What to submit

Submit a .zip package with the name "YourLastName_YourFirstName_HW4.zip", the package contains the following:

1. **(NaiveBayes.jar, NaiveBayes.exe or NaiveBayes.py)** for your implementation, these files should be runnable from command line. The files **take only one argument which is the Data "spambase.data"** processed version of the dataset. Your code should do the data splitting, training, validation and score reporting. The output of your program should be printed on the console in the following format.

   Line1: Fold_1, false positive, false negative, overall error rates for Fold 1

   Line2: Fold_2, false positive, false negative, overall error rates for Fold 2

   Line3: Fold_3, false positive, false negative, overall error rates for Fold 3

   Line4: Fold_4, false positive, false negative, overall error rates for Fold 4

   Line5: Fold_5, false positive, false negative, overall error rates for Fold 5

   Line6: Avg, avg false positive, avg false negative, avg overall error rates for all folds

2. In case you implemented your code in java or C#, submit also the code **(NaiveBayes.java or NaiveBayes.cs)** alongside the executable files.

3. **A Readme.txt** file, in which mention:
   a. The version of Python you used (if python used), framework used in case of C# or Java.
   b. Any additional libraries needed to run your code.
   c. Any part of your code that is not working.

4. A **Report.pdf or Report.doc** file, in which you report:
   a. **Your scores** in the format specified in the homework document.

   b. Provide **some statistical analysis for the folds** showing the ratio of positive samples and the ratio of negative samples in train data and dev data for each iteration, and how does that affect the scores you get over that iteration.

   c. **Compare** your results with the result of **just choosing the majority class**, provide some analysis of the results you get and summarize your intuition behind the scores you get (**Would it make sense if choosing majority class gets higher results? and if so why do you think that happened?, the same goes for if you got higher scores using naive bayes. Does it make sense and why?)**