

# **Machine Learning Assignment**

In Q1 to Q11, only one option is correct, choose the correct option:

1. A) Least Square Error
2. A) Linear regression is sensitive to outliers.
3. B) negative.
4. A) Regression.
5. C) Low bias and high variance
6. B) Predictive model
7. D) Regularization.
8. D) SMOTE
9. C) Sensitivity and Specificity.
10. A) True
11. B) Apply PCA to project high dimensional data.

Q12, more than one options are correct, choose all the correct options:

12. A) and B)

Q13 and Q15 are subjective answer type questions, Answer them briefly

13. Explain the term regularization?

In general, regularization means to make things regular or acceptable. In the context of machine learning, regularization is the process which regularizes or shrinks the coefficients towards zero. In simple words, regularization discourages learning a more complex or flexible model, to prevent overfitting.

Regularization is a technique used to reduce the errors by fitting the function appropriately on the given training set and avoid overfitting.

14. Which particular algorithms are used for regularization?

1. There are three main regularization techniques, namely: Ridge Regression (L2 Norm), Lasso (L1 Norm), Dropout. Ridge and Lasso can be used for any algorithms involving weight parameters, including neural nets. Dropout is primarily used in any kind of neural networks e.g., ANN, DNN, CNN or RNN to moderate the learning. Let's take a closer look at each of the techniques

15. Explain the term error present in linear regression equation?

Within a linear regression model tracking a stock's price over time, the error term is the difference between the expected price at a particular time and the price that was actually observed. The error term stands for any influence being exerted on the price variable, such as changes in market sentiment.

The error term is also known as the residual, disturbance, or remainder term, and is variously represented in models by the letters  $e$ ,  $\epsilon$ , or  $u$ .

$$Y = \alpha X + \beta \rho + \epsilon;$$

$\alpha, \beta$ =Constant parameters,  $X, \rho$ =Independent variables,  
 $\epsilon$ =Error term