CSE 4238 - Soft Computing Lab
Assignment # 02 (NLP)

Course Instructor: Nibir Chandra Mandal

There are two parts in this assignment-
- I.   Coding Part
- II.   Report Writing Part

Here, each part contains individual marks.

Problem Definition:

You have been assigned to solve a text classification problem using the model described in section 2. Find the accuracy, precision, recall, and F1 score for your model. In addition, report the error curve against epochs.

I.   Coding Part

1) Dataset

Find the dataset of your task from the following link Dataset. You will get a CSV file. Keep 20% data for testing. You can also keep 20% of data for validation *[Optional]*. During the splitting, set the seeding value as your StudentID.

2) Model Creation

Create the following models described in the table using torch or TensorFlow. Choice of your hyperparameters (e.g., number of hidden layers, learning rate, activation function, dropout layers, etc) should be logical. You should tune your hyperparameter using the validation dataset. ***DO NOT USE THE SAME HYPERPARAMETERS AS YOUR FRIENDS.***

Pre-Processing Steps	Feature Selection	Model
i) Stop word removal ii) Stemming	BoW	Neural Network
	TF-IDF	Neural Network
	One-Hot Embedding	CNN

II.   Report Writing Part

    01. Add a cover page.

    02. Make necessary tables and graphs/charts to show your result and discuss them.

    03. Compare the results of all three models using proper diagrams.

    04. Upload your code in Github and share the link in your report.

Notes:

1. The report doesn't have any page limit.
2. Use any formatting tool latex/word/google doc to prepare your report.

Additional Guidelines:

★ Use Google Colaboratory to complete the coding part.

★ Use the Pytorch/Tensorflow/Keras library for implementation.

★ Write the report by following the instructions of the Report Writing Part.

★ Submit the report in pdf format using the naming convention as follows studentID.pdf. *Example: 170104000.pdf*

★ Submit your codes in ipynb format using the same naming convention stated above.

★ Keep patience if it takes a longer training time.

★ **All of your assigned tasks must be unique. Plagiarism is strictly prohibited and punishable.**