Attempt #1

Jul 15, 3:13 PM

Marks: 30

**Question 1**

Correct Answer

Marks: 3/3

Decision Trees can be used to to solve \_\_\_\_\_\_\_\_.

Regression problems

Classification problems

Regression as well as Classification problems

You Selected

Neither Regression nor Classification problems

The decision tree algorithm can be used to solve both kinds of problems - classification as well as regression.

**Question 2**

Correct Answer

Marks: 3/3

How is entropy related to the purity of the node?

Higher the value of entropy, higher is the purity of the node

Lower the value of entropy, higher is the purity of the node

You Selected

Depends on the tree

Entropy is not a measure of purity of the node

Entropy is the measure of purity (or impurity) in a decision tree. The lower the entropy, the higher is the purity of the node and vice versa.

**Question 3**

Correct Answer

Marks: 3/3

Which of the following is/are the advantages(s) of decision tree?

A) It requires little data preparation

B) It can handle both categorical and numerical data

C) A small change in the training data will result in a large change in the tree

Only A

A and B

You Selected

B and C

A and C

The advantages of a decision tree are:

1. It requires little data preparation
2. Decision trees can handle both numerical and categorical data

One of the main disadvantages of the decision tree is that a small change in the training data can result in a large change in the tree and consequently the final predictions.

**Question 4**

Correct Answer

Marks: 3/3

 \_\_\_\_\_\_\_ is a technique that reduces the size of decision trees by removing branches of the trees to avoid overfitting in a fully grown decision tree.

Cross-validation

Pruning

You Selected

Test-Train Splitting

Bootstrapping

Pruning is a process of reducing the size of the tree so that it can become simpler and generalize better on unseen data. In the pruning process, the nodes that are adding the least information to the model are pruned/removed in order to avoid overfitting and improve the model performance.

**Question 5**

Correct Answer

Marks: 3/3

The process of using multiple data samples, generated by random sampling with replacement, from the original dataset to create multiple models and aggregating their predictions is called \_\_\_\_\_\_\_.

Pruning

Clustering

Bagging

You Selected

Out of bag errors

In bagging (bootstrap + aggregation), various datasets are created, by random sampling with replacement, and on these different datasets, models are trained. Then, the prediction of all the models is combined to get the final prediction by the bagging model.

**Question 6**

Correct Answer

Marks: 3/3

In general, which of the following is/are the advantage(s) of ensemble techniques?

A) Better Prediction

B) Lower time of execution

C) Simpler than the base model

Only A

You Selected

B and C

A and B

Only C

The ensemble models combine many base models' predictions to get a final prediction, therefore:

1. They are generally more accurate than their base models.

2. As they combine many models they have a higher time of execution.

3. They are more complex than the base model.

**Question 7**

Correct Answer

Marks: 3/3

In a regression problem, for a new test data point, the final prediction by a Random Forest is done by taking the \_\_\_\_\_\_\_\_\_

mode of the individual predictions

minimum of individual predictions

average of individual predictions

You Selected

median of individual predictions

Random forest regression is a bit different from the classification. In classification where we take the mode of the predictions made by the different models, in regression the mean of the predictions are taken.

**Question 8**

Correct Answer

Marks: 3/3

Which of the following can be component(s) of a Time Series?

Trend

Seasonality

Noise

All of the above

You Selected

Trend, seasonality, and noise are components of a time series. The trend shows the variability of the feature over a period of time, if it is increasing with time, then it is a positive trend and if it is decreasing, then it is a negative trend.

Seasonality shows the seasonal behavior of the variable. If it is showing some pattern for a specific period of time, then it is seasonal.

Noise is random fluctuations in the time series that cannot be captured by the model.

**Question 9**

Correct Answer

Marks: 3/3

Which of the following is true for an autoregressive model?

In an autoregressive model, the value from a time series is regressed on future values from that same time series

In an autoregressive model, the value from a time series is regressed on previous values from that same time series

You Selected

In an autoregressive model, the value from a time series is regressed on previous values from a different time series

In an autoregressive model, the value from a time series is regressed on future values from a different time series

In an autoregression model, the target value is the current timestamp value and the input features are the previous timestamp values.

**Question 10**

Correct Answer

Marks: 3/3

When do you say that a Time Series is white noise?

When it has a constant mean of zero, a constant variance and there is a correlation between its values at different times

When it has a constant mean of zero, a constant variance and no correlation between its values at different times

You Selected

When it has a constant mean of one, a constant variance of zero and no correlation between its values at different times

None of the above

A time series is said to have white noise when the mean of the series is zero, the variance is constant and no auto-correlation is there.