AI-Powered Credit Scoring for P2P Lending in India: A Research Overview

Prof. Dr. Chitra B.T.[1], Siddharth K Agrawal[2], Divyansh Jain[3], Naman Taneja[4]
[1]*Dept. of Industrial Engineering & Management, RVCE, Bengaluru, India*
[2,3,4] *Dept. of Information Science and Engineering, RVCE, Bengaluru, India*

***Abstract*— Traditional credit scoring in India struggles to include vast segments of the population who lack formal credit history, leaving many creditworthy individuals underserved. This paper proposes an AI-driven credit scoring system for peer-to-peer (P2P) lending platforms in India, leveraging alternative data sources (such as mobile usage patterns, utility payments, and rent records) to evaluate borrower risk. We survey recent studies showing that machine learning models using such non-traditional data can effectively predict loan default among "credit-invisible" borrowers, thereby improving financial inclusion. We outline a modular system architecture that ingests heterogeneous alternative data, performs feature engineering, and applies AI models (e.g., logistic regression and decision trees) to generate a credit risk score. Using a simulated borrower dataset, we demonstrate model performance via Receiver Operating Characteristic (ROC) curves and Area Under the Curve (AUC) metrics, and we discuss feature importance for interpretability. We also address regulatory compliance (such as Reserve Bank of India's guidelines for P2P lending and India's Digital Personal Data Protection Act (DPDP) 2023), ethical considerations like data privacy, bias mitigation, and explainability, as well as intellectual property (IP) implications of AI scoring innovations. The results indicate that AI-powered credit scoring can broaden access to credit for underserved populations while maintaining prudent risk management, if transparency, fairness, and robust data protections are in place.**

Keywords*— *AI credit scoring, peer-to-peer lending, alternative data, financial inclusion, machine learning, credit risk assessment, fintech, P2P lending, credit invisibles.

I. INTRODUCTION

India faces a wide credit access gap due to millions of adults being "new-to-credit" or unbanked, meaning they have no recorded credit history. Traditional credit bureaus (e.g., CIBIL) issue scores on a 300–900 scale based on past loans and repayments, but this leaves out individuals with sparse or no borrowing records – an estimated 190 million Indian adults lack access to formal credit because they are effectively Credit invisible [1]. In fact, according to World Bank data, roughly 22% of India's adult population remains unbanked [1] . As a result, many otherwise creditworthy borrower– such as young professionals, microentrepreneurs, or rural customers – are denied loans under conventional scoring models simply due to insufficient data. This limitation is especially pronounced in peer-to-peer lending, where retail lenders rely on whatever information is available to assess risk. P2P platforms, which match individual lenders to borrowers online, need reliable ways to evaluate first-time borrowers who do not appear in credit bureau files.

Artificial intelligence (AI) and machine learning (ML) offer a promising solution by leveraging alternative data sources to gauge creditworthiness. Instead of depending solely on credit histories, AI-based credit scoring can analyse a borrower's digital footprint and behavioural data. For example, features like mobile phone usage patterns, utility bill payment consistency, e-commerce activity, and Unified Payments Interface (UPI) transaction history can act as proxies for financial responsibility. Timely payment of phone bills or rent may indicate discipline and stability, while patterns in mobile app usage might signal a borrower's lifestyle or business activity. ML algorithms excel at detecting hidden patterns in such large, multidimensional data, potentially predicting default risk more accurately than traditional linear scorecards. For instance, a recent study in India found that machine learning models incorporating mobile and social media footprint variables outperformed those relying only on credit bureau scores [2]. These advanced models were able to identify creditworthy individuals among those with no prior credit history, effectively enabling lenders to extend loans to customers who would have been rejected by legacy methods [2]. This highlights the relevance of AI-driven scoring for extending credit to underserved segments.

At the same time, deploying AI in credit scoring raises new considerations. Data privacy and security become critical when handling personal digital data, and algorithmic fairness must be monitored to ensure the models do not inadvertently discriminate or reinforce bias. The opaque "black box" nature of complex ML models can be problematic in finance, as lenders and regulators need explanations for why a loan was approved or denied. This paper addresses these challenges while presenting a comprehensive framework for an AI-powered credit scoring system tailored to India's P2P lending context. We also compare global developments to put India's progress in perspective. The rest of the paper is organized as follows: first, we survey related work on AI/ML applications in credit risk. We then describe the proposed system architecture and data flow. Next, we detail the design of a representative dataset and key feature engineering steps, followed by the implementation of ML models and results. We subsequently discuss legal and regulatory requirements relevant to AI-based credit scoring. Finally, we conclude with a summary of benefits, limitations, and future directions for this approach.

