Assignment 4 Naïve Bayes

COMP4901K and MATH 4824B

Fall 2020

Prerequisites

- You need to have some background knowledge about Naïve Bayes (NB). If not, you can check out: Lecture 7 and `https://en.wikipedia.org/wiki/Naive_Bayes_classifier`
- You need to install the NLTK, Pandas, Numpy, Scipy, and scikit-learn packages
- You need to download the assignment4.zip from the canvas

1 Tasks

T1 Fill in the compute_prior function to calculate the prior array (i.e., $P(y)$ in the tutorial).

T2 Fill in the compute_likelihood function to calculate the likelihood matrix (i.e., $P(x|y)$ in the tutorial).

T3 Fill in the predict_NB function to compute a prediction array given a BOW matrix, the prior array in T1, and the likelihood matrix in T2.

2 Submission

You need to submit **two** files, program output and your python script. After you finish the assignment, make sure you include your name and student ID in the beginning of your code.

author : Your_name

student_id : Your_student_ID

Copy all the program output to a text file named StudentID_assignment4_output.txt. Rename your python script solution as StudentID_assignment4.py. Zip them altogether and submit it to Canvas.

Similar with the previous assignments, there's no strict restrictions on the output formats. You could define your format.