

REAL TIME CREDIT RISK MONITORING WITH ML

¹ D. Shirisha, ² Koyagur Sai Kumar, ³ Bhathini Srishanth, ⁴ Chandragiri Akash, ⁵ Chaganti Ajith Kumar

¹ Assistant Professor, ^{2,3,4,5} Students

Department of ECE

Siddhartha Institute of Technology & Sciences, Narapally

shirishadubbaka.ece@siddhartha.co.in, 24tq1a0428@siddhartha.co.in, 24tq1a0403@siddhartha.co.in,

24tq1a0407@siddhartha.co.in, 24tq1a0406@siddhartha.co.in

ABSTRACT

Credit risk is the possibility that a borrower will fail to repay a loan or meet other credit obligations on time. Banks, non-banking finance companies, and digital lenders assess this risk when a loan is sanctioned, but the financial condition of a borrower can change considerably during the life of the loan. Loss of income, rising debt from other lenders, or changes in spending behaviour can move a borrower from low risk to high risk within a few months. This paper presents a real-time credit risk monitoring system that uses machine learning to update the risk level of existing borrowers continuously as new repayment and transaction data arrives.

The system uses a combination of static borrower attributes, such as age, income, employment type, loan amount, and tenure, and dynamic behavioural data, including monthly repayment status, days past due, credit utilisation, account balance trends, and the number of recent credit enquiries. Publicly available credit datasets are used for training, and a streaming simulator replays transaction and repayment events to reproduce the conditions of a live environment. The raw data is cleaned to handle missing values, inconsistent categories, and extreme outliers, and features describing recent behaviour are computed over rolling windows of one, three, and six months.

Because defaulting borrowers form a small minority of the data, class imbalance is handled with resampling techniques and class weights. Several classifiers are trained and compared, including Logistic Regression, Decision Tree, Random Forest, Gradient Boosting, XGBoost, and LightGBM. The models are evaluated using accuracy, precision, recall, F1-score, ROC-AUC, and the Kolmogorov-Smirnov statistic, which is commonly used in credit scoring. The results show that gradient boosting methods provide the best discrimination between borrowers who will and will not default within the next three months, with days past due, utilisation, and balance trend being the strongest indicators.

The trained model is deployed as a scoring service that receives new events and recalculates the probability of default for the affected borrower within seconds. Borrowers are placed into risk bands, and an alert is generated when a borrower moves to a higher band. A web dashboard shows the overall portfolio risk distribution, recent migrations between bands, individual borrower profiles with the factors behind their current score, and the performance of the model over time.

The system is intended to help lenders identify deteriorating accounts early, so that they can contact borrowers, restructure loans, or adjust credit limits before the account becomes non-performing. It supports rather than replaces the judgement of credit officers, and every automated alert is reviewed before action is taken. Future work

includes fairness assessment across borrower groups, integration of account aggregator data, and automatic detection of model drift.

I. INTRODUCTION

Lending is central to economic activity, and the ability of a lender to manage credit risk determines both its profitability and its stability. When a borrower stops paying, the lender loses interest income, incurs collection costs, and must set aside provisions against the expected loss. A rise in non-performing assets can weaken the lender's capital position and reduce its ability to extend credit to other customers. For these reasons, regulators require lenders to monitor the quality of their loan portfolio closely and to classify accounts according to how long payments have been overdue.

Traditional credit risk management places most of its effort at the time of loan approval. The applicant's income, employment, credit history, and existing obligations are examined, and a credit score is assigned. After disbursement, however, monitoring is often limited to tracking whether instalments are paid on time. By the time an account is classified as overdue, the borrower may already be in serious financial difficulty, and the options available to the lender are limited. Periodic reviews of large accounts are carried out manually, but for the very large number of small retail and digital loans, individual review is not practical.

Machine learning makes it possible to learn from historical data which patterns of behaviour tend to precede default, and to apply this knowledge to every borrower in the portfolio at frequent intervals. Changes such as increasing use of available credit, falling account balances, a growing number of enquiries with other lenders, or small delays in payment can each be weak signals on their own, but together they can indicate rising risk well before a payment is missed. Processing these signals as they arrive, rather than in monthly batches, allows the lender to respond while there is still time to prevent a default.

