Project by: Utsav Dwivedi and David Watt

This project is an implementation of the naive bayes spam filtering algorithm, a form of machine learning. It uses calculations to assign a probability to a bag of words that estimates whether each individual word is either spam (“unwanted”) or ham (“wanted”). Some improvements we made to the model were some aesthetic improvements, we made it look aesthetically pleasing. To run this program you will need an IDE that can run Java files such as Intellij. You also need to clone the github. To clone the github, you will first need to download git, there are many online tutorials for that. You then open the command prompt and type git clone “github link.” You can then run our main class to see the program

.References

<https://en.wikipedia.org/wiki/Bag-of-words_model>

<https://en.wikipedia.org/wiki/Naive_Bayes_spam_filtering>

<https://github.com/>



Github link: <https://github.com/Utsav-Dwivedi/csci2020u_utsavdwivedi>

Improvements: Some improvements were made to the UI. These were done using CSS to change the colours of the table and texts allowing for friendlier UI. Also, a button was added to the UI to implement the directory chooser.

How to run: Download the zip and extract it. Open the project in intellij and proceed to click run in the Main.java class. This will launch the UI and the above screen will show. Proceed to click on the 'Test' button to select the directory which contains the ham and spam files. This is where the program will cease running.