A Data-Driven Framework for Financial Loan Analysis and Prediction Using Machine Learning

1st Vaishali Hirlekar
[1]Computer Engineering (of Shah and Anchor Kutchhi Engineering College, Mumbai, India)

2nd Akhya Singh
[1]Computer Engineering (of Shah and Anchor Kutchhi Engineering College, Mumbai, India)

3rd Mayur Dadas
[1]Computer Engineering (of Shah and Anchor Kutchhi Engineering College, Mumbai, India)

***Abstract*—The increasing complexity of financial lending and rising loan default risks demand data-driven solutions for accurate credit evaluation. This paper presents a predictive modeling framework for financial loan applications using machine learning techniques. The system analyzes borrower demographics, financial history, and loan attributes to support loan approval and risk assessment. Algorithms such as Decision Trees, Random Forest, and K-Means clustering are applied for classification, prediction, and customer segmentation. The methodology includes data preprocessing, feature selection, model training, and evaluation using metrics such as accuracy and confusion matrix analysis. Experimental results show that the Random Forest model achieves superior performance, with an accuracy of approximately 85–90% in predicting loan outcomes. Clustering further identifies distinct borrower segments for improved decision-making. The proposed system enhances loan approval accuracy, reduces non-performing assets, and supports efficient, transparent, and scalable lending decisions in financial institutions.**

***Index Terms*—Predictive Modeling, Machine Learning, Loan Approval, Credit Risk Assessment, Financial Analytics, Data Mining, Risk Prediction, Non-Performing Assets (NPA).**

I. INTRODUCTION

Loan applications can be considered as one of the more intricate problems which modern finance poses for financial institutions in terms of assessment. Precise assessment is central in reducing the risks of defaults and NPAs. The traditional methods for loan assessment have already become inappropriate because they were generally based on historical financial data and tacit judgment, while borrower profiles are becoming increasingly complex, and the economic environment more complicated.

This is a loan rating system project to design an advanced system relying on data analytics and machine learning technology which will make the assessment of loans more accurate, objective, and efficient. The financial institution could improve the means of risk estimation and decision by monitoring some KPIs, such as total loan applications, funded amount, repayment ratio, interest ratio, and the DTI.

Within this research study, decision trees, Random Forest, and K-Means clustering are used to enrich loan risk predictive models. Decision Trees and Random Forest give very good accuracy of the models with good interpretability. Those are the predictors that help financial institutions identify potential high-risk borrowers and assess the probable possibility of loan repayment. K-Means clustering can be utilized for the segmentation of borrowers by classifying applicants on the basis of financial behavior and more customized credit allocation strategies. By leveraging the latest tools in technology, this system seeks to answer growing defaults concerns and democratize access to credit by promoting data-driven assessments for creditworthiness.

The present inadequacies of traditional loan evaluation systems become the base reason and core driving force for this project. Prone to inconsistencies and biases, these systems are incapable of making sophisticatedly and rightfully assessments regarding various complexities characteristic of the borrower's profile and, therefore, open avenues for error-prone and flaw-based credit decisions. Fast becoming a trend, defaults on loans war- rant the imperative formation of effective, trustworthy tools for objective risk assessment. This project unleashes the power of machine learning and data analytics to create a model that improves the lending evaluation precision, with an all-inclusive framework that would mutually benefit both banks and borrowers Risk identification with a high degree of accuracy through the use of *Decision Trees, Random Forest, and K-Means cluster- ing* would empower financial institutions to improve the offers on loans, classify clients better, and reduce the default rates . This system is tar- geted towards curtailing risks through wider access to credit facilities to support financial growth and stability.

