LOAN PREDICTION USING LOGISTIC REGRESSION*

BY

SHAILESH VARMA♣

Student, School of Computer Science and Engineering, Ajeenkya D Y Patil University, Pune

shailesh.varma@adypu.edu.in

ARUNA VERMA

Assistant Professor, School of Computer Science and Engineering,

Ajeenkya D Y Patil University, Pune, India

aruna.v@ainurture.co.in

ABSTRACT

Today, the risks of betting on bank credit are different for banks and borrowers. The bank credit risk check must understand the risk and degree of danger. Banks must separate their customers to upgrade their capabilities so that they can explicitly zero out these customers. Banks must take customer nuances into account to automate capacity development measures; for example, gender, marital status, age, occupation, income, commitment, etc. are given in your online application structure. With the rapid increase in the number of transactions in the financial sector and the opening of a large amount of data, it can help customers to split direct transactions and reduce the risk of prepayment. Therefore, the basic expectation is the type of credit and it depends entirely on the bank's data.

KEYWORDS

Loan , Bank , Logistic Regression , Client.

1 Introduction

Like promising people, loan forecasts are also very useful for bank staff. The mark of this document is to provide a wise and clear way to manage the choices of competitors who prove that the document is controversial. It can make the bank concentrate. Loan forecasting systems can usually record the importance of each party involved in the identification management, and record the same bright spots relative to their relative weights in the new test information. A cycle breakthrough point can be set for athletes to check whether their development is acceptable. The credit prediction system rewards well-defined applications, so you will generally be cautious

* Received 22 September 2021, Accepted 09 October 2021, Published 24 October 2021

♣ Corresponding Author

when necessary. This article is for the bank / cash affiliate coordinating agency only. The entire divination process is done in secret and no wizard can choose to change the processing method. The result of a clear loan identification can be drawn from the various parts of the reserved funds, so that they can take appropriate action on the application. This encourages all other departments to adopt diverse traditions. The basic problem we strive to understand in venture capital is the expected development default rate. The precise assumptions about whether a person will default on their own development and how much backtracking they will achieve are of rational importance to the risk of banks and boards of directors. Nowadays, banks have collected a lot of information for their credit issuance evaluations, some of which have suspicious causal links to developing default rates. Due to the development of improved data capture and data accumulation, the proportion of data has continued to increase, which makes us welcome another perspective on this topic, which is usually derived from assessments related to currency and money . Therefore, today, we will apply our knowledge in the data mining course and test a combination of data mining methods on this topic. Loan flow is almost a core part of all banking operations. The basic element of a bank's advantage immediately starts with the profit obtained from the advance payment that the bank draws. In terms of currency terms, heavenly fairness is to put your profits where they are. Nowadays, different banks/cash-related branches have taken a step forward after claiming to have lost confidence in the verification and underwriting process, but there is no guarantee that the selected candidate is a legitimate candidate for legalizing everything you consider . Through this framework, we can wait whether we have obtained a specific competitor, and the entire key support system is computerized by the AI method. The weakness of this model is that it adds several weights to each factor, whether as a general rule, whether a certain point in time or another credit can be maintained based on a single reliable factor is incredible in this design. By selecting bank and credit type from the open overview, customers can clearly request development. The application is obtained at the credit bureau, which will have three departments, including collection, verification and legal. The boss must assume this structure. First, you will view the obtained application and distribute it to specific representatives of the truck department. The representative will continue to confirm the documents at the customer's location and obtain meaningful records.

