

Predictive Model Plan – Student Template

Use this template to structure your submission. You can copy and paste content from GenAI tools and build around it with your own analysis.

1. Model Logic (Generated with GenAI)

Use a GenAI tool (e.g., ChatGPT, Gemini) to generate the logic or structure of your predictive model.

- You may include pseudo-code, a step-by-step process, or a simplified code snippet.
- Briefly explain what the model is designed to do.

Paste your GenAI-generated output below or describe the logic in your own words:

.....

We propose using a Random Forest Classifier to predict customer delinquency based on key financial and behavioral features such as credit score, debt-to-income ratio, and missed payments. This model processes customer inputs, computes non-linear patterns through an ensemble of decision trees, and outputs a probability of delinquency, allowing Geldium to proactively intervene with at-risk customers.

.....

Key Input Features:

- *Credit_Score*
- *Credit_Utilization*
- *Missed_Payments*
- *Debt_to_Income_Ratio*
- *Account_Tenure*

2. Justification for Model Choice

Explain why you selected this specific model type (e.g., logistic regression, decision tree, neural network). Consider:

- Accuracy
- Transparency
- Ease of use or implementation

- Relevance for financial prediction
- Suitability for Geldium’s business needs

We selected the **Random Forest Classifier** due to its strong performance in classification tasks with mixed data types and its robustness to missing values and outliers. Unlike logistic regression, which assumes a linear relationship between features and outcome, Random Forests can capture complex, non-linear patterns—critical for modeling financial risk where relationships between income, credit usage, and default are not always straightforward. Additionally, Random Forests offer interpretable feature importances, which supports transparency and compliance with financial regulations. This model is scalable and relatively easy to deploy, making it a strong candidate for Geldium’s operational needs, particularly in balancing predictive accuracy with explainability for internal and regulatory stakeholders.

3. Evaluation Strategy

- Outline how you would evaluate your model’s performance. Include:
- Which metrics you would use (e.g., accuracy, precision, recall, F1 score, AUC)
 - How you would interpret those metrics
 - Any plans to detect or reduce bias in your model
 - Ethical considerations in making predictions about customer financial behavior

To assess the model’s performance, we propose the following strategy:

Key Evaluation Metrics:

Metric	Purpose
Accuracy	Overall correctness of predictions.
Precision	Fraction of predicted delinquents that were truly delinquent (minimizes false positives).
Recall	Fraction of actual delinquents correctly identified (minimizes false negatives).
F1 Score	Harmonic mean of precision and recall—especially useful for imbalanced datasets.
AUC-ROC	Measures the ability to distinguish between classes at various thresholds.

Bias and Fairness Checks:

- **Demographic Subgroup Analysis:** Evaluate model performance across age, income levels, and tenure groups to ensure no group is disproportionately flagged.
- **Fairness Metrics:**
 - *Equal Opportunity:* Ensure similar recall across groups.
 - *Disparate Impact Ratio:* Check if model unfairly favors or disfavors one group over another.
- **Bias Mitigation Strategies:**
 - *Rebalancing class weights.*

- *Using fairness-aware post-processing methods (e.g., reject option classification).*

Ethical Considerations:

- *Transparently communicate model use and limitations to stakeholders.*
- *Avoid automated decision-making without human oversight, especially in cases of predicted delinquency.*
- *Regularly audit and update the model to adapt to changing financial behaviors and ensure ongoing fairness.*