CHAPTER 2 : REVIEW OF LITERATURE

2.1 INTRODUCTION

This chapter is an attempt to put forth a repository of research studies carried out in the areas of credit scoring, credit scoring models, their importance, usages and limitations, the techniques used in developing credit scoring models, criteria for evaluating credit scores, application of credit scoring in different fields of studies and the application of credit scoring in the area of small business lending. It also investigates the available literature regarding the determinants of repayment behaviour of small business customers. The aim of this chapter is to identify significant research gaps that would pave the way to identify the need for the study and to formulate the objectives of the study.

2.2 CREDIT CHALLENGES FACED BY SMALL BUSINESSES

Despite its huge contribution to the nation’s GDP, exports and employment, the small business sector does not get the due support from the government, formal financial institutions like banks, etc. The inability of small businesses to survive beyond a year of start of business is largely attributed to lack of timely financial support (Amoako, 2013). Of the problems faced by SMEs, financing of both short term and long term needs is the most critical (De, 2009). Credit helps small businesses to undertake productive investments like investment in latest technology; and expansion of business (Improving the Competitiveness of SMEs in Developing Countries, 2001). Financial problems faced by small businesses is the root cause of all problems faced by these enterprises (Aruna, 2015). This has handicapped the growth and development of the sector and the competitiveness in the national and international markets. Lack of adequate credit force SMEs to remain curtailed low-key, self-sustained and unambitious.

a.) Challenges faced by Financial Institutions

Lending to small businesses has its own set of constraints from both the lenders and borrowers point of view. While lenders find it unprofitable to lend to SMEs, SMEs
find borrowing from formal credit sources very costly and cumbersome. Some of the major problems faced by lenders are:

i. **Information Asymmetry**
The incompleteness of information regarding the quality of a proposed business by a bank’s customer raises the problem of adverse selection (Type two error) is a major cause of banks’ disinterest in lending to small businesses (Stiglitz & Weiss, 1981). It becomes too costly for banks to gather all the information required before granting a loan.

ii. **Lack of sizeable security**
Small businesses are sensitive to the changes in the economy at large and also local conditions with makes it highly unpredictable in terms of business cycles, growth of the business, sustainability, etc. Therefore, banks find it safe to lend, with some form of security. However, small businesses especially in the service sector find it hard to provide sizable security to the banks due to low capitalisation, insufficient assets and high mortality rates (Bhattacharya, Faiz, & Zohir, 2000).

iii. **Absence of accounting and financial records**
Many studies have shown that it becomes very hard for small businesses to gain credit from financial institutions due to lack of financial records (Williams, Haka, Bettner, & Carcello, 2008).

iv. **Little or no credit history**
Credit footprints are nearly impossible to detect for small businesses in India. The sector largely consists of unregistered businesses that are difficult to be traced. Over indebtedness has become a serious issue in case of micro finance and small business lending. Credit information is costly to acquire and not the worth the income from small businesses (Myers & Majluf, 1984).

v. **Moral hazard**
This is a very interesting aspect that many bankers consider important. Interest rates are determined based on the risk of the projects undertaken. Charging high interest rates will force the customers to invest in even higher risky projects to earn higher returns (Carbó-Valverde, Rodríguez-Fernández, & Udell, 2008). Also constant
monitoring of the activities of the enterprise is very costly for the banks (Lean & Tucker, 2001).

**vi. Small ticket size**
Small businesses being small in nature have smaller credit needs than big businesses. This discourages banks from opting to lend to small businesses and look out for large businesses with the same effort. It costs the banks just as much as it takes to underwrite a small loan as it does to a big loan (Banks not lending to small businesses, 2015).

**vii. High transaction costs**
Loans to SMEs are small in size and hence increase the cost of doing business through administration costs, maintenance costs, cost of getting information, etc (Kumar, Batra, & Sharma, 2009).

**viii. Business/Financial Illiteracy**
Emerging nations rank in the lower half of World Banks list of ease of doing businesses indicators. The biggest contribution to it is the borrowers’ financial illiteracy (Chironga, Dahl, Goland, Pinshaw, & Sonnekus, 2012). Entrepreneurs who are unable to financially plan their revenues and expenses are exposed to higher levels of business risk and are on a straight path to failure (kamo, 2015).

**b.) Challenges faced by MSMEs**
Customers also face lot of challenges that makes borrowing from banks and other formal sources very cumbersome. It discourages small business firms to take credit from formal sources. Some of the reasons are:

**ix. Cumbersome procedures**
According to the observations made in a national level study on banks, the organisational systems and internal processes have made the banks move far from task orientation and have generated a bias against small loan portfolios (Sebastian & Basanth, 2005). The entire process of getting even a small loan is time consuming and requires the borrower to visit the banker several times and involves a lot of paperwork. Thus it becomes unattractive and the borrower looks out for easier sources, like local money lenders.
x. Lack of customised products
Unlike in developed countries, loans to small businesses in developing countries are standardised. Cash flow cycles, nature of business (cash/credit), business payback period, etc are never considered in determining the interest rates and loan repayment instalments (Vasu & Jayachandra, 2014).

xi. High interest rates
Due to higher risks banks perceive, high interest rates are charged from small businesses. Unpredictability, sensitivity to local conditions, uneven cash-flows all add up to why banks charge high interest rates (Banks not lending to small businesses, 2015).

xii. Insufficient credit
Usually small businesses are severely starved of funds. Revenues generated are either insufficient or just sufficient to run the business without any growth or expansion. Firms need some sort of external finance for further development of the business. Looking at the position of the business, financial institutions find it difficult to entirely meet the proposed needs of the business (Singh, Pareek, & Kapoor, 2014).

xiii. Delay in disbursement of Funds
Unavailability of credit at the right time can create a serious issue of liquidity and can also lead to shutting down of the business (Tripathi, Tripathi, & Rikin Dedhia, 2016).

xiv. No Standard procedures to evaluate creditworthiness
Financial institutions follow out dated evaluation procedures to study the creditworthiness of a business. These archaic processes do not reflect the credit standing of businesses in a dynamic economy (Ingole, 2014).

xv. Lack of business freedom
Firms that are very small have flexibility to change their business to meet the changing commerciality of businesses in the local scenario. They are always on a look out for more profitable avenues. However, many times lenders do not permit the borrowers to venture into these businesses that they may perceive as risky (Kumar, Batra, & Sharma, 2009).
Inability to get further credit

When a firm has utilised its maximum credit capacity, it becomes extremely difficult to seek further any formal sources of credit even in times of emergency, then firms resort to informal sources of credit (Aruna, 2015).

2.2 CREDIT SCORING

Credit Scoring is a statistically derived measure of risk. Credit scoring is a process that uses a host of information to generate a statistical output in the form of a linear scale represented as a score point or a probability percentage indicative of the predictability of default, delinquency and creditworthiness. Credit scores take into consideration many variables like past repayment history, past business trends, current macro-economic variables, etc.

This section is a brief of the literature reviewed in areas regarding the importance and benefits of credit scoring, limitations of credit scoring and different types of credit scores.