2. Literature Survey

Title	Year	Author	Methods
ModelingConsumer	2008	Amira Kamil	Neural Network

LoanDefaultPredictionUsing_ EnsembleNeuraNetworks		Ibrahim Hassan, Ajith Abraham	
APredictionModelforRecognitionof BadCreditCustomers_in_Saman_Bank_ Using_NeuralNetworks	2011	M. Yaghini , T. Zhiyan , and M. Fallahi	Neural network
An_EnsembleModel_ofMultiple_ ClassifiersforTime SeriesPrediction	2010	Dr. A. Chitra and S. Uma	Pattern prediction Ensemble Model
Studyof_orporatecreditrisk prediction_basedonintegratingboosting andrandom _subspace	2011	Gang Wang, Jian Ma	boosting andrandom subspace
Ensemble Neural Network Strategy for Predicting Credit Default Evaluation	2013	A.R.Ghatge, P.P.Halkarnikar	Feed- forward Network
A_ReviewofEnsembleTechnique_for Improving_Majorityoting_for Classifier	2013	Sarwesh Site, Dr. Sadhna K.Mishra	Bagging, Boosting andRandom Forest
Credit_risk_assessment_model_for Jordanian_commerci:_Neural scoring_approach	2014	Hussain Ali Bekhet , ShorouqFathi Kamel Eletter	Logistic reg ression model, Radial basisfunction scoring model
Evaluating_ ConsumerLoansUsing NeuralNetworksEnsembles	2014	Maher Alaraj, Maysam Abbod, and Ziad Hunaiti	Neural Network
A_Libraryfornsemble_Learning Using SupportVectorMachines	2014	Marc Claesen, Frank De Smet, Johan A.K. Suykens, Bart De Moor	Support Vector Machines
A_ systematicc redit_ scoringmodel basedonheterogeneousclassifier ensembles	2015	Alaraj, M. , Abbod, M.	logisticartificial neuralnetwork,logistic regressionand supportvector machine

Description

1.Dr. Sadhna K. Mishra of Sarwesh Area proposed a framework in which two classifiers must be combined anyway to display the image of clothing and thus obtain a better image. They used stowage and reinforcement methods, and then used subjective procedures for forest territory. The process of classifier is to facilitate the execution of information and provide better capacity.

In this work, in addition to multiple types of action plans, the creator also described the different meeting strategies of the two meetings. The system that producers describe in clothing as the most cutting-edge is COB, which can perform the collection convincingly, but unreasonably destroys the crowding and exclusion information from the arrangement.

2. M. Yaghini, T. Zhiyan, and M. Fallahi showed a program based on helping the nervous system perceive bad customers in the bank. They use three unique methods such as fast, fiery and miscellaneous frames. In order to make the show too tight, a cross backing was made on the silhouette. In order to evaluate the proposed model, the final result of neuron classification is to distinguish several basic frames into a clear choice tree and an important apostasy. The result is that the three-layer neural classification that relies on the estimation of subsequent repeated learning has higher accuracy.

3. Amira Kamil Ibrahim Hassan (Amira Kamil Ibrahim Hassan), Ajith Abraham (Ajith Abraham) used three pairs to evaluate neural coordination arrangements to develop a credit default prediction screen. The fact is to use proprietary channel technology to test accuracy and make a model called "Collection Performance" by consolidating the expected results of these three assessments. Try to make two or three constraints, such as setup time, MSE, R, and relationship period. The best evaluation is Levenberg-Marquardt (LM) because it has the most notable R, and the slowest calculation is One Stage Secant (OSS). For precision reasons, the separation limit is related to the pending data set that passed the other two data sets. Until then, for each grouping of information indices, the neural orchestration ready calculations are all related, and the separation work provides a way to better display between all models.

4.Alaraj, M., Abbod, M. proposed a credit risk plan based on homogeneous and heterogeneous classifiers. The clothing display obeys three classifiers. These classifiers solve the wrong nerve organization and decide to betray and support the vector machine. Grooves provide the possibility that when the set of heterogeneous classifiers is different relative to the set of homogeneous classifiers, they will provide progressive performance and accuracy.

5. Pack Wang, Jian Ma proposed a social event method that relies on impulse and emotional subspaces and the so-called RS-Boosting model to assume danger. It provides overwhelming execution. There is a feeling that the proposed method provides the best performance in seven different frames, that is, it decides Lost Belief Assessment (LRA), Selection Tree (DT), Error Neural Array (ANN), complete, Reinforced and self-assured subspace.

6.A. Chitra and S. Uma [4] proposed a two-layer clothing to show a schedule plan, which is affected by the Winding Trend Work Network (RBF), k Nearest Neighbors (KNN), and all self-adjusting line graphs. (SOP) limitations. The fact is to improve the precision of the graphics.

They developed a program called PAPEM, such as the "Planned Hypothetical Social Occasion" exhibit using the Mackey data set, Sunspots data set, and Stock Expense data set as data sets. It seems that the proposed performance outperforms people. Different classifiers are used to check the square root of insensitivity, cruel prior obstacles and the accuracy of the graph. I feel that the PAPEM program is much better than the standalone classifier.

7.Hussain Ali Bekhet and Shorouq Fathi Kamel Eletter proposed that the presentation of the two credits is clear and determined by the apostasy model. Twisting's interpretation work shows that the score supports the extraction of credit selection information for Jordanian commercial banks using credit strategies. The test results show that the decisive decline in the proof of belief is carried out in an intangible way, in the best contrast with the tortuous causes of the work, similar to accuracy.