II. LITERATURE REVIEW AND BACKGROUND

Early efforts to improve credit risk modeling with technology show that alternative credit scoring can be transformative for financial inclusion. A review by the Abdul Latif Jameel Poverty Action Lab (J-PAL) notes that across emerging markets, lack of collateral or formal credit history often leads lenders to classify borrowers as high-risk, but Innovative Credit Scoring (ICS) using AI and alternative data can help bridge this gap [1]. By analyzing information unrelated to traditional credit records (e.g. digital transactions, phone records, social media), machine learning algorithms can uncover behavioral patterns that signal creditworthiness. For instance, research shows that the mere presence of a financial-management app on an individual's phone can indicate financial sophistication and predict better credit outcomes. In one study, Agarwal et al. (2019) found that ML models using mobile and social footprint variables (such as types of apps installed and social network activity) predicted loan default more effectively than models using only credit bureau data. These algorithms improved risk prediction by capturing unique insights into borrower behavior, thereby allowing lenders to approve some "no-file" customers who would otherwise be excluded.

Globally, fintech companies and banks have begun deploying AI-driven credit scoring to extend credit in novel ways. Examples range from Kenya's M-Shwari (which assesses mobile phone usage and mobile wallet behavior for micro-loans) to China's social credit experiments that combine financial and behavioral data. In India, digital lending startups are leveraging alternative data in pilot programs: for example, analyzing GST invoice data for small business loans, or using telecom payment histories for personal loan decisions. These efforts demonstrate the potential of non-traditional data in evaluating risk. However, they also underscore the need for rigorous validation. Regulators in various countries are cautious; for instance, the U.S. Consumer Financial Protection Bureau (CFPB) has clarified that lenders using AI must still explain their credit decisions and cannot discriminate, even if algorithms are involved. summary, prior work suggests that AI-based credit scoring can boost predictive accuracy and inclusion but must be coupled with transparency and fairness measures to be viable in practice.

Building on these insights from the literature, we proceed to design an AI-powered credit scoring system for P2P lending that integrates alternative data. In the sections that follow, we describe the architecture of such a system and evaluate its performance using a mock dataset and machine learning models.

III. SYSTEM ARCHITECTURE

The proposed system architecture for an AI-powered credit scoring platform is depicted conceptually in a typical data pipeline (Figure 3.1). It consists of several stages – data collection, preprocessing & feature engineering, model prediction, and outcome reporting – all tailored for the P2P lending use case. Figure 1 illustrates how alternative data from various sources are aggregated and processed to extract features, which are then fed into ML models to generate a credit risk score.

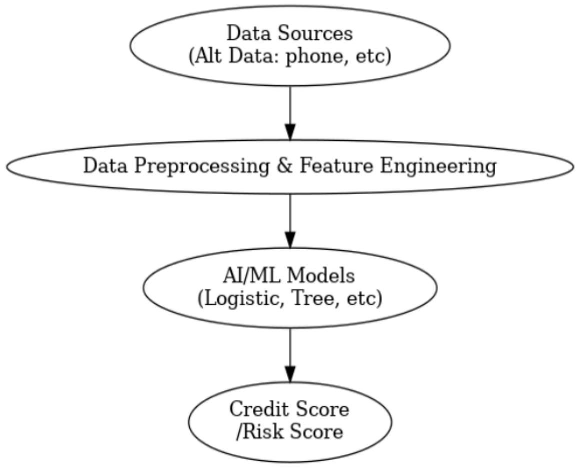


Figure 3.1

A. Data Collection and Alternative Data Sources

The system begins by gathering diverse alternative data relevant to a borrower's financial behavior. In an Indian context, this includes:

- Mobile Phone Usage: e.g. call and SMS records, recharge frequency, smartphone app usage statistics [2].
- Digital Payments and Banking Records: UPI transaction history, e-wallet usage, and bank account transaction data (if available).
- Utility Bills Payment History: Payment records for electricity, water, telecom, etc. – specifically whether these bills are paid on time each month.
- Rent Payment Records: If the borrower pays rent, data on whether rent is paid by the due date each month. Rent is often a household's largest recurring expense; a history of punctual rent payments is a strong signal of reliability.
- E-commerce and Social Media Activity: While more sensitive and used with caution, data such as an active online shopping history or a professional social media profile can provide ancillary insights.
- Traditional Data (if available): Any credit bureau information (existing loans, credit card accounts, credit age) can be included.