2.2.1 Importance and Benefits of Credit Scoring

Credit scoring has evolved over the years as a solid, structured and definite measure of credit risk. The dependability of lenders on credit scoring has played a crucial role in increasing credit facilities to deserving borrowers and curbing the growing uncertainty in lending. The first probable use of a statistical scoring model to distinguish ‘good’ from ‘bad’ customers was by Durand (1941). He used data from sources like banks and other financial companies. One of the most critical decisions that a banker has to take is to distinguish the creditworthy customers from those who are not, and credit scoring is the right tool in today’s lending practices. Bank lending is higher and simultaneously credit risk lower for countries adopt Credit Information Systems (CIS), whether the systems are public or private (Jappelli & Pagano, 2002).

Unlike popular believes that credit scoring would replace the existing credit appraisal techniques, in reality they complement the already existing system (Schreiner, 2000). Frame, Srinivasan, & Woosley, (2001) in their study concluded that the use of credit scoring has helped in reducing the cost of information between the borrowers and lenders. The growing awareness of scoring of borrowers incentivizes them to repay the credit and hence improves the creditworthiness of the customers (Vercammen,
Credit scoring has a cutting edge over traditional techniques of assessment of credit risk. The success of a traditional judgmental process depends on the success rate of the decision-maker, his experience and a host of other factors limits the success of the method. The judgmental process involves subjectivity, inconsistency and individually motivated preferences (Sullivan, 1981) and (Bailey, 2004). Credit scoring also involves fewer variables that are statistically important as compared to a judgmental process that is usually cluttered with many variables (Crook, 1996). Credit scoring reduces the time taken in the process of credit assessment. A study in the US claimed that the time taken to assess and approve a small business loan reduced from two weeks to one and a half hour after the adoption of a credit scoring model (Allen, 1995). In another study of a Canadian bank by Kevin Leonard (1995), it was observed that the loan approval time reduced form nine days to three days. Credit scores take into consideration the characteristics of both good and bad payers, while the judgmental processes tend to take into consideration only bad payers (Chandler & Coffman, 1979). Credit scores also assist bankers in increasing the sale of additional products (Abdoua & Pointon, 2011).

In her thesis (Amari, 2002) lists a few benefits of credit scoring. Some of them are: reduction of processing costs and time; efficient processing time; probability of fewer errors; inclusion of variables that objectively support the credit risk evaluation; a supporting system for the decision making process; consideration of interrelated variables; flexibility in changing the cut off scores based on the changing environment; Post Audit comparisons; Modelling using real data;

### 2.2.2 Limitations of credit scoring

Credit scoring has become a globally accepted phenomenon today, but the use of credit scoring models pose certain limitations that need to be looked at before completely committing to credit scoring models.

The greatest difficulty of credit scoring is not the technical aspect of the model but the organizational aspect. Credit scoring on paper can be very useful. Implementation of the model with the constraints of the existing technology, data base and management
systems can be very difficult. Integration of the scoring model with the organizational systems is very critical. (Chakravarty & Jha, 2012). Many times the use of credit scoring can be threatening to the credit appraisal personnel. They refuse to replace traditional processes with credit scoring systems (Yadav, 2011). In many organizations the use of credit scoring is costlier than the traditional system of appraisal (Collins, Harvey, & Nigro, 2002). Lack of information about customers also hampers the utilization of all characteristics in the usage of a credit scoring model. This lack of information may lead to selection of customers who only have information on all characteristics (Fledmen, 1997). Many credit scoring models are not standardised and are customised for a specific market. This may lead to denying a customer credit facilities due to change in job or location (Amari, 2002). Overdependence on credit scoring model can have a dangerous outcome, especially when the user does not understand the complexities of a credit scoring model. (Gool, Baesens, Sercu, & Verbeke, 2009). The small business lending or microfinance lending is very different as compared to other forms of lending. The predictability of the future based on the past information is not always true in a small business lending. Thus, credit scoring may not be practical (Rhyne, 2001). Certain important factors which cannot be quantified or evaluated are ignored by credit a scoring model, which is a serious disadvantage of a credit scoring system (Freytag, 2009). Not just data, digital data is what is required in a credit scoring process. This problem, especially in the developing nations, where historically data has always been collected on paper and requires lot of energy, cost and time to digitize the data, is critical (Schreiner, 2004). A grave problem that credit scoring poses is the outcome of the results. Most of them generate a dichotomous result predicting the borrower as an eventual defaulter or not. In reality, there are a many other outcomes that could take place like, delay in repayments, defaults on only interest repayments, partial repayments, etc that cannot be deciphered from the output (Heffernan, 2005).

2.3.1 Types of Credit Scores

Presentation of different classifications of credit scoring models is important for the study. The credit scoring methodology used in the study has been determined after a thorough research into the types of credit scoring models that have been used for different studies. Scholars in the field of credit scoring have over the years developed
different models for meeting the needs of borrowers and other parties interested in the lending processes. A thorough review of literature will tell us that there have been different ways of classification of credit scoring models. Classifications of models are based on different parameters like usage, methodology, purpose, etc.

**a) Generic Vs Custom Scores**

One very broad classification of credit scoring models is based on the applicability of the models. Models that are developed by bureaus or institutions based on industry data applicable to all participants or subscribers are called ‘Generic models’. Models built for a specific group of borrowers, who have specific requirements or need a more in-depth underwriting when compared to other borrowers, are called ‘Custom Models’ (Mays & Lynas, 2011). FICO (Credit scores like Fair Isaac Corporation), Vantage and CIBIL etc can be classified as generic models. On the other hand, most of the infamous models that are specific to certain institutions or groups of customers are classified as custom models. Even though conceptually customized scores are better off in the evaluation of a customer, certain important issues like feasibility, development and implementation can be a hindrance to the application of customized models (Chandler, 2001)

**b) Credit scoring methods based on techniques used**

Few scholars have narrowed down to a 3 category style of credit scoring methods (Schreiner, 2003; Caire 2004; Caire et al., 2006; Sur, 2008) – Statistical Methods, Judgmental Methods (or Expert Methods) and Hybrid Methods (a combination of the first two). As the names suggest, the first category uses statistical tools and techniques for building the scores. The second category relies on the judgmental capacity of the evaluator. The third category is a mixed usage of the first two methods.

**c) Other scores based on usage**

There are a whole lot of methods that are named and classified based on the usage of the models. Although, traditionally a scoring model is meant to determine the creditworthiness of customers, models have been developed for different purposes. The FDIC (Federal Deposit Insurance Corporation), USA lists a few scores used for various purposes (Credit Card Activities Manual):
i. **Application Scoring**

In this system of scoring variables like income level, salary scale, period of employment, etc are used to make up a credit score that evaluates the terms and conditions of giving credit.

ii. **Attrition Scoring**

As the name suggests, this score helps identify customers who are likely to close their dealings with the bank. The score weighs the benefits against the cost of maintaining the accounts and balances of the customer.

iii. **Bankruptcy Scoring**

This scoring is very similar to a credit score but does not evaluate the creditworthiness but the probability of customers defaulting and declaring bankrupt.

iv. **Collection Scoring**

The chance of taking delivery of payments from the customers is calculated in this system. The ranking of customers will help the lenders in determining the strategies to be adopted in collections.

v. **Behaviour Scoring**

A behaviour scoring system involves the exclusive study of behavioural variables like payment patterns, credit patterns, delinquency patterns etc related to the customers. A behaviour score is a supplement to other evolutionary systems to constitute managerial strategies.

vi. **Fraud detection Scoring**

These scores are used to identify fraudulent activities of the customers. They are mainly used in the credit card industry to introduce control techniques.

vii. **Recovery Scoring**

It ascertains the amount recoverable after charge off. This helps in the formulation of strategies and deployment of resources to strengthen the recovery system.

viii. **Repayment Projection Scoring**

Based on the patterns of payments by customers this score will assist the management in predicting the payments and developing strategies to mitigate losses arising from repayment anomalies.
ix. Response Scoring
These scores study the responsiveness of a customer or a prospective customer towards marketing actions. They are used to identify customers with high responsiveness and high creditworthiness.

x. Revenue Scoring
It rank orders customers based on the quantum of revenue expected to be generated in the first twelve months of dealing with a customer. It can also be used for pricing and rewarding strategies combined with the perceived levels of risk exhibited by the customer.

