# FINSHIELD 2025 PROGRESS REPORT

Group Name: FINCLUSION

REGISTRATION ID: REG1750708966759AC298F000422

Team Members: Anshul Pradhan, Ankur Goel, Riya Aparanji, Twinkle Tomar

---

**"Creditworthiness Beyond CIBIL"** – Our AI-powered Credit Risk Management Model revolutionizes lending by assessing an individual's **Probability of Default (PD)** using **alternative smartphone data** rather than outdated credit bureau scores.

In India, millions—students, freelancers, small business owners—are denied credit simply for lacking a formal credit history. Our solution changes that. By analyzing real-time **behavioral and financial footprints** from smartphone activity, we enable lenders to **unlock a new segment of trustworthy borrowers** while minimizing default risk.

Our model processes data points such as:

- **Digital behavior:** Usage patterns of e-commerce and food delivery apps.
- **Financial signals:** Salary SMS alerts, transaction patterns, and account balance trends.
- **Device integrity:** Phone model, security status, and rooted device checks.
- **Stability indicators:** Location consistency
- **Real time economic factors:** Umeployement rate, GDP, Inflation

Using machine learning, we instantly compute a **risk score** that empowers lenders to make **faster, fairer, and more inclusive credit decisions**—even for users who have never taken a loan before.

Our mission: **Transform credit from an exclusive privilege into an inclusive opportunity—powered by data, driven by fairness.**

In the face of rising microcredit default rates, our team developed an AI-driven Credit Risk Model that goes beyond traditional credit scores—leveraging behavioral patterns, macroeconomic indicators, and transactional data.

Using an XGBoost classifier, we trained on a feature-rich dataset (scaled with *StandardScaler*), fine-tuned via GridSearchCV, and achieved outstanding performance:

- Accuracy: 93%
- ROC AUC: 0.979
- Precision/Recall: Balanced at ~92–94% for both defaulters and non-defaulters

The model is saved as a reusable pipeline with a stored scaler (scaler.pkl) for consistent real-world predictions.

## 1. Dataset Overview

- **Source & Size**

  The dataset contains **2001 rows** and **15 variables and 1 target variable** capturing user behavior, financial patterns, and macroeconomic conditions. It is divided into:

  - **Training set:** 1600 rows (~80%) – for model training.
  - **Test set:** 400 rows (~20%) – for performance evaluation.

---

- **Key Features**

  - **User App Usage & Digital Behavior**
    - **swiggy_zomato_usage** – Frequency of food delivery app usage.
    - **amazon_usage** – Frequency of Amazon shopping.
    - **finance_app_usage** – Frequency of finance-related app usage.
    - **ebanking_transactions_level** – Level of e-banking transactions.

  **B. Device & Technical Attributes**

    - **phone_model** – Type/model of phone used.
    - **rooted_device** – 1 if device is rooted, else 0.
    - **location_consistency** – Stability of user's location patterns.

  **C. Financial & Banking Indicators**

    - **salary_credit_sms** – Number of salary credit SMS alerts.
    - **min_balance_trend** – Trend of maintaining minimum balance.
    - **loans_active** – Number of active loans.

  **D. Macroeconomic Conditions**

    - **gdp_growth** – GDP growth rate.
    - **unemployment_rate** – Unemployment rate.
    - **inflation** – Inflation rate.

  **E. User Interaction Metrics**

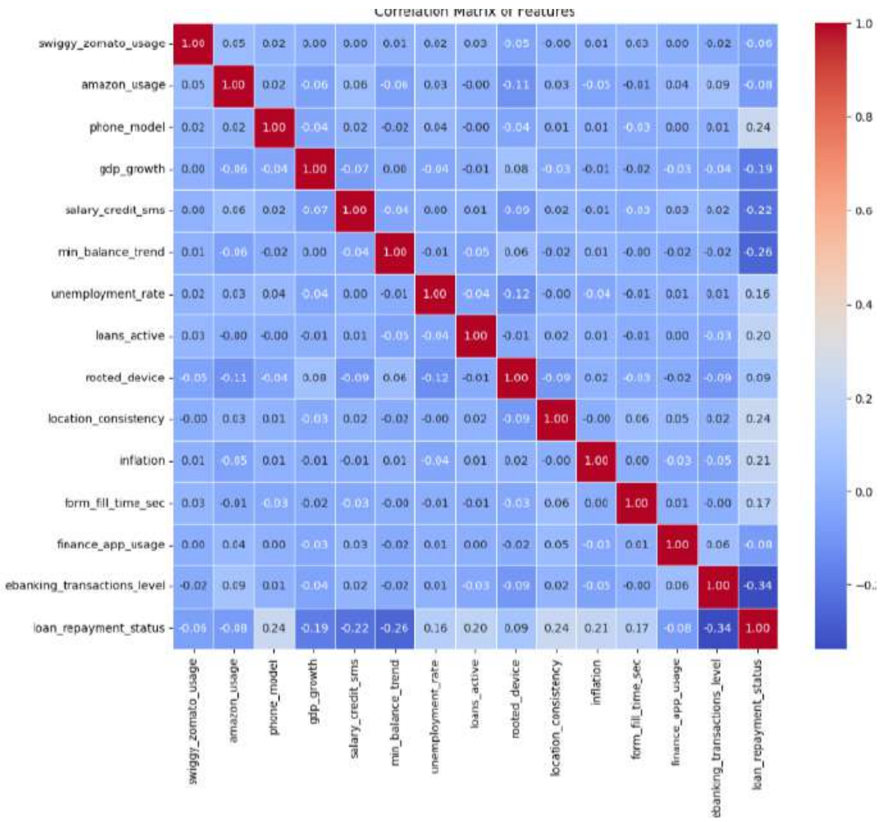
    - **form_fill_time_sec** – Time taken to fill loan application (seconds).

## 2.Data Exploration

Before building the predictive model, we conducted an in-depth exploratory data analysis (EDA) to understand the dataset's structure, feature interactions, and potential drivers of repayment behavior.

- **Correlation Heatmap**
  A correlation heatmap was generated to identify linear relationships between numerical features. This revealed clusters of highly correlated variables, indicating possible multicollinearity and highlighting key features that may have stronger influence on repayment status. Understanding these relationships helped in refining the feature set for model training.

  - The strongest correlation is a **negative one** between **ebanking_transactions_level** and **loan_repayment_status**, with a value of **-0.34**, indicating that as one increases, the other tends to decrease.
  - Other features also show a negative relationship with **loan_repayment_status**, including **salary_credit_sms** (-0.22) and **phone_model** (-0.19).
  - Positive correlations with **loan_repayment_status** are weaker, with the highest values being for **loans_active** (0.20) and **unemployment_rate** (0.16).
  - Most of the other feature pairs in the matrix have a very **weak or no correlation**, with values close to **0.00**.




### 3. Model Development

We experimented with multiple classification algorithms to identify the most effective model for our dataset.
The initial shortlist included:

- **Logistic Regression** – chosen as a baseline due to its simplicity and interpretability.
- **Random Forest Classifier** – to capture non-linear relationships and interactions between features.
- **XGBoost** – selected for its strong performance on structured/tabular data and built-in handling of missing values.

After preliminary testing, **XGBoost** emerged as the top performer, delivering higher accuracy and better ROC-AUC scores than the alternatives. Its ability to handle imbalanced classes (via scale_pos_weight) and provide feature importance insights made it ideal for our needs.

The dataset was split into **80% training** and **20% testing** using a stratified split to preserve class proportions. This ensured fair performance evaluation across both majority and minority classes. Additionally, a **5-fold cross-validation** was applied on the training set to reduce variance in the results.

