

Machine Learning Introduction Lab

(Note: I used Google Colab. Uploaded the .csv in sample_data)

Supervised Learning Algorithms - Linear Regression

Background: The term *Regression* is used when you try to find the relationship between variables. In Machine Learning which involves statistical modeling, this relationship is used for *prediction*.

1. Do the following:

- A. Create two lists
 - a. One which has the age of a car in years
 - b. Other has the speed the car runs for that age
- B. Plot this data as a Scatter Plot
- C. Find out the Coefficient of Correlation(r)
- D. Check if it is a good correlation for prediction or not
- E. Fit a Regression Line to the Scatter Plot
- F. Predict the speed of the car if the age of the car is 10 years old

Code Reference: [IntroMLLab_1.ipynb](#)

2. Create a CSV File of this data (Age & Car Speed) and do the above

Code Reference: [IntroMLLab_2.ipynb](#)

3. Create any bi-variant dataset and use Linear Regression to predict the values. Examples are:

- G. Temperature versus Number of Ice-cream sold
- H. Square Feet Area versus Price of House
- I. Student GRE Score versus Rank of University for PG Admission

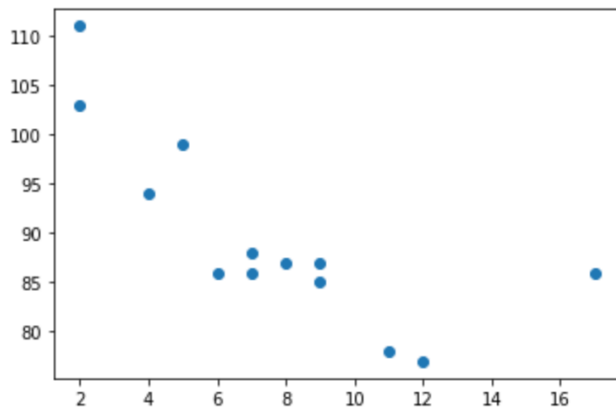
4. Create a dataset which results in a bad fit for Linear Regression. Analyze your output. Examples are:

- J. Student CET Rank versus Salary on Graduation
- K. Number of refrigerators in the shop versus number of ice creams sold

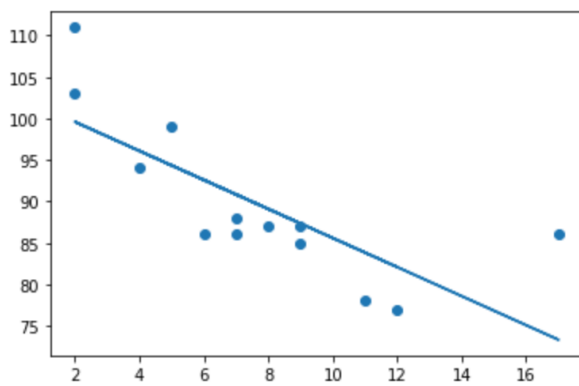
Code Reference: [IntroMLLab_3.ipynb](#)

Output Reference

1B - Scatter Plot



1E - Regression Line



4 - Bad Regression

