Overview

This project implements an Agent-Client Recommendation System that aims to recommend the most suitable financial advisors to clients by ranking agents based on their similarity to clients. Using XGBoost, we identified the key features determining the suitability of each agent for every client and used these features to generate a suitability score for each agent. Thereafter, agents were ranked based on cosine similarity, to recommend those who were most similar to the clients' needs.

Environment Setup

To ensure smooth execution, run the following command and install the required dependencies to set up the environment, **please use python 3.10**:

pip install -r requirements.txt

This will install:

- **pandas** for data handling.
- **numpy** for numerical computations.
- **scikit-learn** for machine learning.
- **xgboost** for training the predictive model.
- **tensorflow** for metric learning.
- **itertools** for agent-client pairing and client similarity computation.
- **pyarrow and fastparquet** to read the datasets.

Running the Notebook and Reproducing Results

Download the notebook (CAT_A_69.ipynb), navigate to the directory containing it and open it on Jupyter Notebook.

Alternatively, if using Google Colab, upload the .ipynb file to Google Colab.

To run the notebook cell by cell, click "Run" on each cell.

Running the Model

Ensure the datasets (nus_agent_info_df.parquet, nus_policy_info_df.parquet and nus_client_info_df.parquet) are downloaded and present in the same directory as the notebook.

If running on Google Colab, run the following code to mount Google Drive:

```
from google.colab import drive
drive.mount('/content/drive')
```

Run the codes in the notebook to load the respective datasets.

If # SECTION 1: AGENT–SIMILAR CLIENTS (METRIC LEARNING),
Siamese Neural Network takes too long to load, please restart the kernel and run all cells.

Key Insights

Our model adopts a hybrid approach, combining metric learning with product expertise and demographic similarity to create a robust recommendation system that leverages both behavioral and contextual data. According to our model and analysis, agent–product rating and background similarity of agents with clients are among the most important determinants of the similarity of agents and clients. This insight could help refine future agent training and marketing strategies.