



DEEP LEARNING ALGORITHM FOR SENTIMENT ANALYSIS OF E-COMMERCE

¹Dr.B.Lakshma Reddy, ²M.R.Sumathi, ³P.Nanthini, ⁴Laxmi Narayan Soni,

¹Professor, ²PG Scholar, ³Assistant Professor, ⁴PhD Scholar,

^{1,2,3,4}Department of Computer Science and Engineering,

¹Rajarajeswari College of Engineering, ²AVS Engineering College, ³Kangeyam Institute of Technology, ⁴AKS University,

¹Bangaluru, Karnataka, ²Ammamet, Salem, ³Tiruppur, ⁴Satna, Madhya Pradesh.

Abstract

Sentiment evaluation of e-commerce critiques is the warm subject in the e-commerce product high-quality management, from which producers are in a position to research the public sentiment about merchandise being offered on e-commerce websites. Meanwhile, clients can be aware of different people's attitudes about the equal products. This paper proposes the deep gaining knowledge of mannequin of Bert-BiGRU-Softmax with hybrid masking, assessment extraction and interest mechanism, which applies sentiment Bert mannequin as the enter layer to extract multi-dimensional product characteristic from e-commerce reviews, Bidirectional GRU mannequin as the hidden layer to gain semantic codes and calculate sentiment weights of reviews, and Softmax with interest mechanism as the output layer to classify the fine or poor nuance. A sequence of experiments are carried out on the large-scale dataset involving over five hundred thousand product reviews. The consequences exhibit that the proposed mannequin outperforms the different deep studying models, which include RNN, BiGRU, and Bert-BiLSTM, which can attain over 95.5% of accuracy and keep a decrease loss for the e-commerce reviews.

Customers of e-commerce systems alternate their ideas with such types of languages. In the age of the current aggressive commercial enterprise world, sentiment evaluation is broadly used in the e-commerce enterprise to enhance effectivity and higher apprehend to make enterprise decisions.

Keyword: Sentiment analysis, e-commerce, deep learning, products

1. Introduction

Data as Negative, Positive, Neutral, Slightly Negative, and Slightly Positive sentiment

Web three has the primary points of the semantic web, synthetic intelligence, connectivity, etc., permitting humans to use social media to speak and specific their opinions about real-world events. In this context, the analysis of users' evaluations is quintessential for agencies to develop worldwide. This makes opinion mining a key participant in the evaluation of evaluations and discussions. Nowadays, businesses analyze this kind of statistics to enhance the best and overall performance of the merchandise and, consequently, continue to exist in a aggressive market.

Opinion mining can be expressed as the cause at the back of any motion or motion that humans use to comply with the motive [1]. Within the massive quantity of records generated on the Internet, vital records is hidden.

Data mining methods are used to extract statistics and resolve a range of problems. Online product evaluations have two vital components below which information are saved on the Internet. Commercial web sites are structures the place customers categorical their sentiment or opinion on countless topics. Sentiment evaluation refers to a huge location of herbal language processing (NLP), computational linguistics, and textual content mining [2]. Theuse of these strategies leads to the extraction and evaluation of the opinion on a given product. Opinion mining defines an opinion as nice or negative, and sentiment evaluation defines the polarity cost of a user's opinion on a specific product or service. The modern tactics of sentiment evaluation are frequently [3] computing device gaining knowledge of algorithms [4], lexicon-based methods, [5] and hybrid models [6, 7].

Negation is a widely wide-spread morphological improvement that influences polarity and, therefore, need to be mirrored in the evaluation of sentiment. Automatic detection of negation in news articles is required for severa textual content processing applications, which include sentiment analysis. Here, we explored the function and significance of users' opinions regarding unique merchandise on the choice the usage of sentiment analysis. We current experimental effects that reveal that sentiment evaluation is fantastic to this end.

The present simple tactics are statistical, primarily based on frequencies of fantastic and poor words. Recently, researchers observed new methods to account for different components of content, such as structural or semantic features. The current work focuses on the identification of documentlevel negation with the aid of the usage of more than one computational methods. In current years, with the exponential boom of smartphone use, many human beings are linked to social networking platforms, like Facebook, Twitter, and Instagram. Social networks have come to be a subject to categorical beliefs or opinions, emotions, thoughts, non-public issues, places, or personalities.

