Predicting And Managing Credit Risk By Implementing Scorecard Using Hybrid Strategy With Trust Rating

M.S.Irfan Ahmed, P.Ramila Rajaleximi

Abstract:Financial institutions possess a great deal of credit risk in assessing credit application for approval. In recent days, to assess, manage and to make decisions on the credit risks of the customers, financial institutions employ internal scorecards. However, major banks use several existing one-dimensional credit scoring model which may lead to inaccurate assessment results. In this paper, a three-dimensional hybrid credit scoring technique has been proposed that includes sequential application scoring along with the dual credit scoring matrix model. Dual credit scoring model uses behavioural credit scoring and credit bureau scoring for computing the trust rating. Also, the behavioural scoring model employs an optimized multiple rank score based feature selection for accurate scoring. On employing the signed approach based trust ratings, the customers are categorized into three risk groups for assessing and managing the customer credit risks. The credit strategies to be followed in making decisions are also presented along with the empirical analysis. The results from the analysis show that the proposed method provides 88% precision with 43.17 K-S statistics value.

Index Terms:Dual Scoring Model, Application Scoring, Behavioural Scoring, Credit Bureau Scoring, Hybrid Scoring Strategy, Sequential-Matrix Model, Trust Rating.

---

1 INTRODUCTION

LOANS and mortgages become essential as it provides financial strength apart from the regular income to make life easier. Loans are provided by the lenders in the form of credit card, home loan, education loan, personal loan, or car loan, etc. to the borrowers with certain constraints. Unfortunately, especially in India, receiving a loan is a tedious process. Conversely, this is not the case for the individuals having a worthy credit score. The lenders have to verify the credit history details before providing financial services to the customers, generating the assessment report about the personal portfolios and creditworthiness of the customer manually. As it is a time-consuming process, the concept of credit score came into existence which is a time-saving process and ease of use. Generally, the credit score is a numerical value that assesses the customer's creditworthiness. The process of measuring the credit score of the individuals from their credit history is termed as credit scoring. The financial institutions make use of such values for making decisions regarding credit applications. Thus in recent days, the credit score becomes the most substantial indicator of the trustworthiness of individuals because of the two main reasons. The first reason is that the model can be used by many people without prior knowledge or background details. The second reason is that unlike evaluating the customer frequently based on the incoming credit application, the model collects the past credit history about the customers, and thus evaluating the credit application is a time-saving and money-saving process. Whenever an application for a loan from a customer is received, banks generally verify and assess the customers' credit bureau score (in India it is termed as CIBIL Score) [1].

- *P. Ramila Rajaleximi is currently pursuing Ph.D. under Research and Development Centre, Bharathiar University, Coimbatore 641046, India. Corresponding Author E-mail: psleximi@gmail.com*
- *M.S. Irfan Ahmed is currently working in Department of Computer and Information Sciences, College of Science and Arts, Taibah University, Al Ula, Madhina, E-mail: msirfan@gmail.com*

Based on the customers' past credit history and the creditworthiness collected from the banks, the credit bureau calculates and update the score and report them to the corresponding banks. If the score is higher, the chance for getting the application to be approved is higher. According to paisabazar.com, India's largest and innovative personal finance platform, approximately 79% of mortgages or credit cards are sanctioned for the customers having a CIBIL score greater than 750 [2]. However, instead of admitting the CIBIL score as it is, the financial institutions will verify other parameters of the customers before approving the application. Some of the parameters are listed below [3]. Employment Status: Apart from credit history assessment, the creditworthiness of the customers is also verified based on their employment status and income. Account Details: Financial institutions verify the customer's creditworthiness by assessing the details of their bank accounts. Payment History: The payment history is verified for any default on payments or amount overdue cases which act as an important factor in reducing the creditworthiness of the customers. Equated monthly instalments (EMI) to Income ratio: The financial institutions also consider the ratio of equated monthly installments towards the monthly income of the customer at the time of approval in assessing the creditworthiness. The chance of getting loan approval increases if the ratio decreases below 50%. If the banks are satisfied based on the above parameters, the eligibility criteria will be verified for the customers' trustworthiness. Some basic requirements considered by the financial institutions in terms of document verification are: Identity Proof: Individuals identity will be verified using Aadhar card, PAN card, Voters ID, Driving license or valid passport. Address Proof: Address for communication will be verified through Aadhar card, Voters ID, Driving license, Utility bills or Valid passport. Proof of Employment: Details about the employment will be verified using Salary slip, Official ID card or Letter from the company. Income Proof: The income will be verified based on the past 3 months' bank statement and salary slip for the last 3 months. Finally, to verify the individual and to reduce the risk of impersonation, passport size photographs are also collected from the customers. Thus, the scoring mechanism is adopted recently by almost all the financial institution in verifying the creditworthiness of the borrower. Though the major advantage of the credit scoring model is to analyze or assess the creditworthiness of the customers, some other advantages that exist in the scoring model are automated decision-making process, time reduction in making decisions, cost reduction, fast adaptation for market conditions [4]. Owing to the advantages of credit scoring, several methods are introduced by the researchers in developing a scorecard which is either statistical or non-statistical models. However, most of the methods implement any one of the scoring methods such as application score, behavioural score or credit bureau scores. This paper presents the sequential matrix based credit scoring strategy with trust ratings in reducing the credit risks. The method uses the application scoring model for sequential along with the behavioural scoring model and credit bureau scoring model for dual scoring matrix-based approach. The dual scoring model additionally uses signed approach based trust ratings for categorizing the credit risks involved in sanctioning the loan approval for a customer or borrower. The paper is organized as follows. Section 2 presents a literature survey related to the proposed work. Section 3 explains the proposed sequential-matrix based hybrid scoring model including application scoring, and matrix-based dual scoring mechanism with behavioural scoring and credit bureau scoring. Section 4 presents the empirical analysis for dual scoring with sign assignment based trust ratings. Finally, the paper concludes the work with the conclusion section.

