README

**Project name**: Naïve Bayes and Logistic Regression Implementation

**Current Version**: 1.0

**Date:** 3/16/2016

**Author**: Chandni Shankar

**Description:**

Implemented the Naïve Bayes classifier and logistic regression classifier to classify a list of documents as ham / spam.

1. Naïve Bayes: Implemented the multinomial Naïve Bayes classifier to calculate the log likelihood and find the maximum log likelihood of data to predict the label for each file
2. Logistic Regression: Implemented the batch gradient ascent algorithm to calculate the maximum condition a-posterior and predicted the label for each file.

**Input**: Training Data file, Test File and stop words file.

**Output**: Accuracy of Naïve Bayes classifier and Logistic regression with and without using the given stop words.

**Files**: NaiveBayesLR.py

**Programming Language used**: python

**Instructions to run**: python NaiveBayesLR.py <path of training file> <path of test file> <path of stop words file>

Example command: python NaiveBayesLR.py C:\Users\chand\Desktop\Machine\_Learning\Assignement2\hw2\_train C:\Users\chand\Desktop\Machine\_Learning\Assignement2\hw2\_test C:\Users\chand\Desktop\Machine\_Learning\Assignement2\cxs152830\_Assignment2\stopWords.txt