Here are the correct answers to the MCQs:

1. Using a goodness of fit, we can assess whether a set of obtained frequencies differ from a set of ___ frequencies.
    - d) Expected
2. Chi-square is used to analyze:
    - c) Frequencies
3. What is the mean of a Chi-Square distribution with 6 degrees of freedom?
    - c) 6
4. Which of these distributions is used for goodness of fit testing?
    - b) Chi-square distribution
5. Which of the following distributions is Continuous?
    - c) F Distribution
6. A statement made about a population for testing purposes is called:
    - b) Hypothesis
7. If the assumed hypothesis is tested for rejection considering it to be true, it is called:
    - a) Null Hypothesis
8. If the Critical region is evenly distributed, then the test is referred to as:
    - a) Two-tailed
9. Alternative Hypothesis is also called:
    - b) Research Hypothesis
10. In a Binomial Distribution, if 'n' is the number of trials and 'p' is the probability of success, then the mean value is given by:

- a) np

1. **R-squared vs. RSS for Goodness of Fit**:
   - **R-squared** is generally considered a better measure of goodness of fit as it represents the proportion of variance in the dependent variable that is predictable from the independent variables. **RSS** measures the total deviation of the response values from the predicted values, but it does not provide a normalized measure. R-squared is scaled between 0 and 1, making it easier to interpret.
2. **TSS, ESS, and RSS**:
   - **TSS (Total Sum of Squares)**: Measures the total variance in the dependent variable.
   - **ESS (Explained Sum of Squares)**: Measures the variance explained by the regression model.
   - **RSS (Residual Sum of Squares)**: Measures the variance not explained by the model.
   - The equation relating these metrics: TSS=ESS+RSSTSS = ESS + RSSTSS=ESS+RSS.
3. **Need for Regularization**:
   - Regularization is needed to prevent overfitting by penalizing large coefficients in the model. It helps in improving the model's generalizability to new data by adding a penalty term to the loss function.
4. **Gini Impurity Index**:
   - The Gini Impurity Index is a measure used in decision trees to determine the purity of a node. It calculates the probability of a randomly chosen element being incorrectly classified if it was randomly labeled according to the distribution of labels in the node.
5. **Unregularized Decision Trees and Overfitting**:
   - Yes, unregularized decision trees are prone to overfitting because they can create very complex trees that perfectly classify the training data, but may not generalize well to unseen data.
6. **Ensemble Technique**:
   - Ensemble techniques combine predictions from multiple models to improve overall performance. Common methods include bagging, boosting, and stacking.
7. **Bagging vs. Boosting**:
   - **Bagging** (Bootstrap Aggregating) involves training multiple models independently using different subsets of the data and then averaging their predictions.
   - **Boosting** involves training models sequentially, with each model attempting to correct the errors of the previous one.
8. **Out-of-Bag Error**:
   - In random forests, out-of-bag error is the average prediction error for each training sample, calculated using only the trees that did not have the sample in their bootstrap sample. It provides an unbiased estimate of the model error.
9. **K-fold Cross-Validation**:
   - K-fold cross-validation involves dividing the dataset into K equally sized folds. The model is trained K times, each time using a different fold as the validation set and the remaining K-1 folds as the training set. The results are then averaged to produce a final estimate.

10. **Hyperparameter Tuning**:
    - Hyperparameter tuning involves selecting the best hyperparameters for a machine learning model. It is done to optimize model performance, as hyperparameters can significantly affect the model's ability to learn from data.
11. **Issues with Large Learning Rate in Gradient Descent**:
    - A large learning rate can cause the model to overshoot the optimal solution, leading to divergence instead of convergence. This can result in oscillations or completely missing the minimum point.
12. **Logistic Regression for Non-Linear Data**:
    - Logistic Regression cannot directly handle non-linear data as it assumes a linear relationship between the independent variables and the log-odds of the dependent variable. For non-linear data, techniques like polynomial features or kernel methods are required.
13. **Adaboost vs. Gradient Boosting**:
    - **Adaboost** focuses on adjusting the weights of incorrectly classified instances, thereby giving more importance to difficult cases.
    - **Gradient Boosting** builds models sequentially, where each new model is trained to correct the residual errors of the previous models using gradient descent.
14. **Bias-Variance Trade-off**:
    - The bias-variance trade-off is the balance between the error introduced by bias (error due to overly simplistic models) and the error introduced by variance (error due to overly complex models). An optimal model minimizes both bias and variance to achieve good generalization.
15. **Kernels in SVM**
    - **Linear Kernel**: Used when the data is linearly separable. It computes the dot product between two vectors.
    - **RBF (Radial Basis Function) Kernel**: Handles non-linear data by mapping input space to a higher-dimensional space using a Gaussian function.
    - **Polynomial Kernel**: Computes the similarity between vectors in a polynomial fashion, suitable for non-linear data with polynomial relationships.