2.3 CREDIT SCORING MODELS
Despite the history of credit dating back to 2000BC, the story of credit scoring begins from about the beginning of the 19th century and the real universal growth in research activities in the field of credit scoring was not until the end of the 20th century (Abdoua & Pointon, 2011). Hence, the available literature is limited. Although in the recent years, credit scoring has gained importance, it is not reflected in the quantum of research especially in developing countries. The following section deals with the review of research available related to classification of credit scoring models, comparison of tools used in credit scoring and criteria for evaluation of credit scores.

2.3.1 Classification of credit scoring models
In one classification, (Thonabauer & Nosslinger, 2004) credit scoring models were classified as three types namely Heuristic models, Statistical models and Causal models. Heuristic models or expert models are based on the experience of experts, whose understanding and decisions are subjective. They depend on the practical and real hands on experience in dealing with complicated situations. These models do not undertake complex mathematical or statistical techniques for evaluations. Classic rating questionnaire, qualitative systems, expert systems and qualitative systems belong to the category of heuristic models. Statistical models involve the formulations of hypotheses to define the problem. The solvency status and creditworthiness parameters are defined. Statistical tools like discriminant analysis, regression, artificial neural networks, etc are used to examine the data. The hypotheses are tested and results obtained. Lastly, Causal models link financial theories to explain the
behaviour of borrowers when historical data of the borrowers is unknown. These models do not study the creditworthiness of the customers but the capabilities of the borrowers to sustain the current credit in hand. Techniques used in these models include option pricing models and cash flow simulation models.

In another classification, (Li & Zhong, 2012) credit scoring models were classified based on the techniques used in building the models. Some of them are: expert scoring models, statistical models and Artificial Intelligence (AI) methods. While, statistical and AI models use complex and advanced techniques, most of the expert models do not use complex techniques.

2.3.2 Credit Scoring Techniques

Over the years a lot of techniques have been incorporated into statistical models. Some of the techniques are Linear Discriminate Analysis (LDA); Linear Regression; Logistic Regression; Bayesian Model; Decision Tree; Markov Model. AI methods use techniques like, Artificial Neural Networks; Support Vector Machine (SVM); Genetic Algorithm and Genetic Programming; K-Nearest Neighbour; Case-Based Reasoning; non-parametric smoothening; Mathematical Programming; Markov Chain Models, recursive partitioning and conditional independent models. However, researchers suggest that there is no one best technique in credit scoring (Abdoua & Pointon, 2011). The following is a brief description of the literature available on some of the popular techniques used in credit scoring.

a) Discriminate Analysis

Discriminant Analysis was first used as a scoring technique by Durand (1941). The technique proved that it was a good predictor of payment defaults. Discriminant Analysis was critically studied over the years. Eisenbeis (1977) examined the use of discriminant analysis in the field of finance, economics and business. One of the important observations he made was that variables that describe the members of the group, are multivariate normally distributed.

This is a big misconception. Definitely, if the variables follow a multivariate ellipsoidal distribution (of which normal distribution is a special case), then the linear discriminant rule will be optimal. However, if discriminant analysis is assumed as yielding a linear combination of all the variables that maximise the separability
criteria, then it is accepted. The assumption of normality is essential only when significant tests are to be done (Hand & Henley, 1997). Most of the credit information is not normally distributed and it is not a critical limitation (Reichert, Chob, & Wagner, 1983). Other studies on the use of discriminant analysis in credit scoring include Myers & Forgy (1963) (who compared discriminant analysis with regression analysis), Lane (1972), Grablowsky & Talley (1981), Moses & Liao (1987), Yobas, Crook, & Ross (2000), Sarlija, Bensic, & Bohacek, (2004), Falangis (2008), Rangau & Ouertani (2010), Gumparthi & V.Manickavasagam (2010).

Some authors have criticised the tool’s application in credit scoring. Some of the them are assumption of linearity, inappropriateness of prior probabilities, Classification errors, etc (Eisenbeis R., 1978). However, discriminant analysis is the most popular and widely used technique in credit scoring (Abdou, Pointon, & Masry, 2009).

b) Linear Regression Analysis
A linear regression model assumes a linear relationship between the probability of default and the factors that influence the behaviour (Mester, 1997). Orgler first used linear regression to examine the existing loans and review commercial loans. Later, Orgler (1971 and 1972) used regression analysis for the construction of score cards that evaluate outstanding loans. Desai, Crook, & Overstreet, (1996) in their study observed that linear regression models were better classifying bad loans as compared to generic models. Other studies include Fitzpatrick (1976), Henley (1995) and Hand & Jacka (1998).

c) Logistic Regression
Theoretically, it can be supposed that a logistic regression is a better statistical tool than linear regression because two discrete classes of risks (good and bad) can be identified (Hand & Henley, Statistical classification methods in consumer credit scoring: a review, 1997). Except for the choice in a parametric model and other assumptions there is no principle difference in the two methods (Hosmer & Lemeshow, 1989). Henley (1995) argued that logistic regression was no better than linear regression. In his case, the estimated probabilities of good risks were in the range of 0.2 to 0.8 and the logistic curve could as well be approximated to a straight line. One of the first works that talked about the use of logistic regression in the area of credit scoring was by Wiginton (1980). The study involved a comparitive
evaluation of logistic regression and discriminant analysis. He concluded that Logistic regression was a superior method. Srinivasan & Kim (1987) included logistic regression as one of the methods for analysing and granting credit. Steenackers & Goovaerts (1989) used logistic regression as a numerical scoring system that appraised personal loans Jones (1993) applied reject inference to logistic regression in the field of credit scoring. Leonard (1993) used logistic regression to evaluate commercial loans. A study in 1996, compared logistic regression with neural networks and found that even though neural network models were good at classifying good and bad loans, logistic regression model was better in accommodating other important aspects (Desai, Crook, & Overstreet, 1996).

d) **Probit Analysis**
Finney (1952) was the first person to pioneer the use of probit analysis but for ‘toxicology problems’. The term ‘probit’ stood for probability unit (Maddala, 2001). It is a technique that generates coefficient values that are the probability unit values of dichotomous coefficients. Normal distributions for threshold values are assumed under probit analysis (Grablowsky & Talley, 1981).

e) **Decision Tree**
Decision trees also called as recursive partitioning (Hand & Henley, 1997) or CARTs (Classification and Regression Trees) which was pioneered by Breiman, Friedman, Olshen, & Stone (1984) is a classification tool that is applied in developing credit scoring models. It is a non-parametric tool that analyses dependednt and categorical variables as a function of continuous variables (Breiman, Friedman, Olshen, & Stone, 1984), (Arminger, Enache, & Bonne, 1997). Its first application in credit management was probably by Rosenberg E & Gleit A (1994). However, the first credit scoring model was built by David Sparks (1972) at University of Richmond. There have been different applications of decision tree in credit scoring by many researchers (Lee, Chiu, Chou, & Lu, 2006), (Bensic, Sarlija, & Zekic-Susac, 2005), (Basens, 2003).

