**Model Documentation: Credit Scoring Model**

**1. Introduction and Overview:**

This document provides documentation for a credit scoring model developed using logistic regression.

**2. Data Source:**

The data used in this model is sourced from an Excel file, "a\_Dataset\_CreditScoring.xlsx." that has details on borrowers’ credit bureau records, captured at the time of loan application and final outcomes on these loans, viz., loan turned good or bad.

**3. Data Preprocessing**:

Data preprocessing includes handling missing values, standardization, and splitting data into training and testing sets.

**4. Model Details:**

Model Type: Logistic Regression

**5. Data Summary:**

*Number of Rows:* 3,000 *Number of Columns:* 30

Columns Dropped: The "ID" column has been dropped as it contains Personally Identifiable Information (PII).

**6. Handling Missing Values:**

Missing values have been imputed with the mean values of the respective columns.

**7. Data Splitting:**

The dataset was split into training and testing sets with an 80:20 ratio.

**8. Feature Standardization:**

Features have been standardized using the StandardScaler.

**9. Exporting Normalization Coefficients:**

The normalization coefficients have been saved for later use in prediction.

**10. Model Training:**

The logistic regression model has been trained on the training data.

**11. Exporting Model:**

The trained logistic regression model has been saved for later use in prediction.

**12. Model Evaluation:**

Confusion Matrix:

[[487 13]

[ 87 13]]

*Accuracy Score*: 0.8333

**13. Probability Predictions:**

Probability predictions have been generated using the trained model.

**14. Model Output File**

The model output, including probability predictions and actual outcomes, has been written to "c1\_Model\_Prediction.csv."

**15. Conclusion:**

This document summarizes the development and evaluation of a credit scoring model using logistic regression. The model can be used to predict creditworthiness based on the provided features.

**16. Reference**

<https://medium.com/@skillcate/credit-scoring-project-using-logistic-regression-c1e88bd7cf25>

**17. Future Enhancements**

- Explore more advanced machine learning models for improved predictive accuracy.

- Implement real-time model scoring and monitoring for up-to-date credit assessments.

**18. Contact**

Chaphowasit Mahayossanan

Email: chaphowasit.work@gmail.com