Sentimental Analysis on Reviews of Protein Supplement Using RNN-LSTM

Md. Abdul Aziz
Department of Electrical and Computer Engineering
North South University
Dhaka, Bangladesh
abdul.aziz01@northsouth.edu

Md Manjarul Islam
Department of Electrical and Computer Engineering
North South University
Dhaka, Bangladesh
manjarul.islam@northsouth.edu

Md. Shah Paran Munshi
Department of Electrical and Computer Engineering
North South University
Dhaka, Bangladesh
shah.paran@northsouth.edu

Samiha Samshi
Department of Electrical and Computer Engineering
North South University
Dhaka, Bangladesh
samiha.samsi@northsouth.edu

Mohammad Monirujjam Khan
Department of Electrical and Computer Engineering
North South University
Dhaka, Bangladesh
monirujjaman.khan@northsouth.edu

***Abstract*— In the most recent period, people are building up a habit of checking reviews before investing their time, money, and emotions in almost every aspect of their lives. To adapt to this newly made custom, we have seen a rise in the use of deep learning models for sentiment analysis from the reviews. Sentiment analysis is now widely used in various social media monitoring, product analysis, feedback analysis, target specific customers, and many other sectors. This paper introduces a deep learning model that will analyze the sentiments from people's reviews and predict whether a person's review is positive, negative, or neutral. This is done with the help of *LSTM* (an artificial recurrent neural network (RNN) as this model is more convenient to use to train the reviews and the outcomes are significantly better than any other model. For the LSTM model, our approach has achieved tremendous success, with 82.8% accuracy. We strongly hope and believe that our model will be able to impact the customer's decision-making. Nonetheless, we will still try to remodify our model to increase the accuracy of our model.**

Keywords*— *Deep learning; rnn; lSTM; review; sentiment analysis; positive; negative; activation function;

I. INTRODUCTION

Nowadays, people are becoming more conscious before buying any products. They will try to check the quality of a product. People were doing physical exercise to make their bodies fit and healthy in the past. But today, they are taking extra protein as a supplement in their body for producing more energy. These protein supplements help people do their exercise more without feeling tired. Due to this result, people are buying more and more supplements. For this, there are many varieties of protein supplements in the market. A person needs to take good protein in their body. Otherwise, it will be harmful for their health. Some people face serious health problems like thirst, bloating, reduced appetite, tiredness, and headache. People should take this protein under the doctor's supervision. Before buying these products, they should check the reviews of this protein supplement. In this paper, we will build a model, that can detect positive or negative reviews from the training data using Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM). In this paper, [1] proposed a system that will analyze a complete sentence word by word-based of RNN. The specialty of this research is that it has followed the GloVe word model advancing the training of the dataset. Moreover, it has proposed a combination of LSTM and GRU to mitigate the shortcomings of RNN regarding long-term memory. The research has shown that their proposed LSTM-GRU model is more efficient in comparing RNN, LSTM, and GRU experiments based on accuracy and training time. CNN-LSTM combined architecture showed to analyze product reviews posted on Twitter [2]. The author proposed a five-layer deep learning network. It is evaluated in weighted embedding layer, convolutional layer, max-pooling layer, LSTM layer, and dense layer. It has also shown an empirical analysis to predict the performances. It evaluates different word embedding vector-like GloVe, LDA2vec, fastText, word2vec, etc., along with weighting functions like TF-IDF, inverse document frequency, etc. The analysis has shown that the proposed combined neural network of CNN-LSTM with TF-IDF can beat the traditional CNN. A novel CNN and LSTM based deep learning sentiment analysis architecture for the Arabic language [3]. The architecture has used one CNN layer and two LSTM layers. The CNN layer is used for feature collection, and the LSTM is used for keeping short memory. SVM classifier has been used for classification. Moreover, it is followed by the FastText word embedding model. Although the authors use this classifier and word embedding system, they have shown the efficiency with others. The result has shown that the proposed model can be accurate to 90.75% in a multi-domain demonstration. The research has also suggested that the model can also be utilized with SVM classifier and FastText skip-gram model in the case of the Arabic language. It is found that the model can be improved up to +20.71% on relevant content. The authors [4] suggested CNN-LSTM based deep neural network with a novel padding method named the sentiment method. The research has intended to explore different types of reviews from tourism, product, blog post, etc. The new approach of a unique padding method can enable the existing architecture to resolute sentiments from specific sized datasets at an improving proportion. The proposed architecture has integrated lexicon with CNN and BiLSTM or LSTM network. The experiment has been conducted on some critical datasets like Stanford Sentiment Treebank. The results show a significant improvement from most of the traditional basic deep learning models. The hybrid architecture of CNN and LSTM to predict sentiments from the various platform in this study [5]. This hybrid system is for the feature extraction ability of CNN and sequential analysis capability of LSTM from a text. It followed to initiate two specific tasks. One is the scalability while evaluating the big data, and the other is that it can be applied to several domains simultaneously. Another impact of this research is that it has used COLSTM as the classifier. It is well-structured for social media text analysis. It can find the best impactful features from a large amount of data. The authors have used four datasets of text collected from different platforms to train the model. Words collected from texts are embedded in a vocabulary dictionary using a word embedding model. It showed a high performance both in accuracy and time than other machine learning algorithm types. Researchers have presented an LSTM based deep learning network for sentiment analysis from the Bangla language [6]. The authors have used a small dataset of Bangla text for training and testing. The goal is to find the effects of hyperparameters on this type of dataset. Moreover, the study tries to build a sentiment classification framework. It has been applied to several domains with different combinations of parameters. The result has shown high performance of 94% accuracy. The paper extends its discussion about the impact of the result on the young generation and common people stating its benefits for psychologists, businesses, etc.

