2021 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE) | 978-1-6654-7849-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/WIECON-ECE54711.2021.9829684

Sentiment Analysis on Bangla Financial News

Kibtia Chowdhury and Swakkhar Shatabda
Department of Computer Science And Engineering,, United International University
Plit-2, Madani Avenue, United City, Badda, Dhaka-1212, Bangladesh
kchowdhury211056@mscse.uiu.ac.bd, swakkhar@cse.uiu.ac.bd

***Abstract*—Sentiment analysis of news and media material can provide important details about events and causal relationships. Bangla text sentiment analysis is getting increasingly popular. As a market indicator, financial market participants use financial sentiment analysis. In Bangla, financial sentiment analysis is a less-explored research field. In this research, we attempt to analyze the emotion of Bangla financial news items. In this study, we created a dataset with titles of news articles classified as positive or negative from a well-known online web portal that provides business and financial news. We employ vector-based feature encoding and different machine learning approaches to illustrate the method's effectiveness. Our findings are encouraging, and they suggest the way forward for Bengali sentiment analysis research in the financial sector.**

***Index Terms*—Sentiment Analysis, Word Embedding, Classification Algorithms, Financial News.**

I. Introduction

More than half of the adult population of the generation receives a newspaper, according to the data: 2.5 billion in print and 600 million in digital form. As a result, the number of readers and users on the internet is bigger than the total number of internet users globally. We've been reading the "Financial Times." The Financial Times is a daily newspaper devoted to business and finance.

In this study, we offered sentiment analysis on Bangla financial news. We used Kaggle and the "Somoy" television station to obtain information. Word2vec, a word embedding approach, was used to preprocess the dataset. Following that, the train test split and k-fold procedures are utilized. In this scenario, we used certain classifier algorithms.

In this paper, we propose sentiment analysis on Bangla financial news. We gathered data from Kaggle and the "Somoy" television channel. The dataset is then preprocessed using word embedding and word2vec. The training test split and k-fold approach are then used. We applied certain classifier techniques in this case.

II. Literature Review

Financial sentiment research has been the subject of various studies. However, we can see that the majority of sentiment analysis work in the financial sector is done in languages other than Bangla. This has been a lesser-visited location in Bangla. However, there has been a significant amount of study on Bengali sentiment analysis in general and in specifically in various disciplines.

Zhao et al. [1]focused on negative emotional data the most. They used two pieces of information: negative financial data and subject selection. Experimental results showed that their method outperformed SVM, LR, NBM, and BERT in two financial sentiment studies. Anita et al. [2] developed an unsupervised approach for effective decision-making in financial news using sentiment analysis. They tried two methods: one that relied on noun-verb pairs and another that was assessed. The noun-verb strategy generated the best results in the testing. Yekrangi et al. [3] offered a financial market sentiment study that included three primary issues: all connected terms, effective words, and the financial corpus. In the sentiment analysis approach, the lexicon performed much better than the WordNet dictionary when applied to the financial corpus.

Based on the goals provided at the SemEval-2017 competition, Guzel et al. [4] offered a solution for polarity categorization of financial news items. Mishev et al. [5] established a methodology for evaluating the usefulness and performance of a variety of sentiment analysis algorithms Using datasets that have been tagged by financial professionals and are publicly available. The best-performing model is the XGB classifier.

Shaika et al. [6] aimed at automatically collecting user thoughts or views from Bangla microblog entries. They classify the data using Support Vector Machine (SVM) and Maximum Entropy (MaxEnt) and then compare the results. Sharmin et al. [7] drew their conclusions from information gleaned from social media. The BBC Bangla and Prothom Alo platforms provided the Bengali corpus. In a stochastic gradient descent technique, the NADAM optimizer was employed. They observed that their model, with a learning rate of 0.002, gave the greatest outcomes, with 96.36 percent. Tuhin et al. [8] recommended two machine learning algorithms for extracting emotions from any Bangla text: the Nave Bayes Classification Algorithm and the Topical Approach. The topical technique outperformed the other at both levels of magnitude, according to a comparison of the two strategies' performance.

As a result, we find a research gap in the analysis of Bangla's financial sentiment. As a result, we have the opportunity to work in this subject. The majority of the work is in the financial and Bangla fields. However, combining Bangla and finance does not work. We present sentiment analysis on Bangla financial news in this study.