f) **Neural network**
Neural network is used in the area of business for various purposes. In the field of finance it is used for bankruptcy classification and detection of fraud (Smith & Gupta, 2000). Credit scoring is an important area of application of neural networks
Neural network model that is usually used for credit scoring is a statistical model that makes use of linear combinations of nested sequences of non-linear transformations of linear combination Variables (Hand & Henley, Statistical classification methods in consumer credit scoring: a review, 1997). Like other models, neural network model has the same amount of flexibility (Ripley, 1994). Rosenberg & Gleit (1994), provided a host of applications of neural networks in credit decisions and detection of frauds. Mittal, Gupta, & K. Jain, (2011) Developed a non-parametric credit scoring model using the neural network technique for the evaluation of credit disbursements to micro enterprises in India. Neural network has a better edge over other statistical models because it can be applied when the analytic relationship between dependent and independent variables is unknown (Yu, Wang, Lai, & Zhou, 2008). Many studies have compared neural networks with other scoring techniques. For example, in a study that compared discriminant analysis, logistic regression and neural network (Islam, Zhou, & Li, 2009), it was observed that neural network model had the highest success rate.

g) Experts Systems
As the name suggests it involves the use of human experts’ knowledge, their interpretations to solve problems (Rosenberg & Gleit, 1994). There are limited studies in this context. The benefit of an experts system is that the expert is in a position to give an explanation as to why the applicant is rejected or accepted (Hand & Henley, 1997). Another benefit of experts system is that the evaluator can bring in new characteristics into the underwriting process (Kumra, Stein, & Assersohn, 2006). A research by Ben-David and Frank (2009), compared machine-learning model models with expert systems models, it was observed that while some of the former were accurate in predicting capabilities, most of them failed (Ben-David & Frank, 2009).

h) Genetic Programming
It is a technique that has recently been put to use in credit scoring and is considered as an extension of genetic algorithm (Koza, 1992). Usually genetic programming generates output programmes that are the LISP language (Koza, 1994), (Núñez-Letamendia, 2002). The use of genetic programming is growing rapidly especially in the areas of bankruptcy predictions (Mckee & Lensberg, 2002), (Etemadi, Rostamy, & Dehkordi, 2009), financial returns (Xia, Liu, Wang, & Lia, 2000), classification
applications (Zhang & Bhattacharyya, 2004), (Ong, Huang, & Tzeng, 2005), (Lensberg, Eilifsen, & McKee, 2006) and scoring (Huang, Chen, & Wang, 2007).

2.3.3 Comparison of Credit Scoring Tools

A study by Guillen and Artis (1992), compares the efficiency of a few credit scoring techniques in correctly classifying the borrowers as good and bad customers.

<table>
<thead>
<tr>
<th>Model</th>
<th>Total correct classification (%)</th>
<th>Correct classification of good (%)</th>
<th>Correct classification of bad (%)</th>
<th>Bad accepted into the good group (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discriminant Analysis</td>
<td>65.4</td>
<td>62.2</td>
<td>78</td>
<td>8.1</td>
</tr>
<tr>
<td>Linear Regression Model</td>
<td>55.1</td>
<td>47</td>
<td>87.5</td>
<td>6.2</td>
</tr>
<tr>
<td>Probit Model</td>
<td>71.9</td>
<td>76.4</td>
<td>54.1</td>
<td>13.1</td>
</tr>
<tr>
<td>Poisson Model</td>
<td>62.4</td>
<td>57.7</td>
<td>81.8</td>
<td>7.3</td>
</tr>
<tr>
<td>Negative Binomial II Model</td>
<td>63.3</td>
<td>58.9</td>
<td>80.6</td>
<td>7.6</td>
</tr>
<tr>
<td>Two-Step Procedure</td>
<td>64.9</td>
<td>61.1</td>
<td>79.8</td>
<td>7.6</td>
</tr>
</tbody>
</table>

*Source: Guillen and Artis (1992), modified*

It can be observed (from Table 2.1) that in spite of probit model achieving the highest total correct classification (71.9%); it is the worst performer in correctly classifying bad customers (54.1%) (Type II error). On the other hand, linear regression is the least in accepting bad customers as good (6.2%) (Type I error), but overall is the worst in correctly classifying customers (55.1%).

Abdou and Pointon (2009), compared traditional scoring techniques with two types of neural network techniques on the basis of ACC (Average Correct Classification) rate. The ACC is derived as the percent of total customers correctly classified as good or bad. It is a measure of efficiency of a credit scoring model.
Table 2.2: Comparison of ACC Rates of different Credit scoring models

<table>
<thead>
<tr>
<th>Scoring Techniques</th>
<th>ACC rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sample1</td>
</tr>
<tr>
<td>Multiple Discriminant Analysis</td>
<td>78.05</td>
</tr>
<tr>
<td>Logistic Regression</td>
<td>82.33</td>
</tr>
<tr>
<td>Probabilistic Neural Network</td>
<td>85.9</td>
</tr>
<tr>
<td>Multilayer Feed-Forward Neural Network</td>
<td>84.07</td>
</tr>
</tbody>
</table>

Sample1 denotes 67/33\% for the training/testing sets; Sample2 denotes 80/20\% for the training/testing sets; Sample3 denotes 90/10\% training/testing sets.

*Source: Abdou and Pointon (2009), modified.*

From the above table it can be seen that the ACC rate was the highest for sample 1 and sample 3 when probabilistic neural network was deployed and highest for sample 2 when multilayer feed-forward neural network was deployed. On the other hand, in all cases multiple discriminant analysis generated the least ACC rate. Thus, it can be deduced that neural network models have a greater accuracy in correctly classifying that traditional models like discriminant analysis and logistic regression.

West (2000) classified ten different scoring models as inferior and superior models based on his study on two different sets of German and Australian data. For the study he used three conventional models (Logistic regression, linear discriminant analysis and CART) and the rest neural network models. The results of his study are tabulated in Table 2.3.

It can be observed that of the conventional models, logistic regression and linear discriminant analysis are best suited for scoring applications and among unconventional models MOE, RBF and MLP models are superior models than the rest.
Table 2.3: Comparison of scoring models and credit data

<table>
<thead>
<tr>
<th>Superior models</th>
<th>German credit</th>
<th>Australian credit</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOE</td>
<td>MOE</td>
<td></td>
</tr>
<tr>
<td>RBF</td>
<td>RBF</td>
<td></td>
</tr>
<tr>
<td>MLP</td>
<td>MLP</td>
<td></td>
</tr>
<tr>
<td>Logistic regression</td>
<td>Logistic regression</td>
<td></td>
</tr>
<tr>
<td>LVQ</td>
<td>LDA</td>
<td></td>
</tr>
<tr>
<td>FAR</td>
<td>K-NN</td>
<td></td>
</tr>
<tr>
<td>LDA</td>
<td>LVQ</td>
<td></td>
</tr>
<tr>
<td>K-NN</td>
<td>FAR</td>
<td></td>
</tr>
<tr>
<td>Kernel density</td>
<td>Kernel density</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Inferior models</th>
<th>German credit</th>
<th>Australian credit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

MOE: mixture of experts; RBF: radial basis function; MLP: multilayer perceptron; LVQ: learning vector quantization; FAR: fuzzy adaptive resonance; all of which are neural networks models. LDA: linear discriminant analysis. K-NN: K-Near Neighbour (Algorithm)

Source: West (2000), modified

Abdou H. (2009), in his study on Egyptian public sector banks, used genetic programming for the appraisal of borrowers. He compared the results of two genetic programming scoring models with traditional models; probit analysis and Weight-of-evidence model.

