IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 14, May 2023

Credit Card Approval Prediction using Machine Learning

Sawan Nihatkar[1], Aditya Shahari[2], Tushar Patil[3], Atharva Kulkarni[4], Ms. Suvarna Potdukhe[5]
BE Students, Department of Information Technology, RMD Sinhgad School of Engineering, Pune, India [1,2,3,4]
Professor, Dept. of IT Engineering, RMD Sinhgad School of Engineering, Pune, India [5]

Abstract: *The increased credit card defaulters have forced the companies to think carefully before the approval of credit applications. Credit score cards typically rely on historical data, which may lose their predictive accuracy during periods of significant economic fluctuations. Logistic regression is a commonly used method for credit scoring, as it is well-suited for binary classification tasks and allows for calculation of coefficients for each feature. To enhance ease of interpretation and practicality, score cards often multiply the logistic regression coefficients by a scaling factor (e.g., 100) and round them. Secondly, the performance of the built model is compared with the other two traditional predictive methods such as Boosting, Random Forest, and Support Vector Machines have been introduced into credit card scoring. However, these methods often do not have good transparency. It may be difficult to provide customers and regulators with a reason for rejection or acceptance.*

Keywords: Explanatory Data Analysis, Bivariate Analysis, Multivariate Correlation, S3 bucket model hosting, Model Deployment, Random Forest, Support Vector Model

I. INTRODUCTION

The banking industry's credit lending business has experienced tremendous growth and fierce competition from various credit startups. At the same time, the growth in credit application and consumption has resulted in an increase in losses due to bad loans. Credit loans are loans made by banks or financial organisations to individuals or customers that are repayable at a certain period with or without interest. Credit loans are often granted for a variety of objectives, including personal usage, educational purposes, medical purposes, travel, and commercial purposes. Financial organisations can get insights about applicants' behaviour, consumption habits, default determinants, and attributes using a variety of sophisticated predictive modelling techniques. Numerous studies have been carried out in order to identify the important factors that can affect loan repayment; these studies are important because they help banks maximise profit.

What do banks look for in general?

- The CIBIL Score and Report: This is one of the most essential factors influencing loan acceptance. A strong credit score and report are indicators of excellent credit health.
- Work Status: In addition to an excellent credit history, lenders look for consistent income and work status.
- Account Details: Lenders carefully evaluate instances that have been sued or written off.
- Payment History: Lenders look for any payment defaults or late instances that may cast an unfavourable light on your total report.
- EMI to Income Ratio: When applying for a loan, banks analyse the percentage of your previous debts to your earnings. If your total EMIs surpass your monthly wage by 50%, your chances of loan acceptance are lowered.

In different ways, commercial banks contribute to economic growth. The interest charged on loan is one of the most important revenue streams for any banking or financial organisation. Banks must bear the greatest credit risk in all of their lending. Banks provide a variety of loan solutions to their consumers. Credit cards, on the other hand, are one of the most important lending instruments that any bank could possibly have. Almost all financial institutions across the world are experiencing difficult times and credit risk when it comes to providing loans to their end consumers. Repayments are rarely guaranteed, and the loan frequently becomes a non-performing credit facility (NPL).

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 14, May 2023

Due to the credit risk aspect included in the credit card, banks and financial institutions are severely analysing eligibility for a credit facility before giving facility to the consumer. This procedure comprises verification, validation, and approval and may result in a delay in issuing a facility, which is detrimental to both the applicant and the bank. Credit officers decide whether borrowers can meet the conditions for a facility, and their judgements and projections are never correct. Credit scoring is a classic way of determining a customer's or entity's trustworthiness when applying for a bank credit facility.

By concentrating on the important client (businesses), which are regarded the key engine in the bank's profitability, utilising the best procedures leads to an increase in accuracy and allows banks to have high profitability via customer satisfaction. In this investigation, k-means, enhanced k-means, fuzzy c-means, and neural networks were employed. The utilised dataset has been labelled, and for neural network classification, developing a new label as a goal becomes the major focus of this study, which helps to minimise clustering execution time and deliver the best results with accuracy.

