CRM in the banking industry. Therefore, this study is a new and useful contribution to this important area.

According to researchers (Storey and Barnett, 2000; and Wong and Wong, 2008), many Knowledge Management (KM) systems failed to give the results or outputs expected. The reasons for this being a lack of careful understanding of knowledge, knowledge management and its processes. Most of the analyses conducted in this part has focused on the nature of this problem(KM system implementation failure) and previous research that have been done (see 2.3.2.1). As a result of this analysis, a new KM approach has been created and tested which is an important contribution of this research (see Figure 8.1)
Using the KM new approach, my work has provided a theoretical and practical understanding of the process KM should be implemented to help banks in general, Jordanian banks in particular, to support managing credit risk. Therefore, this study will complement the existing KM international banking literature, which is currently

significantly skewed towards developed countries. Also, a guide on how to implement this approach has been produced (see Section 3.2).

The produced KM approach is not restrictive in its structure to CRM area in Jordanian banks. Therefore, this approach could be also generalised to be used in other countries.

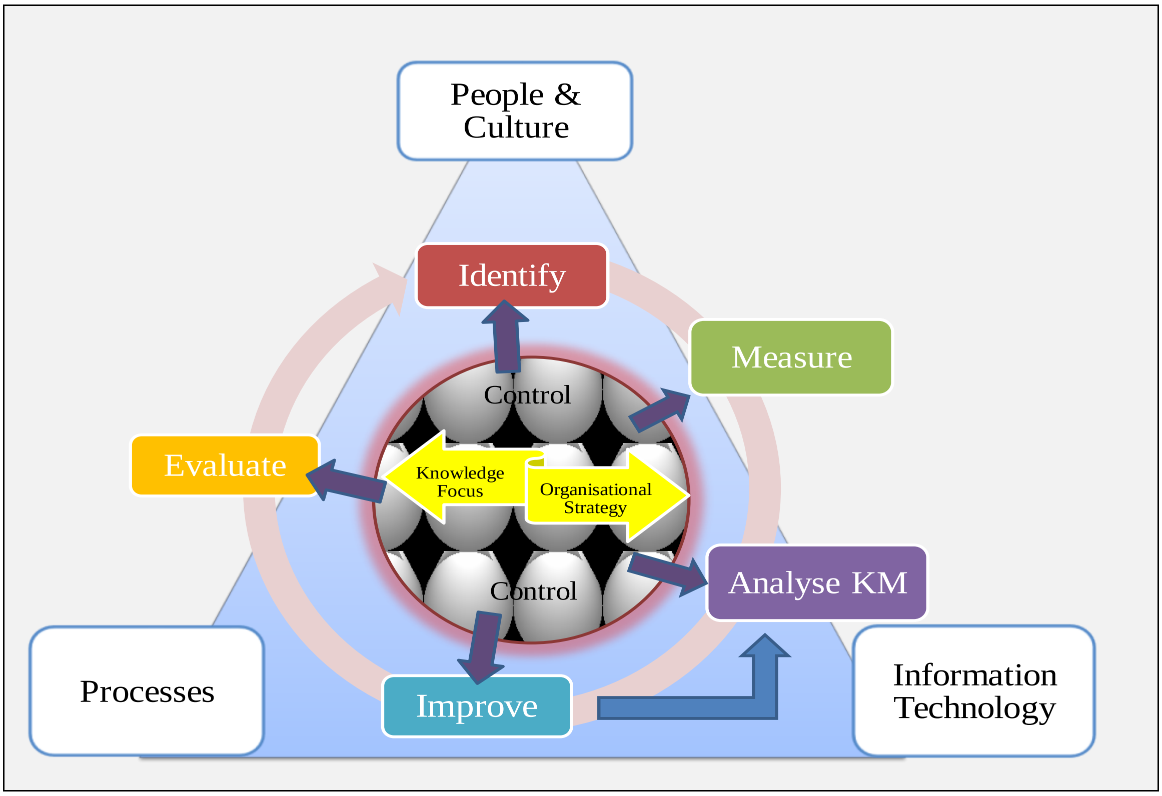


Figure 8.1: CR-KM Approach Cycle

8.3.2 Contribution to KM Measurements Model

First part of this research focused primarily on methods that are able to give KM performance indicators incorporating both tacit and explicit knowledge in an established framework. The main output of this analysis is that there is no well constructed model that is able to meet the requirements of this research (for full details, see Chapter four). In the same context, according to different researchers (Jennex and Zakharova, 2005, Chong, 2006) most of KM-CSFs were identified through qualitative research with their importance established through structured analysis. According to them, further research is needed to consolidate these factors into a single CSF model which could be used as a measurement model (this has been achieved in this research). By filling this gap, the developed MCA-CSA framework, which incorporates both tacit and explicit knowledge and KM enablers as part of its measurement process, is an important contribution to knowledge in this area. For the first time we have a framework that reflects results of research in KM and provides a flexible system that can be easily incorporated in any organisation

The first step in the MCA-CSA model is to ask respondents, usually key staff, to assign weights for each factor and criterion in the model. Changing the weights makes using this model not specific for Jordanian banks. Therefore, the findings of this contribution (the scale model) could be generalised beyond Jordanian banks.

8.3.3 Contribution to the literature on KM Status in Jordanian Banks

Despite the vast literature on KM roles in improving business performance in the banking industry of the United States and Europe, and the rising empirical research in the context of developing countries (for example, see Berger and Humphrey, 1997; Goddard et al., 2001), very few studies have been conducted on KM efficiency in the banking industry in Jordan (see 2.2.2). Using the MCA-CSA model, I have been able to analyse the KM status in CR department in Jordanian banks, and suggest solutions to overcome them. Thus, this study contributes to new knowledge to the literature in this area.

Furthermore, this study has contributed to knowledge in linking KM status to CRM (which has been represented by percentage of bad loans to gross loans in 2007). Results have shown the positive relationship between KM and the efficiency of managing credit risk. As this conclusion has been reached using data from CR departments in Jordanian banks, then this result could only be claimed to be specific to banks in Jordan. Nevertheless, it is an important and useful study for Jordanian banks and increases the contribution to literature in this area.

8.3.4 Contribution to the Literature on examining the variables that influence the risk of loan default in Jordanian banks

The final contribution to knowledge is identifying the variables that influence the risk of loan default in Jordanian banks, and building a Credit Classification System using internal implicit knowledge from Jordanian banks:

1. Using Logistic regression and DM techniques, I have been able to identify the variables that influence the risk of loan default in Jordanian banks (this has not been done before). The results of (see Chapter 6 and 7) can be incorporated into Jordanian banks lending policy. Thus, these results emerge as a key contribution of this research.

According to the World Bank report (2010), Most of countries in the Middle East and North Africa display ***similar*** economic structures. As a result, the conclusions that have been reached (see 6.11 and 7.11) can be useful for Middle Eastern countries with similar economic structure.

2. The impact of using transformation for variables potential risk variables on the accuracy (predicting defaulters and non defaulters) of the produced model has been tested. The results have shown the improving on the accuracy of the produced model.

8.4 Research Limitations

There are some limitations to the research process. First, the produced CR classification system has been built using historical data from three banks in Jordan. Lack of electronic copy of the needed data in most of the national Jordanian banks makes it impractical to build a credit classification system using data from all banks in Jordan at this time.

Alongside that, one of the missing data in this dataset (this variable has not been provided by Jordanian banks) is the time to default. Knowing this variable helps in building a survival analysis. Survival analysis can be applied to approximate the time to default or to early repayment. It is possible that if the time to default is long, the acquired interest will pay off or even exceed losses resulting from default (Stepanova and Thomas, 2002).

8.5 Critical evaluation and further work

This research has been carried out to create a KM approach to support managing credit risk in banks in general and Jordanian banks in particular. Using this approach will support banks in developing internal models to better quantify their credit risks and improve the quality of information. In the same context, using this approach, beyond the standard Credit Risk Models as applied in large banks, will help where there is not enough historic or reliable quantitative data as needed for the CR models.

Alongside with that, the produced CR classification system has been built using historical data from three banks in Jordan. Further research to build a credit classification system using internal data from all banks in Jordan is recommended.

In the same context, one of the missing data in this dataset (this variable has not been provided by Jordanian banks) is the time to default. Knowing this variable helps in building a survival analysis. Survival analysis can be applied to approximate the time to default or to early repayment. It is possible that if the time to default is long, the acquired interest will pay off or even exceed losses resulting from default (Stepanova and Thomas, 2002).

Additionally, an area of research deserving attention concerns a KM comparison among different countries. It might be interesting therefore to carry out similar research over all Arabian countries (as in Europe) to compare banking sector status in implementing KM and building a CR classification system across different Arabian regions.

Another consideration which requires further research is the evaluation of the relative cost of Type I and Type II errors. A Type 1 error occurs when the model predicts non-defaulter customer and the customer actually is. A Type 2 error occurs when the model predicts a defaulter customer and the customer actually is not. Type 1 and Type 2 errors have different costs. Type 1 errors may lead to wrong decisions that may cause financial injuries. Type 2 errors may cause just additional investigations. Thus Type 1 errors have bigger cost than Type 2 errors. Relative cost of Type I and Type II errors must be considered in performance metrics.

Finally, an interesting area for future research is to further investigate the relationship between degrees of KM structure implementation within Jordanian banks and corresponding increases in their performance.

References

Abecker, A. ,Bernardi, A., Hinkelmann, K., Kuhn, O. and Sintek, M. (1998), "Toward a Technology for Organizational Memories", *IEEE Intellegent Systems,* 13(3), pp. 30-34.

Alavi, M., Leidner, E. (2001), "Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues", *MIS Quarterly*, 25(1), pp.107-136.

Alavi, M., and Leidner, E. (1999), "Knowledge Management Systems: Issues, Challenges, and Benefits," *Communications of the AIS* , 1(7), 37 pages.

Al-Jarrah, I. M. (2002), "Efficiency in Arabian Banking", PhD, University of Wales Bangor.

Alissa, S., (2007), "Rethinking Economic Reform in Jordan: Confronting Socioeconomic Realities", *Carnegie Middle East Centre*, Washington DC, No.7, 23 pages.

Allee, V. (1997), "The knowledge Evaluation: Building Organisational Intelligence", Boston: Butterworth-Heinemann, 159 pages.

Altman, I., Resti, A., and Sironi, A., (2006), "Default Recovery Rates: A Review of the Literature and Recent Empirical Evidence," *Journal of Finance Literature,* Winter, pp. 21-45.

Al-Ajlouni, A. (2008), "Developing Strategies for Islamic Banks to Face the Future Challenges of Financial Globalization". [Online]: http://ssrn.com/abstract=1116186, [Accesses April 17th 2009].

Al-Omari, G. (2003), "Developing a Model of Using KM to Achieve High Business Value at Jordanian Commercial Banks", PhD Thesis, Amman Arab University for Graduate Studies.

Al-Osaimy, M. and Bamakhramah, A. (2004), “An EarlyWarning System for Islamic Banks Performance”, *Islamic Econ.*, 17(1), pp. 3-14.

Al-Tamimi, H., and AL-Mazrooei, M. (2007), “Banks’ risk management: a comparison study of UAE national and foreign banks”, *The Journal of Risk Finance*, 8 (4), pp. 394-409.

Anantatmula, V. , and Kanungo, S. (2005), “Establishing and structuring criteria for measuring Knowledge Management Efforts”. *38th Hawaii International Conference on System sciences*, Track 8,Vol.8, Page 251.b

Apostolou, D., and Mentzas, G., (1998). “Managing corporate knowledge: a comparative analysis of experiences in consulting firms”, *Proceedings of the Second International Conference on Practical Aspects of Knowledge Management*, Basel, Switzerland.

Argote, L., Ingram, P., Levine, J., & Moreland, R., (2000), “Knowledge Transfer in Organizations,” *Organizational Behavior and Human Decision Processes*, 82(1), pp. 1-8.

Armisted, C. (1999), “Knowledge Management and Process Performance”, *Journal of Knowledge Management*, 3(2), pp. 143-154.

Arora, R., (2002). “Implementing KM: a balance score card approach”, *Journal of Knowledge Management*, 6(3), pp. 240–249.

Baesens, B., Setiono, R., Mues, C., and Vanthienen, J. (2003), “Using Neural Network Rule Extraction and Decision Tables for Credit-Risk Evaluation”, *Management Science*, vol. 49, pp. 312-329.

Bakshi, G., Dilip, M., and Frank Z., (2001), “Understanding the Role of Recovery in Default Risk Models: Empirical Comparisons and Implied Recovery Rates”, *Finance and Economics Discussion Series*, Federal Reserve Board of Governors, Washington D.C.

Barth, R., Caprio, G., and Levine, R., (2004), "Bank Regulation and Supervision: What Works Best? ", *Journal of Financial Intermediation*, 13(2): 205-248.

Barth, R., Caprio, G., and Levine, R., (2001), "Financial Regulation and Performance: Cross-Country Evidence", *Central Bank of Chile, Santiago*, pp. 113- 141.

Bani Hamdan, K., (2002), "the Analysis of the Relation Between Human Resource Information Systems and Intellectual Capital and its Impact on Achieving the Competitive Advantage: Practical Research in Jordanian Insurance Companies", Unpublished PhD, Baghdad University.

Barkley, L., Deborah, M., and Julia, R. (2001), "Certified Capital Companies: Strengths and Shortcomings of the Latest Wave of State-Assisted Venture Capital Programs", *Economic Development Quarterly*, 15(4), pp. 350-366.

Barney, K., Graves, F., and Johnson, D. (1999), "The Farmers Home Administration and Farm Debt Failure Prediction," *Journal of Accounting and Public Policy*, Vol. 18, pp. 99-139.

Basel Committee on Banking Supervision, (1999), "Best Practices for Credit Risk Disclosure", Basel Committee on Banking Supervision, Basel, [Online]: http://www.bis.org/bcbs/index.htm, [Accessed May 5th 2006].

Belbaly, N. (2005), "Knowledge Management Systems and NPD Performance: An Integrative Approach". *San Diego International Systems Conference (SICS).*

Bennett, R., and Gabriel, H. (1999), "Organizational Factors and Knowledge Management within large marketing departments: an empirical study", *Journal of Knowledge Management*, 3(3), pp. 212-225.

Beveren , V. (2002), " A model of Knowledge Acquisition that Refocuses Knowledge Management", *Journal of knowledge Management*, 6(1), pp. 18-22.

Bhatt, G. (2002), "Management strategies for individual knowledge and organizational knowledge", In *Journal of Knowledge Management*, 6 (1), pp. 31-39.

Boisot, M. (1998), “Knowledge Asset: Securing a Competitive Advantage in the Information Economy”, *Oxford University Press,* New York, NY.

Boisot, M. (1995), ‘‘Is your firm a creative destroyer? Competitive learning and knowledge flows in the technological strategies of firms’’, *Research Policy*, Vol. 24, pp. 489-506.

Barney, K., Graves, F., and Johnson, D. (1999), "The Farmers Home Administration and Farm Debt Failure Prediction," *Journal of Accounting and Public Policy*, Vol. 18, pp. 99-139.

Brooking, A. (1999), “Corporate Memories, Strategies for Knowledge Management”, *Thompson Business Press,* London.

Borghoff, U., and Pareschi, R. (1997), “Information Technology for Knowledge Management”. *Journal of Universal Computer Science,* 7(8), pp. 835-842.

Bourne, D. (1991), "Do we need Bramshill**?",** *Policing*, 7(3), pp.258-70.

Carter, C., and Scarbrough, H., (2001). “Towards a second generation of KM? The people management challenge”, *Education and Training,* 43(5), pp. 215–224.

Central Bank of Jordan, (2008), “Yearly Statistical Series”, *Central Bank of Jordan,* Amman. [Online]: http://www.cbj.gov.jo/pages.php, [Accessed June 14th 2009].

Central Bank of Jordan, (2007), “Yearly Statistical Series”, *Central Bank of Jordan,* Amman. [Online]: http://www.cbj.gov.jo/pages.php, [Accessed June 14th 2009].

Central Bank of Jordan, (2006), “Yearly Statistical Series”, *Central Bank of Jordan,* Amman. [Online]: http://www.cbj.gov.jo/pages.php, [Accessed June 14th 2009].

Central Bank of Jordan, (2005), “Yearly Statistical Series”, *Central Bank of Jordan*, Amman. [Online]: http://www.cbj.gov.jo/pages.php, [Accessed June 14th 2009].

Chabane, A., Laurent J., and Julien S., (2004), “Double Impact: Credit Risk Assessment and Collateral Value”, *Revue Finance*, Vol.25, pp. 157-78.

Charitou, A., Neophytou, E., and Charalambous, C. (2004), “Predicting Corporate Failure: Empirical Evidence for the UK”, *European Accounting Review*, Vol. 13, pp. 465-497.

Cheung, P-K, Chau, P, and Au, A., (2008), “Does knowledge reuse make a creative person more creative?”, *Decision Support Systems*, Vol. 45, pp. 219–227.

Chong, S. (2006), “A comparison of perceived importance versus implementation in Malaysian ICT companies”, *The Learning Organization*, 13(3), pp. 230-256.

Chong, S.C. and Choi, Y.S. (2005), “Critical Factors in the Successful Implementation of Knowledge Management”, *Journal of Knowledge Management Practice*, Vol. 6, [Online]: http://www.tlainc.com/articl90.htm, [Accessed February 16th 2006].

Chung, N., Kang, B., Kwan, S., Joung, S., and Lee, K. (2002), “Knowledge Management Performance Index Considering Knowledge Cycle Process”, *ECAI 2002 Workshop #5 on Knowledge Management and Organizational Memories*.

Covin, T., Hall, N., Smalt, S., and Stivers, B. (1997), “Harnessing corporate IQ”. *CA Magazine*, 130(3), pp. 26 – 29.

Creswell, J. (2003), “Research design: Qualitative, quantitative, and mixed methods Approaches”, *Thousand Oaks, CA: Sage Publications*, 246 pages.

Crauise O'Brien, R. (1995), "Employee involvement in performance improvement: A consideration of tacit knowledge, commitment and trust", *Employee Relations*, 17(3), p. 110- 120.

Daniels, H. and Dissel,H. (2002), "Risk Management Based on Expert Rules and Data-Mining: A Case Study in Insurance", *ECIS 2002*, Gdansk, Poland. [On line]: http://is2.lse.ac.uk/asp/aspecis/20020047.pdf. , [Accessed 2nd of April 2008].

Davenport, T. (1998), "Putting The Enterprise Into The Enterprise System." *Harvard Business Review*, 76(4), pp. 121-131.

Davenport, T. (1993), "Process Innovation: Reengineering Work through Information Technology", *Harvard Business School Press*, Boston, MA, 327 pages.

Davenport, T., De Long, D. & Beers, M. (1998), "Successful knowledge management projects", *Sloan Management Review*, 39(2), pp.43-57.

Davenport, T. and Prusak, L. (1998), "Working Knowledge: How Organizations Manage What They Know", *Harvard Business School Press*, Boston, MA.

Davis, M. (1998), "Knowledge Management, Information Strategy", *The Executive's Journal*, Fall, 1998, p11-22.

Debowski, S. (2006), "Knowledge Management", *Milton, QLD: John Wiley*, ISBN 0470805382.

Desai, S., Crook, N., and Overstreet, A. (1996), "A Comparison of Neural Networks and Linear Scoring Models in the Credit Union Environment," *European Journal of Operation Research*, Vol. 95, pp. 24-37.

Darroch, J. (2003), "Developing a measure of knowledge management behaviours and practices", *Journal of Knowledge Management*, 7(5), pp. 41-54.

Desouza, C. (2003), “Barriers to effective use of knowledge management systems in software engineering”, *Commun. ACM 46*, Vol. 1, pp.99–101.

Despres, C. & Chauvel, D. (2003),”KM AS A Chemin Faisant: The Valtech Experience”, *Electronic Journal of Knowledge Management*, 1(1), pp. 13-22.

De Gooijer, J. (2000), “Designing a knowledge management performance framework”, *Journal of Knowledge Management*, 4(4), pp. 303-310.

Digman L. (1990), “Strategic Management: Concepts, Decisions and Cases”, 2nd ed., Boston, Massachusetts: BPI Irwin, 148 pages.

Dixon, N. (2002), “Common knowledge: How companies thrive by sharing what they know”, *Boston: Harvard Business Press*, 187 pages.