Table 2.4: Comparison of classification results of different scoring models

<table>
<thead>
<tr>
<th>Scoring model</th>
<th>Correctly classified results (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Testing</td>
</tr>
<tr>
<td>Weight-of-evidence model</td>
<td>52.16</td>
</tr>
<tr>
<td>Probit analysis</td>
<td>82.69</td>
</tr>
<tr>
<td>Genetic programming – best program (GPp)</td>
<td>82.93</td>
</tr>
<tr>
<td>Genetic programming – best team (GPt)</td>
<td>83.89</td>
</tr>
</tbody>
</table>

Source: Abdou (2009), modified

It is clear from the above table that the two genetic programming techniques provide better results for both testing sample and the overall sample, as compared to traditional techniques like probit analysis and Weight-of-evidence model.
A comparative study across results from various researchers was conducted by Crook, Edelman, & Thomas (2007), to analyse and determine from different studies as to which was the best scoring model in correctly classifying borrowers, as defaulters and non-defaulters. Table 2.5 is a tabulated result of their analysis.

Table 2.5: A comparison of percentage of correct classification from different studies

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear Regression or LDA</td>
<td>77.5</td>
<td>66.5</td>
<td>79.3</td>
<td>71.4</td>
<td>69.3</td>
<td>79.3</td>
<td>80.8</td>
</tr>
<tr>
<td>Logistic Regression</td>
<td>-</td>
<td>67.3</td>
<td>81.8</td>
<td>73.5</td>
<td>-</td>
<td>79.3</td>
<td>-</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>75</td>
<td>-</td>
<td>77</td>
<td>-</td>
<td>-</td>
<td>77</td>
<td>78.4</td>
</tr>
<tr>
<td>Math Programming</td>
<td>74.7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>79</td>
<td>-</td>
</tr>
<tr>
<td>Neural Nets</td>
<td>-</td>
<td>66.4</td>
<td>82.6</td>
<td>73.7 (77.0)(d)</td>
<td>72</td>
<td>79.4</td>
<td>81.7</td>
</tr>
<tr>
<td>Genetic Programming</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>82.8</td>
</tr>
<tr>
<td>K-NN</td>
<td>-</td>
<td>-</td>
<td>76.7</td>
<td>-</td>
<td>-</td>
<td>78.2</td>
<td>-</td>
</tr>
<tr>
<td>Support Vector Machines</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>79.7</td>
<td>-</td>
</tr>
</tbody>
</table>

a. Figures are an average across two data-sets.
b. Figures are an average of eight data-sets.
c. Figures are an average over two data-sets.
d. Hybrid LDA and NN.

Source: Source: Crook et al. (2007), modified

The above table does not allow for comparisons along rows but only along columns. It was concluded that there is no best model and the selection of a credit scoring model depends on various aspects like type of problem studied, data size and structure, the variables employed, the markets to which they are applied to and the cut-off points used.

2.3.4 Criteria for Evaluating Credit Scores

While credit scoring models try to evaluate the creditworthiness of borrowers, there needs to be a criterion on which the efficiency of the model incorrectly classifying borrowers as defaulters and non-defaulters is evaluated. There are numerous criteria
used for evaluation some of them are- confusion matrix, estimated misclassification cost criterion, ROC (Receiver Operating Characteristics) curve, MSE (Mean-Square Error), RMSE (Root Mean-Square Error), MOE (Mean of Error), GINI coefficient, Goodness-of-fit test (calibration) and discrimination (C-statistic/AUC).The following is a review of literature on some of the most popular criteria used by researchers to measure the efficiency of credit scoring models.

a) **Confusion Matrix**
Confusion matrix also called as the ACC (Average Correct Classification) rate is the percentage of total cases correctly classified by the scoring model as good and bad (Abdoua & Pointon, 2011). It is the most widely used criterion in fields of study like marketing, health and especially finance (for credit scoring) (Paliwal & Kumar, 2009). It is occasionally called as ‘Classification Matrix’ (Abdou H., 2009). A classification matrix is a simple presentation of cases correctly classified as good, cases correctly classified as bad, cases wrongly classified as good and cases wrongly classified as bad. Some researchers have compared the confusion matrix with other criteria like MSE and RMSE (Fletcher & Goss, 1993), (Kumar & Rao, 1995). One researcher drew a comparison between the confusion matrix and Mahalanobis distance and Kolmogorov-Smirrov (Yang, Wang, Bai, & Zhang, 2004).

b) **Estimated misclassification cost criterion**
It is a comparison of relative costs incurred when unworthy applicants are accepted versus when worthy applicants are rejected. This criterion is crucial in evaluating the minimum cost that would be incurred because of misclassification of cases when selecting a credit scoring model (Abdou & Pointon, 2011). The cost criterion is represented as a ratio of misclassification costs of Type II error and Type I error. Usually the costs of Type II error are much higher than Type I error (Lee & Chen, 2005).

c) **ROC Curve**
It is a two dimensional representation of the proportion of bad cases classified correctly (called sensitivity) on vertical axis against proportion of good cases classified as bad on horizontal axis at different cut-off scores. It is also called as the ‘Lorentz diagram’ (Abdou & Pointon, 2011). The ROC curve is an illustration of the behaviour of classifiers without any regard to misclassification costs (Thomas, Crook,
& Edelman, 2002), (Yang, Wang, Bai, & Zhang, 2004). ROC although has been in use in fields like health, engineering, medicine, etc, majority of it use has been in the field of accounting, finance and banking (Basens, 2003), (Blochlinger & Leippold, 2006).

2.4 APPLICATION OF CREDIT SCORING

‘Credit scoring’ is a term that the world has not known for long. Although credit scoring began in 1941, the use of credit scoring was limited to appraisal of customers until the turn of the century. In the short period of existence the uses of credit scoring has expanding to different fields. The following section deals with the review of studies regarding the application of credit scoring in different fields.

The first application of credit scoring was by Durand (1941), who recognised the need for a numerical value that could quantify the creditworthiness of the applicant and assist the lender in the underwriting process. With increase in pressure on the banking sector to generate more credit facilities, the prominence of the tool also increased. Wider range of products has forced lenders to customise their models. There has been a tremendous growth in the use of credit scoring in the banking sector (Henley & Hand, 1996)(Abdou H., 2009b) especially in the last couple of decades (Foglia, Laviola, & Reedtz, 1998), (Emel, Oral, Reisman, & Yolalan, 2003), (Thanh Dinh & Kleimeier, 2007). In the banking sector, credit scoring has been in use for different purposes. The most widely used purpose is the appraisal of consumer credit applicants (Orgler Y. E., 1971), (Malhotra & Malhotra, 2003), (Kim & Sohn, 2004), (Lee & Chen, 2005), (Sustersic, Mramor, & Zupan, 2009), (Einav, Jenkins, & Levin, 2013), credit card based scoring ((Greene, 1998), (Banasik, Crook, & Thomas, 2001), (Quah & Sriganesh, 2008), mortgage applicants, (Cameron & Trivedi, 1996), (Heuson, Passmore, & Sparks, 2001), (Somers & Whittaker, 2007), (Haughwout, Peach, & Tracy, 2008), evaluating new customers (Steenackers & Goovaerts, 1989), (Hsieh, 2004), (Sarlija, Bensic, & Zekic-Susac, 2009) and small business customers (See 2.6).

