SCHOOL OF COMPUTING
Faculty of Engineering

Project Proposal Form MCST1043
Sem: 2 Session: 2024/25

## SECTION A: Project Information.

| | |
|---|---|
| Program Name: | **Masters of Science (Data Science)** |
| Subject Name: | **Project 1 (MCST1043)** |
| Student Name: | Yang Mu |
| Metric Number: | MCS241045 |
| Student Email & Phone: | yangmu@graduate.utm.my number:1162302346 |
| Project Title: | Big data driven: Forecast of global real estate market ups and downs in some regions |
| Supervisor 1: | |
| Supervisor 2 / Industry Advisor(if any): | |

## SECTION B: Project Proposal

**Introduction**:

The global real estate market is a complex system that is affected by many factors, including economic growth, interest rate changes, demographic changes, and government policies. (Mariia, 2021) In recent years, with the rapid development of big data and artificial intelligence technologies, it has become possible to use big data analysis technology to predict the trend of the global real estate market, and it has gradually become a research hotspot. However, most existing studies focus on forecasts for a single region, making it difficult to gain a comprehensive view of the overall trends and influencing factors of the global real estate market.

**Problem Background**:

Global real estate market data is rich, but there are data source dispersion, rigid conditions inconsistent problems such as uneven, analysis Angle, lead to predict and analyze the global real estate market faces many challenges. For instance, there is a lack of a unified global real estate market forecasting and analysis method. (Lee Pei Wen, 2017) For example, the lack of a unified standard for describing and managing property data leads to difficulties in integrating and analyzing data from different

sources. In addition, most of the existing prediction models based on a single region or a single index, is difficult to accurately reflect the overall trend of the global real Estate market.

(Ruoying Tan & Tze-Haw Chan 2021)

**Problem Statement**:

How to build a unified framework for the global real estate market based on big data?. The framework needs to formulate unified standards for describing and managing global real estate market data, and develop effective forecasting models on this basis, so as to improve the accuracy of forecasting changes in the global real estate market.

**Aim of the Project**:

The project aims to use advanced big data analytics techniques to exploit local areas and build global real estate market forecasting model data. The aim is to build global real estate market forecasting model data, enhance the accuracy and reliability of the prediction results, serve as the theoretical support for this project, and conduct in-depth research on the significance of its experimental results. (Petros et al , 2024).

**Objectives of the Project**:

**Data Collection and Organization:** To use advanced big data analytics techniques to exploit local areas and build global real estate market forecasting model data. To support for this project, and conduct in-depth research on the significance of its experimental results.

**Data Framework Construction:** To formulate a unified and standardized data framework and describe and manage the collected data. This framework ensures the usability and analyzability of data, while providing a solid foundation for its analysis and model establishment (Petros et al, 2024).

**Data framework construction**: To build a unified data framework to describe and manage collected housing data, ensuring its comparability and analysis, while providing a standardization basis for data integration and analysis (Petros et al, 2024).

**Forecasting model development**: To develop effective forecasting models, such as time sequence models and machine learning models, to predict the top, downs and trend of housing market in some areas, and to evaluate and verify the model, analyzing the accuracy and truth of the model.

**Analysis and visualization of results**: To visualize forecast results, analyze and

interpret forecast results, and to put forward recommendations, such as forecasting and risk analysis of future real estate market trends, as well as investment strategies.

**Scopes of the Project**:

Data source: Mainly uses public global real estate transaction data, such as data from government agencies, real estate websites, research institutions, etc.

Prediction indicators: Mainly focus on the rise and fall trend of housing prices, and may include other indicators, such as transaction volume, rental level, etc.

Prediction method: Mainly uses time series analysis, machine learning models and other methods for prediction.

Result analysis: Compare and analyze the prediction results with the actual situation, explain the reasons behind the prediction results, and put forward some suggestions.

**Expected Contribution of the Project**:

Promote standardization of global real estate market data management: By establishing a unified metadata framework, this project will provide standardized specifications for global real estate market data management, promote the integration analysis of data from different sources, improve data quality and credibility, and provide the basis for more accurate market forecasting. Improve the accuracy and reliability of global real estate market forecasting: This project will develop a forecasting model based on big data analytics and combine multiple data sources and analysis methods to improve the accuracy and reliability of forecasting, provide investors and decision makers with more effective reference basis to help them better understand market trends and potential risks. (Ruoying Tan & Tze-Haw Chan, 2021)

Provide new perspectives and tools for business analysis: The research results of this project can provide new perspectives tools for business analysis. For example, enterprises can use the metadata framework and prediction model built by this project to analyze real estate market trends, formulate more effective investment strategies, optimize business processes, and improve profitability. Urban planning and development: It can provide data support for urban planning and development, help urban planners better understand real estate market trends, and formulate more reasonable urban planning and development strategies. (Bo et al, 2022)

**Project Requirements**:

Software: Python, Pandas

Hardware: Computer, network, data

Technology/Technique/ Methodology/Algorithm: Data collection and preprocessing

- Collect and organize metadata from COVID-19 scientific datasets.
- Construct a metadata framework for COVID-19 scientific datasets,encompassing external features, content features, and sharing features.
- Utilize Protégé software to build an ontology for scientific datasets, defining core concepts and attribute relationships.
- Store the constructed knowledge graph using the Neo4j graph database, enabling query retrieval and reasoning for entities and their relationships.
- Data analysis
- Model Evaluation

**Type of Project (Focusing on Data Science)**:

[ ] Data Preparation and Modeling

[ √ ] Data Analysis and Visualization

[ √ ] Business Intelligence and Analytics

[ ] Machine Learning and Prediction

[ ] Data Science Application in Business Domain

**Status of Project**:

[ √ ] New

[ ] Continued

If continued, what is the previous title?

## SECTION C: Declaration

**I declare that this project is proposed by**:

[ √ ] Myself

[ ] Supervisor/Industry Advisor ( )

Student Name: Yang Mu

Yang Mu — **Signature**

07/04/2025 — **Date**

## SECTION D: Supervisor Acknowledgement

The Supervisor(s) shall complete this section.

**I/We agree to become the supervisor(s) for this student under aforesaid proposed title.**

Name of Supervisor 1:

**Signature** **Date**

Name of Supervisor 2 (if any):

**Signature** **Date**

## SECTION E: Evaluation Panel Approval

The Evaluator(s) shall complete this section.

**Result:**

[ ] FULL APPROVAL
[ ] CONDITIONAL APPROVAL (Minor)
[ ] CONDITIONAL APPROVAL (Major)*
[ ] FAIL*

***** Student has to submit new proposal form considering the evaluators' comments.

**Comments:**

Name of Evaluator 1: .....

**Signature** **Date**

Name of Evaluator 2: .....

**Signature** **Date**