II. LITERATURE SURVEY

Deepak Ishwar Gouda [1] The distribution of loans is practically every bank's fundamental operation. The majority of the bank's assets are directly derived from the profit gained on loans provided by the bank. The primary goal of banking environment is to put their funds in safe hands wherever they are. Today, many banks/financial companies approve loans after a lengthy process of verification and validation, but there is no guarantee that the chosen applicant is the most deserving of all applicants. We can forecast if a given applicant is safe or not using this approach, and the entire feature validation process is automated using machine learning techniques. Loan Prediction is extremely beneficial to both bank employees and applicants. The goal of this Paper is to give a quick, uncomplicated approach to choose qualified applicants. It may give the bank with specific benefits. The Loan Prediction System can automatically determine the weight of each feature involved in loan processing, and the same characteristics are processed with respect to their associated weight on new test data. A time restriction might be imposed for the applicant to determine whether or not his or her loan can be approved. The Loan Prediction System allows you to skip to a single application and check it on a priority basis. SVM, on the other hand, employs a statistical learning model for prediction categorization. To assess the suggested technique, a dataset from the UCI repository with 21 characteristics was used. Experiments revealed that, rather than independent classifier performances (NB and SVM), the integration of NB and SVM resulted in an effective loan prediction classification. During the data analysis, the following major factors were concentrated: annual income versus loan purpose, customer trust, loan tenure versus delinquent months, loan tenure versus credit category, loan tenure versus number of years in current job, and chances for loan repayment versus house ownership.Finally, the current work resulted in inferring the constraints on the customer who is applying for the loan, followed by a prediction regarding the repayment. Furthermore, the results revealed that customers preferred to take out short-term loans over long-term loans.

Ms. Kathe Rutika Pramod [2] Small loans are an important part of our daily lives because they allow aspiring entrepreneurs to get started on ideas that could grow into businesses; they allow curious students to afford higher education that would otherwise be unavailable without a stable income; and, most importantly, they allow ordinary people who have no friends or relatives for support to obtain short-term financial assistance and get back on their feet to fight for the American Dream. However, as with any loan, there is the possibility of default. Default is a financial phrase that describes the failure to satisfy a loan's legal obligation - paying back the principal and interest. It's a widespread issue in the financial industry and one of the key dangers of lending. Overall, default is a fact of life, and most financial institutions have a well-established practice for mitigating its impact and absorbing the loss. But what if, instead of a single bank, the loan is made up of cash contributed by several investors? Lending Club is one of several peer-to-peer lending companies that have contributed to this unusual predicament. In simple terms, a peer-to-peer lending organization operates as a middleman between borrowers and investors. The firm develops a platform on which borrowers may generate modest unsecured personal loans, and investors can search for these loans and choose which ones to invest in. There is little question that Lending Club already has a process in place for approving loans submitted on their website. This study will investigate the process and outcome of developing a new machine learning model that can forecast loan default; however, the model will focus on minimising the total loss in investment of bad loans in order to reduce the burden placed on individual investors. In addition, the article will investigate privacy-preserving mechanisms for sensitive information

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 14, May 2023

obtained from the borrower's credit report. The ultimate purpose is to analyse a simpler version of RAPPOR (Randomised Aggregately Privacy Preserving Ordinal Response) to see if data hashed by this technique can still be used to predict loan default as indicated.

E. Chandra Blessie [3] Almost every bank's business strategy revolves around finance raising and lending for real estate, consumer, mortgage, and commercial loans. The primary source of credit risk is lending money to unsuitable clients. The majority of the bank's assets are obtained directly from the profits made on its loans. However, banking companies face a dual challenge in distinguishing possible deliberate defaulters from applicants, as well as the biased nature of a few bank employees who have been at the behest of developers of defaulting companies for many years. The basic purpose of the banking community is to invest their cash securely. Many NBFCs and banks now approve loans after thorough verification and authentication. NB and Support Vector Machines (SVM) approaches were used to develop a loan prediction model. Nave Bayes is an independent speculation technique that incorporates probability theory in data categorization. SVM, on the other hand, employs a statistical learning model for prediction categorization. To assess the suggested technique, a dataset from the UCI repository with 21 characteristics was used. Experiments revealed that, rather than separate classifier performances (NB and SVM), the integration of NB and SVM resulted in an efficient categorization of loan forecasts. The loan sanctioning prediction process is based on the NB approach combined with the K-Nearest Neighbour (KNN) and binning algorithms. Income, age, occupation, current loan with duration, amount, and approval status were the seven criteria examined. (1) Pre-processing (managing missing values with KNN and data refining using binning technique), (2) Classification using NB approach, and (3) Frequent dataset updates result in adequate improvement in the loan prediction process. Experimentation led to the conclusion that integrating KNN and binning algorithms with NB resulted in enhanced loan sanctioning process prediction. Using the R programme, I proposed a risk analysis approach for authorizing a loan for the clients. Data selection, pre-processing, feature extraction and selection, creating the model, prediction, and assessment are among the modules. The dataset for this method's assessment was obtained from the UCI repository. The pre-processing operation contains the following sub-processes to fine-tune the prediction accuracy: identification, ranking, and removal of outliers, elimination of imputation, and balancing of dataset via proportionate bifurcation related testing and training process. Additionally, the feature selection process improves prediction accuracy. When tested, the DT model had a prediction accuracy of 94.3%.