This paper presents a real-time credit risk monitoring system that combines static borrower information with continuously updated behavioural data. The system builds rolling behavioural features, handles the imbalance between defaulting and non-defaulting borrowers, compares several classification models, and deploys the best model as a scoring service connected to a monitoring dashboard. The aim is to provide lenders with an early warning of deteriorating accounts together with a clear explanation of why each borrower has been flagged.

II. LITERATURE SURVEY

Credit scoring has a long history in both statistics and finance. Early scorecards were built using discriminant analysis and later logistic regression, which models the probability of default as a function of borrower characteristics. Logistic regression became the industry standard because it is simple to estimate, produces probabilities that can be interpreted directly, and allows each variable to be converted into points on a scorecard that loan officers can understand. Studies in this area developed techniques such as weight of evidence transformation and information value to select

and encode variables. Researchers noted, however, that logistic regression assumes a linear relationship between the transformed inputs and the log-odds of default, and that it may miss interactions between variables. Its performance also depends heavily on careful manual feature preparation, and its accuracy can decline when the population of borrowers changes over time.

The application of machine learning to credit scoring has been examined in many comparative studies. Decision trees, random forests, support vector machines, neural networks, and gradient boosting methods were tested on public and proprietary credit datasets. A recurring finding was that ensemble methods, particularly gradient boosting implementations such as XGBoost and LightGBM, achieved higher discrimination than logistic regression, as measured by the area under the ROC curve and the Kolmogorov-Smirnov statistic. Support vector machines and neural networks performed competitively in some studies but required more tuning and were harder to explain. Authors pointed out that improvements in discrimination, even of a few percentage points, can translate into substantial reductions in credit losses for a large portfolio, which explains the strong interest of lenders in these methods despite concerns about their complexity.

Class imbalance is a central difficulty in credit risk modelling because defaulting borrowers usually form only a small fraction of the portfolio. Studies showed that classifiers trained on imbalanced data tend to favour the majority class and achieve high accuracy while failing to identify many defaulters. Techniques such as random undersampling, random oversampling, the synthetic minority oversampling technique, and cost-sensitive learning with class weights were proposed to address this problem. Comparative research found that no single technique was best in all cases, but that combining moderate resampling with class weights generally improved recall of defaulters without an unacceptable loss of precision. Several authors stressed that accuracy is a misleading measure for imbalanced problems and recommended precision, recall, F1-score, ROC-AUC, and precision-recall curves instead, together with evaluation on data from a later time period.

Behavioural scoring, which uses information on how a borrower manages an existing account, has been studied as a complement to application scoring. Research showed that variables such as payment history, credit utilisation, changes in balance, and frequency of new credit enquiries are strong predictors of future default, often stronger than the information available at the time of application. Studies on early warning systems proposed monitoring these variables continuously and generating alerts when they cross certain levels or change rapidly. With the growth of digital banking and real-time payment systems, more recent work has explored stream processing frameworks that compute behavioural features and update risk scores as each transaction arrives. These studies highlighted practical requirements such as low latency, consistent feature computation between training and deployment, and handling of late or out-of-order events.

Interpretability and fairness have become important themes in recent literature. Regulators and consumer protection guidelines require lenders to explain adverse decisions, and research has examined model-agnostic explanation methods such as

SHAP values and local surrogate models to show which factors contributed to a particular prediction. Other studies have investigated whether credit models treat different demographic groups unequally and have proposed fairness measures and mitigation techniques. The problem of model drift has also received attention, since changes in economic conditions can alter the relationship between borrower behaviour and default, making it necessary to monitor model performance and retrain periodically. The present system draws on these directions by combining behavioural features, imbalance handling, gradient boosting models, explanation of individual scores, and ongoing monitoring of model performance.