Domeniconi, C., Peng, J., and Gunopulos, D., (2002), “Locally adaptive metric nearest-neighbor classification”, *IEEE Trans. Pattern Anal. Mach. Intell.* 24(9), pp.1281–1285.

Drucker, P., (1993), “Managing for Results”, reprint ed., Collins, London.

Drucker, P., (1991), “The new productivity challenge”, *Harvard Business Review*. November - December, pp. 69 - 79.

Dudoit, S., Fridlyand, J., and Speed, T, (2002), “Comparison of discrimination methods for the classification of tumors using gene expression data”, *Journal of the American Statistical Association*, Vol.97, pp.77–87.

Duffy, J., (2000), “Knowledge management: to be or not to be”, *Information Management Journal*, 34 (1), pp. 64-70.

Ermine, J. (2000), “Chalanges and Approaches for Knowledge Management in Companies”, *Workshop Knowledge Mnageement: Theory and Application*, Fourth European Conference on Principles and Practice of Knowledge Discovery in Databases, Lyon, France.

Edvinsson, L. and Malone, M. (1997), “Intellectual Capital: Realizing Your Company’s True Value by Finding Its Hidden Roots”. Newyork: Harper Business.

Eisenbeis, A. (1977), “Pitfalls in the Application of Discriminant Analysis in Business, Finance and Economics”. *Journal of Finance*, Vol. 32, pp. 875-900.

Fayyad, U., Uthurusamy, R., and Smyth, P. (1996), “The KDD Process for Extracting useful Knowledge from Volumes of Data”, *Communications of the ACM*, 39(11), pp. 27-34.

Fensterstock, A. (2005), “On the Advantages of statistical-based models for credit risk analysis”, The Credit and Financial Management Review – Third Quarter.

Fielden, T. (2001), "Knowledge management state of mind", *Infoworld*, 23(1), pp.47-58.

Flyvbjerg, B., (2006), “Five Misunderstandings about Case-Study Research”, *Qualitative Inquiry*, 12(2), pp. 219-245.

Fourie, L. and Shilawa, J. (2004), “The value of concept maps for Knowledge Management in the banking and insurance industry: A German case study”, *Proceedings of the First International conference on concept mapping*, Pamplona, Spain.

Gallagher, S., & Hazlett, S. A. (1999), “Using the Knowledge Management Maturity Model (KM3) as an Evaluation Tool” *Conference on Knowledge Management Concepts and Controversies* 10-11, University of Warwick, Coventry, United Kingdom.

Gao, F., Li, M., and Nakamori, Y., (2002). “Systems thinking on knowledge and its management: systems methodology for knowledge management”, *Journal of Knowledge Management*, 6(1), pp. 7–17.

Glendon, K. , and Kundtz, J. (2000), “Knowledge Management: Process, People and Culture as an enabler to improving Healthcare Performance”. Healthcare Information and Management Systems Society (HIMSS), Vol.2, pp. 108-115.

Grant, R. (1996), “Towards a Knowledge-based Theory of the Firm”, *Strategic Management Journal*, Vol. 17, pp. 109- 122.

Gruben, W. C. and R. P. McComb (1997), "Liberalization, Privatization and Crash: Mexico's Banking System in the 1990s", Federal Reserve Bank of Dallas, *Economic Review (First Quarter)*, pp. 21-31.

Hand, J., Mannila, H., and Smyth, P. (2001), “Principles of Data Mining”, *Cambridge, MA: MIT Press*, 546 pages.

Hariharan, A. (2002), “Knowledge Management: A Strategic Tool”. *Journal of Knowledge Management Practice*, 3(3), pp. 50-59.

Helfk, M.,Pocatera, C., and Weerdmeester, R. (2003), “Knowledge Management Maturity Model”. Information Societies Technology Program, [online]: http://www.kmsk.or.kr/admin/symposium/r_upload/A1-3_kankanhalli.pdf

Henley, W., and Hand, D. (1996), “A k-nearest neighbour classifier for assessing consumer risk”, *Statician*, 44(1), pp.77–95.

Heywood, I., Oliver, J. and Tomlinson, S. (1993), “Building an exploratory multi criteria modelling environment for spatial decision support”, *International Journal of Geographical Information Science*, 7(4).

Holmes, C., Adams, M. (2002), “A Probabilistic Nearest Neighbour Method for Statistical Pattern Recognition”, *Journal of Royal Statistical Society. Series B*, Vol. 64, pp. 295-306.

Holsapple, C., and Joshi K., (2002). “Knowledge management: a threefold framework”, *Information Society*, 18(1), pp. 47–64

Horak, B.J. (2001), "Dealing with human factors and managing change in knowledge management: a phased approach", Topics in Health Information Management, 21(3), pp.8-17.

Hui-Chung Yeh, Min-Li Yang, and Li-Chuen Lee. (2007), "An Empirical Study of Credit Scoring Model for Credit Card", Second International Conference on Innovative Computing, Information and Control (ICICIC 2007), pp.216- 220.

Hussein A., Al-Tamimi, H., and Al-Mazrooei,F., (2007), "Banks' risk management: a comparison study of UAE national and foreign banks", *The Journal of Risk Finance*, 8(4), pp. 394-409.

Inkpen, A. (1996), "Creating knowledge through collaboration", *California Management Review,* 39(1), pp. 123-140.

Isik, I.,Gunduz and Omran, M. (2004), "Managerial and Scale Efficiency in the MENA Banking: A Panel Study of the Jordanian Banking Sector", *Arab Monetary Fund*, Abu Dhabi.

Jarrow, A., (2001), "Default Parameter Estimation Using Market Prices", *Financial Analysts Journal*, 57(5), pp. 75-92.

Joao, B. (2008), "Credit Scoring with boosted decision trees", *MPRA Paper*, Munich: Munich University, pp. 1-14.

Jarrar, Y., (2002) "Knowledge management: learning for organisational experience", *Managerial Auditing Journal*, 17(6), pp. 322–328.

Jennex, M. and Olfman, L. (2004), "Assessing knowledge management success/effectiveness models", *Proceedings of the 37th Hawaii international conference on system sciences.* [online]: http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=1265571.

Joao, B. (2008). Credit Scoring with boosted decision trees. *MPRA Paper*, Munich: Munich University, pp. 1-14.

Kanwar, A. (2005), "Risk Management for Banks", *Market Forces*, 1(1). [online]: http://www.pafkiet.edu.pk/LinkClick.aspx?fileticket=rdQ0TdINVdg%3D&tabid=148&mid=1538

Kaplan,R. and Norton, (1996), “The Balanced Scorecard”, *Boston, Massachusetts: Harvard Business School Press*, 323 pages.

Karkoulian, S., Halawi, A. and McCarthy, V. (2008), ‘‘Knowledge management formal and informal mentoring: an empirical investigation in Lebanese banks’’, *The Learning Organization*, 15(5), pp. 409-420.

Kilian Q. , Blitzer J., and Saul. L. (2006), “Distance Metric Learning for Large Margin Nearest Neighbor Classification”, *In Advances in Neural Information Processing System (NIPS)*, Vol.17.

Kim, M., Kim, Y. and Yu, S. (2004), “Linking Organisational Knowledge Management Drivers to Knowledge Management Performance: An Exploratory Study”, *Proceeding of the 37th Hawaii Conference on System Sciences*.

Kleindorfer P, Kunreuther H, Schoemaker P. (1993), “Decision Sciences: An Integrative Perspective”, *Cambridge University Press*, Cambridge, 461 pages.

Klimko, G. (2001), “Knowledge Management and Maturity Models: Building Common Understanding”. *Proceedings of the 2nd European Conference on Knowledge.*

Kochikar, V. P. (2000), “The Knowledge Management Maturity Model- A Staged Framework for, Leveraging Knowledge”. *Infosys Technologies Ltd*. [online]: http://www.infy.com/knowledge_capital/Knowledge/KMWorld00_B304.

Kridan, A., and Goulding, S. (2006), “The Importance of Knowledge Management Implementation in the Banking Sector”, *The Journal of Information and Knowledge Management Systems*, 36 (2), pp.211-222.

Lai, H., and Chu, T. (2002) “Knowledge management: a review of industrial cases”, *Journal of Computer Information Systems*, 42(5), pp. 26–39

Lang, J.C. (2001), “Managerial concerns in knowledge management”, *Journal of Knowledge Management*, 5(1), pp. 43 – 59.

Langen, M. (2000), “Holistic development of KM with the KM maturity model", *paper presented at the APQC Conference*, Las Vegas, NV, 7-8 December.

Lee, L (2000), "Knowledge Sharing Metrics for large Organisations". *MIT Press 2000*, [online]: http://www.efios.com/pdf/bhp-metrics.pdf

Lee, T and Chen. A. (2002), " two-stage hybrid credit scoring model using artificial neural networks and multivariate adaptive regression splines. Expert Systems with Applications", 28(4), pp.743–752.

Liebowitz, J. (2000), "Knowledge mapping: an essential part of knowledge management", IRMA Conference 2000, pp. 650-653.

Lim, K (1999), "Do Firms Have to Perform Basic Research In Order to Capture Knowledge Spillovers: Evidence from the Semiconductor and Pharmaceutical Industries", *CCC Doctoral Consortium 1999*, New York.

Lubit, R. (2001), "Tacit Knowledge and Knowledge Management: The Keys to Sustainable Competitive Advantage", *Organizational Dynamics*, 29(4), pp. 143-156.

Lynn, L. (1998), "The Management of Intellectual Capital: The issues and the practice". *Management Accounting Issues Paper 16 Management Accounting Practices Handbook*, Society of Management Accountants of Canada, Hamilton, Ontario.

Malhotra, N.K. (2003), "Marketing Research, an applied orientation", 4th ed., *Pearson Education*, Upper Saddle River, New Jersey.

Markus, L. (2001), "Toward a Theory of Knowledge Re-use: Types of Knowledge Re-use Situations and Factors in Re-use Success" Journal of MIS vol 18 No. 1 Summer pp57-93.

Marshal, C. and Prusak, L. (1996), "Financial Risk and need for superior knowledge management", *California Management Review*, 38(3), pp 77-101

Merkel D. (2004), "Dangers of Data Mining", [on line] www.RealMoney.Com, http://www.thestreet.com/p/rmoney/davidmerkel/10162833.html, [Accessed October 2nd 2009].

Mertins K, Heisig P, Vorbeck J. (eds.) (2001), “Knowledge Management. Best Practice in Europe”, Springer-Verlag, Berlin.

Miller, F.J. (2002), “1 = 0 (Information has no intrinsic meaning)”, *Information Research*, 8(1), paper no. 140. Available from http://InformationR.net/ir/8-1/paper140.html. [Accessed June 25th 2007].

Ministry of Planning and International Cooperation MOP. (2007), “Access to Finance in Jordan”, [online]:

http://www.mop.gov.jo/uploads/Access%20to%20Finance%20Discussion%20Chapter%20final_arabic.pdf. [Accessed March 21th 2006].

Mitchell,T. (1999), “Machine Learning and Data Mining”, *Communications of the ACM*, 42(11), pp. 30-36.

Mondale, J., Scott, V., and Venters, W. (2006), “Knowledge Management as an Image of the Organization: Industry standards and processes of knowing in credit risk management”, Proceedings of the European Conference on Information Systems, Goteburg, Sweden.

Moffett, S., McAdam, R., & Parkinson, S. (2003), “An empirical analysis of knowledge management applications”, *Journal of Knowledge Management*, 7(3), pp. 6 – 26.

Moor, D and Smits, M. (2002), “Key Performance Indicators for Knowledge Management in a Community of Practice”. METIS Project Report TI/RS/2002/142, Telematica Instituut, Enschede.

Nonaka, I. (1994), "A dynamic theory of organizational knowledge creation", *Organizational Science*, 5(1), pp.14-37.

Nonaka, I. (1991), "The Knowledge Creating Company", *Harvard Business Review on Knowledge Management,* 69(6), pp. 21-45.

Nonaka, I., Takeuchi, H., (1995), "The knowldge-creating company: How Japanese companies create the dynamics of innovation", *Oxford university press*, New York.

Oakland J. (1995), "Total Quality Management-Text with Cases", *BH Ltd*, Oxford, UK.

Offsey, S., (1997). "Knowledge management: linking people to knowledge for bottom line results", *Journal of Knowledge Management,* 1(2), pp. 113–122.

Orr, E. and Persson, M. (2003), "Performance Indicators for Measuring Performance of Activities in Knowledge Management", *Thesis. School of Economics and Commercial Law*, University of Gothenburg- Department of Informatics, [online]: http://www.handels.gu.se/epc/archive/00002927/. . [Accessed February 2nd 2008].

Pasiouras, F., S. Tanna and C. Zopounidis (2007), "Regulations, Supervision and Banks' Cost and Profit Efficiency around the World: A Stochastic Frontier Approach", School of Management, University of Bath, Bath.

Pemberton, J. D. and Stonehouse, G. H. (1999), "Learning and knowledge management in the intelligent organisation", *Participation and empowerment: An International Journal*, Vol.7**,** pp. 131-144.

Pemberton, D., and Stonehouse, H. (2000), "Organisational learning and knowledge assets - an essential partnership", *The Learning Organization: An International Journal*, 7(4), pp. 184 - 194.

Penrose, E. (1959), "The Theory of The Growth of The Firm2, *Oxford University Press*, New York, NY.

Polanyi, M. (1958), "Personal Knowledge: Towards a Post-Critical Philosophy", University of Chicago Press, Chicago.

Polanyi, M. (1966), "The Tacit Dimension", *Routledge & Kegan Paul,* London.

Polgar, S., and Thomas, S. (1995), "Introduction to research in the health sciences", 3rd edition, *London: Harcourt Brace*.

Power, D. (2009), "Types of Decision Support Systems (DSS)", [Online]: http://www.gdrc.org/decision/dss-types.html, [Accessed January 12th 2010].

Richard, R. (2003), "Trade, International Competitiveness, and Entrepreneurial Cooperation: the case of Jordan", *Arab Bank Review*, 5(2).

Rubenstein-Montano B, Liebowitz J, Buchwalter J, McCaw D, Newman B, Rebeck K, (2001), "SMARTVision: a knowledge-management methodology", *Journal of Knowledge Management*, 5(4), pp. 300–310

Rogers, E. (2001), "A Theoretical Look at Firm Performance in High-Tech Organizations: What Does Existing Theory Tell Us?", *Journal of High Technology Management Research,* Vol. 12, pp. 39-61.

Roos, G., Roos, J., Dragonetti, N and Edvinsson, L. (1997). "Intellectual Capital; Navigating in the New Business Landscape", New York University Press, New York: Macmillan.

Ryan, S. and Prybutok, V. (2001), "Factors affecting knowledge management technologies: a discriminative approach", *Journal of Computer Information Systems,* 41(3), pp. 31-7.

Saaty, S. (1994), "Highlights and Critical points in the Theory and Application of the Analytic Hierarchy Process", *European Journal of Operational Research,* Vol.74, pp.426-447.

Satchidananda, S., & Simha, B. (2006), "Comparing Decision Trees with Logistic Regression for Credit Risk Analysis", *SAS APAUGC,* pp. 1-7.

Sanghani, P., (2009) "Knowledge Management Implementation: holistic framework based on Indian study", *Pacific Asia Conference on Information Systems (PACIS).* Available at http://aisel.aisnet.org/pacis2009/69

Siam, A. (2007), "Effects of Collateral Pledges in Reducing Credit Risks – Confronting Banks in Jordan, as Lending Institutions", *Journal of Social Sciences,* 3 (3), pp. 134-137.

Smits, M., Moor, A., (2004), "Measuring Knowledge Management Effectiveness in Communities of Practice", *Proceedings of the 37th Hawaii International Conference on System Sciences.*

Snowden, D. (2000). "Knowledge elicitation: indirect knowledge discovery: Part Two of Basics of Organic Knowledge Management", *Journal of Knowledge Management,* Vol 3(9).

Stake, E. (1995), "*The art of case study research*", *Thousand Oaks, CA: Sage Publications*.

Stewart, T. (1999), "Intellectual Capital: The New Wealth of Organizations", *Bantam Dell Publishing Group,* New York, NY.

Storey, J., and Barnett, E. (2000). "Knowledge management initiatives: learning from failure." *Journal of Knowledge Management,* 4(2), pp 145–156

Sunassee, N., and Sewry, D. (2003), "An Investigation of Knowledge Management Implementation Strategies", *Proceedings of SAICSIT 2003,* pp. 24 – 36, ISBN:1-58113-774-5.

Sveiby, E. (2000), "What is knowledge management?", [online]:

http://www.sveiby.com.au/KnowledgeManagement.html,[Accessed March 23rd 2006].

Tabladillo M. (2009), “Microsoft: Data Mining – Why Use It?”, [on line] www.msdn.microsoft.com , http://msdn.microsoft.com/en-us/library/ee412262.aspx , [Accessed January 2nd 2010].

Tambyrajah, A. and Al-Shawabkeh, A. (2007), “Developing Performance Indicators for Knowledge Management”, *Proceeding of the 8th European Conference on Knowledge Management*, Barcelona, pp. 972 – 981.

Turban, E. (2006), “Information Technology For Management”, *John Wiley & Sons Inc*, ISBN: 0471705225.

Uzzi, B. and Lancaster, R. (2003), “Social capital and the cost of business loan contracting”,.Proceedings, Federal Reserve Bank of Chicago, issue Apr, pp. 237-260.

Yeh, I-Cheng. and Lien, Che-hui(2009). “The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients”, *Expert Systems with Applications*, Vol.36, pp. 2473–2480 [online]: www.sciencedirect.com. [Accessed July 25th 2009].

Waldvogel, A., and Whelan, N. (2008), “Towards better financial risk learning”, *Journal of Risk management in financial institutions*, 4(4), pp. 382-393.

Wang H. and Wang S. (2008), “A knowledge management approach to data mining process for business intelligence”, *Industrial Management and Data Systems - Emerald Insight*, Vol. (1008), pp. 1- 11.

Webb, R. (2002), “Statistical Pattern Recognition”, *John Willey & Sons*, 2nd edition.

Wenger, E. (1998), "Communities of Practice: Learning, Meaning, and Identity", *Cambridge University Press*, 318 pages.

West, D. (2000), "Neural network credit scoring models", *Computers and Operations Research*, Vol.27, pp.1131–1152.

Wiig, M., (1999), "What future knowledge management users may expect", *Journal of Knowledge Management*, 3(2), pp. 155-65.

Wiig, K., De Hoog, R., and Van Der Spek, R., (1997). "Supporting knowledge management: a selection of methods and techniques", *Expert Systems with Applications*, 13(1), pp. 15–27

Wong, K. (2005), "Critical success factors for implementing knowledge management in small and medium enterprises", *Industrial Management & Data Systems*, 105(3), pp. 261-279.

Wong, W., and Aspinwall, E., (2004). "Knowledge Management Implementation Frameworks: A Review", *Wiley InterScience Knowledge and Process Management*, 11(2), pp. 93-104.

World Bank, (2009), "Middle East & North Africa Region, 2008 Economic and Perspectives – Regional Integration for Global Competitiveness", Washington D.C., [Online]: http://web.worldbank.org/WBSITE/EXTERNAL/COUNTRIES/MENAEXT/0,,contentMDK:21822147~pagePK:146736~piPK:226340~theSitePK:256299,00.html, [Accessed March 15th 2010].

World Bank, (2006), "World Development Indicators 2006", World Bank, Washington, DC.

World Bank, (2002), "Building knowledge economies", [online]: http://lnweb18.worldbank.org/ECA/ECSSD.nsf/a3b026a6ee1e272585256ad2007130d3/9e9735587b22d64285256bce005ddbad/$FILE/Building%20Knowledge%20Economies-final%20final.pdf. [Accessed March 19th 2007].

www.crisp-dm.org, (2010), [On line]: http://www.spss.com/SE/analyticsday/Data%20Mining%20for%20beginners.pdf, [Accessed February 17th 2010].

www.IBM.com , (2009), “Knowledge discovery and data mining”, [on line]

http://domino.research.ibm.com/comm/research.nsf/pages/r.kdd.html , [Accessed December 7th 2009].

www.kdnuggets.com, (2009), [On line]:

http://www.kdnuggets.com/polls/2007/data_mining_methodology.htm, [Accessed April 20th 2009].

Yahya, S. and Goh, W. (2002), “Managing human resources toward achieving knowledge management”, *Journal of Knowledge Management,* 6(5), pp. 457-68.