Yu Li [4] The scoring method based on logistic regression algorithm has the following advantages: (1) simple algorithm and mature technology; (2) robust estimation of probability under given data conditions; (3) strong explanatory power of variables and models; and (4) easy detection and deployment of models. But, at the same time, the traditional credit scoring system has some flaws: (1) The model has a restricted number of variables and may be exploited by fraudsters. (2) the logistic regression algorithm must fulfil certain assumptions, but the actual business may not meet the related assumptions; (3) the logistic regression method's differentiation capacity is difficult to enhance. Other supervised learning methods in machine learning, such as neural networks, nearest neighbor methods, and support vector machines, have been rapidly developed over the last 40 years, in addition to the logistic regression algorithm [1]. Logical regression dates back to the 1950s. Three traditional decision tree implementations. In this article, the observation period is defined as the most recent credit report received within two months of the application date. The performance term is 15 months from the date of application. There are two types of response variables: (1) poor samples are 90 days or more late within 15 months; (2) excellent samples are no more than 89 days late within 15 months. To assure the predictive modelling impact and the speed of algorithm operation, stratified sampling was used to provide model samples, which included 45,000 excellent samples, 5,000 poor samples, and a 9:1 good sample to bad sample ratio. The credit risk prediction using machine learning is investigated in this study. When the model discrimination, model interpretability, and model stability of the logistic regression model and the XGBoost model are compared, it is clear that the XGBoost model has significantly higher model discrimination and model stability than the logistic regression model, which can effectively improve the identification ability of personal fast credit risk.

Archana Gahlaut [5] Credit allows an individual or corporation to "buy ahead of ability" or "desire to pay." Banks make loans to people based on their requirements in the agricultural, industrial, and commercial sectors. Furthermore, when people use their brilliant minds and entrepreneur skills in the presence of credit, it results in overall economic growth, thereby strengthening the country's economy. With today's developing country growth, it might be risky for banks to

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 14, May 2023

extend credit to all of its clients without knowing whether or not they will be able to repay it on time. The same might be a huge problem for all banks nowadays, since the overall loss will be enormous if clients do not return credit loans with interest on time, resulting in bankruptcy. We offered various data mining models to lessen the probability of failure. These models will aid in determining a customer's ability to repay credit loans on time by using credit scoring and categorizing them as a 'Good credit'- customer has a good score and no faulty or defaulter past credit records- or a 'Bad credit'- customer has a bad score and may have faulty past records. Banks will be able to provide decent credit, which will eventually result in a profit in their annual income. Many studies in the banking and insurance analytics sectors have discussed related issues within the data mining framework. For example, Jin et al. used a data-driven approach and a data mining approach to predict loan risk and compared data mining models: decision trees, support vector machines, and neural networks, using a 10-fold cross-validation approach and a high average percent hit ratio to demonstrate the better prediction. To assess the quality, a cumulative lift curve analysis is performed. [1] The Support Vector Machine fared the best.

ZHANG Lei-lei [6] Credit scoring has piqued the curiosity of many researchers in the literature. The credit rating manager frequently analyses the consumer's credit based on intuition. The manager, on the other hand, can precisely analyse the applicant's credit score with the help of the credit categorization model. Support Vector Machine (SVM) classification is a recent research topic that effectively addresses classification issues in a variety of disciplines. In order to create a model with more explanatory power, this article applies support vector machines (SVM) to the problem. For the Australian and German credit datasets from UCI, we used backpropagation neural network (BNN) as a benchmark and obtained prediction accuracy of around 80% for both BNN and SVM methods. Credit scoring is a way of estimating the risk of credit applications. In general, it combines statistical methodologies and historical data to provide a score that financial organizations may use to assess the risk of credit applications.[1] Credit scoring models have been widely employed for credit admittance evaluation due to the fast expansion of the credit business. Several quantitative methods for credit admission decision making have been developed over the last two decades. Credit scoring algorithms are designed to categories applicants as either approved or refused based on their factors such as age, income, and marital status. [2-4] lending officers are faced with the challenge of growing lending volume without increasing their risk of default. SVMs evolved from statistical learning theory, with the goal of addressing only the problem of interest without having to tackle a more complex problem as an intermediary step.