III. SYSTEM ANALYSIS

The system is analysed as a continuous loop that starts with borrower events and ends with updated risk scores and alerts. Static information for each borrower, including age, income, employment type, loan amount, interest rate, and tenure, is stored in a borrower profile database. Dynamic events such as instalment payments, missed payments, credit card transactions, balance updates, and credit bureau enquiries arrive as a stream, which in this project is reproduced by a simulator that replays historical records in time order. Each event updates the rolling behavioural features of the affected borrower, including days past due, average utilisation over one, three, and six months, trend in account balance, number of recent enquiries, and ratio of payments made to payments due. The updated feature vector is passed to the trained classification model, which returns a probability of default within the next three months. This probability is mapped to a risk band, and a change to a higher band triggers an alert. Explanation values for the main contributing features are computed for each flagged borrower and displayed on the dashboard.

Existing System

In the existing approach used by many lenders, credit risk is assessed mainly at the time of application using a bureau score and a set of eligibility rules. After the loan is disbursed, monitoring relies on the delinquency status of the account, which is updated when an instalment is missed and becomes visible in monthly reports. Accounts are classified by the number of days past due, and collection activity starts only after a payment has been missed. Behavioural information such as rising credit utilisation, declining balances, or new borrowing from other lenders is available in principle but is rarely combined into a forward-looking risk measure. Portfolio reviews are performed in periodic batches, typically once a month, so a borrower whose situation deteriorates early in the month may not be identified until several weeks later. Large corporate accounts receive manual review by credit officers, but retail and small digital loans are too numerous for individual attention. Rule-based alerts, where they exist, use fixed thresholds and produce many false alarms, leading staff to ignore them.

Disadvantages of Existing System

- Risk is assessed mainly at loan approval and rarely updated afterwards.
- Monitoring depends on missed payments, which are a late indicator of trouble.

- Portfolio reviews are carried out in monthly batches rather than continuously.
- Behavioural signals such as rising utilisation are not combined into a risk score.
- Retail and digital loans are too numerous for manual review.
- Rule-based alerts use fixed thresholds and produce frequent false alarms.
- Imbalance between defaulters and non-defaulters is not addressed in simple models.
- Reasons behind a change in risk are not explained to credit officers.
- Model performance is not monitored as economic conditions change.
- Collection efforts start only after the account has already become overdue.

Proposed System

The proposed system monitors the credit risk of existing borrowers in real time using machine learning. Borrower profiles containing static attributes are combined with a stream of repayment, transaction, balance, and enquiry events. A feature service updates rolling behavioural indicators for each borrower as events arrive, including days past due, utilisation averages, balance trend, enquiry count, and payment ratio. Historical data is cleaned, encoded, and scaled, and class imbalance is handled through SMOTE oversampling combined with class weights. Logistic Regression, Decision Tree, Random Forest, Gradient Boosting, XGBoost, and LightGBM classifiers are trained and compared using precision, recall, F1-score, ROC-AUC, and the KS statistic on a time-based test set. The best model is deployed as a scoring service that recalculates the probability of default whenever a borrower event is received. Borrowers are placed in risk bands, alerts are raised on upward migration, and SHAP values explain the main factors behind each score. The dashboard presents portfolio risk, alerts, borrower details, and model performance.

Advantages of Proposed System

- Updates the risk score of each borrower as soon as new events arrive.
- Detects deteriorating accounts before a payment is actually missed.
- Combines static borrower data with rolling behavioural indicators.
- Handles class imbalance to improve detection of likely defaulters.
- Compares several classifiers and deploys the most accurate one.
- Uses credit-industry measures such as ROC-AUC and the KS statistic.
- Explains the main factors behind each risk score using SHAP values.
- Places borrowers in clear risk bands and raises alerts on migration.
- Monitors model performance to detect drift over time.
- Supports early contact, restructuring, or limit adjustment by credit officers.

