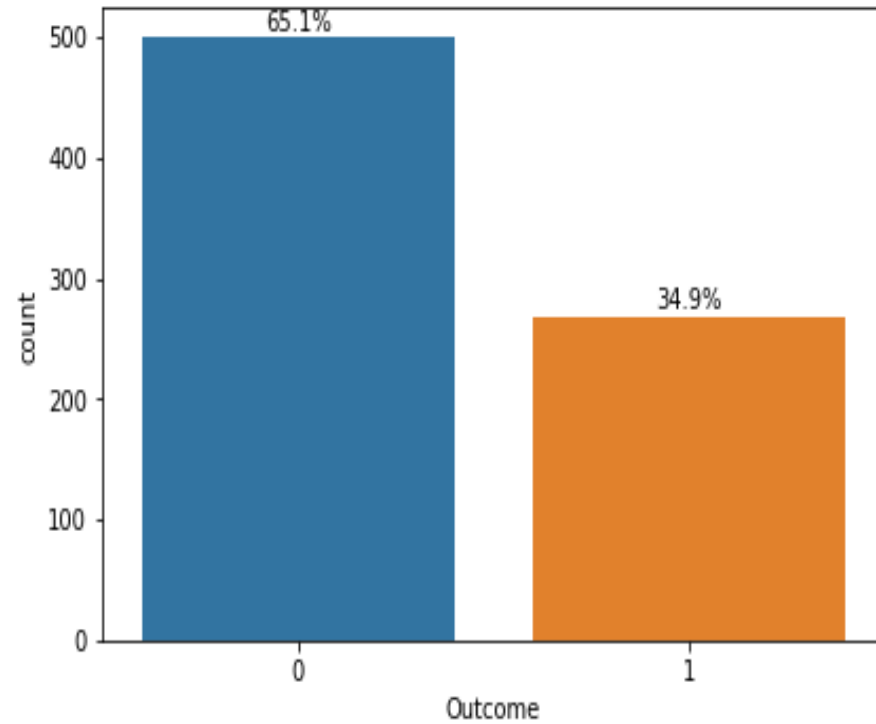# CC4051 Machine Learning

