

Classification Exercise

Yegin Genc

Question

1 - Using the historical data (ClassifyRisk_historical.csv) create a knn classification algorithm, calculate and report the accuracy of the values. Only use the numerical values in your model first.

2 - Update your model to include categorical values.

3 - Using your model predict the risk type for the applications in 'classifyrisk.csv'. Note: Keep in mind that `classifyrisk.csv` does not have 'labels'. You will have to train/test the model on the historical data. You can select the optimal k value given the historical data. Then you can treat the entire historical data as your training data and the classifyrisk data as your test dataset.

You will upload your .rmd file and the updated classifyrisk.csv file with risk column populated with your predictions.