IV. METHODOLOGY

The methodology begins with the collection of public credit datasets containing borrower attributes, monthly repayment history, and default outcomes. The data is cleaned by removing duplicate records, correcting inconsistent category labels, imputing missing income and employment values, and capping extreme values. Categorical variables are encoded and numerical variables scaled. Behavioural

features are computed over rolling windows of one, three, and six months, including maximum days past due, average utilisation, change in balance, count of enquiries, and ratio of amount paid to amount due. The target is defined as default within the next three months. The data is split by time so that the test period follows the training period, and SMOTE with class weights is applied only to the training portion. Each classifier is trained and tuned using stratified cross-validation within the training period. Precision, recall, F1-score, ROC-AUC, and the KS statistic are calculated on the test period, and the probability cut-offs for each risk band are chosen from the score distribution. The selected model is packaged with its feature pipeline and exposed through a scoring API.

SYSTEM ARCHITECTURE

The architecture of the system is arranged in layers that carry information from borrower events to decisions by credit officers. The Data Source Layer holds borrower profiles and receives repayment, transaction, balance, and enquiry events. The Event Streaming Layer delivers these events in time order to the processing components. The Feature Computation Layer updates rolling behavioural indicators for each borrower as events arrive. The Preprocessing Layer applies the same encoding and scaling used during training. The Model Training Layer fits and compares classifiers with imbalance handling and time-based validation. The Real-Time Scoring Layer computes the probability of default for each updated borrower. The Risk Banding and Alert Layer assigns risk bands and raises alerts when borrowers migrate to higher bands. The Explanation Layer computes SHAP values to show the main factors behind each score. The Dashboard Layer presents portfolio distribution, alerts, borrower profiles, and model performance charts to credit officers.