II. METHODOLOGY

A. Proposed System

Here, we want to introduce a model based on a deep learning algorithm that analyzes the sentiments from people's reviews of various products. We built a system that fetches a chunk of public reviews data, and by analyzing those data, it predicts the percentage of positive, negative, or neutral. Data come from maybe restaurant reviews, protein supplement reviews, movie reviews or Twitter post-analysis. Our target is to use this system to fetch any textual reviews of any protein supplement. We will search our system for a particular object and the public opinions regarding this object. Furthermore, our system will classify the reviews into positive, negative classes. Thus, we will get better products with better reviews. In this paper, [7] suggested an LSTM based architecture for English and Spanish data. The study has aimed to overcome two specific obstacles of deep learning methods. First, it initiates an optimal value analyzing the data with cross-validation, and second, it uses dropout-based regularization to find efficient architecture. The paper has presented some crucial insights regarding this type of deep learning algorithm and its promising efficiency. The author [8] used LSTM for the sentiment analysis of text data. The system has used a large dataset of 50,000 text reviews from IMDB and another 50,000 text reviews from Amazon. The architecture follows one layer of word embedding, one layer of LSTM network, and one layer of the dense layer where Softmax is applied. It has approached to differentiate reviews into two parts: positive and negative. The result showed a high accuracy of 85% to classify sentiments. The researcher [9] proposed an LSTM based cryptocurrency prediction system. They aimed to predict the cryptocurrency rate and its relation to social media, analyzing the posts published in several domains. For this study, the authors collected 24,000 Weibo posts and 70,000 comments on them. Moreover, it has created a Crypto Vocabulary with the text and rate from time to time. The LSTM network classified the text reviews as positive, negative, and neutral. The study shows up to an 18.5% improvement in precision and 15.4% in recall than an autoregressive-based model. The authors [10] presented another LSTM based sentiment analysis. The study has used the GOP Debate Twitter dataset for training and analysis. It experiments with multiple parameters and their combination for higher performance. The research proposed an LSTM version that reduces six different parameters, namely slim LSTM. The results showed a good performance on detecting negative sentiments but low on positive. According to the authors, this positive sentiment detection can be improved with the artificially balanced dataset. Moreover, the study stated that adding multiple dense layers after the LSTM layer does not impact the performance. Among the optimizers evaluated, RMSprop and Adam performed well. Again, the experiment achieved optimum performance using 16 as batch size and 0.4 as a training validation ratio. Overall, the authors commented that standard LSTM and LSTM6 are more efficient than others. After all the paper reviewing, we found LSTM model is best for sentimental analysis-related topics. So, we proposed a model using LSTM.