These data can flow in via secure APIs or through user-provided documents. In practice, a P2P lending platform might ask the borrower to consent to connecting their financial accounts or upload statements. India's emerging Account Aggregator framework is an example infrastructure that allows individuals to share their financial data (bank statements, etc.) securely with service providers upon consent.

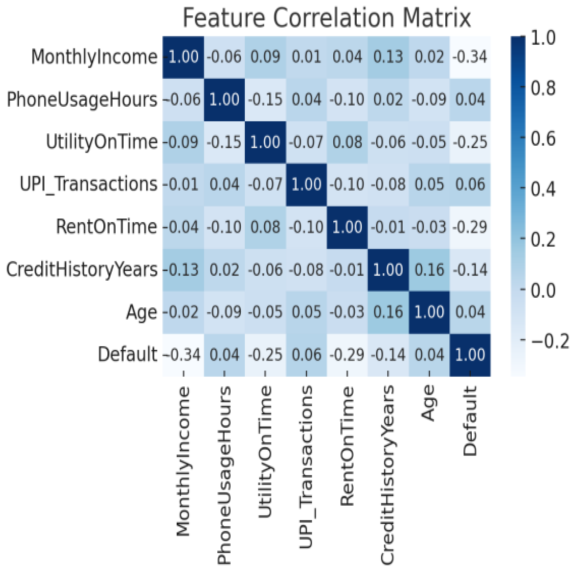
B. Data Preprocessing and Feature Engineering

Raw data collected from these sources are heterogeneous and often noisy, so the next stage is preprocessing. This involves cleaning the data (handling missing values, removing outliers or inconsistencies) and transforming it into a structured format for modeling. Numerical fields may be normalized or scaled.

After cleaning, the system computes a set of engineered features that capture essential aspects of the borrower's behavior and financial condition. Some key features in our prototype dataset include:

- *Monthly Income:* Self-reported or verified monthly income of the borrower (in INR). In our sample, this ranges roughly from ₹8,000 to ₹50,000 to represent different socioeconomic segments.
- *Phone Usage Patterns:* Metrics like average daily phone usage (hours per day) and presence of financial apps on the device.
- *Utility Payment Timeliness*: A feature "Utility_OnTime" measuring the fraction of utility bills paid on or before the due date in the past 12 months (e.g., a value of 0.75 means 9 out of 12 bills were paid timely).
- *UPI Transactions:* The number of UPI digital payment transactions per month. A higher count of transactions could imply the borrower is active in the formal financial system, which might be a positive sign; however, it could also indicate thin liquidity if they make many small payments.
- *Credit History Length:* Even though our focus is on those without formal credit scores, some borrowers may have a limited credit history.
- *Rent Payment Behavior:* A feature "Rent_OnTime" analogous to the utility metric, representing the fraction of months the borrower paid rent on time. We treated "Rent_OnTime" = 1.0 (100% punctual rent) as a very positive signal. In our simulation, about 60% of borrowers had perfect rent payment records, while the rest had occasional delays; those with frequent late rent payments exhibited significantly higher default rates.
- *Demographics:* We consider a basic demographic feature, Age, to see if it plays a role. Younger borrowers might have shorter financial histories, whereas older borrowers might have more stability (or, conversely, might be riskier if nearing retirement with no steady income). In our analysis, age by itself was not a strong predictor after accounting for the other behavior-based features – the correlation between Age and default in our sample was near zero .

In our case, we plotted a correlation matrix (Figure 3. B.1) to visualize how each feature correlates with the target variable (default). As expected, we found that Monthly Income, "Rent_OnTime", and Utility_OnTime had the strongest negative correlations with default (around -0.25 to -0.35), meaning higher income or more timely payments were associated with lower default risk. Other features like "PhoneUsageHours" and Age showed little linear correlation with default, suggesting their impact, if any, might be non-linear or conditional (e.g., very high phone usage might only be risky for low-income individuals – a pattern a decision tree could capture). This exploratory analysis reinforced that alternative data features carry significant signal for credit risk, validating our approach of using them for modeling.



