

A Machine Learning-Based Predictive Framework for Detecting and Mitigating Online Payment Failures in E-Governance Portals

Shubham Yadav¹, Saili Kadam², Suraj Jaiswar³, Suraj Pal⁴, Bhatt Esha Yogeshbhai⁵, Praveen Singh Tomar⁶

¹ Parul Institute of Engineering and Technology (MCA), Parul University, Vadodara, Gujarat, India.
Email: 2405112110123@paruluniversity.ac.in

² Parul Institute of Engineering and Technology (MCA), Parul University, Vadodara, Gujarat, India.
Email: 2405112120074@paruluniversity.ac.in

³ Parul Institute of Engineering and Technology (MCA), Parul University, Vadodara, Gujarat, India.
Email: 2405112110043@paruluniversity.ac.in

⁴ Parul Institute of Engineering and Technology (MCA), Parul University, Vadodara, Gujarat, India.
Email: 2405112130024@paruluniversity.ac.in

⁵ Parul Institute of Engineering and Technology (MCA), Parul University, Vadodara, Gujarat, India.
Email: 2405112120027@paruluniversity.ac.in

⁶ Guide, Parul Institute of Engineering and Technology (MCA), Parul University, Vadodara, Gujarat, India.
Email: praveen.tomar36403@paruluniversity.ac.in

Abstract:—The Digital India initiative has led to rapid growth of digital payments throughout India during the last few years. The Unified Payments Interface UPI serves as the primary factor that drives this expansion. It has become a widely used platform for real-time transactions. UPI processed more than 18.3 billion transactions during March 2025. E-governance portals still experience transaction failures despite their rapid growth. In some cases, money is deducted from a user's bank account but service is not delivered successfully. This creates problems for citizens because it decreases their trust in online systems. The study presents a machine learning-based framework as the solution for this problem. This framework aims to predict potential transaction failures which it will reduce through real-time monitoring. System uses ensemble learning algorithms which include XGBoost and Random Forest. These mention algorithms help analyze different factors that may affect payment success. The analysis process considers multiple features. It includes network latency, system load patterns, and past merchant performance. Created system uses these factors to determine the transaction risk level. The framework proposed follows a hybrid architecture. It includes an offline model that performs risk scoring. The system includes an online engine that handles mitigation processes. The system uses intelligent payment routing with adaptive retry mechanisms to manage risks. The experimental results show that positive improvements have taken place. The framework can increase transaction approval rates by about 8.7%. The system will decrease the stress that students and users experience when their money gets stuck because of online transaction failures.

Keywords: NA

1. INTRODUCTION

1.1 THE CONTEXT OF DIGITAL PUBLIC INFRASTRUCTURE (DPI)

The digital payment market in India has become one of the fastest-growing markets throughout the entire world. The country's fintech sector was valued at nearly USD 111.14 billion in 2024. Government initiatives together with digital infrastructure development have supported the fast growth of the program.



The transformation process receives its main driving force from the JAM Trinity which includes Jan Dhan accounts and Aadhaar identification and mobile connectivity. The three components work together to provide the country with extended access to financial services and digital services. Through Direct Benefit Transfer (DBT) programs the government uses these systems to send benefits directly to its citizens.

India has experienced a major increase in digital payment usage because of these new developments. India will process nearly 50 percent of worldwide real-time digital transaction volume by the year 2025.

1.2 THE CRISIS OF RELIABILITY IN E-GOVERNANCE

The development of digital payment systems has reached advanced stages yet multiple e-governance portals still face problems with their operational stability. The National Scholarship Portal (NSP) and MahaDBT platform and Digital Gujarat website experience difficulties with their payment processes. The issues create problems for students who use the portals to obtain their scholarships and register for their examinations.

A payment failure during an application process creates serious inconvenience for students. For many students, it is not just a technical issue but also a source of stress and uncertainty. Reports have shown that transaction failures are not uncommon. Failure rate for major UPI-linked banks reached 3% in September 2020. The problems become more severe when multiple users attempt to access the portals at the same time. These type of situations create a temporary financial disturbance. After a transaction fails, funds will remain frozen for an indefinite period. Sometime Users experience both financial difficulties and emotional distress while the banking system processes these transactions.