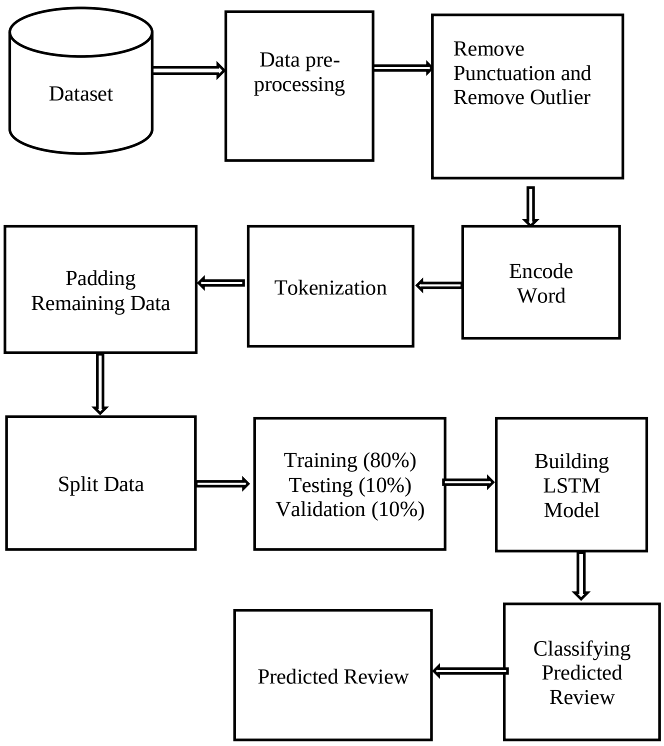

Fig.1. Block diagram of full proposed system.

B. Data Collection and Cleaning

The reviews are collected from Amazon, Flipkart, Facebook, and Google [11]. In total 4890 sentences are organized from these sites, in which 2655 are positive reviews and 2235 are negative. The obtained text is manually tagged according to the emotion of the words. The positive is given the number 1, while the negative is given the number 0. Table

I shows a sample of the final preprocessed dataset. Table II has more information on the dataset.

TABLE I. PRE- PROCESSED SAMPLE DATASET

Sentence	Tokenized word	Label Value
Best mass gainer in segment, good mixability	best, mass, gainer, segment, good, mixability	1
Give bad acne on my torso	Give, bad, acne, my, torso	0

TABLE II. DATASET DATASET INFORMATION

Data	Label 1 (Positive)	Label 0 (Negative)	Total
Total	2655	2235	4890
Training Data	2124	1788	3912
Testing Data	265	223	488
Validation Data	265	223	488

In this paper [11], the authors have provided us with a substantial textual dataset of both Bangla and Romanized Bangla called BRBT(Bangla and Romanized Bangla Text) consisting of 9337 post samples which are the very first of its kind and then been tested in Long Short Term Memory (LSTM). After preprocessing the dataset the authors split the set into a training, testing, and validation sets, then the datasets were categorized as positive, negative, and ambiguous. In total, 32 different experiments were conducted using the same LSTM model using different datasets, and for most of the experiments, accuracy rates are higher except for the dataset with categorical loss, modified text, ambiguous converted to 2 has a low accuracy of less than 55%.

C. Algorithom

The Long Short Term Memory (LSTM) model challenges addressing long-term information storing and short-term input, avoiding latent variable models. This paper [12] contains information about the basic architecture block of LSTM. First, split the dataset into two halves to divide it into inputs and outputs [13]. Because it can handle a wide range of data, LSTM is the best choice for both classical and statistical linear models. LSTM will forecast future data based on prior data. First, it is critical to identify how many days before each forecast will be accessible. This range includes window sizes as well. The forecasts will differ based on the values that are taken into account. Our model is divided into five layers which are interconnected with each other like an embedding layer.

