A logo for college computing

Description automatically generated

**Assessment Cover Page**

|  |  |
| --- | --- |
| *Student Full Name* | Derly Milded Montealegre Gonzalez |
| *Student Number* | 2024151 |
| *Module Title* | Machine Learning (10 ETCS) |
| *Assessment Title* | CA 1 – Project |
| *Assessment Due Date* | 6 March 2024 |
| *Date of Submission* | 21 April 2024 |

**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 taken 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.

Contents

[Introduction 2](#_Toc164334786)

[References 3](#_Toc164334787)

PROJECTION OF FUTURE HOUSING RELOCATIONS IN BEIJING

# Introduction

In this project, one of the most evident urban phenomena in the last decades in Beijing is being detected, this is the process of urban renewal of the city center, which CLOU has been managing to renovate inside the Capital Square Beijing and what they want to achieve with this is to reposition and improve the public space. Clou is managing through various processes to illuminate the office lobbies with a modern design language, introducing the concept "*The City Lantern*".

Objective In fact, what this project aims to achieve is to generate a notable impact on the real decline and abandonment that is being observed in houses and apartments. Additionally, in this Data it is observed that some houses were built since 1960 and through this it is intended to analyze the number of times these houses have been remodeled to date, according to the structure of the building.

Definition of the problem, the objective of this project is to be able to determine the areas of the perimeter that are experiencing the most abandonment according to the year of construction and the number of reforms, or any type of repairs that have been carried out in the buildings in order to find the necessary results on which area of the buildings is showing the most deterioration. In this way, steps can be taken to analyze what could be the possible causes that could be prevented to make the infrastructure look more modern and more useful over the years.

Regarding the methodology that has been worked in class and according to the information observed, we would say that it is a segmentation methodology, which seeks to group the possible years of greater activity and in this way obtain more relevant information on whether there are similarities or if they have one or more aspects in common with previous years. In addition, the main idea is to improve the functionality of the homes by updating their interiors, such as the facade of the buildings, as this could improve the appearance of the homes and make better use of the property.

To obtain more accurate data on the amount of money spent on housing in recent years, it would be necessary to determine which perimeter areas are abandoned based on the year of construction and the number of renovations. To do this, sorting algorithms are used to classify perimeter areas into categories of abandonment. For example, there are areas with high abandonment, areas with moderate abandonment, and areas with no abandonment. This will allow us to know the estimated cost of the changes that have been made to the property. Then, with this information, a percentage of money could be considered for contingencies that may arise over time. However, for this budget to be as tight as possible, it should be taken into account that it will vary depending on the location of the place, the size of the house, the number of people living in the place and the amount of remodeling that has been done or has been done to date.

We can see the association between the categorical variables after cleaning up our database. To illustrate how the variables affect the several building types tower, bungalow, plate/tower, and plate in relation to our objective variable price, we created a box diagram. It is evident from this graph that the bungalow-style buildings are more expensive than the norm.A graph of a graph showing different colored squares

Description automatically generated with medium confidence

Here, we can observe that the display frequency steadily stabilises at a greater price and finishes at 150,000 RMB.A graph of a graph of a number of blue bars

Description automatically generated with medium confidence

1. **Predicción de precios de viviendas:** Si el objetivo del proyecto es predecir los precios de las viviendas en función de ciertas características, puedes utilizar algoritmos de regresión, como Regresión Lineal, Bosques Aleatorios o Máquinas de Vectores de Soporte (SVM), para desarrollar un modelo predictivo. Entrenas el modelo utilizando características como tamaño de la vivienda, ubicación, número de habitaciones, etc., y el precio de venta como variable objetivo.
2. **Predicción de precios de viviendas:**
   * Objetivo: Predecir el precio de las viviendas en función de características como tamaño, ubicación, número de habitaciones, etc.
   * Algoritmo: Regresión Lineal, Bosques Aleatorios, Máquinas de Vectores de Soporte (SVM).
   * Justificación: Utilizando un conjunto de datos históricos que incluye características de las viviendas y sus precios de venta, puedes entrenar un modelo de regresión para predecir los precios de las viviendas en función de estas características. La precisión del modelo se puede evaluar utilizando métricas como el RMSE (Root Mean Squared Error) o R^2 (Coeficiente de determinación).

