Wargen Guittap 5004493060
Austin Merrick 2001003855

Assignment #4 Report

- The dataset you used, its source and characteristics.

Data source:
https://archive.ics.uci.edu/ml/datasets/banknote+authentication

We use the follow as features:
  1. variance of Wavelet Transformed image (continuous)
  2. skewness of Wavelet Transformed image (continuous)
  3. curtosis of Wavelet Transformed image (continuous)
  4. entropy of image (continuous)

To predict:
  banknote is genuine or forged

- Relevant evaluation metrics for the test dataset

Where K = 5

Accuracy = 1.0

Where K = 21

Accuracy = 0.9927007299270073

Where K = 101

Accuracy = 0.9854014598540146

- Comparison of performance metrics

For our dataset, it looks like K-nearest neighbors return a better accuracy than logistic regression with gradient descent. The higher the K value the least accurate the result.