In fields of finance and accounting, the scope of credit scoring has widened. Some of the areas in which credit scoring has been used are, financial distress (Hu & Ansell, 2007), (Hu Y., 2008), (Mukkamala, Vicera, & Sung, 2008), financial decisions and returns (Xia, Liu, Wang, & Lia, 2000), (Yu, Wang, Lai, & Zhou, 2008), bankruptcy classification (Mckee & Lensberg, 2002), (Lensberg, Eilifsen, & McKee, 2006),
There has also been application of credit scoring in fields like marketing, engineering, manufacturing, health and medicine.

2.5 CREDIT SCORING AND SMALL BUSINESSES

This section deals with the review of studies undertaken that are in relation to the use of credit scoring for appraising small business customers.

Historically, the use of credit scoring was limited to appraisal of home loans, mortgages, and consumer credit. However, lending to small businesses was considered a risky game by most lenders. Banks preferred to lend to large businesses that processed systematic future plans with large valued collaterals, accounted financials and proven credit history. Small businesses on the other hand were finding it very difficult to impress the lenders. In recent years however, lenders find it very difficult to grow in the existing markets. Formal lenders are looking for opportunities to expand their customer base. Small businesses whose financial needs were mainly met by local money lenders are now being targeted by bank and other formal lenders. Therefore, looking out for creditworthy borrowers is one of the important priorities of the lender. Credit scoring has proved to be an effective tool in evaluating the credit risk associated and reduce the cost of lending to small business borrowers (Feldman, 1997). Credit scoring helps to reduce screening and monitoring costs of small business loans, enhance competition among banks, and adjust interest rates in view of the borrowers’ credit risk and to increase the availability of credit for risky and marginal firms. (Ono, 2006). The approval rate for small businesses tends to be higher for banks that use credit scoring when compared with banks that do not use them (Aduda, Magutu, & Wangu, 2012). Banks tend to be willing to lend to a customer on an unsecured basis when the credit scoring rate is high and reliable (Chakrabarty, 2013). Credit Scoring improves the psychological, social and financial well-being of the borrowers (Copisarow, 2000). A study on evaluating the effects of implementing a scoring model in appraising small businesses showed that there was at least a 50 percent reduction in underwriting costs (Kumar & Motwani, 1999). A study in
Atlanta, showed that implementation of credit scoring for small businesses reduced the effect of problems arising from asymmetric information from borrowers especially in low and moderate-income areas (Frame, Srinivasan, & Woosley, 2001). Another study by DeYoung et al., (2008) suggested that automated processes in lending helped lenders seek new and previously unserved markets. Carter & McNulty (2005) claim that credit scoring plays an important role in expanding credit facilities to small business borrowers.

Tsaih & Lien (2004) in their research work developed a credit scoring model that could be easily altered to meet the changing environmental needs. A study was conducted by Bensic, Sarlija & Zekic-Susac (2005) on a sample data of small businesses borrowers, different statistical tools like logistic regression, neural networks and decision tree were used in building credit scoring models. The study suggested that neural -networks was the best model in rightly evaluating the sample borrowers. Leonard (1992), in his study suggests the use of SBCS (Small Business Credit Scoring) help in the increase of credit availability, reduction in credit risk and in lowering the prices.

2.6 DETERMINANTS OF REPAYMENT BEHAVIOUR

This is the final section under review of literature. It pertains to review of literature dealing with determinants of repayment behaviour. It encompasses determinants related to the personal aspects of the borrower, business related aspects and bank related aspects.

Credit scores are numeric expressions of the worthiness of credit of a customer. The dependability on a credit scoring model is directly proportionate to the ability of the score to reflect the past behaviour or actions of the customer. Loan characteristics play an important role in determining the loan repayments (A.H. Roslan & Karim, 2009); (Njoku & Odii, 1991). Picking the right determinants is very important in credit scoring. Rouse, (1991) believed that credit scores that predominantly use the borrowers’ characteristics might discriminate against people living in poor conditions and may never clear a screening test. Mondar et al. (2002), however suggests that credit scoring models intend to reduce subjectivity and are localized in implementation and can be flexible. Although factors that determine the personal
borrowing are very different from factors that determine commercial borrowing, According to Basel norms, lending to small businesses should be considered as retail lending and hence most of the factors tend to remain similar for personal and small business lending. However, certain scholars point out that both the characteristics of the borrower and the business are important and their due share must be given (Derban, Binner, & Mullineux, 2005). Studies also showed that SMEs are better understood for evaluation when classified under retail borrowers (Dietsch & Petey, 2004). In a study in Piura, it was observed that informal lending performed better than formal lending because informal lenders have a better knowledge of the borrower personally and the practice of small lending initially with gradual increase over the good performance of the borrower (Guirkinger, 2008). Therefore to compete with a large informal sector the formal sector must be robust in their processes and proactive in their approach.

A detailed review of literature shows that there is little research in studying the factors that determine loan repayment and even sparser when it comes to studies in developing countries or low income countries (Derban, Binner, & Mullineux, 2005). The following is a review of available literature on determinants of repayment.

2.6.1 Personal Aspects of the Borrower

a) Age
Studies show that age is an important and significant variable. For example, in a study in Malaysia, the older borrowers were more responsible and disciplined (above 25 years) and younger borrowers were of the opinion that they could get credit from other sources at a later time since they were young (Mokhtar, Nartea, & Gan, 2012). In a study in Kenya, the results showed that the chances of younger customers defaulting were higher than older customers (Angaine & Waari, 2014). Similarly in Ethiopia, Kenya and Dekehare Sub-Zone age was a significant factor (Pasha & Negese, 2014); (Mulili, 2013); (Asgedom & Muturi, 2014). Other studies also point out that age is an important factor (Wongnaa & Awunyo-Vitor, 2013); (Kohansal & Mansoori, 2009);

Certain studies give an opposite result where age has no significant impact (Chong, Morni, & Suhaimi, 2010).
b) Gender

Women are the centre of sustainable development, therefore ensuring gender equality in all sectors would help the society as a whole (Denton, 2002). According to the latest data base on financial inclusion, women relatively have less chances of having a bank account as compared to men. In developing countries, women are seventeen percent less likely to have taken a loan and 20 percent less likely to have a bank account (Asli, Leora, Dorothe, & Peter, 2015). Women are considered poorer than men on an average (Burjorjee, Deshpande, & n, 2002). The availability of collateral to finance their needs is also less.

Few studies show that gender has no influence on the repayment behaviour. Studies in Bangladesh showed that although female gender had a positive influence on repayment behaviour, it wasn’t very significant (Sharma & Zeller, 1997); (Godquin, 2004). Other researchers in the US, Ethiopia and Malaysia emphasized that gender was not a significant in determining repayment (Bhatt & Tang, 2002); (Brehanu & Fufa, 2008); (Chong, Morni, & Suhaime, 2010).