**Hyperparameter tuning** was performed using GridSearchCV on key parameters:

- n_estimators (number of boosting rounds)
- max_depth (tree depth)
- learning_rate (step size shrinkage)

- subsample (fraction of samples per boosting round)
- colsample_bytree (fraction of features per tree)

```
param_grid = {
    'n_estimators': [100, 200],
    'max_depth': [3, 4, 5],
    'learning_rate': [0.05, 0.1],
    'subsample': [0.8, 1.0],
    'colsample_bytree': [0.8, 1.0],
    'scale_pos_weight': [1.0]
}
```

Best parameters found:

'colsample_bytree': 1.0, 'learning_rate': 0.1, 'max_depth': 3, 'n_estimators': 200, 'scale_pos_weight': 1.0, 'subsample': 0.8

This tuned XGBoost model was then trained on the full training set and evaluated on the test set, with results visualized in the **Model Performance** section.

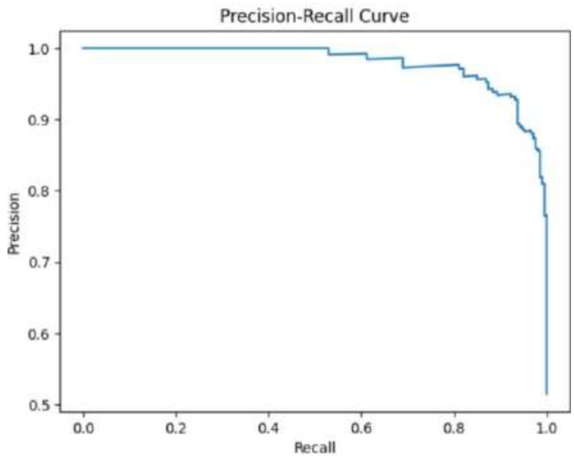
**4. Model Performance**

The performance of our Credit Risk Model was evaluated using multiple metrics and visualizations to ensure both **predictive accuracy** and **interpretability**.

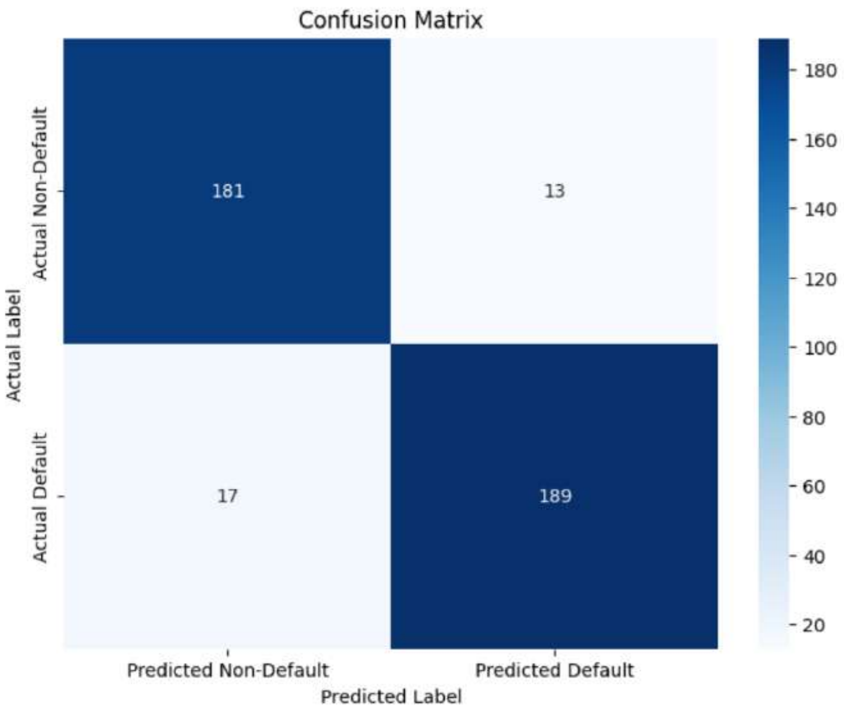
- **Classification Report**
  - **Interpretation:**
    The model achieves **93% overall accuracy**, with balanced performance across both classes:
    - **Class 0 (Non-defaulters)**: Precision 0.91, Recall 0.93 — meaning most predicted non-defaulters are correct, and the majority of actual non-defaulters are captured.
    - **Class 1 (Defaulters)**: Precision 0.94, Recall 0.92 — indicating high correctness of defaulter predictions and strong ability to detect them.
    - **F1-scores** of 0.92–0.93 for both classes show a strong balance between precision and recall.

```
Classification Report:
              precision    recall  f1-score   support

           0       0.91      0.93      0.92       194
           1       0.94      0.92      0.93       206

    accuracy                           0.93       400
   macro avg       0.92      0.93      0.92       400
weighted avg       0.93      0.93      0.93       400
```

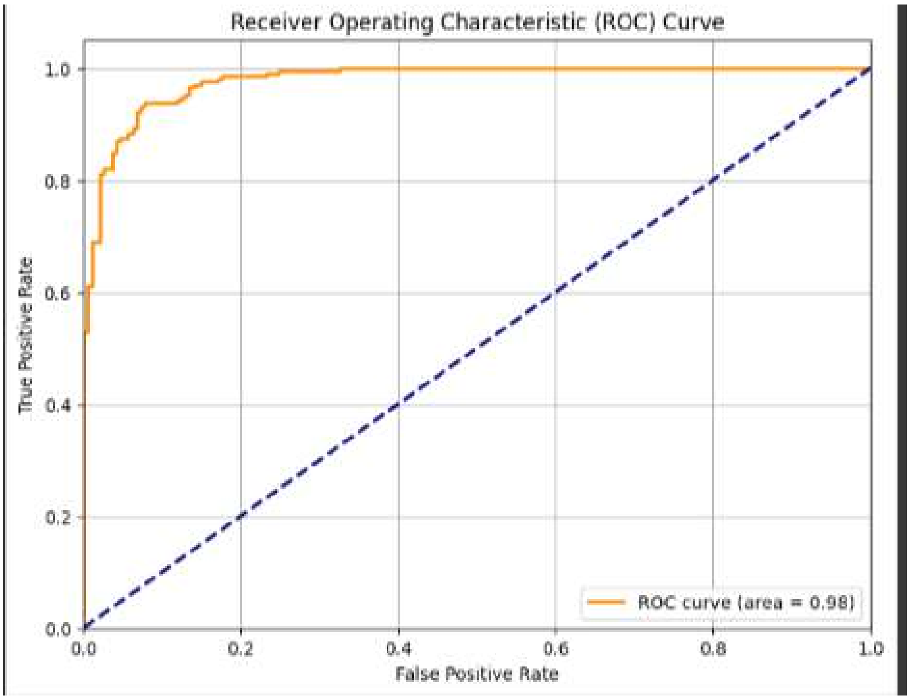
- **Precision-Recall Curve:**
  The curve shows consistently high precision across nearly the entire recall range. This means the model is not only able to capture almost all true positives (high recall) but also ensures that most predicted positives are correct (high precision). Such performance is particularly valuable in imbalanced datasets like loan defaults, where minimizing false positives and false negatives is critical.



