ECE 9063 Data Analytics Foundations

Assignment 2: Neural Networks

**Student Name:** Jianping Ye

**Student Number:** 250887769

**Instructor:** Katarina Grolinger

**Problem Statement**

The used car market is a perfect place for finding cars in decent conditions and with fair prices. It is also the reason that the market has been growing in recent years. However, it is difficult to choose the opportune moment to buy or sell as the price fluctuates constantly. And there are many factors contributing to the price fluctuations. For instance, cars have diverse conditions and the market trend is not stationary all the time. It will be beneficial for both buyers and sellers if we could make a model to predict the value of cars such that they can make a more confident decision. With the help of a suitable model, buyers will be able to make sure the car is worthy of its price, and sellers can get a more accurate price estimation in accordance with other cars having similar conditions. In this report, the forecasting problem is defined as follow: predict the price of a used car in the current year given a set of relevant information.

**Dataset Description**

Link to the data: <https://www.kaggle.com/adityadesai13/used-car-dataset-ford-and-mercedes>

These datasets list scraped data of used cars in the British market and are separated into files specific for each car manufacturer. In this report, the dataset selected is “Audi.csv”. It contains 9 attributes and 10668 samples. The dataset is suitable for this assignment as it has adequate attributes and samples. With over 10,000 samples, it is easier to strike a balance between computational time and reliability of the model . The attributes are listed below:

* Model: The model code of the car
* Year: registration year of the car
* Price: price on the market
* Transmission: type of gearbox, either manual, automatic, or semi-auto
* Mileage: distance used so far
* fuelType: type of fuel the engine uses, either diesel, petrol, hybrid, or other
* tax: road tax
* mpg: miles per gallon
* engineSize: size of engine in litres

Noticeably, model, transmission, and fuelType have nominal data that needs to be transformed into numerical values. All the attributes in the dataset are considered in the model as they are all important factors while estimating the price of cars in the real-world.

**Neural Network Architecture Overview**