International Journal of Research in Advanced Engineering and Technology
www.allengineeringjournal.in
ISSN: 2455-0876
Received: 26-05-2023, Accepted: 14-06-2023, Published: 30-06-2023
Volume 9, Issue 2, 2023, Page No. 28-37

Forecast the rating of online products using novel deep learning technique

Md. Ahsan Habib, Asma Akter
Department of Computer Science and Engineering, Mawlana Bhashani Science and Technology University, Dhaka, Bangladesh

Abstract
The term "forecast" refers to a process that takes historical data as input for making informed estimates to determine future trends. Nowadays, marketing on online platforms is growing very fast to cope with the challenges of the decade. Because of this, online product rating is an essential parameter to measure the admissibility of products to consumers. From this, an online consumer decides the quality and nobility of the available products. It helps the consumer make a decision to buy or not. Being a successful businessman or stakeholder requires the ability to judge online product ratings. After analysis, the product rating helps a producer reach a decision on whether to modify their products or not. In this decade, online marketing has become more commonplace day by day, where consumers buy their products and give them a numerical rating, like a star. Producers need to analyze this rating to drive better revenue in their business. In our paper, we proposed a hybrid model comprised of long-short term memory (LSTM) and a convolutional neural network (CNN) that achieves 95% accuracy in online product review analysis. We applied the model mentioned above to the dataset named "GrammarandProductReviews." provided by Datafiniti. We have also applied some supervised machine learning algorithms such as Random Forest, Support Vector Machine (SVM), Logistic Regression, and XGBoost Algorithm with TF-IDF vectorize to analyze the customer product text review. Finally, we understand the findings obtained from the model presented by the various researchers. From all the studies, we observed that the combination of long-short term memory (LSTM) and a convolutional neural network (CNN) shows better performance (95% accuracy) than any other model.

Keywords: Forecast, product reviews, deep learning, CNN, LSTM, word-embedding

Introduction

In this digital era, requesting and revealing product ratings is an important and ongoing part of doing business on an online platform. Consumers rely on ratings and reviews, which have only been around for about two decades. According to consumer research from Deloitte and Toche LLP in 2019, the majority of online consumers use product reviews or ratings to evaluate online products ("Why ratings and reviews are important for your business," 2020). With the revolution of online marketplaces such as Amazon, Flipkart, etc. and review hosting websites such as Yelp, consumers are leaving more reviews now than ever. According to the article ("Why ratings and reviews are important for your business," 2020) [25], the number of reviews submitted increased nearly 11% year-over-year. Consumers may seek to understand whether their product fully meets their specific demand by analyzing product reviews. Review can give a producer's SEO (search engine optimization) a real boost. As an online retailer, product reviews are important for sales. Here are some statistics, ("Why Product Reviews are Important for Buyers and Sellers," 2018) [24].

- 63% of customers prefer to make a purchase when there are user reviews.
- 88% of online shoppers include product ratings and reviews in their purchase decisions.
- 70% of customers consult reviews or ratings before making a final purchase.

There are several ways to analyze customer product reviews. Deep learning is a growing field of automation in different sectors. Deep learning is a branch of machine learning (ML) and artificial intelligence (AI) that simulates the way humans gain certain types of knowledge. Deep learning also consists of statistics and predictive modeling. The advanced machine learning algorithm, which is known as an "artificial neural network," underpins most deep learning models. As a result, deep learning may sometimes be referred to as "deep neural learning" or "deep neural networking." Deep learning has recently been used in image recognition tools, natural language processing (NLP), and speech recognition software. The best contribution of this paper is to notify online market product reviews about the satisfaction of consumers with the corresponding products. A comprehensive view of the satisfaction of their consumers. To the best of our knowledge, there is very little research work on this dataset called "GrammarandProductReviews." So the reason for selecting the dataset in our study is to evaluate how it performs with the proposed model as well as with different machine learning algorithms. The following contributions are accounted for in this study:

- A guide for the embedding of online product reviews named "GrammarandProductReviews" is provided.
- Deep learning approaches for various characters have been used for analyzing product reviews or ratings. Analyzing efficiency is improved based on this integrated use.
- A novel hybrid paradigm is suggested that incorporates a range of product reviews and rating representations. High classification progress has been achieved when the proposed model derives improved functionality.
- Validate the proposed model's efficiency compared with other supervised machine learning algorithms.

The rest of the paper is structured as follows Section 2 discusses similar work on analyzing product review in online marketplace. Description and properties of dataset are discussed in section 3 The methodology proposed for section 4 accounts. Experimental findings and analysis are discussed in section 5. This article is eventually concluded in the last section 6.

Related work

This section presents previous work related to our proposed model. Many great researchers have left a significant trace in the fields of deep learning and natural language processing. In previous reviews, linguistic features and feature engineering were important aspects of the analysis. In the paper (Turney, 2002) [20] the unsupervised algorithm for review classification is represented. The model works on positive reference and reference words to find the semantic orientation. The paper shows an accuracy of 74% on the opinion dataset, whereas on the movie review dataset it is about 66%. In (Pang and Lee, 2008) [15], they worked on different supervised machine learning algorithms such as Naïve Bayes, Maximum Entropy, and SVM and their effectiveness for analyzing reviews. It works on the movie review dataset, which shows that SVM outperforms any supervised machine learning algorithm. The paper (Mitchell and Lapata, 2010) [12] proposed a framework for representing the meaning of word combinations in vector space. The authors had implemented different ways of combining words to represent the meaning of the sentence, such as vector averaging, vector addition, and vector multiplication. The experimental results on the BNC corpus and Wordsim353 dataset show that models are significantly correlated with human ratings.

In their paper (Bengio *et al*., 2003) [2], they proposed a neural network model to learn distributer representation for words. The proposed model represents the concept of statistical language modeling to learn the distributed representation of words. A comparative experiment was conducted on Brown Corpus and Associated Press (AP) news, and the results that they got using the proposed approach yielded better perplexity than the state-of-the-art method. (Socher *et al*., 2011) [18] have proposed a recursive autoencoder-based model to analyze sentences from review text. The model achieved state-of-the-art performance on MPQA and the newly created dataset. They (Ruder *et al*., 2016) [17] have proposed a hierarchical model for aspect-based sentiment analysis tasks using bidirectional LSTM. After an experiment, their result shows the model has an edge over the non-hierarchical baseline. The author also shows language and domain independence. (Wang *et al*., 2016) [23] have presented a model that consists of a neural network architecture that tries to represent both convolutional neural network (CNN) and recurrent Neural Network(RNN) architectures. The proposed model performs well on three benchmark datasets and achieves higher classification than any other existing model.

