Project:

It has 2 steps to be completed, the first step will be designing a machine learning system based on given scenario. And second step will be building API for the same loan predictor model from the last class. Below are the details.

Project 1:

Based on the scenario provided, design a machine learning design system, this has to be just a high level design, that means just share on idea how will you handle data collection, processing, model training, evaluation, and all the other things (you may refer the second session slides). Make sure the answer is not more than **2 pages**.

Scenario:

Imagine you are a machine learning engineer working for a real estate company that wants to build a machine learning system to predict housing prices in a given area. Your goal is to create a model that can accurately predict the price of a house based on various features such as square footage, number of bedrooms and bathrooms, location, and other relevant factors.

Task:

A report detailing your data collection process, data preprocessing steps, feature selection/engineering techniques, model selection criteria, model training process, evaluation metrics, deployment strategy and monitoring strategy.

Project 2:

Given the loan predictor model and also the test data. You need to build your own API to host the model and usable from your endpoint. Below is example of the output that is expected from your endpoint.

```json
{
  "status": "success",
  "error_code": 0,
  "prediction": {
    "label": 1,
    "value": "approve"
  },
  "time_taken": "250ms"
}
```

These are the possible values for each key in the JSON output

Key	Values
status	success/failed
error_code	0/1
prediction.label	1/0
prediction.value	approve/reject
time_taken	Any range

Refer to the documentation in readme to recall on how many unique value each feature of the model has.

Once the API is ready push to your github and submit the repo URL by DM to me. Put your name and url for github. **ANY PLAGIARISM WILL NOT BE TOLERATED, MAKE SURE YOUR WORK IS YOUR EFFORT. :)**

Bonus points:

1. Use pydantic model to define the schema for output and input of endpoint.
2. Use proper code structuring to make the code more maintainable.

Submission Deadline: 25 March 2024 11:59pm