Notes

1. You may use any libraries, but you should demonstrate an understanding of the machine learning concepts behind them.
2. We may inspect your code, so make sure that your code is clean and well structured.
3. Please be concise when describing your work.
4. If necessary, please state your assumptions for solving the problem.

Problem Statement

NLP has been deployed in many domains to make our life better. So much so that it is easy to overlook their potential to benefit society by promoting equity, diversity, and fairness. Nonetheless, as NLP systems become more human-like in their predictions, they can also perpetuate human biases. Previous work showed that sentiment analysis systems can perpetuate and accentuate inappropriate human biases, e.g., systems that consider utterances from one race or gender to be less positive, or customer support systems that prioritize a call from an angry male over a call from the equally angry female.

Currently, we are exploring research related to the bias in sentiment analysis task. To do this, an ability to understand and build a sentiment analysis model is important.

Given a tweet, **your task is to build a machine learning model to classify the tweet into a positive/negative class.** You need to build:

1. *Classical models using scikit-learn.* Please pick 5 models (Naive Bayes, SVM, kNN, Logistic Regression, Random Forest, Extra Tree Classifier, etc). Please use text/tweet preprocessing techniques, such as lemmatization, stemming, etc.
2. *Deep learning models using PyTorch/TensorFlow*. Please implement 2 models:
    - A neural network model (such as CNN, RNN, LSTM) that utilizes word embedding.
    - A BERT-like model, such as BERT-BASE, ALBERT, XLNET, etc.

Important Notes: Model performance (e.g. accuracy, F1) is not important in this task. This task is addressed to know how fast you can learn a new material and implement it using available resources. Don't spend too much time improving the model performance.

Dataset

The given dataset (SemEval18 Task 1) consists of a set of English tweets for training, development/validation, and testing. Each file has 4 columns: ID, Tweet, Affect Dimension, and Intensity Classes. Tweet and Intensity Classes are the only relevant columns in our task. There are seven Intensity Classes (corresponding to various levels of positive and negative sentiment intensity that best represents the mental state of the tweeter):
3: very positive mental state can be inferred
2: moderately positive mental state can be inferred
1: slightly positive mental state can be inferred
0: neutral or mixed mental state can be inferred

-1: slightly negative mental state can be inferred
-2: moderately negative mental state can be inferred
-3: very negative mental state can be inferred

Expected Submission

We expect you to submit your source code and a brief write up of your solution. If you prefer, you can do both in an IPython (or any other) notebook.

- Please submit all source code you have written. This includes any simple script you used to explore the dataset.
- Please include a README file describing how to run your code.
- Please do not include any executable files, libraries, or output from the build process.
- Please provide a detailed step of your machine learning pipeline in the write up (including preprocessing, analysis, results, graphs, or anything else you think is necessary to demonstrate your understanding).
- Please package everything in a zip file and send it to us. You can also provide us a GitHub repository link containing the source code and the write up.