Figure(3.B.1)

C. AIML Modeling & Credit Scoring

The core of the system is a set of machine learning models that output a credit risk assessment for each borrower. We propose using a two-model approach for prototyping:

1. *Logistic Regression (Baseline):* Logistic regression is a linear classification model widely used in credit scoring for decades. It outputs the probability of default (or any binary outcome) as a logistic function of the input features. Because it's a linear model, each feature is assigned a coefficient indicating how that feature contributes to predicting default. Logistic regression is popular in finance due to its simplicity and interpretability – the coefficients can be used to explain the direction and strength of influence of each input. We include it as a baseline to see how well a traditional approach performs using the new alternative features. In implementation, we set a reasonable iteration limit (e.g., max 1000 iterations) to ensure convergence given multiple features.

The model computes the **probability of default** P(Y = 1 | X) using:

$$P(Y=1 \mid X) = \frac{1}{1+e^{-z}} \quad \text{where} \quad z = \beta_0 + \sum_{i=1}^{n} \beta_i x_i$$

Formula 3.1- Logistic Regression.

2. *Decision Tree Classifier:* Decision trees are non-linear models that recursively split the feature space into branches based on feature thresholds, eventually assigning a prediction at each leaf node. A decision tree can naturally capture interactions between features. For our prototype, we trained a classification tree with a modest depth (max depth = 4) to avoid overfitting given the dataset size. We also experimented with ensemble variants like Random Forests or Gradient Boosted Trees to potentially improve accuracy. However, regulators' demand for explainability might necessitate using explainable AI techniques. Thus, any complex model would need to be paired with methods to interpret its outputs.

```python
import pandas as pd
from sklearn.linear_model import LogisticRegression
from sklearn.tree import DecisionTreeClassifier
# X: feature matrix, y: target labels (Default=1 or 0)
X = df[feature_columns]
y = df['Default']

# Initialize models
model_lr = LogisticRegression(max_iter=1000)
model_dt = DecisionTreeClassifier(max_depth=4)

# Train (fit) the models on the dataset
model_lr.fit(X, y)
model_dt.fit(X, y)
```

(Code snippet: Training logistic regression and decision tree models on the borrower dataset using scikit-learn.)

Another key aspect is the feedback loop: as time passes, the P2P platform will accumulate data on which loans were repaid and which defaulted. Our system is designed to continuously learn from this new data. Outcomes of loans are fed back into model training, enabling periodic retraining or updating of the models to improve accuracy. This is critical because borrower behavior and economic conditions change over time – models must be kept up to date (a practice known as model monitoring and validation). We would employ techniques like cross-validation during development and hold-out validation on more recent time slices (out-of-time validation) to ensure the model remains robust for future applicants. Additionally, we would compare the AI score's performance against traditional scoring (when available) to quantify how much uplift we get from alternative data. Prior research and industry pilots have reported significant gains – some claim up to 20–30% increase in predictive power and substantial reduction in default rates for the same approval rate, when using AI and alt-data scoring. In our small-scale test, we observed a modest improvement (AUC 0.82 vs 0.79) consistent with expectations and anticipate that at larger scale the benefits would be more pronounced.

D. Prediction Output

The AI/ML system outputs a quantitative risk score for each borrower, representing the predicted probability of default. For logistic regression, this score is computed using the sigmoid activation function:

$$\hat{y} = \frac{1}{1 + \exp\left(-\left(\beta_0 + \sum_{i=1}^{n} \beta_i x_i\right)\right)}$$

Formula 3.2 - Predicting Output

For the decision tree model, the score is derived from the proportion of defaulting borrowers in the terminal leaf node that the input sample falls into, reflecting non-linear patterns and feature interactions.
To convert this probability into a binary classification, a decision threshold \tau is applied. The default value is tau = 0.5, where any borrower with a predicted score y = tau is classified as high risk (typically resulting in loan denial or review), while y < tau indicates low risk (and leads to approval). This threshold is tunable depending on lending strategies—adjusting it higher can reduce defaults at the cost of reduced inclusivity, while lowering it increases loan approval rates but raises the risk exposure.