Yin, K. (2003). “Case Study Research Design and Methods”, *Thousand Oaks, California: Sage Publications.*

Zack, M. (1999), “Developing a Knowledge strategy”, *California Management Review*, 41(3), pp. 125-146.

Zurada, J. And Lonial, S. (2005), “Comparison Of The Performance Of Several Data Mining Methods For Bad Debt Recovery In The Healthcare Industry”, *The Journal of Applied Business Research,* 22(2).

Appendix A

Step	-2 Log likelihood	Cox & Snell R Squared	Nagelkerke R Squared
1	945.703[a]	.354	.654
2	736.379[b]	.402	.741
3	652.705[b]	.420	.774
4	571.523[c]	.436	.805
5	512.276[c]	.448	.827
6	443.914[c]	.462	.852
7	425.786[c]	.465	.859
8	415.640[c]	.467	.862
9	405.643[c]	.469	.866
10	392.845[c]	.472	.870
11	374.533[c]	.475	.877
12	362.965[c]	.477	.881
13	363.227[c]	.477	.881
14	350.341[c]	.480	.885
15	350.712[c]	.480	.885
16	345.406[c]	.481	.887

TableA1. Goodness of Fit

Table 2. Classification Table

Observed			Predicted PERF 0	Predicted PERF 1	Percentage Correct
Step 1	PERF	0	252	112	69.2
		1	0	2391	100.0
		Overall Percentage			95.9
Step 2	PERF	0	272	92	74.7
		1	41	2350	98.3
		Overall Percentage			95.2
Step 3	PERF	0	280	84	76.9
		1	22	2369	99.1
		Overall Percentage			96.2
Step 4	PERF	0	292	72	80.2
		1	19	2372	99.2
		Overall Percentage			96.7
Step 5	PERF	0	305	59	83.8
		1	21	2370	99.1
		Overall Percentage			97.1
Step 6	PERF	0	314	50	86.3
		1	28	2363	98.8
		Overall Percentage			97.2
Step 7	PERF	0	315	49	86.5
		1	25	2366	99.0
		Overall Percentage			97.3
Step 8	PERF	0	314	50	86.3
		1	19	2372	99.2
		Overall Percentage			97.5
Step 9	PERF	0	317	47	87.1
		1	23	2368	99.0
		Overall Percentage			97.5
Step 10	PERF	0	314	50	86.3
		1	21	2370	99.1
		Overall Percentage			97.4
Step 11	PERF	0	319	45	87.6
		1	20	2371	99.2
		Overall Percentage			97.6
Step 12	PERF	0	320	44	87.9
		1	18	2373	99.2
		Overall Percentage			97.7
Step 13	PERF	0	321	43	88.2
		1	17	2374	99.3
		Overall Percentage			97.8
Step 14	PERF	0	321	43	88.2
		1	19	2372	99.2
		Overall Percentage			97.7
Step 15	PERF	0	325	39	89.3
		1	19	2372	99.2
		Overall Percentage			97.9
Step 16	PERF	0	320	44	87.9
		1	21	2370	99.1
		Overall Percentage			97.6

a. The cut value is .500

Table 2A: Classification Results

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	CAS by AGE	-.304	.014	461.280	1	.000	.738
	Constant	5.818	.243	574.470	1	.000	336.292
Step 2[b]	INC#5	4.270	.355	144.801	1	.000	71.538
	CAS by AGE	-.306	.016	373.521	1	.000	.737
	Constant	-.872	.499	3.053	1	.081	.418
Step 3[c]	CAS2	.161	.019	71.478	1	.000	1.175
	INC#5	3.913	.364	115.797	1	.000	50.038
	CAS by AGE	-.614	.043	202.345	1	.000	.541
	Constant	.556	.536	1.076	1	.300	1.743
	OCC(1)	-19.347	11791.245	.000	1	.999	.000
	OCC(2)	1.972	.523	14.213	1	.000	7.188
	OCC(3)	-.350	.469	.558	1	.455	.704
	OCC(4)	1.191	.385	9.593	1	.002	3.292
	OCC(5)	-.575	.318	3.271	1	.071	.563
	OCC(6)	2.040	.628	10.540	1	.001	7.690
	OCC(7)	-.736	.633	1.353	1	.245	.479
	OCC(8)	1.794	.428	17.538	1	.000	6.012
	OCC(9)	-.597	.543	1.209	1	.271	.550
	CAS2	.180	.022	68.277	1	.000	1.197
	INC#5	4.086	.390	109.668	1	.000	59.482
	CAS by AGE	-.636	.049	167.608	1	.000	.530
	Constant	-.267	.601	.197	1	.657	.766
Step 5[e]	OCC			53.860	9	.000	
	OCC(1)	-12.285	42479.936	.000	1	1.000	.000
	OCC(2)	-2.168	1.486	2.127	1	.145	.114
	OCC(3)	-2.370	1.232	3.700	1	.054	.094
	OCC(4)	5.110	1.280	15.947	1	.000	165.736
	OCC(5)	-6.306	1.361	21.460	1	.000	.002
	OCC(6)	-.991	2.131	.216	1	.642	.371
	OCC(7)	-37.082	10939.059	.000	1	.997	.000
	OCC(8)	1.048	1.097	.913	1	.339	2.852
	OCC(9)	-3.445	1.239	7.725	1	.005	.032
	CAS2	.148	.023	40.331	1	.000	1.160
	INC#5	2.954	.509	33.623	1	.000	19.181
	CAS by AGE	-.575	.052	122.966	1	.000	.563
	INC* OCC			48.081	9	.000	
	INCby OCC(1)	-4.308	31413.043	.000	1	1.000	.013
	INCby OCC(2)	2.425	1.009	5.778	1	.016	11.298
	INCby OCC(3)	1.028	.646	2.532	1	.112	2.795
	INCby OCC(4)	-1.310	.370	12.558	1	.000	.270
	INCby OCC(5)	2.424	.565	18.387	1	.000	11.293
	INCby OCC(6)	1.207	.864	1.952	1	.162	3.344
	INCby OCC(7)	17.900	5469.529	.000	1	.997	5.942E7
	INCby OCC(8)	.237	.538	.194	1	.660	1.267
	INCby OCC(9)	1.023	.448	5.204	1	.023	2.781
	Constant	1.280	.840	2.318	1	.128	3.595
Step 6[f]	OCC			68.747	9	.000	
	OCC(1)	-18.535	42524.516	.000	1	1.000	.000

	OCC(2)	-7.571	1.839	16.951	1	.000	.001
	OCC(3)	-6.705	1.785	14.115	1	.000	.001
	OCC(4)	11.373	1.948	34.090	1	.000	86956.892
	OCC(5)	-7.198	1.454	24.497	1	.000	.001
	OCC(6)	.473	3.040	.024	1	.876	1.605
	OCC(7)	-34.375	11099.460	.000	1	.998	.000
	OCC(8)	-3.184	1.840	2.995	1	.084	.041
	OCC(9)	-3.175	2.343	1.837	1	.175	.042
	CAS2	.132	.025	27.179	1	.000	1.141
	INC#5	42.224	6.404	43.467	1	.000	2.176E18
	LnINC	-28.420	4.449	40.807	1	.000	.000
	CAS by AGE	-.527	.054	93.516	1	.000	.590
	INC* OCC			67.514	9	.000	
	INCby OCC(1)	-.616	31488.478	.000	1	1.000	.540
	INCby OCC(2)	5.610	1.171	22.943	1	.000	273.139
	INCby OCC(3)	3.440	1.022	11.333	1	.001	31.196
	INCby OCC(4)	-3.498	.639	29.962	1	.000	.030
	INCby OCC(5)	3.022	.654	21.367	1	.000	20.527
	INCby OCC(6)	.809	1.242	.424	1	.515	2.245
	INCby OCC(7)	17.089	5549.730	.000	1	.998	2.641E7
	INCby OCC(8)	2.665	.993	7.199	1	.007	14.367
	INCby OCC(9)	.677	1.056	.412	1	.521	1.969
	Constant	-35.974	6.136	34.376	1	.000	.000
Step 7[g]	OCC			66.465	9	.000	
	OCC(1)	-17.162	42030.369	.000	1	1.000	.000
	OCC(2)	-7.579	1.879	16.263	1	.000	.001
	OCC(3)	-5.748	1.947	8.720	1	.003	.003
	OCC(4)	11.300	1.795	39.609	1	.000	80804.612
	OCC(5)	-6.354	1.520	17.470	1	.000	.002
	OCC(6)	.718	3.004	.057	1	.811	2.050
	OCC(7)	-35.146	11104.475	.000	1	.997	.000
	OCC(8)	-3.142	1.930	2.649	1	.104	.043
	OCC(9)	-3.976	2.318	2.942	1	.086	.019
	CAS2	.137	.026	28.482	1	.000	1.147
	INC#5	39.341	6.213	40.091	1	.000	1.218E17
	LnINC	-26.698	4.377	37.200	1	.000	.000
	LogEDU	6.583	1.591	17.113	1	.000	722.793
	CAS by AGE	-.541	.056	92.710	1	.000	.582
	INC* OCC			65.423	9	.000	
	INCby OCC(1)	-1.153	31310.527	.000	1	1.000	.316
	INCby OCC(2)	5.769	1.193	23.365	1	.000	320.139
	INCby OCC(3)	3.380	1.120	9.102	1	.003	29.372
	INCby OCC(4)	-3.454	.590	34.239	1	.000	.032
	INCby OCC(5)	2.691	.686	15.409	1	.000	14.749
	INCby OCC(6)	.647	1.239	.273	1	.601	1.911
	INCby OCC(7)	17.276	5552.237	.000	1	.998	3.183E7
	INCby OCC(8)	2.832	1.052	7.243	1	.007	16.979
	INCby OCC(9)	.925	1.037	.794	1	.373	2.521
	Constant	-36.473	5.939	37.715	1	.000	.000
Step 8[h]	OCC			62.822	9	.000	
	OCC(1)	-16.248	42313.613	.000	1	1.000	.000

	OCC(2)	-7.495	1.847	16.470	1	.000	.001
	OCC(3)	-6.114	1.990	9.436	1	.002	.002
	OCC(4)	11.243	1.901	34.972	1	.000	76324.958
	OCC(5)	-6.666	1.524	19.133	1	.000	.001
	OCC(6)	1.040	3.183	.107	1	.744	2.829
	OCC(7)	-36.103	10851.886	.000	1	.997	.000
	OCC(8)	-3.356	1.943	2.982	1	.084	.035
	OCC(9)	-4.056	2.303	3.101	1	.078	.017
	CAS2	.139	.026	27.763	1	.000	1.149
	INC#5	40.537	6.460	39.373	1	.000	4.029E17
	LnINC	-27.397	4.520	36.733	1	.000	.000
	LogEDU	6.946	1.634	18.078	1	.000	1039.027
	CAS by AGE	-.520	.057	82.007	1	.000	.594
	INC* OCC			62.292	9	.000	
	INCby OCC(1)	-1.492	31818.232	.000	1	1.000	.225
	INCby OCC(2)	5.610	1.164	23.228	1	.000	273.277
	INCby OCC(3)	3.626	1.156	9.843	1	.002	37.577
	INCby OCC(4)	-3.441	.625	30.290	1	.000	.032
	INCby OCC(5)	2.847	.685	17.273	1	.000	17.240
	INCby OCC(6)	.554	1.312	.179	1	.673	1.741
	INCby OCC(7)	17.650	5425.943	.000	1	.997	4.628E7
	INCby OCC(8)	2.982	1.071	7.754	1	.005	19.726
	INCby OCC(9)	.899	1.011	.792	1	.374	2.458
	AGE by MAR	-.288	.090	10.336	1	.001	.750
	Constant	-37.805	6.182	37.398	1	.000	.000
Step 9[i]	OCC			61.353	9	.000	
	OCC(1)	-15.841	42858.127	.000	1	1.000	.000
	OCC(2)	-7.837	1.877	17.425	1	.000	.000
	OCC(3)	-6.342	2.032	9.737	1	.002	.002
	OCC(4)	11.133	1.919	33.666	1	.000	68373.410
	OCC(5)	-7.002	1.565	20.011	1	.000	.001
	OCC(6)	1.135	3.299	.118	1	.731	3.110
	OCC(7)	-35.769	11026.748	.000	1	.997	.000
	OCC(8)	-3.537	1.971	3.221	1	.073	.029
	OCC(9)	-4.124	2.533	2.652	1	.103	.016
	CAS2	.118	.027	18.982	1	.000	1.125
	INC#5	43.663	6.455	45.760	1	.000	9.173E18
	LnINC	-28.745	4.495	40.887	1	.000	.000
	LogEDU	6.888	1.597	18.595	1	.000	980.095
	INCby LAM	-.203	.060	11.325	1	.001	.816
	CAS by AGE	-.522	.058	81.916	1	.000	.594
	INC* OCC			62.037	9	.000	
	INCby OCC(1)	-1.587	32331.713	.000	1	1.000	.205
	INCby OCC(2)	5.785	1.188	23.705	1	.000	325.309
	INCby OCC(3)	3.742	1.183	10.004	1	.002	42.189
	INCby OCC(4)	-3.412	.622	30.095	1	.000	.033
	INCby OCC(5)	2.987	.703	18.031	1	.000	19.819
	INCby OCC(6)	.554	1.362	.165	1	.684	1.740
	INCby OCC(7)	17.512	5513.374	.000	1	.997	4.029E7
	INCby OCC(8)	3.070	1.085	8.010	1	.005	21.550
	INCby OCC(9)	1.047	1.120	.874	1	.350	2.848

	AGE by MAR	-.309	.091	11.624	1	.001	.734
	Constant	-39.997	6.116	42.775	1	.000	.000
Step 10[j]	OCC			61.925	9	.000	
	OCC(1)	-15.247	42445.136	.000	1	1.000	.000
	OCC(2)	-7.713	1.938	15.844	1	.000	.000
	OCC(3)	-6.424	2.063	9.702	1	.002	.002
	OCC(4)	11.063	1.890	34.252	1	.000	63747.809
	OCC(5)	-7.675	1.674	21.032	1	.000	.000
	OCC(6)	.463	3.339	.019	1	.890	1.589
	OCC(7)	-36.594	10792.790	.000	1	.997	.000
	OCC(8)	-3.877	2.037	3.623	1	.057	.021
	OCC(9)	-4.667	2.161	4.663	1	.031	.009
	CAS2	.213	.039	29.109	1	.000	1.237
	INC#5	43.784	6.373	47.197	1	.000	1.035E19
	AGE#5	3.493	1.013	11.883	1	.001	32.883
	LnINC	-28.543	4.468	40.818	1	.000	.000
	LogEDU	6.633	1.632	16.513	1	.000	759.715
	INCby LAM	-.255	.061	17.426	1	.000	.775
	CAS by AGE	-.732	.087	70.036	1	.000	.481
	INC* OCC			62.145	9	.000	
	INCby OCC(1)	-1.934	32001.091	.000	1	1.000	.145
	INCby OCC(2)	5.622	1.228	20.964	1	.000	276.571
	INCby OCC(3)	3.632	1.205	9.084	1	.003	37.774
	INCby OCC(4)	-3.389	.609	30.982	1	.000	.034
	INCby OCC(5)	3.245	.747	18.859	1	.000	25.658
	INCby OCC(6)	.832	1.384	.362	1	.547	2.299
	INCby OCC(7)	17.810	5396.395	.000	1	.997	5.429E7
	INCby OCC(8)	3.171	1.124	7.956	1	.005	23.833
	INCby OCC(9)	1.128	.938	1.447	1	.229	3.091
	AGE by MAR	-.349	.096	13.257	1	.000	.706
	Constant	-45.363	6.268	52.380	1	.000	.000
Step 11[k]	OCC			60.474	9	.000	
	OCC(1)	-16.813	42628.807	.000	1	1.000	.000
	OCC(2)	-6.912	1.881	13.507	1	.000	.001
	OCC(3)	-6.907	2.059	11.257	1	.001	.001
	OCC(4)	12.151	2.055	34.960	1	.000	189215.650
	OCC(5)	-7.429	1.785	17.322	1	.000	.001
	OCC(6)	1.698	3.430	.245	1	.620	5.465
	OCC(7)	-35.367	10869.160	.000	1	.997	.000
	OCC(8)	-3.661	1.996	3.362	1	.067	.026
	OCC(9)	-3.376	2.081	2.632	1	.105	.034
	CAS2	.184	.039	22.387	1	.000	1.202
	INC#5	64.031	8.843	52.428	1	.000	6.434E27
	AGE#5	8.296	1.506	30.344	1	.000	4005.922
	LnINC	-37.410	5.294	49.935	1	.000	.000
	LogEDU	6.538	1.653	15.641	1	.000	690.638
	INCby LAM	-.165	.070	5.512	1	.019	.848
	CAS by AGE	-.657	.085	59.526	1	.000	.519
	INC* OCC			61.484	9	.000	
	INCby OCC(1)	-1.098	32106.070	.000	1	1.000	.334
	INCby OCC(2)	5.291	1.196	19.562	1	.000	198.487

	INCby OCC(3)	3.947	1.214	10.567	1	.001	51.792
	INCby OCC(4)	-3.821	.672	32.305	1	.000	.022
	INCby OCC(5)	3.058	.812	14.183	1	.000	21.284
	INCby OCC(6)	.296	1.442	.042	1	.837	1.344
	INCby OCC(7)	17.257	5434.580	.000	1	.997	3.122E7
	INCby OCC(8)	3.117	1.109	7.893	1	.005	22.577
	INCby OCC(9)	.584	.892	.429	1	.513	1.793
	INCby AGE	-.777	.175	19.660	1	.000	.460
	AGE by MAR	-.358	.098	13.432	1	.000	.699
	Constant	-72.218	9.840	53.863	1	.000	.000
Step 12[l]	OCC			39.950	9	.000	
	OCC(1)	-16.101	42622.505	.000	1	1.000	.000
	OCC(2)	-5.617	1.851	9.213	1	.002	.004
	OCC(3)	-6.148	1.987	9.574	1	.002	.002
	OCC(4)	19.184	4.978	14.849	1	.000	2.145E8
	OCC(5)	-7.837	1.777	19.443	1	.000	.000
	OCC(6)	.153	3.691	.002	1	.967	1.165
	OCC(7)	-36.649	10939.415	.000	1	.997	.000
	OCC(8)	-2.730	1.872	2.128	1	.145	.065
	OCC(9)	-2.562	2.284	1.258	1	.262	.077
	CAS2	.199	.040	25.128	1	.000	1.221
	INC2	1.804	.685	6.924	1	.009	6.071
	INC#5	-21.152	31.888	.440	1	.507	.000
	AGE#5	11.011	1.883	34.181	1	.000	60516.943
	LnINC	8.224	16.775	.240	1	.624	3730.998
	LogEDU	6.999	1.686	17.225	1	.000	1095.111
	INCby LAM	-.221	.072	9.454	1	.002	.801
	CAS by AGE	-.700	.089	62.236	1	.000	.497
	INC* OCC			44.094	9	.000	
	INCby OCC(1)	-1.573	32102.598	.000	1	1.000	.207
	INCby OCC(2)	4.601	1.182	15.151	1	.000	99.633
	INCby OCC(3)	3.489	1.161	9.029	1	.003	32.749
	INCby OCC(4)	-6.073	1.601	14.386	1	.000	.002
	INCby OCC(5)	3.246	.795	16.692	1	.000	25.699
	INCby OCC(6)	.982	1.572	.391	1	.532	2.671
	INCby OCC(7)	17.801	5469.707	.000	1	.997	5.380E7
	INCby OCC(8)	2.595	1.031	6.336	1	.012	13.402
	INCby OCC(9)	.301	1.012	.088	1	.766	1.351
	INCby AGE	-1.108	.216	26.203	1	.000	.330
	AGE by MAR	-.340	.099	11.696	1	.001	.712
	Constant	7.129	30.642	.054	1	.816	1247.425
Step 13[l]	OCC			47.080	9	.000	
	OCC(1)	-16.321	42528.537	.000	1	1.000	.000
	OCC(2)	-5.865	1.800	10.621	1	.001	.003
	OCC(3)	-6.351	1.967	10.429	1	.001	.002
	OCC(4)	17.806	3.563	24.976	1	.000	5.410E7
	OCC(5)	-7.797	1.785	19.082	1	.000	.000
	OCC(6)	.436	3.617	.014	1	.904	1.546
	OCC(7)	-36.271	10615.973	.000	1	.997	.000
	OCC(8)	-2.927	1.863	2.469	1	.116	.054
	OCC(9)	-2.667	2.257	1.396	1	.237	.069