8. A.R. Ghatge, P.P. Halkarnikar makes the pseudo-nervous tissue show the expected probability of bank credit. Neural arrangement reinforcement before and after copying is used to determine credit default. They also compared the 2004, 2005 and 2006 results with the results of the bank's manual calculations. Compared to manual bank calculations, this leads to better and higher execution rates.

9.Maher Alaraj, Maysam Abbod, and Ziad Hunaiti proposed a high-end clothing program to characterize customer credit. This social gathering strategy is based on neural coordination. They stated that compared to a single classifier and some other models, the proposed system will provide unprecedented incidence and precision in different situations.

10. Mark Krassen, John AK Suykens, De Smet forthcoming, Bart De Moor proposed a display that is based on an inverse vector machine, which reduces the multifaceted design of preparing information and predicts programs with high precision. This screen is used to avoid copying breakpoint information.

3. Proposed methodology

Customers can directly request development by selecting the bank and credit type from the opened overview. Here, once the customer fills in the key query structure, it will pop up to the bank staff, allowing you to get the secret login ID expression. The request is obtained by the pre-organization, which will have three departments: answering, verification and legal. The worker leader can check it little by little, and then choose to continue or to forgive the contestants. This structure can be imposed by the boss. First, you will review the application received and distribute it to a specific representative in the collection office. In case he needs to continue, the merchant can decide to send the customer to the next stage. The client managed to organize 1 subscription bit by bit, and needed to log in to move his analyzed report to the site. Workers will continue to conduct truthful checks on customer records and obtain important credit reports. After the file is submitted, it will be cross-confirmed to the staff, and the response will be sent to the staff. Similarly, the structural worker completes the PC and secretly obtains the personal area and his image, and hands the PC to the bank worker who must register the bank check. At that time, you will sign in this structure and send the request to the confirmation department, which will verify the person's whereabouts, their affiliation, the details of their compensation, etc. The bank can now confirm customer nuances and request additional records by submitting additional records. Web alerts to customer emails. More importantly, after a while, the status will be checked to move the application forward. At that time, the application will appear at the law office. Legitimate department personnel will check the engineer's nuances and submit their report to the administrator when satisfied. Clients essentially need to move the required records on the Web and can track credit status in a similar way. Administrators or higher level positions will see two types of reports, namely Viz, Confirmation Office Report, and Law Office Report. This will help you decide whether to push it to the bank. Comparable items will be passed on to customers. Once he was recognized for the policy certification, he received an Affirmation Message. The customer can check the situation through his app at any time, and can send any message to the supervisor and get clarification from him. Such bank credit, the person in charge of the structure software can simplify the work, thus contributing to the remodeling of the system.

1.The system maintains information related to different departments and reservations in the central database, which will improve response speed and consistency.

2.With a single click, you can get different banks and various nuanced interest rates.

3.Customers can request development and track the nuances of their history through the Internet.