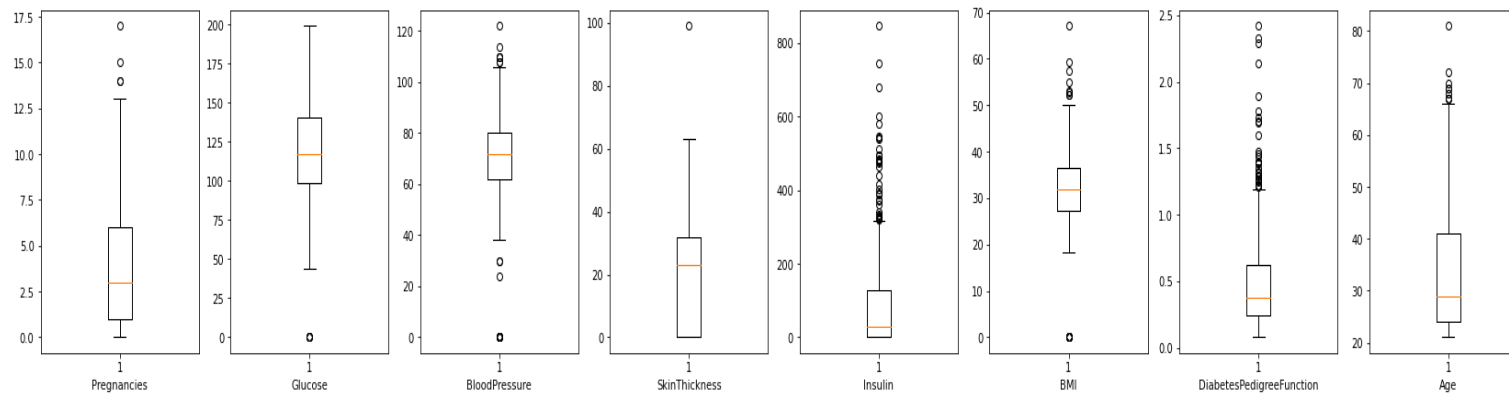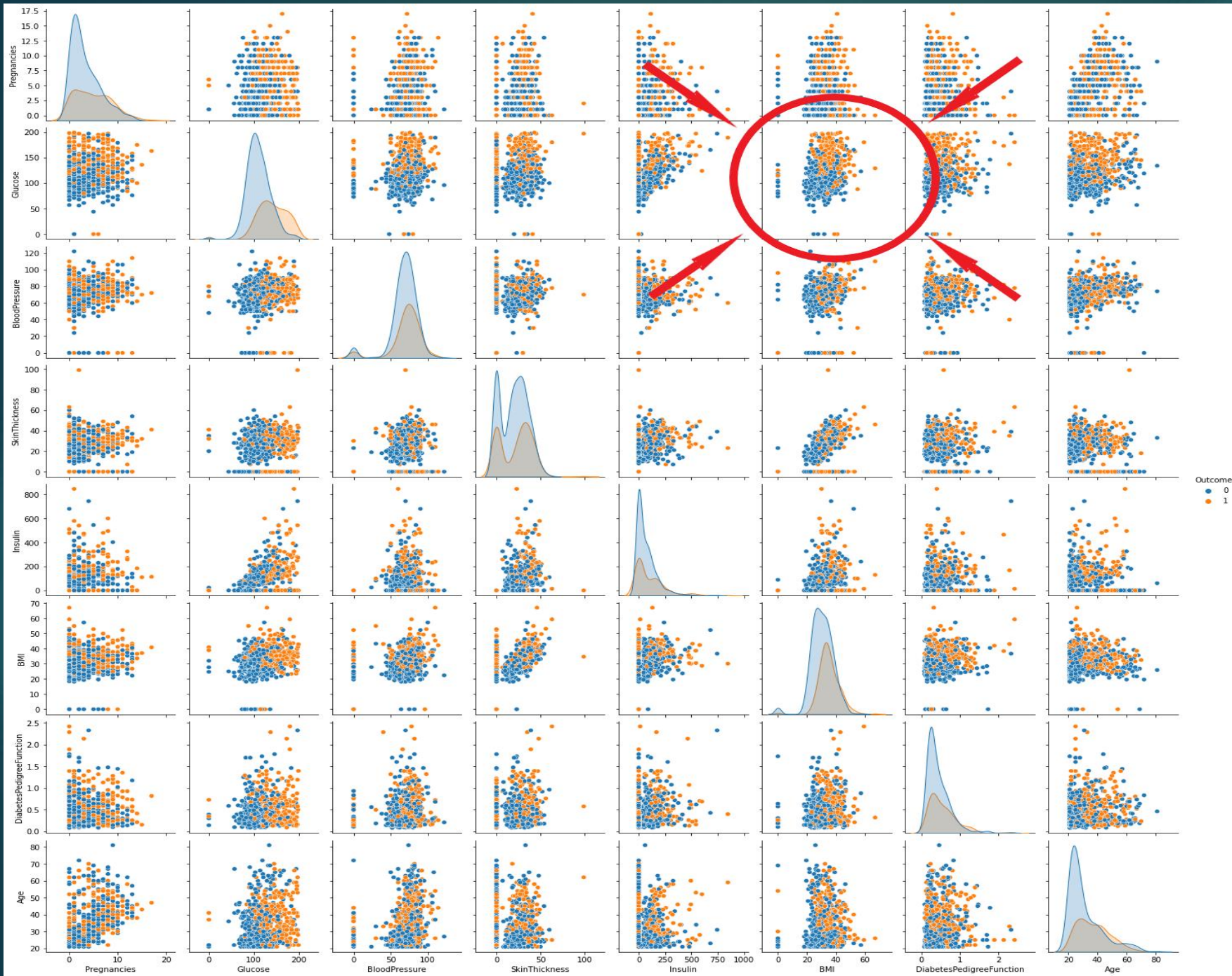ASSIGNMENT1 2020/2021