SVMs are founded on the notion of structural risk minimization, which is closely connected to regularization theory. [15-17]

This concept includes capacity management to prevent over-fitting and so provides a partial solution to the bias-variance trade-off quandary. The mathematical programming approaches and kernel functions are two critical components in the implementation of SVM. Rather than addressing a non-convex, unconstrained optimization problem, the parameters are determined by solving a quadratic programming problem with linear equality and inequality constraints. Because of the versatility of kernel functions, the SVM may explore a wide range of hypothesis spaces.

Mohammad Ahmad Sheikh [7] Banks have numerous goods to sell in our banking system, but the major source of income for each bank is its credit line. As a result, they may profit from the interest on the loans they credit. A bank's profit or loss is heavily influenced by loans, namely whether clients repay the loan or fail on it. The bank can lower its Nonperforming Assets by forecasting loan defaulters. As a result, research into this phenomenon is critical. Previous research in this age has revealed that there are several techniques for studying the subject of loan default control. Logistic Regression models have been run, and various performance indicators have been obtained. The models are compared using performance metrics such as sensitivity and specificity. The final findings demonstrated that the model produces diverse outcomes. Model is marginally better because it includes variables (personal attributes of a customer such as age, purpose, credit history, credit amount, credit duration, and so on) other than checking account information (which shows a customer's wealth) that should be taken into account to correctly calculate the probability of loan default. As a result, using a logistic regression technique, the ideal clients to target for loan giving may be simply identified by analysing their chance of loan default. The suggested model predicts whether a bank will give a loan to a customer. Classification is the goal for constructing the model, hence Logistic Regression with a sigmoid function is employed to create it. Pre-processing is the most time-consuming component of the model, followed by Exploratory Data Analysis, Feature

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 14, May 2023

Engineering, and Model Selection. Feeding the model with two distinct datasets, followed by the model. Logistic regression is a statistical machine learning technique/algorithm used to categorize data by analysing outcome variables on extreme ends and attempting to draw a logarithmic line that differs between them. Logistic Regression may be used to make predictions in this manner.

Md. Golam Kibria [8] The growing number of credit card defaulters has compelled businesses to exercise caution before approving credit applications. Credit card firms typically use their discretion to choose whether or not to offer a credit card to a consumer who meets specific requirements. Some machine learning algorithms were also used to help make the decision. The primary goal of this study is to develop a deep learning model based on UCI (University of California, Irvine) data sets that can aid in credit card acceptance decisions. Second, the created model's performance is compared to that of two more standard machine learning algorithms: logistic regression (LR) and support vector machine (SVM). Our findings reveal that our deep learning model performs somewhat better overall. The emergence of the internet has resulted in a considerable increase in credit card usage. It is becoming one of the most popular payment options. Credit card fraud is expanding at an alarming rate as the global economy grows [1]. It is also clear that the number of credit card defaulters has grown dramatically. As a result, credit card companies are becoming more cautious when giving credit cards to clients. Furthermore, the downturn of financial institutions in the United States and Europe during the subprime mortgage crisis in the United States and the European sovereign crisis in Europe has raised concerns about proper risk management [2]. As a result, scholars and practitioners have paid close attention to these difficulties. To handle credit card-related challenges, a variety of statistical and machine learning approaches have been developed (see [1]-[7]). It has been discovered that machine learning approaches outperform other traditional statistical techniques when it comes to credit rating [8–11]. Deep learning, in particular, is a popular and accurate classification approach that outperforms other machine learning models (for example, logistic regression (LR), linear discriminant analysis (LDA), multiple discriminant analysis (MDA), k-nearest neighbour (k-NN), decision trees, and so on). [12]. Specifically, the DL achieves the highest F1-measure score of .886, indicating the overall performance of the model, based on the F1-measure. The F1-measure value for SVM is.863, whereas it is.861 for LR. These two methods gave nearly identical F1-measure results. The SVM beat the other two algorithms in terms of false positive rate, with 12.80% for SVM, 16.10% for LR, and 16% for DL. Based on all accuracy indicators except FP in Table V, we can infer that the deep learning model outperforms the other two models.

