1. Explain the linear regression algorithm in detail. – Linear regression is a supervised machine learning algorithm. In regression model a target prediction value is predicted based on one or more predictor/independent variables. Linear regression is such regression where there exists a linear relationship between predictor and predicted variables and hence a straight line can be drawn through them. It follows equation y = mx +c. There are two types of linear regression, simple(where there is one independent variable) and multiple (where more than one independent variable exists).
2. What are the assumptions of linear regression regarding residuals? – Linear Regression assumptions regarding residuals/error are as follows: - i. Error terms are normally distributed with mean at 0. ii. Error terms are independent of each other. iii. Error terms have constant variance.
3. What is the coefficient of correlation and the coefficient of determination? – Coefficient of correlation also known as the pearson's r is used in statistics to measure how strong a relationship exists between variables. Coefficient of determination also denoted as r2 is interpreted as the proportion of the variance in the dependent variable that is predicted from the independent variable. It is calculated as 1-(RSS/TSS).
4. Explain the Anscombe's quartet in detail. – It's a set of 4 graphs with 11 data points each which signifies the need to visualize data before interpreting. All 4 sets are very identical when compared using simple summary but very different when graphed.
5. What is Pearson's R? – Pearson's R also called the coefficient of correlation is used in statistics to measure how strong a relationship exists between variables.
6. What is scaling? Why is scaling performed? What is the difference between normalized scaling and standardized scaling? - Feature scaling is a method used to normalize the range of independent variables or features of data. Scaling is preformed to make sure when different features are compared against one another who have a very different range, the relation is not miss interpreted. It basically brings down all the features to comparable terms. Normalized scaling – Brings all data in the range 0 and 1. Standardized scaling basically brings all the data into a standard normal distribution with mean 0 and standard deviation 1.
7. You might have observed that sometimes the value of VIF is infinite. Why does this happen? - An **infinite VIF** value indicates that the corresponding variable may be expressed exactly by a linear combination of other variables.
8. What is the Gauss-Markov theorem? – it states that in a in linear regression model in which the errors are uncorrelated have equal variance and expectation value of zero, (**BLUE**) of the coefficients is given by the ordinary least squares (OLS) estimator, provided it exists.
9. Explain the gradient descent algorithm in detail. – Its an optimization algorithm used to minimize the cost function. Its better to have small steps and move towards the direction of higher incline. This is carried out until u reach the bottom of the curve. It is basically used to reach the lowest point in the curve which is how it minimizes the cost function.
10. What is a Q-Q plot? Explain the use and importance of a Q-Q plot in linear regression. - A **Q-Q plot** is a scatterplot created by **plotting** two sets of quantiles against one another. If both sets of quantiles came from the same distribution, we should see the points forming a line that's roughly straight.