978-1-6654-7849-6/21/$31.00 ©2021 IEEE

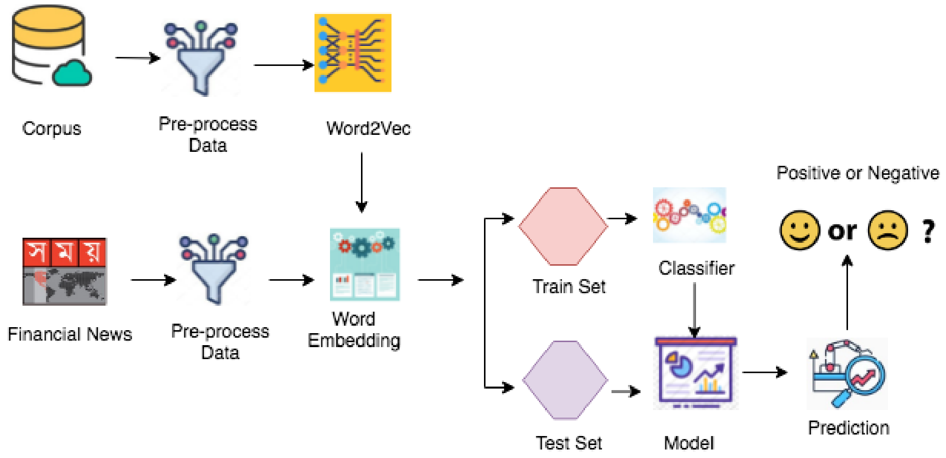
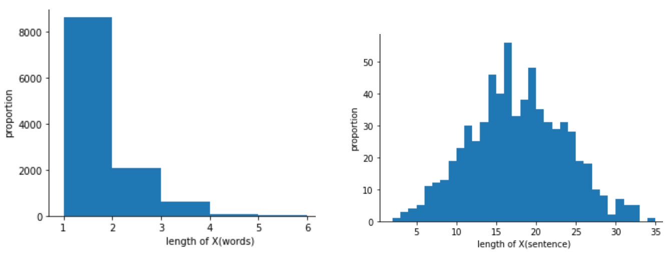


Fig. 1. System Diagram



(a) (b)

Fig. 2. Histograms showing the distribution of (a) word lengths and (b) sentence lengths in the dataset.

III. Our Proposed Method

We depicted the complete project effort in the system diagram given in Figure 1. We started by gathering two datasets, one of which is a corpus dataset and the other of which is a business news dataset. We pre-processed the corpus and training data before applying the word2vec vector after gathering the datasets. The training dataset is then loaded into the data frame. The training dataset is then loaded into the data frame. We already know that positive equals is 1 and negative is 0. For our experiments, we used two approaches. The first is the train-test split approach, and the second is the k-fold cross validation method. The classifier algorithm is then used to find a projected result. We showed a classification report in each trial using different metrics (precision, recall, F1 Score).

A. Dataset Collection

We used web scrapping to collect our dataset. We gathered dataset from a Bengali web based news portal. We gathered business news datasets from the "Somoy" Channel (https://www.somoynews.tv/). A total of 615 data points were gathered between December 1, 2018, and April 30, 2021. There are 350 data items that are positive and 265 data items that are negative. We denoted positive and negative data instances as two classes: 1 (positive) and 0 (negative). A summary of the dataset is given in Table I.

TABLE I
Summary of the financial dataset.

Financial News	Numbers of Data	Output Label
Positive	350	1
Negative	265	0

B. Pre-Processing

Word Embedding is a word description form that approves machine learning algorithms to catch words with the same elements. It helps in natural language processing, document clustering, feature creation, text classification tasks. Here we use word2vec techniques because we know word2vec is the popular technique. There are lots of surveys done by many people [9], [10], [11]. All those surveys looked at the technique of word embedding. Therefore, we pre-processing our two datasets financial news and corpus with word2vec word embedding. We collected a total 4152 corpus dataset from kaggle(https://github.com/Shayokh144/Bangla_Dataset_for _Opinion_Mining). We shown our two datasets words count description in Table II. Also, the histograms of the dataset is given in Figure 2.

TABLE II
Dataset Description.

Dataset	Numbers of Data	Words	Charaters
Financial News	615	4098	29311
Corpus	4151	106225	663730

C. Classifier Algorithm

We have employed a number of classifier algorithms in our experiments. In this section we briefly state about the details.

1) Naive Bayes: The naive Bayes classifier uses likehood and priors to find the posterior probability and thus generates classification. The following equation is the basic principla behind Naive Bayes Classifier.