2 Related Works

Credit scoring becomes the most significant part of any financial institutions due to the growing competition in the credit industry. Several methods exist in developing a scorecard for assessing creditworthiness; however, research on credit score is getting increased to improve the performance of the classification and its accuracy. Several machine learning techniques have been employed to improve classification accuracy. The classifiers such as Neural Networks (NN), Support Vector Machines (SVM), Random Forests (RF), Decision Trees (DT), and Naïve Bayes (NB) are widely used in credit scoring [5]. Neural networks and artificial neural networks along with machine learning techniques are used in credit scoring model for binary classification [6], [7], [8]. Other classifiers such as genetic programming [9], decision tree [10], back-propagation [11] are also employed in assessing the credit risk. Hybrid credit scoring model was also proposed that enhance the performance of the classification by incorporating genetic programming and support vector machines [12]. Logistic regressions are most significant in developing a scorecard for assessing credit risk [13], [14]. Several variations of Logistic regression such as stepwise and best-subset based logistic regression and multinomial and ordinal logistic regression are also implemented in credit scoring [15]. A fuzzy credit scoring model was also proposed by employing fuzzy logistic regression which is considered as an appropriate prediction approach for credit risk classification [16]. Recently behavioural scoring has become most popular as the model tries to assess the behaviours of the borrowers based on the past credit history. A set of constraints was proposed as a part of a semi-supervised constrained clustering algorithm to assess the rational behaviour of the defaulter. In addition, the method tries to differentiate the defaulter behaviours as 'can't pay' and 'won't pay' [17], [18]. The hybrid model was suggested by several researchers that make use of either the combination of behavioural score and credit bureau score or the combination of application score and behaviour score to improve the accuracy of risk prediction [12], [19]. All these methods employ either a single model or dual model in predicting credit risk which lacks in accuracy. Thus, to improve the performance of the credit scoring model and to increase the classification accuracy, the proposed method employs application score, behavioural score, and credit bureau score.

3 Sequential-Matrix based Hybrid Scoring

The method has been suggested to predict and manage the credit risks by developing an in-house scorecard. The method uses banks internal data to construct the sequential-matrix based hybrid approach with trust rating. The application score is the first level filter or primary filter in identifying the credit risk of the customer and for the secondary filter, the method employs dual scoring based matrix approach that uses behavioural score computed from the bank's internal data and credit bureau score computed by the credit bureau. Fig. 1 shows the proposed sequential-matrix based behavioural scoring architecture. The applicants are initially scrutinized based on the applications and upon passing the given cutoff score, the applications are moved to the next level matrix approach to examine behavioural score and the score given by the credit bureau. This combination of sequential-matrix approach is more versatile than the individual sequential approach and matrix approach. The application score is computed using experts' judgment employing the parameters with weights. Once the application satisfies the given cutoff score in the primary level, the applications are moved to matrix scoring in the secondary level in which statistical scores are evaluated to accept or reject the applications using signed approach based trust ratings [20], [21]. The detailed framework of the proposed model is given in Fig. 2.

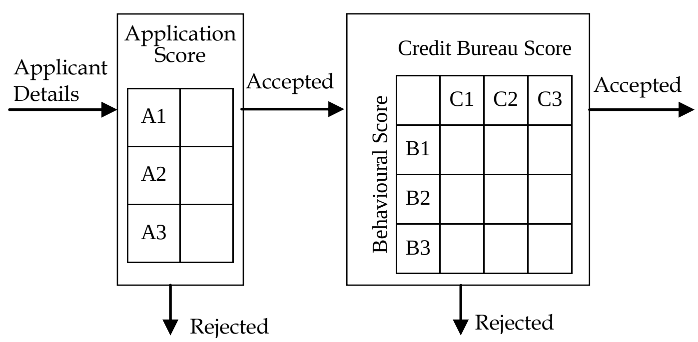

***Fig. 1. Sequential-Matrix based Credit Scoring Architecture*.**

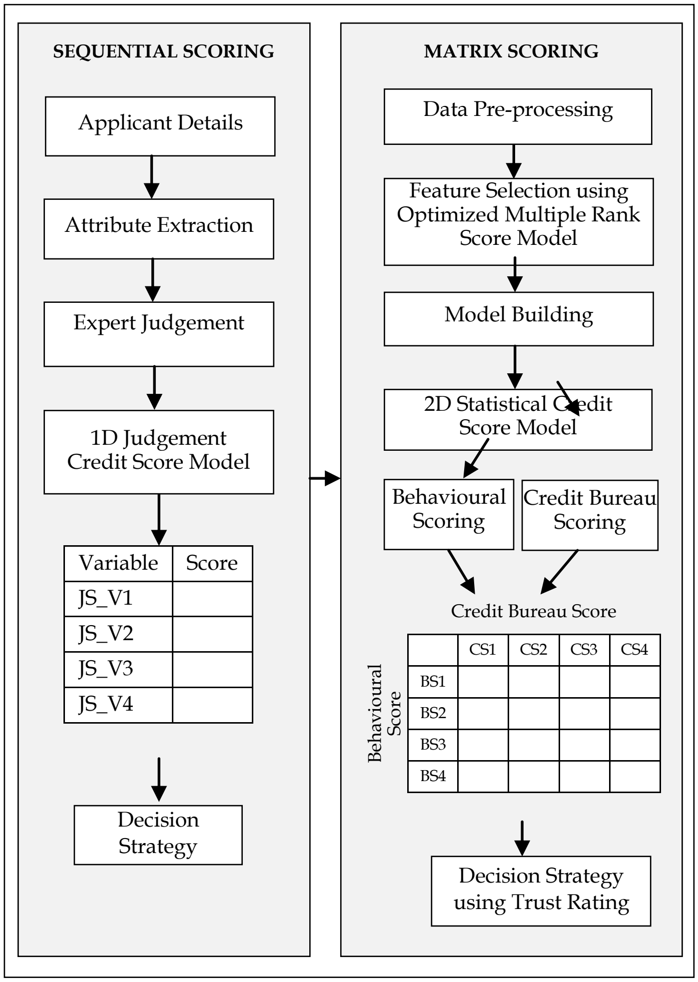

Fig. 2. Detailed Framework of the Proposed Model.