	CAS2	.196	.039	25.146	1	.000	1.216
	INC2	1.496	.218	47.213	1	.000	4.466
	INC#5	-5.714	2.443	5.471	1	.019	.003
	AGE#5	10.757	1.786	36.289	1	.000	46967.622
	LogEDU	6.928	1.677	17.071	1	.000	1020.138
	INCby LAM	-.215	.071	9.206	1	.002	.807
	CAS by AGE	-.692	.087	63.628	1	.000	.501
	INC* OCC			49.850	9	.000	
	INCby OCC(1)	-1.436	32047.454	.000	1	1.000	.238
	INCby OCC(2)	4.738	1.160	16.688	1	.000	114.180
	INCby OCC(3)	3.607	1.153	9.791	1	.002	36.846
	INCby OCC(4)	-5.639	1.149	24.094	1	.000	.004
	INCby OCC(5)	3.227	.801	16.239	1	.000	25.200
	INCby OCC(6)	.862	1.538	.314	1	.575	2.368
	INCby OCC(7)	17.635	5307.987	.000	1	.997	4.558E7
	INCby OCC(8)	2.706	1.029	6.921	1	.009	14.968
	INCby OCC(9)	.339	.998	.115	1	.734	1.403
	INCby AGE	-1.083	.208	27.090	1	.000	.338
	AGE by MAR	-.342	.099	11.860	1	.001	.710
	Constant	-7.552	4.155	3.304	1	.069	.001
Step 14[m]	OCC			50.153	9	.000	
	OCC(1)	-15.761	41954.834	.000	1	1.000	.000
	OCC(2)	-5.949	1.879	10.020	1	.002	.003
	OCC(3)	-5.866	2.239	6.866	1	.009	.003
	OCC(4)	20.422	3.805	28.801	1	.000	7.399E8
	OCC(5)	-8.022	1.816	19.517	1	.000	.000
	OCC(6)	-.177	3.411	.003	1	.959	.838
	OCC(7)	-36.217	10360.491	.000	1	.997	.000
	OCC(8)	-2.968	1.947	2.323	1	.127	.051
	OCC(9)	-2.567	2.509	1.047	1	.306	.077
	CAS2	.196	.040	24.049	1	.000	1.216
	INC2	2.144	.340	39.654	1	.000	8.535
	INC#5	-1.602	2.676	.358	1	.549	.201
	AGE#5	11.379	1.860	37.436	1	.000	87508.996
	LogEDU	22.536	5.789	15.156	1	.000	6.126E9
	INCby LAM	-.225	.074	9.375	1	.002	.798
	CAS by AGE	-.688	.089	60.428	1	.000	.503
	Education by Ind_Income	-1.087	.367	8.762	1	.003	.337
	INC* OCC			53.681	9	.000	
	INCby OCC(1)	-1.763	31858.047	.000	1	1.000	.172
	INCby OCC(2)	4.774	1.181	16.336	1	.000	118.384
	INCby OCC(3)	3.481	1.316	6.998	1	.008	32.508
	INCby OCC(4)	-6.525	1.223	28.485	1	.000	.001
	INCby OCC(5)	3.342	.809	17.056	1	.000	28.262
	INCby OCC(6)	1.172	1.436	.667	1	.414	3.229
	INCby OCC(7)	17.711	5180.246	.000	1	.997	4.918E7
	INCby OCC(8)	2.776	1.077	6.643	1	.010	16.056
	INCby OCC(9)	.308	1.121	.075	1	.784	1.360
	INCby AGE	-1.162	.221	27.596	1	.000	.313
	AGE by MAR	-.346	.101	11.828	1	.001	.707
	Constant	-17.413	5.235	11.065	1	.001	.000

Step 15[m]	OCC			59.607	9	.000	
	OCC(1)	-15.209	41835.470	.000	1	1.000	.000
	OCC(2)	-5.388	1.617	11.100	1	.001	.005
	OCC(3)	-5.340	2.059	6.730	1	.009	.005
	OCC(4)	19.869	3.616	30.199	1	.000	4.255E8
	OCC(5)	-7.742	1.739	19.827	1	.000	.000
	OCC(6)	-.263	3.327	.006	1	.937	.769
	OCC(7)	-36.109	10336.659	.000	1	.997	.000
	OCC(8)	-2.413	1.689	2.042	1	.153	.090
	OCC(9)	-2.238	2.377	.886	1	.347	.107
	CAS2	.196	.040	24.234	1	.000	1.217
	INC2	2.137	.342	39.127	1	.000	8.474
	AGE#5	11.573	1.820	40.448	1	.000	106213.819
	LogEDU	23.846	5.395	19.539	1	.000	2.271E10
	INCby LAM	-.220	.074	8.838	1	.003	.803
	CAS by AGE	-.687	.088	60.618	1	.000	.503
	Education by Ind_Income	-1.177	.339	12.095	1	.001	.308
	INC* OCC			66.109	9	.000	
	INCby OCC(1)	-2.057	31799.598	.000	1	1.000	.128
	INCby OCC(2)	4.487	1.067	17.681	1	.000	88.812
	INCby OCC(3)	3.219	1.240	6.735	1	.009	25.011
	INCby OCC(4)	-6.360	1.166	29.762	1	.000	.002
	INCby OCC(5)	3.210	.770	17.394	1	.000	24.777
	INCby OCC(6)	1.199	1.397	.737	1	.391	3.316
	INCby OCC(7)	17.650	5168.329	.000	1	.997	4.629E7
	INCby OCC(8)	2.496	.956	6.820	1	.009	12.134
	INCby OCC(9)	.164	1.054	.024	1	.877	1.178
	INCby AGE	-1.186	.216	30.122	1	.000	.305
	AGE by MAR	-.349	.100	12.115	1	.001	.705
	Constant	-19.926	3.184	39.157	1	.000	.000
Step 16[n]	OCC			58.518	9	.000	
	OCC(1)	-14.295	41397.323	.000	1	1.000	.000
	OCC(2)	-5.925	1.643	13.002	1	.000	.003
	OCC(3)	-5.718	2.026	7.961	1	.005	.003
	OCC(4)	18.688	3.582	27.212	1	.000	1.306E8
	OCC(5)	-7.818	1.788	19.126	1	.000	.000
	OCC(6)	-.059	3.322	.000	1	.986	.943
	OCC(7)	-36.122	10463.263	.000	1	.997	.000
	OCC(8)	-2.678	1.723	2.416	1	.120	.069
	OCC(9)	-2.391	2.250	1.129	1	.288	.092
	CAS2	.194	.041	22.678	1	.000	1.214
	INC2	2.087	.342	37.151	1	.000	8.058
	AGE#5	12.370	1.907	42.090	1	.000	235509.590
	LogEDU	22.895	5.402	17.963	1	.000	8.770E9
	INCby LAM	-.200	.076	6.854	1	.009	.819
	CAS by AGE	-.679	.090	57.146	1	.000	.507
	Education by Ind_Income	-1.190	.339	12.343	1	.000	.304
	INC* OCC			64.163	9	.000	
	INCby OCC(1)	-2.537	31802.244	.000	1	1.000	.079
	INCby OCC(2)	4.711	1.065	19.570	1	.000	111.168

INCby OCC(3)	3.465	1.239	7.815	1	.005	31.962
INCby OCC(4)	-5.968	1.158	26.537	1	.000	.003
INCby OCC(5)	3.184	.794	16.086	1	.000	24.136
INCby OCC(6)	1.173	1.381	.721	1	.396	3.233
INCby OCC(7)	17.617	5231.631	.000	1	.997	4.477E7
INCby OCC(8)	2.627	.982	7.158	1	.007	13.826
INCby OCC(9)	.225	.984	.052	1	.819	1.252
INCby AGE	-1.177	.222	28.183	1	.000	.308
AGE by MAR	-1.089	.341	10.193	1	.001	.337
Education by MAR	.754	.334	5.098	1	.024	2.126
Constant	-20.836	3.289	40.124	1	.000	.000

a. Variable(s) entered on step 1: CAS * AGE .
b. Variable(s) entered on step 2: INC#5.
c. Variable(s) entered on step 3: CAS2.
d. Variable(s) entered on step 4: OCC.
e. Variable(s) entered on step 5: INC* OCC .
f. Variable(s) entered on step 6: LnINC.
g. Variable(s) entered on step 7: LogEDU.
h. Variable(s) entered on step 8: AGE * married_YN .
i. Variable(s) entered on step 9: INC* LAM .
j. Variable(s) entered on step 10: AGE#5.
k. Variable(s) entered on step 11: INC* AGE .
l. Variable(s) entered on step 12: INC2.
m. Variable(s) entered on step 14: Education * INC.
n. Variable(s) entered on step 16: Education * married_YN .

Table 3A. The estimation results of Logit analysis

Education * Ind_Income * performingnonperforming Crosstabulation

performingnonperforming				Ind_Income 1	2	3	4	5	6	Total
0	Education	1	Count	6	0	0	0			6
			% within Education	100.0%	.0%	.0%	.0%			100.0%
		2	Count	43	16	0	0			59
			% within Education	72.9%	27.1%	.0%	.0%			100.0%
		3	Count	37	138	3	0			178
			% within Education	20.8%	77.5%	1.7%	.0%			100.0%
		4	Count	12	73	21	13			119
			% within Education	10.1%	61.3%	17.6%	10.9%			100.0%
		5	Count	0	0	2	0			2
			% within Education	.0%	.0%	100.0%	.0%			100.0%
		Total	Count	98	227	26	13			364
			% within Education	26.9%	62.4%	7.1%	3.6%			100.0%
1	Education	1	Count	0	1	0	0	1	0	2
			% within Education	.0%	50.0%	.0%	.0%	50.0%	.0%	100.0%
		2	Count	2	22	15	9	5	0	53
			% within Education	3.8%	41.5%	28.3%	17.0%	9.4%	.0%	100.0%
		3	Count	74	358	127	102	217	0	878
			% within Education	8.4%	40.8%	14.5%	11.6%	24.7%	.0%	100.0%
		4	Count	41	136	562	284	309	1	1333
			% within Education	3.1%	10.2%	42.2%	21.3%	23.2%	.1%	100.0%
		5	Count	21	13	14	47	30	0	125
			% within Education	16.8%	10.4%	11.2%	37.6%	24.0%	.0%	100.0%
		Total	Count	138	530	718	442	562	1	2391
			% within Education	5.8%	22.2%	30.0%	18.5%	23.5%	.0%	100.0%

Table 4A. Relation between Education, Income, and Defaulting

Education * married_YN * performingnonperforming Crosstabulation						
				married_YN		
performingnonperforming				1	2	Total
0	Education	1	Count	2	4	6
			% within Education	33.3%	66.7%	100.0%
		2	Count	34	25	59
			% within Education	57.6%	42.4%	100.0%
		3	Count	117	61	178
			% within Education	65.7%	34.3%	100.0%
		4	Count	73	46	119
			% within Education	61.3%	38.7%	100.0%
		5	Count	0	2	2
			% within Education	.0%	100.0%	100.0%
		Total	Count	226	138	364
			% within Education	62.1%	37.9%	100.0%
1	Education	1	Count	0	2	2
			% within Education	.0%	100.0%	100.0%
		2	Count	3	50	53
			% within Education	5.7%	94.3%	100.0%
		3	Count	184	694	878
			% within Education	21.0%	79.0%	100.0%
		4	Count	446	887	1333
			% within Education	33.5%	66.5%	100.0%
		5	Count	15	110	125
			% within Education	12.0%	88.0%	100.0%
		Total	Count	648	1743	2391
			% within Education	27.1%	72.9%	100.0%

Table 5A. Relation between Education, Marital status interaction, and Defaulting

Ind_Income * LoanAmount * performingnonperforming Crosstabulation												
				LoanAmount								
performingnonperforming				1	2	3	4	5	6	7	8	Total
0	Ind_Income	1	Count	33	45	19	1	0	0			98
			% within LoanAmount	25.8%	29.6%	29.2%	7.7%	.0%	.0%			26.9%
		2	Count	89	92	36	8	2	0			227
			% within LoanAmount	69.5%	60.5%	55.4%	61.5%	50.0%	.0%			62.4%
		3	Count	6	10	8	2	0	0			26
			% within LoanAmount	4.7%	6.6%	12.3%	15.4%	.0%	.0%			7.1%
		4	Count	0	5	2	2	2	2			13
			% within LoanAmount	.0%	3.3%	3.1%	15.4%	50.0%	100.0%			3.6%
		Total	Count	128	152	65	13	4	2			364
			% within LoanAmount	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%			100.0%
1	Ind_Income	1	Count	23	48	65	1	1	0	0	0	138
			% within LoanAmount	6.9%	9.3%	5.6%	.3%	1.4%	.0%	.0%	.0%	5.8%
		2	Count	73	173	280	4	0	0	0	0	530
			% within LoanAmount	21.9%	33.4%	24.2%	1.4%	.0%	.0%	.0%	.0%	22.2%
		3	Count	115	145	353	92	9	2	1	1	718
			% within LoanAmount	34.5%	28.0%	30.5%	31.5%	12.3%	15.4%	33.3%	50.0%	30.0%
		4	Count	44	74	244	73	7	0	0	0	442
			% within LoanAmount	13.2%	14.3%	21.1%	25.0%	9.6%	.0%	.0%	.0%	18.5%
		5	Count	78	78	214	122	56	11	2	1	562
			% within LoanAmount	23.4%	15.1%	18.5%	41.8%	76.7%	84.6%	66.7%	50.0%	23.5%
		6	Count	0	0	1	0	0	0	0	0	1
			% within LoanAmount	.0%	.0%	.1%	.0%	.0%	.0%	.0%	.0%	.0%
		Total	Count	333	518	1157	292	73	13	3	2	2391
			% within LoanAmount	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Table 6A. The relation between Income, Loan amount interaction, and defaulting

performingnonperforming * LoanAmount Crosstabulation			LoanAmount								
			1	2	3	4	5	6	7	8	Total
performingnonperforming	0	Count	128	152	65	13	4	2	0	0	364
		% within performingnonperforming	35.2%	41.8%	17.9%	3.6%	1.1%	.5%	.0%	.0%	100.0%
	1	Count	333	518	1157	292	73	13	3	2	2391
		% within performingnonperforming	13.9%	21.7%	48.4%	12.2%	3.1%	.5%	.1%	.1%	100.0%
	Total	Count	461	670	1222	305	77	15	3	2	2755
		% within performingnonperforming	16.7%	24.3%	44.4%	11.1%	2.8%	.5%	.1%	.1%	100.0%

Table 7A. The relation between Loan amount and Defaulting

PERF * CAS Crosstabulation			CAS						
			1	2	3	4	5	6	Total
PERF	0	Count	20	21	8	29	63	223	364
		% within PERF	5.5%	5.8%	2.2%	8.0%	17.3%	61.3%	100.0%
	1	Count	1752	202	57	18	223	139	2391
		% within PERF	73.3%	8.4%	2.4%	.8%	9.3%	5.8%	100.0%
	Total	Count	1772	223	65	47	286	362	2755
		% within PERF	64.3%	8.1%	2.4%	1.7%	10.4%	13.1%	100.0%

Table 8A. The relation between Customer-Bank-Age and Defaulting

performingnonperforming				NewAge 1	2	3	4	5	6	Total
0	LoanAmount	1	Count	0	2	26	68	29	8	133
			% within NewAge	.0%	22.2%	44.1%	40.2%	29.6%	28.6%	36.5%
		2	Count	0	3	18	74	46	12	153
			% within NewAge	.0%	33.3%	30.5%	43.8%	46.9%	42.9%	42.0%
		3	Count	1	4	12	21	20	5	63
			% within NewAge	100.0%	44.4%	20.3%	12.4%	20.4%	17.9%	17.3%
		4	Count	0	0	2	6	2	3	13
			% within NewAge	.0%	.0%	3.4%	3.6%	2.0%	10.7%	3.6%
		5	Count	0	0	1	0	1	0	2
			% within NewAge	.0%	.0%	1.7%	.0%	1.0%	.0%	.5%
		Total	Count	1	9	59	169	98	28	364
			% within NewAge	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
1	LoanAmount	1	Count		22	284	3	0	0	309
			% within NewAge		7.6%	18.6%	.8%	.0%	.0%	12.9%
		2	Count		66	326	39	28	7	466
			% within NewAge		22.7%	21.4%	10.8%	15.6%	20.0%	19.5%
		3	Count		158	667	207	95	13	1140
			% within NewAge		54.3%	43.7%	57.3%	53.1%	37.1%	47.7%
		4	Count		40	203	71	37	9	360
			% within NewAge		13.7%	13.3%	19.7%	20.7%	25.7%	15.1%
		5	Count		5	37	30	16	5	93
			% within NewAge		1.7%	2.4%	8.3%	8.9%	14.3%	3.9%
		6	Count		0	4	10	2	0	16
			% within NewAge		.0%	.3%	2.8%	1.1%	.0%	.7%
		7	Count		0	3	0	1	1	5
			% within NewAge		.0%	.2%	.0%	.6%	2.9%	.2%
		8	Count		0	1	1	0	0	2
			% within NewAge		.0%	.1%	.3%	.0%	.0%	.1%
		Total	Count		291	1525	361	179	35	2391
			% within NewAge		100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Table 9A: The relation between Age and Loan Amount

Appendix B

Code	Job Cluster (OCC)
9	Business, Management and administration
8	Law and Engineering
7	Employee
6	Education, Training and Media
5	Health services
4	Marketing, Sales and Finance, customer services
3	Government jobs and Military
2	Vocational
1	Non-workers,
0	Un-Known

1B. Occupation

Code	Loan Amount (LAM)
1	less than or equal 5000
2	5001-10000
3	10001-30000
4	30001-50000
5	50001-100000
6	100001-200000
7	200001-350000
8	greater than 350000

2B. Loan Amount

Code	Customer Bank Age (CAS)
1	less than or equal One year
2	1 – 2
3	2 – 3
4	3 – 4
5	4 – 5
6	More than Five Years

3B. Customer-Bank-Age

Code	Gender (GND)
0	Female
1	Male

4B. Gender

Code	Age (AGE)
1	less than 20
2	20 – 29
3	30 – 39
4	40 – 49
5	50 – 59
6	>=60

5B. Age

Code	Education (EDU)
1	Uneducated
2	High School
3	Diploma
4	Bachelor
5	Postgraduate

6B. Education

Code	Income (INC)
1	Less than or equal JOD400
2	401 – 600
3	601 – 800
4	801- 1000
5	Greater than 1000

7B. Income

Code	Married (MAR)
1	Not Married or divorced
2	Married

8B. Marital Status

Code	Perform (Perf)
0	Non Perform
1	Perform

9B. Perform

Appendix C

1) Model one using multiple logistic regression:

$$\sum_{k=1}^{n} B_k X_k = -.365LAM - 0.975CAS - 1.419AGE + \alpha * OCC + 1.071EDU + 1.261INC - .805MAR$$

Appendix D

Impact of Knowledge Management

Research Questionnaire

Abdallah Al-Shawebkeh

KM research Group

School of Computing and Mathematical Sciences

Dear Employee

First of all, I would like to thank you for the time that you will spend in completing this important questionnaire. The questionnaire is designed out to measure the performance of implementing Knowledge Management (KM) in Jordanian Banks. Please answer all the (64) questions carefully. Your response will make a valuable contribution in understanding the extent to which KM activities have already been implemented in Jordanian Banks, and pointing out problems and areas where improvements are needed.

The data collected in this questionnaire is completely confidential and individual identifiers are not required to be revealed.

Thank you

Part One

Section One: Please circle the answer

1) What is your gender? Male , Female

2) What is your age? [ ] years

3) What is your highest qualification?