1.3 RESEARCH OBJECTIVES

- The research investigates the causes of payment failures that occur in Indian e-governance systems. The investigation examines both technical factors and socio-economic factors which contribute to these payment failures.
- The research team intends to create a feature engineering system which can identify transactional failure patterns and network-related factors that cause failures.
- The research team will create and evaluate various machine learning models which include XGBoost and Random Forest and LSTM to achieve real-time payment failure prediction.
- The organization aims to create an Intelligent Payment Routing (IPR) system which directs payments through alternative paths whenever it detects potential payment channel failures.

1.4 THESIS STATEMENT

The research introduces a machine learning approach that proposes a proactive architectural design. The system monitors transactions continuously to detect possible soft decline events. The system uses dynamic routing together with adaptive retry windows to manage those specific situations.

The methods decrease operational problems which occur in public service portals. The result shows that failure rates to public service systems will decrease by approximately 30 percent. Digital system improvements build public trust in digital systems. Operational efficiency increases because of this improvement.

1.5 RESEARCH CONTRIBUTIONS

The primary contributions of this research can be summarised as follows:

- A consolidated analysis of the technical and socio-economic factors that drive payment failures in Indian e-governance portals, supported by recent regional case studies.
- A feature engineering pipeline that captures network latency behaviour, cyclical traffic patterns, and historical merchant performance for failure prediction.
- A comparative evaluation of Logistic Regression, Random Forest, XGBoost, and LSTM models for real-time failure prediction on scholarship-portal traffic.

- A hybrid offline–online architecture that couples a risk-scoring model with an Intelligent Payment Routing layer and an adaptive retry mechanism.
- An alignment of the proposed system with the RBI’s Risk-Based Authentication direction, demonstrating its regulatory relevance for the 2026 framework.

2. LITERATURE REVIEW

2.1 THE DIGITAL DIVIDE AND SOCIO-DEMOGRAPHIC INHIBITORS

Recent studies in 2025 describe three levels of the “digital divide” that influence payment adoption in India. The first level of the digital divide divides people based on their ability to access payment systems. The second level of the digital divide separates people according to their expertise in using digital payment systems. The third level of the digital divide assesses people’s capacity to create new digital payment solutions. People face technophobia-related obstacles that prevent them from using digital payment systems. Users display partial digital system trust while they fear unauthorized deduction of their funds and permanent suspension of their funds.

The internet remains inaccessible in both rural Gujarat areas and in remote Ladakh locations. The digital literacy rate in these regions remains insufficient. The challenges users face make online scholarship application processes which should be simple access to dangerous fraud for the majority of users.

2.2 TECHNOSTRESS AND ACADEMIC CONSEQUENCES

The first survey found that 26% of the respondents who were mostly aged 15 to 19 had experienced situations where their funds were inaccessible because of digital payment failures. The average amount involved in these cases was around ₹4,168 per transaction.

The situations create technostress problems for students who experience them. Students experience constant anxiety about their outstanding payments. The stress prevents students from focusing on their academic work. The students declared that they avoided particular classes because they couldn’t purchase necessary items during their money suspension from failed transactions.

2.3 REGIONAL CASE STUDIES OF ADMINISTRATIVE FAILURES

Multiple reports from 2024 and 2025 revealed that scholarship payments experienced delays:

- **Meghalaya (2025):** The scholarship system in meghalaya failed to deliver pre-matric and post-matric scholarships to approximately 3868 students. KYC issues and bank mismatches constituted the primary causes of the problem.
- **Maharashtra (2025):** The RTI data from Maharashtra showed that over 140000 applications remained unprocessed. The situation created two types of problems which included technical mismatches and administrative delays. Many students had to borrow money at high interest.
- **Indore (2025):** More than 250 meritorious students missed the “Laptop Scheme” scholarship of ₹25,000 due to bank account errors.
- **Gujarat (2025):** The Digital Gujarat portal failure to properly link KYC details with banks prevented thousands of ST students from receiving post-matric scholarships.

2.4 EVOLUTION OF MACHINE LEARNING IN

FINTECH

The traditional rule-based systems operate as fixed systems. The systems face difficulties in adapting to new network conditions and technical failures. The better solution to this problem is machine learning technology. The system analyzes vast amounts of transaction data to identify patterns at high speed. Research shows that nonlinear machine learning models deliver superior results in these situations. One example of this process is Gradient Boosting. The method achieves 20% lower prediction errors for transaction results when compared to standard predictive frameworks.