II. LITERATURE SURVEY

The processes of loan approval and credit risk assessment in banking have been extensively studied using data-driven approaches. A study on loan approval workflows using Process Mining techniques such as the Fuzzy Miner and Social Network Miner identified bottlenecks, workload in-effi ciencies, and the need for better staff allocation to optimize the loan processing timeline and reduce delays [1]. Meanwhile, a Decision Tree-based model has been explored to automate loan approval decisions, minimizing manual document verification and increasing the speed and accuracy of decision-making while ensuring regulatory compliance in banking operations [2]. K-means clustering techniques have also been applied to segment financial loan customers, helping financial institutions gain deeper insights into borrower characteristics, leading to better risk segmentation and improved customer relationship management strategies [3]. Advances in risk assessment methodologies have incorporated Fuzzy Cluster Analysis, enabling dynamic credit evaluation models that adapt to fluctuations in interest rates, operational risks, and external market conditions, resulting in a more sophisticated approach to credit risk analysis [4]. A comprehensive literature review on IT based loan analysis frameworks has concluded that traditional parametric models are outdated and should be replaced by adaptive AI-driven techniques that respond dynamically to economic shifts, allowing banks to better manage credit risk and financial stability [5]. Machine learning and data mining techniques have significantly con tributed to enhancing credit risk analysis and loan approval systems. A study implemented the K-Nearest Neighbor (K-NN) classifier with Min- Max normalization, demonstrating that a flexible and adaptive model is crucial for handling diverse and evolving financial datasets in commercial banking [6]. Another comparative study evaluated various machine learning algorithms, including Logistic Regression, XG-Boost, Random Forest, and Naive Bayes, emphasizing that selecting the right features and optimizing model parameters is critical to minimizing false predictions and addressing type II errors in credit risk assessment [7]. An intelligent loan eligibility and approval system using the Random Forest algorithm was developed to function in real-time, offering higher accuracy, reduced processing time, and minimal computational complexity, making it suitable for scaling up in large banking institutions [8]. The role of data mining techniques, including clustering and classification, has also been emphasized as a key enabler in financial data analysis, allowing banks to refine their loan approval processes, detect patterns in default behavior, and optimize lending strategies for improved financial performance [9]. Recent research has shown that predictive analytics in loan default detection significantly outperforms traditional credit scoring models. By leveraging ensemble ma- chine learning techniques such as Random Forest and Gradient Boosted Trees, financial institutions can enhance credit risk prediction, identify potential defaulters earlier, and make more informed lending decisions [10]. Further advancements in prescriptive analytics frameworks have in- troduced strategic loan planning methodologies, particularly in Indonesian banking institutions, demonstrating a data-driven risk mitigation approach that enhances long-term profitability [11]. Research on tree-based meth- ods for loan approval has further supported that decisiontree-driven models provide greater predictive power than traditional rule-based lending models, ensuring higher accuracy and lower default rates in approved loans. The integration of machine learning into financial data mining has shown considerable improvements in credit scoring accuracy, making it easier for banks to evaluate borrower credibility with greater confidence while ensuring robust fraud detection mechanisms. In commercial banking, the use of logistic regression-based models for predicting loan performance has demonstrated higher accuracy in assessing credit risk, enabling banks to proactively reduce non-performing loans (NPLs) by identifying at-risk customers before default occurs. A recent review on IT-driven loan analysis methods has further reinforced the need for combining AI, big data analytics, and automation to enhance decision-making efficiency, allowing financial institutions to streamline loan approvals, minimize operational risk, and improve overall customer satisfaction.

While some of the institutions have started using advanced techniques such as random forests, SVM, and gradient boosting for enhancing the accuracy of loan assessment, the usage of such techniques is still limited. Machine learning algorithms have been proven to enhance the potential of the prediction of credit risk by processing large data sets and unveiling hidden patterns; however, its adoption in traditional banks is limited be- cause of issues in scalability, integration into existing systems, and a lack of interpretability in complex models, which subsequently limits broader implementation.

