Q41

Rahul Atre

2023-11-27

Q41 [Classification]

Do exercise 9, from DUDADS, chapter 21, using the wine.csv dataset.

Exercise 9: Build boosted models for the Wine dataset.

```
set.seed(144) #Setting seed for reproducibility

wine_df = read.csv("wine.csv")
wine_df = na.omit(wine_df)
wine_df$Class = as.factor(wine_df$Class) #Convert to a factor

#Let us assume that the response variable we would like to predict is Class

#Let us do a 80/20 split on the training/test data
split = sample.split(wine_df$Class, SplitRatio = 0.8)
training_set = wine_df[split, ]
test_set = wine_df[!split, ]

#Apply Gradient Boosting to Wine dataset

# Fit a boosting model to the training data
wine_boosting_model = gbm(Class ~ ., data = training_set, distribution = "gaussian")

wine_prediction = round(predict(wine_boosting_model, test_set))
```

```
## Using 100 trees...
```

```
wine_accuracy = sum(wine_prediction == test_set$Class) / nrow(test_set)
wine_accuracy
```

```
## [1] 0.8888889
```

Therefore, with the boosting model, we obtain an accuracy rate of 88.88% on the Wine dataset that predicts the class values.