CS561/571 - Executive Assignment

ASSIGNMENT-5: Naive Bayes

Date: October 07, 2023

Deadline: October 22, 2023

Total Credit: 30

- Markings will be based on the correctness and soundness of the outputs.
- Marks will be deducted in case of plagiarism.
- Proper indentation and appropriate comments are mandatory.
- *All code needs to be submitted in '.py' format.* Even if you code it in '.IPYNB' format, download it in '.py' format and then submit
- You should zip all the required files and name the zip file as:
    - <roll_no>_assignment_<#>.zip, eg. 1501cs11_assignment_01.zip.
- Upload your assignment (the zip file) in the following link:
    - https://www.dropbox.com/request/STJmj7CVJEvbPEcZlkFN
- **Note: Do not send your zip files to us in email**

Problem Statement: Given a dataset containing details about various details about the details of employees such as 'workclass', 'education', 'occupation' etc. The task is to predict the 'salary' of the employee given the other details. The salary field contains two values '<= 50k' and '>50k'.

Dataset:

- **Train set:** Link
- **Test set:** Link

Implementation Details:

- Implement two naive bayes classifiers:
    - Gaussian Naive Bayes
    - Multinomial Naive Bayes
- You can use existing packages such as Scikit-Learn to create the classifier
- Report accuracy on the test set for both the classifiers

Documents to submit:

- Model code
- Detailed document describing results such as time taken for the execution and accuray results

For any queries regarding this assignment, contact:

Ramakrishna Appicharla (ramakrishnaappicharla@gmail.com) and,
Ratnesh Kr. Joshi (ratneshkr.joshi@gmail.com)