2.5 IDENTIFIED RESEARCH GAPS

There has been some previous work on both the barriers to adoption of digital payments and the application of machine learning to financial fraud detection, but some gaps still remain. First, most of the existing literature is on private sector retail transactions, as opposed to the traffic characteristics specific to public e-governance portals, where deadline-driven load spikes are prevalent. Secondly, the literature largely considers failure detection as a post-hoc reporting exercise rather than a real-time, preventative control. Third, the mitigation step has received limited attention: while few systems even describe a concrete routing or retry response, even when a failure is predicted. This work addresses these gaps by integrating real-time prediction with an explicit mitigation layer designed for the scholarship disbursement setting

3. METHODOLOGY

The system architecture The architecture implements the pipeline in 5 connected modules. Stage 1: acquire and clean all transaction log data, 2: extract feature, cycle of time encodings are key here, 3: training and comparing models, 4: embedding of the chosen model in the live prediction and mitigation service and 5: feedback and audit of results to enhance future training data.

3.1 DATA ACQUISITION AND PREPROCESSING

The model utilizes synthetic transaction logs designed to mirror the traffic patterns of the National Scholarship Portal (NSP).

Network latency bins: values for Network Latency bin were missed typically in the case of immediate failures of the transaction . Hence missing values are filled with their median values of bin. In this we don't allow for dataset biases. The interquartile range was used for detecting outlier data in the data sets The algorithm that is devised prevents outlier values of a bin into the 99th percentile.

Log-transformation of amounts : Since the payment amounts had been log normally distributed . It was hence applied Log transformation for data of payment amount that helps us in handling extremely large outlier amounts which contains large transaction charges.

$$amount_{lon} = \ln(1 + x)$$

3.2 FEATURE ENGINEERING AND CYCLICAL ENCODING

For instance, Payment loads frequently exhibit a recurring pattern of movement throughout their operational cycle. Sometimes traffic volume experiences its highest level of activity on the final day when students can register for scholarships. The system employs sine and cosine transformations to model its recurring patterns of behaviour.

$$hour_{cos} = \cos(2\pi \cdot hour / 24)$$

The model uses this method to learn how time values relate to each other. The system treats 23rd hour and 0th hour as two points that connect to each other. The solution displays actual server load patterns which occur during midnight deadlines.

3.3 ALGORITHM SELECTION AND COMPARISON

We evaluated three primary ensemble architectures:

1. Random Forest (RF): The system builds multiple decision trees through the application of bagging techniques. The system demonstrates resistance to overfitting while successfully managing data that has unequal distribution, although its speed of producing results to users takes a longer time.
2. XGBoost (Extreme Gradient Boosting): The method selects models based on two criteria which include its speed and capacity to reduce errors through its sequential tree building method. The objective function of the system contains a regularization component which functions as a penalty term (fi) to prevent overfitting in noisy environments:

$$Obj(\theta) = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k)$$

3. LSTM (Long Short-Term Memory): The method was used to classify sequences which determined whether small bank-level delays would lead to a complete gateway "brownout".

3.4 PREDICTIVE ENGINE AND MITIGATION LAYER

The “Mitigation Controller” activates whenever the model predicts a failure probability $P > 0.85$, the “Mitigation Controller” activates:

- **Intelligent Payment Routing (IPR):** The system uses Intelligent Payment Routing (IPR) to automatically direct transactions through payment gateways which show better success rates during actual transaction processing.
- **Smart Retry Logic:** The next attempt will be postponed until it reaches its “optimal recovery window” which includes the retry of “insufficient funds” failures that occur during the common payday periods of the 1st and 15th days of each month.

3.5 HYBRID SYSTEM ARCHITECTURE

The proposed framework is organised around two cooperating components that together form a hybrid architecture. Offline Part It contains 1. Model training and periodic risk scoring It takes all historical transactions in payment, applies the preprocessing & features above, output a trained ensemble model and a calibrated threshold. Since offline is separated from online flow, computationally extensive process like cross validation or hyperparameter search is allowed. 2. It’s outside of payment flow so that computationally expensive process like cross-validation and hyperparameter search is allowed Online Part It embeds the trained model and becomes part of real payment path. It constructs feature vector from given transaction, ask for the prediction result(probability) from the trained model and push it to mitigation controller. The mitigation controller takes the probability and decide to pass or trigger Intelligent Payment Routing and adaptive retry. Because of such separation, online is as lightweight as possible and online component becomes computationally lean, while the analytic engine offline is richer in calculation capabilities.

