

## MACHINE LEARNING

1. C
2. B
3. C
4. B
5. B
6. A,D
7. B,C
8. A,C
9. A,B
- 10.

R-squared ( $R^2$ ) is a statistical measure that represents the proportion of the variance for a dependent variable that's explained by an independent variable or variables in a regression model. R Squared has no relation to express the effect of a bad or least significant independent variable on the regression.

Compared to R Squared which can only increase, Adjusted R Squared has the capability to decrease with the addition of less significant variables, thus resulting in a more reliable and accurate evaluation. Adjusted  $R^2$  is able to penalize as it considers the degree of freedom factors. Degree of freedom is given by:

$$d.o.f. = n - k - 1$$

where,

k is no of independent variables

n is the number of observation

Adjusted R Squared, however, makes use of the degree of freedom to compensate and penalize for the inclusion of a bad variable.

Adjusted R Squared can be expressed as :

$$Adj R^2 = 1 - (1 - R^2) * (n - 1) / d.o.f.$$

$$\text{That is, } Adj R^2 = 1 - (1 - R^2) * (n - 1) / (n - k - 1)$$

The value of Adjusted R Squared decreases as k increases also while considering R Squared acting a penalization factor for a bad variable and rewarding factor for a good or significant variable. Adjusted R Squared is thus a better model evaluator and can correlate the variables more efficiently than R Squared.

- 11.

In Ridge regression, we add a penalty term which is equal to the square of the coefficient. The cost function of ridge function looks like:

Cost function=  $MSE + \alpha * (\text{sum of square of coefficients})^2$

Lasso regression stands for Least Absolute Shrinkage and Selection Operator. It adds penalty term to the cost function. This term is the absolute sum of the coefficients. As the value of coefficients increases from 0 this term penalizes, cause model, to decrease the value of coefficients in order to reduce loss.

Cost function=  $MSE + \alpha * |\text{sum of magnitude of coefficients}|$

The difference between ridge and lasso regression is that it tends to make coefficients to absolute zero as compared to Ridge which never sets the value of coefficient to absolute zero.

12.

A variance inflation factor(VIF) detects multicollinearity in regression analysis. Multicollinearity is when there's correlation between predictors in a model; it's presence can adversely affect your regression results. The VIF estimates how much the variance of a regression coefficient is inflated due to multicollinearity in the model.

$$VIF = 1 / (1 - R^2)$$

VIF of 2.5 or above but less than 9 is suitable for regression modeling.

13.

Scaling is one of the important pre-processing that is required for standardizing/normalization of the input data. When the range of values are very distinct in each column, we need to scale them to the common level. The values are brought to common level and then we can apply further machine learning algorithm to the input data. One way to scale the values is to bring the values of all the column between 0 to 1 or we can bring them to common level having values between -3 to 3.

The other is gradient descent:

The gradient descent algorithm which is used to reach the optimal solution in most of the cases, it reached the optimal solution much faster if all the features are at the same scale. That's why scaling helps to reach the optimal solution.

14.

The metrics that we can use to check the goodness of fit for linear regressions are as follow:

Mean Absolute Error and Mean Square Error

Root Mean Squared Error

Relative Absolute Error and Relative Squared Error

$R^2$  and Adjusted  $R^2$

15.

SENSITIVITY OR RECALL: 0.9523

SPECIFICITY: 0.8275

ACCURACY: 0.88