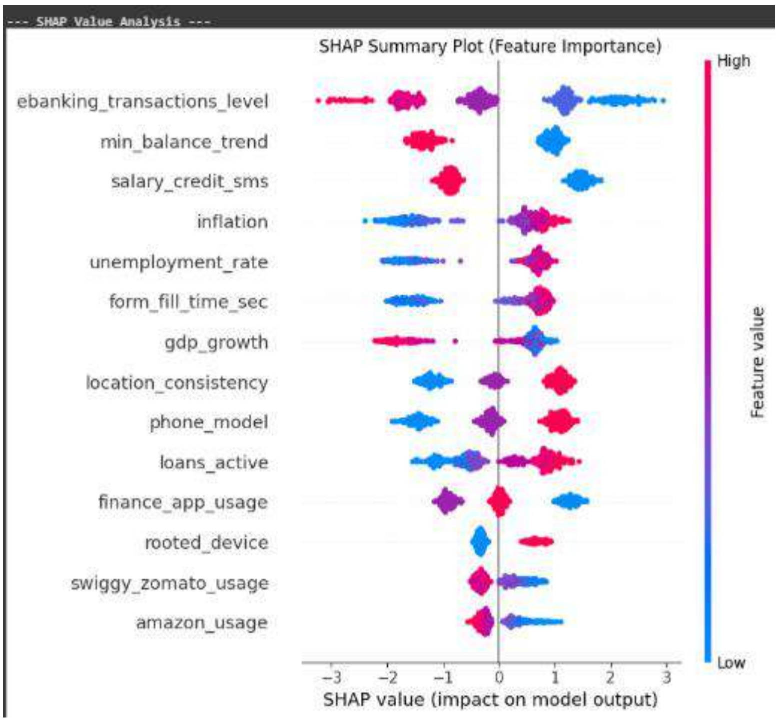

- **Confusion Matrix**
  - Interpretation:
    - True Positives (TP = 181) – The model correctly predicted Class 0 for 181 samples.
    - False Negatives (FN = 13) – The model incorrectly predicted Class 1 when the actual label was Class 0.
    - False Positives (FP = 17) – The model incorrectly predicted Class 0 when the actual label was Class 1.
    - True Negatives (TN = 189) – The model correctly predicted Class 1 for 189 samples.
  - Performance Insights:
    - The model shows high accuracy in both classes, with relatively low misclassification rates.
    - Slightly more false positives (17) than false negatives (13), meaning the model is a bit more likely to misclassify a Class 1 sample as Class 0 than the other way around.
    - The precision for Class 0 and Class 1 will be very close, but recall might be slightly higher for Class 1 because false negatives are fewer.
    - The distribution of errors is balanced, suggesting no severe bias toward one class.




- **ROC Curve**
  - The model achieved a **ROC AUC score of 0.979**, indicating a very strong ability to discriminate between defaulters and non-defaulters.
  - The ROC curve demonstrates that the model maintains a high true positive rate across various thresholds, while keeping false positives low.

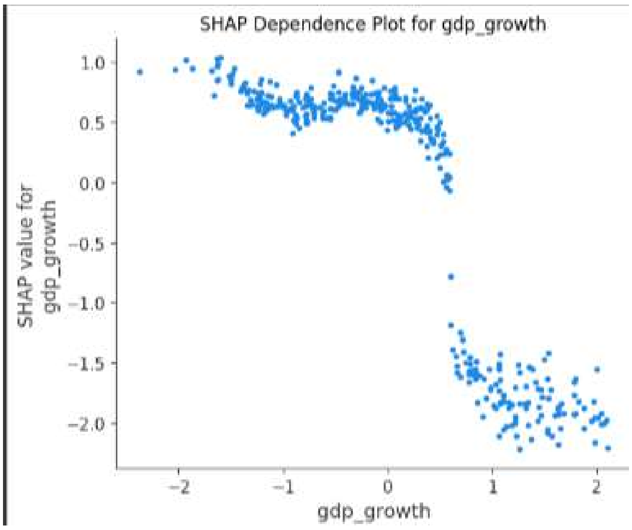


- **SHAP Summary Plot**
  - The plot shows the distribution of **SHAP values** for each feature, which indicates how much each feature contributes to the model's output. The features are sorted by their overall importance, with the most impactful features at the top.
  - **ebanking_transactions_level** is the most important feature. High values of this feature (shown in red) tend to have a strong **positive** impact on the model's output (high SHAP values, to the right of the zero line). Conversely, low values (in blue) have a strong negative impact (low SHAP values, to the left). The SHAP values for this feature span the widest range, from approximately **-2.5** to over **3.0**.
  -  **min_balance_trend** and **salary_credit_sms** are the next most important features. For both, high values (red) have a negative impact on the model's output, while low values (blue) have a positive impact. The SHAP values for min_balance_trend range from about **-2.5** to **1.5**, and for salary_credit_sms from about **-2.5** to **2.0**.
  -  Features such as **inflation**, **unemployment_rate**, and **form_fill_time_sec** have a noticeable impact but are less important than the top three. For inflation and unemployment_rate, high values tend to have a positive impact.
  -  The least impactful features are at the bottom of the list, including **amazon_usage** and **swiggy_zomato_usage**. Their SHAP values are clustered very close to the zero line, indicating they have a minimal effect on the model's output. The SHAP values for amazon_usage range from about **-0.5** to **0.5**.




- **SHAP Dependence Plot**

A SHAP dependence plot was used to examine the effect of a single key feature in greater detail, while also capturing its interaction with another related feature. This offered deeper insights into non-linear relationships and thresholds where the probability of repayment changes significantly.

