**CCT College Dublin**

**Assessment Cover Page**

|  |  |
| --- | --- |
| **Module Title:** | Strategic Thinking |
| **Assessment Title:** | CA 1 – Capstone Project Proposal |
| **Lecturer Name:** | James Garza |
| **Student Full Name:** | Federico Ariton, Alaa Shammary |
| **Student Number:** | sba22090 , sba23361 |
| **Assessment Due Date:** | 29th October 2023 |
| **Date of Submission:** | 29th October 2023 |

**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. |

Contenido

[Introduction: 3](#_Toc149436884)

[Objectives: 3](#_Toc149436885)

[Development of a Predictive Pricing Model: 3](#_Toc149436886)

[Data collection and Analysis: 3](#_Toc149436887)

[Market insights: 3](#_Toc149436888)

[Compare the predicted price against the available price. 4](#_Toc149436889)

[To be able to make a better-informed decision by comparing the available price of a car vs the predicted price of a similar model to decide whether the buyer is getting a bargain or the car is overpriced. 4](#_Toc149436890)

[User-friendly tool: 4](#_Toc149436891)

[Problem Definition: 4](#_Toc149436892)

[Scope: 4](#_Toc149436893)

[Business Understanding: 4](#_Toc149436894)

[Data Understanding: 4](#_Toc149436895)

[Model Development: 5](#_Toc149436896)

[Evaluation: 5](#_Toc149436897)

[Deployment: 5](#_Toc149436898)

[Documentation: 5](#_Toc149436899)

[Inclusion: 5](#_Toc149436900)

[Used vehicles: 5](#_Toc149436901)

[Irish Market: 5](#_Toc149436902)

[Exclusion: 5](#_Toc149436903)

[Real-time Market Data: 5](#_Toc149436904)

[New vehicles: 5](#_Toc149436905)

[Vehicles outside the Irish Market 5](#_Toc149436906)

[Team Roles: 5](#_Toc149436907)

[Timeline: 6](#_Toc149436908)

[Ethical Considerations: 6](#_Toc149436909)

[References: 6](#_Toc149436910)

Title: "Analysing and Predicting Prices of Used Vehicles"

# Introduction:

The proposed capstone project aims to address the critical issue of understanding and predicting car pricing in the used vehicles market. The used car market is a big part of the car industry, and it affects both individuals as well as businesses. “There is a need for a used car price prediction system to effectively determine the worthiness of the car using a variety of features.” (Gokce, 2020).

This project is important because it deals with real-world problems like car price inflation, how some people have more information than others, and the need for accurate pricing. This project will help potential car buyers, sellers and dealers make informed decisions, ultimately contributing to a more transparent and efficient marketplace.

# Objectives:

Development of a Predictive Pricing Model:

Create an advanced pricing model that considers various factors influencing used car prices, Manufacture, Model, Production year, Category, Fuel type, Mileage, and Engine size.

Data collection and Analysis:

Collect and analyze extensive historical data to train and validate the pricing model, ensuring its accuracy and reliability. (Simplilearn, 2021)

Market insights:

Studying and understanding how the used car market works. To understand why people, prefer certain cars, why prices change in different places, and how factors like the economy or the time of year affect car prices. By doing this, we can make sure our pricing model is smart and can adapt to all these changes in the market.

Compare the predicted price against the available price.

To be able to make a better-informed decision by comparing the available price of a car vs the predicted price of a similar model to decide whether the buyer is getting a bargain or the car is overpriced.

User-friendly tool:

Develop a user-friendly interface or application that allows consumers and dealers to estimate car prices easily based on the model's predictions.

# Problem Definition:

The used car market suffers from pricing uncertainty, often leaving buyers and sellers unsure of the fair market value of vehicles especially when prices of vehicles have inflated about %77.5 since the start of the pandemic according to Dondeal. This has led to overpricing, affecting consumers' purchasing power and businesses' profitability. Therefore, this project aims to develop an accurate pricing model that addresses the issues of inflation and overpricing of vehicles and provides users with reliable pricing estimates, that can leave the buyer and the seller with a better-informed decision.