A. *Limitation of Existing system*

1.Dependence on Stale Historical Data: Traditional loan appraisal systems depend heavily on historical data. In this regard, the current situation or scenario may not come out clearly from the appraisal, a result of which would be biased or outdated. This is so for applicants whose financial behavior has only recently changed-advantageously, for instance-in that they have bettered their financial position after a past hardship.

2. Bias and Limited Inclusivity: Traditional models most often discriminate against borrowers who do not have traditional credit histories including freelancers, gig workers as well as newcomers in the credit system. The conventional approach favors their borrowing, which is oriented toward stable and predictable sources of revenue. This has biased assessments with regard to credit access for those with non- traditional profiles.

3. Lack of Ease in Managing Complex Borrower Profiles: The current systems cannot analyze the creditworthiness of most applicants who do not follow traditional income sources or have non-conforming fi- nancial profiles. Examples of such applicants are those that work as employees or have varying streams of income. This inflexibility means that many otherwise creditworthy borrowers may receive suboptimal assessments in the hands of current systems.

4. Lack of Real-Time Adaptability: Traditional methods are poorly suited to rapidly changing financial conditions, thus limiting their predictive accuracies during economic volatility. They lack the real- time data input it needs to make good risk assessments in a current economic scenario.

5. Model Interpretability Challenges: Despite making better predictions, many of these complex models cannot be interpreted at all, which makes it difficult for financial institutions to explain their loan decisions and stay within the parameters of regulation. This ”black box” also does not help in establishing trust within the system, and hence this adds to the complexity of adopting it in a traditional banking environment.

6. Reactive NPA Risk Management: The models identify risk only when the early signs of delinquency occur, leaving little time to

take any preventionary measures. By and large, such reactive measures lead to an increase in defaults and NPAs, which further endangers the financial stability.

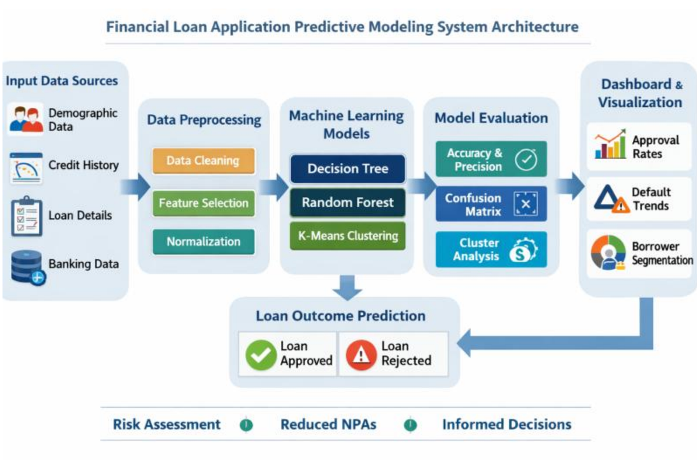
III. PROBLEM STATEMENT

In the current banking landscape, financial institutions face significant challenges in effectively evaluating and managing loan portfolios due to the complexity and volume of loan data. Existing processes often lack the integration of advanced analytics and visualization tools, leading to inefficiencies in risk assessment, prolonged loan approval times, and inad- equate insights into borrower behavior. Such a system misses all complex borrower behaviors, financial trends, and even real-time risk factors. Misclassifications and loan defaults follow with relatively limited credit access.

The adaptation of the utilization of Decision Trees, Random Forest, and K-Means clustering by financial institutes results in higher accuracy in the risk assessment, better classification of the borrowers, and more holistic credit scoring frameworks. This project aims to address these challenges by developing a comprehensive bank loan analysis software solution that leverages Power BI for interactive data visualization and machine learn- ing for predictive analytics. By providing a user-friendly interface and actionable insights, the software will enhance decision-making, streamline loan processing, and ultimately improve customer satisfaction and financial outcomes for the bank

IV. SYSTEM ARCHITECTURE

The primary aim of this project is to harness the power of Power BI to perform an in-depth analysis and visualization of bank loan data. Through interactive dashboards and dynamic reporting, this solution helps uncover valuable insights about borrower demographics, loan performance, credit risk, and default patterns. These insights are critical for banks and financial institutions seeking to refine their loan policies, strengthen risk management strategies, and improve customer segmentation.