English the same, Please help me correct my introduction and objective as this will be the approach I will follow:

Home Price Prediction: If your project goal is to predict home prices based on certain characteristics, you can use regression algorithms, such as Linear Regression, Random Forests, or Support Vector Machines (SVM), to develop a model. predictive. You train the model using characteristics such as house size, location, number of rooms, etc., and the sales price as the target variable.

Home Price Prediction:

Objective: Predict the price of homes based on characteristics such as size, location, number of rooms, etc.

Algorithm: Linear Regression, Random Forests, Support Vector Machines (SVM).

Rationale: Using a historical data set that includes home characteristics and their sales prices, you can train a regression model to predict home prices based on these characteristics. Model accuracy can be evaluated using metrics such as RMSE (Root Mean Squared Error) or R^2 (Coefficient of Determination).

Mi usuario para mi plataforma en la universidad

<https://moodle.cct.ie/login/index.php>

usuario: 2024151

clave: Derly27\*

vas a donde dice my course

o hay en la parte de abajo ya te sale este trabajo es de Mach Learn y el otro es de StatsTechdata. Hay en la plataforma vas a econntrar varios ejemplos de como hacer estos ejercicios.

Estoy intendando sacar este cuadro pero no me sale, es como si mi Data fuera muy pesada entonces estaba pensando en utilizar otro método.

A screenshot of a color chart

Description automatically generated

PREGUNTAS PARA COLOCAR MAS PALABRAS

What are the most important features for predicting X as a target variable? o Which classification approach do you prefer for the prediction of X as a target variable, and why? o How to classify the loyal and churn customers using Support Vector Machines? o Why is dimensionality reduction important in machine learning?

De aqui saque mi data

Housing price of Beijing from 2011 to 2017, fetching from <https://www.kaggle.com/datasets/ruiqurm/lianjia>

Este es uno de los link de donde saque mi Data y realizo la limpieza con la misma base de datos

<https://www.kaggle.com/code/aadhavvignesh/regression-with-scikit-learn-practical-ml-1>

este es otro ejemplo con la data que seleccione

<https://www.kaggle.com/code/eraw0x/house-prices-in-beijing-eda-arima#data-cleaning-and-preparation-for-analysis>

Links para trabajar en mi Text

<https://www.humanizeai.io/>

<https://app.gptzero.me/app/ai-scan?tab=0>

<https://quillbot.com/>

<https://www.deepl.com/translator>

<https://moodle.cct.ie/course/view.php?id=41>

<https://chat.openai.com/c/7c9da5d8-da7a-4cb7-88e9-7ead9d0654b9>

# References

Housing price of Beijing from 2011 to 2017, fetching from <https://www.kaggle.com/datasets/ruiqurm/lianjia>

Beijing second-hand house Beijing rent Beijing real estate network Beijing Lianjia network. (2024). Recovered from: <https://bj.lianjia.com/chengjiao>.

Clostermann, Zhong, Zhao, Li, Cheng, Ding. (2023). Capital Square Beijing Renovation. ARQA. Recovered from: <https://arqa.com/en/architecture/capital-square-beijing-renovation.html> [March 25,2024].

Zhicheng. (2020). Chinese growth of 6.1%, the lowest in 30 years. PIME Asianews. Recovered from: <https://www.asianews.it/noticias-es/El-crecimiento-chino-del-6,1,-el-m%C3%A1s-bajo-en-30-a%C3%B1os-49052.html>.

EURE (Santiago) vol.37 no.111 Santiago (mayo 2011). Recovered from: <https://www.scielo.cl/scielo.php?pid=S0250-71612011000200010&script=sci_arttext&tlng=pt>.

This article is based on the (December 2010) version of the PRMS 100 Explanatory Report. Recovered from: <http://www.seremi13minvu.cl/opensite_20110103164100.aspx>.

American Psychological Association. (2020). Publication manual of the American Psychological Association (7th ed.). Recovered from: <https://doi.org/10.1037/0000165-000>

https://github.com/derlyai/Machine-Learning--10-ETCS--.git