

Project Information:

This project calculates the probability of a file being spam based on various calculations. These probabilities are then compared to actual Spam and Ham files, which allows us to calculate an accuracy and a precision for this project.

A screenshot of the project running can be found below:

Assignment 1 - Solution

Spam Dectetor Statistics

File Name	Actual Class	Spam Probabilities
00001.1a31cc283af0060967a233d26548a6ce	Ham	0.0
00002.5a587ae61666c5aa097c8e866aedcc59	Ham	0.0
00003.19be8acd739ad589cd00d8425bac7115	Ham	0.0
00004.b2ed6c3c62bbdfab7683d60e214d1445	Ham	0.0
00005.07b9d4aa9e6c596440295a5170111392	Ham	0.0
00006.654c4ec7c059531accf388a807064363	Ham	0.0
00007.2e086b13730b68a21ee715db145522b9	Ham	0.0
00008.6b73027e1e56131377941ff1db17ff12	Ham	0.0
00009.13c349859b09264fa131872ed4fb6e4e	Ham	0.0
00010.d1b4dbbad797c5c0537c5a0670c373fd	Ham	0.0
00011.bc1aa4dca14300a8eec8b7658e568f29	Ham	0.0
00012.3c1ff7380f10a806321027fc0ad09560	Ham	0.0
00013.245fc5b9e5719b033d5d740c51af92e0	Ham	0.0
00014.8e21078a89bd9c57255d302f346551e8	Ham	0.0
00015.d5c8f360cf052b222819718165db24c6	Ham	0.0
00016.bc1f434b566619637a0de033cd3380d1	Ham	0.0
00017.8b965080dffada165a54c041c27e33f	Ham	0.0
00018.3b6a8c5da4043f2a6a63a1ae12bd9824	Ham	0.0
00019.c6b272a04ec32252f7c685f464ae3942	Ham	0.0
00020.83ef024f76cc42b8245a683ed9b38406	Ham	0.0
00021.ba795c59691c8f5d8a02425fdd9bf0ea	Ham	0.0
00022.b7c5c97a3a140eed207b9e90d4e650a1	Ham	0.0
00023.0e033ed93f68fcb5aab26cbf511caf0e	Ham	0.0
00024.066b89ecd18c7688e91833f97cf415ca	Ham	0.0
00025.84faba510a966c90f6ca7658260a7e4c	Ham	0.0
00026.1757d50d495d41e8a5eb30a2f371019c	Ham	0.0
00027.c9e76a75d21f9221d65d4d577a2cfb75	Ham	0.0
00028.4e9595edd918f1a5fa26f8740cfdb358	Ham	0.0
00029.807838f09bfb11b71e179a75334a5a62	Ham	0.0
00030.cc523265aefc37ee6ce3015d8ff6aa24	Ham	0.0
00031.7caef7fe7af2114d0e4bf6aa0faf3a03	Ham	0.0
00032.75f27327d5f41f09e0b2160c62097643	Ham	0.0

Accuracy:

0.47501665

Percision:

0.82051283

Display Stats

The improvements that I made to the interface are:

1. I added a button which when clicked will call the `getAccuracy()` and `getPrecision()` commands and display them to the Accuracy and Precision text fields on the interface.

To Successfully run this process:

1. Open a command prompt window.
2. Navigate to a GIT monitored repository
3. Clone the Assignment 1 Repository by using the `git clone` command
4. Open IntelliJ or any other IDE
5. Navigate to the file that you just cloned.
6. Run the `Main.java` file.
7. When the file window appears, please navigate to the directory which contains two directories `test` and `train` in them and click the 'select folder' button.
8. Once this is completed a Window will appear with a table of statistics
9. To view accuracy and precision click the Display Stats button.

Github Repo:

<https://github.com/NathanCabral/csci2020u> NathanCabral