

**Name:** Imtiag Hossin Joy, **Roll:** 1907092

**Paper1:** *Bangla Aspect-Based Sentiment Analysis Based on Corresponding Term Extraction*

ABSA (Aspect-Based Sentiment Analysis) is text analyzing method that isolates aspects present in text and find the sentiment relating to each aspect. The research has been done regarding two fields - Cricket, Restaurant and the proposed technique for term extraction is PSPWA (Priority Sentence Part Weight Assignment). In this research, only the aspect terms extraction has been done of ABSA model. A public imbalanced dataset was used for the research and several algorithms that were used. Punctuations, numerical words have been removed from dataset in preprocessing steps. TF-IDF (Term Frequency and Inverse Document Frequency) has been used for tokenization. Constants features and Correlated features are removed by threshold values 0.005 and 0.75 respectively. In ABSA model NOUNS are considered as priority word and part of sentence holding it is considered as priority sentence part and it is found out based on the length of a sentence. Weight assignment is done by considering some rules and Min-Max classifier is used for normalization. 80% of the dataset was used as training dataset and 20% of the dataset as a test dataset and 20% of the training data was used as validation data. F-1 score was used for performance measurement of the proposed model and CNN got the highest score of 0.59 and 0.67 for cricket and restaurant respectively. The performance of the proposed model is better than the previous works in this field.

**Paper2:** *Aspect Based Sentiment Analysis in Bangla Dataset Based on Aspect Term Extraction*

The main focus of this research is to perform ABSA's aspect extraction in Bengali language. Aspect refers to an attribute of an entity and ABSA has become an advanced tool that can analyze online reviews and predicts opinions on aspect-level. The four subtasks under ABSA are Aspect Term Extraction, Aspect Term Polarity, Aspect Category Detection and Aspect Category Polarity. The research has been done in two fields – Cricket, Restaurant where datasets of each field consist of five different aspects each. Removing the special characters and punctuation is done in data preprocessing to make the datasets less noisy. BOW (Bag of Words) is used for converting texts into numbers and TF-IDF (Term Frequency-Inverse Document Frequency) is used to measure the importance of a term in a document relative to a collection of documents. Trial and error process is used to determine the collection of algorithms used for the research. 80% of dataset was used for training and 20% for testing. The results of the research are better than the previous work and better results can be achieved if the datasets can be trained using POS tagging.

**Paper3:** *BAN-ABSA: An Aspect-Based Sentiment Analysis Dataset for Bengali and Its Baseline Evaluation*

ABSA task consists of two subtasks - aspect term extraction and sentiment classification. In this research, supervised learning approaches are used for both subtasks completion. In the dataset, there are four aspects - Politics, Sports, Religion, and Others and for every aspect respective ternary sentiment classification: Positive, Negative, and Neutral were used. All the comments are collected from the Facebook page of well-known Newspapers. Each comment was annotated by 3 annotators for both aspect and sentiment polarity and majority voting was followed for the annotation of the comment. ICC (Intra-class Correlation) was 0.77 which means there is a high level of agreement among the measurements. Comments containing English words, punctuation marks were removed in data pre-processing phase and then the dataset was tokenized. Among the deep neural networks that have been used for the research CNN achieved better accuracy than the others but BiLSTM's context catching capability was better for its bi-directional approach. As the prepared dataset is more enrich and contains more variation than other datasets of previous works all the models performed better in this dataset. BiLSTM is the proposed architecture to be used in ABSA and even though the results may not be as high as English ABSA, the overall performance can be improved by adding attention on attention in the architecture.