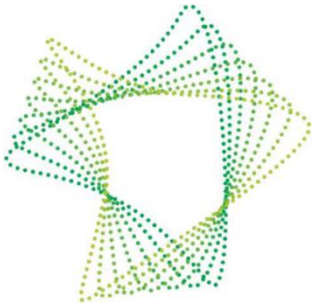

AI Practice

Case Study Interview – Intern

Instructions for completing the case study

This case study aims to assess your holistic knowledge and ability at building an end-to-end Machine Learning (ML) project. Please complete as much of the case as possible within 36 hours *of receiving it and submit your work within this timeframe. Please submit your final code to GitHub and share the repository link with the following github handles:* ***victorong-github, yandm****.*

Instructions for interview

You will meet two or more current members of the AI Practice team for a two to three hours sessions, where you will be invited to run through your solution, perform live coding and answer any question they may have regarding the choices you made.

The assessment will focus on the quality of your deployment and the robustness of your infrastructure design. The final output should be a proof of concept, with emphasis on implementation quality rather than model performance metrics.

Start of Assignment

Background

HDB aims to develop a system that combines traditional machine learning with Large Language Models (LLMs) to analyse and predict resale flat prices across Singapore. The system will provide comprehensive market insights and location-specific comparisons through analysis of transaction data and property characteristics, supplemented with AI-generated natural language explanations.

The housing market is influenced by various factors at both micro and macro levels. Each town has distinct price patterns based on its characteristics, amenities and demographic composition. There is an increasing need for detailed comparative analyses between different towns to better understand local market dynamics and inform decisions about BTO flat development.

The system should be able to make predictions on the price of different archetypes of BTO unit based on their location.

System Requirements

The system should recommend estates for BTO development based on specified requirements and predict prices for different flat types. For example, the system should respond to prompts such as:

"Please recommend housing estates that have had limited Build-To-Order (BTO) launches in the past ten years. For each estate, provide an analysis of potential BTO prices for both 3-room and 4-room flats, considering low, middle, and high floor levels. For each price category, include the recommended household income needed to afford the flat."

The system will generate predictions by calculating resale prices with a standard discount (e.g., 20%) and match these to different income groups.

Your tasks

Your tasks should minimally include the following:

Data Engineering & Architecture

- Extract HDB transaction data from data.gov.sg
- Design and implement database schema
- Document additional data sources if used

Model Training & LLM Integration

- Implement and train price prediction models
- Deploy & create an endpoint
- Specify API service design
- Design and/or implement LLM-powered features

Monitoring & Testing

- Implement and/or recommend monitoring metrics
- Implement automated testing and establish operational processes, (e.g., model versioning and edge case testing)

[Optional] Frontend

- Develop a frontend interface

Where possible, provide additional recommendations to ensure a high-quality solution. If time or resource constraints prevent you from implementing certain features or recommendations, outline your proposed approach for future implementation.

End of Assignment
