README

Project Name: *Bong Sentigem: Sentiment Analysis in Bengali*

Team members

Ananya Bothra (UIN: 830000041)
Esha Dutta (UIN: 529005704)
Projna Paromita (UIN: 427007589)
Sourjya Banerjee (UIN: 729007624)

What is included in the zip file

1. Codes
2. Data - inside the ‘data’ directory

Package requirements

python	3.7.9
matplotlib	3.3.2
numpy	1.19.1
pandas	1.1.2
scikit-learn	0.23.2
scikit-optimize	0.8.1
scipy	1.4.1
tensorflow	2.1.0
xgboost	1.3.3

The running instruction for the codes

1. ML_models.py - run the following instruction: “*python ML_models.py*”

   This code includes the following models for sentiment analysis:
   a. Naive Bayes

b. Support Vector Machine (SVM)
c. Random Forest (RF)
d. Decision Tree (DT)
e. XG boost (XGB)
f. K-nearest neighbor (KNN)

The output results of this file is provided in the 'ml_result.txt'.

2. RNN.ipynb - This is a jupyter notebook file that can be run using jupyter lab or google colab. Run the notebook as given.

This code includes the model implementing the Long Short-Term Memory (LSTM) model for sentiment analysis.

3. SNN.ipynb - This is a jupyter notebook file that can be run using jupyter lab or google colab. Run the notebook as given. This code saves the processed data and model. If someone intends to use the saved files, just change the following variables from 0 to 1:

   read_from_store_data = 0
   read_from_stored_model = 0

This code includes the model implementing the Siames Neural Network for sentiment analysis.

Both the .ipynb files have the outputs from the run at the output text.

****All these codes, along with the saved pre-processed data and models could be found in the following github link:*

https://github.com/eshadutta/NLP_Bengali_Sentiment_Analysis