(Zhang *et al*., 2016) [27] proposed a model that consists of a bidirectional gated neural network model to compute the sentiment of tweets. The author used bidirectional GNN to compute an intermediate representation. Pooling is then applied to his representation. Then a three-way GNN is applied to compute the interaction between the target aspect and the surrounding context. Experimental results show better performance than baselines. (Tang *et al*., 2015) [19] have proposed a model that consists of GRNN to capture the intrinsic relationship between sentence in a review for document level sentiment classification. CNN and LSTM for sentence representation and GRNN for document representation. Experimental results on IMDB and Yelp dataset shows an edge over state of the art model. (Mikolov *et al*., 2013) [11] proposed a model that consists of two model *viz*. continuous skip-gram model and continuous bag of words model for computing continuous vector representation of words. The author had proposed a new technique for learning high-quality word vectors from the huge corpus. (Goldberg and Zhu, 2006) [6] have proposed a semi-supervised graph-based model for sentiment classification tasks. The model creates separate graphs for labeled and unlabeled reviews. Experiment were done on 'scale dataset v1.0' movie dataset, which has four classes (0, 1, 2, and 3). From the experiment, we find that models with PSP similarity achieve better performance.

(Rafay *et al*., 2020) [16] developed a model that performs rating prediction on business reviews. They achieved a tremendous result using both binary and multiclass data from their dataset. They perform the Multinomial Naïve Bayes algorithm, the Deep Learning algorithm, and the convolution Long Short Term Memory algorithm. The author achieved 84% accuracy on CLSTM and 83% with Glove. (Nikolenko *et al*., 2019) [14] proposed an aspect-based ratings prediction algorithm called AspeRa, which predicts rating based on the review text. Their proposed model achieves Aspe(Glove) -87% accuracy on Amazon's Instant Videos dataset Aspe(Glove) 73% on Amazon's Toys and game dataset.(Viard and Fournier-S'niehotta, 2018) [22] proposed a rating prediction algorithm that uses the XGBoost algorithm for their experiment, and their model achieves 78% accuracy. (Kumar *et al*., 2018) [9] developed a model that combines EEG signals and sentiment analysis of product reviews. They used the Artificial Bee Colony (ABC) algorithm on an EEG dataset and got 72% accuracy. (Cheng *et al*., 2018) [4] developed a model that removed the innate views from customer reviews and used them to improve the algorithm in different aspects. They proposed a model that consists of Stochastic Gradient Descent (SDG), which shows better performance than any other machine learning algorithm. (Zhang *et al*., 2019) [26] created a model that uses attention Convolution Collaborative Filtering (Att-ConCF) to improve the effectiveness of the feature. They use a Convolutional Neural Network (CNN) and achieve 77% accuracy.

Description of dataset

1. Dataset Properties

This segment comprises product reviews taken from the Kaggle dataset. In this paper, we use "GrammarandProductReviews." A labeled and standard dataset provided by Datafiniti. We have collected this from Kaggle ("Grammar and Online Product Reviews," n.d.). The experimental data collection has extracted a total of 71,045 product reviews from 1,000 different products. Our dataset contains 25 columns, but in this work we concerned two columns, "reviews. rating" and "reviews.text". The term review text contacts public opinion about a product, and review rating range from 1 to 5. The programming was completely implemented using Google Collab with Pandas Library, a versatile Python language development environment with advanced editing, checking, and numerical computation environments. Table 1. shows the number of instances of the dataset.

Table 1: Distribution of Rating in Dataset

Rating	Dataset Instances
1	3701
2	1833
3	4369
4	14598
5	46543

From Table 1, we see our dataset is imbalanced. The instances of rating 5 are too high compared to other rating instances. To imbalance the dataset, we use oversampling to remove imbalances in the dataset. Oversampling copies the data until all the distributions are not equal to the highest one. Figure 1 shows the pie chart distribution of the dataset before oversampling.

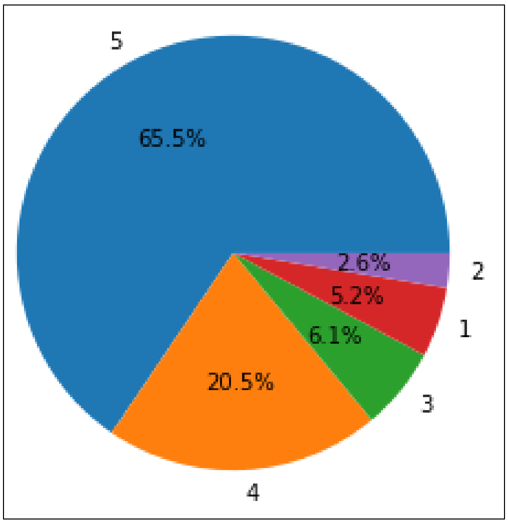

Fig 1: Pie Chart of Rating Distribution before Oversampling.

In Figure 2 show the Pie chart distribution of the dataset after oversampling. The figure shows us a balance dataset.

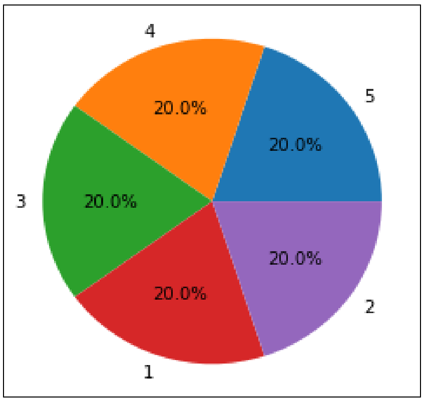

Fig 2: Pie Chart of Rating Distribution after Oversampling

In our dataset, you already know that we have 25 columns, and one of them is review text, which contains different public opinions for different products. We analyzed these public reviews and Figure 3 shows the words that people have used the most in their public reviews.

