A logo for college computing

Description automatically generated

**Assessment Cover Page**

|  |  |
| --- | --- |
| *Student Full Name* | Miqueias S. dos Santos |
| *Student Number 2* | 2016287 |
| *Module Title* | Strategic Thinking (HDip in Data Analytics) |
| *Assessment Title* | CA 1 – Capstone Project Proposal |
| *Assessment Due Date* | 29th March 2024 |
| *Date of Submission* |  |

**Declaration**

By submitting this assessment, I confirm that I have read the CCT policy on academic misconduct and understand the implications of submitting work that is not my own or does not appropriately reference material take from a third party or other source.

I declare it to be my own work and that all material from third parties has been appropriately referenced.

I further confirm that this work has not previously been submitted for assessment by myself or someone else in CCT College Dublin or any other higher education institution.

Table of Contents

**Type chapter title (level 1)1**

Type chapter title (level 2)2

Type chapter title (level 3)3

**Type chapter title (level 1)4**

Type chapter title (level 2)5

Type chapter title (level 3)6

**Introduction**

The House Price Prediction project aims to develop a powerful machine learning model for predicting house prices, with the potential to significantly impact the real estate market. Utilizing both current and historical data, the project seeks to provide valuable future insights into the industry's dynamics and evolution. By addressing pressing issues like rising prices, homelessness, and promoting fair buyer opportunities, the project aims to contribute to strategic initiatives.

The overarching goal of this project is to play a pivotal role in enhancing transparency within the real estate market. Through the accurate prediction of house prices, stakeholders can be empowered with timely information, facilitating well-informed decision-making. The project also aspires to assist governmental bodies by offering insights that can aid in policy formulation and strategic planning.

In essence, the House Price Prediction project strives to be a comprehensive resource that goes beyond mere price forecasting. Its objectives encompass addressing societal challenges, supporting fair practices, and providing insights that are valuable for both industry stakeholders and government entities.

1. **Objectives**
   1. **Development of the Model**

Create a machine learning model that can predict house prices based on key features like location, numbers of room, and construction year.

* 1. **Enhance Market Transparency**

Utilize the model to offer clear insights into house prices, empowering buyers and sellers to make well-informed decisions.

* 1. **Price Negotiations**

Improve efficiency and fairness in negotiations by using the model's predictions to establish a clearer understanding of property values.

* 1. **Inform Governmental Decisions**

Explore using predictive insights to assist government decisions, particularly in identifying areas for social housing construction to address housing challenges.

* 1. **Contribute to Strategic Policies**

Explore the model's potential role in shaping strategic initiatives and governmental policy formulation, addressing issues like rising house prices, homelessness, and ensuring fair opportunities for buyers. These objectives align with the broader goal of developing a house price prediction model, showcasing its practical applications and impacts in the real estate industry and beyond.

1. **Problem Definition**

The House Price Prediction project addresses a critical challenge in the Irish real estate market—the significant growth of the house price-to-income ratio. According to Statista (published on Jan 8, 2024), between 2015 and 2023, the ratio in Ireland increased by nearly 17 index points. This substantial increase is a barrier to homeownership, limiting affordability and disrupting community stability [1]. The project aims to develop a machine learning model for house price predictions, striving to address challenges linked to escalating housing prices and fostering a more accessible and equitable real estate landscape.

1. **Scope**

The two-semester House Price Prediction capstone project encompasses a comprehensive exploration of machine learning methodologies to accurately forecast house prices, focusing on the Irish real estate market. The project includes:

* 1. **Data Collection and Preprocessing**

Gather relevant real estate data, including property details, location information, and historical pricing. Clean and preprocess the data to ensure its quality and suitability for machine learning model training.

* 1. **Data Sources**

The data intended for use in the project is publicly available and can be accessed at The Property Services Regulatory Authority - Property Price Register.

* 1. **Model Development**

Create and fine-tune a robust machine learning model, incorporating features such as location, number of rooms, and year built. Implement advanced techniques to enhance the model's predictive accuracy.

* 1. **Transparency and Efficiency Enhancement**

Utilize the developed model to provide transparent insights into house prices, empowering buyers, and sellers with valuable information. Integrate the model into the real estate market to optimize efficiency in price negotiation.