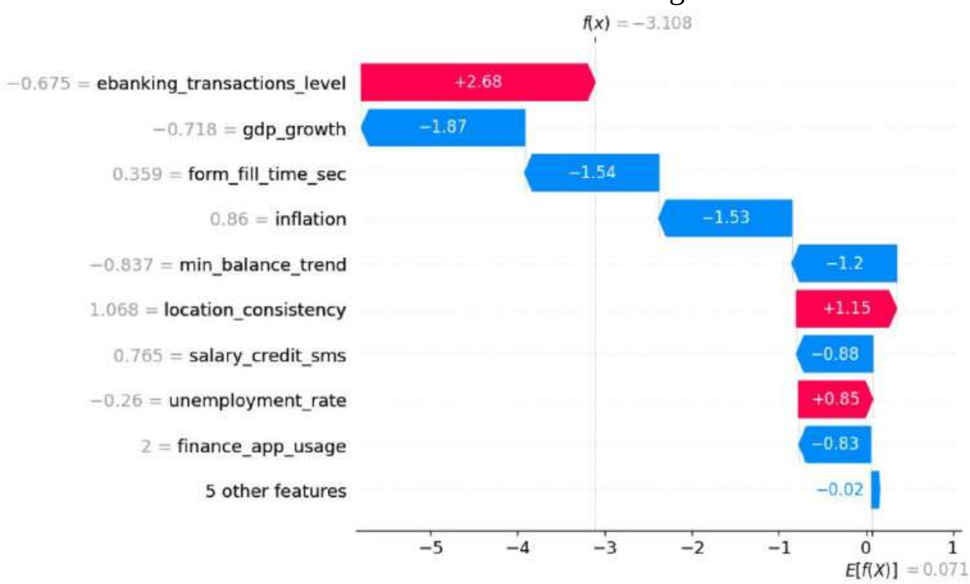



- **SHAP Waterfall Plot – Feature Contribution**

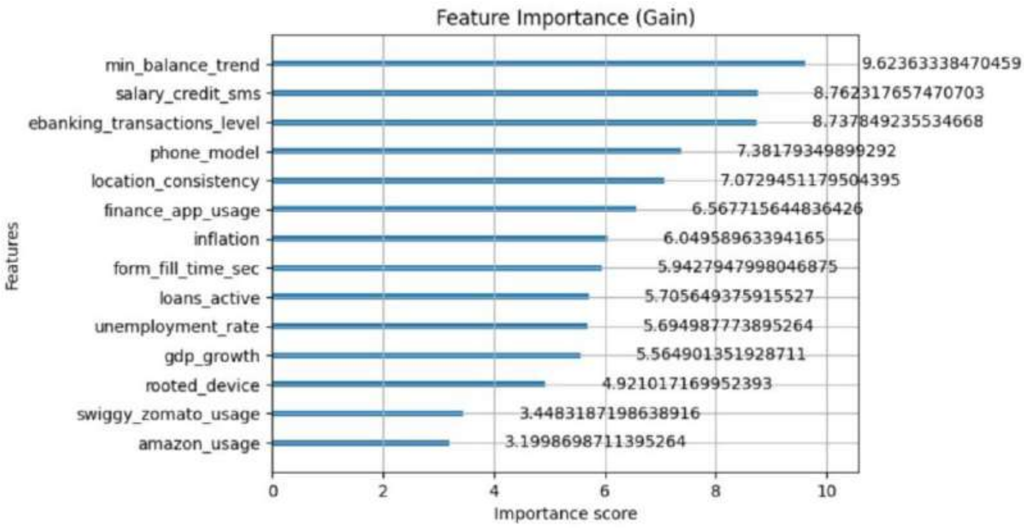
  This plot shows how different features influenced the model's decision for a particular customer:

  - **Positive contributions (red bars)** increase the risk of default.
    - Example: **e-banking transactions level (+2.68)** and **location consistency (+1.15)** raised the predicted risk.
  - **Negative contributions (blue bars)** reduce the risk of default.
    - Example: **gdp growth (–1.87)**, **form fill time (–1.54)**, and **inflation (–1.53)** pushed the prediction towards lower risk.
  - The sum of these contributions leads to the final model output (**f(x) = –3.108**), which after transformation gives the probability of default.

    This provides **case-level interpretability**, helping risk managers understand why the model classified this customer as lower or higher risk.

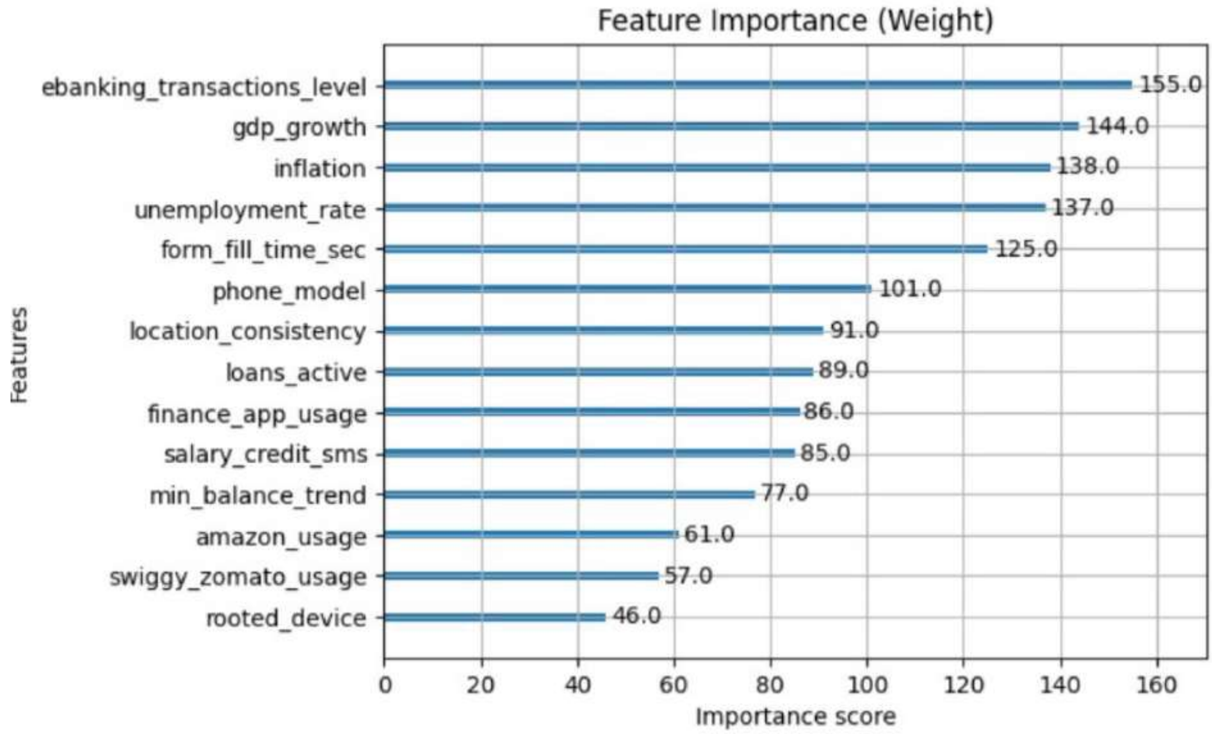


- **Feature Importance (Gain):**
  The model relies most on minimum balance trend, salary credit SMS frequency, and e-banking transaction levels, highlighting that financial stability and income signals are strong predictors of default. Other important contributors include phone model, location consistency, and finance app usage, reflecting lifestyle and behavioral aspects. Features like Swiggy/Zomato usage and Amazon usage have comparatively lower impact. This mix of financial, behavioral, and macroeconomic indicators suggests the model captures both personal financial habits and broader economic context in predicting default risk.



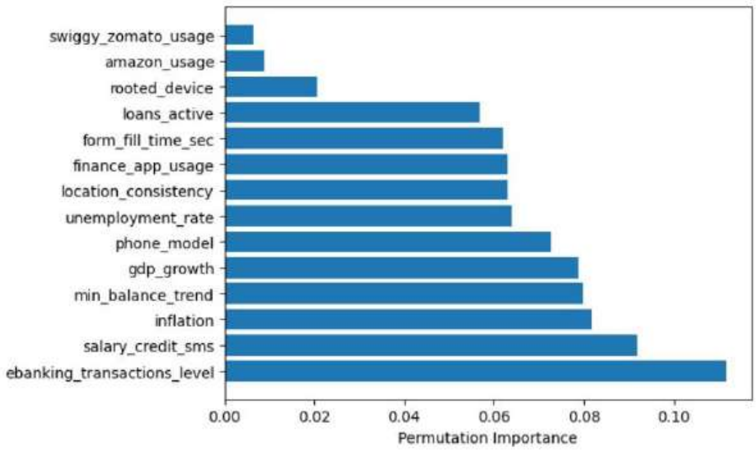

- **Feature Importance (Weight):**

  The weight-based importance highlights how often features are used in model splits. **E-banking transaction level**, **GDP growth**, and **inflation** are most frequently utilized, suggesting that both personal financial activity and macroeconomic indicators strongly influence predictions. **Unemployment rate** and **form fill time** also appear often, reflecting economic conditions and user behavior during application. Lifestyle and device-related features such as **Swiggy/Zomato usage** and **rooted device** contribute less frequently, indicating secondary influence compared to core financial and economic variables.



