Sentiment Analysis of Bangla E-commerce Dataset: Using Different Machine & Deep Learning Models

Misbahul Sheikh
Computer Science and Engineering
Ahsanullah University of Science and Technology
Dhaka, Bangladesh
misbahul.cse.20200204039@aust.edu

Naushin Mamun
Computer Science and Engineering
Ahsanullah University of Science and Technology
Dhaka, Bangladesh
naushin.cse.20200204010@aust.edu

Nafisa Tasnim Neha
Computer Science and Engineering
Ahsanullah University of Science and Technology
Dhaka, Bangladesh
nafisa.cse.20200204020@aust.edu

***Abstract*—In the rapidly growing field of e-commerce, understanding customer sentiment is crucial for business success. This paper presents a comprehensive sentiment analysis of Bangla e-commerce dataset utilizing various machine learning and deep learning models. The study aims to evaluate the performance of Random Forest Classifier (RF), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Logistic Regression (LR), and Long Short Term Memory (LSTM) in classifying customer reviews. We preprocess the dataset using standard techniques such as tokenization, lemmatization, vectorization, etc. to ensure high-quality input for the models. Our experimental results demonstrate the effectiveness and comparative performance of these models, providing valuable insights into their applicability for Bangla text sentiment analysis.**

***Keywords*—Sentiment Analysis, Bangla, E-commerce, Random Forest, KNN, SVM, Logistic Regression, LSTM.**

I. Motivation

In recent years, the rapid expansion of e-commerce in Bangladesh has led to an unprecedented increase in online transactions and customer interactions. As businesses seek to enhance customer satisfaction and engagement, understanding consumer sentiment has become paramount. Sentiment analysis, a branch of Natural Language Processing (NLP), provides valuable insights into customer opinions by analyzing textual data. Despite its global application, there is a notable gap in sentiment analysis research focused on the Bangla language, particularly within the context of e-commerce.

This study aims to bridge this gap by employing various machine learning and deep learning models— RF, KNN, SVM, LR, and LSTM—to analyze sentiment in Bangla e-commerce datasets. By leveraging these models, we strive to develop robust techniques for accurately interpreting customer sentiments, ultimately contributing to improved business strategies and customer experiences. The integration of NLP and machine learning in this research not only addresses the linguistic challenges unique to Bangla but also showcases the potential of advanced computational methods in enhancing the e-commerce landscape in Bangladesh.

II. Literature Review

Some past works related to this topic were studied for this analysis.

A paper [1] on Bengali sentiment analysis for e-commerce product reviews. K-nearest neighbors (KNN) achieved the highest accuracy at 96.25% among all the models, outperforming Support Vector Machine (SVM) at 94.35%, Random Forest (RF) at 90.84%, Logistic Regression (LR) at 90.33%, and XGBoost at 90.56%. The study utilized an oversampled dataset to address the imbalance, significantly enhancing the performance metrics.

Another paper [2] focuses on sentiment analysis of Bangla conversations from Bangla movies and short film scripts using machine learning techniques. Including support vector machine, multinomial naïve Bayes, and logistic regression. The SVM achieved the highest accuracy of 86%.

This paper [3] presents a sentiment analysis of Bangla conversations, using Support Vector Machine (SVM), Multinomial Naïve Bayes (MNB), k-nearest Neighbors (k-NN), Logistic Regression (LR), Decision Tree (DR), and Random Forest (RF). The SVM and MNB demonstrated the highest accuracy among the six classifiers, achieving 85.59% and 85.15%, respectively.

The paper [4] investigates sentiment analysis on Bangla Facebook comments using various machine learning and deep learning models. The study focuses on classifying comments into three basic emotions: happiness, anger, and sadness. The authors tested several embedding layers including Word2Vec, CBOW, and Glove, and implemented models such as CNN, LSTM, and a hybrid CNN-LSTM. The CNN-LSTM hybrid model with Word2Vec embedding achieved the highest accuracy of 90.49% and an F1-score of 92.83%.

III. Methodology

A. Data Collection

The Data have been collected from kaggle.com website. This dataset contains 1000 unique reviews with 4 classes: positive, very positive, negative, and very negative.

B. Data Preprocessing

We perform several tasks in the pre-processing stage. We remove non-Bengali text, URLs, all types of digits, numbers, punctuation, emojis, and Bengali stopwords from our dataset. Then we apply tokenization to the comments. After that, we lemmatize the comment review column from the above task. Pre-processing of data normalizes the text.

C. Feature Extraction

Feature extraction is a technique to reduce the dataset dimensionality. We have used the TF-IDF (Term Frequency-Inverse Document Frequency) for feature extraction.

We have converted the Comment-Type column from multi-class to binary class. It simplifies the problem of distinguishing between two classes. Negative and Very Negative comments are represented as 1 (Negative) and Positive and Very Positive comments are represented as 0 (Positive).