UP201703990- ÂNGELO GOMES

# Question 1: dataset Pima
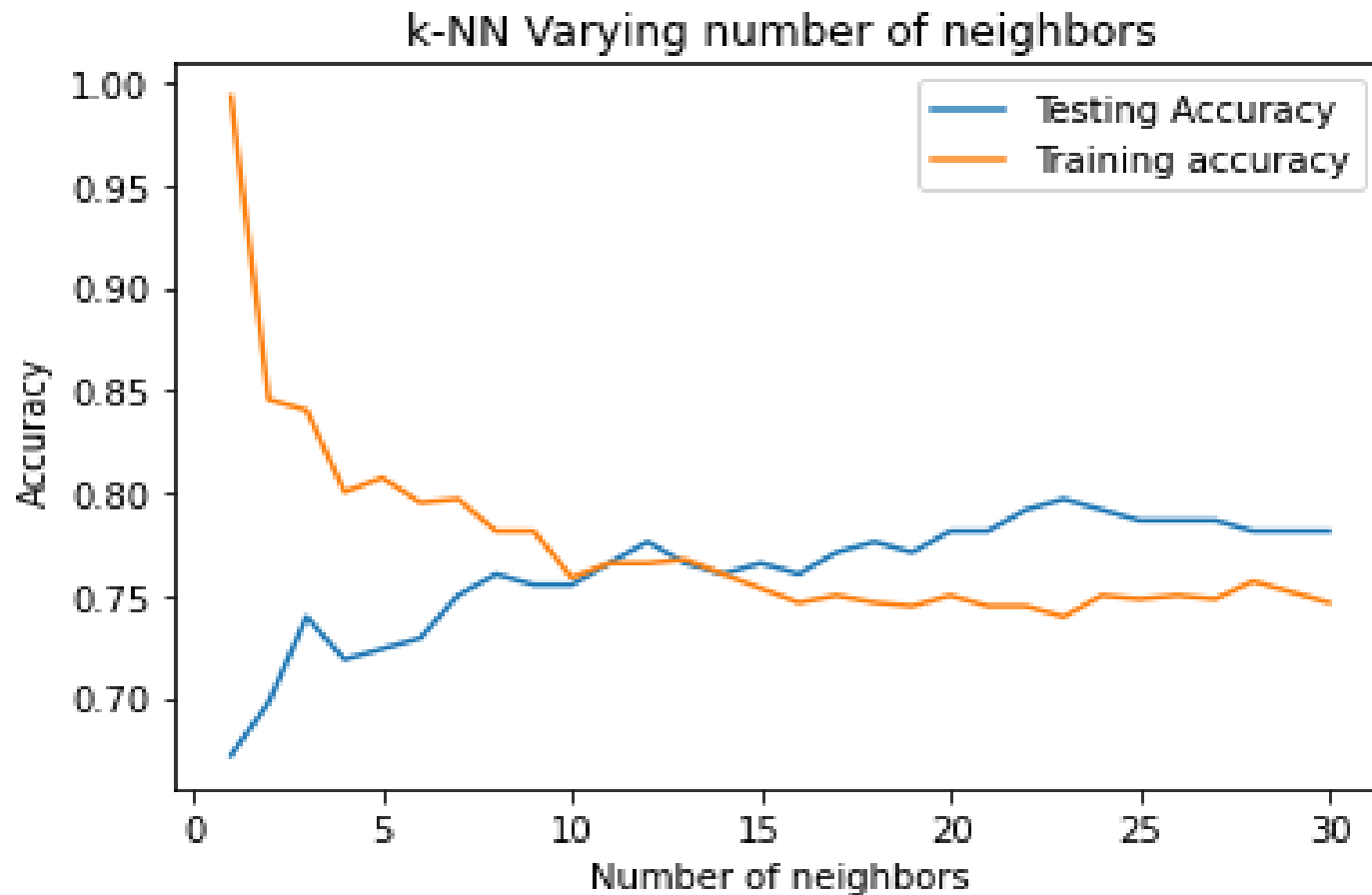
METHODS USED:

- KNN
- LOGISTIC REGRESSION
- QUADRATIC DISCRIMINANT ANALYSIS

Two attributes used to separate classes

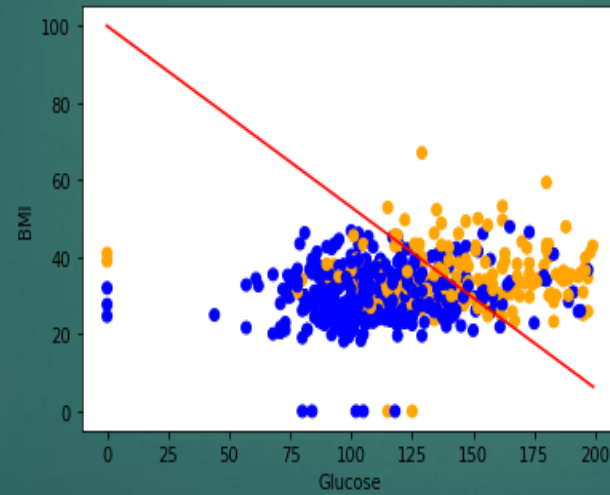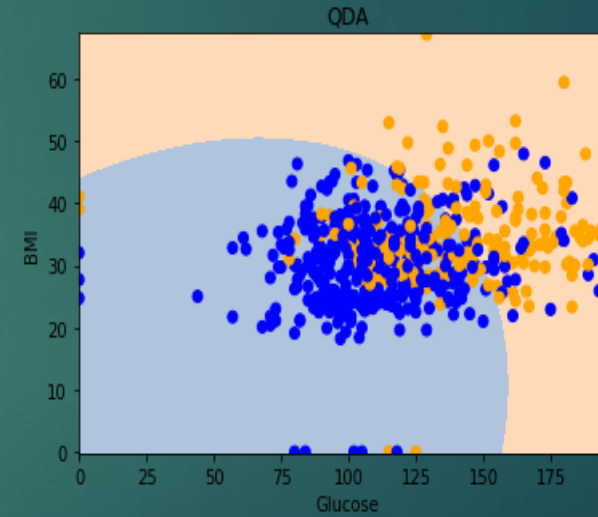# KNN method accuracy for various K values

# Boundaries

# Testing results

## KNN

```
KNN classification
              precision    recall  f1-score   support

           0       0.82      0.89      0.85       125
           1       0.75      0.63      0.68        67

    accuracy                           0.80       192
   macro avg       0.78      0.76      0.77       192
weighted avg       0.79      0.80      0.79       192

****CONFUSION MATRIX****
 [[111  14]
 [ 25  42]]

Accuracy is:  0.796875
```

## LOGISTIC REGRESSION

```
Logistic Regression classification
              precision    recall  f1-score   support

           0       0.80      0.89      0.84       125
           1       0.74      0.58      0.65        67

    accuracy                           0.78       192
   macro avg       0.77      0.74      0.75       192
weighted avg       0.78      0.78      0.77       192

****CONFUSION MATRIX****
 [[111  14]
 [ 28  39]]

Accuracy is:  0.78125
```