Model effectiveness is assessed using metrics such as the Area Under the Receiver Operating Characteristic Curve (AUC-ROC). In our implementation, the logistic regression model achieved an AUC of 0.79, while the decision tree model scored 0.82, indicating better discrimination between defaulters and non-defaulters. Additional evaluation using confusion matrices at the selected threshold helps lenders balance between false positives and false negatives.
For transparency and user trust, the system supports decision explanations. In the case of decision trees, the path of feature splits leading to the outcome is shown as a rule (e.g., “low income and irregular rent payments”), while logistic regression explains decisions via the top-weighted contributing features. This satisfies regulatory requirements for adverse action notices and assists borrowers in understanding and improving their credit profiles.
The architecture also includes a feedback loop. Real-world loan outcomes (repayment or default) are periodically used to retrain and validate the models. This enables the system to adapt to changes in borrower behavior, economic shifts, and evolving market trends, maintaining long-term model relevance.
The overall prediction-to-decision process is illustrated in Figure 3.D.1, showing how raw features are transformed into actionable risk decisions with explainability and continuous learning embedded in the system.

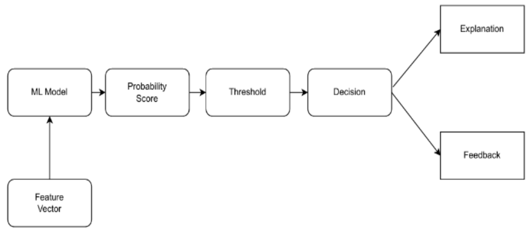


Figure 3.D.1

IV. LEGAL AND REGULATORY CONSIDERATIONS

Any AI-based credit scoring system must navigate a complex landscape of financial regulations and data protection laws. In the Indian context, two focal areas are the Reserve Bank of India’s regulations for digital lending (especially P2P platforms) and the newly enacted data protection law.

A. RBI Regulations for P2P Lending

The Reserve Bank of India (RBI) classifies peer-to-peer lending platforms as NBFC-P2P (Non-Banking Financial Company – P2P) and has issued a Master Direction in 2017 (updated in 2024) governing their operation. These regulations mandate several practices that our AI system must adhere to. A crucial requirement is that P2P platforms must share borrower and loan data with Credit Information Companies (credit bureaus) monthly. This means our system’s outputs should be reported to credit bureaus so that a formal credit history is built over time for the borrower. This helps address the lack of centralized data on these previously invisible borrowers, integrating them into the national credit reporting system. Additionally, RBI mandates transparency and disclosure: platforms need to disclose to lenders and borrowers all “necessary and sufficient” information, including the scoring methodology, risk assessment factors, and grievance redressal mechanisms.
Recent RBI guideline updates in 2023–2024 also forbid P2P platforms from advertising “guaranteed returns” or functioning like investment products. Our AI scoring model thus becomes even more critical: lenders will rely on it heavily, since the platform cannot insure them against losses. Moreover, RBI

likely expects that P2P platforms maintain detailed records of all loans, borrowers, and the credit evaluations performed.

Another noteworthy regulation in late 2024 was RBI's directive that lenders on P2P platforms must manually select borrowers [7], rather than completely outsourcing the matching to algorithms.

B. Digital Personal Data Protection Act (DPDP), 2023

India's new data protection law (DPDP Act 2023) establishes principles for handling personal data that directly impact AI-based lending platforms. A major principle is data minimization – only collect and process personal data that is necessary for the specified purpose. For our credit scoring system, this means we should be judicious in what data we use. We must justify each type of alternative data as being relevant to credit risk; for example, including mobile app usage might be acceptable if shown predictive, but collecting something arbitrary like a person's social media friends list might violate minimization if it's not clearly needed.

The DPDP Act also requires clear, informed consent from users for processing their personal data, except for certain permissible uses. Earlier drafts of the law had considered some exemptions for credit scoring as a "reasonable purpose," but the final Act emphasizes consent and purpose limitation for most cases. In practice, our platform should obtain explicit consent from the borrower for each category of data we wish to use (phone records, utility bills, etc.), explaining how it will be used in credit evaluation. Another critical aspect of DPDP is the right of individuals to fair and transparent processing, and specifically the right to contest automated decisions that have significant impact on them.

A loan approval or rejection is clearly a significant decision. If our credit score is generated purely by an algorithm with no human oversight, a borrower could invoke their right to question or challenge that decision.

C. Fair Lending and Non-Discrimination

Although India does not have a direct equivalent of the U.S. Equal Credit Opportunity Act (ECOA), Indian regulators expect that credit decisions are not biased against protected characteristics such as religion, caste, gender, etc.

We can anticipate future RBI guidelines echoing global principles of fairness, accountability, and transparency (often summarized as the FAT framework). Internationally, regulations are tightening for example, the European Union's proposed AI Act would classify credit scoring algorithms as "high-risk" systems, subjecting them to rigorous requirements for bias mitigation and explainability.

