

Answer Sheet

Connor Deakin

MAT 345

Project 2

- Data Set -

A

I used python for this entire project. I started by making one python script that takes every fourth email file and moves it into a testing directory. This script is called `seperate_testing_data.py`. After seperating the testing data, I made another script that performs all of the calculations needed for the assignment. This is script is `calculate.py`.

- Hammiest Words -

spambayes

users

razor

zzzzteana

sadev

- Spammiest Words -

rates

adv

money

insurance

systemworks

Accuracy: 0.8933333333333333

Precision: 0.9302325581395349

Recall: 0.32