# Intro to Artificial Intelligence
# Assignment 4

August 18, 2021
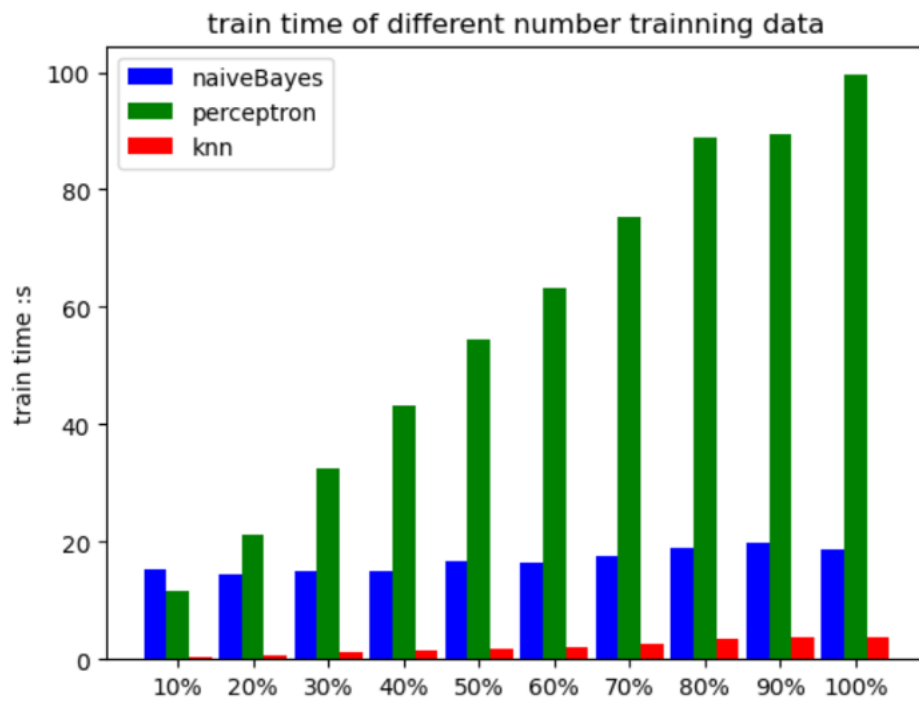
**Zesheng Zhang zz354**
**Zongqi Zhong zz443**
**Jing Du jd1340**

# 1 Report

Use the following the commands:

python dataClassifier.py -d faces

python dataClassifier.py -d digits

The following figure shows, using the digits dataset, the variation of training
time for the three algorithms as the training data changes

The following figure shows, using the digits dataset, the change in the accuracy
of the three algorithms as the training data changes



accurracy of different number trainning data
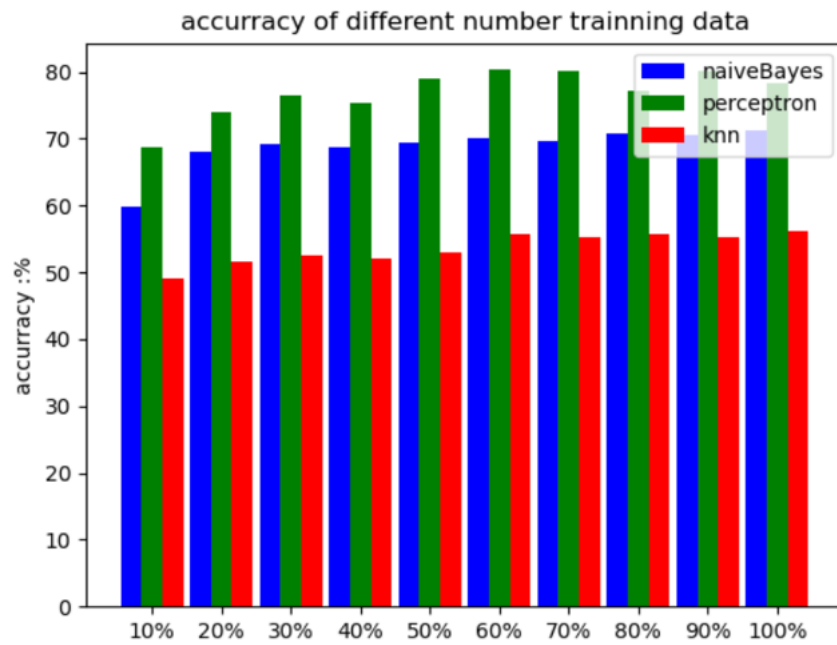
The following figure shows the variation of the training time of the three algorithms with the variation of the training data, using the faces dataset



