

BITS PILANI, DUBAI CAMPUS
DUBAI INTERNATIONAL ACADEMIC CITY, DUBAI

SECOND SEMESTER 2023 – 2024

COURSE: BITS F464 (Machine Learning)

COMPONENT: Practice Tutorial 1

DATE: 25 September 2023

Q1: Find the least square regression line, $y = ax + b$, for the following data. Also estimate the value of “y” if $x=10$.

X	Y
0	2
1	3
2	5
3	4
4	6

Answer:

$$y = 0.9x + 2.2$$

For $x=10$, $y=11.2$

Q2: Consider the following data for rating of Money Heist Season 1 and 2 from different users out of 5.

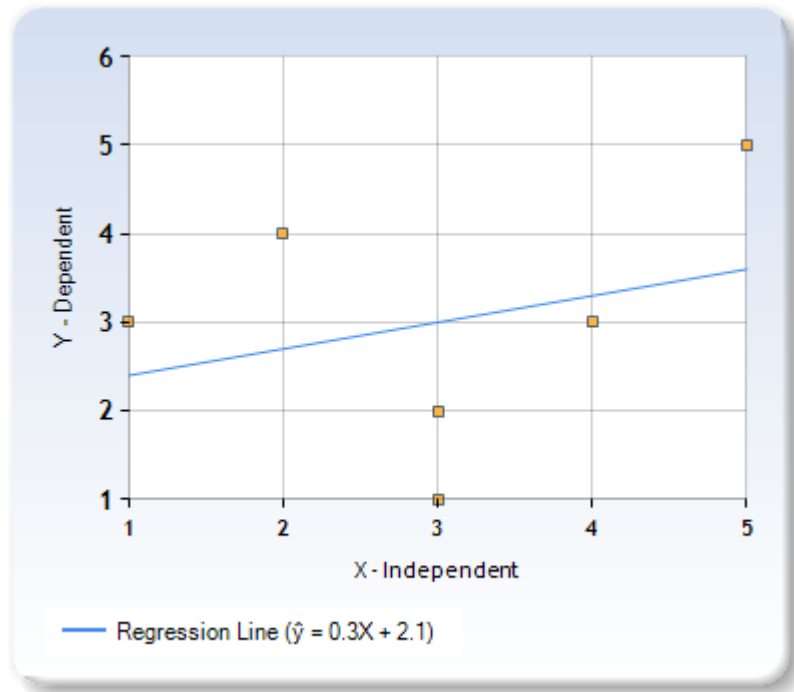
User	Season 1 ratings	Season 2 ratings
John	4	3
Ronny	2	4
Suresh	3	2
Rahul	5	5
Sameera	1	3
Diksha	3	1

- i. Find the regression line equation for given data.
- ii. Draw the regression line that best fit the given data.
- iii. Interpret and explain the equation of regression line.
- iv. Predict the season 2 rating for Dion if season 1 rating by Dion is 4.
- v. Also, calculate MAE and R^2 values for the obtained linear regression model.

Answers:

(i) $y = 0.3X + 2.1$

(ii)

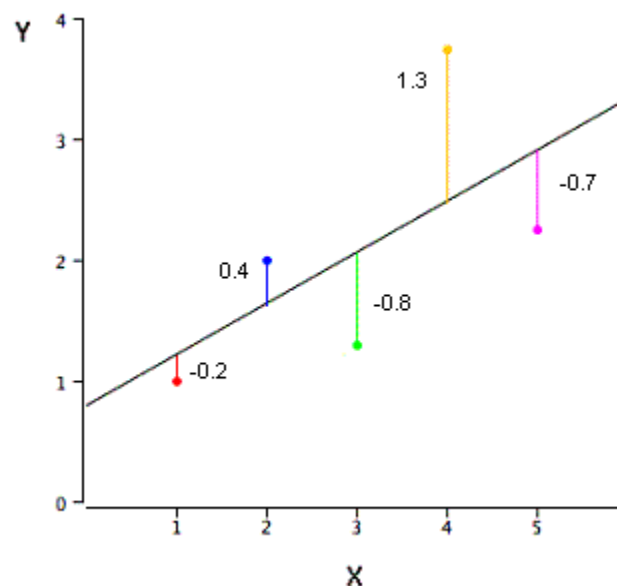


(iii) $SEE < 1$, Obtained regression line is good.

(iv) For $x=4$, $y = 3.3 \Rightarrow 3$

(v) $MAE = 1.1$, $R^2 = 0.09$

Q3: The graph below represents a regression line predicting Y from X. This graph shows the error of prediction for each of the actual Y values. Use this information to compute the standard error of the estimate in this sample.



Answer:

SEE = 1

Q4: Consider the dataset USA housing (USA_housing.csv) and perform following operations on it.

- i. Read the csv file and bulid the dataframe USAHousing.
- ii. Perform the exploratory data analytics using joinplot, and pariplot.
- iii. Build multiple linear regression model for predicting Price of house by considering Avg. Area Income, Avg. Area House Age, Avg. Area Number of Rooms, Avg. Area Number of Bedrooms, Area Population, as an independent variable.
- iv. Also apply regularization on your model to reduce the complexity of your model.