

Data Partitioning and Modeling

The data was partitioned into train and test datasets.

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

The trained model was then applied to the test dataset.

This is important because want to test our model on data that was not used to train it.

When partitioning the data using sampling, it is important to set the random seed because it ensures that we get reproducible results.

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