- Embedding layer changes over our text tokens (whole numbers) into embedding of explicit size.
- LSTM Layer: An LSTM layer is defined by a hidden state size and the number of layers.
- Fully Connected Layer: A fully-connected output that maps the output of the LSTM layer to the desired output size.
- Sigmoid Activation Layer: that turns all output esteems in an incentive somewhere in the range of 0 and 1.
- Output: Sigmoid yield from the last timestep is considered as the last yield of this organization.

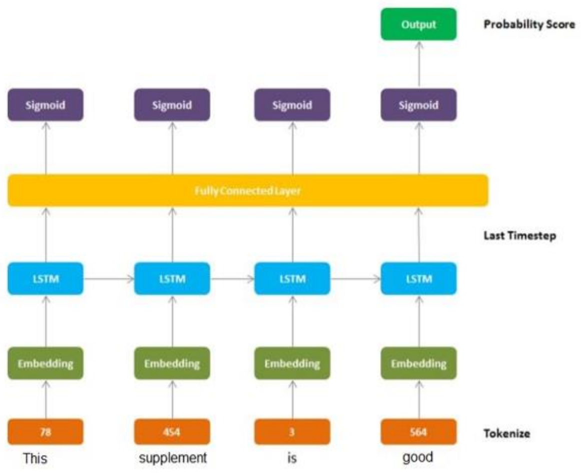

Fig.2. LSTM architecture for sentimental analysis.

Three gates make up an LSTM. Forget Gate, Input Gate and Output Gate are the three gates. Input, Forget, and Output Gates are mathematically represented by (1), (2), and (3) respectively.

$$f_t = \sigma(W_f\,[h_t\text{-}1,\, x_t] + b_f) \quad (1)$$

$$i_t = \sigma(W_i[h_t\text{-}1,\, x_t\,] + b_i) \quad (2)$$

$$o_t = \sigma(W_o\,[h_t\text{-}1,\, x_t] + b_o) \quad (3)$$

where "σ" represents a sigmoid function, "w_i","w_f" and "w_o" represents the weight for the input, forget and output neuron gate, "x_t" is the input at timestamp 't' and "b_i", "b_f" and "b_o" represents biases for input, forget and output gate respectively.

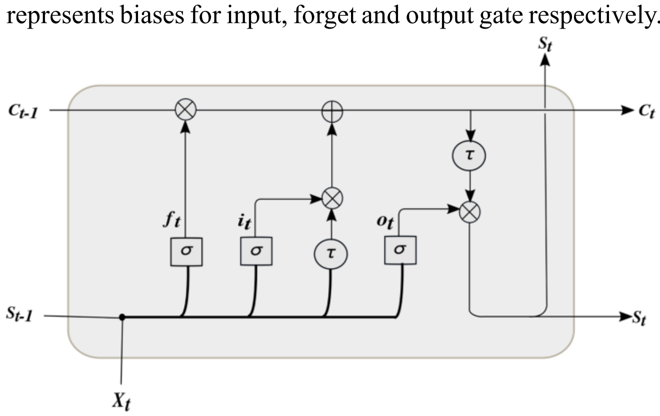

Fig.3. Basic architecture of a block of LSTM.

In the above Fig.3 "c_{t-1}", "c_t", "h_{t-1}", and "h_t" are the candidate cell state at timestamp (t), cell state at timestamp(t), the input of the previous hidden state, and input for the next hidden state. The mathematical equations for the candidate, state, and final output are shown below.

$$\overline{c_t} = \tanh\left(W_c[h_{t-1}, x_t] + b_c\right) \quad (4)$$

$$c_t = f_t * c_{t-1} + i_t * \overline{c_t} \quad (5)$$

$$h_t = o_t * \tanh(c_t) \quad (6)$$

.

