Chapter 18
Banking Credit Scoring Assessment Using Predictive K–Nearest Neighbour (PKNN) Classifier

Saroj Kanta Jena
BML Munjal University, India

Anil Kumar
BML Munjal University, India

Maheshwar Dwivedy
BML Munjal University, India

ABSTRACT

Credit scoring models is a scientific methodology adopted by credit providers to assess the credit worthiness of applicants. The primary objective of such models has been to predict the potentiality of the loan applicant. A proper evaluation of the credit can help the service provider to determine whether to grant or to reject credit. Therefore, the objective of the study is to predict banking credit scoring assessment using Predictive K–Nearest Neighbour (PKNN) classifier. For the purpose of analysis two different credit approval datasets: Australian credit and German credit have been used. The results from the study show that the proposed model used for classification works better on credit dataset. Here, the study firstly attempted to find the optimal ‘K’ value of the neighbourhood so that the classifier is tuned to forecast the credit worthiness and secondly, validated our proposed model on two credit approval datasets by checking the performance of the proposed models on the basis of classification accuracy.

DOI: 10.4018/978-1-5225-0997-4.ch018
Banking Credit Scoring Assessment Using Predictive K-Nearest Neighbour (PKNN) Classifier

1. INTRODUCTION

The business of financial institutions depends largely on the growth in the demand for credit and other financial services (Kim, 2004). The business community today pays more attention towards credit scoring assessment methods to better understand their customers and their purchasing capacity and the academic world is also interested in indulging in newer methods to develop and evaluate credit risk models. In this scenario, credit scoring plays an important role towards comprehending if it is feasible or not to grant loan to the applicant. Typically, all such scoring models are designed to predict the probability whether a loan applicant is a good credit or a bad credit through a systematic method to evaluate and assess credit. The primary objective remains to quantify and manage the financial risks associated with credit decisions. The most important takeaway of such an iterative process enables financial firms to increase the processing speed of loan applications. Financial institutions through such assessment models can take advantage through the systematic automation of the lending. The credit limit of the customer can also be determined by considering their credit scores.

The key focus amongst academicians lately has been to utilize credit scoring models so as to prescribe the right customer segment of a product for the marketer. Reduction of the cost of analysing the credit in order to enable the manager to take faster credit decisions in return ensures higher credit collections which is the key to better performance. Therefore, financial institutions are using different classification models to find the credit scoring of the customer. The central idea behind such classification schemes has been to determine and identify good loan applicant groups from bad loan applicant groups.

The nature of decision making towards identifying the right application for disbursing loan remains a complex task. Here, the study has attempted to categorize loan applicants broadly into either good credit or bad credit applications. The model considers the different features of an applicant like age of the applicant, income sources and marital status to categorize them into the specified classes. Financial institutions could then accept the application for loan for those identified good credit members who could in probability honour the repayment of their financial obligations. The classification rules could be individually designed by the creditors based on the features from previously accepted and rejected applications received for all previous loans. In doing so, it is vitally essential to find the accuracy of the proposed models especially designed for credit scoring and improvement in order to reduce the potential of risk associated with the credit system.

In previous literatures, many predictive models have been proposed by different authors. Though Linear Discriminant Analysis (LDA) methods have traditionally been used for credit scoring, it cannot distinguish the applicants owing to the type of the data which is mostly categorical in nature. It has been observed that the covariance matrices are not the same for both the classes. The method of K-Nearest Neighbours (KNN) for predicting credit risk assessment has also been studied by several researchers. Therefore to predict the classification of a new sample point the K-Nearest Neighbours (KNN) algorithm can be applied to databases where the data points are separated into several separate classes. But very limited studies are available in the literature where the researchers used K-Nearest Neighbours (KNN) on the Australian and German credit datasets for credit approval. Therefore, the objective of the study is to use Predictive K-Nearest Neighbours (PKNN) classifier for Australian and German banking credit assessment analysis so that they predict and segment the customers properly.