Ambika and Santosh Biradar [9] People desire to apply for loans through the Internet as data volumes increase due to banking sector digitalization. Artificial intelligence (AI), as a common approach to information exploration, is gaining popularity. Individuals from diverse industries are using AI calculations to solve problems based on their sector knowledge. Banks are experiencing substantial difficulties in loan approval. Every day, there are several apps that bank staff must maintain, and the likelihood of errors is great. Most banks benefit from loans, but selecting eligible consumers from a large number of applications is hazardous. A single error might result in a significant loss for a bank. Loans have simplified our lives by giving us with financial leverage that goes beyond our wages. Loans, whether Credit Card, Home Loan, Personal Loan, or Auto Loan, are credit granted to us by lenders based on certain critical characteristics. Obtaining a loan in India, on the other hand, might be a time-consuming procedure for the uninitiated, but not for those with a decent credit score. Banks use your CIBIL Score and Report to analyze your credit history and credit eligibility whenever you ask for a loan. The higher your score, the more likely your loan application will be accepted. This application is operational and meets all Banker criteria. This component is easily pluggable into a variety of different systems. It functions properly, meets all banking criteria, and may be linked to a variety of different systems. There were several computer faults, content problems, and weight correction in computerized forecast systems. In the near future, banking software might be more dependable, accurate, and dynamic, and it could be integrated with an automated processing unit. There have been several incidents of computer glitches, content mistakes, and most importantly, the weight of features has been addressed in automated prediction system to make it more secure, dependable, and dynamic weight adjustment. The system is trained using old training datasets, however future software may be designed such that new testing data can be included in training data after a certain period of time. Machine learning aids in understanding the elements that have the greatest influence on certain results. Other methods, like as neural networks and discriminate analysis, can be employed alone or in combination to improve prediction reliability and accuracy.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 14, May 2023

Aboobyda Jafar Hamid [10] Customer segmentation and profitability, high risk loan applicants, predicting payment default, marketing, credit analysis, ranking investments, fraudulent transactions, optimizing stock portfolios, cash management and forecasting operations, most profitable Credit Card Customers, and Cross Selling are just a few examples of how data mining can be used in the financial sector. When seeking to borrow money, there are many different sorts of loans to choose, and it's critical to understand your alternatives. Loan categorization is the process of evaluating loan collections and grouping or grading loans based on perceived risk and other loan attributes. Loans come in a variety of forms, including: Open-ended loans are loans that can be extended indefinitely. The most well-known kinds of open-ended loans are credit cards and lines of credit. You have a credit limit that you may use to purchase items with one of these two sorts of loans. Your available credit will drop whenever you may purchase automatically. As you spend, your cash on hand grows, allowing you to use the credit more and more. Closed-ended loans cannot be borrowed again once they have been repaid. When you make payments on closed-ended loans, the loan balance decreases. However, you do not have any existing credit to use on closed-ended loans. Secured loans are those that are backed by an asset. In the event of loan default, the lender may seize the asset and utilize it to satisfy the amount. Secured loan well-being rates may be lower than unsecured loan well-being rates. Before you may get a secured loan, the asset may need to be examined. Unsecured loans may be more difficult to get and have higher concern rates. Unsecured loans rely only on your credit history and income to fulfil the lending conditions. If you fail to repay an unsecured loan, the lender must exhaust collection options such as debt collectors and file a claim to recover the payment. Prediction and description are the two most significant aims for data mining. Prediction entails utilizing certain variables in a data set to forecast the unknown values of other variables. Description focuses on identifying patterns in data that can be comprehended by humans. Data mining is the process of identifying hidden patterns in massive amounts of data in order to make write judgements. The generated information must be novel, not obvious, relevant, and applicable in the field where it was obtained. It is also the technique of obtaining usable information from unstructured data. Data Mining is a fascinating and important field of study that aims to extract information from massive amounts of amassed data sets. Data mining is growing more popular in the banking industry since there is a need for effective analytical methodologies for finding unknown and important information in bank data. Skills and expertise are crucial requirements for completing Data Mining tasks since the success or failure of Data Mining is heavily dependent on the person leading the process owing to the lack of a standard framework.

III. METHODOLOGY

This software predicts whether or not a credit card applicant will be accepted. Each time there is a hard inquiry, your credit score suffers. This programme predicts your chances of approval without hurting your credit score. Applicants who wish to check out if they will be approved for a credit card without impacting their credit score can use this app. Gradient boosting was the final model employed for this assignment. Recall was utilised as a metric.