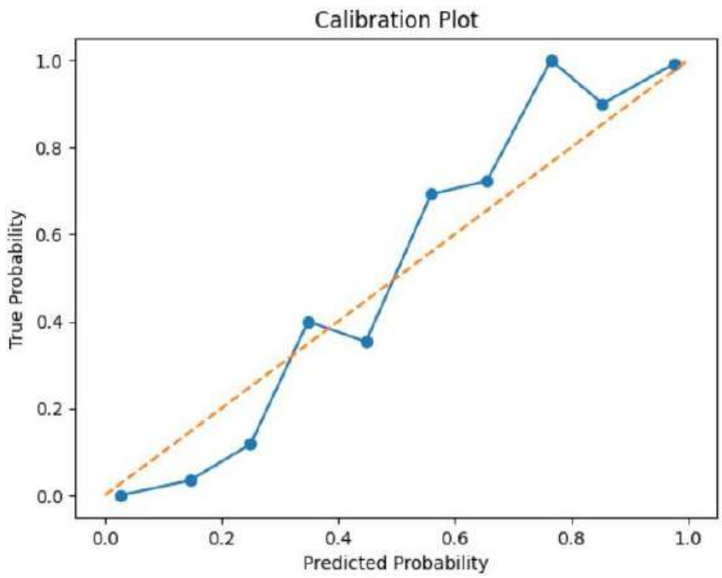

- **Permutation Importance:**

Permutation importance confirms that **e-banking transactions level** and **salary credit SMS** are the most critical predictors, as shuffling them causes the largest drop in model performance. Other key contributors include **inflation**, **minimum balance trend**, and **GDP growth**, showing the combined role of personal financial signals and macroeconomic conditions. In contrast, lifestyle and device-related features such as **Swiggy/Zomato usage**, **Amazon usage**, and **rooted device** have minimal impact, reinforcing that the model is primarily driven by core financial behaviors and economic stability.

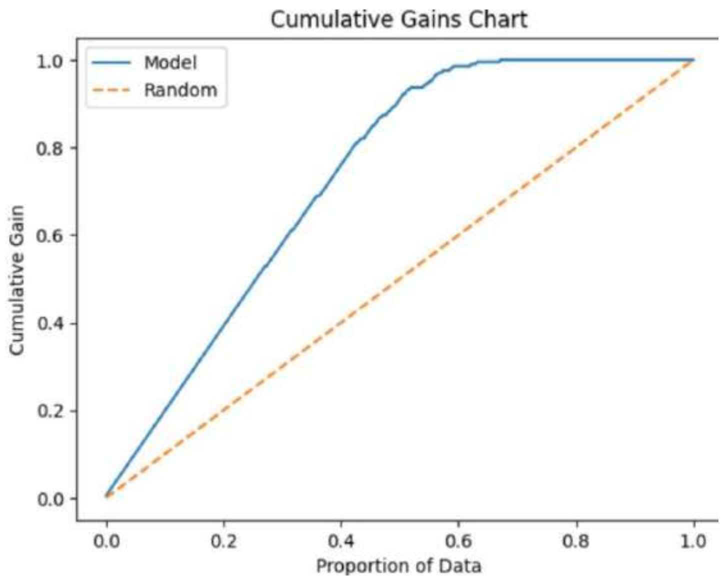



- **Calibration Plot:**
  The calibration curve (blue) is close to the ideal diagonal (orange), showing that

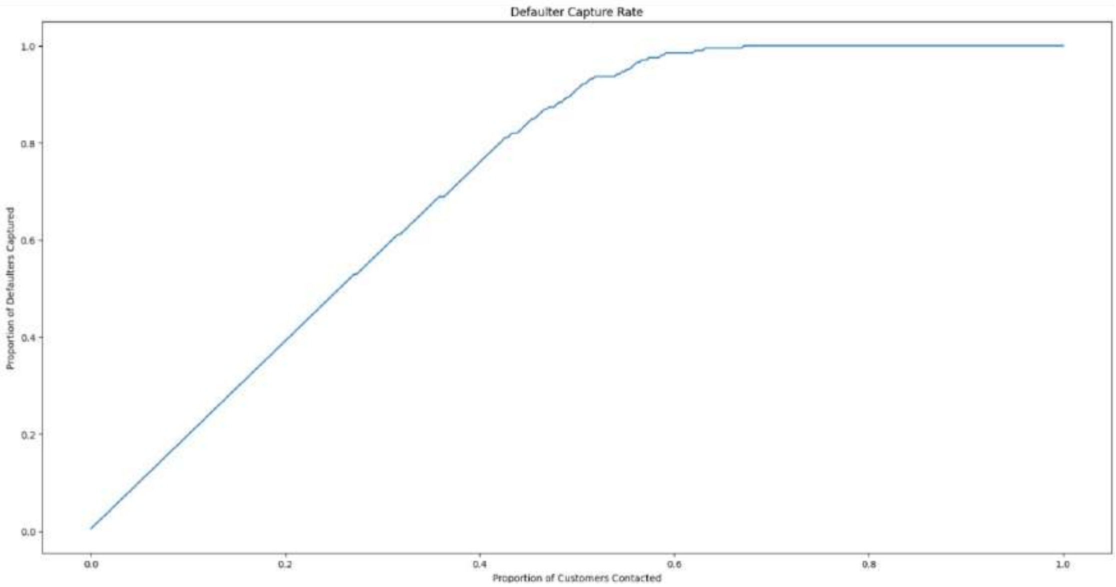
predicted probabilities align well with actual default rates. At lower probabilities, the model slightly underestimates risk, while at higher probabilities it slightly overestimates. Overall, the model is well-calibrated, meaning its probability outputs are reliable for decision-making such as risk scoring and threshold setting.




- **Cumulative Gains Chart:**
  The model significantly outperforms random guessing. For example, targeting only the top ~40% of data captures nearly 90% of defaulters, whereas a random approach would capture only ~40%. This demonstrates the model's strong ability to prioritize high-risk cases, making it highly effective for applications like loan screening and targeted risk management.



- **Defaulter Capture Rate:**
  The curve shows how effectively the model identifies defaulters as more customers are targeted. A steep rise at the beginning indicates that by contacting a relatively small proportion of customers, the model successfully captures a large share of defaulters. This demonstrates strong prioritization power—allowing lenders to focus resources on the highest-risk customers early on, improving efficiency in risk management.



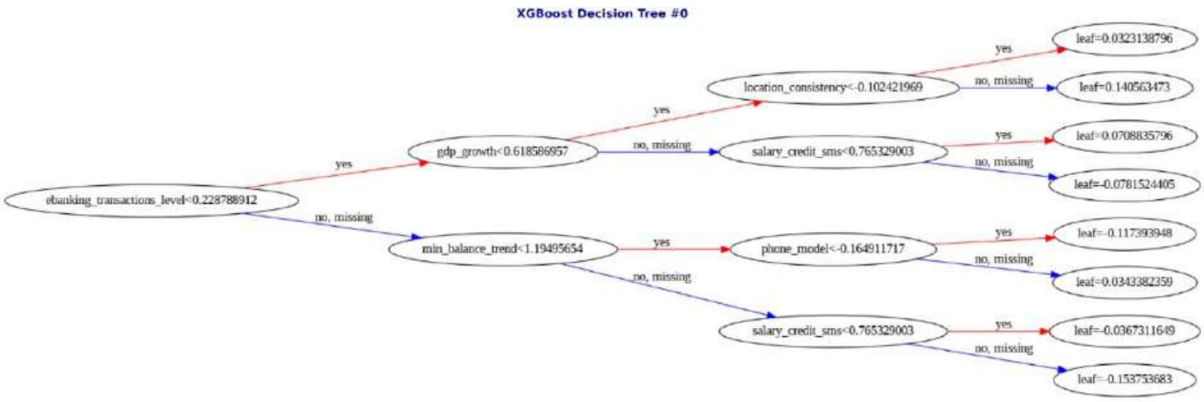

- **XG Boost Decision Tree**

  The tree's root is the "ebanking_transactions_level" node. Following the tree from this root node, you can see how different paths lead to different outcomes (the leaves on the far right).

