# CS523 - Applied AI Project Proposal

Andrei Cozma, Hunter Price

October 18, 2021

## 1 Problem Description

In this project we will be following Pathway #2: "AI as an Application". This project aims to build a system that will help both buyers and sellers make more informed decisions when they navigate the fast-paced used car market. This system will serve 3 separate groups with both classification and regression machine learning models:

- FOR BUYERS: Classification model that recommends cars based on the person's preferences of features, such as the year, the transmission, the engine size, miles per gallon, as well as the person's budget.
- FOR SELLERS: Regression model that provides a price estimation for a car as a function of its characteristics such as model, year, price, transmission, mileage, fuel type, and engine size.
- ENTERPRISES: Regression model that used car dealerships and sellers can use to evaluate a car's worth based on its features to understand which features of a car weigh more heavily on its final price. For example, mileage tends to influence the worth of a car to a way higher degree than the engine size, mileage, or type of transmission. The insight of this variance can help sellers price their cars more appropriately.

## 2 Problem Motivation

The price of a car changes significantly over time. This price fluctuation is hard to predict and can lead to significant financial loss to its owners. Further, people tend to be biased when they are selling or shopping for a used car, and they let their emotions get in the way of making a sound decision. The used car market is a dark market in terms of information. It often requires a lot of timely searching and making comparisons between many different sources for the buyer to make an informed decision. Used car dealerships are often in the business of making money, and some individual sellers may not always be honest about the car's actual worth. Our project could help people buy and sell used cars more effectively by providing them with more accurate estimations that are derived from a much larger set of samples. A very large set of samples can effectively produce more accurate averages and estimates for the car's worth because the sample size can compensate for the randomness and variability in the data. This eliminates the need for the end-user to perform many timely stages of search and comparisons on the prices of cars equipped with similar desired features.

# 3 Planned Activities

The project will consist of six main activities:

- Collecting the dataset: We will use data publicized on kaggle.com
- Data pre-processing: The data format will need to be cleaned up and pre-processed before it can be used, especially if we want to combine multiple sets of data.
- Data exploration: We will quickly analyze the data with statistical software to understand how it can be structured to suit our needs. We will also learn about how to choose the right parameters for the model.
- Building the model: We will implement a model that can predict the price of a used car given its characteristics. Similarly, we can implement a model that can suggest a vehicle based on the buyer's budget and preferred features.
- Evaluation: We will then use the model to predict the prices of the testing set and compare the predictions to the actual prices to evaluate the model's efficiency.
- Final report: Write a detailed final report.

# References


[1] Ashish Chandak, Prajwal Ganorkar, Shyam Sharma, Ayushi Bagmar, and Soumya Tiwari. Car price prediction using machine learning. *International Journal of Computer Sciences and Engineering*, 7:444–450, 5 2019.

[2] Chuancan Chen, Lulu Hao, and Cong Xu. Comparative analysis of used car price evaluation models. *AIP Conference Proceedings*, 1839(1):020165, 2017.

[3] Enis Gegic, Becir Isakovic, Dino Keco, Zerina Masetic, and Jasmin Kevric. Car price prediction using machine learning techniques. *TEM Journal*, 8(1):113–118, 2 2019.

[4] Nabarun Pal, Priya Arora, Puneet Kohli, Dhanasekar Sundararaman, and Sai Sumanth Palakurthy. How Much Is My Car Worth? A Methodology for Predicting Used Cars' Prices Using Random Forest. 886:413–422, 2019. Series Title: Advances in Intelligent Systems and Computing.

[5] K. Samruddhi and R. Ashok Kumar. Used car price prediction using k-nearest neighbor based model. *International Journal of Innovative Research in Applied Sciences and Engineering (IJIRASE)*, 4:629–632, 2020.