## QUADRATIC DISCRIMINANT ANALYSES

```
Quadratic Discriminant Analysis classification
              precision    recall  f1-score   support

           0       0.80      0.90      0.85       125
           1       0.75      0.58      0.66        67

    accuracy                           0.79       192
   macro avg       0.78      0.74      0.75       192
weighted avg       0.78      0.79      0.78       192

****CONFUSION MATRIX****
 [[112  28]
 [ 13  39]]

accuracy is:  0.7864583333333334
```
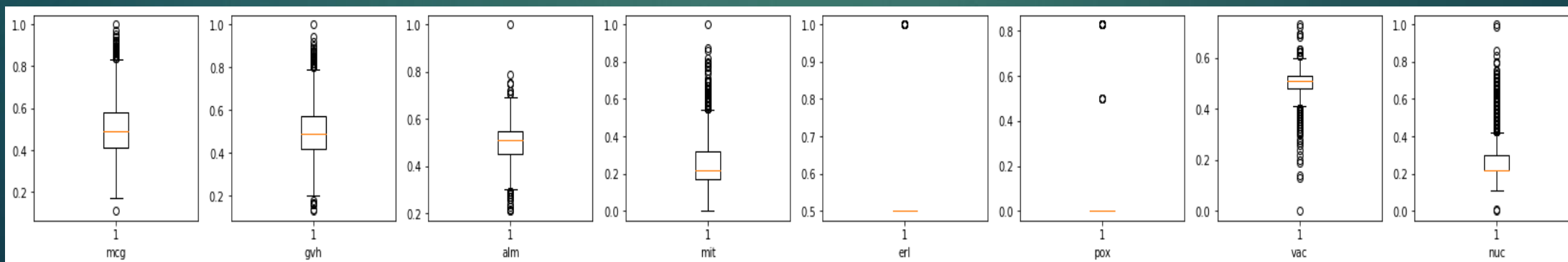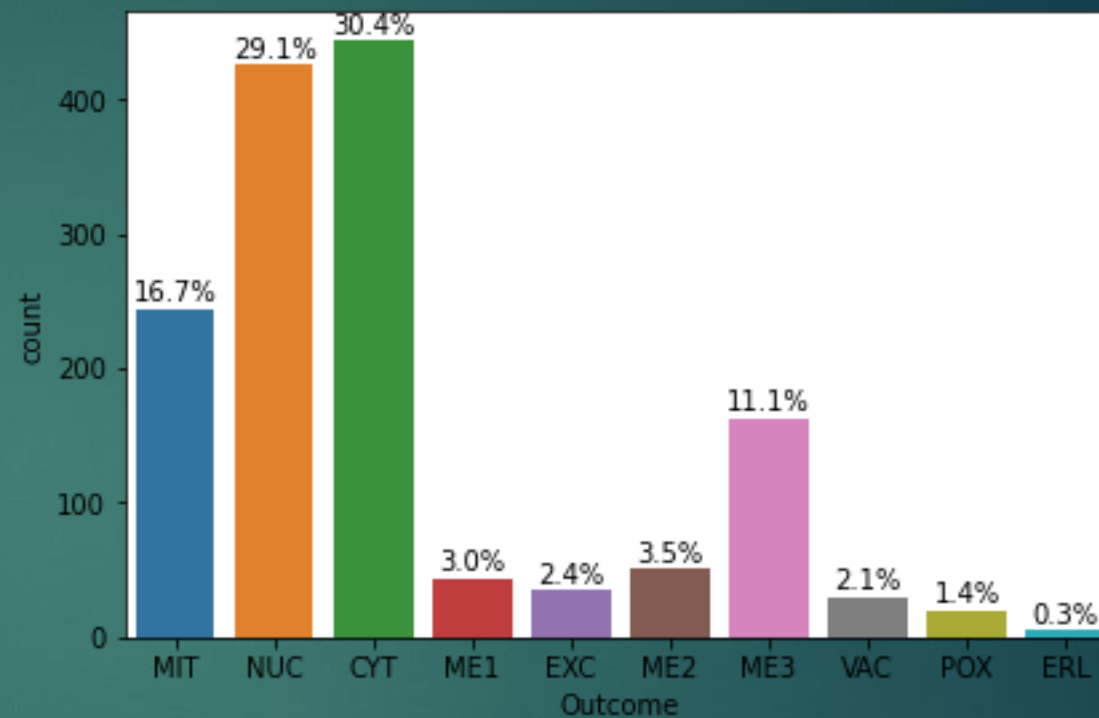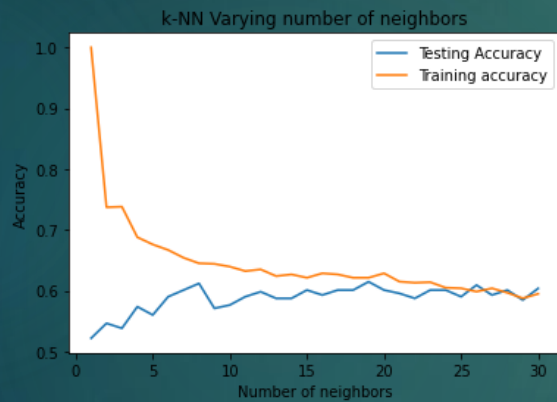
# Question 2: dataset Yeast

METHODS USED:
- KNN
- LOGISTIC REGRESSION
- DECISION TREE

# Results