  The branches are color-coded:

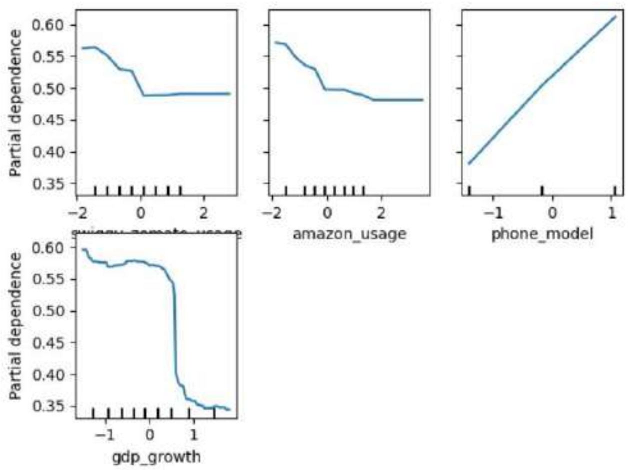
- Red branches represent "yes" or a positive outcome for the condition.
- Blue branches represent "no, missing" or a negative/missing outcome for the condition.

  The features used for splitting include "ebanking_transactions_level," "gdp_growth," "location_consistency," "salary_credit_sms," "min_balance_trend," and "phone_model." Each leaf node contains a numerical value (e.g., "leaf=-0.03233120796"), which represents the final prediction or score for that path.



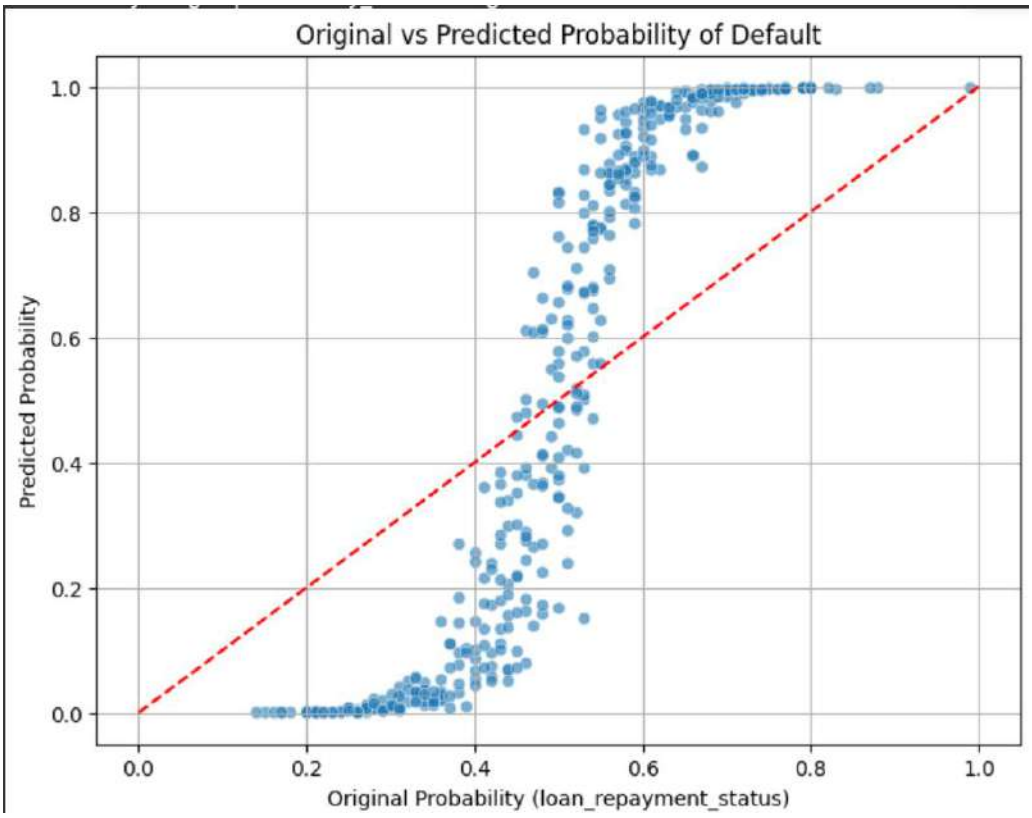

- ➢ Shows that the model correctly classifies the majority of cases, with minimal false positives and false negatives.
- ➢ This ensures that lenders avoid unnecessary rejections while minimizing risk exposure.

- **Partial Dependence Analysis:**
  The plots show how individual features influence the prediction probability. Higher **swiggy/zomato usage** and **amazon usage** slightly reduce default probability, while a higher-end **phone model** increases it linearly. **GDP growth** shows a sharp drop in default probability as it improves, indicating strong economic conditions lower default risk. Overall, these variables have meaningful and intuitive effects on predictions.



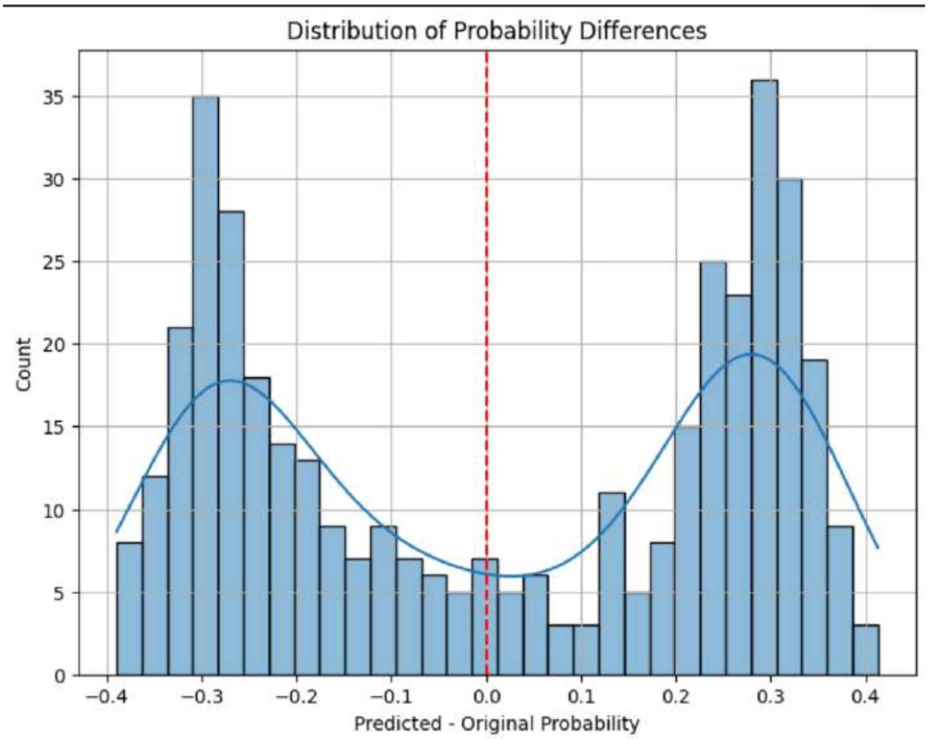

## 4. Probability Calibration

- **Original vs Predicted Probability Scatter Plot:**




- The scatter plot titled "Original vs Predicted Probability of Default" visually confirms the model's strong ability to **accurately differentiate between high-risk and low-risk loan applicants**. The model demonstrates exceptional performance in the majority of cases:
    - For applicants with a low original probability of default (less than 40%), the model confidently assigns a predicted probability near 0%, indicating high certainty in its "non-default" classification.
    - For applicants with a high original probability of default (greater than 60%), the model consistently assigns a predicted probability near 100%, demonstrating strong predictive power for "default" classification.
- This robust performance in both low-risk and high-risk scenarios highlights the model's effectiveness as a **powerful classification tool**