3.1 Application Scoring

Generally, the scores are computed to estimate the customers' behaviour on repaying the loan. Specifically, it focuses on negative behaviour such as bankruptcy. The estimation can be carried out either using a judgmental way or statistical way. In case of application score estimation, the scores map the customers' characteristics specified in the application and the customers' creditworthiness [22]. Based on the data collected from the application using several parameters such as demographic, financial, social, etc., the credit risks are estimated and quantified. However, the application can be processed upon satisfying the three months' period after opening the account. In the proposed method, judgmental scores are computed based on the attributes extracted from the clients' application. The judgmental scorecard is designed based on the experts' judgment and the knowledge gained by the financial institution from the past experience. It is advisable to employ the judgmental scorecard if the number of parameters in the data to be processed is very minimum. This method is simple and the parameters to be processed are around 10 to 15 in numbers. The experts will identify the important parameters that are useful in making decisions are identified and the possible values for each parameter are also identified in which the scores for each value in the parameters especially for categorical values are assigned. Once the scorecard is computed based on the experts' judgment and institutions' experience, the neutral values for each parameter are assigned and the total accept or reject cutoff is calculated by summing the neutral values given for the parameters. The proposed method uses a cutoff score for accepting the application. The neutral values for the parameters employed in the application are assigned by the experts. The new application of the customer is analyzed. Based on the developed scorecard, the scores for the values given in the application are allocated based on which the total scores are computed by summing the values of the parameters. If the total value scored by the applicant is above the given cutoff, then the application is accepted for further process. However, if the total value scored by the application is below the cutoff value, then the application is rejected and it is removed from further processing. The sample experts' judgment scorecard based on the application of the customer prepared by the experts of a bank is given in Table 1.

TABLE 1
SAMPLE EXPERTS' JUDGMENT SCORECARD

A1: Time at present address	
less than 1 year	1
1 to 3 years	2
3 to 5 years	3
5 to 10 years	4
more than 10 years	5
Neutral	*3*
A2: Home Status	
Rental	1
Own	3
Neutral	*1*
A3: Age of Borrower	
Less than 17 Years	0
18 to 25 years	3
25 to 35 years	5
35 to 45 years	4
45 to 60 years	2
More than 60 years	1
Neutral	*3*
A4: Martial status	
Married	3
Unmarried	1
Neutral	*1*
A5: Work Status	
Govt. / PSU Employee	*5*
Private Employee	4
Own business	3
Part-time working	2
Retired	1
Neutral	*3*
A6: Years of Work Experience	
Less than 3 years	*2*
3 to 10 years	*3*
more than 10 years	1
Neutral	*2*
A7: Spouse Working status	
Yes	2
No / Unmarried	0
Neutral	*2*
A8: Annual Income	
less than Rs.5,000/-	1
Rs.5,000 – Rs.10,000/-	*2*
Rs.10,000 – Rs.20,000/-	3
Rs.20,000 – Rs.40,000/-	4
more than 40,000/-	5
Neutral	*3*
A9: Time with Bank	
less than 3 months	0
3 months to 2 years	1
2 years to 5 years	*2*
5 years to 10 years	3
greater than 10 years	4
Neutral	2
A10: Bank Account History	
Cheque bounce history	0
Minimum balance not maintained	1
Overdraft allowed in account	*0*
No remarks	3
Neutral	1
A11: Security related	
Brand new	4
Up to 1 year old	3
More than 1 year up to 2 years old	2
More than 2 years old	*1*
Neutral	*2*
A12: Loan Repayment mode	
Salary account deduction by employer	5
Salary account in branch	4
ECS mode from other banks	*3*
Cheques given from other banks	2
Cash payment	1
Neutral	*3*
A13: Nature of Security	
EMT/Hypothecation of durable utility article or vehicle	3
Third party personal Guarantee/Co-obligation	1
Clean loan	*1*
Neutral	*3*
Cutoff: 27	

In the sample experts' judgmental scorecard, 13 attributes are identified with their own characteristics based on which the cutoff score is computed as 27. The applications having the scores greater than the cutoff score is accepted and the applications getting scores lesser than the cutoff score is rejected. Thus, the first level of scoring is helpful in making decisions on further processing of an application. It acts as a primary filter or selection based on the details given in the application. Once the application is selected for the second level of processing, the complete details about the customer are collected and the customer is verified based on the proofs provided by them.