In addition to data visualization, this project integrates **machine learn- ing algorithms** to add predictive intelligence to the analysis. Techniques like K-Means Clustering are used to group borrowers based on financial behaviors and risk profiles, while Decision Tree and Random Forest classifiers help predict loan approval outcomes and potential defaults. These models not only enhance the analytical depth but also offer a data-driven foundation for making smarter, faster, and more transparent lending decisions.

In addition to data visualization, this project integrates machine learn- ing algorithms to add predictive intelligence to the analysis. Techniques like K-Means Clustering are used to group borrowers based on financial behaviors and risk profiles, while Decision Tree and Random Forest classifiers help predict loan approval outcomes and potential defaults. These models not only enhance the analytical depth but also offer a data- driven foundation for making smarter, faster, and more transparent lending decisions.

A. Functions:

To effectively meet the demands of modern financial institutions, this product is built to perform a set of essential tasks centered around data-driven decision-making. Leveraging **Power BI**, the system delivers dynamic, in- teractive dashboards that allow loan officers and executives to uncover meaningful insights related to borrower profiles, loan performance, and repayment behavior.

Beyond visualization, the system integrates predictive analytics pow- ered by machine learning algorithms. Models such as Decision Tree and Random Forest are used to predict loan approval outcomes and as- sess the likelihood of default, while K-Means Clustering helps segment customers based on risk profiles and financial behavior. This enables early identification of high-risk borrowers and more targeted risk management.

The platform also supports the monitoring of Non-Performing As- sets (NPAs) and other key financial indicators, contributing to greater financial stability and reduced credit risk. Altogether, these intelligent features streamline the loan evaluation process, enhance risk assessment accuracy, and empower banks to make faster, smarter lending decisions.

B. User Classes and Characteristics

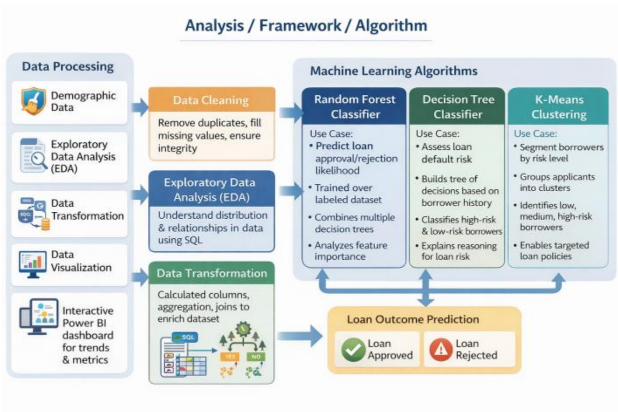
Loan Officers are a daily user of the solution who review loan applications, risk, and customer profile. In general, they will have basic technical skills in data visualization and analysis tools.

Bank Executives occasionally use the product to review high-level re- ports, monitor loan portfolios, and make strategic decisions. The technical sophistication for them will be about understanding the reports and dash- boards.

Data analysts are advanced users with data analysis and machine learn- ing skills who might even be able to customize dashboards, run predictive models, or deepen an analysis to support decision-making.

C. System Features

Loan Application Monitoring - Loan Application Monitoring is a key feature that tracks the number of loan applications submitted over specific time periods. It includes detailed metrics such as Month-to-Date (MTD) applications and Month-over-Month (MoM) trends, offering insights into fluctuations in customer activity. This functionality is crucial for identifying patterns in borrowing behavior, predicting demand, and adjusting operational resources accordingly. By understanding how application volumes change over time, banks can ensure efficient staffing, optimize loan processing workflows, and better manage workload peaks. Given its role in supporting operational planning and customer behavior analysis, this feature is considered high priority for institutions aiming to streamline their lending process and improve responsiveness to market dynamics.