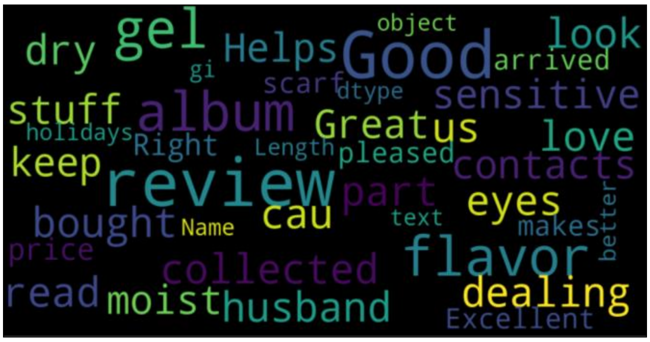

Fig 3: Words People have Used most in Their Public Opinions.

Fake reviews are defined as" deceptive reviews with the intention of misleading consumers in their purchase decision -making, often by reviewers with little or no actual experience with the products or services being reviewed." Fake reviews are the text that is not written by the actual consumers. Figure 4 shows who is giving fake reviews in our dataset.

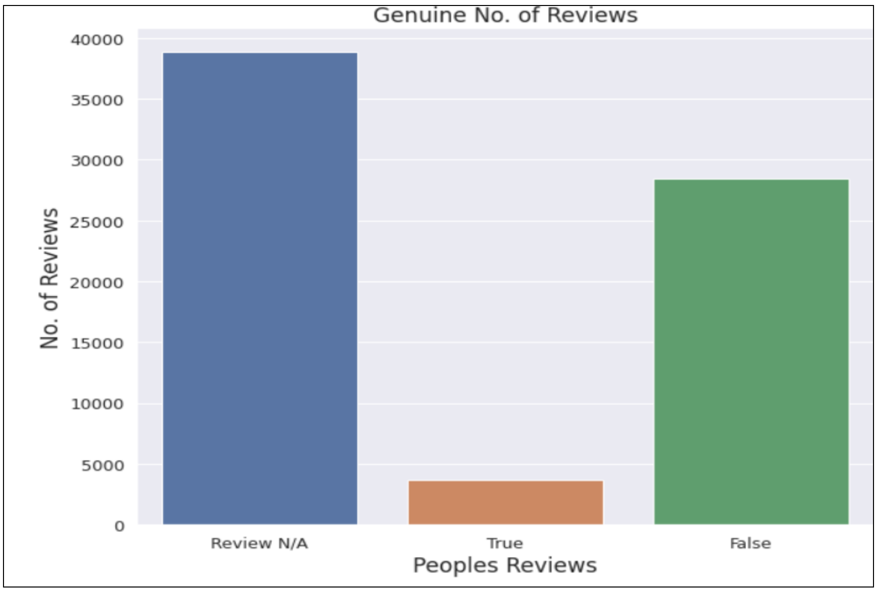

Fig 4: Fakes Reviews Graph of the Our Experimental Dataset.

The term "correlation" shows the changes between two variables. A correlation matrix is simply a table which displays the correlation coefficients for variables. The matrix depicts the correlation between all the possible pairs of values in a table (Luna, 2021) [10]. A correlation matrix is a powerful tool to summarize a large dataset, which consists of rows and columns that show the variable. Figure 5 shows the correlation matrix of the dataset.

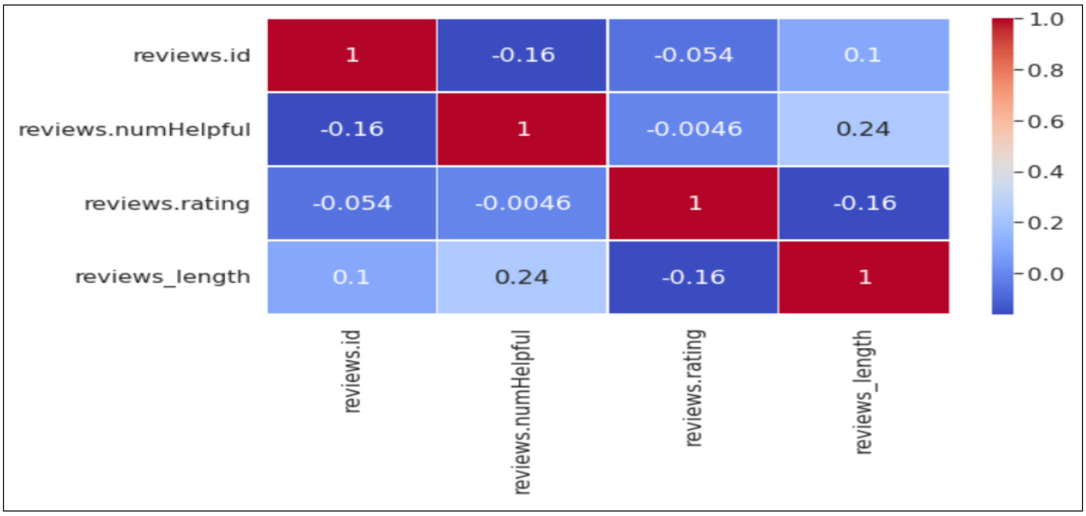

Fig 5: Correlation Matrix of Our Experimental Dataset.

From Figure 6. Let's have a look at what do the length of the reviews tell us about the rating of our experimental dataset.

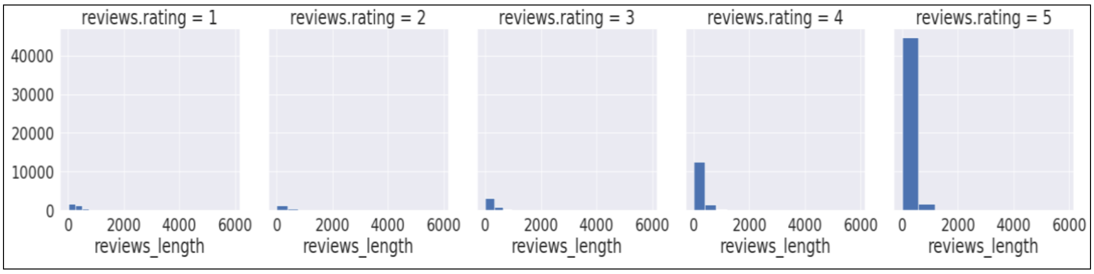

Fig 6: Length of the Reviews Tell about the Rating.

