

Car Price Project

Prepared by: Hisham Abulfeilat

ID: 19110095

School of Computer & Informatics

11/8/2021

The Problem

An automobile company aspires to enter the US market by setting up their manufacturing unit there and producing cars locally to give competition to their US and European counterparts.

They have contracted an automobile consulting company to understand the **factors** on which the **pricing of cars depends**. Specifically, they want to understand the factors affecting the pricing of cars in the American market, since those may be very different from other markets. The company wants to know:

- Which variables are significant in predicting the price of a car
 - How well those variables describe the price of a car
-
- Based on various market surveys, the consulting firm has **gathered** a large dataset of different types of cars across the American market.

Goals

- To model the price of cars with the available independent variables.
- It will be used by the management to understand how exactly the prices vary with the independent variables.
- To help management manipulate the design of the cars, the business strategy etc. to meet certain price levels.
- Good way for management to understand the pricing dynamics of a new market.

Questions

- What is T (Task), E (Experience), P (Performance) for this specific problem?
 - The T is to compare & predict the price of a car with features of the car like horsepower and citympg.
 - The E is the cars and their features which is the dataset (training data) we gathered.
 - The P is the accuracy.
-
- What is the size of the training data? How many features do we have?
 - The size is 205 cars which we have 205 training instances or observation
 - We have 23 features

Methods Used

- Linear Regression:
- Scikit-learn package

Conclusions

Horsepower/Price

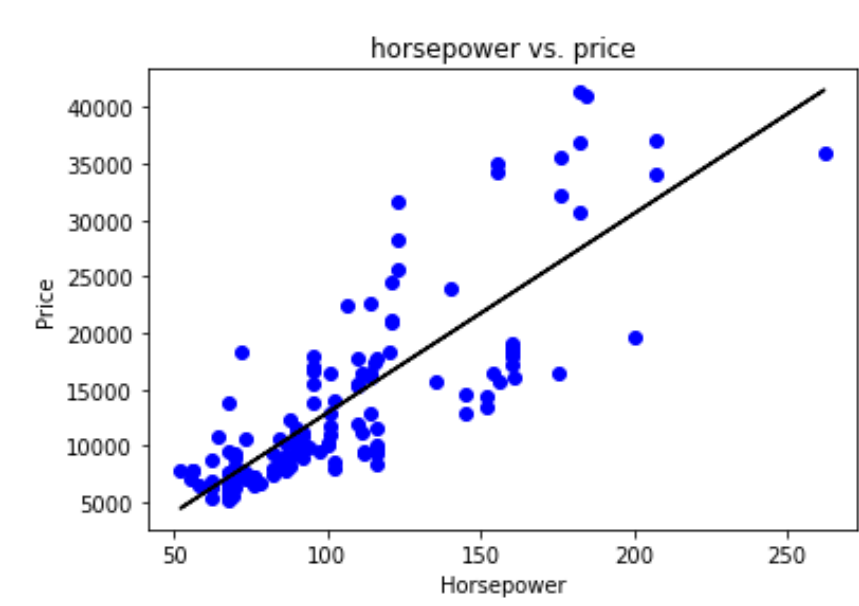
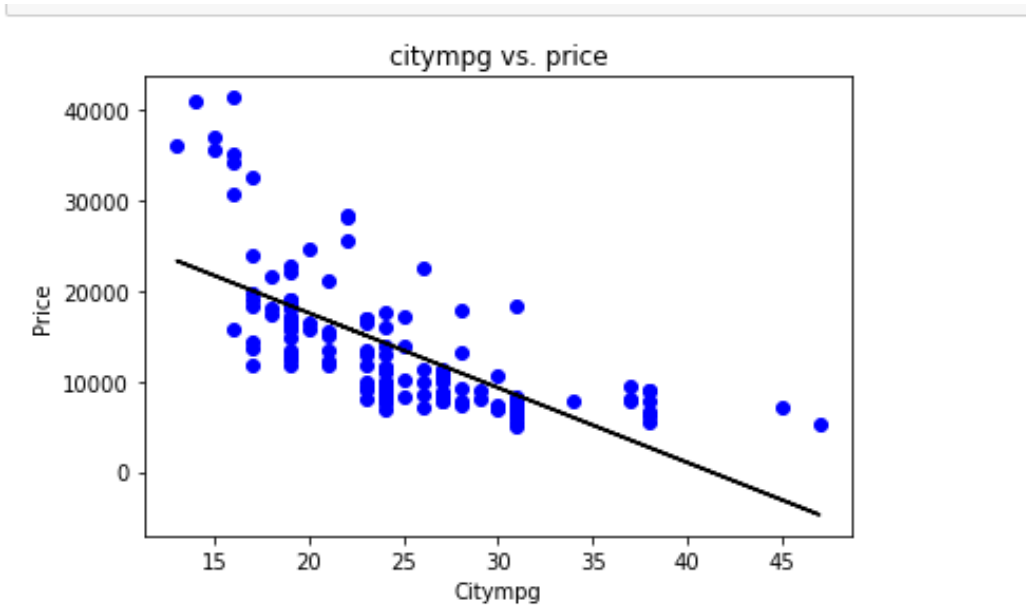
1:

	Actual	Predicted
115	16630.0	12400.619826
94	7299.0	7458.028160
24	6229.0	7281.507029
85	6989.0	10811.929648
202	21485.0	18931.901671

Citympg/Price

2:

	Actual	Predicted
159	7788.0	2684.525997
33	6529.0	9278.917095
94	7299.0	8454.618208
95	7799.0	8454.618208
59	8845.0	12576.112644



Future Work

- In future we can compare the price with other features to come up with top 5 features that affect car price
- Obstacle faced:
- Data size is a little big so processing the information is hard
- Wrong data
- Citympg increases while price decreases which doesn't make sense

Different models

- We can model the effect of both horsepower and citympg using Multilabel Classification model.
- Multi-label classification is a predictive modeling task that involves predicting zero or more mutually non-exclusive class labels.
- Neural network models can be configured for multi-label classification tasks

- Thank you for listening