3.2 Behavioural Scoring

Behavioural scoring is highly helpful in forecasting the likelihood that the next loan installment will be late. With this type of scoring, the contrast between the good and the bad will get maximized. It is an internal scoring system in which deeper analysis will be made on the clients' credit behaviour and to monitor or forecast the performance of the accounts across probability of default or delinquency over the particular time horizon [22]. This model is used in conjunction with application scoring or credit bureau scoring by many of the financial institutions with detailed analysis of internal data for decision making. It is also highly helpful in improving credit portfolio management and understanding or managing customers. Some of the steps in computing scoring include data pre-processing, attribute selection and model building. Data Pre-processing: Data collection and pre-processing is the primary step in developing a scorecard. The internal data related to customers' behaviour is collected from financial institutions. This data includes attributes related to customer borrowing characteristics and payment characteristics. However, the internal data collected from the customers and other sources are often incomplete with missing attribute values, inaccurate with errors or outliers that deviate from the majority of data and inconsistent with discrepancies in the attribute values. However, the quality of data is a more important factor and it is the base for any analysis and mining process [19], [23]. Thus accuracy, completeness, and consistency are the three main pillars of data quality. Therefore, to improve the quality of the data, the collected data undergoes pre-processing with several tasks such as data integration, data cleaning, and data transformation. Data about the customers' behaviour are generally collected from various sources and each source will have data in a different format. Data integration integrates multiple databases or files from different sources into a common form. Data cleaning endeavours to clean the incomplete data by filling the missing values and by removing outliers. The data transformation is applied to the collected dataset to transform them to the common format using normalization, where the attribute values are scaled to fall in a particular range, discretization in which attribute values are replaced with interval labels for numerical values and conceptual labels for categorical values. Concept hierarchies are also generated for nominal values. Data transformation is applied during model building. Attribute Selection: Attribute selection is one of the significant steps in machine learning. This step influences the performance of the model. Though the data collected has been pre-processed, it contains several attributes out of which many of the attributes may not be significant for the study. These irrelevant attributes that do not contribute or degrade the accuracy of the underlying study are to be removed before further processing. In this proposed model, the optimized multiple rank score model has been employed for fetching significant attributes [24]. This feature selection mechanism comprises of two main phases. In the first phase, various ranker search methods with different attribute strategies are employed. Each ranker generates different ranks for the same attribute and thus the output of this phase is multiple sets of attribute ranks. In the second phase, the rank score accumulation is carried out in which n number of attribute rankings generated from the previous phase will be served as an input. Once the final scores are identified then the attributes having minimum value than the given threshold are considered as less relevant and can be removed. Finally, the output includes only the optimal features that are considered as highly relevant after removing the features that are less relevant. Model Building: Once the collected dataset is pre-processed and the pertinent attributes are selected, the next step is to build the model. This model building is the primary concept that contains several steps that include statistical evaluation for developing the scorecard. The following are the various steps to be followed in developing a scorecard. Attribute Discretization Attribute Calibration Attribute Value Score Computation *Attribute Discretization* - Each attribute or features will have many values due to which the model building will become a tedious task. Thus, the attribute values having similar behaviour must be grouped. The advantages of grouping the attribute values include a better understanding of the attribute relationship, easy implementation and gaining the deeper insight on behaviours that act as a risk predictor. The attribute discretization will be normally carried out using binning, normalization and concept hierarchies. Binning is a proper method where the numerical values are converted to the categorical or range of values. For example, age attribute values are converted to categorical values such as 20-30, 30-40, … 80-90. Normalization is another procedure that converts a large range of values to a small range. For example, the income attribute values can be normalized in such a way that the output produces a new range say, [0-1] or [1-10], etc. Concept hierarchies are employed to convert the nominal values that have a finite number of distinct values. For example, the nominal values for the occupation attribute can be grouped and can be labelled as professional jobs or white collar jobs. *Attribute Calibration* - The next step in model building is to calibrate the score computation process using logistic regression. According to literature, logistic regression is considered to be the perfect statistical method for binary classification problem [12, 25]. The weight of evidence and information value are two statistical techniques used in logistic regression for the attribute categories. These steps are essential to verify the binning process or discretization performed in the previous step. Once the attribute values are grouped as categories for all the selected attributes in the dataset, the weight of evidence (WoE) is computed for all the possible attribute values. This step is highly helpful in model building to verify the previous attribute discretization step and to proceed with the next step. Basically, WoE deals with computing the weight or strength of the particular attribute category value that corresponds to the good or bad customer behaviour class. Good customer behaviour indicates the customers who do not have default and bad customer behaviour class indicates the customers who have defaulted the load. This computation takes the proportion of good customer behaviour class to bad customer behaviour class for each grouped categorized attribute values. The positive WoE values indicate that the particular attribute grouping or categorization includes higher good applicants than bad applicants which specify the low credit risk with that group of customers. However, the negative WoE values indicate that the particular attribute grouping or categorization includes higher bad applicants than good applicants which specify the greater credit risk with that group of customers. Thus, the formula to compute the weight of evidence (WoE) is given in Eq. (1)

$$WeightofEvidence = In\left(\frac{Distributio\ n_{Good}}{Distributio\ n_{Bad}}\right) \quad (1)$$

where $Distribution_{Good}$ specifies the percentage of good customers in the particular attribute group and $Distribution_{Bad}$ specifies the percentage of bad customers in the particular attribute category. At this stage, the attribute categories having similar WoE values can be combined for the continuous variables and the attribute categories can be combined with new WoE values as a label for the categorical values for performance improvement of the model [26]. The next step is to compute the information value for each attribute value category. This step identifies the most significant attribute that is used in the prediction model. This information value depicts the predictive power of the attribute category and it selects and ranks the variable based on the computed predictive strength of the attribute. The formula to compute the Information value is given in Eq. (2).

$$Info\_Val = (Distribution_{Good} - Distribution_{Bad}) \times WoE \quad (2)$$

The attribute selection using information value can be made using the predictive strength of the attribute and the decisions can be made as given in Table 2 [27].

TABLE 2
INFORMATION VALUE VERSUS DECISION RULE

Information Value Range	Decision Rule
< 0.02	Useless for prediction and can be removed
0.02 to 0.1	Weak prediction and can be selected
0.1 to 0.3	Medium prediction and can be selected
0.3 to 0.5	Strong prediction and can be selected
>0.5	Very good prediction and can be selected

Attribute *Value Score Computation* -Logistic regression is applied to the attribute values and the regression coefficients are computed for each attribute. The regression coefficient scales the value in which the score values are confirmed to the particular range. Thus, the scorecard is developed in such a way that it corresponds to the scores for each attribute values. And, once the scorecard is developed, the credit score of a client can be computed by summing the scores for the attribute values of the particular client. This work scales in such a way that a total credit score of 300 points corresponds to G/B odds of 2 to 1 with an increase of 20 points with respect to Points to Double the Odd. The scores for each attribute value are computed using the formula given in Eq. (

$$Score = In(odds) \times factor + offset \quad (3)$$

$$Score = \sum_{i,j=1}^{n,m} \left(-(WoE_j \times \beta_i + \frac{a}{n}) \times factor + \frac{offset}{n}\right) \quad (4)$$

$$In(1) \times factor + offset = 300$$

$$In(2) \times factor + offset = 320$$

$$factor = \frac{20}{In(2)} = 28.85$$

$$Offset = 300 - factor \times In(2) = 280$$

where WoE is the weight of evidence for each grouped attribute, n isthe number of attributes, m is the number of categories in each attribute, β isthe regression coefficient for each attribute and a is the logistic regression intercept. In support with making decisions, the credit scores are generally divided into different risk categories. For example, the scores below 350 are considered as extremely high risk, the scores between 350 and 400 as high risk, scores between 400 and 450 as medium risk, scores between 450 and 500 as low risk and scores above 500 is the extremely low risk.

