

Optical Recognition of Handwritten Digits dataset

-Successfully implemented Decision tree algorithm and obtained an accuracy of 76% for the test data.

-Used rpart library for implementing in R

- Also, tried to implement the Decision tree in python.

-Challenges faced: Implementation of decision tree without using libraries is the most difficult part of the assignment as it has 10 classifications. Hence libraries are used to implement the algorithm.

-As it was first time working in R, it took some time to understand the language syntax and features.

Amazon reviews sentiment analysis dataset

-Created an algorithm to find the sentiment score of each review by counting the number of positive words and negative words.

-Some anomalies are present in this algorithm like high number of negative words for some five star reviews and vice versa.

-Challenging task was the implementation of sentiment score algorithm and the determination of positive and negative words.