# Scope:

## Business Understanding:

In this phase, the project objectives are defined, and data mining goals are established. This involves understanding the problem, its relevance, and how it can add value to the business.

## Data Understanding:

This phase focuses on data collection, exploration, and initial data analysis. It is crucial for understanding the available data, its quality, and its suitability for the project. The team will collect and clean a large dataset of historical car sales data. This dataset should include information about used car prices, Manufacture, Model, Production year, Category, Fuel type, Mileage, and Cylinders.

Step that we are going to implements:

“Collect initial data, Describe data, Explore data, Verify data quality” (Hotz, 2022).

Data Preparation:

Data preprocessing and cleaning are performed in this phase. Data is transformed and made ready for analysis and modeling. That helps ensure that our machine learning model has clean, relevant and representative data for can lead with the most accurate price prediction.

Step that we are going to implements:

“Select data, Clean data, Construct data, Integrate data and Format data” (Hotz, 2022).

Model Development:

Machine learning models will be developed to predict car prices. Various regression algorithms such as linear regression, decision trees, random forests, and K-Nearest Neighbors (KNN) will be explored and tested.

## Evaluation:

Models are evaluated against predefined criteria to ensure they meet project objectives. Model performance and effectiveness are assessed.

Step that we are going to implements:

“Evaluate results, review process and determine next steps” (Hotz, 2022).

## Deployment:

The final models and the user interface will be deployed in this phase to make the car price prediction system accessible to end-users.

Documentation:

A comprehensive documentation of the project, including data sources, methodologies, model selection, and code, will be prepared.

Inclusion:

Used vehicles: The project is based on used vehicles information and figures.

Irish Market: The project will be on vehicles in the Irish market.

## Exclusion:

Real-time Market Data: The project will not provide real-time pricing information. It will focus on historical data analysis and predictions based on historical trends.

New vehicles: The project will exclude new vehicles, with the focus based on used vehicles available in the market.

Vehicles outside the Irish Market: The project will not contain information or figures of the worldwide market.

Team Roles:

Temporary roles as tasks will be distributed once work in depth starts.

Business Understanding: (Alaa Shammary, Federico Ariton)

Data Understanding: (Alaa Shammary, Federico Ariton)

Data Preparation: (Alaa Shammary, Federico Ariton)

Model Development: (Alaa Shammary, Federico Ariton)

Evaluation: (Alaa Shammary, Federico Ariton)

Deployment: (Alaa Shammary, Federico Ariton)

Documentation: (Alaa Shammary, Federico Ariton)

## Timeline:

Business Understanding: (1st October - 29th October 2023)

Data Understanding: (1st November – 30th November)

Data Preparation: (1st December – 30th December)

Model Development: (10th January – 25th February)

Evaluation: (1st March – 30th March)

Deployment: (1st April – 15th April)

Documentation: (15th April – 30th April)

This is an initial timeline and can be changed based on the project need.

# Ethical Considerations:

Ethical considerations include safeguarding user privacy and ensuring that data collected is used solely for the purpose of the project. Data permissions will be sought from all relevant sources, and any personally identifiable information will be anonymized. The project will adhere to all data protection and privacy regulations. There are no medical aspects involved in this capstone project.

# References:

Gokce, E. (2020). Predicting Used Car Prices with Machine Learning Techniques. [online] Medium. Available at: <https://towardsdatascience.com/predicting-used-car-prices-with-machine-learning-techniques-8a9d8313952>.

Simplilearn (2021). *What is Collection of Data? Methods, Types & Everything You Should Know*. [online] Simplilearn.com. Available at: <https://www.simplilearn.com/what-is-data-collection-article#:~:text=Data%20collection%20is%20the%20process%20of%20collecting%20and%20analyzing%20information>

markdonlon2 (2023). *Prices continue to grow, with total car price inflation now at 77.5% since the start of the pandemic*. [online] DoneDeal Blog. Available at: <https://www.blog.donedeal.ie/post/prices-continue-to-grow-with-total-car-price-inflation-now-at-77-5-since-the-start-of-the-pandemic>