Paper Title: SENTIMENT CLASSIFICATION IN BANGLA TEXTUAL CONTENT: A COMPARATIVE STUDY

Paper Link: https://arxiv.org/ftp/arxiv/papers/2011/2011.10106.pdf

1 Summary:

1.1 Motivation:

This paper explores the field of sentiment analysis with a focus on low-resource languages, specifically Bangla which has a lack of enough datasets and well-defined train/test splits. The study pointed out that limited resources and inconsistent benchmarks have hindered sentiment analysis for Bangla. Sentiment Classification has been used significantly in Natural Language Processing,the authors of this study examined several publicly available sentiment-labeled datasets and created classifiers using both classical and deep learning algorithms. The classical algorithms used were SVM and RF, while deep learning algorithms included CNN, FastText, and Transformer-based models. The study compared these models in terms of their performance and time-resource complexity. Notably, the findings suggested that Transformer-based models, a relatively unexplored area for Bangla sentiment analysis, outperformed all other models. In addition to model comparisons, the authors created a weighted list of lexicon content based on valence scores per class and analyzed the content for highly significant entries per class in the datasets. One common approach in sentiment analysis has been to use sentiment lexicons as features for designing sentiment classifiers, and machine learning algorithms, including classical and deep learning approaches, have been employed to address the problem.
The results and resources provided in this paper can serve as a benchmark for future studies in Bangla sentiment analysis. The paper's motivation revolves around the feasibility of effectively conducting sentiment analysis in Bangla, despite resource limitations, and the potential advantages of utilizing transformer-based models in this context- by conducting a comparative study with publicly available datasets and providing valuable insights and resources by utilizing classical and deep learning algorithms which include training complexity, model performance, and implication in real deployment.

1.2 Contribution:

As in low resource languages like Bangla, there are not many available datasets there, and it's harder for the analyst to research on this because most of the textual information is in English. So In this paper, it highlights the importance of understanding sentiment in non-English languages and its potential applications in various domains. The paper contributes to the field of sentiment analysis by addressing the challenges of sentiment analysis in low-resource languages like Bangla, comparing different machine learning approaches, introducing transformer-based models, and providing publicly available datasets for future research which will help in the future.

1.3 Methodology:

The study includes data description and data analysis-which included token length distribution and class-wise Lexical n-gram discrimination. And for models, for the classical algorithms, Random Forest (RF) and Support Vector Machines (SVM) and for deep learning algorithms, CNN, FastText, and transformers based models such as BERT, DistilBERT and XLMRoBERTa.

1.4 Conclusion:

Among the classical algorithms, Support Vector Machines (SVM) generally outperform Random Forest (RF), except for the ABSA cricket dataset where RF achieves a slightly higher accuracy. Deep learning models consistently outperform classical algorithms, with both skip-gram (word2vec) and glove embedding with Convolutional Neural Networks (CNN) showing good performance. FastText performs well in three datasets, while CNN is slightly better overall. However, transformer-based models consistently outperform all others. BERT performs well on three datasets, while XLM-RoBERTa excels on two. The paper discusses the significance of computing valence scores (ϑ) for developing a sentiment lexicon in an automated and meaningful manner.

2 Limitations:

2.1 First Limitation

Bangla is a complex language with different dialects and nuances, making sentiment analysis challenging. The language's rich morphology, idiomatic expressions, and cultural context can lead to difficulties in accurately capturing sentiment which are not included in the study. Rather the deep learning algorithms that are used mostly in pointing out the positive/negative sentiment only as it said.

2.2 Second Limitation

Words in Bangla can have multiple meanings (polysemy), and different words can convey similar sentiments (synonymy). Sentiment analysis models must grapple with these linguistic complexities, and there is no such study or application of algorithms about this in this paper.

3 Synthesis:

The paper notes that an end-to-end comparison across datasets may not be feasible due to differences in the test sets. However, it concludes that the highest weighted F1 score is achieved on the YouTube comments dataset, while the relatively small size of the SAIL dataset may contribute to lower performance. The study covers a range of machine learning algorithms, including classical (e.g. SVM) and deep learning (e.g. CNN, transformer models). The authors curated publicly available Bangla sentiment datasets, cleaned the data, and split it into training, development, and test sets. They make these datasets available for reproducibility, which is essential for future research in this domain. The paper concludes that deep learning algorithms offer higher performance, albeit with increased resource and time complexity. The absence of official splits in most datasets prevents direct comparison with previous results, leading the authors to release their data splits to promote reproducibility and offer their results as benchmarks for future research. As for future work, the paper suggests investigating cross-model performance and consolidating datasets to potentially enhance model performance. Additionally, exploring aspect-based sentiment analysis within the dataset's diversity is mentioned as a potential avenue for future study.