3.3 Matrix based Dual Scoring

Generally, the credit bureaus such as Equifax, TransUnion, and Experian collect the credit-related information about all the customers and keep track of all their accounts. Based on the collected information, they maintain a credit score for each customer. The financial institution collects the scores for the corresponding customer based on the customer's loan application. The advantage of having the credit bureau score is that the scores are centralized. In India, CIBIL scores are the most commonly used by the banks which are a part of TransUnion. Thus, the CIBIL score will be computed and the banks that receive the customer's application can request the score to make credit decisions. However, the main disadvantage is that the credit bureaus collect the information independently and due to which there is a possibility for missing the transaction details about the customers which lead to wrong credit score. Thus, the main objective of credit scoring is to predict the credit risk involved in approving the loan. Thus, instead of trusting the credit bureau score fully, the bank can collect and maintain the customer's behavioural data and can be used along with the credit bureau score. The dual scoring model highly analyzes the credit risk of the customer and helps in designing credit strategies based on good and bad rates.

4 EMPIRICAL ANALYSIS

4.1 Behavioural Scoring Model

For empirical analysis of the behavioural scoring model, the German credit dataset from the UCI repository has been chosen to conduct the initial experiments. The dataset contains 20 attributes in which 17 are numerical and 13 are categorical with 1000 instances including 700 good and 300 bad customers. In the proposed model, the scaling is chosen as a total credit score of 300 points corresponds to G/B odds of 2 to 1, the number of instances in the German credit (GC) dataset as 900 is chosen randomly with 600 good and 300 bad. However, in reality, the bank can collect the internal details about accounts opened for not less than 12 months.

The transactions carried out in the past 12 months will be served as the payment history. The dataset is pre-processed and the significant attributes are selected using an optimized multiple rank score model [24] that uses classifier attribute evaluators [28] such as ZeroR, Decision Table, M5Rules, Correlation and Relief algorithm [29].The significant attributes such as status of existing checking account, duration in month, credit history, purpose, credit amount, savings account/bonds, present employment since, installment rate in percentage of disposable income, personal status and sex, property, other installment plans, housing, foreign worker that are greater than the given threshold as 0.5 are extracted for further processing. Once the significant attributes are selected, the next step is to start the model building with attribute discretization. However, as the GC dataset contains discretized attribute values, the numerical values are discretized. The next step to be processed is attribute calibration by computing the weight of evidence as in Eq. (1) and the Information value as in Eq. (2). The attributes such as savings account/bonds, present employment since and installment rate in percentage of disposable income are considered as weak and useless predictors and hence the three attributes are removed before further processing. Thus, 10 useful predictors are identified from the GC dataset after computing the information value. Once the useful predictors are identified, the attribute values are replaced with the weight of evidence and the logistic regression is applied over the data. Logistic regression employs the selected useful predictors to forecast the probability of binary outcome such as good or bad. The selected useful attribute predictors along with its information value and regression coefficients and the intercept are computed and are presented in Table 3.

TABLE 3

INFORMATION VALUE AND COEFFICIENTS FOR THE ATTRIBUTE OF GERMAN CREDIT DATASET

S.No	Variable	Information Value	Coefficients
1	Status of Existing Checking Account	0.131	0.0102
2	Duration in Month	0.220	0.0074
3	Credit History	0.332	0.0083
4	Purpose	0.175	0.0104
5	Credit Amount	0.127	0.0051
6	Personal Status and Sex	0.060	0.0118
7	Property	0.101	0.0032
8	Other Installment Plans	0.060	0.0073
9	Housing	0.086	0.0061
10	Foreign Worker	0.040	0.0101
	Intercept		0.6861

After computing the regression coefficients of the attributes, the next step is to compute the attribute score points for each attribute values as given in Eq. (4). Table 4 presents the Good/Bad odd ratio, WoE and the calculated attribute points. To analyze the behavioural scoring model, K-S Statistics, accuracy and precision are computed for the proposed model. The 900 records are drawn at random from the GC dataset are taken for the analysis in the ratio 2:1 that comprises of 600 good and 300 bad records. The Kolmogorov–Smirnov test is a statistical measure that quantifies the maximum distance between the cumulative distributions of the two class labels. This measure also helps in identifying the cut-off value in making a decision either to accept or reject the application.

TABLE 4

WEIGHT OF EVIDENCE AND ATTRIBUTE POINTS FOR ATTRIBUTES OF GERMAN CREDIT DATASET

Attributes and their Types	Attribute Values or Behaviours	WoE	Good/ Bad Odd Ratio	Points
Status of Existing Checking Account	Below 0 DM	-27.59	0.6824	36
	At least 0 and at most 200 DM	-2.69	0.9582	43
	Above 200 DM	19.89	1.0093	50
	No checking account	91.63	2.3032	71
Duration in Month	Below 10.5	79.85	2.2578	61
	At least 10.5 and below 12.5	21.91	1.3251	49
	At least 12.5 and below 19	3.51	1.0059	45
	At least 19 and below 25	2.99	0.8962	45
	At least 25 and below 37.5	-45.95	0.5629	34
	Above 37.5	-97.34	0.4392	23
Credit History	Critical account existing (not at this bank)	77.01	1.9946	63
	Existing credits paid back duly till now	-8.65	0.8237	42
	Delay in paying off in the past	-19.67	0.9733	39
	No credits taken/ all credits paid back duly	-151.41	0.2725	9
	All credits at this bank paid back duly	-119.21	0.4001	16
Purpose	Domestic appliances	44.76	1.3619	58
	No Vacation	-48.84	0.6882	30
	Radio/television	-12.86	0.8223	40
	Car (new)	-37.78	0.4977	33
	Car (used)	73.63	2.8316	66
	Business	-23.05	1.2416	37
	Repairs	-47.00	0.6060	30
	Education	-20.76	0.6853	38
	Furniture/equipment	-69.31	1.1157	23
	Retraining	125.28	3.7524	82
Credit amount	Below 1203.5	10.44	0.8024	46
	At least 1203.5& below 1554.5	18.23	1.0086	47
	At least 1554.5& below 2321.5	21.40	1.1033	47
	At least 2321.5& below 3373	30.01	1.5066	49
	At least 3373 &5553	-1.03	1.1406	44
	Above 5553	-76.51	0.6466	33
Personal Status and Sex	Male - Divorced/ Separated	-59.78	0.4760	24
	Female - Divorced/ Separated/Married	-26.05	0.8353	35
	Male -Single	18.97	1.3672	51
	Male - Married/Widowed	13.10	0.9314	49
Property	Real estate	42.20	1.2851	48
	Building society savings agreement/ life insurance	-2.14	1.0064	44
	Car or other, not in attribute 6	-0.49	1.0614	44
	Unknown/ no Property	-58.04	0.6180	39
Other Installment Plans	None	12.37	1.3444	47
	Bank	-48.77	0.8159	34
	Stores	-50.21	0.6310	34
Housing	Own	19.70	1.1887	48
	Free	-47.00	0.6973	36
	Rent	-41.99	1.1972	37
Foreign Worker	Yes	-3.26	1.1453	43
	No	121.64	3.9262	80