Why use recall as a metric: Because the goal of this challenge is to reduce the risk of a loan default, the metrics to employ are determined by the present economic situation:

People feel prosperous and employed during a bull market (when the economy is rising). Money is typically inexpensive, and the danger of default is low due to economic stability and low unemployment. Because the financial institution can handle the risk of default, it is not too cautious when it comes to credit. The financial institution can manage a few problematic customers as long as the majority of credit card holders are excellent customers (those who pay back their credit on time and in full). In this instance.

People lose employment and money in the stock market and other investment venues during a bear market (when the economy is declining). Many people have difficulty meeting their financial commitments. As a result, financial institutions tend to be more conservative when it comes to granting credit or loans. The financial organisation cannot afford to extend credit to many consumers who will be unable to repay it. The financial institution would prefer to have fewer good clients, even if it means turning down some good ones. In this instance, high accuracy (specificity) is preferredAll paragraphs must be indented.

It should be noted that there is always a trade-off between precision and memory. Choosing the appropriate measurements is dependent on the task at hand. According to the findings of this experiment, the three most predictive factors in predicting whether an application would be authorised for a credit card are income, family member number, and

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 14, May 2023

employment length. Other factors such as age and employment status are also useful. The kind of residence and automobile ownership are the least relevant characteristics.

When reviewing the application profile, the idea is to pay more attention to the most predictive aspects and less attention to the least predictive features.