train time of different number trainning data
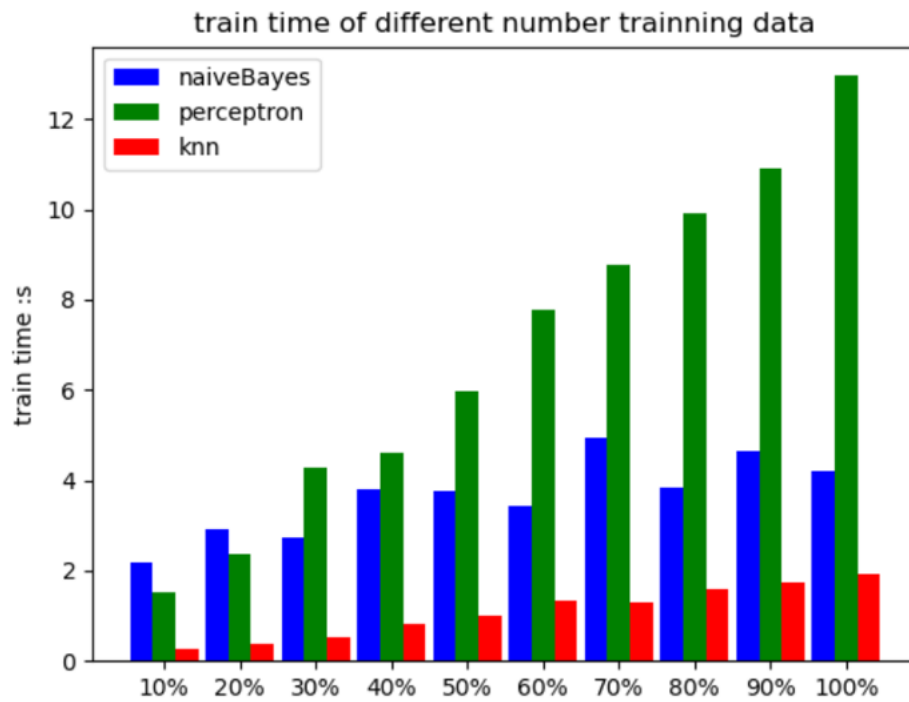
The following figure shows the change in the accuracy of the three algorithms as the training data changes using the Faces dataset
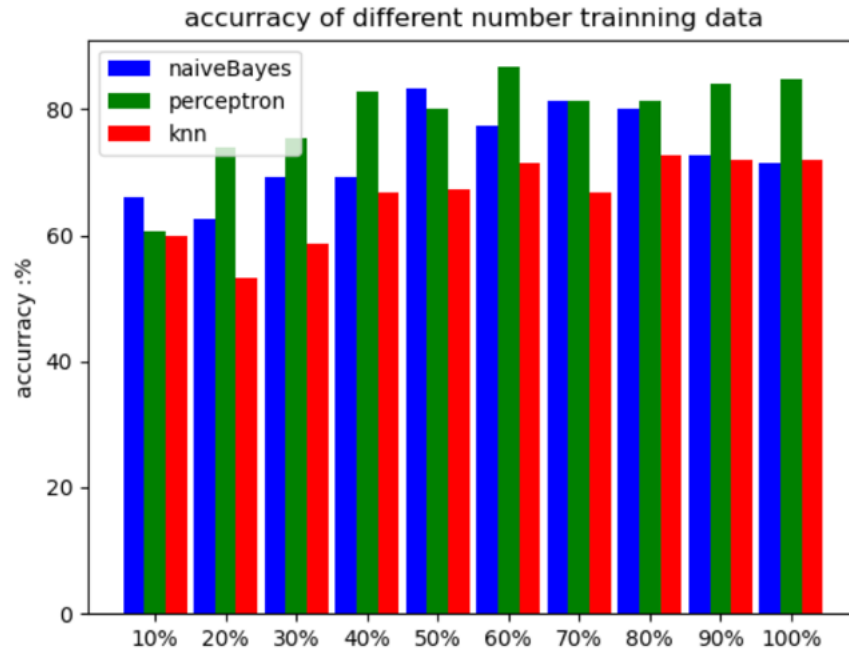


Analysis:

Overall, as the number of training data increases, the accuracy of each method increases, and the classification time increases

The difference is viewable when comparing the accuracy and training time of three algorithms.

Accuracy: perceptron > naiveBayes > knn

Training time: perceptron > naiveBayes > knn

Description of Knn algorithm: The algorithm used is k-nearest neighbor, first calculate the center of each label, then use the dot product method to compare the similarity between the test sample and the center of each type of label, and select the label with the greatest similarity as the result