QUIZ TOPIC - REGRESSION

1. **Suppose you have to predict the salary of an employee from their years of experience where the dataset has a salary range from 10000 to 50000. In which of the intervals your regressive model should predict?**

- ○ A. 10000 to 20000
- ○ B. 10000 to 40000
- ○ C. 25000 to 50000
- ◉ D. 10000 to 50000 ✔

2. **In simple linear regression, if you change the input value by 1 then output value will be changed by:**

- ○ A. 1
- ◉ B. The slope parameter ✔
- ○ C. The intercept parameter
- ○ D. None

3. **You can compute the residual by-**

- ◉ A. actual y-coordinate value - predicted y-coordinate value ✔
- ○ B. predicted y-coordinate value - actual y coordinate value
- ○ C. actual y-coordinate value / predicted y-coordinate value
- ○ D. None

4. **How to see the value of residuals geometrically**

- ◉ A. The perpendicular distance between a data point and the regression line ✔
- ○ B. The euclidian distance between a data point and the regression line
- ○ C. The horizontal distance between a data point and the regression line
- ○ D. The vertical distance between a data point and the regression line

5. **The equation of the regression line is y = 5x + 3. Predict y when x = 8.**

- ◉ A. 43 ✔
- ○ B. 53
- ○ C. 23
- ○ D. None

6. **The equation of the regression line is y = 8x - 2. Compute the residual for the point (4, 28)**

- A. 2
- B. 1
- C. -2 ✔
- D. 4

7. **What would be the best regression model for more than one independent variable?**

- A. Simple Linear Regression
- B. Multiple Linear Regression ✔
- C. Logistic Regression
- D. All of the Above

8. **Suppose you have observed that you data has an exponential growth tendency. Then what regression model you should use-**

- A. Simple linear regression
- B. Multiple linear regression
- C. Polynomial regression ✔
- D. Logistic regression

9. **Can we perform linear regression with a neural network?**

- A. Yes, we can ✔
- B. No, we can not
- C. Partially we can
- D. None

10. **If you get a poor accuracy using a simple linear regression model. What will be the cause behind it-**

- A. The data was not linear
- B. The data has outliers
- C. Both A or B depending on the context ✔
- D. None

11. **If your data grows in a non-linear fashion. Which model won't perform well?**

- A. Polynomial regression
- B. Random forest regression
- C. Simple linear regression ✔
- D. None

12. **Suppose you got a training accuracy of 90% and a test accuracy of 50%. What happened with your model-**

- (●) A. The model was over fitted with the training data ✔
- ( ) B. The model was under fitted with the training data
- ( ) C. The model is absolutely fine
- ( ) D. None

13. **What is a support vector?**

- ( ) A. The average distance between all the data points
- ( ) B. The distance between any two data points
- (●) C. The distance between two boundary data points ✔
- ( ) D. The minimum distance between any two data points

14. **What is a kernel?**

- ( ) A. A function that calculates the distance of two boundary data points
- (●) B. A function that maps the value from one dimension to the other ✔
- ( ) C. A function that predicts the output value of a regression
- ( ) D. None

15. **Which of the following is not a kernel?**

- ( ) A. Polynomial Kernel
- ( ) B. Gaussian Kernel
- ( ) C. Sigmoid Kernel
- (●) D. None ✔

16. **What does epsilon represent in Support Vector Regression?**

- ( ) A. Boundary threshold
- (●) B. Error threshold ✔
- ( ) C. Distance threshold
- ( ) D. None

17. **In Regression, a decision tree splits the dataset based on-**

- (●) A. Information entropy ✔
- ( ) B. Information gain
- ( ) C. Both A and B
- ( ) D. None

18. **Which one is a different algorithm?**

- (●) A. Logistic Regression ✔
- ( ) B. Support Vector Regression
- ( ) C. Polynomial Regression
- ( ) D. None

19. **Which one is not a better algorithm in the sense of overfitting?**

- (●) A. Simple linear regression ✔
- ( ) B. Decision tree
- ( ) C. Random forest
- ( ) D. All of the above

20. **If the actual value of a data point is 50 and the predicted value is 55, what will be the Mean Absolute Error(MAE)**

- ( ) A. -5
- (●) B. 5 ✔
- ( ) C. 2.5
- ( ) D. -2.5

21. **Which of the following is a regression algorithm?**

- (●) A. Linear Regression ✔
- ( ) B. Logistic Regression
- ( ) C. Both A and B
- ( ) D. None

22. **Suppose you have to predict the salary of employees from their experience. This is a-**

- ( ) A. Classification task
- (●) B. Regression task ✔
- ( ) C. Clustering task
- ( ) D. None

23. **Regression is a-**

- (●) A. Supervised Learning Algorithm ✔
- ( ) B. Unsupervised Learning Algorithm
- ( ) C. Reinforcement Learning Algorithm
- ( ) D. None

24. **Which of the following is/are true about Normal Equation?**

- ○ The equation itself choose the learning rate
- ○ B. Becomes slower with a large number of features ✖
- ○ C. Iteration is not required
- ◉ D. All of them ✔

25. **Which methods are used to find the best fit line in linear regression?**

- ○ A. Logarithmic Loss
- ○ B. Area Under Curve
- ○ C. Both A and B
- ◉ D. Least Square Error ✔

26. **What will happen when you increase the size of training data?**

- ○ A. Bias decreases and Variance increases
- ○ B. Bias increases and Variance increases
- ◉ C. Bias increases and Variance decreases ✔
- ○ D. Bias decreases and Variance decreases

27. **If you fit 2 degree polynomial in linear regression-**

- ○ A. The model will overfit the data
- ◉ B. The model will underfit the data ✔
- ○ C. The model will perform perfectly
- ○ D. None

Explanation: Higher degree polynomials have chances to underfit at a lower degree.

28. **Which of the following evaluation metrics can be used for Regression?**

- ○ A. AUC-ROC
- ◉ B. Mean-Squared-Error ✔
- ○ C. Accuracy
- ○ D. f1 score

Explanation: Regression gives continuous output. So, we use Mean-Squared-Error or MSE as evaluation metric. Rest are used in classification.

29. **Linear regression is-**

- ◉ A. sensitive to outliers ✔
- ○ B. not sensitive to outliers
- ○ C. not affected by outliers
- ○ D. None

Explanation: The regression line changes due to outliers. So, it is sensitive to outliers.

30. **What is true about Residuals?**

- A. Higher is better
- B. Lower is better ✔
- C. A or B depending on the context
- D. None