Data Science IBM

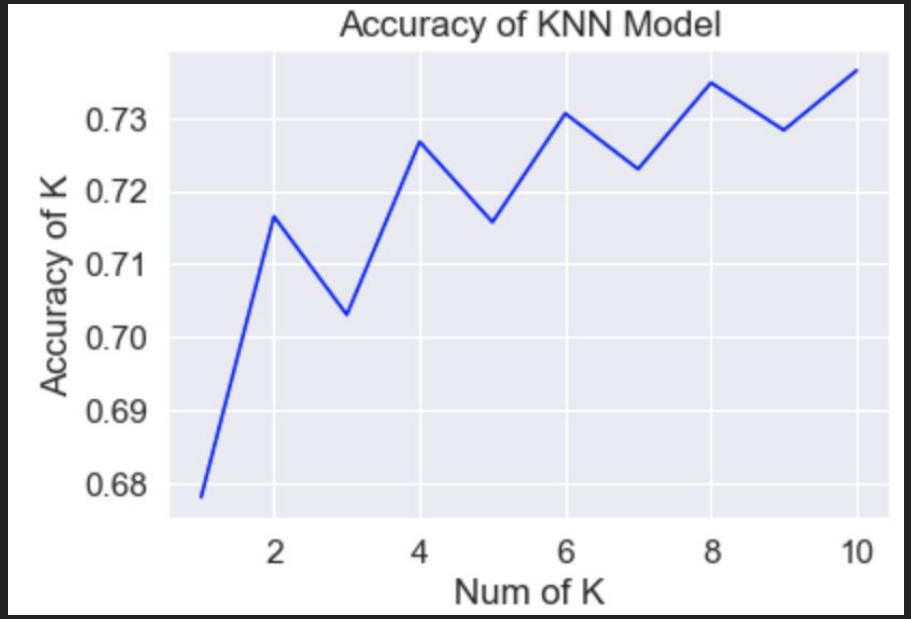
Leonardo Inza

Introduction/Business Problem

- The problem is that sometimes we are making a trip and in the middle of the road there was a collision and you have to wait two hours until all the accident area had taken care. So is there a way to predict this accident?
- The main idea is to predict an accident by some data that include weather, road condition, location, severity of the accident, type of coallision, etc.

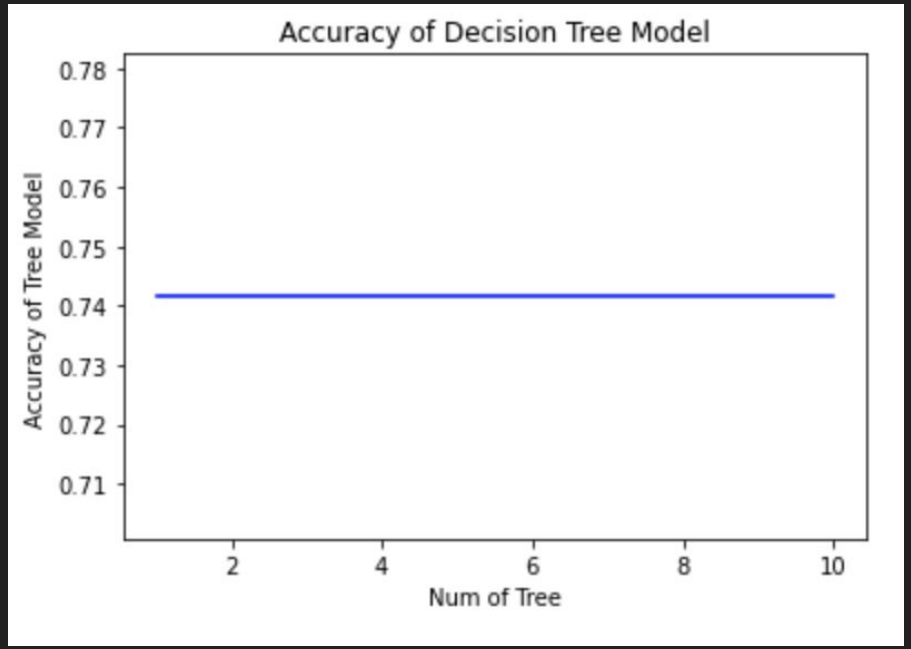
KNN Accuracy Plot

- After executing knn model wit k number from 1 to 10 we saw that best k number was 10
- We can see a graph of the accuracy values with test data of knn model with k number from 1 to 10
- While bigger is k, better is the accuracy