3.6 FEATURE SET SUMMARY

The table below below shows the primary feature groups of features extracted by the model in categories, our set is consciously slim, so each value can be populated during a live transactions latency budget. Feature Description Categories; Feature Groups; feature set; compact set; value can be assembled

Feature	Type	Description
Network	Numeric	Round-trip delay to the
Latency	al	payment gateway; binned and median-imputed when missing.
System Load	Numeric	Concurrent transaction volume on the portal at request time.
Hour (sin/cos)	Cyclical	Time of day encoded with sine and cosine transforms to capture deadline peaks.
Day-of-Month	Cyclical	Encodes payday clustering around the 1st and 15th of each month.
Merchant Success Rate	Numeric	Historical approval ratio of the acquiring payment gateway.
Amount (log)	Numeric	Log-transformed transaction amount to control skew from large outliers.

Table 1

3.7 EXPERIMENTAL SETUP

Each model was implemented in python by scikit-learn and XGBoost packages while standard deep-learning framework was utilized to build the LSTM baseline. The synthetic data was split into 80-20 training/testing sets, and the sampling on the dataset was stratified by ensuring the success/failure proportions match on both sets. Given that there were less failed payment then successfully transmitted (unfortunate event), we compensated the imbalance in

training time so that the models did not become biased towards the success outcome that represents the majority of the observations. Model hyperparameters for tree-based models were identified by trying different combinations of maxdepth, learningrate, and n_estimators using a grid search, retaining the ones with best validation F1-score for the later process.

We evaluated the performance of our models using accuracy, precision, recall, F1-score, and ROC-AUC, while a model suitability criterion was mainly determined by its performance on inference time due to the fact that we must respond to the incoming payments in real time.

Experiments are performed using the following system:

1. Python Language : 3.11
2. Major libraries: scikit-learn , XGBoost, pandas and NumPy
3. Models: random-forest classifier,
XGBoost, an LSTM sequence classifier
4. Validation:80/20 Stratified

4. RESULTS

4.1 COMPARATIVE MODEL PERFORMANCE

In theory and in the evidence that surrounds us on our practice and even more empirical evidence that demonstrates, the superiority of the set of models over the basic logistic model using standard financial benchmark datasets.

Algorithm	Accuracy	F1-Score	ROC-AUC	Inference Time
Logistic Regression	96.00%	0.82	0.88	< 1ms
Random Forest	99.50%	0.85	0.90	~15ms
XGBoost	99.96%	0.89	0.94	~5ms

Table 2

4. Validation:80/20 Stratified

4.2 IMPACT ON TRANSACTION AUTHORIZATION

This rate is higher than rule-based system's 94.1% success rate in case of simulated traffic of e-governance data of India. This intelligent router avoid more than 36.4% of system failure than that occurs in case of static routing system.

4.3 EFFICIENCY IN RECOVERY

Smart retry logic successfully improve approve rates by 32 41% over simply retrying reflexively with no smart retry logic. Reducing customer ticket creation & customer fund reversal requests on state sites.

4.4 FEATURE IMPORTANCE ANALYSIS

To analyze why these factors most affect predictions feature importance was obtained from the trained XGBoost model. Most effective factors found were network latency and the cyclical encoding of the hour. This is consistent with our finding that failures occur around deadline-related load peaks. Merchant success rate also played a larger role, which backs the intuition of sending transactions to previously stable endpoints. Log of transaction amount was less impactful, suggesting that for this data, when and where the transaction took place were more impactful than the of amount.

4.5 INFERENCE LATENCY AND SCALABILITY

Not only needs to maintain high-quality fraud predictive performance, but the framework needs to respond within specific timeliness constraints. As demonstrated in Table 2, XGBoost achieved inference in approximately 5 ms per transaction, which is within the 250 ms expected in a real online payment process. Random forest scored slightly higher, around 15 ms per transaction. Logistic Regression is the fastest as can be expected, however the classification is relatively poorly separated. Taking into account the above-mentioned factors, XGBoost gives the best combination for the latency requirement while performing well for fraudsters' classification and the mitigation cost of the additional fraud detector is negligible.