High School, Diploma , University Degree , Master , PhD

4) How long have you been in this bank? [ ] years

Section Two: Please circle the answer that most exactly applies to you

5) The bank offered me a specialized training before starting my current job?
a) Yes b) No

(If the answer is NO, please go to question 8)

6) The specialized training offered before starting my current job was very useful?
a) Strongly agree b) Agree c) Neutral d) Disagree e) Strongly disagree

7) The type of training offered was related to my current job.
a) Strongly agree b) Agree c) Neutral d) Disagree e) Strongly disagree

8) The bank offered me an IT- training for my job requirements?

a) Yes b) No

(If the answer is NO go to question 10)

9) I am satisfied with the IT training given for my job requirements

a) Strongly agree b) Agree c) Neutral d) Disagree e) Strongly disagree

10) Whenever new software is installed I am given training

a) Yes b) No

(If the answer is NO go to question 12)

11) I am satisfied with the IT training given to me when new software is installed

a) Strongly agree b) Agree c) Neutral d) Disagree e) Strongly disagree

Section Three: Please, for each of the following questions, tick boxes indicating the level of your response

Q (No)	Question	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
12	I am satisfied with my training plan as provided by the Human Resource Department					
13	I am satisfied with the reference documents available for me in relation to my work					
14	My management supports and encourages me to attend seminars and conferences that relate to my work					
15	The bank provides me well structured training to recognize Knowledge that is valuable to my work.					
16	The presentations given to me by the bank on global case studies are very relevant to my work					
17	I am satisfied with my training on how to extract knowledge from my bank's knowledge base					
18	I am satisfied with teamwork inside the bank					
19	I am satisfied with the attendance in group's presentations					
20	I am satisfied with the social activities in the bank					
21	I am satisfied with the amount of hours I spend with external experts per month					
22	I am satisfied with regular meetings between departments to discuss market trends and developments					

23	I am satisfied with the amount of hours we spend in work meetings					
24	I am satisfied with the number of shared reports produced					

Q(NO)	Question	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
25	I am satisfied with the number of patents that we published					
26	Managers encourage sharing knowledge between employees					
27	I am satisfied with the support from top level management					
28	IT managers have built strong interpersonal relationship with employees					
29	Top management gives appropriates rewards to motivate employees to share knowledge					
30	I am satisfied with the workspace that I have been allocated in the bank					
31	Top-management supports research and development well					
32	I am satisfied with knowledge availability					
33	I use the knowledge-base frequently					
34	I call the help-desk frequently					
35	I know exactly who to ask when I need information for a specific task					
36	I use knowledge about customers to do my tasks					

37	By following routine business processes I gain knowledge about the business.					
38	Peer evaluation/review and collaboration is an established practice in the bank					

Q(NO)	Question	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
39	The bank provides enough time for work breaks					
40	Standard processes for retrieving knowledge are defined clearly					
41	Organizational Knowledge is updated regularly					
42	Case notes on successful and unsuccessful services are documented and archived					
43	I know who to ask when I need specific knowledge					
44	Requirements for sharing tacit knowledge have been provided					
45	Professional knowledge such as customer information, is managed systematically					
46	Processes and workflows are well documented					
47	Formats to document best practice and case studies are clearly defined					
48	Sources of explicit knowledge are well defined					
49	Internal best-practices are well recorded					
50	The bank's database is user friendly					

51	Help-instructions in the bank's database are sufficient					
52	I use the knowledge base frequently					

Q(NO)	Question	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
53	I download knowledge from the knowledge-based frequently					
54	Data access is fast					
55	The Information Systems process the data very well.					
56	I can use the Information Systems very easy.					
57	I can rely on the Information Systems in the bank					
58	We use frequently teleconferencing in the bank					
59	The bank makes good use of groupware (such as lotus note) to share documents on services and notes					
60	The bank uses Decision Support Systems extensively.					
61	I am satisfied with the availability of knowledge in Information Systems.					
62	The bank uses the expert systems extensively.					
63	The bank uses the internet in executing transactions					

Part Two

Dear Manager,

In the Table below, could you please fill in your perceived relative importance of each factor and criterion according to their importance in managing knowledge in CR department? .This should be done in relative to the below instructions.

Expected consequences of each option are assigned a numerical score on strength of preference scale for each option for each factor or criterion. The scores for all elements being compared under each factor or criterion must add up to100. More favoured options score higher on the scale and less favoured score lower. Scales extending from 0 to 100 are often, where 0 represents a real or hypothetical least favoured option and 100 is associated with a real or hypothetical most favoured option.

Factor	Perceived Importance	Criterion	Perceived Importance
1. F1		C 1.1	
		C 1.2	
		C 1.3	
		C 1.4	
		Total	100%
2.F2		C 2.1	
		C 2.2	
		Total	100%
3. F3		C 3.1	
		C 3.2	
Total	100%	Total	100%

Appendix E

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	F1[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: Badloan

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.300[a]	.090	.086	.02807

a. Predictors: (Constant), F1

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.019	1	.019	23.769	.000[a]
	Residual	.189	240	.001		
	Total	.208	241			

a. Predictors: (Constant), F1

b. Dependent Variable: Badloan

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	
		B	Std. Error	Beta			Tolerance
1	(Constant)	.114	.011		10.049	.000	
	F1	-.016	.003	-.300	-4.875	.000	1.000

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions	
				(Constant)	F1
1	1	1.987	1.000	.01	.01
	2	.013	12.443	.99	.99

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions (Constant)	F1
1	1	1.987	1.000	.01	.01
	2	.013	12.443	.99	.99

a. Dependent Variable: Badloan

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	F2[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: Badloan

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.311[a]	.097	.093	.02797

a. Predictors: (Constant), F2

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.020	1	.020	25.640	.000[a]
	Residual	.188	240	.001		
	Total	.208	241			

a. Predictors: (Constant), F2

b. Dependent Variable: Badloan

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.115	.011		10.274	.000
	F2	-.016	.003	-.311	-5.064	.000

a. Dependent Variable: Badloan

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	F3[a]	.	Enter

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.300[a]	.090	.086	.02807

a. Predictors: (Constant), F3

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.019	1	.019	23.804	.000[a]
	Residual	.189	240	.001		
	Total	.208	241			

a. Predictors: (Constant), F3

b. Dependent Variable: Badloan

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.110	.011		10.362	.000
	F3	-.015	.003	-.300	-4.879	.000

a. Dependent Variable: Badloan

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	KM[a]	.	Enter

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.337[a]	.114	.110	.02771

a. Predictors: (Constant), KM

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.024	1	.024	30.805	.000[a]
	Residual	.184	240	.001		
	Total	.208	241			

a. Predictors: (Constant), KM

b. Dependent Variable: Badloan

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.126	.012		10.366	.000
	KM	-.019	.003	-.337	-5.550	.000

a. Dependent Variable: Badloan

Collinearity Diagnostics[a]

Model Dimension	Eigenvalue	Condition Index	Variance Proportions (Constant)	F1	F2	F3
1 1	3.968	1.000	.00	.00	.00	.00
2	.016	15.625	.99	.03	.04	.09
3	.010	19.862	.01	.26	.08	.85
4	.005	27.025	.00	.70	.88	.05

a. Dependent Variable: Badloan

Dec-07						
Bank Name	Net Loans	Bad Debt Provision	Suspend Interest	Gross Loans	Bad Loans	Percentage of bad loans to gross loans
Jordan Bank	600.2	19.6	4.7	639.2	28.6	0.044743429
Ntional Bank	483.8	58	21.8	545.3	83.8	0.153676875
HSBC	205.6	13.4	5.3	252.7	16.9	0.066877721
Rafdien				12.6	1.9	0.150793651
Arab Bank	1703	60.5	69.1	1832.6	184.1	0.100458365
Arab Egyption Bank	159.4	12.2	5.3	176.9	70	0.395703787
Cairo Amman Bank	413.2	21.3	6.1	440.6	27.6	0.062641852
Standard Charter	181.2	7.9	7.7	196.8	15.5	0.078760163
City Bank	28.1	0.1	0	28.2	0	0
Jordan Kwuit Bank	804.7	3.4	0.5	808.6	3.6	0.00445214
Jordan Gulf Bank	268.7	13.9	19.4	302	41.8	0.138410596
Arab Investment Bank	168.5	6.9	3.5	178.9	11.2	0.062604807
Jordan Islamic Bank	604.2	6.8	33.7	644.7	16.9	0.026213743
Housing Bank	1312.7	27.7	12.6	1353	48	0.035476718
JIFB	221.9	35.7	7.3	264.9	59	0.222725557
ABC	209.6	5.9	7.5	223	11.1	0.049775785
Union Bank	427	4.9	3	434.9	12	0.02759255
Societe General	96.3	5.1	15.6	117	21.3	0.182051282
Export bank	472.2	6.6	2.2	481	13.1	0.027234927
Arabic Islamic Bank	274.1	1.1	20	295.2	1.9	0.006436314
Kwuit National	70.8	0.3	0	71.1	0.4	0.005625879
Lebnon	71.7	0.5	0	72.2	1.2	0.016620499
Audi	125.3	0.6	0	125.9	0.4	0.003177125

Percentage of Bad loans to gross loans (December-2007)

Appendix F

CSFs Knowledge		Employee Training	Trustworthy teamwork & employee involvement	Top-management leadership support	Employee aware about knowledge inside organisation
Internalisation Socialisation	Tacit	1) Organisation offers education to help employee recognise what knowledge is value. (Chong and Choi,2005) 2) Employees engaged in double loop learning (Chong and Choi, 2005) 3) Formal training on IT for users. (Chong and Choi,2005) 4) Organisation offers learning mechanisms (post mortem phase, after action review, etc.). (Desouza, 2005) 5) Number of subscriptions to journals (Lee, 2000) 6) Education to help employees recognise what knowledge is valuable provided through training tools like corporate university (Chong and Choi, 2005) 7) Encouraging employees to attend training seminars and conferences. (Galvin,1996,DarrochYEAR)	1) Number of users participating in knowledge sharing activities (Orr and Persson, 2003) 2) Number of hits on document repository. (Lee, 2000) 3) Attendance at group presentation. (Lee, 2000) 4) Number of links per respondent. (Lee, 2000) 5) Direct communication links. (Moor et al., 2002) 6) Number of links per respondent (Knowledge sharing density). (Lee, 2000) 7) % regulated socialisation. (Moor et al., 2002) 8) Teaming is large part of organisation activity today (Glendon and Kundtz, 2000) 9) Frequency of advice sharing (Orr and Persson, 2003) 10) We share knowledge necessary for tasks? (Chung, 2002) 11) We improve tasks efficiency by sharing knowledge (Chung, 2002) 12) Appropriate networks established (Moor et al., 2002) 13) Number of social interactions (is directly correlated with the degree of trust and commitment). (Lee, 2000) 14) Number of hours spends with external experts per month. 15) Regular meeting between departments to discuss market trends and developments. (Gabriel,1999; Darr, YEAR)	1) Leaders encouraging employees to share knowledge. (Kim, 2004) 2) Support from top-level managers.(Kim, 2004) 3) Organisation has identified networks or groups of practitioners with sufficient expertise to assess value of IC in critical K areas. (Glendon and Kundtz, 2000) 4) CKO builds strong interpersonal relationships with employees 5) Workspace set up to make it easy for people to talk to each other. (Davenport & Prusak, 1998; Darr, YEAR)	1) Amount of user satisfaction with system routines and knowledge availability. (Orr and Persson, 2003) 2) Frequency use of KB.(Orr and Persson, 2003) 3) Number of downloads. (Orr and Persson, 2003) 4) Number of calls to help desk (prefer decrease) (Orr and Persson, 2003) 5) Employees know exactly who to ask when they need information for their tasks. (Davenport & Prusak, 1998, Darr, YEAR)

Externalisation Combination	Explicit	1) Presentations on KM concepts & global case studies . (Orr and Persson, 2003) 2) Number of presentations. (Lee, 2000) 3) Training on how best to use system (writing, editing, formatting skills, etc.) in order to input items to knowledge repositories.	1) % hours assigned to project meeting (Moor et al.,2002) 2) Number of shared documents published .(Lee, 2000) 3) Number of patents published. (Lee, 2000).	1) Top management receives regular reports on KM initiatives and results (Chong and Choi,2005) 2) Top management use knowledge activity report in rewarding and motivating employees (Chong and Choi,2005) 3) KM group inside organisation (Chong and Choi,2005) 4) Knowledge DB administrator compared to previously.(Chong and Choi,2005)	1) Knowledge about customers used in tasks (Chong and Choi,2005) 2) Customer data-base searched extensively to obtain knowledge for tasks. (Chung, 2002) 3) Employee has ability to administer knowledge necessary for tasks systematically and store for future use (Chung, 2002)

Table F1: CSFs in People and Culture Category

CSF / Knowledge		Standard processes for knowledge contribution and content management	Organisation has ability to access, structure and categorise content of knowledge
Socialisation Internalisation	Tacit	1) Peer evaluation/ review and collaboration established practice through much/ parts of organisation (Lee, 2000) 2) KM embedded in routine business processes (Chong and Choi,2005) 3) % of non-assigned working time (Moor and Smits, 2002) 4) % of regulated socialisation (Moor and Smits, 2002) 1) Standard processes for retrieving knowledge 2) Standard processes for membership on CoP defined. (Lee, 2000) 3) Using techniques such as quality circles (Nonaka and Takeuchi, 1995) 4) Encouraging mentoring or coaching (Nonaka and Takeuchi, 1995)	1) Sources of tacit knowledge defined. (Glendon and Kundtz, 2000) 2) Social networks mapped, key personnel identified. (Glendon and Kundtz, 2000) 3) Requirements for sharing tacit knowledge evaluated.(Glendon and Kundtz, 2000)
Externalisation Combination	Explicit	1) % of hours assigned to project meetings (Moor and Smits, 2002) 2) Knowledge augmented with pointers to people (Lee, 2000) 3) Standard processes defined for (Chung, 2002): -Accepting content -Maintaining quality -Keeping content current -Deleting or achieving**???updating** obsolete content ? 4) Implementation-project based on knowledge reuse (Orr and Persson, 2003) 5) Knowledge augmented with pointers to people (Lee, 2000) 6) Organisation-wide knowledge and information updated regularly and well maintained (Chung, 2002) 7) Write case notes on successful and unsuccessful products and processes (Bennett and Gabriel,1999, Darr, YEAR	1) Professional knowledge such as customer knowledge & demand for casting managed systematically (Chung, 2002) 2) Processes & workflows documented (Glendon and Kundtz, 2000) 3) Org has begun codifying knowledge through use of shared networks, CoP and best practice (Lee, 2000) 4) Formats to document best practice and case studies defined. 5) Knowledge augmented with pointers to people (Lee, 2000) 6) Sources of explicit knowledge defined (Glendon and Kundtz, 2000) 7) No. of KB-communities (Orr and Persson, 2003) 8) No. of areas of expertise in KB-communities .(Orr and Persson, 2003) 9) No. of taxonomies in DB (Orr and Persson, 2003) 10) Internal best practices recorded (Gabriel, 1999)

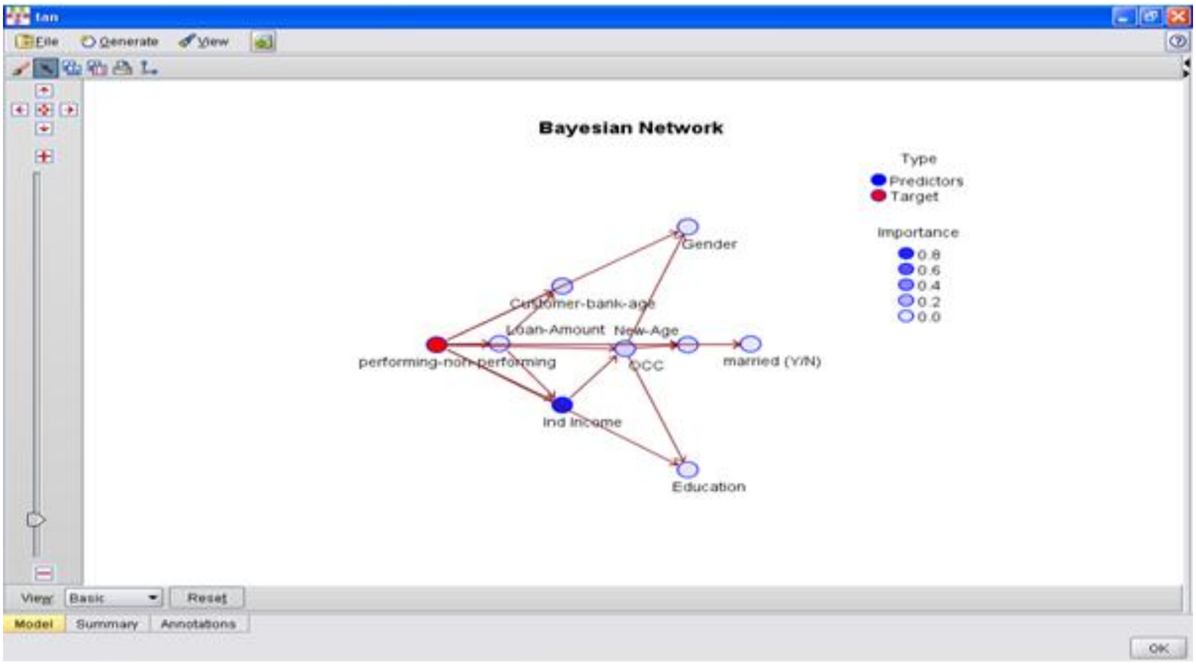
Table F2: CSFs in Process Category

Knowledge \ CSFs		Robust & user-friendly technology which is available to employees put in place	Tools for managing knowledge cycles activities have been established
Internalisation Socialisation	Tacit	1) Frequency use of KB (Orr and Persson, 2003) 2) No. of downloads (Orr and Persson, 2003) 3) No. of hits on document repository (Orr and Persson, 2003) 4) % of repeat visits to KB (Orr and Persson, 2003) 5) Size of discussion DB (Lee, 2000)	1) Size of discussion database (Lee, 2000) 2) Frequency use of video conferencing (Bennett and Gabriel, 1999) 3) Frequency use of teleconferencing (Bennett and Gabriel, 1999) 4) Good use of groupware, such as Lotus notes, to share information on products and notes (Bennett and Gabriel, 1999)
Externalisation Combination	Explicit	1) No. of sites accessed (Lee (2000). 2) No. of categories in KB (Moor et al., 2002). 3) No. of items in KB (Moor et al., 2002). 4) Total bytes of projects' document (Moor et al., 2002). 5) Scalability & ease of future upgrades to include new features (Lee, 2000). 6) Enables knowledge-contribution by employees (Moor et al., 2002). 7) Extent to which employees consider : - DB to be user friendly. - Help-instructions in DB sufficient. - Finding colleagues with correct competence easy (Orr and Persson, 2003).	1) Single point for saving and retrieving knowledge (Orr and Persson, 2003) 2) Availability of knowledge in IS's (Orr and Persson, 2003) 3) Infrastructure to facilitate knowledge capture, indexing & retrieval established to support knowledge sharing. (Chong and Choi,2005)