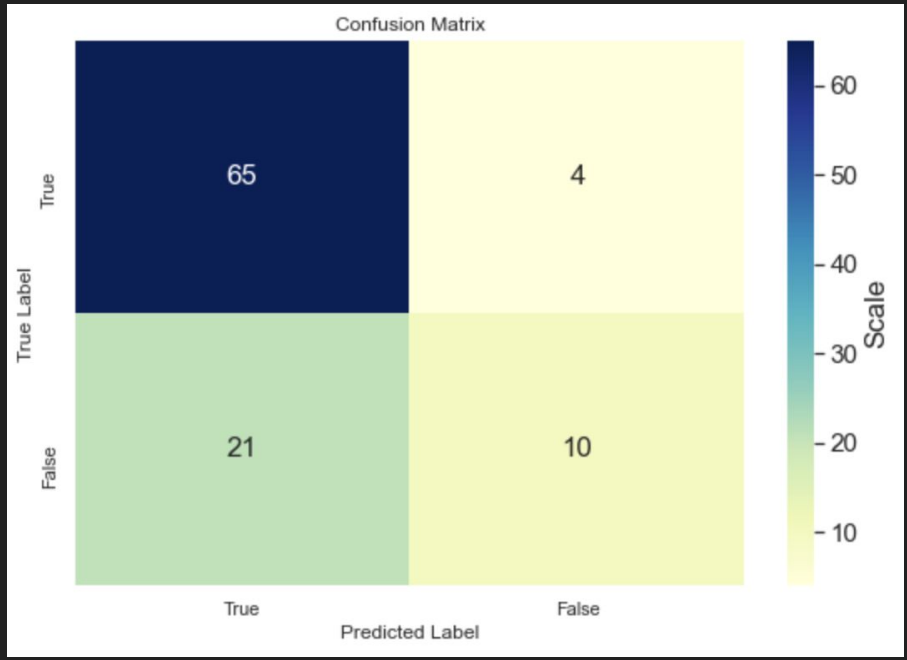
Decision Tree Accuracy Plot

- After executing the tree model with number from 1 to 10 we saw that best number of trees was all the same
- We can see a graph of the accuracy values with test data of desicion tree model that are all the same so we can choose any value from 1 to 10
- The value of accuracy is: 0.7416