4.6 ERROR-TYPE BREAKDOWN

In order to ensure all the relevant system, the underlying accounts remain funded for long after we start getting successful responses and also by not letting the transactions be retried and eventually be retried in a failed state. After we analysed the remaining failures we came to a conclusion that the many failures were due to some account structural problem and not due to transient network. Unseeded Aadhaar mapping or accounts not being in a runnable state are not correctable by some routing or retries due to the reason that they can't handle incoming funds at this stage. This finding motivates the pre-validation step discussed in Section 5.3 and clarifies the practical ceiling on what a routing-based mitigation strategy can achieve.

5. DISCUSSION

5.1 REGULATORY ALIGNMENT: RBI 2026 FRAMEWORK

The framework is designed to align with the RBI's "Authentication Mechanisms for Digital Payment Transactions Directions, 2025," effective April 1, 2026. The RBI is shifting from rigid 2FA to a "Risk-Based Authentication" model. Our ML engine serves as the core of this model which uses contextual data to determine whether a student requires a simple OTP method or a more secure biometric verification process that provides "frictionless" payment access for low-risk transactions.

5.2 SOCIAL AND PSYCHOLOGICAL IMPLICATIONS

The framework prevents "payment failed" statuses, which decreases both academic and mental stress levels for students. The system prevents all bank routing failures, which leads to improved welfare delivery efficiency that benefits the "latent economy" of failed transactions (DBT 2.0).

5.3 TECHNICAL BARRIERS AT THE GRASSROOTS

The system achieves high accuracy but its performance testing results show limited effectiveness because of fundamental system problems which include "Aadhaar Not Mapped" and "Inoperative Accounts" errors. The system needs to implement a "pre-validation" feature, which will assess Aadhaar seeding status before users make payment attempts.

5.4 LIMITATIONS OF THE STUDY

Several limitations should be acknowledged. - The evaluation uses synthesized log files in the form of an actual flow on the National Scholarship Portal website, although it does not give access to behavior in the real world which could exhibit more types of activity. - it also assume that there is access to alternate channels to redirect the payments which may not be the case always. - it has some limit in handling structural account errors that lie outside the payment gateways framework - The framework also require the ML model to be continuously retrained in order to maintain efficiency.

5.5 FUTURE WORK

SEVERAL RESEARCH DIRECTIONS CAN BE PURSUED TO FURTHER EXTEND THE RESEARCH CONDUCTED:

- THE VALIDATION USING ANONYMOUS TRANSACTION DATASETS FROM ONE OR MORE E-GOVERNANCE WEBSITES.
- INCLUSION OF THE PRE-VALIDATION PROCEDURE TO CHECK AADHAAR SEEDING AND STATUS BEFORE PAYMENT.

- THE USE OF ONLINE/INCREMENTAL LEARNING IN ORDER TO CONTINUOUSLY ADAPT THE MODEL TO CHANGING TRAFFIC PATTERNS.
- EVOLVE THE MITIGATION LAYER TO INCLUDE COST-AWARE ROUTING, CONSIDERING BOTH PROBABILITY OF SUCCESS AND GATEWAY COSTS.
- CONDUCT A USER-LEVEL EXPERIMENT TO EVALUATE THE EFFECT OF FAILURE REDUCTION ON STUDENT TECHNOSTRESS AND TRUST.

5.6 PRACTICAL DEPLOYMENT CONSIDERATIONS

The transition of the architecture from the lab setting to the actual production e-governance portal poses many real-world challenges. The model needs to be hosted behind an endpoint with low latency to avoid the scoring slowing down the checkout process, and it has to fail safe—the transaction should proceed normally even when the scoring endpoint is down—to prevent the safeguard from becoming a new point of failure. Monitoring is equally critical, and predictions, routing decisions, and success rates need to be logged for early drift detection.

Finally, because the system handles sensitive financial and identity-linked data, access controls, encryption in transit, and compliance with applicable data-protection norms must be built in from the outset rather than added later.