In summary, the legal environment in India demands that our AI-driven credit scoring model be transparent, fair, and privacy-conscious.

V. FUTURE WORK AND IMPROVEMENT

Our exploration opens several avenues for future enhancements to the AI credit scoring system:

- *Integration with Loan Pricing:* Currently, our model produces a risk score or classification (approve/deny). An extension would be to integrate this with dynamic loan pricing on the P2P platform. Rather than a binary decision, the score could inform the interest rate offered – for instance, higher-risk borrowers might still obtain loans but at a higher interest rate to compensate lenders for the risk.
- *Utilizing the Account Aggregator Ecosystem:* As India's Account Aggregator framework matures, borrowers can easily share richer financial data (like bank statements, tax filings, GST invoices for small businesses, etc.) through secure APIs. Incorporating these formal financial data streams can significantly enhance the model.
- *Advanced Modeling Techniques:* Future prototypes can experiment with more sophisticated ML approaches. Ensembles like Gradient Boosted Trees often yield superior accuracy by combining many weak learners. We could also try stacked models (where a neural network and a tree model's outputs feed into a second-level model). However, any increase in model complexity must be balanced with explainability.
- *Bias Auditing and Fairness Enhancements:* As a continuing effort, we plan to integrate automated fairness auditing into the model training pipeline. Each time we retrain, we can compute metrics like disparate impact ratio, equal opportunity difference, etc., for key demographics (if known) to ensure the model's decisions do not skew unfairly.
- *Regulatory Sandboxing and Collaboration:* Given the regulatory sensitivity of AI in lending, a prudent next step would be to collaborate with regulators or participate in sandbox initiatives. The RBI, for instance, has encouraged innovation through its Regulatory Sandbox programs in the

past. Working in a sandbox could allow us to test the AI scoring system with real (but controlled) data under regulatory oversight, obtaining feedback and ensuring compliance. Additionally, engaging with policymakers to help shape standards for explainable AI in credit scoring could be valuable.

- *Continuous Learning and Model Governance:* Implementing a continuous learning system requires careful model governance. Future improvements include setting up an automated feedback loop where model performance is tracked over time.

In essence, the future work will focus on enhancing the system's accuracy, scope, and reliability while maintaining the core principles of fairness and transparency. By incorporating richer data, better models, and collaborating with the regulatory environment, we aim to make the AI-powered credit scoring system a robust component of the digital lending infrastructure.

VI. CONCLUSION

In summary, AI-driven credit scoring for P2P lending presents a compelling case of technology addressing a real market need in India: how to extend credit in a data-driven yet fair manner to those historically left out of the system. Our research and prototype indicate that with prudent use of alternative data and rigorous attention to ethical and legal standards, such systems can materially benefit both borrowers (through access and potentially better terms) and lenders (through improved risk management). The impact could be substantial – imagine a young entrepreneur with no collateral but a healthy digital financial trail being able to get a business loan because our AI model recognizes their reliability, or a rural borrower who always pays their utility bills on time finally obtaining credit based on that behaviour. These are the outcomes we strive for.

As with any innovation, caution and continuous improvement are key. We must remain vigilant about model biases, data privacy, and evolving threats (like cybersecurity risks or new fraud tactics). But armed with the right data, algorithms, and oversight, AI-powered credit scoring can become a cornerstone in building a more inclusive financial system in India and similar economies. The work ahead will involve translating this prototype into deployed solutions, rigorously tracking real-world outcomes, and refining the approach considering feedback.

REFERENCES

[1] RiskSeal Blog – "Alternative Credit Scoring in India - How to Issue Loans to the Unbanked Population"

[2] J-PAL (Poverty Action Lab) Blog – "Using alternative data and artificial intelligence to expand financial inclusion: Evidence-based insights"

[3] Spice Route Legal – "Regulatory Round-Up – 2024" (AI in finance sector, RBI initiatives)

[4] RBI – Master Directions for NBFC-P2P Lending Platform(2017, updated 2024)

[5] U.S. Consumer Financial Protection Bureau – Guidance On Credit card Denials using AI(2023)

[6] Equifax – Explainable AI Credit Scoring and NeuroDecision® patent info.

[7] Majmudar & Partners – "Peer-to-peer lending platforms Under pressure following RBI's guidelines"