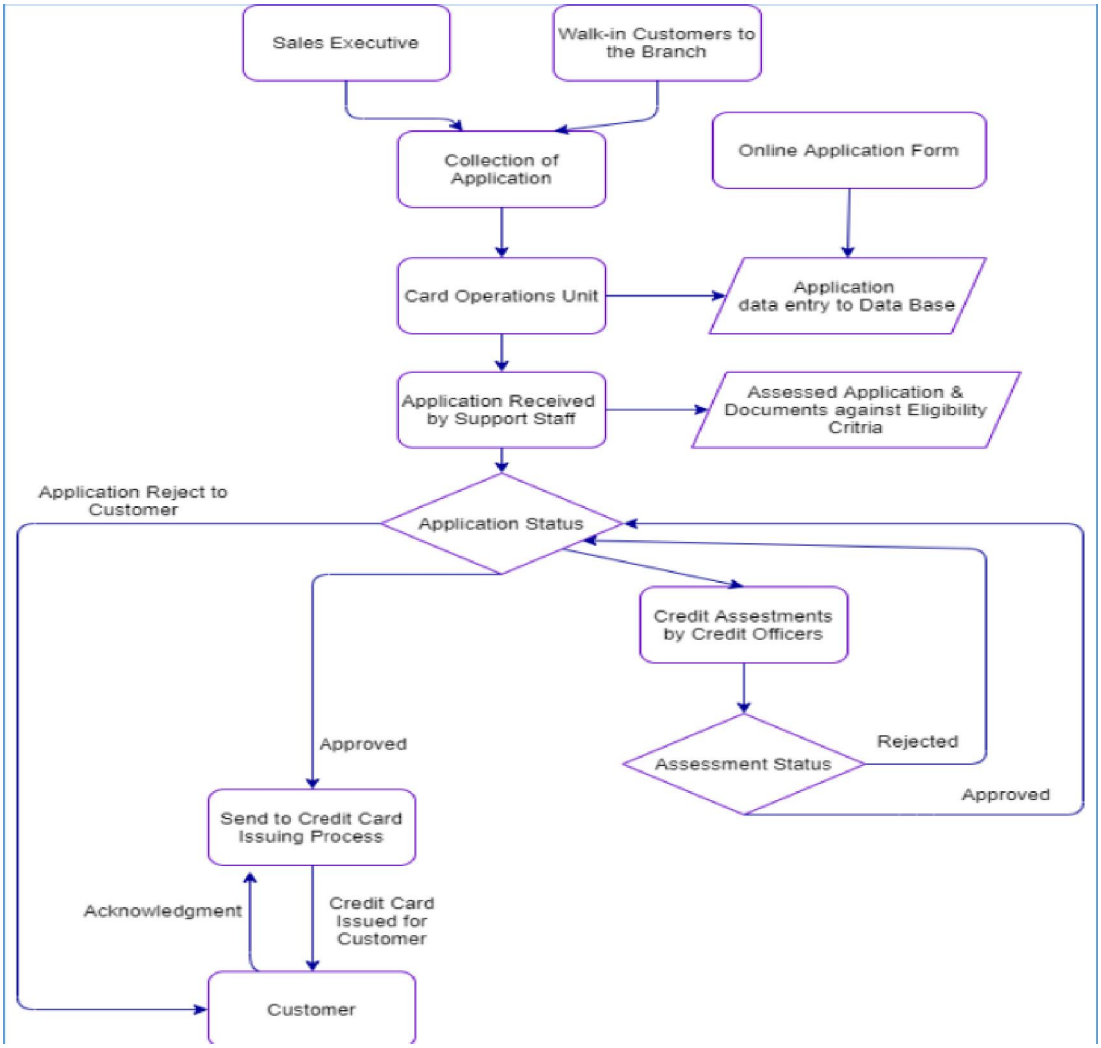

Fig. 1 Credit Card Approval System

[A] Gradient Boosting Algorithm

Gradient Boosting is a machine learning algorithm that works by combining multiple weak learners to create a strong learner. The algorithm iteratively trains decision trees on the residuals (errors) of the previous tree, gradually improving the model's accuracy.

In Gradient Boosting, each decision tree focuses on the residuals of the previous tree, rather than the raw data. This allows the model to learn from its mistakes and adjust its predictions accordingly. The algorithm also uses a technique called boosting, which assigns more weight to misclassified data points in order to correct for errors.

The main advantage of Gradient Boosting is its ability to handle complex, non-linear relationships between variables. It is particularly effective for problems where the input data is high-dimensional and there are many interacting variables.

Gradient Boosting is a powerful machine learning algorithm that can be used to improve the accuracy of credit card approval models. It works by iteratively training decision trees on the residuals of the previous tree, gradually improving the model's accuracy.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 14, May 2023

In the context of credit card approval, Gradient Boosting can be used to predict the probability of default for each applicant. This is done by training the model on historical data that includes information about each applicant's credit history, income, employment status, and other relevant factors. Once the model has been trained, it can be used to make predictions about the creditworthiness of new applicants. The output of the model is a probability score, indicating the likelihood that an applicant will default on their credit card payments. Based on this score, the lender can make an informed decision about whether to approve or deny the applicant's request for credit.

By using Gradient Boosting to improve the accuracy of credit card approval models, lenders can make more informed decisions about who to approve for credit. This can help to reduce the risk of default and improve the overall profitability of the lending business.

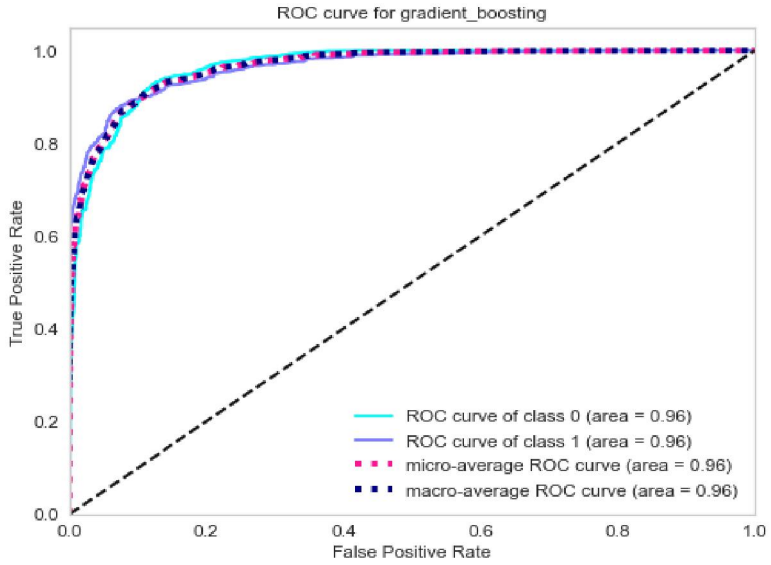

Fig. 2 ROC Curve for Gradient Boosting

[B] Online User Interface

The online UI for credit card approval prediction is designed to provide a user-friendly interface for users to input their personal and financial information in order to predict whether their credit card application will be approved or not. The UI is designed with simplicity and ease of use in mind, with clear labels and instructions to guide users through the process.

The UI begins with a welcome screen, inviting the user to enter their personal information, including their name, age, and contact details. The user is then prompted to provide details about their occupation, education, income, family size, and marital status. The UI also includes fields for employment status and work experience, which are important factors in determining creditworthiness.

The next section of the UI focuses on the user's financial information, including their monthly income, monthly expenses, and any outstanding debt or loans. The UI also includes fields for the user's credit score and credit history, which are key indicators of creditworthiness and play a major role in determining whether a credit card application will be approved.