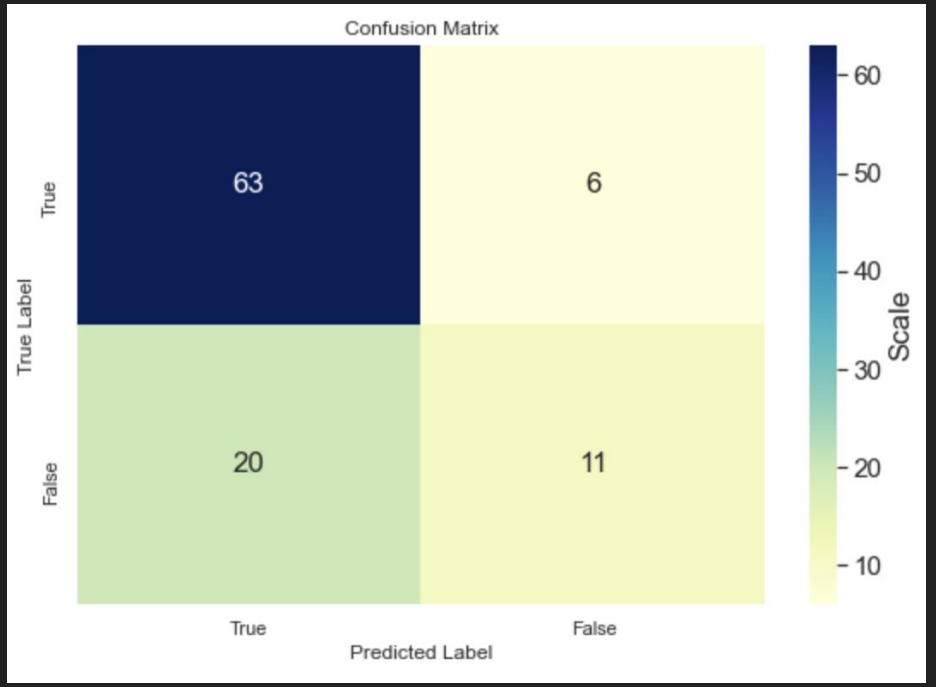
ANN Confusion Matrix

- In the plot we can see the percentage of all possible options in the confusion matrix
  - if we sum 65% with 10% we get 75% that the model predict one value and the reality was correct
  - So we can say that our model have an accuracy of it
  - If we sum 21% with 4% is 25% that the model predict one value and the reality was the other values.



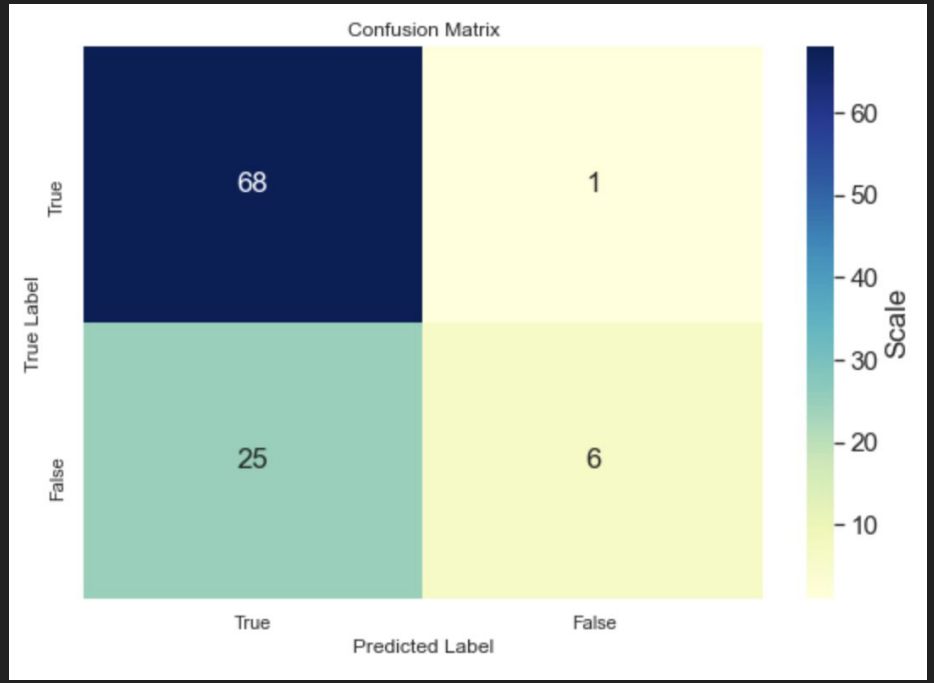
KNN Confusion Matrix

- In the plot we can see the percentage of all possible options in the confusion matrix
  - if we sum 63% with 11% we get 74% that the model predict one value and the reality was correct
  - So we can say that our model have an accuracy of it
  - If we sum 20% with 6% is 26% that the model predict one value and the reality was the other values.



Decision Tree Confusion Matrix

- In the plot we can see the percentage of all possible options in the confusion matrix
  - if we sum 68% with 6% we get 74% that the model predict one value and the reality was correct
  - So we can say that our model have an accuracy of it
  - If we sum 25% with 1% is 26% that the model predict one value and the reality was the other values.



Results with Test data

Model	Accuracy
ANN	0.751
KNN (k = 10)	0.737
SVM (num tree = 1 to 10)	0.7416