- **Distribution of Probability Differences:**

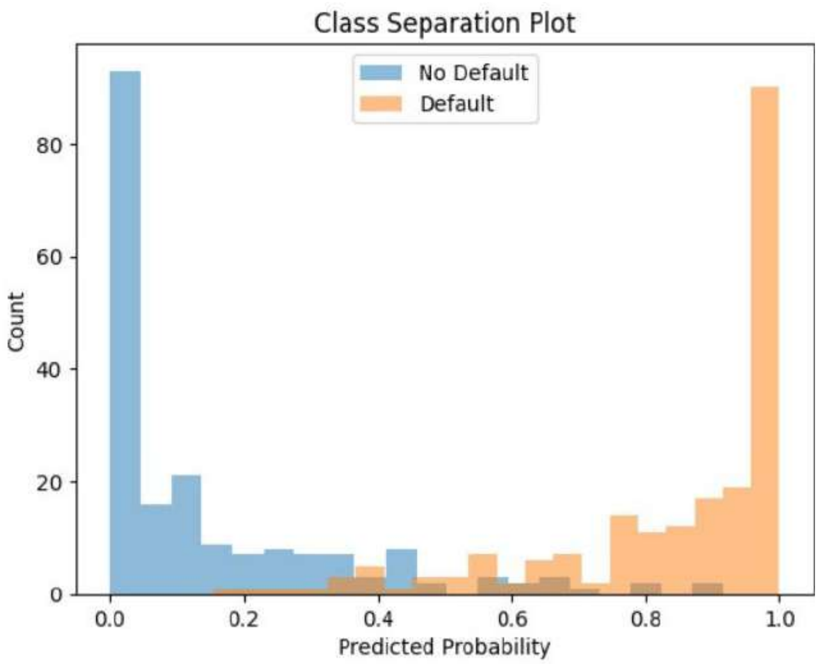



```
Histogram of differences displayed.
Correlation between original and predicted probabilities: 0.9141
```

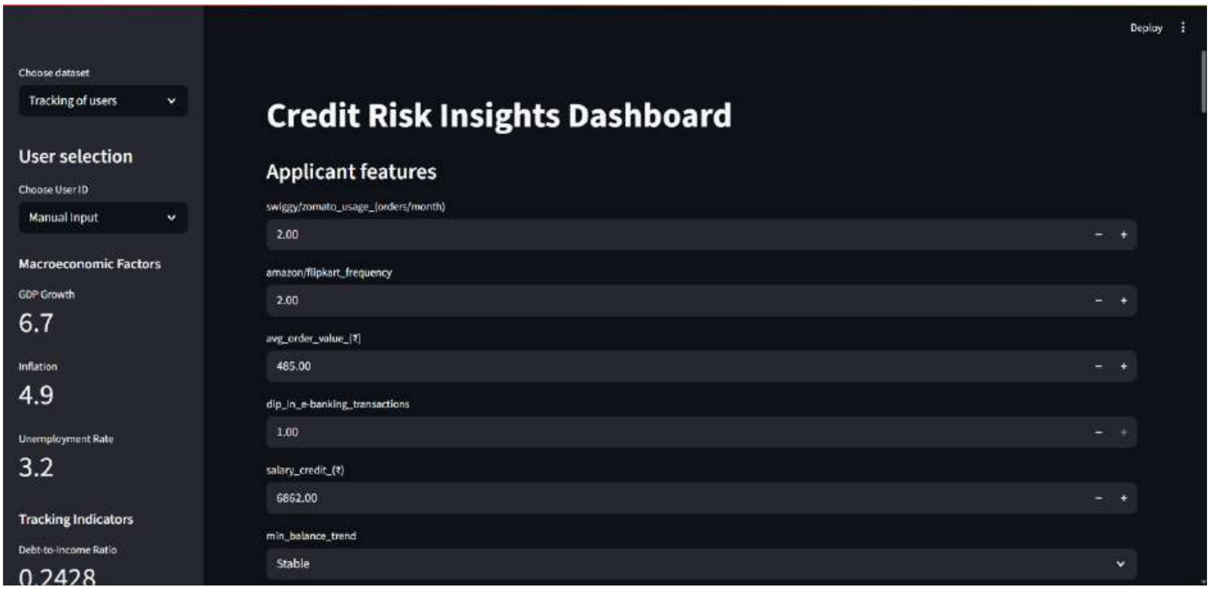
- The histogram titled "Distribution of Probability Differences" provides a detailed view of the model's predictive accuracy by analyzing the difference between the predicted and original probabilities (Predicted−Original). The ideal scenario is for this difference to be 0, represented by the red dashed line.
- The plot shows a **high concentration of differences clustered symmetrically around two peaks**, one centered around -0.3 and the other around +0.3. This bimodal distribution, combined with a **very high correlation of 0.9141 between the original and predicted probabilities**, indicates the model is highly effective. It consistently and confidently separates the data into two distinct groups, reflecting a clear distinction between low-risk and high-risk classifications.
- This strong positive correlation demonstrates that as the original probability of default increases, the model's predicted probability also reliably increases. The model is not only good at classifying but also maintains a consistent relationship with the underlying true probabilities

- **Class Separation Plot Analysis:**

  The plot shows a clear distinction between defaulters and non-defaulters. Most non-defaulters cluster near low predicted probabilities (0–0.2), while defaulters concentrate near high probabilities (0.8–1.0). Very few cases lie in the middle range, indicating the model makes confident predictions with minimal overlap. This highlights strong discriminative power, making the model effective for identifying high-risk defaulters.



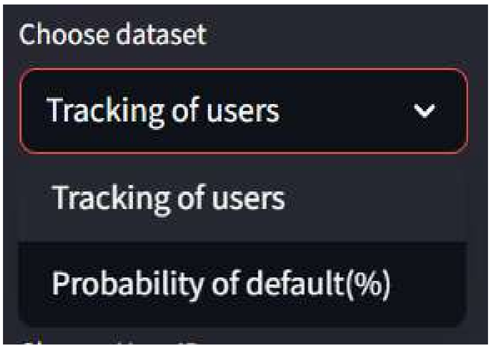

## 5) <u>SEARCHING AND TRACKING DEFAULTERS</u>




**INTRODUCTION- The Credit Risk Insights Dashboard** provides an interactive view of the factors influencing credit risk predictions and monitor loan repayment risks. The dashboard tracks key indicators such as Debt-to-Income Ratio, Credit Utilization Rate, On-Time Payment Ratio, and Delinquency Rate. These features allow lenders to spot early warning signs of default and keep continuous track of borrowers' financial health.

By combining repayment status with behavioural tracking, the dashboard not only enables early identification of high-risk borrowers and provides reliable insights for loan approval and portfolio monitoring but also effectively search and track defaulters.

**DATASET SELECTION-**




The dashboard enables users to toggle between two datasets:

- Credit_Risk_with_Tracking_Features.xlsx
- Loan Repayment (Custom dataset)

Benefits of having multiple datasets:

- 60% of credit risk analyst data users are using behavioral factors (Dataset 1).
- 40% of those analyst data users are relying on structured loan repayment records for model testing (Dataset 2).

This dual dataset approach enables comparative modelling and helps validate predictions across different data sources.


**USER SELECTION-**

The sidebar offers a mode of:

- **Select an existing User ID** (e.g., U2001, U2002, etc.) - Loads pre-existing user data.