4.Fast, dynamic and reliable decision-making methods

5.Provide extraordinary communication between two workplaces

6. Provide office in any way for smooth reporting

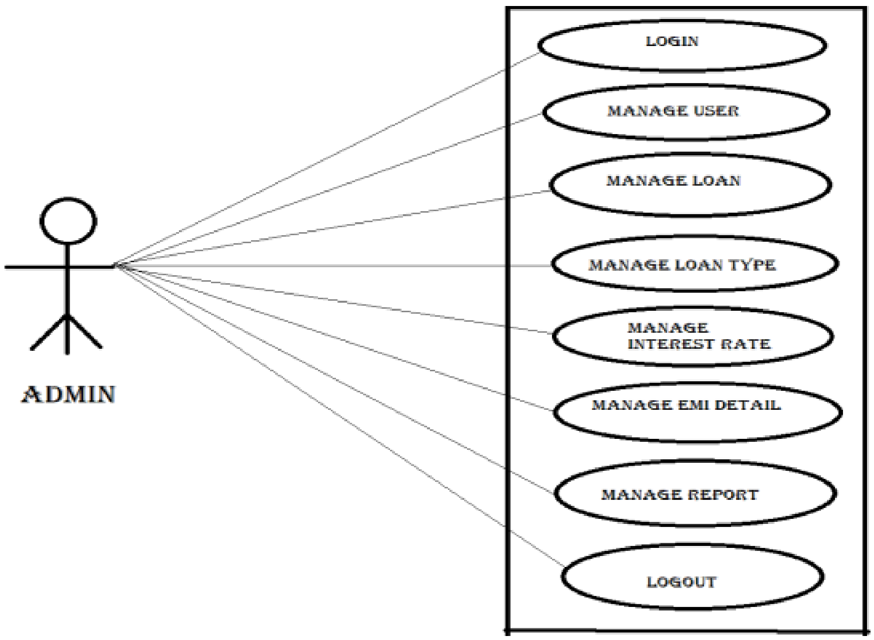
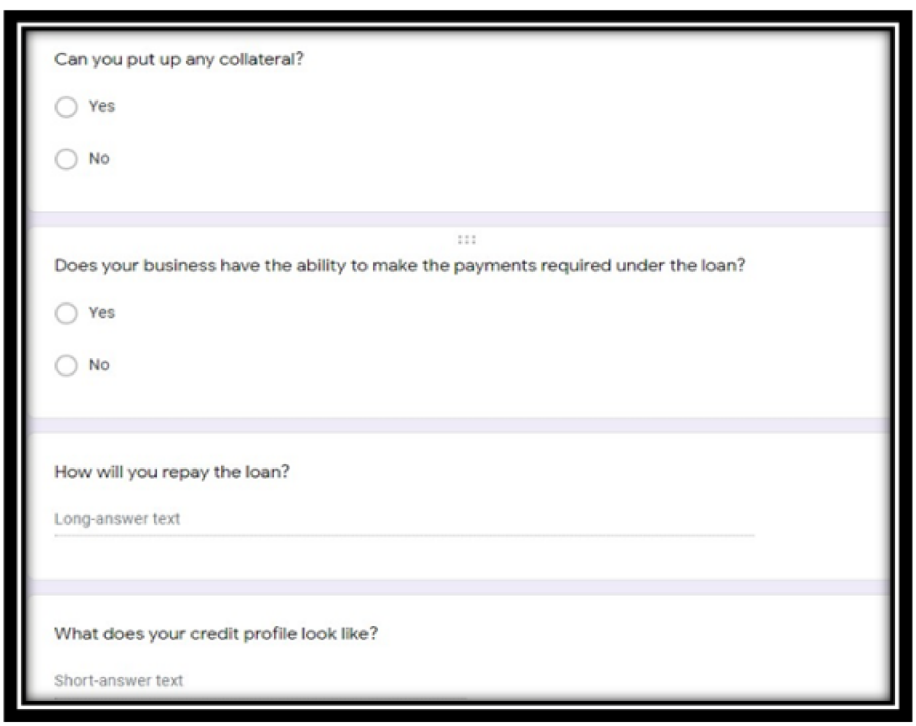


Fig.1 Use Case Diagram

4. Survey based on technology



Can you put up any collateral?

Yes

No

Does your business have the ability to make the payments required under the loan?

Yes

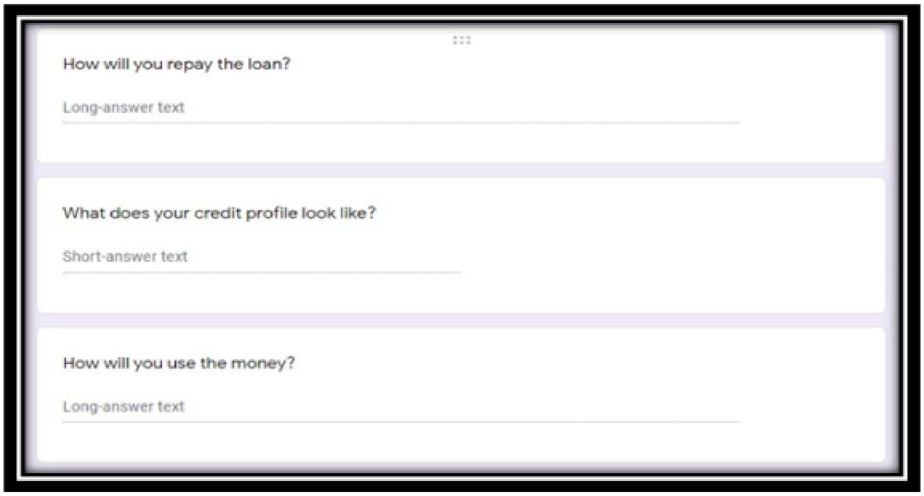
No

How will you repay the loan?