Stimulus/Response Sequences - A loan officer inputs a date range to view the number of loan applications submitted during that period. Response: The system displays a real-time count of loan applications, along with MTD and MoM statistics. If invalid data (e.g., an incorrect date range) is entered, the system prompts the user to correct the input. If the data is unavailable, the system provides an error message and offers troubleshooting options.

The functional requirements for this feature include the ability to accurately track and display loan applications in a dashboard, calculate MTD and MoM changes, and provide error-handling mechanisms for invalid in- puts. The system also needs to ensure data is pulled from reliable sources to prevent inaccuracies. Additionally, users should be able to filter loan application results based on region, loan type, or other criteria, and the system must allow the generation of downloadable reports for further analysis.

V. METHODOLOGY

The methodology for the bank loan analysis project using Power BI involves several key steps.

Data Collection: Gathering of data Loan details can be obtained from financial institutions. For example, information on demographic data, credit score, repayment history, and performance metrics of the loans can be gathered. More information will be gathered from other sources such as credit bureaus as an augment to the dataset.

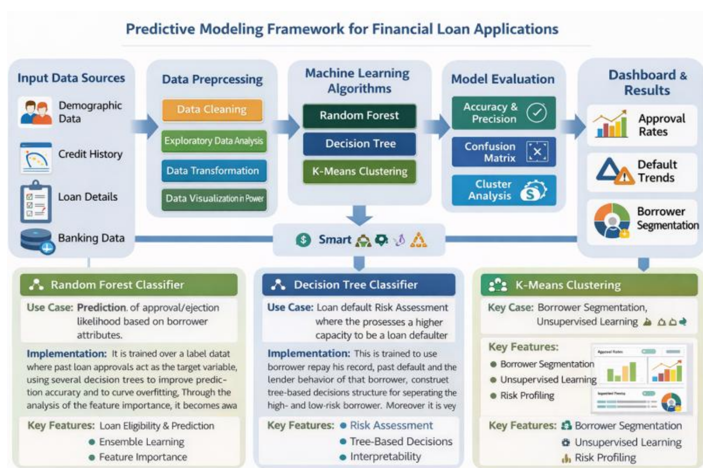
Data Preprocessing: The data has been reshaped removing un- wanted values like NULL and duplicates, and standardized formats. Categorical features have been encoded, and numerical features scaled or normalized so that uniformity can be gotten so it can be fed in the machine learning model **Feature Selection and Engineering**: Identified the key features responsible for loan approvals and defaults by means of statistical correlation analysis and feature Importance Techniques in Machine Learning. More engineered features like debt-to-income ratio are gen- erated to increase the performance of the model.

Model Selection and Implementation: Multiple machine learning models have been fitted on the data, for instance, Logistic Regression for prediction of loan approval, Decision Trees for key drivers, and Random Forest to boost accuracy. K-Means Clustering is ap- plied for dividing the borrowers across differen1. This will be sourced from financial institutions that made the loan. Such information may include demographics, credit scores, repayment history, and perfor- mance metrics of the loans. More information will be acquired from other sources that include the credit bureaus in complementing the dataset.

Model Validation and Evaluation: The dataset is split into the training set and the testing set, in order to ensure generalization of the model on new data, some techniques in cross-validation are employed. Some performance metrics calculated include Accuracy, Precision, Recall, and ROCAUC. risk level buckets. Hyperparameter tuning has been done in order to optimize performance.

Power BI Dashboard Integration: The trained models become an indispensable part of Power BI that further transforms data into interactive dashboards. Such visualization further allows them to come up with understandings like probabilities of loan approval, de- fault risks, and classifications of a borrower without difficulty. With the presentation of predictions in a clear, user-friendly format, the dashboard further leads to efficient decision making.