The score that makes the maximum difference is considered to be the cut-off value. Thus, the application having a behavioural score less than the cut-off value is rejected. From the analysis done for the German credit dataset, the computed cutoff value is 437 accurately. Table 5, presents the K-S statistics computed for the 900 records of the GC dataset. According to K-S Statistics, the maximum difference between the cumulative good and cumulative bad distribution is 43.17 while representing the credit score at 20 points scale. Thus, the value corresponds to the score 440 approximately is the predicted cut-off score. The good count and bad count for the credit scores given in table 5 is depicted as a graph in Fig. 3. The cumulative good distribution and cumulative bad distribution values along with the K-S statistics are plotted in a graph and is presented in Fig. 4.

TABLE 5
K-S STATISTICS COMPUTATION

Credit Score	#Good	#Bad	Good Distr.	Bad Distr.	Cumulative Good Distr.	Cumulative Bad Distr.	K-S Statistics
320	0	0	0.00	0.00	0.00	0.00	0.00
340	0	1	0.00	0.33	0.00	0.33	0.33
360	4	3	0.67	1.00	0.67	1.33	0.67
380	6	14	1.00	4.67	1.67	6.00	4.33
400	16	32	2.67	10.67	4.33	16.67	12.33
420	40	77	6.67	25.67	11.00	42.33	31.33
440	81	76	13.50	25.33	24.50	67.67	***43.17***
460	183	62	30.50	20.67	55.00	88.33	33.33
480	144	31	24.00	10.33	79.00	98.67	19.67
500	72	3	12.00	1.00	91.00	99.67	8.67
520	41	1	6.83	0.33	97.83	100.00	2.17
540	13	0	2.17	0.00	100.00	100.00	0.00

Fig. 3. Good Counts and Bad Counts for the Credit Scores.

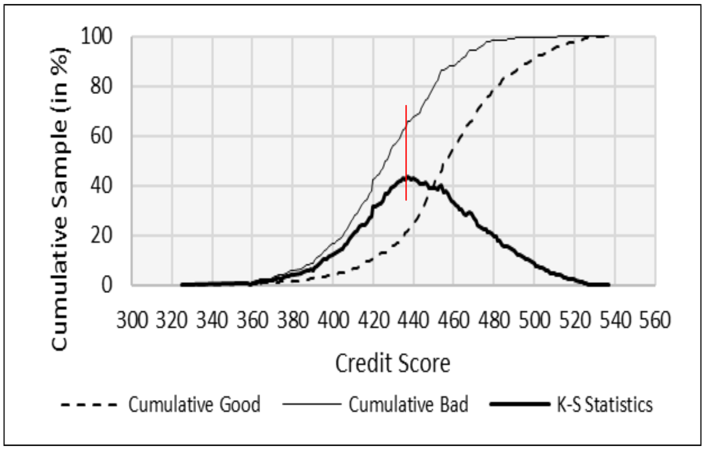

Fig. 4. K-S Statistics.

In Fig. 4, the red line represents the maximum difference between the cumulative distribution of the two class labels good and bad. The sensitivity, specificity, accuracy and precision values are computed for the German credit dataset by varying the number of records from 100 to 900. The values are given in Table 6. Thus, the average sensitivity value is 0.8, specificity value is 0.72, the average precision is 0.88 and the average accuracy is 0.77.

TABLE 6
PRECISION AND ACCURACY FOR GERMAN CREDIT DATASET USING BEHAVIOURAL SCORING

Dataset Size	Sensitivity	Specificity	Precision	Accuracy
100	0.82	0.91	0.95	0.85
200	0.78	0.91	0.98	0.80
300	0.78	0.73	0.86	0.76
400	0.80	0.73	0.88	0.78
500	0.80	0.70	0.87	0.77
600	0.81	0.70	0.90	0.78
700	0.80	0.70	0.91	0.78
800	0.81	0.68	0.88	0.78
900	0.81	0.67	0.83	0.76
Average	*0.8*	*0.72*	*0.88*	*0.77*

The precision and accuracy of various classification methods are compared with the classification accuracy of the proposed behavioural scoring technique for the German credit dataset and is given in Table 7. Fig. 5 depicts the comparison of various analysis using TP Rate, FP Rate, precision, and accuracy. Thus, based on the analysis, the credit score range and ratings are categorized and are shown in Table 8 for making decisions in approving the application. The ratings for the behavioural score are also categorized by assigning signs for the credit score. Thus, low negative (--), negative (-), neutral (0), positive (+) and high positive (++) are assigned for bad (<400), poor (400-449), fair (450-499), good (500-549) and excellent (>550) ratings.