The final section of the UI asks the user to provide their phone number and email address, which will be used to contact them with the results of their credit card application. The UI also includes a disclaimer and privacy policy to ensure that users are aware of how their personal information will be used.Once the user has provided all of the necessary information, the UI uses a machine learning algorithm to predict the likelihood that their credit card application will be approved. The algorithm takes into account a variety of factors, including the user's personal and financial information, credit score, and credit history, to generate an accurate prediction.

The results of the credit card approval prediction are displayed on the UI, along with an explanation of how the prediction was generated. If the prediction is positive, the user is given information on how to apply for a credit card and what steps they can take to improve their credit score. If the prediction is negative, the user is given advice on how to improve their creditworthiness and increase their chances of getting approved in the future.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 14, May 2023

Overall, the online UI for credit card approval prediction is designed to provide a simple and intuitive way for users to determine their creditworthiness and increase their chances of getting approved for a credit card. By taking into account a wide range of personal and financial factors, the UI provides a comprehensive and accurate prediction that can help users make informed decisions about their credit card applications.

IV. CONCLUSION

The problems with the current system will be reduced by putting the system into practice, and the Credit card approval system will offer a simple, flexible, and easy approach to handling all of the User data and predicting the possibility of credit card approval that can be quickly accessed and readable by higher authorities.

REFERENCES

[1]. D. Kornack and P. Rakic, "Cell Proliferation without Neurogenesis in Adult Primate Neocortex," Science, vol. 294, Dec. 2001, pp. 2127-2130, doi:10.1126/science.1065467.
[2]. M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.
[3]. PhilHyoJinDo ,Ho-Jin Choi, "Sentiment analysis of real-life situations using location, people and time as contextual features," International Conference on Big Data and Smart Computing (BIGCOMP), pp. 39–42. IEEE, 2015.
[4]. Bing Liu, "Sentiment Analysis and Opinion Mining," Morgan & Clay pool Publishers, May 2012.
[5]. AakankshaSaha, Tamara Denning, VivekSrikumar, Sneha Kumar Kasera. "Secrets in Source Code: Reducing False Positives using Machine Learning", 2020 International Conference on Communication Systems &Networks (COMSNETS), 2020
[6]. K. Hanumantha Rao, G. Srinivas, A. Damodhar, M. Vikas Krishna: Implementation of Anomaly Detection Technique Using Machine Learning Algorithms: International Journal of Computer Science and Telecommunications (Volume2, Issue3, June 2011).
[7]. Mohammad Ahmad Sheikh Amit Kumar Goel Tapas Kumar "An Approach for Prediction of Loan Approval using Machine Learning Algorithm"
[8]. Xinxin Li and Lina Wang "Financial Data Center Configuration Management System Based on Random Forest Algorithm and Few-Shot Learning"
[9]. Md. Golam Kibria Mehmet Şevkli "Application of Deep Learning for Credit Card Approval: A Comparison with Two Machine Learning Techniques"
[10]. Aboobyda Jafar Hamid Tarig Mohammed Ahmed "DEVELOPING PREDICTION MODEL OF LOAN RISK IN BANKS USING DATA MINING"
[11]. Archana Gahlaut Tushar Prince Kumar Singh "Prediction analysis of risky credit using Data mining classification models"
[12]. Theodore B. TrafalisHuseyinInce Support Vector Machine for Regression and Applications to Financial Forecasting.
[13]. ZHANG Lei-lei，HUI Xiao-feng，WANG Lei "Application of Adaptive Support Vector Machines Method in Credit Scoring"
[14]. Deepak Ishwar Gouda1, Ashok Kumar A, Anil ManjunathaMadivala, Dilip Kumar R, Dr.Ravikumar "LOAN APPROVAL PREDICTION BASED ON MACHINE LEARNING"
[15]. E. Chandra Blessie, R. Rekha "Exploring the Machine Learning Algorithm for Prediction the Loan Sanctioning Process"
[16]. Ms. Kathe Rutika Pramod Ms. Panhale Sakshi Dattatray Ms. Avhad Pooja Prakash Ms. DapsePunam Laxman Mr. Ghorpade Dinesh B. "Prediction of Loan Approval using Machine Learning Algorithm: A Review Paper".
[17]. Yu Li "Credit Risk Prediction Based on Machine Learning Methods"
[18]. Ambika and Santosh Biradar "Survey on Prediction of Loan Approval Using Machine Learning Techniques"