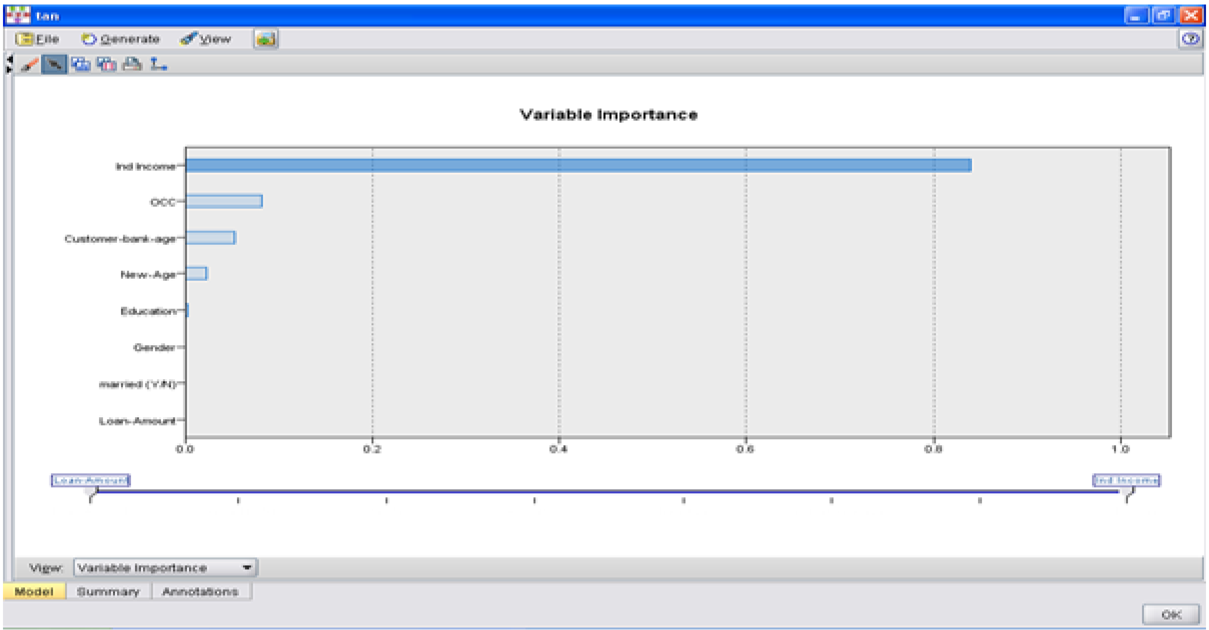
Table F3:CSFs in the IT Category

Appendix G

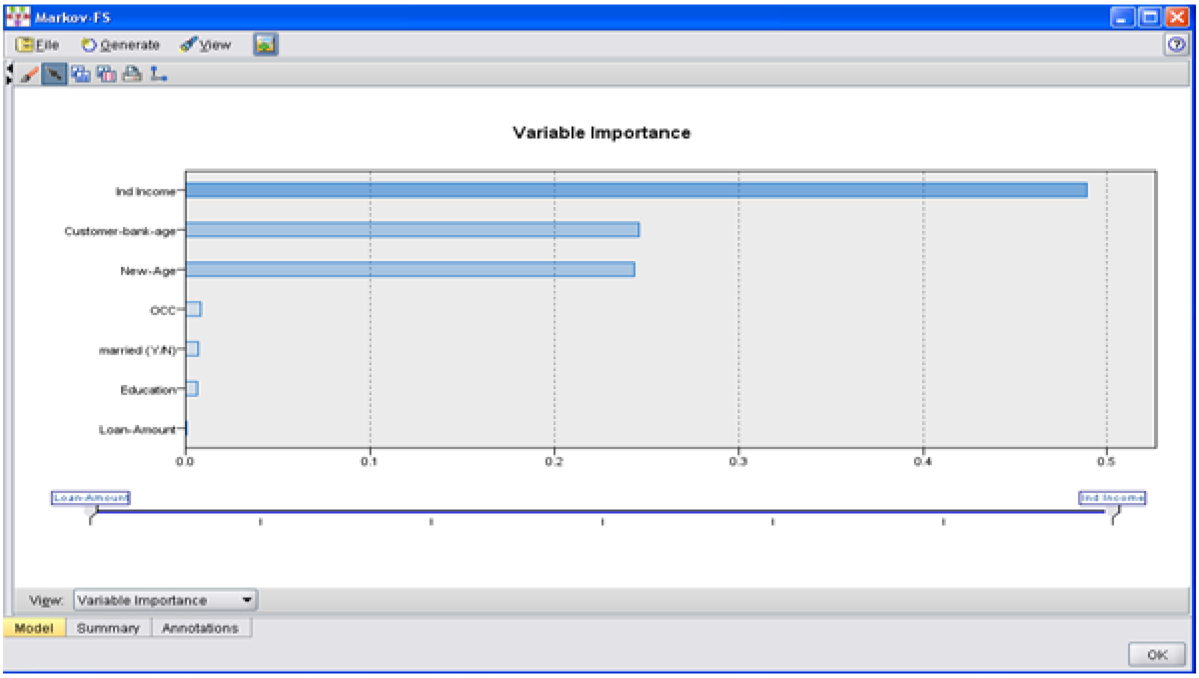
1. NAT Bayesian Network



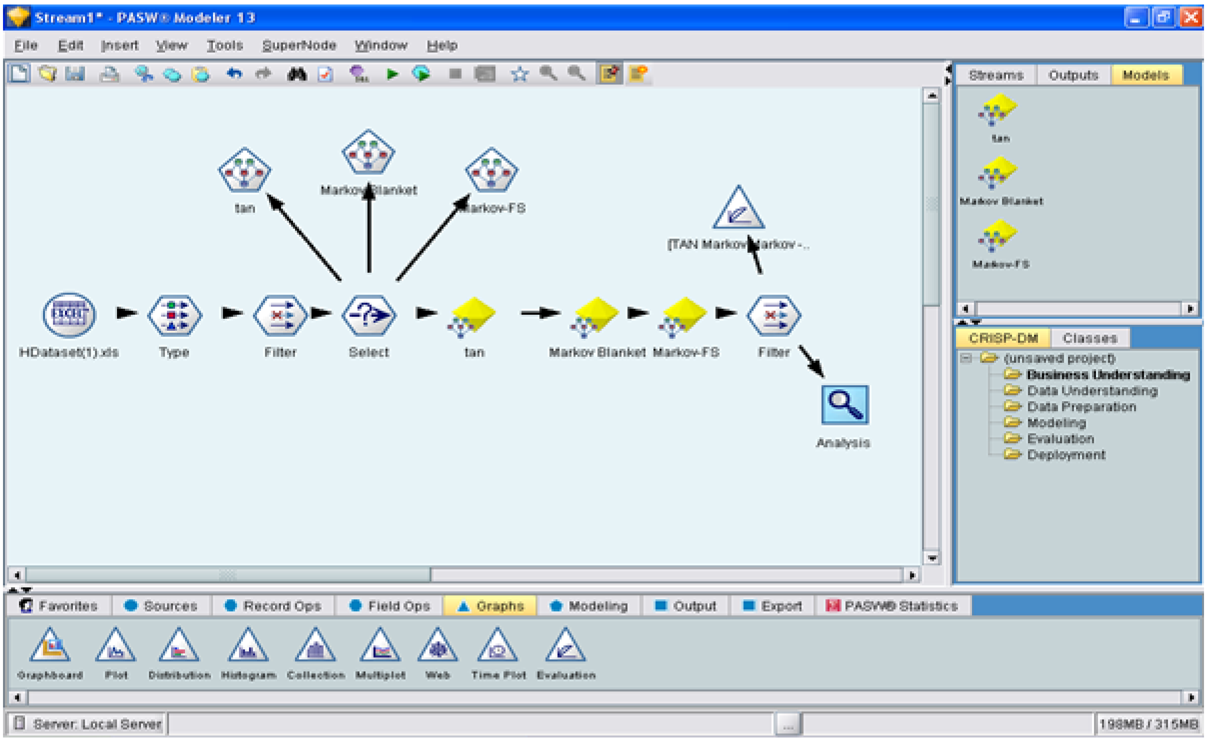
2. NAT variable importance



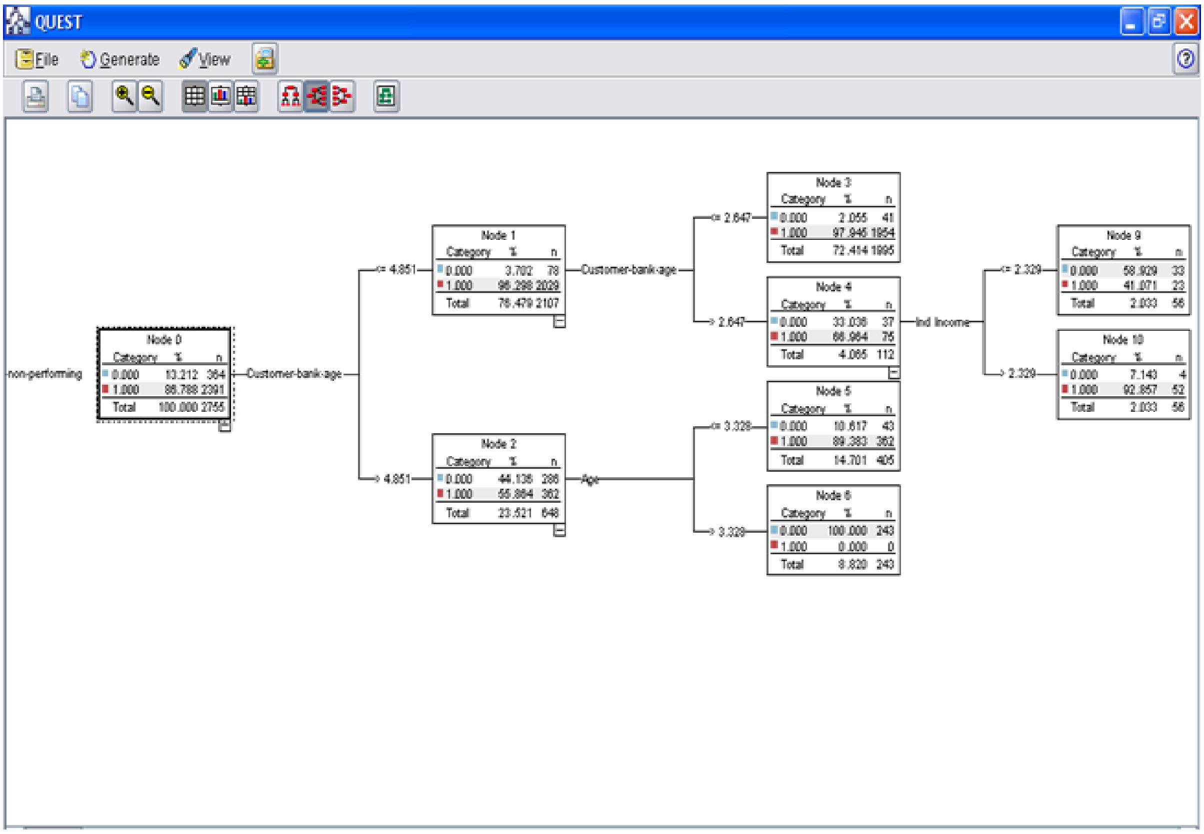
3. Markov FS variable importance

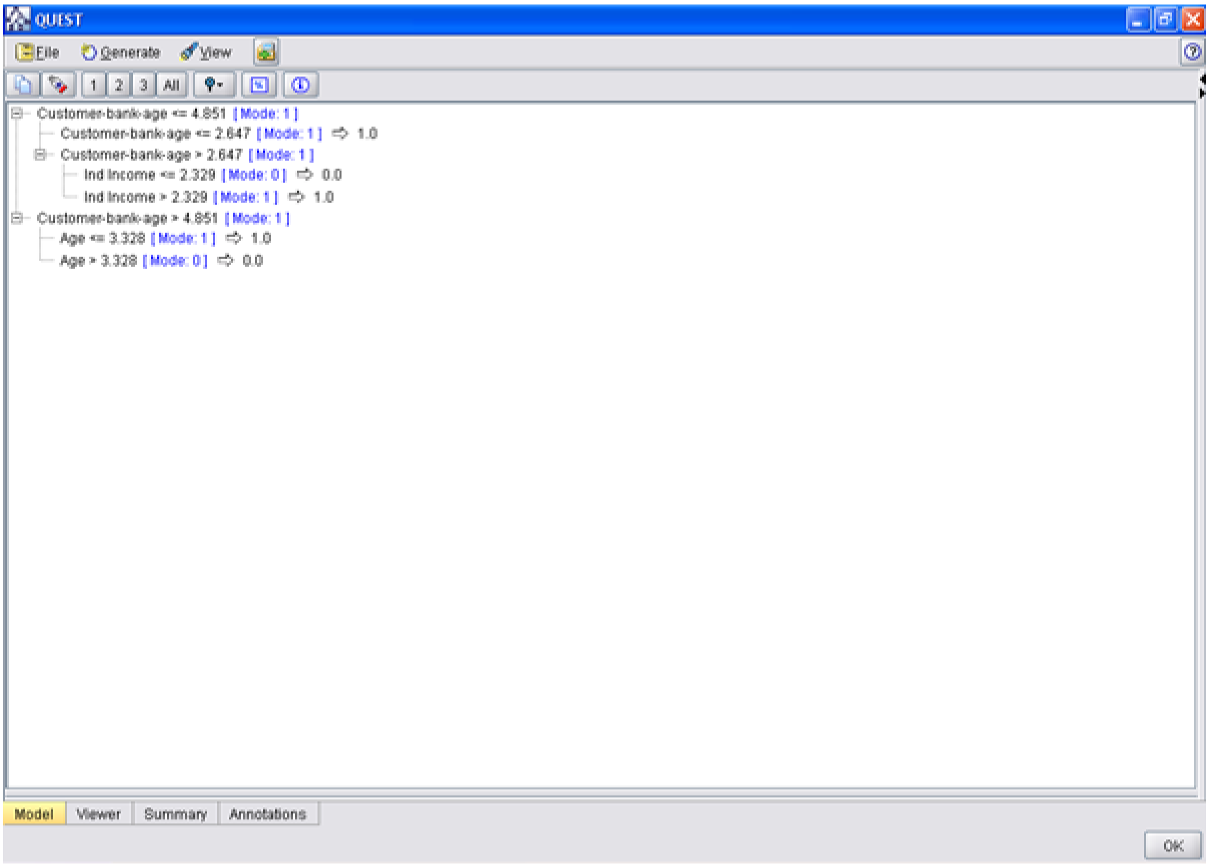


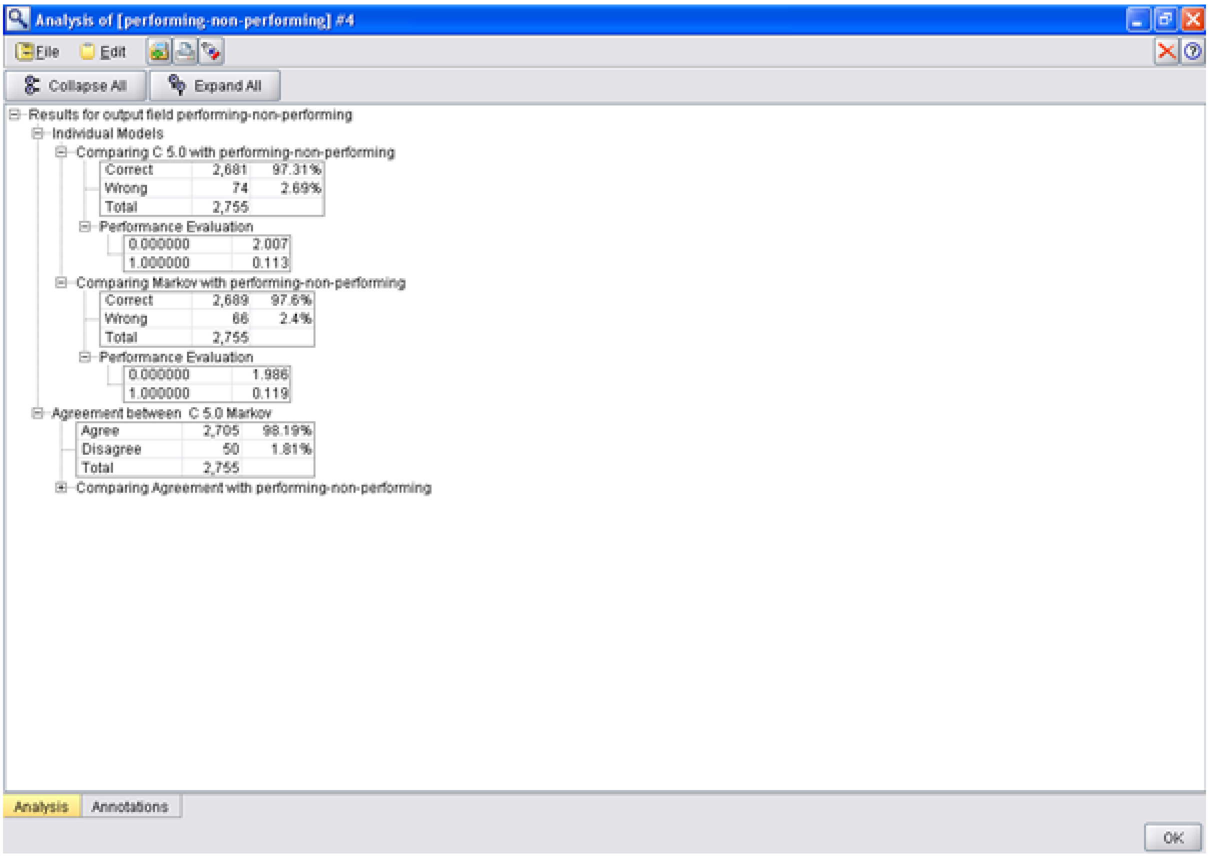
4. Stream



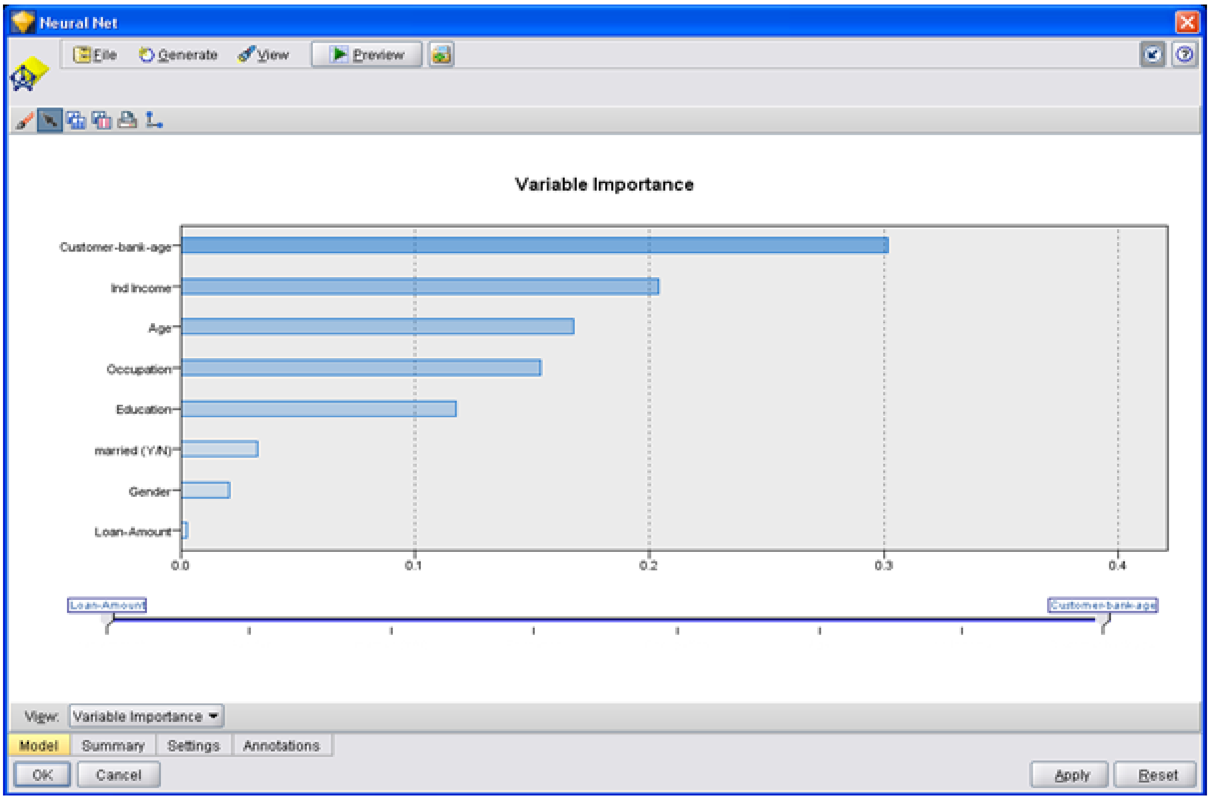
5. Decision Trees: QUEST

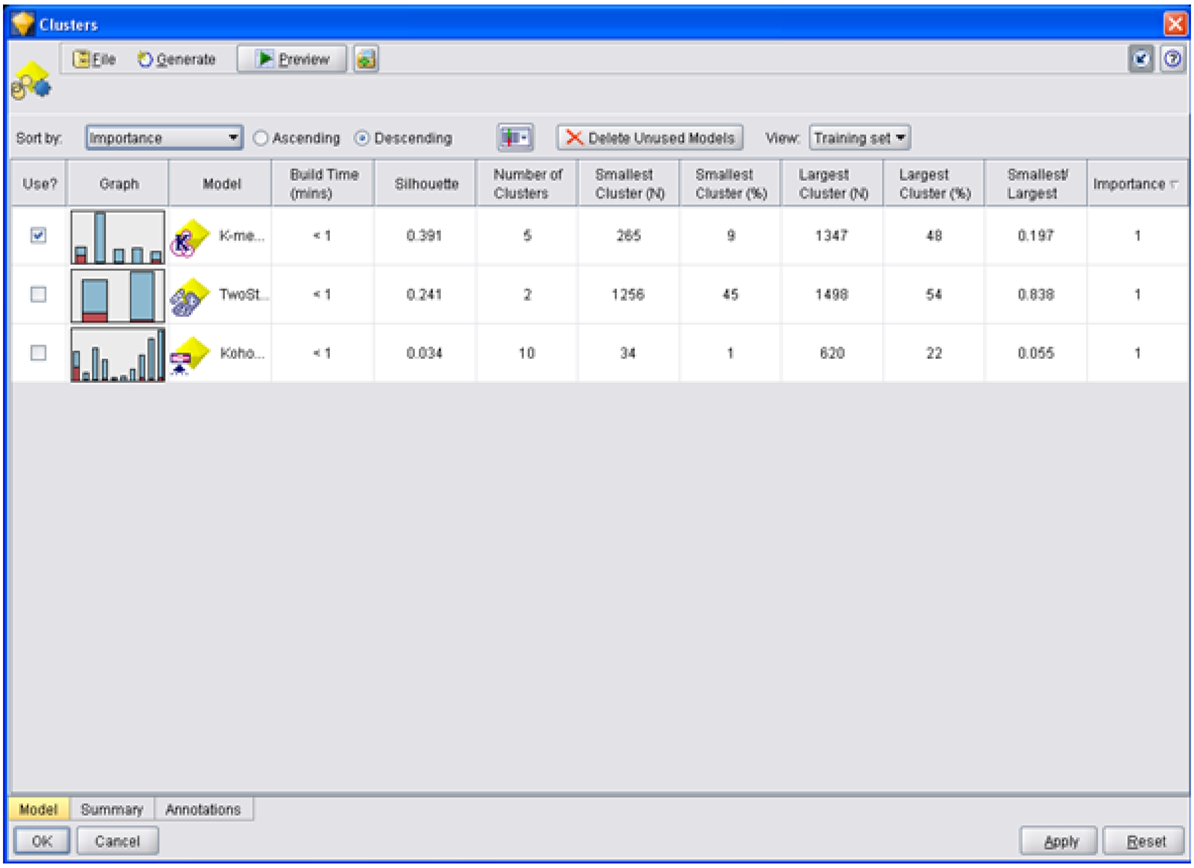


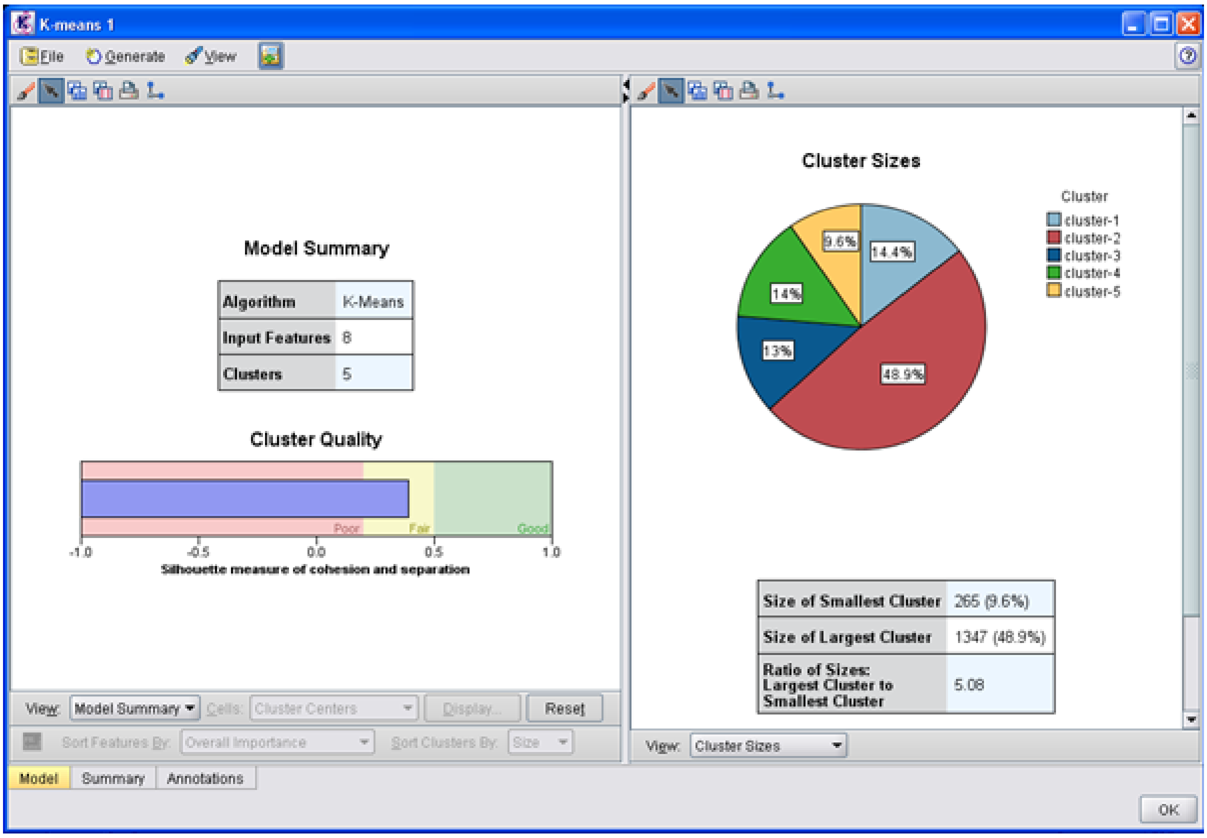

QUEST
File Generate View
1 2 3 All
Customer-bank-age <= 4.851 [Mode: 1]
Customer-bank-age <= 2.647 [Mode: 1] ⇨ 1.0
Customer-bank-age > 2.647 [Mode: 1]
Ind Income <= 2.329 [Mode: 0] ⇨ 0.0
Ind Income > 2.329 [Mode: 1] ⇨ 1.0
Customer-bank-age > 4.851 [Mode: 1]
Age <= 3.328 [Mode: 1] ⇨ 1.0
Age > 3.328 [Mode: 0] ⇨ 0.0
Model Viewer Summary Annotations
OK

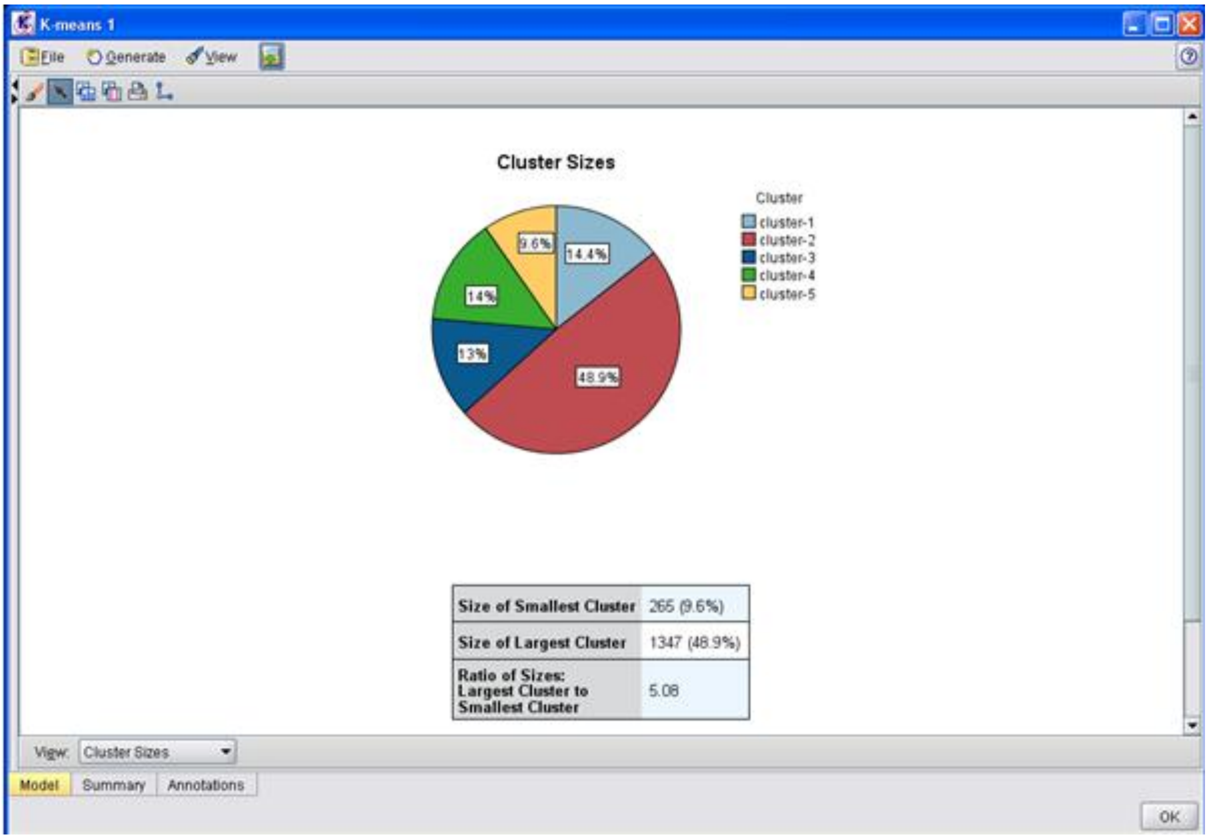

Analysis of [performing-non-performing] #4
File Edit
Collapse All Expand All
Results for output field performing-non-performing
Individual Models
Comparing C 5.0 with performing-non-performing
Correct 2,681 97.31%
Wrong 74 2.69%
Total 2,755
Performance Evaluation
0.000000 2.007
1.000000 0.113