There are severa research making use of sentiment analysis, some of which used real-time information from Twitter for extracting patterns through using the Twitter-streaming utility programming interface (API) [8, 9]. The sentiment analyzers are divided into two types: SentiWordNet [10] and WordNet [11]. Sentiment evaluation makes use of effective and poor ratings to classify opinions. By growing a mannequin to analyze phrase sequence disambiguation [12], the Twitter streaming API used to be used to collect information regarding the Indonesian presidential elections [13]. Needless tweets have been removed, and the closing statistics have been investigated for sentimental components with the aid of dividing every tweet into severa sub-tweets and calculating the sentiment polarity of the subtweets for predicting the outcome of the elections. The imply absolute error metric was once used to consider the results, it cited that the prediction error was once 0.6 higher than the preceding find out about [14]. To predict the Swedish election result with the aid of the use of Twitter data, a gadget used to be developed [15]. To predict the consequence of the European elections, a new approach was once designed that studied the similarity of the shape with the effect of the vote. Another approach was once created to take a look at Brazilian municipal el

In this methodology, sentiment evaluation was once utilized alongside with a stratified pattern [17] of customers to evaluate the traits of the findings with the proper voters.

Many researchers have used computing device getting to know and synthetic brain to analyze the sentiment of tweets [18, 19]. In [20], the Naive Bayes, aid vector desktop (SVM) [21], and statistics entropy-based [22] fashions have been utilized to classify product reviews. A hybrid computer getting to know algorithm primarily based on Twitter opinion mining used to be proposed in [23]. Heydari et al. [24] proposed time sequence mannequin for fraudulent sentiment reviewer analysis. Hajek et al. [25] developed a deep feedforward neural community and convolution mannequin to realize faux advantageous and bad overview in an Amazon dataset. Long et al. [26] utilized LSTM with multi-head interest community for predicting sentimentbased textual content the use of China social media dataset. Dong et al. [27] proposed supervised computing device linear regression for predicting sentiment of clients introduced in on line purchasing facts the use of sentiment evaluation gaining knowledge of approaches.

Some common approaches, which are assisted by means of desktop gaining knowledge of techniques, are based totally on elements of the used language. Using the area of film opinions, Pang et al. [18] studied the overall performance of a range of computer gaining knowledge of algorithms, together with Naive Bayes, most entropy, and SVMIn this aspect, the majority of NLP techniques are situated on the utilization of n-grams however the use of a bag-of-words approach is additionally frequent [38, 39]. Numerous research have

validated great effects when using the bag-of-words as a textual content illustration for object categorization [40–44].

Researchers have taken benefit of NLP issues to enhance deep getting to know fashions based totally on neural networks with greater than three layers, in accordance to the journal Nature. Most of these research observed that deep mastering fashions precisely realize sentiment in quite a number situations. The CNN [45, 46], RNN [47], deep neural community [48], recursive neural deep mannequin [49], and the attention-based bidirectional CNN-RNN [50] fashions are some consultant examples. Some researchers mix models, which are then referred to as hybrid neural networks.

The main contributions of the proposed research are the following:

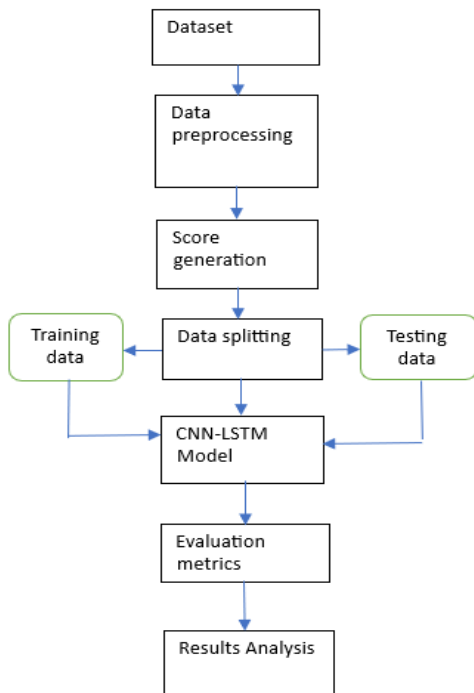


Figure 1: Framework for the proposed methodology.

- (1) Labeling the review texts as negative if the generated sentiment score is <0 or positive if the score is >1 .
- (2) The combination of all product reviews into a single data frame to obtain more sentiment-related words.
- (3) Improving the accuracy by developing a hybrid deep learning model combining the CNN and LSTM models for the product-related sentiment classification.
- (4) Comparing the classification performance of the CNN-LSTM and LSTM models.

2. Materials and Methods

The proposed methodology for predicting the review-related sentiments is based totally on the deep mastering algorithms introduced here. The phases of the proposed machine are the following: dataset collection, information preprocessing, producing the sentiment score, polarity calculation, making use of the CNNLSTM model, contrast metrics, and evaluation of the results. Figure 1 indicates the framework of the proposed methodology used in the current study.

2.1. Datasets. To consider the proposed system, the dataset [53] was once gathered from opinions on the Amazon internet site in JSON file format. Each JSON file involves a range of critiques (Table 1). The dataset consists of opinions of laptops, cellular phones, tablets, televisions, and video surveillance products. The records preprocessing consists of a number of steps, such as lowercase processing with meta-features like the reviewer's ID, the product ID, and the assessment text.

2.2. Data Preprocessing.

We implemented different preprocessing steps aiming at cleaning the review texts so that

Product name	Review count
Smartphones	1,946
Mobile phones	1,918
Tablets	1,894
Televisions	1,596
Video surveillance products	1,597

Table 1: The number of reviews per product category.

they are easy to process. The following preprocessing methods were performed on the dataset as a whole.

2.2.1. Lowercase. It entails converting whole words of the review text into lowercase words.

2.2.2. Stopword Removal. Stopwords are extensively used phrases in a language, such as “the,” “a,” “an,” “is,” and “are”. As these phrases do no longer lift any facts huge for the model, they have been eliminated from the content material of the review.

2.2.3. Punctuation Removal. All punctuation marks in the overview texts have been removed.

2.2.4. One-Word Review Elimination. Reviews that protected solely one phrase have been eliminated.

2.2.5. Contraction Removal. This system replaces a phrase in the beginning written in the quick shape with the respective full form; for instance, “when’ve” will become “when have.”

2.2.6. Tokenization. Each sentence of the assessment texts was once divided into small portions of phrases or tokens.

2.2.7. Part-of-Speech Tagging. This step is used to tag every phrase current in the sentence with a POS tag, for example, “VB” for a verb, “AJJ” for an adjective, and “NN” for a noun.

2.2.8. Score Generation. The assessment textual content was once evaluated for sentiment, and a rating used to be generated. For calculating the sentiment score, the dataset used to be matched with opinion lexicon [53] that consists of 5,000 fantastic phrases and 4,500 bad phrases with their respective scores. The sentiment rating used to be calculated for every overview textual content primarily based on the ratings of the lexicon. The evaluate textual content was once labeled as fine if the rating was once >0 ; otherwise, it used to be labeled as negative.

2.2.9. Word Embeddings. We calculated numerical vectors with each preprocessed sentence in the product assessment dataset the use of the “Word embeddings” method. To create phrase indices, we first grew to become all of the evaluate textual content phrases into sequences. The Keras textual content tokenizer [54] is being used to acquire these indices. We made positive that no time period or phrase receives a zero index in the tokenizer, and that the vocabulary measurement is adjusted properly. Then, for every single phrase in the education and trying out sets, a one of a kind index is generated,



Figure 2: The structure of the CNN-LSTM model.

which is employed to create numeric vectors of all evaluate texts of the dataset.

2.3. The CNN-LSTM Model. Figure two provides the structure of the CNN-LSTM mannequin used for sentiment classification of customers' opinions the use of an Amazon dataset.

2.3.1. Embedding Layer. This is the preliminary layer of the CNNLSTM mannequin that is used to seriously change every phrase in the coaching dataset into an actual-valued vector, which means that a set of sentiment-related phrases are developed and converted into a numerical form. This technique is acknowledged as phrase embedding. The embedding layer consisted of three components: the vocabulary measurement (maximum features; 15,000 words), the embedding dimensions (50), and the enter sequence size (400 words).