On the other hand there are many studies to prove that women have a higher tendency to outperform in repayment. For instance, in a book the author points out that the Grameen bank in Bangladesh initially had both men and women as customers, but later decided to focus exclusively on women due to repayment problems from men (Armendáriz & Morduch, 2005). Muhammad Yunus in his study claims that women have a better long term vision and are better in managing resources and credit (Yunas, 2004). In Malawi it was studied that 92 percent of women paid on time as compared to only 83 percent of men (Hulme, 1991) and in Malaysia, women were 95 percent and men were 72 percent likely to repay their loans (Gibbons & Kasim, 1991). Even in group lending women are at the forefront on repayment (Kevane & Wydick, 2001).

c) Marital Status

Marital Status is common yet significant information that has many implications. On one hand, it is an implication of maturity, responsibility and reliability (Baklouti, 2013). More often than not single borrowers tend to be less responsible than married borrowers and hence face a higher probability of default (Dunn & Kim, 1999). On the other hand, it can be assumed that the probability of default is higher for married borrowers rather than single borrowers. Married people are financially stressed due to
a number of dependents (Dinh & Kleimeier, 2007). Therefore, the relationship between the marital status and the loan repayment behaviour is uncertain and unpredictable.

d) **Number of Dependents**
The size of the family (number of dependents) determines the number of mouths to feed. This information is required by the bankers to determine the amount of funds that will be available for loan repayment. Results of studies that studied the significance of family size are mixed. Some find that it is insignificant. For example, a report published by the ‘World Bank’ states that the size of the family was statistically insignificant in determining defaults (Wakuloba, 2007).

On the contrary, certain studies give significant importance to the size of the family. Another research study also pointed out that default risk increases with the size of the family (Bandyopadhyay, 2007) The probability of default increases with the increase in the number of dependents (Pollio & Obuobie, 2010); (Kitawa & Terye, 2015).

e) **Education Level**
The education levels of customers have always proved to be worthy determinants of repayment behaviour. The simple equation is that higher the level of education higher is the income generating capacity, and thus the ability to repay the loans (Dufhues, Buchenrieder, Quoc, & Munkung, 2011). Bad loans are triggered by lack of education (Asantey & Tengey, 2014). Studies have shown that training, basic education and health services improve the repayment performance (Khandker, Khalily, & Khan, 1995); (Matin, 1997); (Godquin, 2004). On the other hand borrowers that had no little training or education had a higher probability of default (Haim, Abidin, Noor, & Majid, 2007). Another study revealed that of the socio-economic factors the level of education was a major factor in influencing the repayment behaviour (Oladeebo & Oladeebo, 2008). Researchers have hypothesized that women as compared with men show a stronger preference for education. Therefore access to credit by women can empower to influence household decisions (Thomas D., 1990); (Behraman & Rosenzweig, 2002).
2.6.2 Business related Aspects

f) Over-indebtedness

Over-indebtedness is defined as the inability of the borrower to repay the entire debt on time (Haas, 2006). According to the latest Banana skin report (2014), Over-Indebtedness was found to be the biggest salient risk for the micro-finance industry. Over-indebtedness can be considered in three different situations. First, the borrower is unwilling to repay the debt despite having the capability to do so and finally defaults on the loan. The second is a situation where the borrower has to undertake exceptional costs and extraordinary actions to fulfil the commitment on repayment of the debt. These are unanticipated at the time of contract of the debt. The third situation is when the borrower is willing to repay but is not equipped with the ability to do so and finally makes a part repayment or a complete default (Gonzalez, 2008).

Over-indebtedness is a regular practice in rural areas. It is used by households as a tool to smoothen out their cash flows (Collins, Morduch, Rutherford, & Ruthven, 2009). Short term credit from informal sources is leveraged to repay debt from formal sources from time to time (Maldonado & González-Vega, 2008). Borrowers are many times overoptimistic and find it difficult to measure over-indebtedness and hence resort to information asymmetry and multiple borrowings. Bankers many times find it easy to pass on the risk of over-indebtedness to the depositors and hence recklessly lend to overly indebted borrowers (Chaves & vega, 1994). Saturated markets are breeding grounds for over-indebtedness. Lender who are either unaware of multiple borrowings or do not take necessary steps to identify over-borrowing are the biggest victims of over-indebtedness (Schicks & Rosenberg, 2011).

There have been many studies in the field of indebtedness. In a study of 900 borrowers in Karnataka, it was found that over-indebtedness was high in mass default towns and low in Non-mass default towns (Krishnaswamy, 2011). In another study to identify the percentage of borrowers inclined to over borrow in Bolivia, it was found that 85% of respondents had at least once in the past four years been overly indebted (Gonzalez, 2008). In a study in Tamil Nadu, more than 20% of the respondents have been regularly resorting to over borrowing and the reasons simply being unable to meet both the ends (Guérin, Roesch, Venkatasuresh, & Kumar, 2011). In a study in Bangladesh, it was found that a major cause of over indebtedness was
unforeseen financial problems. The study points out that there is an increase in the number of cases every year (Chaudhury & Matin, 2002). One study had a totally contrary view from the above studies. The study concluded that multiple borrowings can be associated with better repayment rates (Krishnaswamy, 2007)

\textbf{g) Credit History}

The incorporation of credit history reduces the influence of cultural and racial biases that creep into the appraisal process (Cotterman, 2002). Studies in Africa showed that a deeper evaluation of the credit history of the borrowers increased the acceptance levels for extending credit without any degradation in the level of defaults (Karlan & Zinman, 2010), (Viganò, 1993).

\textbf{h) Collateral}

Collateral based lending was first introduced by British landlords in the seventeenth century to secure the money they lent to borrowers. Land was used as collateral and the borrower was allowed to use the land until a default occurred (Caouette, Altman, & Narayanan, 1998). Collateral based lending is useful in lending to firms that are young, small and have little or no credit history (Menkhoff, Neuberger, & Suwanaporn, 2006). Also collateral reduces the risk of strategic default in situation where cash flows can be diverted (Bond & Rai, 2002). The requirement of collateral reduces with the increase in the length of the relationship between the borrower and lender (Voordeckers & Steijvers, 2006). But, long credit relationships do not reduce the collateral requirement by reducing information asymmetry (Menkhoff, Neuberger, & Suwanaporn, 2006). In one study, distance of the borrower from the lender is a determinant of collateral requirements. The use of collaterals for local borrowers is higher as compared to distant borrowers (Jiménez, Salas, & Saurina, 2009).

Although, collateral based lending is considered safe and a hedge against credit risk, it poses its own risk called collateral risk. Sometimes, the use of collateral can be costly for both lenders and borrowers (Berger, Saunders, Scalise, & Udell, 1998). Some of the problems that arise are illiquidity (inability to realize money without loss of value and time), asset legality and asset quality (Caouette, Altman, & Narayanan, 1998). Certain studies show very unusual results. For example, one study showed that a bank had low levels of non-performing loans despite of high share of uncollateralized loans.
(Becchetti & Conzo, 2011). Another study does not support the idea of collateral security improving the credit behaviour of customers (Schuh & Stavins, 2015).

**i) Personal Guarantee**

A financial Guarantee given by a third party can partially cover for the lack of collateral (Chen Y., 2006). A study in USA showed that as part of financial innovation, bankers were requesting borrowers to use guarantees as a safe security when collateral was unavailable (Canesi, Alpaos, & Marella, 2016). A Japanese study on SME borrowers revealed that personal guarantee is effective in creditworthiness of the borrower (Ono & Uesugu, 2009). Another research on Italian SMEs also derived that same conclusion (Zecchini & Ventura, 2009). Personal Guarantees can be also be used as incentives against moral hazard problems and reduce ex-ante credit risk (Pozzolo, 2004).

