MACHINE LEARNING

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

1. Which of the following methods do we use to find the best fit line for data in Linear Regression?
   A) **Least Square Error** B) Maximum Likelihood
   C) Logarithmic Loss D) Both A and B

2. Which of the following statement is true about outliers in linear regression?
   A) **Linear regression is sensitive to outliers** B) linear regression is not sensitive to outliers
   C) Can't say D) none of these

3. A line falls from left to right if a slope is_____?
   A) Positive B) **Negative**
   C) Zero D) Undefined

4. Which of the following will have symmetric relation between dependent variable and independent variable?
   A) Regression B) **Correlation**
   C) Both of them D) None of these

5. Which of the following is the reason for over fitting condition?
   A) High bias and high variance B) Low bias and low variance
   C) **Low bias and high variance** D) none of these

6. If output involves label then that model is called as:
   A) Descriptive model B) **Predictive model**
   C) Reinforcement learning D) All of the above

7. Lasso and Ridge regression techniques belong to_____?
   A) Cross validation B) Removing outliers
   C) SMOTE **D) Regularization**

8. To overcome with imbalance dataset which technique can be used?
   A) **Cross validation** B) Regularization
   C) Kernel 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** B) Sensitivity and precision
   C) Sensitivity and Specificity D) Recall and precision

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

11. Pick the feature extraction from below:
    A) Construction bag of words from a email
    B) Apply PCA to project high dimensional data
    C) Removing stop words
    D) Forward selection

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

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.
    C) We need to iterate.
    D) It does not make use of dependent variable.

MACHINE LEARNING

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

13. Explain the term regularization?

Ans. Regularization can be defined as regression method that tends to minimize or shrink the regression coefficients towards zero. The need for regularization arises when the regression co-efficient becomes too large which leads to overfitting, for instance in the case of polynomial regression, the value of regression can shoot up to large numbers when the degree of the polynomials increases. To avoid the risk of overfitting we use regularization. One important pre-requisite of Regularization is the predictors to be scaled. It's a scaled variant, keep that in mind. Regularization penalizes our model and tries to reduce bias(complexity of the model or flexibility).

14. Which particular algorithms are used for regularization?

Ans. There are 2 algorithms used for regularization:-

- **A.** LASSO Regression (L1 Form): Least Absolute Shrinkage and selection operator.
- **B.** RIDGE Regression (L2 Form).

Lasso Regression (L1 form): This regression penalizes(nullify) the model based on the sum of magnitude of the coefficients.

In other words, it will give zero importance to those features which has no relation with result.

Formula ➔ $\lambda * \Sigma | B_j |$

λ **–** Shrinking Factor

Ridge Regression (L2 Form): This regression will give very less importance to those features which has no relation with the end result. Unlike Lasso regression, it will not give zero importance but it will give very less importance.

Formula ➔ $\lambda * \Sigma |(B_j)^2|$

λ **–** Shrinking Factor

Note: Lasso and Ridge can be imported from:

```
from sklearn.linear_model import Ridge, Lasso, LassoCV, RidgeCV
```

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

Ans. There are 3 error scores which are used for Model Evaluation.

1. Mean Absolute error or MAE
2. Mean Squared error or MSE
3. Root Mean Squared error or RMSE

Mean Absolute error or MAE: It represents average error.

Mean Squared error or MSE: Similar to MAE but noise is exaggerated, and larger points are "punished". It is harder to interpret than MAE as it is not in base units, however, it is generally more popular.

Root Mean Squared error or RMSE: Most popular metric, similar to MSE, however the result is square-

MACHINE LEARNING

rooted to make it more interpretable as it is in base units. It is recommended that RMSE be used as the primary metrics to interpret the model.

Note: MAE, MSE can be imported from:

```
from sklearn.metrics import mean_squared_error, mean_absolute_error
```