III. Result and Discussion

Utilizing the pack of-words highlight just as by evaluating the execution of various lambda on the validation set, we found that the outcome with the best precision was accomplished when lambda was set to 0.0003. The comparing validation mistake was 11.035%. As our dataset contain 4,890 data and we divided them into three parts testing, training and validation. For training purposes, we took 80% of the data from the dataset and the remaining 20% of data is kept for the text testing and validation of the model. The training output is given in Fig. 4

```
Epoch: 1/4... Step: 100... Loss: 0.609600... Val Loss: 0.629905
Epoch: 1/4... Step: 200... Loss: 0.628857... Val Loss: 0.581476
Epoch: 1/4... Step: 300... Loss: 0.662385... Val Loss: 0.589381
Epoch: 1/4... Step: 400... Loss: 0.545402... Val Loss: 0.575211
Epoch: 2/4... Step: 500... Loss: 0.475501... Val Loss: 0.567665
Epoch: 2/4... Step: 600... Loss: 0.532206... Val Loss: 0.509516
Epoch: 2/4... Step: 700... Loss: 0.465704... Val Loss: 0.515248
Epoch: 2/4... Step: 800... Loss: 0.464423... Val Loss: 0.445131
Epoch: 3/4... Step: 900... Loss: 0.410016... Val Loss: 0.456189
Epoch: 3/4... Step: 1000... Loss: 0.235831... Val Loss: 0.515811
Epoch: 3/4... Step: 1100... Loss: 0.256774... Val Loss: 0.434333
Epoch: 3/4... Step: 1200... Loss: 0.251274... Val Loss: 0.454153
Epoch: 4/4... Step: 1300... Loss: 0.247446... Val Loss: 0.562225
Epoch: 4/4... Step: 1400... Loss: 0.217698... Val Loss: 0.497040
Epoch: 4/4... Step: 1500... Loss: 0.183696... Val Loss: 0.539584
Epoch: 4/4... Step: 1600... Loss: 0.169013... Val Loss: 0.570934
```

Fig.4. Training the dataset and validation loss visualization

In this Fig.4, we can figure out how the validation loss is decreasing from Epoch 1/4 to 4/4. In Epoch 1/1, the validation loss was 0.6299, and in Epoch 4/4, the loss dropped to 0.5709. We also locate that our model's precision is lower subsequent to stemming, lemmatizing, and erasing stop words. Our speculation is that some stop words may, in any case, have assumption inclination. By disregarding these words, potential data was lost, leading to lower precision.

This paper [14] focused on applying CNN to conduct on Hindi movie reviews collected from online newspapers and websites. The dataset was categorized as positive, negative and neutral. Variable numbers of convolution layers are used in the tests, as well as different numbers and sizes of filters and the end result clearly shows that the model consists of 2 convolutional layers with filter sizes 3 and 4.

TABLE III: Training and validation analysis

Data	1000	3000	4000	5000
Training Accuracy	0.68	0.72	0.74	0.75
Validation Accuracy	0.62	0.66	0.68	0.69
Training loss	0.620	0.549	0.525	0.490
Validation loss	0.640	0.625	0.586	0.582

From the TABLE III, we can easily figure out that as much the training goes, the training and validation accuracy goes higher, and training and validation loss accuracy goes lower. After training, when we test our data, we kept for testing, and we get some accuracy for test loss and test accuracy to see in Fig.5.

```
# avg test loss
print("Test loss: {:.3f}".format(np.mean(test_losses)))

# accuracy over all test data
test_acc = num_correct/len(test_loader.dataset)
print("Test accuracy: {:.3f}".format(test_acc))
```

```
Test loss: 0.515
Test accuracy: 0.820
```

Fig.5. Test loss and test accuracy output.

As we know that our data set was smaller so that we got low accuracy because NLP (Natural language Processing) is a big process to understand with a small dataset. It is a process of understanding human sentimental, human thought and belief. For understanding human sentimental, we need at least 100k data. Nevertheless, we have only 25k data, including positive negative and neutral reviews of protein data. However, we get a test loss of around 51% and Test accuracy around 82%.The researcher [15] presented a comparison among several machine learning algorithms to predict stock price. This includes linear regression, support vector machine, and LSTM. The study also incorporates Twitter sentiment analysis. In the result, the authors have shown that LSTM can be more fruitful than the others. However, the study has some shortcomings. It shows a high positive accuracy when there is a lot of polarizing comments in a Twitter post. That is, if there are about 80% comments or posts featuring a positive note, then the stock price will likely be higher. Other than this, the architecture foreshadows positive and negative sentiments.