Real time credit risk monitoring with ML

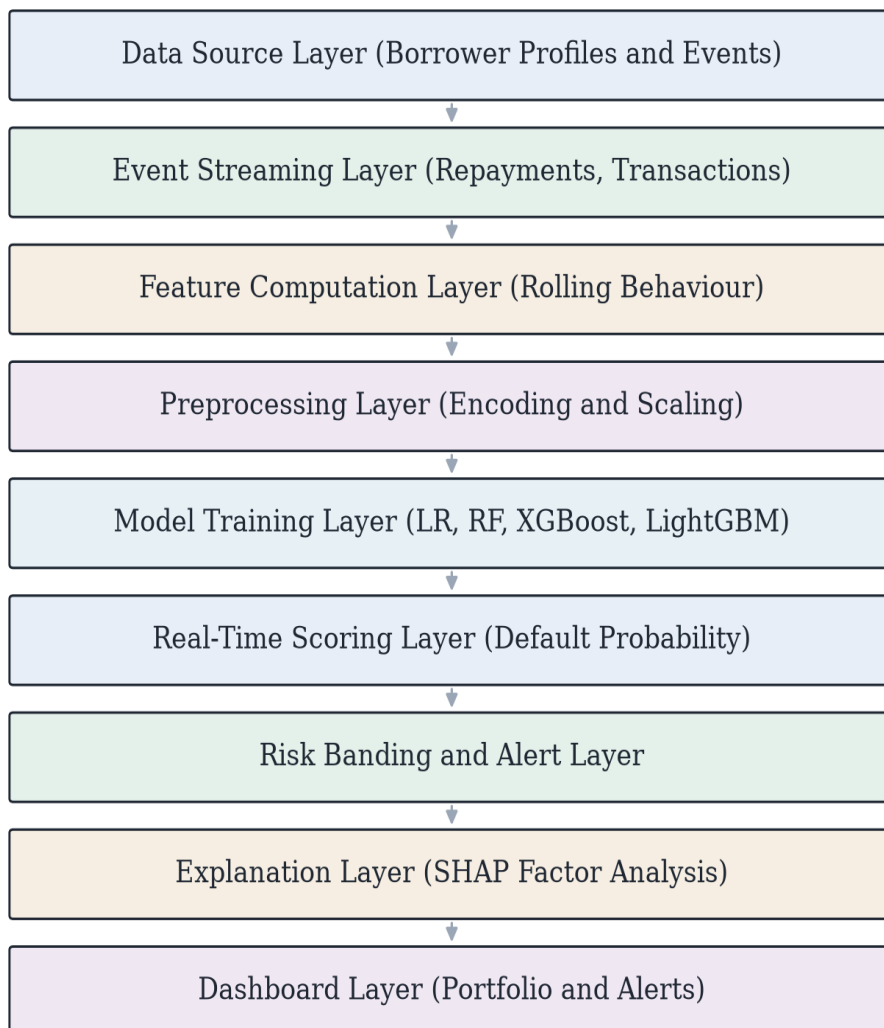
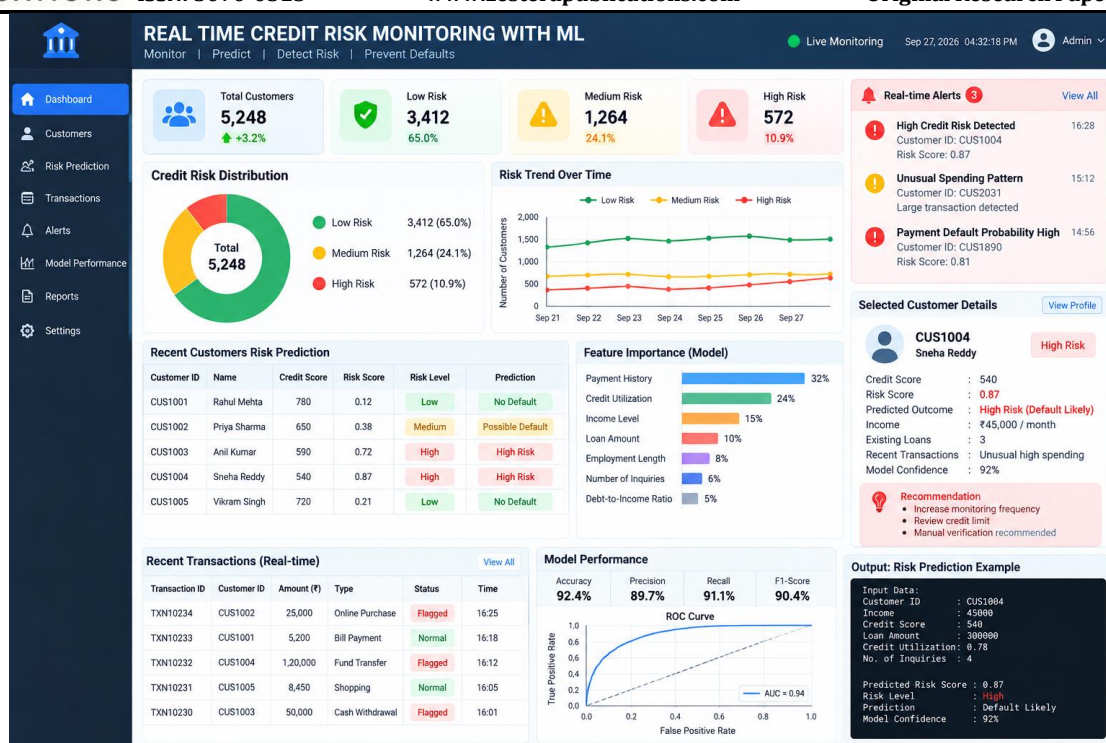


Fig. 1 System Architecture

V. RESULT AND OUTPUT



VI. CONCLUSION

This paper has presented a real-time credit risk monitoring system that uses machine learning to update the risk level of existing borrowers as new information becomes available. By moving from periodic batch reviews to continuous scoring, the system allows lenders to identify deteriorating accounts earlier and to act while there is still an opportunity to prevent default.

The combination of static borrower attributes with rolling behavioural features proved to be the key to effective prediction. Days past due, credit utilisation, balance trends, and recent enquiries together gave a much clearer picture of risk than application data alone. Among the models compared, gradient boosting methods achieved the best discrimination, and the use of resampling with class weights improved the detection of defaulters without producing an excessive number of false alerts. Evaluation on a later time period gave a realistic estimate of performance.

