

Data Partitioning and Modeling

The data was partitioned into train and test datasets.

The training data set was used to create the decision tree model.

The trained model was then applied to the testing dataset.

This is important because we need both training and testing dataset. And they normally

When partitioning the data using sampling, it is important to set the random seed because this will let us get the same partitioning result every time, and hence get the reproducible result.

A screenshot of the resulting decision tree can be seen below:

