MACHINE LEARNING

1. **In which of the following you can say that the model is overfitting?**
   Answer -
2. **Which among the following is a disadvantage of decision trees?**
   Answer – B
3. **Which of the following is an ensemble technique?**
   Answer – C
4. **Suppose you are building a classification model for detection of a fatal disease where detection of the disease is most important. In this case which of the following metrics you would focus on?**
   Answer – A
5. **The value of AUC (Area under Curve) value for ROC curve of model A is 0.70 and of model B is 0.85. Which of these two models is doing better job in classification?**
   Answer – B
6. **Which of the following are the regularization technique in Linear Regression?**
   Answer – A & D
7. **Which of the following is not an example of boosting technique?**
   Answer – B & C
8. **Which of the techniques are used for regularization of Decision Trees?**
   Answer – A & B
9. **Which of the following statements is true regarding the Adaboost technique?**
   Answer – D
10. **Explain how does the adjusted R-squared penalize the presence of unnecessary predictors in the model?**
    Answer – The new predictor enhances the model above what would be obtained by probability. Conversely, it will decrease when a predictor improves the model less than what is predicted by chance.
    If R-squared does not increase significantly on the addition of a new independent variable, then the value of Adjusted R-squared will actually decrease. The adjusted R-squared compensates for the addition of variables and only increases if

    - Use predicted R-squared to determine how well a regression model makes predictions. This statistic helps you identify cases where the model provides a good fit for the existing data but isn't as good at making predictions. However, even if you aren't using your model to make predictions, predicted R-squared still offers valuable insights about your model.

11. **Differentiate between Ridge and Lasso Regression**.

Answer – **Lasso** is a modification of linear regression, where the model is penalized for the sum of absolute values of the weights. Thus, the absolute values of weight will be (in general) reduced, and many will tend to be zeros. During training, the objective function become:

$$\frac{1}{2m}\sum_{i=1}^{m}(y - Xw)^2 + alpha\sum_{j=1}^{p}|w_j|$$

As you see, Lasso introduced a new hyperparameter, *alpha*, the coefficient to penalize weights.

Ridge takes a step further and penalizes the model for the sum of squared value of the weights. Thus, the weights not only tend to have smaller absolute values, but also really tend to penalize the extremes of the weights, resulting in a group of weights that are more evenly distributed. The objective function becomes:

$$\sum_{i=1}^{n}(y - Xw)^2 + alpha\sum_{j=1}^{p}w_j^2$$

12. What is VIF? What is the suitable value of a VIF for a feature to be included in a regression modelling?

Answer – A variance inflation factor (VIF) is a measure of the amount of multicollinearity in regression analysis. Multicollinearity exists when there is a correlation between multiple independent variables in a multiple regression model. This can adversely affect the regression results. Thus, the variance inflation factor can estimate how much the variance of a regression coefficient is inflated due to multicollinearity.

KEY TAKEAWAYS

- A variance inflation factor (VIF) provides a measure of multicollinearity among the independent variables in a multiple regression model.
- Detecting multicollinearity is important because while multicollinearity does not reduce the explanatory power of the model, it does reduce the statistical significance of the independent variables.
- A large VIF on an independent variable indicates a highly collinear relationship to the other variables that should be considered or adjusted for in the structure of the model and selection of independent variables.

What Is a Good VIF Value?
As a rule of thumb, a VIF of three or below is not a cause for concern. As VIF increases, the less reliable your regression results are going to be.

What Does a VIF of 1 Mean?
A VIF equal to one means variables are not correlated and multicollinearity does not exist in the regression model.

What Is VIF Used for?

VIF measures the strength of the correlation between the independent variables in regression analysis. This correlation is known as multicollinearity, which can cause problems for regression models.

The Bottom Line

While a moderate amount of multicollinearity is acceptable in a regression model, a higher multicollinearity can be a cause for concern.

Two measures can be taken to correct high multicollinearity, First, one or more of the highly correlated variables can be removed, as the information provided by these variables is redundant. The second method is to use principal components analysis or partial least square regression instead of OLS regression, which can respectively reduce the variables to a smaller set with no correlation, or create new uncorrelated variables. This will improve the predictability of a model.

13. Why do we need to scale the data before feeding it to the train the model?

Answer – After building an understanding of how to do data scaling and which data scaling techniques to use, we can now talk about where to use these data scaling techniques.

If an algorithm uses gradient descent, then the difference in ranges of features will cause different step sizes for each feature. To ensure that the gradient descent moves smoothly towards the minima and that the steps for gradient descent are updated at the same rate for all the features, we scale the data before feeding it to the model. Having features on a similar scale will help the gradient descent converge more quickly towards the minima.

Specifically, in the case of Neural Networks Algorithms, feature scaling benefits optimization by:

- It makes the training faster
- It prevents the optimization from getting stuck in local optima
- It gives a better error surface shape
- Weight decay and Bayes optimization can be done more conveniently