The explanation layer addresses an important practical requirement. Credit officers can see which factors caused a borrower to move to a higher risk band, which makes the alerts easier to trust and helps them decide on an appropriate response, such as contacting the borrower, offering restructuring, or reducing an unused credit limit. All alerts remain subject to human review, and the system is designed to support rather than replace professional judgement.

Future work can include assessment of fairness across borrower groups, integration of consented financial data from account aggregator services, and automatic detection of model drift with scheduled retraining. Extending the system to estimate loss given default and exposure at default would allow it to report expected loss for the whole

portfolio, and stress testing under different economic scenarios would help lenders prepare for adverse conditions.

REFERENCES

- [1] Kumar, R. D., Prudhviraaj, G., Vijay, K., Kumar, P. S., & Plugmann, P. (2024). Exploring COVID-19 through intensive investigation with supervised machine learning algorithm. In *Handbook of Artificial Intelligence and Wearables* (pp. 145-158). CRC Press.
- [2] Swathi, B., Vijay, K., Sushanth Babu, M., & Dinesh Kumar, R. (2024, November). Machine Learning Techniques in Cloud Based Intrusion Detection. In *The International Conference on Artificial Intelligence and Smart Environment* (pp. 557-564). Cham: Springer Nature Switzerland.
- [3] Sv satyakrishna, shirisha rang, bhargavi nalacheruve. (2024) Prospective investigation on colorectal cancer with SMOTE on machine learning Algorithm
- [4] Dr.G.Vishnu Murthy, BhargaviNalacheruve 1Professor, Department of computer Science & engineering, Anurag University, TS, India. 2Student, Department of computer Science & engineering, Anurag University, TS, India.
- [5] V. N. S. Manaswini, K. K, C. Nigam, S. S. Ali, R. Niranjana, and Suman, "Real-Time Object Detection in Drone Surveillance Using YOLOv5," in *Proc. 2025 3rd Int. Conf. IoT, Communication and Automation Technology (ICICAT)*, Gorakhpur, India, 2025, pp. 1–6, doi: 10.1109/ICICAT68430.2025.11414670.
- [6] B. Soundarya, V. N. S. Manaswini, M. Ayyakrishnan, R. D. Kumar, "Contextual Analysis of Big Data Analytics in Intelligent Transportation Frameworks," in *Intersection of Artificial Intelligence, Data Science, and Cutting-Edge Technologies: From Concepts to Applications in Smart Environment*, Lecture Notes in Networks and Systems, vol. 1353, Cham: Springer, 2025, doi: 10.1007/978-3-031-88304-0_79.
- [7] R. D. Kumar, V. N. S. Manaswini, "Applications of blockchain in smart cities: detecting fake documents from land records using blockchain technology," in *Blockchain for Smart Cities*, Elsevier, 2021, pp. 105–117, doi: 10.1016/B978-0-12-824446-3.00017-X.
- [8] Tejavath Veeramma, Badarla Anil, Guguloth Ravinder, "An advanced movie recommender using collaborative filtering and sentiment analysis," *International Research Journal of Modernization in Engineering Technology and Science*, vol. 7, no. 7, July 2025, doi: 10.56726/IRJMETS81618.
- [9] Ravi Kumar Banoth, Ramana Murthy B V, "Automatic crop recommendation system using LightGBM and decision tree machine learning models," *Journal of Machine and Computing*, vol. 5, no. 1, pp. 343, Jan. 2025, doi: 10.53759/7669/jmc202505026.
- [10] Ravi Kumar Banoth, Dr. B.V. Ramana Murthy, "Smart agriculture through IoT and machine learning for analyzing carbon footprints," in *Proc. Int. Conf. Computer Science and Communication Engineering (ICCSCE)*, Apr. 2025.
- [11] Ravi Kumar Banoth, B. V. Ramana Murthy, "Soil image classification using transfer learning approach: MobileNetV2 with CNN," *SN Computer Science*, vol. 5, art. no. 199, 2024, doi: 10.1007/s42979-023-02500-x.