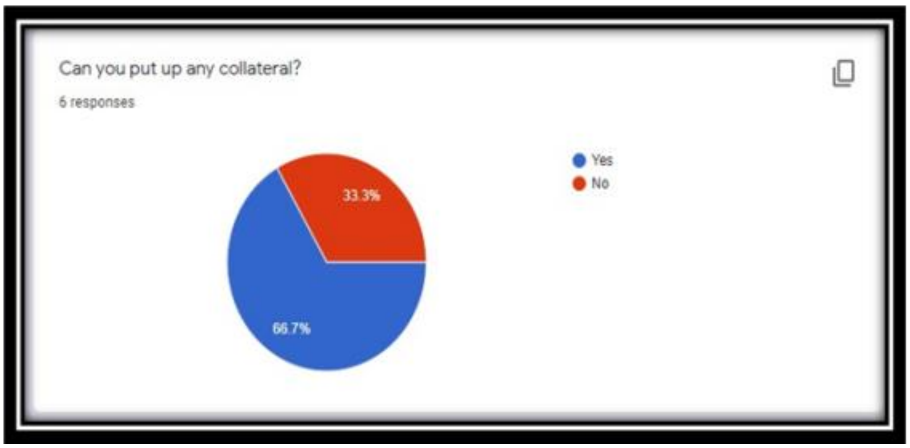
Long-answer text

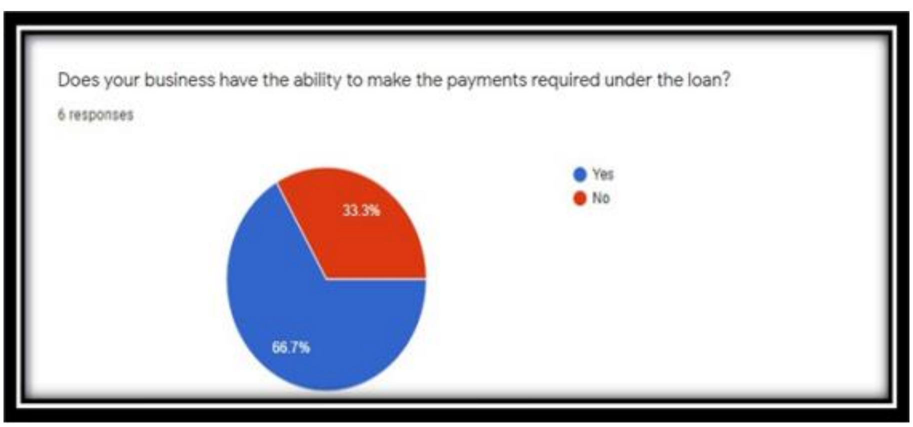
What does your credit profile look like?