IV. Conclusion

In this paper, we present a module where you can find sentimental analysis of protein supplements using modern Recurrent Neural Networks (RNN) like Long Short-Term Memory (LSTM) and Bidirectional Recurrent Neural Networks for sentimental analysis of protein supplements. There are three parameters to determine the sentimental: positive, negative, or neutral. Using this app, the user will get to know that (i) a protein is good for his/her health, (ii) unhealthy for him/her (iii) does not applicable for his/her health. The comparable methodology can then be recreated for surveys and remarks in different zones like film audits and online media posts. Assume somebody might want to discover a film generally prestigious for its ideal exhibition of 'adoration' by working conclusion investigation and extremity identification on IMDB or spoiled tomato film audits, he would get what he wanted. That would be our expectations later on: gathering suppositions from individuals, separating data from conclusions, and producing recommendations from data.

REFERENCES

[1] Hassan, A., Amin, M.R., Al Azad, A.K. and Mohammed, N., 2016, December. Sentiment analysis on Bangla and romanized Bangla text using deep recurrent models. In 2016 International Workshop on Computational Intelligence (IWCI) (pp. 51-56). IEEE.

[2] Hassan, A., Amin, M.R., Mohammed, N. and Azad, A.K.A., 2016. Sentiment analysis on Bangla and romanized bangla text (BRBT) using deep recurrent models. arXiv preprint arXiv:1610.00369.

[3] Yu, B., Zhou, J., Zhang, Y. and Cao, Y., 2017. Identifying restaurant features via sentiment analysis on yelp reviews. arXiv preprint arXiv:1709.08698.

[4] Dos Santos, C. and Gatti, M., 2014, August. Deep convolutional neural networks for sentiment analysis of short texts. In Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers (pp. 69-78).

[5] Wang, J., Yu, L.C., Lai, K.R. and Zhang, X., 2016, August. Dimensional sentiment analysis using a regional CNN-LSTM model. In Proceedings of the 54th annual meeting of the association for computational linguistics (volume 2: Short papers) (pp. 225-230).

[6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[7] Lawrence, S., Giles, C.L. and Fong, S., 2000. Natural language grammatical inference with recurrent neural networks. IEEE Transactions on Knowledge and Data Engineering, 12(1), pp.126-140.

[8] https://www.lamag.com/article/best-protein-powder-supplements-review-top-ranking-products/

[9] Al-Amin, M., Islam, M.S. and Uzzal, S.D., 2017, February. Sentiment analysis of Bengali comments with Word2Vec and sentiment information of words. In 2017 international conference on electrical, computer and communication engineering (ECCE) (pp. 186-190). IEEE.

[10] Pal, S., Ghosh, S. and Nag, A., 2018. Sentiment analysis in the light of LSTM recurrent neural networks. International Journal of Synthetic Emotions (IJSE), 9(1), pp.33-39.

[11] Ain, Q.T., Ali, M., Riaz, A., Noureen, A., Kamran, M., Hayat, B. and Rehman, A., 2017. Sentiment analysis using deep learning techniques: a review. Int J Adv Comput Sci Appl, 8(6), p.424.

[12] Tripathi, Milan. "Sentiment Analysis of Nepali COVID19 Tweets Using NB, SVM AND LSTM." Journal of Artificial Intelligence 3, no. 03 (2021): 151-168.

[13] Sivaganesan, D. "Novel Influence Maximization Algorithm for Social Network Behavior Management." Journal of ISMAC 3, no. 01 (2021): 60-68.

[14] Rani, S. and Kumar, P., 2019. Deep learning based sentiment analysis using convolution neural network. Arabian Journal for Science and Engineering, 44(4), pp.3305-3314.

[15] R. Ni and H. Cao, "Sentiment Analysis based on GloVe and LSTM-GRU," 2020 39th Chinese Control Conference (CCC), 2020, pp. 7492-7497, DOI: 10.23919/CCC50068.2020.9188578.