**j) Business Experience**

Experience of the customer in the area of business is very important in the determination of loan repayment. Bankers can resort to experience of the customers for information and to seek lesser collateral. Businesses are most likely to fail in the first year of operation and hence the risk of lending to businesses with less experience is very high (Nawai & Shariff, 2013). Once a business has survived for the first few years, there is a significant reduction in the chances of failure (Everett & Watson, 1998).

Experience also includes the number of years the customer has engaged in the activity of business as an entrepreneur or even as an employee. There is a positive and a significant relationship between the work experience and repayment rate (Baklouti, 2013). Training also plays an important role in the performance of the business and the repayment behaviour. Roslan and Zaini (2009) stated that borrowers, who did not have any training of their business, had higher probability of default.

**k) Loan Size**

The size of the loan is yet another important factor in determining the repayment of loan. Studies have pointed out at two different views. Pham & Lensink (2008) in their study in Vietnam say that larger the size of the loan, larger is the probability of default. Also in another study it was found that the incidence of loan repayment decreases and leads to high default rates (Hietalahti & Linden, 2006). Some studies
have an opposite view. For instance, Oladeebo & Oladeebo (2008) point out larger loans will facilitate larger investments, leading to increased use of technology and finally increased output and money to repay loan. Ndanitsa et al., (2011) suggested that the problems of default decreases with increased loans size. In a Study by Martin (1997) the size of the loan did not have any impact.

1) Distance from the bank
Availability of Information of small businesses has been an eternal challenge for the banks. Lending to small businesses has historically been a costly affair due to the extra efforts that are required to obtain relevant information (Berger & Udell, 1995). It was crucial that to poses information, proximity and the intensity of relationship with the customer. This was important to improvise the precision in credit assessment (Petersen & Ragan, 2002). Borrower proximity allows for collection of soft information, which could lead to trade-offs in interest rates and the availability of credit (Agarwal & Hauswald, 2010). The credit decisions to be made must be close to where the information was gathered. This would be a costly operation and would be uneconomical. Hence, such operations cannot be carried out and would imply a reduction in the credit provided to small businesses. (Berger, Saunders, Scalise, & Udell, 1998);(Strahan & Weston, 1998).

There is an inverse relation between the distance from the client and the lender. A study revealed that for every one kilometre increase in the distance from the bank the probability of repayment would decrease by 0.92 percent (Oke & Agbonlahor, 2007). Also in group lending, the physical distance among the members of the group is an important factor. A study show that the greater the fraction of group members within a distance covered in a ten minutes’ walk, the lesser is the chance of default (Karlan & Goldberg, 2011). There is an increased chance of encountering repayment problems when the increase in the distance between the lender and the borrower (Wydick, 1999).
2.6.3 Bank related Aspects

m) Repayment Incentives
Customers must be efficiently managed through actions that encourage them to repay their loans and discourage them from delayed repayments or defaults. Dynamic incentives are used as mechanisms to get rid of moral hazard. Dynamic incentives make it costly for the customer to default and encourage them to repay as early as possible (Gine, Jakiela, Karlan, & Morduch, 2006). For example, a delinquent customer is eliminated from further rounds of lending. Incentives for frequent repayment schedules act as signals to indicate the progress of borrowers’ projects. Hence a dynamic incentive is a successful enforcement mechanism for loan monitoring (Godquin, 2004). Repeated loans as a part of dynamic incentives helps reduce information asymmetry and loan spreads (Petersen & Rajan, 1994); (Berger & Udell, 1995) and (Boot, 2000). However there is a study to suggest that repeated loans were not significant in effecting repayment behaviour (Ndanitsa, Musa, & Umar, 2011).

n) Flexible repayment Schedules
Flexing of the repayment schedule is a trade-off benefit against default risk. In a study in West Bengal, when the repayment timings were changed from weekly to monthly, it was observed that clients faced 45 percent lesser stress financially, 54 percent less likely to lack confidence in repayment and 60 percent less likely to spend time thinking about repayment. (Field, Pande, Park, & Papp, 2010). Businesses with uneven cash flows are benefited from flexible repayments. It allows for redistribution of repayments during low income periods to high income periods (Weber, Musshoff, & Petrick, 2014). In a study on group lending, clients were allowed for bi-weekly repayment; the results indicated group dropouts’ reduction and increased repayment performance (McIntosh, 2008). Flexibility also reduces the chances of multiple borrowings and increases access to credit facilities (Hamp & Laureti, 2011). Flexible financial repayments enable borrowers to receive higher loan and hence increased income for borrowers and increased business for lenders (Mamun, Wahab, Malarvizhi, & Mariapun, 2011).
Some studies state that flexible repayment scheduling has little or no effect on the repayment behaviour. In a study on random set of urban borrowers, the repayment frequency has no effect on delinquency (Field & Pande, 2008).

\textit{o) Other Determinants}

Length of the relationship with the bank is said to another important aspect that is taken into consideration before lending to a borrower. A borrower is more likely to get a line of credit if he/she has had a long relation with the lender (Cole, 1998), (Baas & Schrooten, 2006).

In a Study in South Africa, it was found that a major contribution to loan defaults was the perception of the customers that micro loans were government’s free grants and that they need not repay (Makina & Malobola, 2004). Awoke (2004), in study it was found that defaults mainly arose from poorly managed procedures and loan diversification. Customers who had other sources of income have better chances of repayment (Wongnaa & Awunyo-Vitor, 2013). Copisarrow, (2000) found that the problems of defaults can be attributed more to the poor implementation and inefficient program designing rather than problems arising from the borrower. Hulme & Mosley (1996) in their popular book “Finance against Poverty”, strongly argue that design features of loans like screening processes, incentives to repay, etc are important factors that contribute to the repayment performance. Diversion of loans for unproductive purposes and for consumption purposes at times of adversity was an important cause of default (Swain & Swain, 2007).

Spouse plays an important role, especially in owner-owned businesses. Businesses that involve spouse in the business activity tend to do better (Lewis, 2011).

Often situational circumstances are more important that a statistical score, failure to take into account a situation before determining credit risk could lead rejection of extending lines of credit which could lead to eventually losing worthy customers (Avery, Calem, & Canner, 2004).

The Ambit of available literature on the determinants of repayment is ambiguous and uncertain. Different studies give importance to different factors. There is no consensus on what factors do and what factors do not influence the repayment behaviour, therefore a few bankers from local banks were interviewed (informal) in-
depth to get an insight into other aspects that are considered in the process of credit appraisal. Some of them were, number of children, concern for children’s education, other sources of income and nature of business.

2.7 RESEARCH GAP

Based on the studied literature in the field of small business lending, credit risk management and repayment behaviour, the following research gaps were identified:

- Research pertaining to small businesses is a niche area that gets little attention in Indian academics.
- Studies that relate to small businesses try to address issues like impacts, efficiency, challenges and other empirics, but there have been few attempts to predict the repayment behaviour of borrowers in India.
- Studies pertaining to credit scoring for small businesses are far less than other areas of banking like consumer credit, mortgage loans, credit card, etc.
- The role of credit scoring in screening borrowers including small businesses is under played in India. Academic Research pertaining to credit scoring methodology can be seldom found.