* 1. **Government Decision Support**

Exploration of the potential application of predictive insights to aid public government decisions, particularly in identifying areas for strategic social housing construction.

* 1. **Strategic Initiatives and Policies Contribution**

Investigation of the model's predictions' applicability in contributing to strategic initiatives and policy formulation, addressing challenges like rising house prices and homelessness.

* 1. **Exclusions**

The project will not delve into predictive analysis beyond the scope of house prices in the real estate market. Implementation of policies; the focus is on providing insights for policy considerations.

* 1. **Boundaries**

The project is limited to the capabilities of machine learning algorithms and the quality of available data. Ethical considerations and data privacy will be strictly adhered to.

* 1. **Planned Methods, Techniques, and Approaches**

Utilization of supervised learning algorithms for model development. Apply feature engineering to enhance model performance. Evaluate various regression techniques for predicting house prices. Ensure ethical handling of data, privacy, and compliance.

1. **Timeline**
   1. **Semester One:**

Weeks 1-4: Data Collection and Preprocessing.

Weeks 5-8: Model Development and Fine-tuning.

Weeks 9-12: Initial Transparency and Efficiency Enhancement.

* 1. **Semester Two:**

Weeks 1-4: Government Decision Support Exploration.

Weeks 5-8: Contribution to Strategic Initiatives and Policies.

Weeks 9-12: Final Model Optimization, Documentation, and Presentation Preparation.

* 1. **Expected Deliverables by End of Semester Two**

A well-trained machine learning model capable of predicting house prices accurately. Transparent insights into real estate market dynamics. Integration of the model into a user-friendly interface. Comprehensive documentation of methodologies, results, and recommendations.

1. **Ethical Considerations**

The House Price Prediction capstone project is committed to upholding the highest ethical standards throughout its lifecycle.

* 1. **Data Privacy and Confidentiality**

Ensure that all collected data adheres to privacy regulations and is anonymized to prevent the identification of individuals.

* 1. **Informed Consent**

If needed, obtain consent from data sources, acknowledging the intended use of their information, and ensure transparency about data handling practices.

* 1. **Bias and Fairness**

Regularly assess and mitigate bias within the machine learning model to prevent unfair treatment of specific demographic groups. Implementing fairness-aware algorithms and techniques to ensure equitable outcomes.

* 1. **Transparency**

Providing clear and understandable explanations of the model's predictions to users, avoiding black-box scenarios. Disclosing the limitations and potential biases of the model to users and stakeholders.

* 1. **Data Permissions**

Acquiring explicit permissions for data usage from relevant authorities or organizations that own or manage the datasets. Adhering to any restrictions or guidelines stipulated by data sources to ensure ethical data utilization.

* 1. **Societal Impact**

Evaluating potential societal impacts of the project, including its influence on property pricing, market dynamics, and housing accessibility. Ensuring that the project's outcomes contribute positively to the welfare of individuals and communities.

During the project development, I am committed to continuous monitoring of ethical considerations, adjusting as needed, and researching when necessary to uphold the highest standards of integrity and responsibility in all project activities.

# References

1. Statista. (2024). Annual house price change in Ireland. Retrieved from [<https://www.statista.com/statistics/1155332/annual-house-price-change-in-ireland/>](https://www.statista.com/statistics/1155332/annual-house-price-change-in-ireland/) (Accessed: 06 March 2024).
2. The Property Services Regulatory Authority - Property Price Register. Available at: <https://www.propertypriceregister.ie/> (Accessed: 07 March 2024).
3. OpenAI. (2023). ChatGPT (Mar 14 version) [Large language model]. Available at: [https://chat.openai.com/chat](https://chat.openai.com/chat%20) (Accessed: 06 March 2024).
4. Central Statistics Office Ireland. (n.d.). CSO StatBank. Available at: <https://data.cso.ie/> (Accessed: 05 March 2024).
5. daft.ie. (2024). Ronan Lyons Report - 2023 Q4. Available at: <https://ww1.daft.ie/report/ronan-lyons-2023q4-dafthouseprice/infographic?d_rd=1> (Accessed: 07 March 2024)