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)} \tag{1}$$

2) Decision Tree: The Decision Tree algorithm divides the data into groups based on a attribute selection parameter. They are widely used fro classification and regression problems. The outcome of the class to which the data belongs is predicted by the tree. We utilized the classification and regression tree (CART) for our work and used entropy as the attribute selection parameter.

$$Entropy = -\sum_{j=1}^{m} p_{ij} log_2 p_{ij} \tag{2}$$

3) Support Vector Machine: Support Vector Machine (SVM) usually utilized to increase a classifier's margin. By altering the position and direction of the hyperplane, the Support Vector brings the data point closer to the hyperplane. SVM works by using a linear hyperplane to divide data into discrete classes. We employ a 'linear' kernel in our experiments.

$$D = \frac{|b1 - b2|}{||w||} \tag{3}$$

4) Random Forest: Random forest takes a portion of decision trees during training and outputs the mean forecast for regression and the mode prediction for classification. The decision trees value the many models in the data. The classification forecast is based on the majority opinion. We used the values of $n\_estimator = 2000$ in our experiments.

$$Entropy = \sum_{i=1}^{c} - p_i * log_2 p_{(i)} \tag{4}$$

5) Ada-Boost: Ada-boost combines multiple weak classifiers into a single strong classifier. In our experiment, we set the value of $n\_estimator = 1000$. This algorithm can be used for both classification and regression problem. Initially, the weight w is,

$$w = 1/n \tag{5}$$

After every iteration the weight W is updated, and after T iteration the final output is calculated,

$$H(x) = sign\left(\sum_{t=1}^{T} \alpha_t h_t(x)\right) \tag{6}$$

D. Performance Evaluation

In this work, we have used accuracy, precision, recall and F1-Score to measure the performances of the classifiers. These are defined using the confusion matrix. In a confusion matrix, True Positive (TP) is the number of positive instances correctly predicted; True Negative (TN) is the number of negative instances correctly predicted; False Positive (FP) is the number of negative instances wrongly predicted as positive and False Negative (FN) are the positive instances wrongly predicted as negative. Now accuracy is defined as following:

$$Accuracy, P = TP + TN/(TP + FP + TN + FN) \tag{7}$$

Precision is the percentage of positive instances from the total predicted positive instances. Precision is calculated as,

$$precision, P = TP/(TP + FP) \tag{8}$$

Recall is the percentage of positive instances from the total actual positive instances. Recall is calculated as,

$$Recall, R = TP/(TP + FN) \tag{9}$$

F1 score is the harmonic mean of precision and recall. The best value of F1 score is 1, and the worst is 0. F1 score is calculated as,

$$F1Score = 2((Precision * Recall)/(Precision + Recall)) \tag{10}$$

In our experiments, we have used both k-cross fold validation and train-test split. In the former case, the data set is divided into k folds. We used 5 fold cross validation in this project.

TABLE III
TRAIN ACCURACY AND CLASSIFICATION REPORT.

Classifier	Train Set	Precision	Recall	F1 Score
NB	0.56	0.66	0.57	0.43
DTA	1.0	1.0	1.0	1.0
SVM	0.63	0.69	0.64	0.59
RFA	1.0	1.0	1.0	1.0
ADA	1.0	1.0	1.0	1.0

IV. EXPERIMENTAL ANALYSIS

As previously stated, we used two trials to determine the sentiment result from financial news: train-test split and cross validation.

A. Train-Test Aplit Analysis

We created this data collection with the financial 615 news for the first experiment in Table III and Table IV. We divided the data into two sets: 80% train and 20% test. We have 492 train and 123 test after the split. DTA has an accuracy of 86 percent, ABA has an accuracy of 88%, and RFA has the best accuracy of 92%.

To assess our model, we employed precision, recall, and the f1 score. Table IV shows that DTA delivers 87% accuracy while ABA gives 89% accuracy for our dataset. RFA offers us the best accuracy of 93% for all of them.

All classifier algorithms had to be trained and tested for correctness. We used a bar chart to compare train tests (Figure 3). The color blue denotes train accuracy, while the color green denotes test accuracy. We could observe the amount of overfitting that might be present in the training.

TABLE IV
CLASSIFICATION REPORT, TEST ACCURACY, ROC ACCURACY FOR TEST SET.

Classifier	Test Set	Precision	Recall	F1 Score	ROC Curve
NB	0.60	0.48	0.60	0.51	0.55
DTA	0.86	0.87	0.85	0.85	0.83
SVM	0.61	0.57	0.62	0.57	0.62
RFA	0.93	0.94	0.93	0.94	0.98
ABA	0.88	0.90	0.89	0.89	0.97

TABLE V
k-FOLD ACCURACY AND CLASSIFICATION REPORT.

Classifier	5 K-Fold	Precision	Recall	F1 Score
NB	0.55	0.52	0.55	0.48
DTA	0.77	0.77	0.77	0.77
SVM	0.55	0.51	0.55	0.48
RFA	0.82	0.83	0.82	0.82
ADA	0.83	0.83	0.83	0.83

B. k-fold Cross Validation Analysis

The results of 5 fold cross validation are shown in Table V. We can see that RFA has an accuracy of 82%, while ABA has the best accuracy of 83%. To assess our model, we employed precision, recall, and the f1 score. Table 5 shows that DTA delivers 77% accuracy for our dataset, whereas ABA & RFA gives us the best accuracy of 83%.

