Task 1: Foundational Modeling for Anomalous Transaction Identification

1. Project Background

"FinSecure", an emerging FinTech company, provides digital wallet and online payment services to its customers. To build user trust and mitigate financial risk, the company needs to develop a preliminary machine learning model to flag potentially risky "anomalous transactions" for further review. As a junior data scientist at the company, your first assignment is to use a simplified and sanitized internal transaction dataset to quickly build and evaluate several foundational classification models. This task will establish a solid foundation before you tackle more complex fraud scenarios in the next task.

2. Task Goal

Your core objective is to build a binary classifier to identify anomalous transactions. The target variable in the dataset is Status, where 1 represents an anomalous transaction and 0 represents a normal one.

This dataset has been curated to help you focus on the modeling process itself. However, please pay close attention to the fact that anomalous transactions are still a minority class. Therefore, you must think critically about how to scientifically evaluate your model's performance, as relying solely on overall accuracy can be misleading.

3. Dataset Information

You are provided with three files for this task:

- **finsecure_train.csv**: The training set. This file contains all the features and the target variable (Status). You **must** use this file to train and validate your models.
- **finsecure_test.csv**: The testing set. This file contains all the features for the test data but **does not include the Status column**. You will use your trained model to generate predictions for each row in this file.
- **sample_submission.csv**: A sample file that shows the required format for your submission. Your final prediction file must have the exact same columns and structure.

Other files:

- **evaluate_mac_1** The command line tool to evaluate your result on macOS. Usage: Press "command + space" to open spotlight search and type in "terminal", then type in the following command: ./evaluate_mac_1 ./submission_1.csv. Please note that "./" denotes the current position of the command line and "submission_1.csv" denotes your submission file name.
- **evaluate_windows_1.exe** The command line tool to evaluate your result on Windows. Usage: Press "command + r" and then type "cmd" in the dialog box to launch a terminal, then type in the following

command: .\evaluate_windows_1.exe .\submission_1.csv. Please note that "." denotes the current position of the command line and "submission_1.csv" denotes your submission file name.

4. Attribute Information

- Index: A unique identifier for each transaction row.
- TimeElapsed: The time elapsed (in seconds) between this transaction and the first transaction in the dataset.
- F1 to F28: 28 anonymized numerical features that represent various underlying patterns of the transaction.
- Amount: The monetary value of the transaction.
- Status: **(Target Variable)** The status of the transaction. 1 for anomalous, 0 for normal.