Proposed methodology

We use online product reviews as an input to our developed model. In our model, we load data directly from the Comma Separated File (CSV) format. After loading the dataset, we divide the data into two steps

- **Phase I:** Convolutional Neural Network (CNN) Feature Extraction
- **Phase II:** Long Short-Term Memory (LSTM) Feature Extraction

1. Convolutional Neural Network (CNN) Feature Extraction.

In addition, Long Short-Term Memory (LSTM) is used to compute the continuous representation of sentences in a sequential manner. The purpose of using LSTM is that it can compute a fixed-length sentence vector for any arbitrary variable length sentence. Then, to analyze the reviews, embedded functions have been applied to the Long Short Term Memory output layer. Text preprocessed using Natural Language Processing (NLP) techniques and word embedding is trained using word2vec.

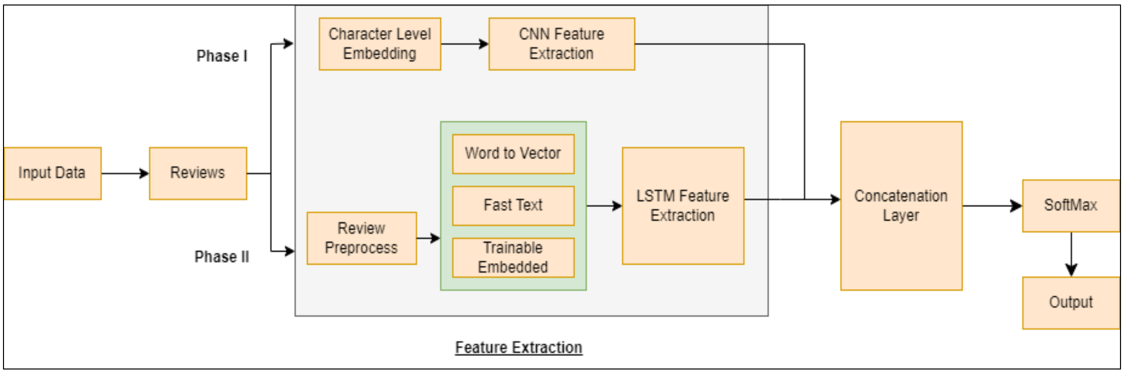

Fig 7: Architecture of the Proposed Novel Deep Learning Model.

a. Character Level Embedding

Character level embedding employs a one-dimensional convolutional neural network (1D-CNN) to determine the numeric representation of words based on character-level composition (Antonio, 2019) [1]. 1D-CNN is a process where several numbers of scanners slide through a word, character by character. At the end of the scanning process, information from different scanners is collected to form the representation of a word.

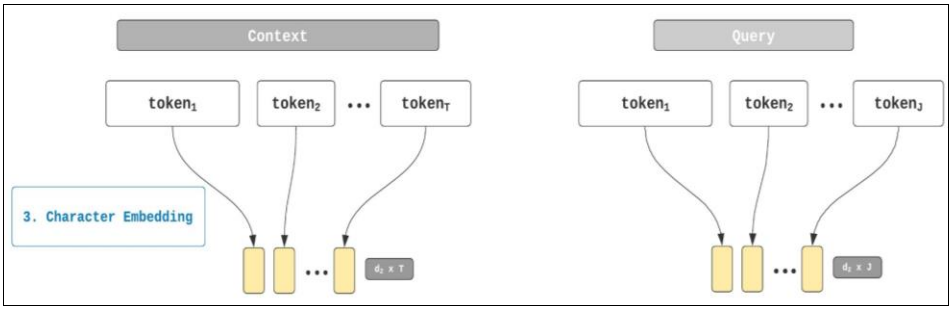

Fig 8: The character Embedding Step Converts Context tokens into a d2-by-T matrix and Query Tokens into a d2-by J matrix (Antonio, 2019) [1]

More broadly speaking, 1D-CNN is an algorithm capable of extracting information from shorter segments of a long input sequence.

b. CNN Feature Extraction:

A deep learning CNN model will pass their feature matrix through a series of convolution layers with filters (Kernels), RELU layer, polling layers, fully connected layers (FC) and apply some activation function such as sigmoid or SoftMax function to analyze reviews with probabilistic values between 0 and 1. An image can be represented as an array of pixel values. Similarly, we can represent text as an array of vectors that can be processed with the help of a CNN. When we are working with sequential data, we work with one-dimension convolutions, but the idea and application remain the same. Using Glove to obtain pre-trained word embeddings for our model is useful. Using Keras to train our data on a CNN architecture and evaluate the accuracy obtained on the validation set. A convolutional neural network is composed of "convolutional" layers and "down sampling" or "subsampling" layers. Convolutional layers comprise neurons that scan their input for patterns. Down sampling layers, or "pooling" layers often set up after convolutional layers in a ConvNet, mainly to reduce the feature map dimensionality for computational efficiency, which can in turn improve actual performance. Basically, these two layers occur in an alternate order, but that's not necessarily always the case. This is followed by an MLP with one or more layers (fully connected layer). Here, we will be training a Convolutional Neural Network to perform reviews analyzing on a dataset containing reviews from online product reviews. We will follow the following workflow:

- Importing the data and preprocess it into a desirable format.
- Using Glove to obtain pre-trained word embedding for our model.
- Using Keras to train our data on a CNN architecture and evaluate accuracy.