	Comment	Product_Type	Comment_Type	binary_class
0	[খুবই, আরাম, দায়ক, শীতের, ঘরে, পরার, পারফেক্ট,...	Fashion	Positive	0
1	[স্লিপার, গুলা, দাম, হিসেবে, ভালো, ধন্যবাদ, সে...	Fashion	Very Positive	0
2	[প্রাইজ, হিসেবে, মান, রেগুলার, ইউজ, সুযোগ, ওয়া...	Fashion	Positive	0
3	[দাম, অনুসারে, ভালোই, টাকা, কিনেছিযদিও, পায়ে, ...	Fashion	Positive	0
4	[আলহামদুলিল্লাহ, জুতার, মান, ভালো, কালার, ছবির...	Fashion	Positive	0
5	[পুরাই, ফালতু, অর্ডার, দিয়েছি, এইটা, কথা, দারা...	Fashion	Very Negative	1
6	[সুন্দর, সেন্ডেল, একটু, যত্নসহকারে, পেকেট, আনল...	Fashion	Positive	0
7	[ঝামেলা, ছাড়াই, জুতা, পেয়েছি, জুতা, খুবভালো, চ...	Fashion	Very Positive	0
8	[জুতো, জোড়া, অবশ্যই, ভালো, মানের, পক্ষ]	Fashion	Very Positive	0
9	[মিলিয়ে, জুতাটি, অনেকভালো, ওজন, দিক, তো]	Fashion	Very Positive	0

Fig. 1. Tokenization Table

	pron	অক	অট	অড	অত	অতট	অধ	অন	অনন	অনল	...	য়ক	য়গ	য়ছ	য়জ	য়দ	য়ন	য়প	য়ম	য়া	॥
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	...	0.341991	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	...	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	...	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.313690	0.0	0.0	...	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	...	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	...	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	...	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	...	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	...	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.414673	0.0	0.0	...	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	...	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
11	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	...	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
12	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.320224	0.0	0.0	...	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
13	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	...	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
14	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	...	0.479821	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

15 rows × 719 columns

Fig. 2. Feature Table

D. Data Spliting

We have split the data into train and test sets and the ratio was (train: test) = (70: 30) for machine learning models. We also split the data into train, test, and validation sets and the ratio was (train: test: validation) = (70: 30: 30) for deep learning models. We kept the validation set the same as the testing set and used 10 epochs.

IV. RESULT ANALYSIS

The performance of our sentiment analysis models was evaluated using key metrics such as accuracy, precision, recall, and f1-score.

Classifier	Accuracy	Precision	Recall	F1-Score
KNN	0.60	0.64	0.60	0.56
RF	0.76	0.76	0.76	0.75
LR	0.78	0.78	0.78	0.77
SVM	0.80	0.80	0.80	0.79
LSTM	0.88	0.88	0.88	0.87

TABLE I
PERFORMANCE METRICS

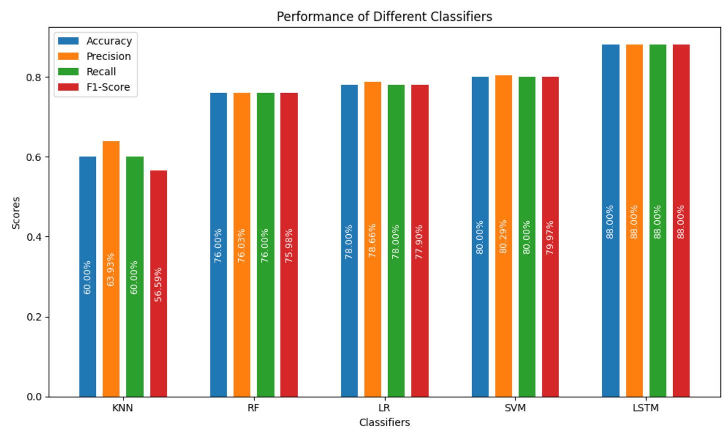

Fig. 3. Result Graph

Deep learning model–Long Short Term Memory (LSTM) provides us a better accuracy of 88% among all the models. This model's precision, recall and f1-score values are 0.88, 0.88, and 0.87 respectively. The second good result is obtained by Support Vector Machine (SVM) with an accuracy of 80%. K-Nearest Neighbors (KNN) gives us the most poor result with an accuracy of 60%.

REFERENCES

[1] Mst. Tuhin Akter, Manoara Begum, and Rashed Mustafa, "Bengali Sentiment Analysis of E-commerce Product Reviews using K-Nearest Neighbors," 2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD), 27-28 February, Dhaka.

[2] Abdullah Al Jamil, Rahman,Rifat, "Sentiment and Emotion Analysis from Textual Data in Bangla Language," Southeast University Journal of Computing Sciences, Volume 01, No 01, June 2021

[3] Mahmudul Hassan, Shahriar Shakil, Nazmun Nessa Moon, Mohammad Monirul Islam, Refath Ara Hossain, Asma Mariam, Fernaz Narin Nur, "Sentiment analysis on Bangla conversation using machine learning approach," International Journal of Electrical and Computer Engineering (IJECE) Vol. 12, No. 5, October 2022, pp. 5562 5572.

[4] Muntasir Hoq, Promila Haque, and Mohammed Nazim Uddin, "Sentiment Analysis of Bangla Language Using Deep Learning Approaches," Springer Nature Switzerland AG 2021 N. Chaubey et al. (Eds.): COMS2 2021, CCIS 1416, pp. 140–151, 2021.