

Assignment 3B

Natural Language Processing (CS563)

Department of CSE, IIT Patna

(Read all the instructions carefully and adhere to them.)

Date: 3rd-April-2021

Deadline:- 30th-April-2021

Instructions:

1. Markings will be based on the correctness and soundness of the outputs.
2. Marks will be deducted in case of plagiarism.
3. Proper indentation and appropriate comments (if necessary) are mandatory.
4. You should zip all the required files and name the zip file as:
<roll_no>_assignment_<#>.zip, eg. 1501cs11_assignment_01.zip.
5. Upload your assignment (the zip file) in the following link:

<https://www.dropbox.com/request/WMtFjtWFir417V1U2iFc>

For any queries regarding this assignment contact:

Dushyant Singh Chauhan (dushyantchauhan27@gmail.com) and

Kshitij Mishra (kmishra.kings@gmail.com)

Sentiment Analysis:

Sentiment analysis is an important task in natural language processing and has a wide range of real-world applications. The typical sentiment analysis focuses on predicting the positive or negative polarity of the given sentence(s). This task works in the setting that the given text has only one aspect and polarity. A more general and complicated task would be to predict the aspects mentioned in a sentence and the sentiments associated with each one of them.

As per the above statement, sentiment analysis can be performed in two ways:

- A. Sentence-level Sentiment Analysis
- B. Aspect-based Sentiment Analysis

(B) Aspect-based Sentiment Analysis

Aspect-based sentiment analysis (ABSA) is a text analysis technique that categorizes data by aspect and identifies the sentiment attributed to each one. Aspect-based sentiment analysis can be used to analyze customer feedback by associating specific sentiments with different aspects of a product or service.

- **Problem Statement:** For a given message (review) analyze customer's feedback by classifying specific sentiments with different aspects of a product or service (wrt. RESTAURANT).
 - **Input:** Given Sentence
 - **Output:** Aspect based sentiments of the sentence (positive, negative, or neutral).
- **Embedding vectors**
 - Use Fasttext embedding for experiment and download the feature vector from [crawl-300d-2M.vec.zip](https://fasttext.cc/docs/en/english-vectors.html).
 - For more information and usage please go through this link (<https://fasttext.cc/docs/en/english-vectors.html>)
- **Methods:** LSTM
- **Dataset:** Download the dataset for aspect-based sentiment analysis from here <https://drive.google.com/drive/folders/18HDLabdejm9-GI1QW-Blk6fyWNmBU3-C?usp=sharing>
 - There will be two files *train.xml* and *test_with_gold.xml* in the dataset.
 - Train your model on *train.xml* and test your model on *test_with_gold.xml*.
- **Evaluation Metrics:** Evaluate your model based on the following metrics:
 - Accuracy
 - F-score
 - Precision and recall
 - Confusion matrix