Take-Home Assignment: Binary Classification Model

Objective:

Your task is to build a **binary classification model** using the dataset provided. The dataset consists of anonymized features named `feature_0` to `feature_48`, and the class label is stored in the `target` column.

The goal is to train a model that generalizes well to unseen data. Your model will be evaluated on a **hidden test set** that we will use internally to assess your solution.

Dataset:

- A CSV file will be shared with you containing the following:
    - `feature_0` to `feature_48`: 49 numerical features.
    - `target`: Binary class label (0 or 1).

Deliverables:

Please submit a git project containing:

1. `notebook.ipynb` or `train.py`:
    - Code to load the data, preprocess it, train the model, and generate predictions.
    - You may use any libraries (e.g., scikit-learn, XGBoost, LightGBM, etc.).
2. `README.md`:
    - Brief explanation of your approach and assumptions.
    - Steps to reproduce the training process and generate predictions.
3. `model.pkl` or any serialized model file.

4. `requirements.txt` listing all dependencies.

Time & Expectation:

This task is designed to be completed in **4–6 hours**, but you may take up to **2 days** to submit your solution. You are encouraged to use this time to explore thoughtful, well-structured, and innovative approaches to the problem.

You are free to use any methodology, tools, or techniques you believe are appropriate. This is your opportunity to **showcase your skills, creativity, and problem-solving ability**—feel free to go beyond the basics if you think it adds value.

Submission:

Please email the git repository to us.

Evaluation Criteria:

- Correctness of implementation.
- Code structure and readability.
- Modeling choices and justification.
- Model performance on the hidden test set.
- Reproducibility of results.