

## MACHINE LEARNING WORKSHEET

1. Which of the following methods do we use to find the best fit line for data in Linear Regression?

A) Least Square Error

2. Which of the following statement is true about outliers in linear regression?

A) Linear regression is sensitive to outliers

3. A line falls from left to right if a slope is \_\_\_\_\_?

B) Negative

4. Which of the following will have symmetric relation between dependent variable and independent variable?

B) Correlation

5. Which of the following is the reason for over fitting condition?

C) Low bias and high variance

6. If output involves label then that model is called as:

B) Predictive modal

7. Lasso and Ridge regression techniques belong to \_\_\_\_\_?

B) Regularization

8. To overcome with imbalance dataset which technique can be used?

D) SMOTE

9. The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for binary classification problems. It uses \_\_\_\_\_ to make graph?

A) TPR and FPR

10. In AUC Receiver Operator Characteristic (AUCROC) curve for the better model area under the curve should be less.

B) False

11. Pick the feature extraction from below:

B) Apply PCA to project high dimensional data

12. Which of the following is true about Normal Equation used to compute the coefficient of the Linear Regression?

- A) We don't have to choose the learning rate.
- B) It becomes slow when number of features is very large.
- D) It does not make use of dependent variable.

Explain the term regularization?

the term 'regularization' refers to a set of techniques that help the machine to learn more than just memorize.

The term 'regularization' refers to a set of techniques that regularizes learning from particular features for traditional algorithms or neurons in the case of neural network algorithms.

This is a form of regression, that constrains/ regularizes or shrinks the coefficient estimates towards zero. In other words, this technique discourages learning a more complex or flexible model, so as to avoid the risk of overfitting.

A simple relation for linear regression looks like this. Here Y represents the learned relation and  $\beta$  represents the coefficient estimates for different variables or predictors(X).

$$Y \approx \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p$$

The fitting procedure involves a loss function, known as residual sum of squares or RSS. The coefficients are chosen, such that they minimize this loss function.

$$RSS = \sum_{i=1}^n \left( y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2 .$$

Now, this will adjust the coefficients based on your training data. If there is noise in the training data, then the estimated coefficients won't generalize well to the future data. This is where regularization comes in and shrinks or regularizes these learned estimates towards zero.

14. Which particular algorithms are used for regularization?

There are three main regularization techniques, namely:

1. Ridge Regression (L2 Norm)
2. Lasso (L1 Norm)
3. Dropout

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

An error term essentially means that the model is not completely accurate and results in differing results during real-world applications. For example, assume there is a multiple linear regression function that takes the following form:

$$Y = \alpha X + \beta p + \epsilon$$

where:

$\alpha, \beta$  = Constant parameters

X, p = Independent variables

$\epsilon$  = Error term

When the actual Y differs from the expected or predicted Y in the model during an empirical test, then the error term does not equal 0, which means there are other factors that influence Y.