Decision Tree Classifier And Regressor

Interview Questions:

1. Decision Tree
2. Entropy, Information Gain, Gini Impurity
3. Decision Tree Working For Categorical and Numerical Features
4. What are the scenarios where Decision Tree works well
5. Decision Tree Low Bias And High Variance- Overfitting
6. Hyperparameter Techniques
7. Library used for constructing decision tree
8. Impact of Outliers Of Decision Tree
9. Impact of mising values on Decision Tree
10. Does Decision Tree require Feature Scaling

Random Forest Classifier And Regresor

1. Ensemble Techniques(Boosting And Bagging)
2. Working of Random Forest Classifier
3. Working of Random Forest Regresor
4. Hyperparameter Tuning(Grid Search And RandomSearch)

Important properties of Random Forest Classifiers

1. Decision Tree---Low Bias And High Variance
2. Ensemble Bagging(Random Forest Classifier)--Low Bias And Low Variance

1. What Are the Basic Assumption?

There are no such assumptions

2. Advantages

1. Doesn't Overfit
2. Favourite algorithm for Kaggle competition
3. Less Parameter Tuning required
4. Decision Tree can handle both continuous and categorical variables.
5. No feature scaling required: No feature scaling (standardization and normalization) required in case of Random Forest as it uses DEcision Tree internally
6. Suitable for any kind of ML problems

3. Disadvantages

1.Biased With features having many categories

2.Biased in multiclass classification problems towards more frequent classes.

4. Whether Feature Scaling is required?

No

6. Impact of outliers?

Robust to Outliers

Types of Problems it can solve(Supervised)

1. Classification
2. Regression

Practical Implementation

1. https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html
2. https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeRegressor.html
3. https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html
4. https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html

Performance Metrics

Classification

1. Confusion Matrix
2. Precision,Recall, F1 score

Regression

1. R2,Adjusted R2
2. MSE,RMSE,MAE