

Intent Classification from Banglish Facebook Comments: Identifying what Customers Want

Sabbir Ahmed Sibli
Nuray Jannat
Shihab Sharar
Annajiat Alim Rasel

Abstract

In Bangladesh as well as many Bengali speaking people all over the world use Bangla phonetic form in English known as Banglish for social media commenting. In this research, We propose a hypothetical model to collect customer's Banglish Facebook comments from e-commerce based pages and predict their latent reaction (positive or negative) for any particular products.

1 Introduction & the Goal of Our Research

Online business is a blooming industry nowadays. Beside the e-commerce giants, many local entrepreneur also coming forward to the e-commerce industry with the help of various social media platforms like Facebook. This strategy has several benefits including expanding online reaches to the target audiences. In social e-commerce, generally sellers give public post of their products where people are open to comment on those posts. These comments can be used to analyse the customers reaction on the particular products. In Bangladesh, Customers here mostly use Bengali phonetic form in English known as **Banglish** on commenting. In the field of Natural Language Processing (NLP), sentiment analysis of customers is a widely common topic. But there are not so broad works available in Bengali language, especially for Banglish term. That's why we have a plan to collect, analyze and classify latent views of the people using Banglish comments. Here in this research, we've shown a proposed model where we are willing to collect Banglish Facebook comments from the ecommerce based pages post and classify the customers positive or negative view from those comments

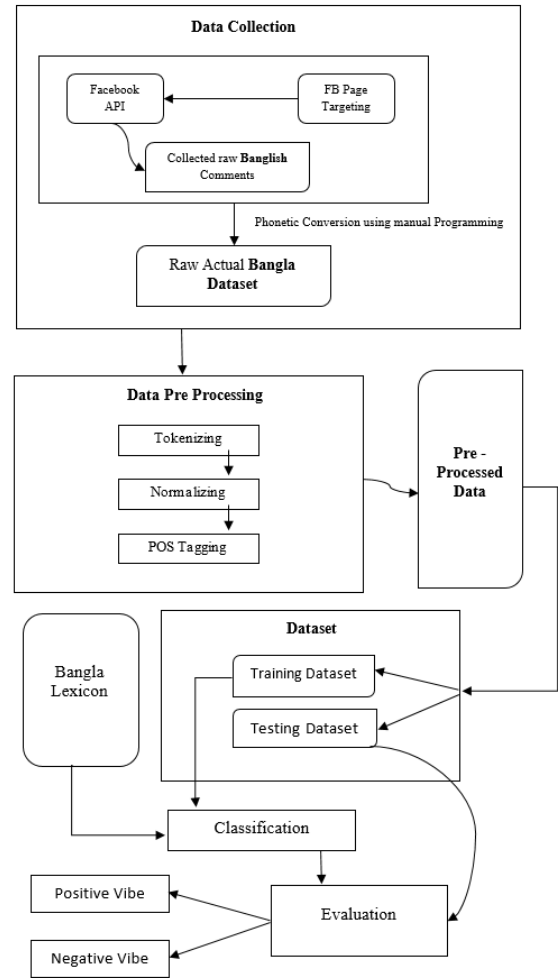


Figure 1: Workflow of our Proposed Model

2 Proposed Model

1 shows the workflow of the proposed model. The model is mainly divided into 3 parts. Firstly data collection, secondly data pre-processing and lastly the classification. In data collection, we plan to use Facebook e-commerce page targeting API(Elouardighi et al., 2017) for collecting Banglish comments. After collecting comments, we have a plan to convert those Banglish sentences into actual Bengali form using manual programming for phonetic conversion. This will end the Data collection phase. Then comes data pre-processing. tokenizing, normalizing and the POS tagging are the consequent steps of data pre-processing. We can use Bangla NLTK tools(Hasan et al., 2007) for doing these. This pre-processed dataset will be splitted into two parts, training data & testing data. Training data will be faded into a categorize classification mixing with a Bangla Lexicon dataset(Chowdhury and Chowdhury, 2014). This will classify which comments contains positive words and which are negative. Lastly, this classified data are compared with Testing dataset for evaluation purpose whereas the comments contains positive vibes or negative vibes of the customers.

3 Takeaway

In this hypothetical research, our goal is to visualize intent classification from Banglish Facebook comments. We have full plan to actualize the whole research. further specially on data collection using data mining techniques instead of manual programming.

References

- Shaika Chowdhury and Wasifa Chowdhury. 2014. Performing sentiment analysis in bangla microblog posts. In *2014 International Conference on Informatics, Electronics & Vision (ICIEV)*, pages 1–6. IEEE.
- Abdeljalil Elouardighi, Mohcine Maghfour, and Hafdalla Hammia. 2017. Collecting and processing arabic facebook comments for sentiment analysis. In *International Conference on Model and Data Engineering*, pages 262–274. Springer.
- Fahim Muhammad Hasan, Naushad UzZaman, and Mumit Khan. 2007. Comparison of different pos tagging techniques (n-gram, hmm and brill’s tagger) for bangla. In *Advances and innovations in systems, computing sciences and software engineering*, pages 121–126. Springer.