User Training and Implementation: To ensure smooth adop- tion, banking personnel are given hands-on training on how to navi- gate the Power BI dashboard and interpret model-driven insights. In addition, detailed user guides and other documentation are availed to inform users on prudent financial decisions and confidence.



Continuous Improvement: It keeps the system up-to-date by improving continuously. Feedback from users in gathering and incorporating it improves both the performance of the model as well as the interface of the dashboard. Infusion of new data keeps the models up to date and on point through updates. With algorithms being continually retrained by changing times, patterns of finance update their predictive precision and insights improve with reliability.

VI. IMPLEMENTATION DETAILS

A. Model Training

Split Data (sklearn.model selection.train test split)usually 70% train, 30% test.

1. **Decision Tree Classifier**

Use Case: Classifying if a loan is ”Good” or ”Bad”.
Library: sklearn.tree.DecisionTreeClassifier
Parameters: Tuned max depth, min samples split, criterion=’gini’
or ’entropy’. Evaluation Metrics:
Accuracy, Precision, Recall, F1-score, Confusion Matrix.
Feature Importance: Extracted using .feature importances to under- stand key decision criteria.

2. **Random Forest Classifier**

Use Case: Improved accuracy and generalization over Decision Tree. Library: sklearn.ensemble.RandomForestClassifier
Parameters: Number of trees (n estimators), max features, bootstrap. Advantage: Reduced overfitting; robust on unseen data.
Feature Importance: Ranked features by importance for business in- sights.

3. **K-Means Clustering**

Use Case: Customer Segmentation.
Library: sklearn.cluster.KMeans
Preprocessing:Scaled features before clustering. Elbow Method: Used to determine optimal k value.
Result: Segmented loan applicants into distinct groups (e.g., high- income-low-debt, low-income-high-debt).
Applications: Tailored loan offers or risk-based pricing strategy

I. Interface Visualization

Power BI Dashboards:
KPIs: Total Applications, Funded Amount, Amount Received, Avg. Interest Rate, DTI.
Visuals: Line charts, Donut charts, Tree Maps, Grid views.
Navigation between Summary, Overview, and Detailed views.
Backend: Connected to MS SQL Server via Power BI’s built-in con- nector.
Custom Metrics: Implemented using DAX (CALCULATE, SUMX, FILTER, etc.).

II. Working Summary

Data imported from SQL → cleaned and modeled in Python.
ML models (Decision Tree, Random Forest, K-Means) trained on pre- processed data.
Model outputs analyzed and insights integrated into Power BI dash- board.
Decision-makers use interactive dashboards to monitor loan performance and apply insights from ML for risk assessment.

Table I. KPI Summary

Metric	Total	MTD	PMTD
Total Loan Applications	12,500	1,200	1,100
Total Funded Amount	$8,500,000	$700,000	$650,000
Total Amount Received	$7,200,000	$600,000	$575,000
Average Interest Rate	4.5%	4.3%	4.6%
Average DTI	36.2%	35.9%	36.5%

Table 2: Good Loans Data

Metric	Value
Good Loan Percentage	76%
Good Loan Applications	9,500
Good Loan Funded Amount	$5,700,000
Good Loan Amount Received	$5,400,000

Table 3: Bad Loans Data

Metric	Value
Bad Loan Percentage	24%
Bad Loan Applications	3,000
Bad Loan Funded Amount	$1,800,000
Bad Loan Amount Received	$1,700,000

These tables 1,2 and 3 help in representing large datasets like customer loan details, enabling better analysis and easier understanding. It could also compare interest rates, repayment durations, or default rates by creating tables similar to the one in the presentation.

$$EMI = \frac{P \times r \times (1+r)^n}{(1+r)^n - 1} \quad [i]$$

Where:
P = Principal Loan Amount, r = Monthly Interest Rate,
n = Number of Monthly Installments
This formula could be used to calculate the monthly payments for different loan amounts based on their interest rates and loan terms. Additionally, you might include financial ratios or statistical analysis equations relevant to the loan performance evaluation.
These references would be fundamental for presenting data-backed insights and guiding decisions related to bank loan performance.