TABLE 7

COMPARISON OF VARIOUS CLASSIFICATION METHODS

Methods	TP Rate	FP Rate	Precision	Accuracy
Proposed Method	80.8	33.3	82.9	76.11
Support Vector Machine	71.2	46.4	69.6	71.22
Logistic Regression	71.8	42.8	70.4	71.78
Naive Bayes	69.9	44.9	68.2	69.88
Bagging	68.3	47.0	66.4	68.33
AdaBoost	67.8	54.9	64.7	67.78
RandomForest	69.0	46.8	67.1	69.00

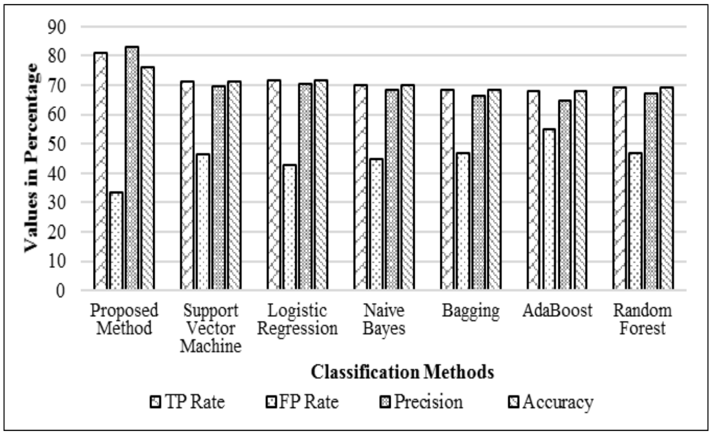

Fig. 5. Comparison of Various Classification Methods.

4.2 Credit Bureau Scoring with Sign Assignment

Generally, the credit bureaus collect the transaction details and maintain the report for all the account holders. Based on the credit details, the credit score is measured for the account holders with respect to the application. Normally, the CIBIL scores range from 300 to 900. However, the score range differs from one bureau to another. TransUnion CIBIL, Experian Equifax and CRIF High Mark are the major credit bureaus that compute the credit scores. Also, the computed credit bureau score for the given account holder might differ for the bureaus as they are independent. Similar to sign assignment employed in behavioural scoring, the credit bureau score ratings are also categorized by assigning signs for the credit score. The credit score range and ratings are categorized and are given in Table 8.

TABLE 8

SCORE RANGE AND RATINGS

Behavioural Credit Score Range	Behavioural Score Rating	Sign Assignment	Credit Bureau Rating	Credit Bureau Score Range
Below 400	Bad	--	Bad	Below 550
400 – 449	Poor	-	Poor	550 – 649
450 – 499	Fair	0	Fair	650 – 699
500 – 549	Good	+	Good	700 – 749
550 and above	Excellent	++	Excellent	750 and Above

4.3 Dual Scoring Model with Trust Rating

The five ratings along with their signs of behavioural scoring and credit bureau scoring are combined to make a dual scoring model with 5 x 5 matrix. Based on the assigned signs, the trust ratings are computed. These trust ratings are grouped into three categories of risk groups which are helpful in making decisions. The 5 x 5 dual scoring model matrix is given in Table 9. The signs presented in the cells represent the trust rating.

TABLE 9

DUAL SCORING MODEL WITH TRUST RATINGS

Dual Scoring Model			Credit Bureau Scoring Model				
			< 550	550–649	650–699	700–749	> 750
			--	-	0	+	++
Behavioural Scoring Model	< 400	--	-	-	-	-	0
	400 – 449	-	-	-	-	0	+
	450 – 499	0	-	-	0	+	+
	500 – 549	+	-	0	+	+	+
	>550	++	0	+	+	+	+

The trust ratings of the cells (- or +) are assigned based on the maximum number of signs or neutral if + and – occurs in equal number and is given in Table 10. The low-risk category is represented in pale green colour cells, pale red colour cells represent the high-risk category and the pale yellow colour cells represent the medium risk category. Thus, the maximum number of – sign represents the low-risk group. An equal number of + sign and – sign depicts neutral and the increase in the number of + sign than - sign is 1, then the cells correspond to the medium risk group. The maximum number of + sign corresponds to the high-risk group. The trust rating is depicted in Table 10.

TABLE 10

TRUST RATING

Signs of Scoring Model	Trust Rating	Signs of Scoring Model	Trust Rating
(--, --) or (-, -)	-	(-, +) or (+, -) or (0, 0)	0
(--, -) or (-,--)	-	(-, ++) or (++, -)	+
(--, 0) or (0, --)	-	(0, +) or (+, 0)	+
(--, +) or (+, --)	-	(0, ++) or (++, 0)	+
(--, ++) or (++, --)	0	(+, +) or (++, ++)	+
(-, 0) or (0, -)	-	(+, ++) or (++, +)	+

Hence, based on the categories identified from the dual scoring model with trust ratings such as low-risk group, medium-risk group and high-risk group, different strategies can be applied to collect the loan amount. Low-risk group corresponds to a very low bad rate and thus the possibility of repaying the loan on time is very high. However, an automatic reminder message can be sent to the low-risk group customers before the due date. Medium-risk group corresponds to low bad rate and the possibility of repaying the loan on time is neutral. Thus, the details about the medium risk group can be monitored by junior authorities and these authorities will be responsible for reminding the due date for repayment. High-risk group corresponds to very high bad rate and thus the possibility of repaying the loan on time is very low. As this group corresponds to the high-risk group, the details about this high-risk group can be monitored by senior authorities directly.

5 Conclusion

Credit scoring is an important aspect of any financial institutions. This paper presents the sequential matrix based hybrid scoring strategy with a trust rating to predict and manage the credit risk. The applications are initially processed to scrutinize them for the next level by computing the application score employing experts' judgment. In the next level, the dual scoring matrix model uses statistical score computations in making decisions. This matrix scoring model uses both the behavioural scoring computed using in-house scorecard by using an optimized multiple rank score model for selecting significant attributes and the credit score given by the credit bureau. The credit ratings in the dual score are assigned a sign and the trust ratings are computed. The model categorizes the trust ratings into three groups for making decisions. The behavioural scoring model has experimented with German credit dataset and the results are analyzed using K-S statistics, precision and accuracy. The K-S value for the proposed behavioural score model is 43.17 whereas the average precision and accuracy are 88% and 77% respectively. The proposed model is highly helpful for the financial institution in reducing and managing the risks associated with the customer credits and also in making decisions.

References

[1] J. Lohokare, R. Dani, and S. Sontakke, "Automated data collection for credit score calculation based on financial transactions and social media," *In International Conference on Emerging Trends & Innovation in ICT (ICEI)*, IEEE, pp. 134-138, 2017, doi: 10.1109/ETIICT.2017.7977024.