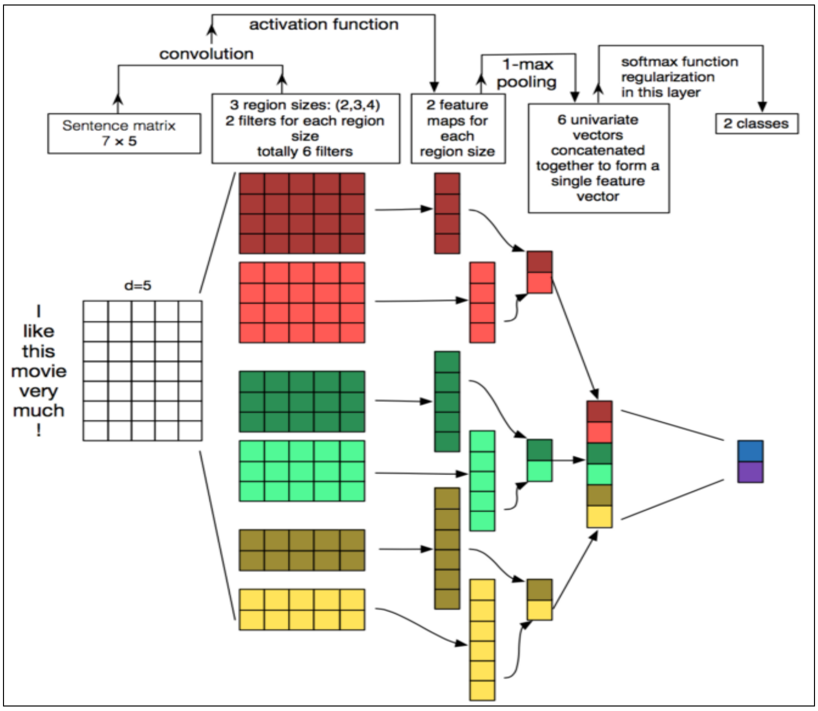

Fig 9: Review Classing using Convolution Neural Network (Newatia, 2019) [13]

2. Long Short-Term Memory (LSTM) Feature Extraction

2.1 Data Preprocess

To get a better result, the tweets are being processed. Tweets can have several languages based on the user, so we have to clean up the irrelevant data. After that, the URLs and the username are removed. Then we will do case-folding from uppercase to lowercase letters. Finally, we will remove the stop words as they are not significant and not related to emotions.

Create a word to integer

The vocabulary consists of a list of words that occurred in our text document, and these words have their own index. It will help us create a vector for a text document. First, we take the sentence, then vectorize it and count the number of occurrences in the text. The final vector will be the length of the text and be called the "featured vector." In this vector, each dimension will be numeric or categorical. We use the Count Vectorize provided by a scikit-learn library for vectorizing.

Encode the data label

So far, we create a list of review and index mapping dictionary from all our product review. In this step, we replace each word in our document by integers.

Encode the label

Our dataset consists of many labels but we use text and rating labels in our work.

Padding the remaining data

To maintain long or short text in our document we will pad or truncate of our text in a specific length. We call it sequence length which is the same as the number of steps for the LSTM layer.

2.2 Word2VecEmbed Features

The Word2Vec embedding technique has been used to get Word2VecEmbed features. We used the Genism ("Gensim," n.d.) library to create the Word2Vec model. We trained the model with our training info. After that, we obtained word vectors from the Word2Vec model that was learned. Here, too, every term is represented with a contour of dimension 300.

2.3 Fast text Embed Features

Here, the embedding of the Fast Text technique was used. The method for receiving fast text embed features is the same as the Word2VecEmbed features.

2.4 Trainable Embed Features

We used the Trainable Embed layer to create the Trainable Embed functionality

2.5 LSTM Feature Extraction

LSTM is a special type of recurrent neural network. In RNN, the output of the previous step is fed as input for the current step. LSTM overcomes the long-term dependencies of RNN. It can not predict the words stored in the long term memory but can give more accurate predictions based on the most recent information. By default, LSTM can retain the information for a long period of time. LSTM has a chain structure that contains four neural networks and different memory blocks called cells.

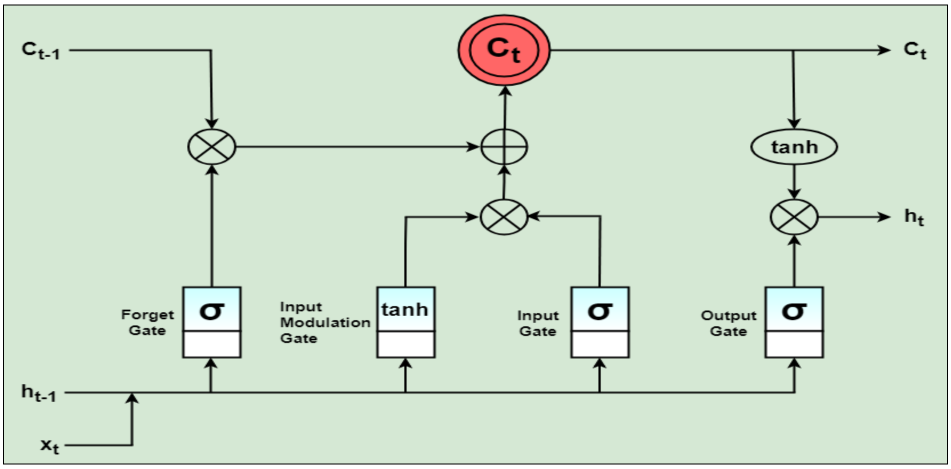

Fig 9: Long Short-Term Memory.

An LSTM network computes a mapping from an input sequence x = (x1, x T) to an output sequence h = (h1, hT) by calculating the network unit activations using the following equations iteratively from t = 1 to T.

$$f_t = \sigma(W_f.[h_{t-1}, x_t] + b_t) \quad (1)$$

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

$$\tilde{C}_t = \tanh(W_c.[h_{t-1}, x_t] + b_c) \quad (3)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (4)$$

$$o_t = \sigma(W_o.[h_{t-1}, x_t] + b_o) \quad (5)$$

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

Where W denotes the weight matrices, Ct is the cell state, and b is the input bias vector. And the i, f and o are the input, forget, and output gate layers. Cel- out activation function in this paper is tanh.

2.6 Concatenation Features

At the last stage of the proposed model, the phase-I and II features were merged. We use the SoftMax function in the layer of each network to map the network's non normalized performance to the likelihood distributions over the predicted output groups. It allows us to make predictions about all the labeled information.

Result and discussion

This section summarizes the experimental findings and performance interpretation of the proposed model. For word vector initialization, we used the pre-trained 300-dimension word vector. These word vectors are trained on product reviews. The dataset is divided into an 80:20 ration for training and testing. Experiments are conducted in a supervised setting where the overall rating (1,2..5) is used as a sentimental class. The loss function used for the classification task is cross-entropy.