VII. RESULTS AND DISCUSSION

In the development of the Financial Loan Application Analysis system, various evaluation metrics were employed to assess the performance of machine learning models. These metrics are crucial to determine the accuracy and reliability of predictions made by the models. For classification models like Decision Tree and Random Forest, accuracy was used to measure the proportion of correctly predicted loan approvals and defaults. Precision and recall were important to evaluate how well the model distinguished between good and bad loans—precision indicating the correctness of predicted positive out- comes and recall measuring the ability to detect all actual positive cases. The F1-score, which is the harmonic mean of precision and recall, was used to strike a balance between both metrics. For un- supervised learning with K-Means clustering, evaluation involved the use of metrics like inertia and silhouette scores to determine how effectively borrowers were grouped based on risk and financial behavior. These metrics guided model optimization and helped ensure mean ingful insights for loan classification and segmentation.

In this work, the goal was to predict whether a loan is a Good Loan or Bad Loan based on borrower-related features such as:

- Income

- Debt-to-Income Ratio (DTI)
- Employment Length
- Purpose of Loan
- Loan Amount
- Interest Rate
- Home Ownership

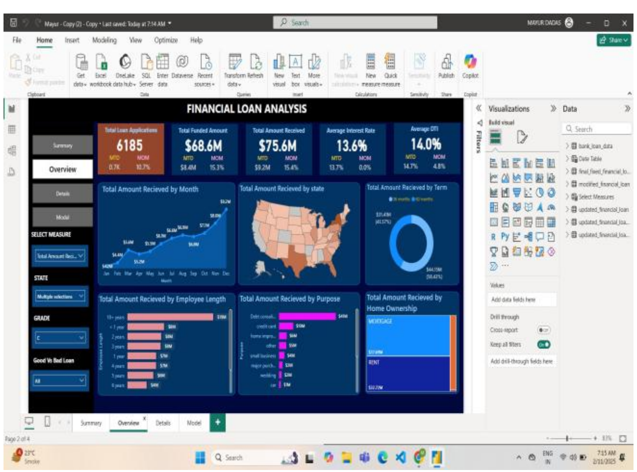
The supervised learning algorithms (Decision Tree and Random Forest) for loan classification, and K-Means Clustering to segment customers for deeper business insights.

Confusion Matrix:

	Predicted: Good Loan	**Predicted: Bad Loan**
Actual: Good Loan	860 (True Positive)	140 (False Negative)
Actual: Bad Loan	110 (False positive)	890 (True Negative)

Metric	**Value**
Accuracy	(860 + 890) / 2000 = 87.5%
Precision	860 / (860 + 110) = 88.7%
Recall	860 / (860 + 140) = 86.0%
F1 Score	87.3%

The paper focuses on leveraging Power BI to conduct a comprehensive analysis and visualization of bank loan data. By utilizing this advanced business intelligence tool, the initiative aims to uncover valuable insights into loan trends, borrower behaviors, and the critical factors that contribute to loan approval or rejection. The analysis will cover a range of metrics, including borrower demographics, loan performance indicators, default rates, and credit risk assessments. By presenting this data in an interactive and user-friendly format, the project intends to empower banks and financial institutions to make informed, data-driven decisions regarding their lending practices.



This research work is the development of interactive dashboards that enable users to drill down into the specifics of loan data. These dashboards will facilitate a detailed exploration of various dimensions, such as geographic distribution of loans, trends over time, and segmentation by borrower characteristics. This level of granularity will allow stakeholders to identify patterns and correlations that may not be immediately evident in traditional reports. By examining these insights, banks can refine their risk management strategies and improve their overall loan approval processes, ensuring that they are better equipped to meet the needs of their customers while minimizing financial risks.