C. ROC Curve For All Classifier

As a result, we can conclude that DTA, RFA, and ABA provide good accuracy in all studies. RFA is the most effective method. For each algorithm, we display a ROC Curve (Figure 4).

V. CONCLUSION

We concluded that classifying financial news using various NLTK classifiers is easier, and the larger the training data set, the more precise the findings. We will endeavor to improve the results of SVM and NB algorithms in the future. We'll also aim to cooperate with LSTM and CNN. We'll use a large dataset to increase accuracy while taking emotions into account. Working an actual time model that can forecast market changes based on mood in Bangla financial news can be more complicated.

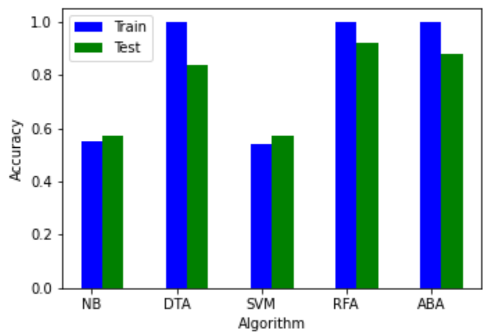


Fig. 3. Bar Chart for Train-Test Method.

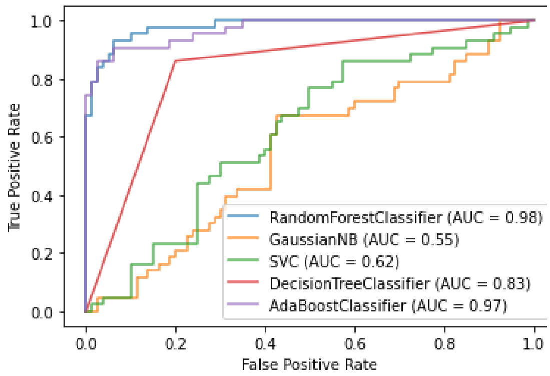


Fig. 4. ROC curve Accuracy For Test Set

REFERENCES

[1] L. Zhao, L. Li, X. Zheng, and J. Zhang, "A bert based sentiment analysis and key entity detection approach for online financial texts," in *2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, pp. 1233–1238, IEEE, 2021.

[2] A. Yadav, C. Jha, A. Sharan, and V. Vaish, "Sentiment analysis of financial news using unsupervised approach," *Procedia Computer Science*, vol. 167, pp. 589–598, 2020.

[3] M. Yekrangi and N. Abdolvand, "Financial markets sentiment analysis: Developing a specialized lexicon," *Journal of Intelligent Information Systems*, vol. 57, no. 1, pp. 127–146, 2021.

[4] G. Rinatovna Eremeeva, E. Vladimirovna Martynova, A. Aidarovna Khakimova, and L. Ernstovna Ilikova, "Sentiment analysis on english financial news," *Journal of Research in Applied Linguistics*, vol. 10, no. Proceedings of the 6th International Conference on Applied Linguistics Issues (ALI 2019) July 19-20, 2019, Saint Petersburg, Russia, pp. 574–582, 2019.

[5] K. Mishev, A. Gjorgjevikj, I. Vodenska, L. T. Chitkushev, and D. Trajanov, "Evaluation of sentiment analysis in finance: from lexicons to transformers," *IEEE Access*, vol. 8, pp. 131662–131682, 2020.

[6] S. Chowdhury and W. Chowdhury, "Performing sentiment analysis in bangla microblog posts," in *2014 International Conference on Informatics, Electronics & Vision (ICIEV)*, pp. 1–6, IEEE, 2014.

[7] S. Sharmin and D. Chakma, "Attention-based convolutional neural network for bangla sentiment analysis," *AI & SOCIETY*, vol. 36, no. 1, pp. 381–396, 2021.

[8] R. A. Tuhin, B. K. Paul, F. Nawrine, M. Akter, and A. K. Das, "An automated system of sentiment analysis from bangla text using supervised learning techniques," in *2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS)*, pp. 360–364, IEEE, 2019.

[9] D. Tang, F. Wei, N. Yang, M. Zhou, T. Liu, and B. Qin, "Learning sentiment-specific word embedding for twitter sentiment classification," in *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 1555–1565, 2014.

[10] F. Ali, D. Kwak, P. Khan, S. El-Sappagh, A. Ali, S. Ullah, K. H. Kim, and K.-S. Kwak, "Transportation sentiment analysis using word embedding and ontology-based topic modeling," *Knowledge-Based Systems*, vol. 174, pp. 27–42, 2019.

[11] B. O. Deho, A. W. Agangiba, L. F. Aryeh, and A. J. Ansah, "Sentiment analysis with word embedding," in *2018 IEEE 7th International Conference on Adaptive Science & Technology (ICAST)*, pp. 1–4, IEEE, 2018.