1. Evaluation Metric

The overall performance of the model is evaluated using accuracy, recall, precision and the F1-score. The considerations obtained from the uncertainty matrix will be matched with the classification findings obtained in associated tests from the review classification to illustrate the accuracy.

Accuracy can be measured using the equation

$$\text{Accuracy} = (TP+TN)/(TP+TN+FP+FN) \quad (7)$$

Precision refers the approximation of the class labels for each class. Precision can be measured using equation 2.

$$\text{Precision} = TP/(TP+FP) \quad (8)$$

Recall value is the weighted average of the right points described correctly in any class. The equation defines this value.

$$\text{Recall} = TP/(TP+FN) \quad (9)$$

F1 is specified in equation 4 and F1 value is close to 1 for a good measure.

$$\text{F1 score} = (2*\text{Precision}*\text{Recall})/(\text{Precision} + \text{Recall}) \quad (10)$$

2. Results and Comparison

Many libraries and resources for constructing deeo learning models are now open. In this paper, we use the accuracy value as the critical success measure to equate our findings. In the accuracy calculation, we consider precision, recall and F1 score to determine the classifiers overall accuracy. Again,the overall accuracy of our model is 95%.

2.1 Machine Learning Algorithm

Different types of supervised machine learning have been used for this experiment, and we compared the results with our proposed model in Table 6. In our dataset, we conducted a train-test division pursuant to 80-20 regulations, of which 80 percent of the data was for training and 20 percent for testing.

a. XGBoost

XGBoost is an algorithm that has recently been dominating applied machine learning and Kaggle competitions for structured or tabular data. This algorithm is the implementation of gradient boosted trees designed for speed and performance (Brownlee, 2016) [3]. We implement this dataset for our model to analyze online product reviews. Table 2 shows the performance table of the dataset using XGBoost.

Table 2: Performance Measurement of the Dataset Using XGBoost.

Review Rating	Precision	Recall	F1
1	0.74	0.42	0.53
2	0.78	0.02	0.05
3	0.50	0.08	0.14
4	0.47	0.12	0.19
5	0.71	0.98	0.82

The Roc curve presentation of the XGBoost algorithm is shown in Figure 10.

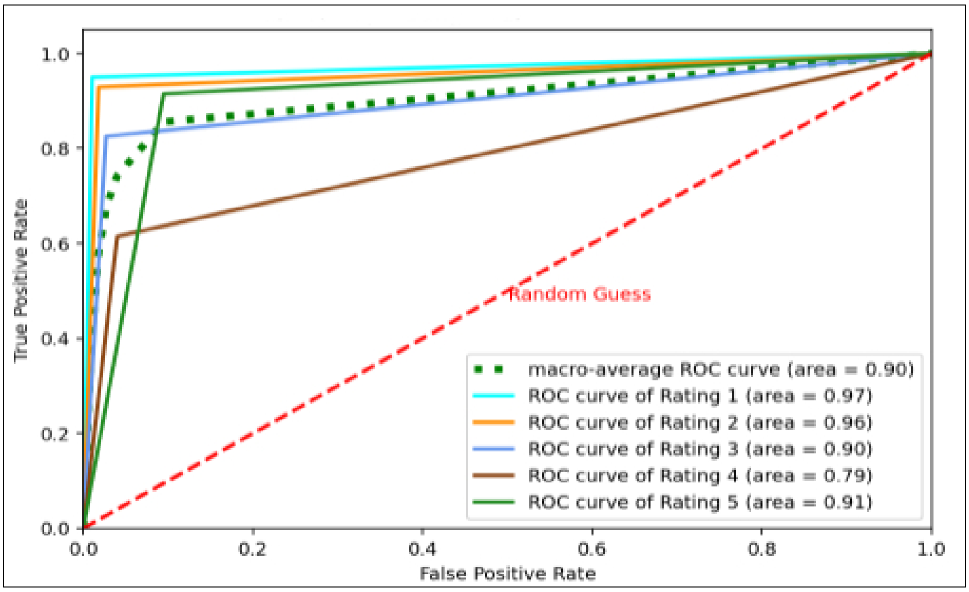

Fig 10: Roc Curve of XGBoost

b. Random Forest Classifier

Random forest is the most popular machine learning algorithm in supervised learning models. It can be used for both classification and regression. This algorithm is based on the concept of both classification and regression problems in ML. It is the process of combining multiple classifiers to solve a complex problem and improve the performance of the model. Table 3 shows the performance of the dataset using Random Forest.

Table 3: Performance Measurement of the Dataset Using Random Forest Classifier

Review Rating	Precision	Recall	F1
1	0.82	0.50	0.62
2	0.99	0.16	0.28
3	0.89	.0.13	0.22
4	0.64	0.18	0.28
5	0.72	0.99	0.83

The Roc curve presentation of the Random forest classifier is shown in Figure 11.

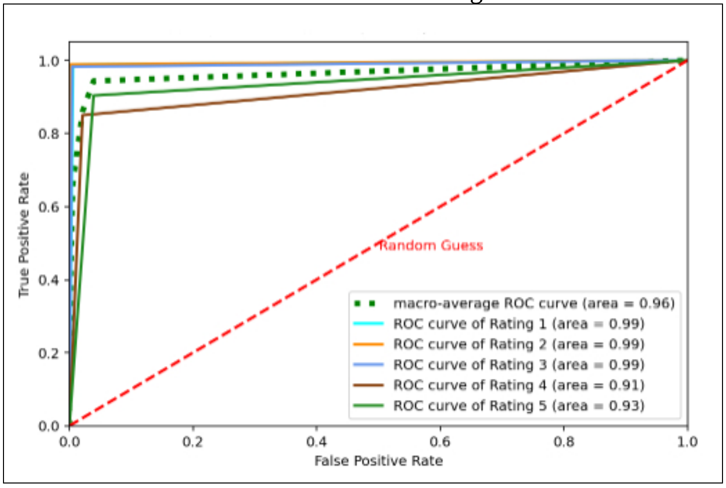

Fig 11: ROC curve of Random Forest Classifier

c. Logistic Regression

Logistic regression is one of the supervised machine learning algorithms used to predict the probability of a target variable. It is one of the simplest ML algorithms that can be used for various classification problems. Table 4 shows the performance of the dataset using Logistic Regression.