[2] Paisabazaar.com, "How to Build or Maintain a Healthy CIBIL Score," available online at https://www.paisabazaar.com/cibil/9-tips-build-maintain-healthy-cibil-score/

[3] CIBIL, Part of TransUnion, "Loan Approval Process," available online at https://www.cibil.com/loan-approval-process

[4] K. Nurlybayeva, G. Balakayeva, "Algorithmic scoring models," *Applied Mathematical Sciences*, vol. 7, no. 12, pp. 571-586, 2013.

[5] M. Ala'raj, and M.F. Abbod, "A new hybrid ensemble credit scoring model based on classifiers consensus system approach," *Expert Systems with Applications*, vol. 64, pp. 36-55, 2016.

[6] C.L. Chuang and R.H. Lin, "Constructing a reassigning credit scoring model," *Expert Systems with Applications*, vol. 36, no. 2, pp. 1685-1694, 2009.

[7] C.F. Tsai, and J.W. Wu, "Using neural network ensembles for bankruptcy prediction and credit scoring," *Expert systems with applications*, vol. 34, no. 4, pp .2639-2649, 2008.

[8] A. Khashman, "Neural networks for credit risk evaluation: Investigation of different neural models and learning schemes," *Expert Systems with Applications*, vol. 37, no. 9, pp. 6233-6239, 2010.

[9] N. Metawa, M.K. Hassan, and M. Elhoseny, "Genetic algorithm based model for optimizing bank lending decisions," *Expert Systems with Applications*, vol. 80, pp. 75-82, 2017.

[10] G. Wang, J. Ma, L. Huang, and K. Xu, "Two credit scoring models based on dual strategy ensemble trees," *Knowledge-Based Systems*, vol. 26, pp. 61-68, 2012.

[11] B.W. Yap, S.H. Ong, and N.H.M. Husain, "Using data mining to improve assessment of creditworthiness via credit scoring models," *Expert Systems with Applications*, vol. 38, no. 10, pp. 13274-13283, 2011.

[12] B.W Chi, and C.C. Hsu, "A hybrid approach to integrate genetic algorithm into dual scoring model in enhancing the performance of credit scoring model," *Expert Systems with Applications*, vol. 39, no. 3, pp.2650-2661, 2012.

[13] G. Dong, K.K. Lai, and J. Yen, "Credit scorecard based on logistic regression with random coefficients," *Procedia Computer Science,* vol. 1, no. 1, pp. 2463-2468, 2010.

[14] A.C. Bahnsen, D. Aouada, and B. Ottersten, "Example-dependent cost-sensitive logistic regression for credit scoring," *In 2014 13th International Conference on Machine Learning and Applications*, IEEE, pp. 263-269, 2014.

[15] C. Bolton, "Logistic regression and its application in credit scoring," Doctoral dissertation, University of Pretoria, 2009.

[16] S.Y. Sohn, D.H. Kim, and J.H. Yoon, "Technology credit scoring model with fuzzy logistic regression," *Applied Soft Computing*, vol. 43, pp. 150-158, 2016.

[17] C. Bravo, L.C. Thomas, and R. Weber, "Improving credit scoring by differentiating defaulter behavior," *Journal of the operational research society*, vol. 66, no. 5, pp. 771-781, 2015.

[18] M. Alborzi, and M. Khanbabaei, "Using data mining and neural networks techniques to propose a new hybrid customer behaviour analysis and credit scoring model in banking services based on a developed RFM analysis method," *International Journal of Business Information Systems*, vol. 23, no. 1, pp. 1-22, 2016.

[19] B.W. Chi, C.C. Hsu, and M.H. Ho, "Enhancing credit scoring model performance by a hybrid scoring matrix," *African Journal of Business Management*, vol. 7, no. 18, pp. 1791-1805, 2013.

[20] K. Thiagarajan, A. Raghunathan, G. Ponnamal Natarajan, G. Poonkuzhali, and Prashant Ranjan, "Weighted Graph Approach for Trust Reputation Managements," *International Conference on Intelligent Systems and Technologies*, Published in Proc. Of World Academy of Science and Technology, vol. 56, pp. 830-836, 2009.

[21] G. Poonkuzhali, K. Sarukesi, G.V. Uma, "Web content outlier mining through mathematical approach and trust rating," in the book titled *Recent Researches in Applied Computer and Applied Computational Science*, Included in ISI/SCI Web of Science and Web of Knowledge, Venice, Italy, pp. 77-82, 2011.

[22] L.C. Thomas, J. Ho, and W.T. Scherer, "Time will tell: behavioural scoring and the dynamics of consumer credit assessment," *IMA Journal of Management Mathematics*, vol. 12, no. 1, pp. 89-103, 2001.

[23] S. Kotsiantis, B.D. Kanellopoulos, and P. Pintelas, "Data Preprocessing for Supervised Leaning," *International Journal of Computer Science*, vol. 2, no. 2, pp. 111-117, 2006.

[24] P. Ramila Rajaleximi, M.S. Irfan Ahmed, and Ahmed Alenezi,, "Feature Selection using Optimized Multiple Rank Score Model for Credit Scoring," *International Journal of Intelligent Engineering and Systems*, vol. 12, no.2 pp. 74-84, 2019.

[25] E.K. Laitinen, and T. Laitinen, "Bankruptcy prediction: Application of the Taylor's expansion in logistic regression," *International Review of Financial Analysis*, vol. 9, no. 4, 327–349, 2000.

[26] G. Zeng, "A necessary condition for a good binning algorithm in credit scoring," *Applied Mathematical Sciences*, vol. 8, no. 65, pp. 3229-3242, 2014.

[27] N. Siddiqi, *Credit Risk Scorecards. Developing and Implementing Intelligent Credit Scoring*. John Wiley & Sons, Inc, 2006.

[28] E. Kalapanidas, N. Avouris, M. Craciun, and D. Neagu, "Machine Learning Algorithms: A Study on Noise Sensitivity," *In Proc. of Balcan Conference in Informatics*, pp. 356-365, 2003

[29] R.J. Urbanowicz, M. Meeker, W. La Cava, R.S. Olson, and J.H. Moore, "Relief-based Feature Selection: Introduction and Review," *Journal of Biomedical Informatics*, vol. 85, pp. 189-2013, 2018.