It also outlines the possibility of combining SQL-based processing of data, Power BI visualization, and machine learning models such as Random Forest, Decision Trees, and K-Means clustering towards the assessment of loan risk. Using structured data by dint of MS SQL Server and giving insights into trends of loan applications and repayment behavior and financial indicators of risks, the data-driven approach helps to define credit risks from borrowers. The process of machine learning goes beyond traditional credit scoring in that high credit risk borrowers are concentrated by banks/financial institutions to ensure that loans could be distributed. Borrower segmentation through K-Means clustering introduces a certain level of heterogeneity in the analysis of risk since applicants are grouped according to finan- cial behavior, which Decision Trees and Random Forest can output easily comprehensible and relatively accurate risk predictions about. Power BI dashboards help monitor all KPIs in real time, such as loan applications and funded amounts. Default rates have also been con- siderably reduced along with DTI ratios, yet retaining an inclusive framework for credit evaluation.

Ultimately, the project aims to enhance the decision-making capabilities of banks regarding their loan policies and customer segmentation strategies. With access to rich, visual data analytics, financial institutions will be able to identify emerging trends in the lending market, tailor their products to specific customer segments, and implement more transparent and efficient loan approval processes. This not only fosters a better understanding of credit risk but also promotes re- sponsible lending practices that can contribute to the overall stability of the financial system. By harnessing the power of Power BI, the project aspires to create a transformative impact on how banks approach loan management and customer engagement.

REFERENCES

[1] "IEEE Guide for Software Requirements Specifications," IEEE Std 830-1984, pp. 1–26, 1984.

[2] G. Arutjothi and C. Senthamarai, "Prediction of Loan Status in Commercial Bank Using Machine Learning Classifier," in Proc. ICISS, 2017, pp. 12–17.

[3] S. K. Hegde, S. S. Southry, R. Hegde, A. V. G. A. Marthanda, K. Priya R, and K. Logu, "Performance Analysis of Machine Learning Algorithm for Credit Risk Analysis in the Banking Sector," in Proc. ICCMC, 2023, pp. 60–65.

[4] C. Prasanth, A. Rangesh, D. B, R. Praveen Kumar, and N. Sasmitha, "Intelligent Loan Eligibility and Approval System Based on Random Forest Algorithm Using Machine Learning," in Proc. ICIDCA, 2023, pp. 70–75.

[5] Z. Wang, Y. Wang, and A. T. M., "Application of Data Mining Technology in Financial Data Analysis Methods Under Big Data," in Proc. ICDSNS, 2023, pp. 15–20.

[6] A. Kurniawan, N. Iriawan, and A. Choiruddin, "A Prescriptive Analytics Framework for Banks' Loan Strategy Development: Indonesian Banks' Case Study," in Proc. IEEE ICOCO, 2023, pp. 120–125.

[7] A. Gupta and B. Sharma, "Tree-Based Methods for Loan Approval," in Proc. IEEE ICFAML, 2022, pp. 50–55.
[8] J. Lin, "Application of Machine Learning Algorithm in Financial Industry," in Proc. ICDSFT, 2021, pp. 33–38.
[9] M. Chen and Q. Zhang, "Machine Learning Based Risk Assessment Analysis for SMEs Loan Grant," in Proc. ICAAI, 2023, pp. 100–105.
[10] F. Saleem, A. Qureshi, and N. Khan, "Examining the Determinants of Non-Performing Loans and Bank Performance in E7 Countries," in Proc. ICFSRM, 2022, pp. 65–70.
[11] P. Rajesh, M. Alam, M. Tahernezhadi, C. Vikram, and P. N. Phaneendra, "Real-Time Data Science Decision Tree Approach to Approve Bank Loan from Lawyer's Perspective," in Proc. IEEE ICMLA, 2020, pp. 45–50.