Table 4: Performance Measurement of the Dataset Using Logistic Regression

Review Rating	Precision	Recall	F1
1	0.72	0.71	0.72
2	0.52	0.13	0.21
3	0.46	0.25	0.33
4	0.49	0.28	0.36
5	0.78	0.94	0.85

The Roc curve presentation of the Logistic Regression algorithm is shown in Figure 11.

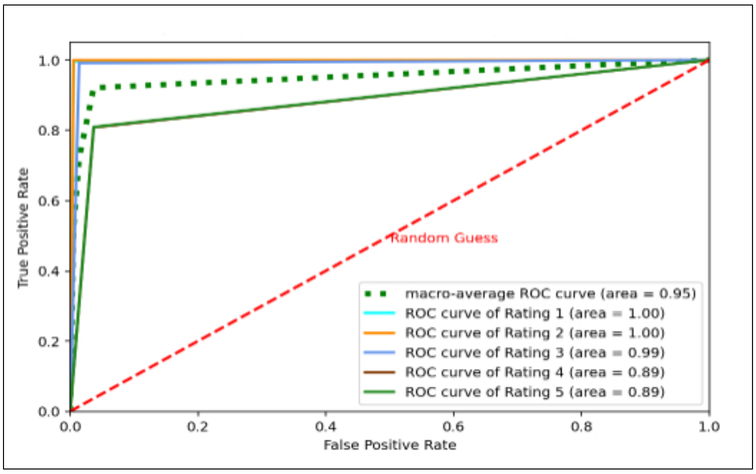

Fig 12: Roc Curve of Logistic Regression

d. GRNN with LSTM

In our experiment, we also analyzed our dataset using an existing model (Verma *et al*., 2017) [21]. The model consists of long-short term Memory and a gated recurrent neural network (GRNN). The sentence is processed to a fixed length vector using LSTM, a variant of RNN, and the interdependencies are captured using GRNN. We have also repeated the experiment for 100 rounds. After 70 rounds, the results seem to converge. We take the report after an interval of 10 rounds.

Table 5: Experimental Results of GRNN with LSTM

Round	Macro F-Measure	MAE	MSE	Accuracy
10	0.39	0.53	1.11	0.66
20	0.44	0.50	1.00	0.66
30	0.43	0.51	1.04	0.65
40	0.42	0.52	1.06	0.65
50	0.43	0.53	1.02	0.63
60	0.42	0.56	1.09	0.61
70	0.42	0.57	1.14	0.61

Figure 13 shows the graphical representation of the performance of the experiment based on the macroF-measure, accuracy, MSE, and MAE criteria.

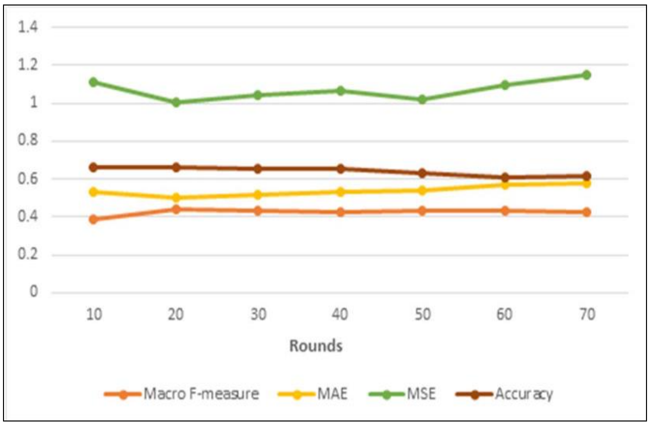

Fig 13: Comparison of Different Performance Measure.

In comparison with the current machine learning models, six different classification models are suggested, such as Random Forest, Support Vector Machine (SVM), XGBoost, Logistic Regression, Gaussian Naïve Bayes and GRNN with LSTM (Verma *et al*., 2017) [21]. We see from the contrast that our concept is better than any other machine learning algorithm.

Table 6: Accuracy Comparison of Different Machine Learning Algorithm with Our Proposed Model

Algorithm Name	Accuracy
Random Forest	71.82%
Support Vector Machine	72.10%
XGBoost	69.22%
Logistic Regression	72.67%
Gaussian Naïve Bayes	71.66%
GRNN with LSTM	66.14%
Proposed Model	95%

2.2 Comparative Analysis

We have applied our methodology to the dataset of the study (Verma *et al*., 2017) [21] and study (Hossain *et al*., 2021) [8]. Both papers use the "Amazon Electronics Dataset," which contains 7299 reviews of electronic products. The papers (Verma *et al*., 2017) [21] GRNN with LSTM and (Hossain *et al*., 2021) [8] analysis machine learning algorithm. Table 7 represents the accuracy comparison of both studies on the same dataset.

Table 7: Comparative Analysis with Other Study

Study	Algorithm	Dataset	Highest Accuracy
Study (Verma *et al*., 2017) [21]	LSTM, GRNN	Amazon Electronics Dataset (7299 Records)	66.14% by LSTM
Study (Hossain *et al*., 2021) [8]	RF, XGBoost, LR	Amazon Electronics Dataset (7299 Records)	94% by LR
Proposed Model	LSTM, CNN	Amazon Electronics Dataset (7299 Records)	95%

Conclusion

This paper basically used a hybrid model which consists of long-short term memory and convolutional neural network. In our research work, we also analyze our dataset using different supervised machine learning algorithms and compare them with our proposed model. We also implement different types of existing model and compare them with our model. The accuracy of our model is 95% which is better than any other existing model. In our work, we also analyze the properties of datasets from different perspectives. We think this paper applies empirically to data science and review research. In his paper, we suggested a new form of deep learning model with different supervised machine learning algorithms for analysis reviews. There is also space for progress in this study. In the future we will explore the applicability of Tree LSTM to predict ratings based on synonym and different types of emoticons and combine them with our proposed model.

