

Consider the `CarPrice_Assignment.csv` data file. This data is public available on the Kaggle website, and has information on cars (characteristics related to car dimensions, engine and more). The goal is to use car information to predict the price of the car. **In Python**, answer the following:

1. (5 points) Load the data file to you S3 bucket. Using the pandas library, read the csv data file and create a data-frame called `car_price`.
2. (5 points) Using the `wheelbase`, `enginesize`, `compressionratio`, `horsepower`, `peakrpm`, `citympg`, as the predictor variables, and `price` is the target variable, split the data into train (80%) and test (20%).
3. (5 points) Using the `MinMaxScaler`, transform the input variables in the train and test dataset to 0-1 scale.
4. (7 points) Using the train dataset, build a MLP with a single hidden layer with 10 neurons and the `ReLU` as the activation function. Also use `optimizer = 'adam'`, `epochs = 100` and `batch_size = 100`. After that, use this model to predict on the test dataset. Report the MSE of this model.
5. (7 points) Using the train dataset, build a MLP with two hidden layers: the first one with 10 neurons and the `ReLU` as the activation function, the second one with 8 neurons and `ReLU` as the activation function. Also use `optimizer = 'adam'`, `epochs = 100` and `batch_size = 100`. After that, use this model to predict on the test dataset. Report the MSE of this model.
6. (3 points) Using the results from parts (4) and (5), what model would you use to predict car prices? Explain.