Comparing Markov with performing-non-performing
Correct 2,689 97.6%
Wrong 66 2.4%
Total 2,755
Performance Evaluation
0.000000 1.986
1.000000 0.119
Agreement between C 5.0 Markov
Agree 2,705 98.19%
Disagree 50 1.81%
Total 2,755
Comparing Agreement with performing-non-performing
Analysis Annotations
OK

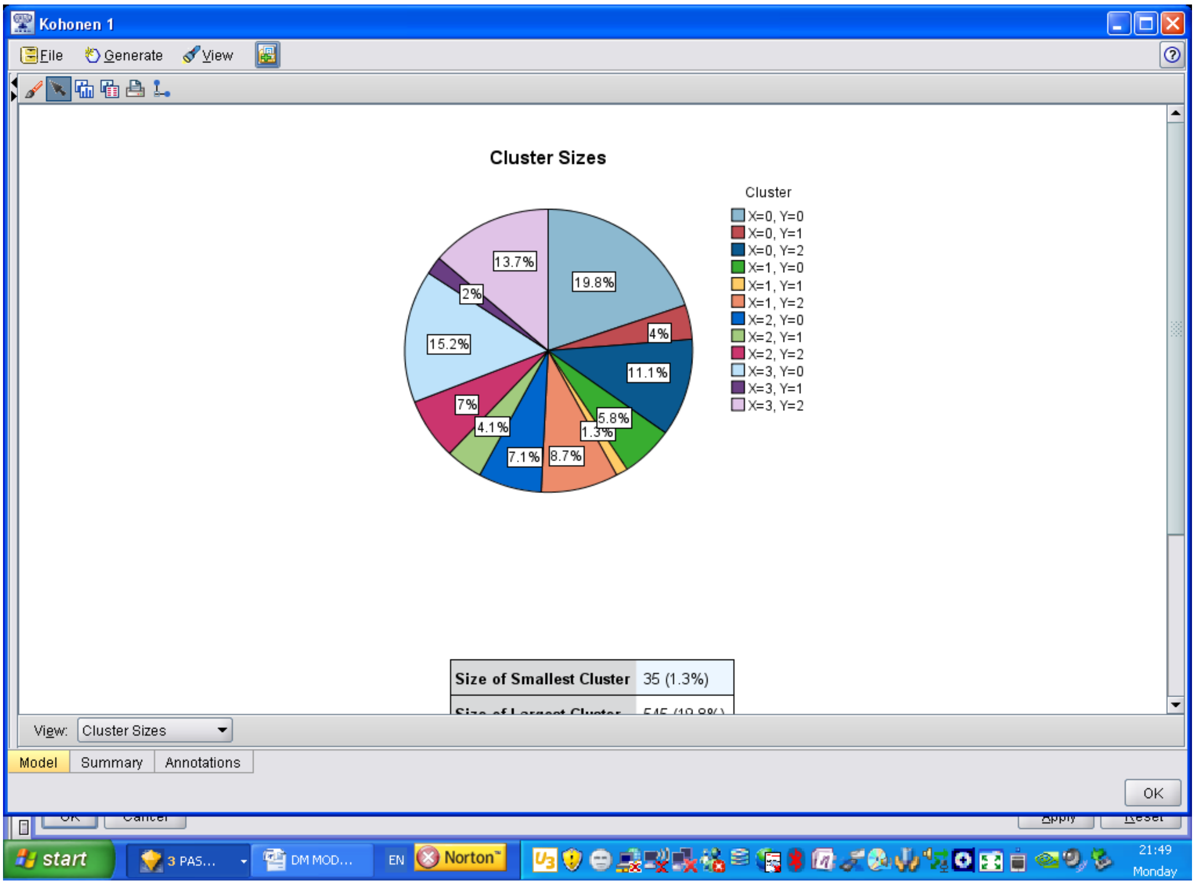
6. NN

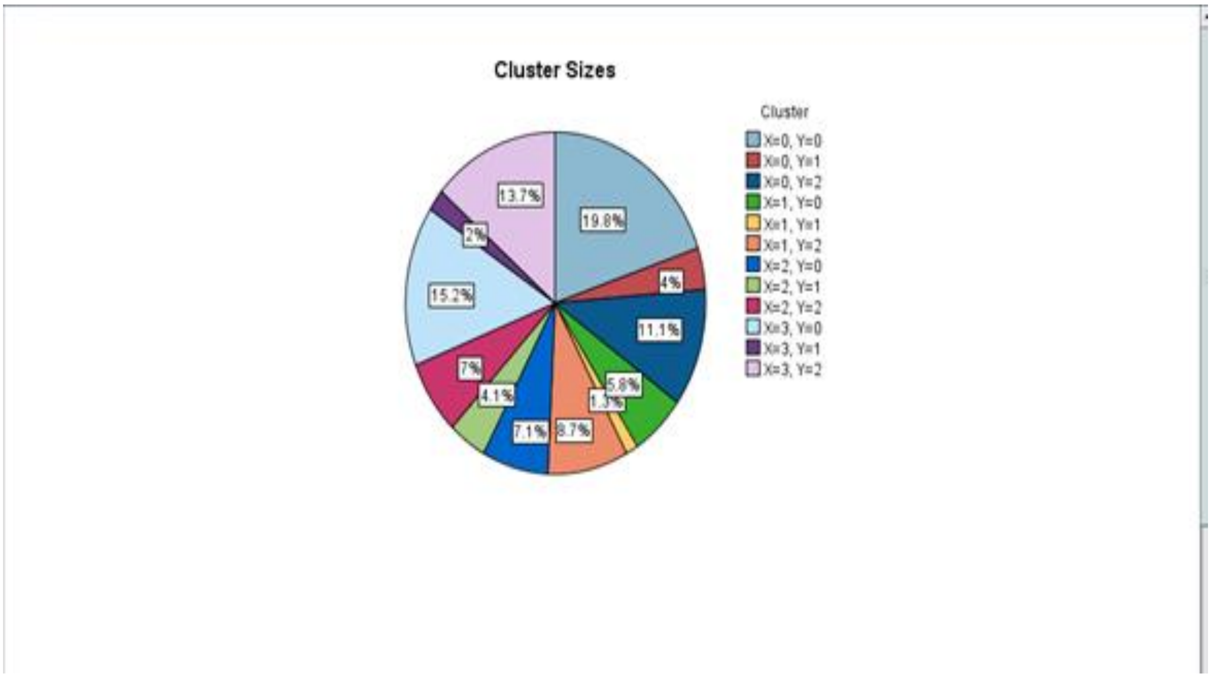

Neural Net
File
Generate
View
Preview
Variable Importance
Customer-bank-age
Ind Income
Age
Occupation
Education
married (Y/N)
Gender
Loan-Amount
0.0
0.1
0.2
0.3
0.4
Loan-Amount
Customer-bank-age
View: Variable Importance
Model
Summary
Settings
Annotations
OK
Cancel
Apply
Reset

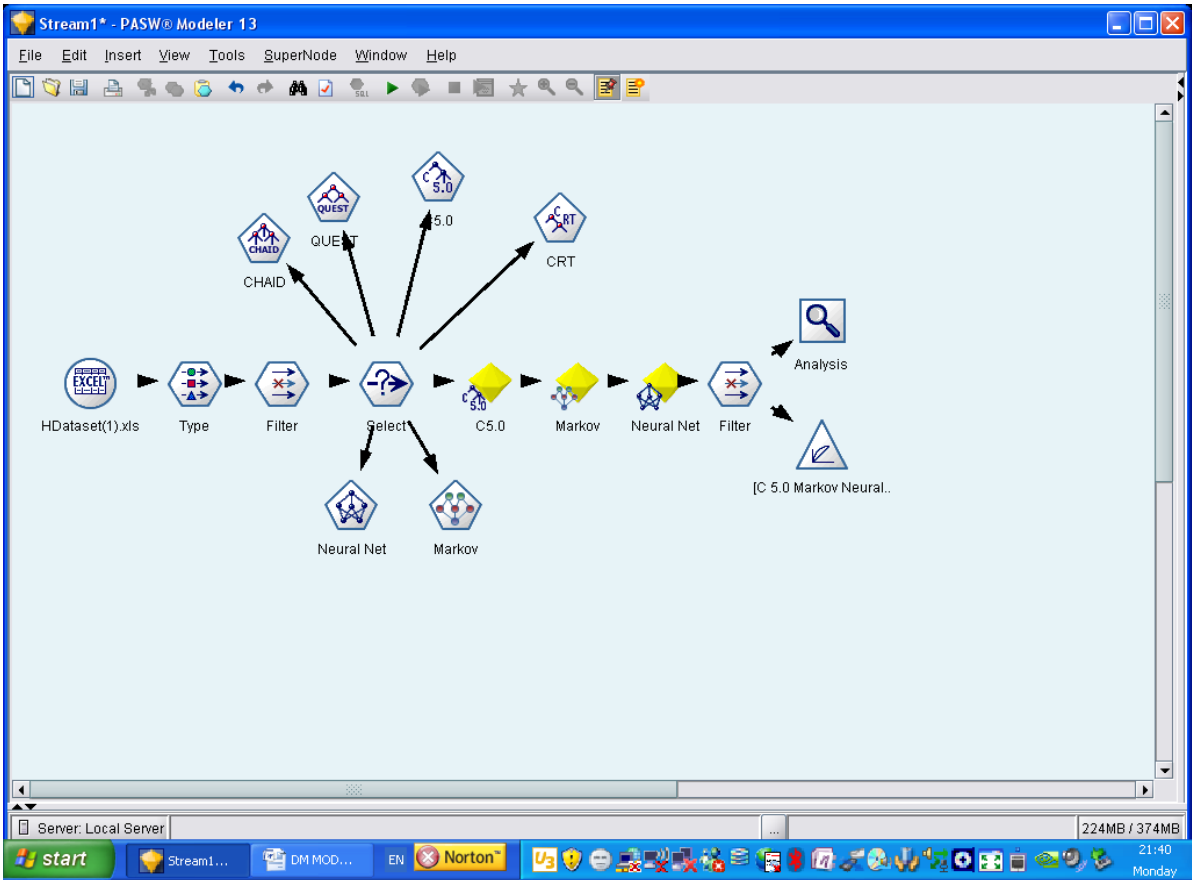

Clusters
File
Generate
Preview
Sort by: Importance
Ascending
Descending
Delete Unused Models
View: Training set
Use? | Graph | Model | Build Time (mins) | Silhouette | Number of Clusters | Smallest Cluster (N) | Smallest Cluster (%) | Largest Cluster (N) | Largest Cluster (%) | Smallest/ Largest | Importance
K-me... | < 1 | 0.391 | 5 | 265 | 9 | 1347 | 48 | 0.197 | 1
TwoSt... | < 1 | 0.241 | 2 | 1256 | 45 | 1498 | 54 | 0.838 | 1
Koho... | < 1 | 0.034 | 10 | 34 | 1 | 620 | 22 | 0.055 | 1
Model
Summary
Annotations
OK
Cancel
Apply
Reset