References

1. Antonio M. Word Embedding, Character Embedding and Contextual Embedding in BiDAF — an Illustrated Guide [WWW Document], 2019. Medium. URL https://towardsdatascience.com/the-definitive-guide-to-bidaf-part-2-word-embedding-character-embedding-and-contextual-c151fc4f05bb (accessed 1.20.22).
2. Bengio Y, Ducharme R, Vincent P, Janvin C. A neural probabilistic language model. J. Mach. Learn. Res.,2003:3:1137–1155.
3. Brownlee J. A Gentle Introduction to XGBoost for Applied Machine Learning. Mach. Learn, 2016. Mastery. URL https://machinelearningmastery.com/gentle-introduction-xgboost-applied-machine-learning/ (accessed 1.22.22).
4. Cheng Z, Ding Y, Zhu L, Kankanhalli M. Aspect-Aware Latent Factor Model: Rating Prediction with Ratings and Reviews. Proc. 2018 World Wide Web Conf. World Wide Web – WWW,2018:18:639–648. https://doi.org/10.1145/3178876.3186145
5. Gensim: topic modelling for humans [WWW Document], n.d. URL https://radimrehurek.com/gensim/# (accessed 7.9.21).
6. Goldberg A, Zhu X. Seeing stars when there aren't many stars: Graph-based semi-supervised learning for sentiment categorization, in: Proceedings of TextGraphs: The First Workshop on Graph Based

Methods for Natural Language Processing. Association for Computational Linguistics, New York City, 2006, 45–52.
7. Grammar and Online Product Reviews [WWW Document], n.d. URL https://kaggle.com/datafiniti/grammar-and-online-product-reviews (accessed 1.19.22).
8. Hossain MdI, Rahman M, Ahmed T, Islam AZMT. Forecast the Rating of Online Products from Customer Text Review based on Machine Learning Algorithms, in: 2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD). Presented at the 2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD), 2021, 6–10. https://doi.org/10.1109/ICICT4SD50815.2021.9396822
9. Kumar S, Yadava M, Roy P. Fusion of EEG Response and Sentiment Analysis of Products Review to Predict Customer Satisfaction. Inf. Fusion, 2018, 52. https://doi.org/10.1016/j.inffus.2018.11.001
10. Luna Z. Feature Selection in Machine Learning: Correlation Matrix | Univariate Testing | RFECV. Geek Cult, 2021. URL https://medium.com/geekculture/feature-selection-in-machine-learning-correlation-matrix-univariate-testing-rfecv-1186168fac12 (accessed 1.20.22).
11. Mikolov T, Sutskever I, Chen K, Corrado G, Dean J. Distributed Representations of Words and Phrases and their Compositionality, 2013. ArXiv13104546 Cs Stat.
12. Mitchell J, Lapata M. Composition in Distributional Models of Semantics. Cogn. Sci.,2010:34:1388–1429. https://doi.org/10.1111/j.1551-6709.2010.01106.x
13. Newatia R. How to Implement CNN for NLP tasks like Sentence Classification. Saarthi.ai, 2019. URL https://medium.com/saarthi-ai/sentence-classification-using-convolutional-neural-networks-ddad72c7048c (accessed 1.20.22).
14. Nikolenko SI, Tutubalina E, Malykh V, Shenbin I, Alekseev A. AspeRa: Aspect-based Rating Prediction Model. ArXiv190107829 Cs, 2019.
15. Pang B, Lee L. Opinion Mining and Sentiment Analysis. Found. Trends Inf. Retr.,2008:2:1–135. https://doi.org/10.1561/1500000011
16. Rafay A, Suleman M, Alim A. Robust Review Rating Prediction Model based on Machine and Deep Learning: Yelp Dataset, in: 2020 International Conference on Emerging Trends in Smart Technologies (ICETST). Presented at the 2020 International Conference on Emerging Trends in Smart Technologies (ICETST), 2020, 8138–8143. https://doi.org/10.1109/ICETST49965.2020.9080713
17. Ruder S, Ghaffari P, Breslin JG. A Hierarchical Model of Reviews for Aspect-based Sentiment Analysis. ArXiv160902745 Cs, 2016.
18. Socher R, Pennington J, Huang EH, Ng AY, Manning CD. Semi-Supervised Recursive Autoencoders for Predicting Sentiment Distributions, in: Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing. Presented at the EMNLP 2011, Association for Computational Linguistics, Edinburgh, Scotland, UK., 2011, 151–161.
19. Tang D, Qin B, Liu T. Document Modeling with Gated Recurrent Neural Network for Sentiment Classification, in: Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. Presented at the EMNLP 2015, Association for Computational Linguistics, Lisbon, Portugal, 2015, 1422–1432. https://doi.org/10.18653/v1/D15-1167
20. Turney P. Thumbs Up or Thumbs Down? {S}emantic Orientation Applied to Unsupervised Classification of Reviews. Comput. Res. Repos. – CORR, 2002, 417–424. https://doi.org/10.3115/1073083.1073153
21. Verma S, Saini M, Sharan A. Deep sequential model for review rating prediction, in: 2017 Tenth International Conference on Contemporary Computing (IC3). Presented at the 2017 Tenth International Conference on Contemporary Computing (IC3), 2017, 1–6. https://doi.org/10.1109/IC3.2017.8284318
22. Viard T, Fournier-S'niehotta R. Augmenting content-based rating prediction with link stream features. Comput. Netw, 2018, 150. https://doi.org/10.1016/j.comnet.2018.12.002
23. Wang X, Jiang W, Luo Z. Combination of Convolutional and Recurrent Neural Network for Sentiment Analysis of Short Texts, in: Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers. Presented at the COLING 2016, The COLING 2016 Organizing Committee, Osaka, Japan, 2016, 2428–2437.
24. Why Product Reviews are Important for Buyers and Sellers [WWW Document], 2018. ShipStation. URL https://www.shipstation.com/blog/product-reviews-important-buyers-sellers/ (accessed 1.17.22).
25. Why ratings and reviews are important for your business [WWW Document], 2020. Bazaarvoice. URL https://www.bazaarvoice.com/blog/why-ratings-and-reviews-are-important-for-your-business/ (accessed 1.17.22).
26. Zhang B, Zhang H, Sun X, Feng G, He C. Integrating an Attention Mechanism and Convolution Collaborative Filtering for Document Context-Aware Rating Prediction. IEEE Access,2019:7:3826–3835. https://doi.org/10.1109/ACCESS.2018.2887100
27. Zhang M, Zhang Y, Vo DT. Gated neural networks for targeted sentiment analysis, in: Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, AAAI'16. AAAI Press, Phoenix, Arizona, 2016, 3087–3093.