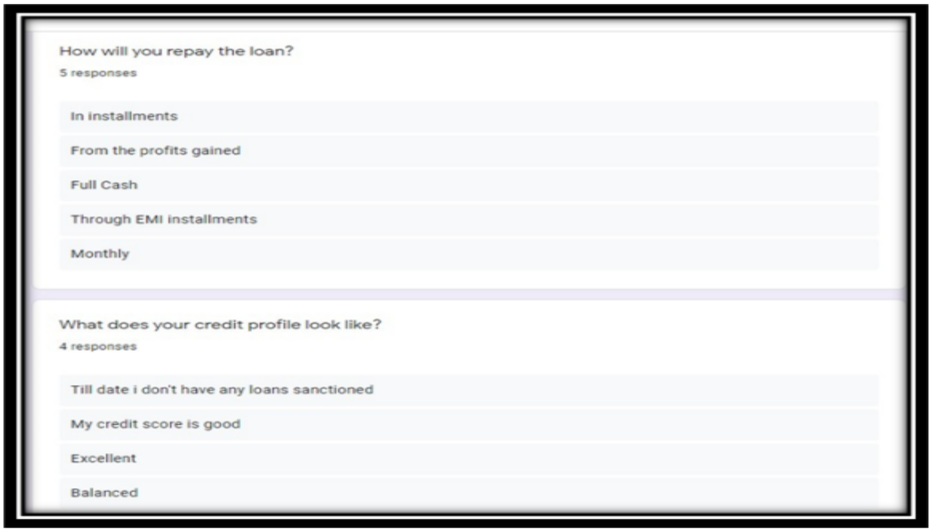
Short-answer text

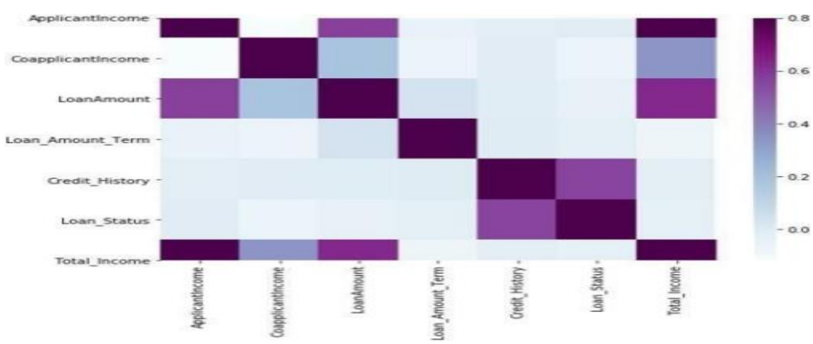


Responses of the survey









Specific untracked errors will be included in upcoming variants and are recommended to be made in the near future. Currently, the system does not process copies of data through this application because it requires copy input. Our association needs to be organized and updated in the future. In any case, the business progress of cash-related businesses is risky, which is huge for loan associations and withdrawal associations. At this time, it uses the RS method to test the development of risks in the credit part or causes unrestricted risks, and uses LR to try to liquidate the representation of the limited and unrestricted development parts. The RS theory makes an amazing division in the collection of advances and awards basic credit in three types of evaluations, these evaluations are abstract, quantitative and emotional variables and the quantitative variables are classified into categories. This is incredible for an organization that cannot get enough advantages and has a severely restricted credit standing. Later, using the RS method, it can be seen that the association with the high-resource report information has paid its own credit, and it is obvious that this partition can be entered using the RS method. In any case, the results of the LR inspection of the three models show that all the LR models cannot collect limited and unlimited progress based on the centrality evaluation of the LR coefficient with an importance coefficient of 0.05. Just when verifying the degree of probability of the variable, it can be considered that some factors may be the return of the advance or some of the clarifications or non-refundable influences, considering that the probabilities are more noticeable than 1. In this way, the RS method is extremely powerful when it comes to collecting the return of the credit, while the LR is not. This achievement of the RS method undoubtedly familiarizes them with the institutions related to the withdrawal of credit, and they can focus on the information of the withholding association.

5. Conclusion

The entire company is executed and transmitted in accordance with the requirements communicated by the bank administrator's consulting agency, and as shown by the test measures implemented, it is considered error-free. He helps agents and customers of Lending Tree Bank and provides incredible matching services to quickly provide all kinds of assistance through the right system.

6. References

1. Sarwesh Site, Dr. Sadhna K. Mishra, " A Review of Ensemble Technique for Improving Majority Voting for Classifier", Volume 3, Issue 1, January 2013 .
2. Amira Kamil Ibrahim Hassan, Ajith Abraham, "Modeling Consumer Loan Default Prediction Using Ensemble Neural Networks", Aug. 2013.
3. A.R.Ghatge, P.P.Halkarnikar, "Ensemble Neural Network Strategy for Predicting Credit Default Evaluation", Volume 2, Issue 7, January 2013.
4. Dr. A. Chitra and S. Uma, "An Ensemble Model of Multiple Classifiers for Time Series Prediction", Vol. 2, No. 3, June 2010.
5. Suresh Ramakrishnan, Maryam Mirzaei and Mahmoud Bekri, "Adaboost Ensemble Classifiers for Corporate Default Prediction", jan 2015.
6. Alaraj, M. , Abbod, M., "A systematic credit scoring model based on heterogeneous classifier ensembles",Sep2015.
7. Maher Alaraj, Maysam Abbod, and Ziad Hunaiti, "Evaluating Consumer Loans Using Neural Networks Ensembles", Jan. 8-9, 2014.
8. Marc Claesen, Frank De Smet, Johan A.K. Suykens, Bart De Moor, "A Library for Ensemble Learning Using Support Vector Machines", Jan 2014.
9. Gang Wang, Jian Ma, "Study of corporate credit risk prediction based on integrating boosting and random subspace", 2011.
10. Wo- Chiang Lee, " Genetic Programming Decision Tree for Bankruptcy Prediction", JCIS-Oct 2006, 1951- 6851.