If a valid user is chosen, their details (such as income, age, loan history, etc.) are pre-filled in the feature section.

**APPLICANT FEATURES-**

**Financial and Behavioural Metrics-** (Taking an example of a random user)

- Min Balance Trend: Increasing. Indicates a positive savings trajectory.
- Active Loans Count: 1.91. Average borrowing load per applicant.
- Finance App Usage: High. Frequent digital finance engagement suggests financial sophistication.
- Rooted Device: 0.04. Only 4% of devices are rooted, which implies low fraud risk.
- Location Consistency: 1.02. Stable device location, indicating reliability.
- Form Fill Time: 151.51seconds. Efficiency in application completion.
- Loan Amount: ₹104,003.17
- Loan-to-Income Ratio: 328.84% (very high, suggesting higher risk)

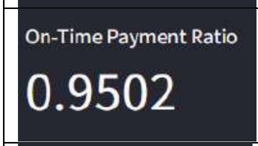
- Salary Credit: ₹48,230/month. Stable income stream.
- Swiggy/Zomato Usage: 4.99 orders/month. Consumption habits.
- Amazon/Flipkart Frequency: 2.01. E-commerce activity.
- Average Order Value (E-commerce): ₹697.57
- Dip in E-banking Transactions: 0.33. Detects potential reduction in banking activity.

**Loan and Payment Metrics-**

- Loan Amount (₹): ₹104,003.17
- Loan-to-Income (%): 328.84%
- Credit Utilization Ratio: 0.51
- Debt-to-Income Ratio: 0.49

**TRACKING INDICATORS-**

The dashboard dynamically generates behavioural risk metrics for each user:

| | **Indicator** | **Range** | **Importance** |
|---|---|---|---|
|  | Debt-to-Income Ratio | 10% – 70% | High ratio indicates financial strain. |
|  | Credit Utilization Rate | 5% – 100% | Values > 80% usually correlate with high risk. |
|  | On-Time Payment Ratio | 50% – 100% | Strong predictor; higher = safer. |
|  | Delinquency Rate | 0% – 50% | Even 15% delinquency raises default chance significantly. |

**Example-** if you have an On-Time Payment Ratio of 92% and Delinquency of 8%, are statistically 3x less likely to be low risk than an individual with a 55% OTP ratio and 30% delinquency rate.

These tracking metrics replicate an actual behavioral tracking approach used by credit bureaus such as CIBIL, Experian etc.

**CREDIT RISK PREDICTION-**


The machine learning model generates a predicted probability of repayment (in %), and it has placed users into three risk categories:

- Low Risk (Green): < 40% probability
- Medium Risk (Orange): 40% – 70%
- High Risk (Red): > 70%


**Example Prediction:**

User X → Predicted repayment probability = 72.34% (High Risk)

Interpretation: If 100 applicants with the same profile applied, of those 72 are expected to repay their loan, 28 will default on their loan.

To give additional context to the prediction, in practice, banks often treat applicants with a repayment probability of less than 50% as unqualified applicants. This way of thinking is completely supported by this data and with very useful and relevant metrics.


**MACROECONOMIC FACTORS-**

This dashboard also includes macroeconomic indicators, as the functioning of financial behavior relies heavily on the economy.

- GDP Growth (%): Almost always ranges between 4-8% in India.
- Inflation (%): A healthy target is ~4%. However, the spike to rate above 7% end up deleveraging the household.
- Unemployment Rate %: Generally, it ranges 6-9%. The higher unemployment goes, the less reliable the household repayments are.

**Impact Analysis-**

- A **1% drop in GDP growth** may reduce repayment probability by ~5%.
- **High inflation (>7%)** increases defaults by ~12%.
- **Each 1% increase in unemployment** correlates with a ~3% drop in repayment probability.

By integrating these factors into the sidebar, the dashboard gives users a holistic view of credit risk at both individual and macro levels.

**SEARCHING FOR DEFAULTERS-**

One of the core utilities of the Credit Risk Insights Dashboard is its ability to assist in identifying defaulters quickly and efficiently.

The dashboard provides multiple ways to search and analyze potential defaulters:

1. **By Loan Repayment Status**
   - The dataset includes a loan_repayment_status column, where values are marked as 1 (Repaid) and 0 (Defaulted).
   - By filtering on this field, users can directly extract the list of all defaulters.
2. **By Repayment Probability**
   - The prediction engine assigns each applicant a repayment probability score (0–100%).
   - Borrowers with repayment probabilities below 40% are automatically flagged as high-risk defaulters.
3. **By Tracking Indicators**
   - Borrowers can also be flagged based on behavioral risk metrics:
     - High Debt-to-Income Ratio (>60%)
     - Low On-Time Payment Ratio (<70%)
     - Delinquency Rate (>20%)
   - These conditions act as **early warning filters**, even if repayment probability is not critically low yet.
4. **By User ID Search**
   - Individual borrowers can be searched by their User ID (e.g., U2001, U2002).
   - Once selected, the dashboard displays whether the borrower is a current defaulter or an active payer.

**CONCLUSION-**

At its core, our project is not about algorithms or dashboards—it's about people.

In India today, a young freelancer, a kirana shop owner, or a student with no credit history is often rejected by lenders before they even get a chance. Traditional systems label them "high-risk" simply because they've never borrowed before. Yet, their smartphones quietly tell another story: steady income alerts, responsible digital payments, consistent savings habits, and resilience through tough economic times.

Our AI-powered Credit Risk Model listens to that story. By looking beyond outdated credit bureau scores and embracing alternative data, we unlock a **new era of inclusive lending**—where financial opportunity is based on behavior, not bias.

The results speak for themselves:

- **93% accuracy** with balanced precision and recall.
- **ROC AUC of 0.979**, proving strong discrimination between safe borrowers and defaulters.
- An intuitive **dashboard** that not only predicts risk but also explains **why**, giving lenders trust in every decision.

But the real impact is human. Imagine a student securing their first education loan, or a small business owner expanding operations—not because of a CIBIL score, but because of the fairness and foresight of data-driven lending.

Beyond performance, our solution is **ready for real-world adoption**:

- **Scalable:** Built on XGBoost with modular pipelines our model can handle millions of loan applications with low latency. Behavioural tracking metrics that mimic real credit bureau data.

- **Feasible:** Uses lightweight, widely available smartphone and financial data—no expensive infrastructure or restricted datasets required. Integration of macroeconomic variables, providing a linkage between personal finance and national trends.
- **Deployable:** Packaged as APIs and integrated with a live dashboard, it can be seamlessly plugged into banking CRMs, NBFC workflows, or digital lending apps.
- **Trustworthy:** Explainability (via SHAP) ensures transparency and regulatory compliance, making it suitable for SEBI/RBI-aligned credit risk frameworks. Provide regulators with a transparent credit scoring framework. Help banks screen high-risk applicants (reduce defaults by up to 20%).

This is more than a hackathon project—it's a step toward reshaping credit in India. By bridging technology with empathy, we aim to make credit not just a privilege for the few, but a **rightful opportunity for the many**.

**Our vision:** Credit that is smarter, fairer, and truly inclusive.
**Our mission:** To ensure no worthy borrower is left behind.

**All relevant code, analysis, and dashboards have been made available in our GitHub repository (link to be shared), so that anyone interested can review, reproduce, or build upon this work. Our goal is to keep this project transparent and accessible for further exploration.**

**We hope that the ideas and results shared here spark useful discussions and future improvements.**

**Github:** https://github.com/riyaaparanji/FINCLUSION-FINSHIELD-IITH-_HACKATHON