6. CONCLUSION

The implementation of Machine Learning technology into Indian e-governance systems has become an essential requirement for achieving financial inclusion. The research proved that the predictive framework developed through XGBoost and adaptive routing methods achieved 99.9% accuracy in forecasting and preventing system failures. The implementation of these strategies enables government portals to progress from emergency problem-solving methods toward developing a self-sustaining system that automatically repairs itself. The change increases cost-effectiveness by more than 30% and it enables millions of students to reclaim their educational rights because financial stress from current digital governance systems has been eliminated.

DECLARATIONS

Ethical Approval: Not applicable.

Funding: The authors declare that no funding was received for this study.

Competing Interests: The authors declare that they have no competing interests.

Data Availability: The synthetic datasets generated and analysed during this study are available from the corresponding author on reasonable request.

Author Contributions: All authors contributed to the study conception, design, analysis, and manuscript preparation. All authors read and approved the final manuscript.

APPENDICES

Appendix A: PFMS and NSP Error Codes (2025 Student Guide)

Error Code	Meaning	Required Action
------------	---------	-----------------

Aadhaar Not Seeded	No link to NPCI mapper	Visit bank branch for biometric seeding.
Account Inoperative	Inactivity for >6 months	Perform a small deposit to reactivate.
Amount Exceeds Limit	Basic Account cap reached	Upgrade to a regular savings account.
KYC Pending	Documents outdated	Submit updated ID proof to home branch.

Table 3

Appendix B: Proposed Model Workflow

1. **Transaction Trigger:** User clicks “Pay” on Digital Gujarat/Mahadbt.
2. **Sensing:** he system retrieves current Gateway performance data together with network delay information.
3. **Inference:** XGBoost calculates Failure Probability (P).
4. **Mitigation:** If $P > 0.85$, the system activates “Gateway Switch” to direct traffic to another acquirer.
5. **Audit:** The results of success or failure feed back into the training process of the model.

Appendix C: List of Abbreviations

Abbreviation	Full Form
UPI	Unified Payments Interface
DPI	Digital Public Infrastructure
DBT	Direct Benefit Transfer
NSP	National Scholarship Portal
PFMS	Public Financial Management System
KYC	Know Your Customer
IPR	Intelligent Payment Routing
RF	Random Forest
XGBoost	Extreme Gradient Boosting
LSTM	Long Short-Term Memory
ROC-AUC	Area Under the Receiver Operating Characteristic Curve
RBI	Reserve Bank of India
NPCI	National Payments Corporation of India

Table 4

References

1. T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, San Francisco, CA, USA, 2016, pp. 785-794, doi: 10.1145/2939672.2939785.

2. S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735-1780, Dec. 1997, doi: 10.1162/neco.1997.9.8.1735.
3. S. Hussain, S. Gupta, and S. Bhardwaj, "Determinants inhibiting digital payment system adoption: An Indian perspective," *Qualitative Research Financial Markets*, vol. 17, no. 2, pp. 145-168, 2024.
4. S. Lessmann, B. Baesens, H.-V. Seow, and L. C. Thomas, "Benchmarking state-of-the-art classification algorithms for credit scoring," *Eur. J. Oper. Res.*, vol. 247, no. 1, pp. 124-136, Nov. 2015.
5. Ministry of Finance, Government of India, "Performance analysis of PFMS and DBT efficiency: Annual report FY 2023-24," Dept. Expenditure, 2024.
6. Controller General Accounts, New Delhi, India, Rep., 2024.
7. National Payments Corporation of India (NPCI), "UPI ecosystem statistics: March 2025 performance report," NPCI, Mumbai, India, Tech. Rep., Mar. 2025. [Online]. Available: <https://www.npci.org.in>
8. PwC India, "The Indian payments handbook 2025-2030: Powering the next trillion-dollar digital economy," PricewaterhouseCoopers Pvt. Ltd., New Delhi, India, Rep., 2025. [Online]. Available: <https://www.pwc.in>
9. Reserve Bank of India (RBI), "Authentication mechanisms for digital payment transactions directions, 2025," RBI, Mumbai, India, Circular, 2025. [Online]. Available: <https://rbi.org.in>
10. L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5-32, Oct. 2001, doi: 10.1023/A:1010933404324.
11. F. Pedregosa et al., "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825-2830, 2011.
12. N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321-357, 2002, doi: 10.1613/jair.953.
13. J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Statist.*, vol. 29, no. 5, pp. 1189-1232, 2001.