K-means 1
File
Generate
View
Model Summary
Algorithm
K-Means
Input Features
8
Clusters
5
Cluster Quality
Poor
Fair
Good
-1.0
-0.5
0.0
0.5
1.0
Silhouette measure of cohesion and separation
Cluster Sizes
Cluster
cluster-1
cluster-2
cluster-3
cluster-4
cluster-5
9.6%
14.4%
14%
13%
48.9%
Size of Smallest Cluster
265 (9.6%)
Size of Largest Cluster
1347 (48.9%)
Ratio of Sizes: Largest Cluster to Smallest Cluster
5.08
View: Model Summary
Cells: Cluster Centers
Display...
Reset
Sort Features By: Overall Importance
Sort Clusters By: Size
View: Cluster Sizes
Model
Summary
Annotations
OK

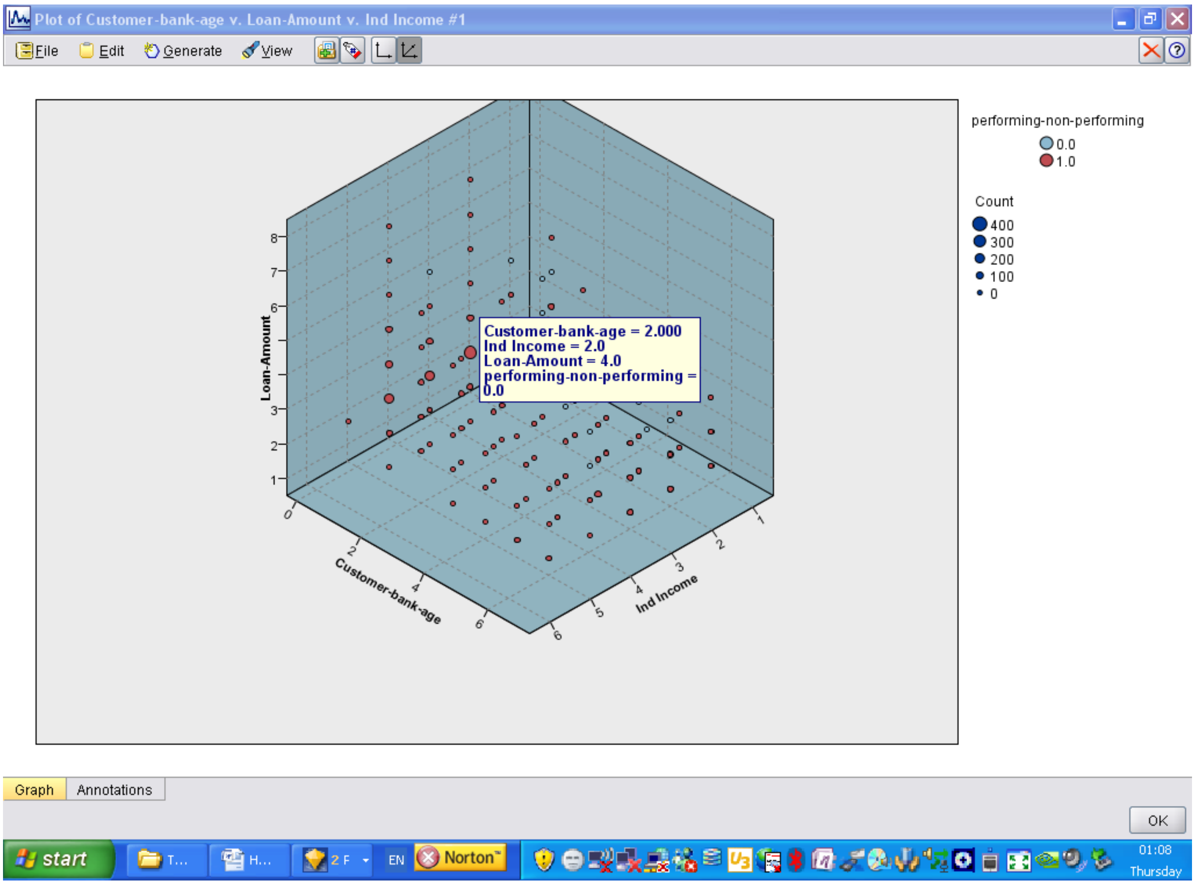

K-means 1
File
Generate
View
Cluster Sizes
Cluster
cluster-1
cluster-2
cluster-3
cluster-4
cluster-5
9.6%
14.4%
14%
13%
48.9%
Size of Smallest Cluster
265 (9.6%)
Size of Largest Cluster
1347 (48.9%)
Ratio of Sizes: Largest Cluster to Smallest Cluster
5.08
View: Cluster Sizes
Model
Summary
Annotations
OK

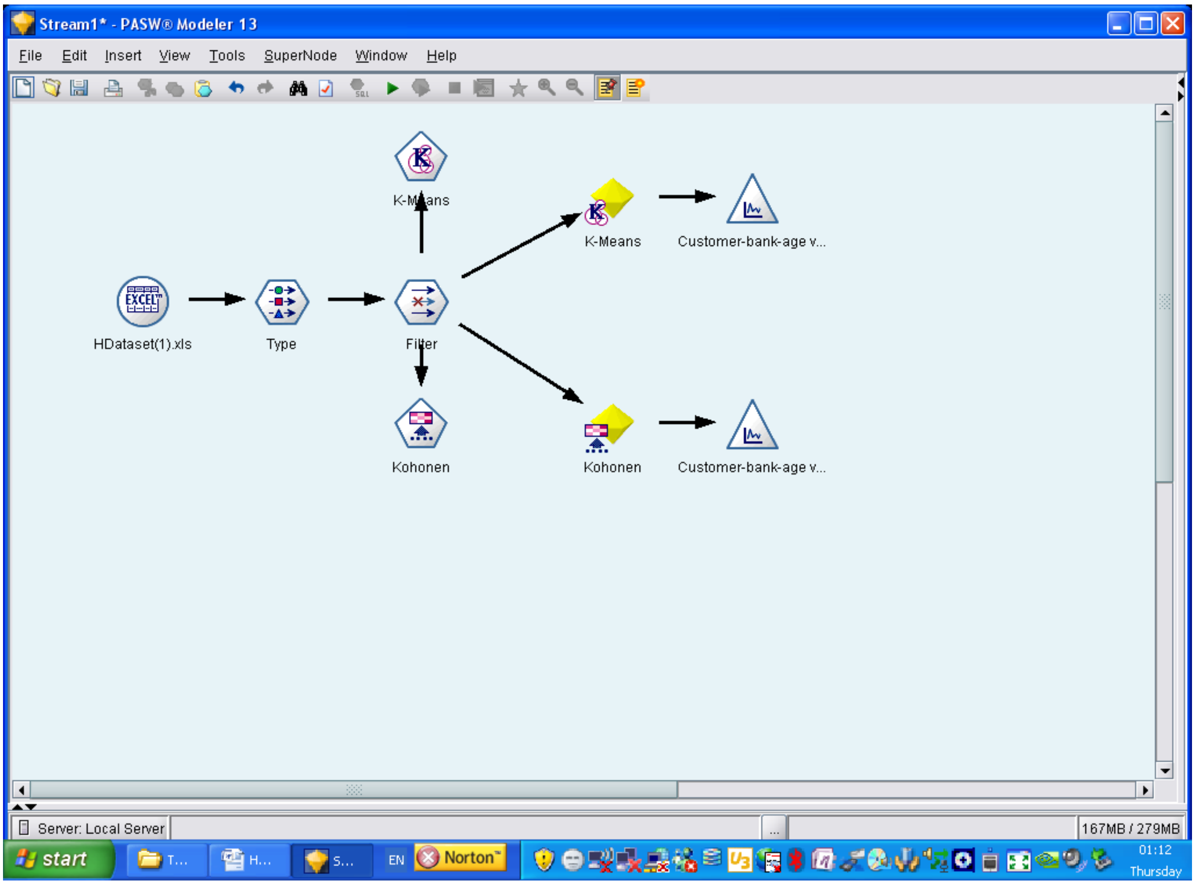

Kohonen 1
File
Generate
View
Cluster Sizes
Cluster
X=0, Y=0
X=0, Y=1
X=0, Y=2
X=1, Y=0
X=1, Y=1
X=1, Y=2
X=2, Y=0
X=2, Y=1
X=2, Y=2
X=3, Y=0
X=3, Y=1
X=3, Y=2
19.8%
4%
11.1%
5.8%
1.3%
8.7%
7.1%
4.1%
7%
15.2%
2%
13.7%
Size of Smallest Cluster
35 (1.3%)
View:
Cluster Sizes
Model
Summary
Annotations
OK
start
Norton
21:49
Monday

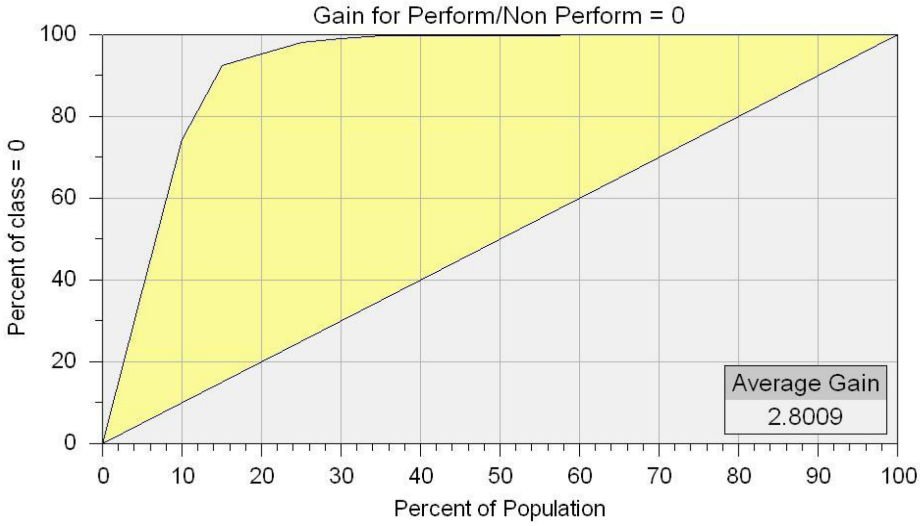

Cluster Sizes
Cluster
X=0, Y=0
X=0, Y=1
X=0, Y=2
X=1, Y=0
X=1, Y=1
X=1, Y=2
X=2, Y=0
X=2, Y=1
X=2, Y=2
X=3, Y=0
X=3, Y=1
X=3, Y=2
19.8%
4%
11.1%
5.8%
1.3%
8.7%
7.1%
4.1%
7%
15.2%
2%
13.7%

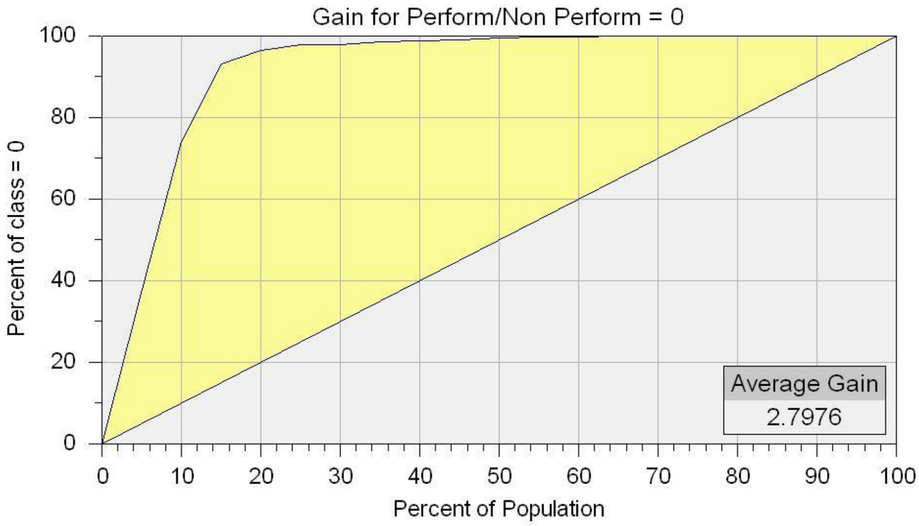


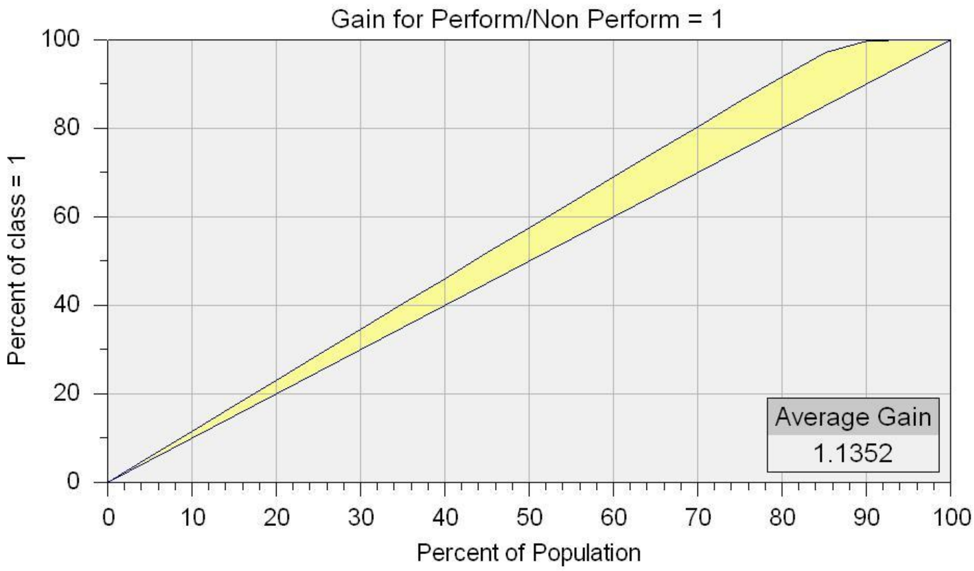
Stream1* - PASW® Modeler 13
File Edit Insert View Tools SuperNode Window Help
CHAID
QUEST
C5.0
CRT
HDataset(1).xls
Type
Filter
Select
C5.0
Markov
Neural Net
Filter
Analysis
[C 5.0 Markov Neural..
Neural Net
Markov
Server: Local Server
224MB / 374MB
start
Stream1...
DM MOD...
EN
Norton
21:40
Monday

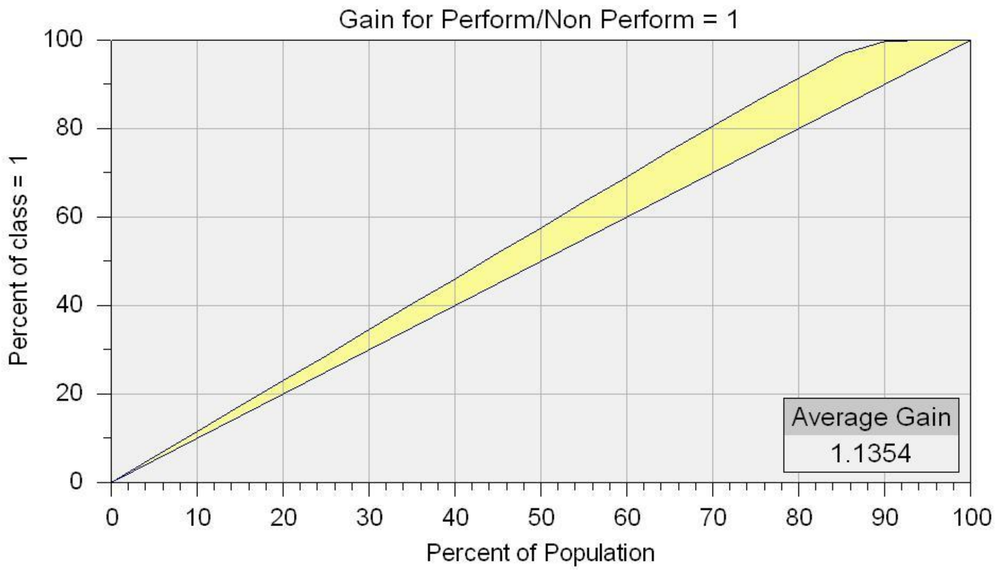

Plot of Customer-bank-age v. Loan-Amount v. Ind Income #1
File Edit Generate View
performing-non-performing
0.0
1.0
Count
400
300
200
100
0
Loan-Amount
Customer-bank-age
Ind Income
Customer-bank-age = 2.000
Ind Income = 2.0
Loan-Amount = 4.0
performing-non-performing = 0.0
Graph
Annotations
OK


Stream1* - PASW® Modeler 13
File Edit Insert View Tools SuperNode Window Help
K-Means
K-Means
Customer-bank-age v...
HDataset(1).xls
Type
Filter
Kohonen
Kohonen
Customer-bank-age v...
Server: Local Server
167MB / 279MB


Gain for Perform/Non Perform = 0
Percent of class = 0
Percent of Population
Average Gain
2.8009


Gain for Perform/Non Perform = 0
Percent of class = 0
Percent of Population
Average Gain
2.7976


Gain for Perform/Non Perform = 1
Percent of class = 1
Percent of Population
Average Gain
1.1352



At 95% the odds ratio for non-performing classification is as follows:

Odds Ratio for Non Performing @ 95% Confidence Interval		
	Unadjusted	**Adjusted**
Predictors	**Odds Ratio**	**Odds Ratio**
less than or equal 5000	3.351808329	0.161766717
5001-10000	2.592441252	0.359478827
10001-30000	0.231864994	1.153729802
30001-50000	0.266228584	20.55285786
50001-100000	0.352819036	50.59233082
100001-200000	1.010656379	854.4004544
200001-350000	9.66288E-14	1.12603E-07
greater than 350000	3.11845E-11	1.80138E-10
less than or equal One year	0.021205388	0.036842618
1 – 2	0.663451185	0.640696124
2 – 3	0.920167095	1.432612929
3 – 4	11.41239448	88.79230581
4 – 5	2.034805018	19.20940495
More than Five Years	25.62324635	44.13706916
Gender	0.927465205	0.840044846

Less than or equal J.D400	6.014661068	3989.016547
401 – 600	5.818252739	3583.875372
601 – 800	0.179245304	187.2603697
801- 1000	0.163310824	122.6334715
Greater than 1000	7.9181E-10	2.41605E-11
Martial_Status	0.227001194	0.334171816
Uneducated	20.0193639	741.1477514
High School	8.533503466	7.167808791
Diploma	1.649051048	1.03790058
Bachelor	0.38550543	0.902939253
Postgraduate	0.100158635	0.164721353
Un-Known	9.7052E+31	1.24149E+68
Non-workers	0.304464738	0.102930631
Vocational	4.494255384	0.948285179
Government jobs and Military	0.445525854	0.362873859
Marketing, Sales and Finance, customer services	8.520712806	2.427592438
Health services	0.168587563	0.235133933
Education, Training and Media	11.2999692	2.407767573
Employee	0.244045645	0.135660478
Law and Engineering	1.0010005	2.800505678
Business, Management and administration	0.540965377	1.999705661
less than 20	4.88616E+34	4.14629E+68
20 – 29	0.182957755	0.713266611
30 – 39	0.109854337	0.129431952
40 – 49	4.873439741	6.112280841
50 – 59	4.552606837	4.276350676

>=60	5.609715471	16.36753828

The above Table can be used to find the confounding of certain predictor variables and their relationship to non-performing outcome classification by analysing it for change in value of odds ratio from unadjusted to adjusted. If certain change is visible, it suggests the impact of that predictor variable on classification.

It can be observed from Table above that when loan amount is 50,000-100,000 or 100,0001-200,000 the odds ratio has changed significantly from 0.352 to 50.592 and 1.010 to 854.400 when used with other predictor variables. This certainly shows that these predictor variables statistically contribute significantly to decision of non performing. Similar changes are observed in customer age 3-4, 4-5 and >5 as odds ratio changes from 11.412 to 88.792, 2.034 to 19.209, and 25.623 to 44.137 respectively. All as predicted coginitvely, income should have effect in perform or non perform outcome, we see the large change in odd ratio for all incomes making them statistically significant when used with other predictor variables. Uneducated loan applier will also suffer from increase in beta value of predictor variable from 20.019 to 741.147. Finally age when >=60 changes the odds ratio from non performing from 5.609 to 16.367.