Predictive Model Plan – YUVRAJ KUMAR GOND

1. Model Logic (Generated with GenAI)

To predict customer delinquency, we developed a supervised machine learning model using a Decision Tree Classifier. The goal is to classify whether a customer is likely to default based on historical financial data. The process begins by importing and preprocessing the dataset — selecting relevant features like missed payments, credit utilization, income, and account age. These features were identified during EDA as key indicators of financial risk.

The model logic involves splitting the dataset into training and testing subsets, fitting a Decision Tree algorithm on the training data, and then making predictions on the test set. The decision tree works by learning a series of logical rules (if-else conditions) that best separate delinquent from non-delinquent customers. This structure allows for easy traceability and interpretability of model decisions, making it ideal for financial applications. GenAI tools helped guide the modeling structure and logic to ensure alignment with business needs and compliance.

---

2. Justification for Model Choice

The Decision Tree model was selected because it provides a strong combination of accuracy, transparency, and operational simplicity. In financial services, where auditability and regulatory compliance are essential, decision trees offer a clear advantage through visual, rule-based outputs that can be easily explained to stakeholders. The model handles both numerical and categorical variables without requiring extensive preprocessing, which streamlines implementation.

Additionally, the interpretability of the model supports ethical AI practices by making it easier to detect and correct biases in predictions. While advanced models like XGBoost or neural networks can provide marginally higher accuracy, they sacrifice transparency and are more complex to monitor. For Geldium, a Decision Tree offers the right balance between performance, explainability, and real-world deployment feasibility in credit risk management.

---

3. Evaluation Strategy

The model was evaluated using standard classification metrics: accuracy, precision, recall, F1-score, and ROC-AUC. Accuracy reflects the overall prediction quality, while precision and recall assess how well the model identifies actual delinquents without generating false alarms. The F1-score is particularly valuable in this imbalanced classification problem, as it balances false positives and false negatives. ROC-AUC provides an overall view of the model’s ability to distinguish between risky and non-risky customers.

To ensure fairness, we plan to conduct subgroup analysis to check whether model predictions are biased against certain demographic or financial groups. If disparities are detected, mitigation strategies such as rebalancing the dataset or applying fairness-aware learning techniques will be considered. Ethical considerations are also central, as we aim to build a system that does not unfairly penalize vulnerable customers and offers transparency in how decisions are made.