Sumon Singh
30.4.22

AIM :- Write a program to implement polynomial regression

1) Dataset Description:

A) Name of dataset: Real_estate_valuation_dataset xlsx

B) Description: This dataset consist of data of houses based on which we evaluate prices of the houses. The dataset is available on Kaggle website

c) Size :- This dataset consist of ~~414~~ 414 rows and 8 columns

D) Attributes :- This dataset has ~~8 features~~ 7 features, all of them are continious values.

a) NO. :- Serial No
b) transaction date
c) house age
d) distance to the nearest MRT station
e) Number of stores
f) Latitude
g) Longitude

Sumon Singh
30.4.22

30.4.22
Sumon Singh

E) Label :- The target data is price of the house which is generated based on features of the house.

2) Model Evaluation :-

A) Name :- Polynomial regression

B) Type :- This is a supervised learning model, used for regression type problem

C) Algorithm :-

Input :- features and target data of training dataset

step 1: Create n-dimension graph and plot all the datapoints ~~of the~~ on the graph

step 2: plot a best fit polynomial line on the datapoint of order 2 or more so that the generated output and the target data has the least error

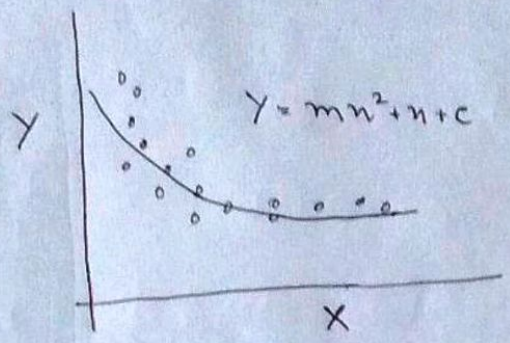
equation: $y = b_0 + b_1 x + b_2 x^2 + \dots b_n x^n$

where y is target, b_0 is intercept and $b_1, b_2 \dots b_n$ are slope

step 3: Based on the generated ~~graf~~ polynomial regression curve new data points are tested.

Sumon Singh
30.4.22

D) Draw :-



X → feature
Y → target

3) Result Analysis :-

* In this dataset we have applied 2 degree polynomial regression and the accuracy is 68%, that means predicted data are correctly recognized by the features.

* we have also tried other model such as Linear regression where accuracy is 57% and polynomial regression more than 2 degree is failing to be a good model so degree 2 polynomial is best for this dataset.