2.3.2. Dropout Layer. The major venture of this layer is to keep away from the overfitting of the mannequin [52]. Here, we assigned the price 0.4 to the dropout price parameter, the place this fee has a vary between zero and 1. The principal characteristic of the dropout layer is to arbitrarily deactivate a set of neurons in the embedding layer, the place each and every neuron denotes the dense exemplification of a sentiment phrase in a evaluate text.

CNN is a deep mastering method used in exceptional areas such as herbal language preprocessing tasks, pc vision, and scientific photograph processing.

2.3.3. Convolution Layer. The 1/3 layer of the CNN-LSTM mannequin is used for the extraction of points from the enter matrix. It makes use of n convolution filters that function over the factors of the enter sequence matrix to locate the convolutions for every sequence. We set the quantity of filters to sixty four and the dimension of the filter kernel to three \times 3.

2.3.4. Max Pooling Layer. This layer performs downsampling beside the spatial dimensionality of the given enter sequences. It considers the most fee of all enter aspects in the pool of every filter kernel. It has assigned to five \times five kernel.

2.3.5. LSTM Layer. LSTM is a kind of RNN succesful of getting to know long-term dependence [52]. We used an LSTM layer and assigned it to 50 hidden gadgets towards the subsequent layer. One of the most extraordinary blessings of using a convolutional neural community as function extraction approach past a normal LSTM is the discount in the aggregating quantity of features. Throughout the function extraction process, a sentiment classification model makes use of these elements (words) for prediction of the product assessment textual content as superb or bad sentiment. LSTM executes precalculations for the enter sequences earlier than presenting an output to the closing layer of the network. In each cell, 4 discrete computations are carried out based totally on 4 gates: enter (it), forget about (ft), candidate (ct), and output (ot). The shape of the LSTM mannequin is introduced in Figure three The equations for these gates are as follows: toes sig $\square Wfxt + Ufht - 1 + bf\square$, it sig $Wixt + Uiht - 1 + bi\square$,

Ot sig $Woxt + Uoht - 1 + bo\square$, $c \sim t \tanh wcxt + Ucht - 1 + bc\square$, (1)

Ct $f\text{otc} - 1 + itoc \sim t\square$, $ht \text{Oto} * \tanh\square Ct$,

$1 - e^{2x} \tanh(x) 2x$, $1 - e$

where sig and tanh are the sigmoid and tangent activation functions, respectively, X is the enter data, W and b signify the weight and bias factor, respectively, Ct is the mobilephone state, $c \sim t$ is the candidate gate, and ht refers to the output of the LSTM cell.

2.3.6. Dense Layer (Fully Connected Layer). This is a hidden layer in the CNN-LSTM model. It consists of 512 synthetic linked neurons that join all neurons of the network. The characteristic utilized to this layer is the rectified linear unit described by using the following equation:

In this methodology, sentiment evaluation was once utilized alongside with a stratified pattern [17] of customers to evaluate the traits of the findings with the proper voters.

Many researchers have used computing device getting to know and synthetic brain to analyze the sentiment of tweets [18, 19]. In [20], the Naive Bayes, aid vector desktop (SVM) [21], and statistics entropy-based [22] fashions have been utilized to classify product reviews. A hybrid computer getting to know algorithm primarily based on Twitter opinion mining used to be proposed in [23]. Heydari et al. [24] proposed time sequence mannequin for fraudulent sentiment reviewer analysis. Hajek et al. [25] developed a deep feedforward neural community and convolution mannequin to realize faux advantageous and bad overview in an Amazon dataset. Long et al. [26] utilized LSTM with multi-head interest community for predicting sentiment based textual content the use of China social media dataset. Dong et al. [27] proposed supervised computing device linear regression for predicting sentiment of clients introduced in on line purchasing facts the use of sentiment evaluation gaining knowledge of approaches.

Researchers have been focusing on growing effective fashions to deal with the eer-increasing complexity of large statistics [28, 29], as properly as increasing sentiment evaluation to a vast vary of purposes [30, 31], from economic forecasting to advertising techniques [32] amongst different areas [33, 34]. However, solely a few of them analyzed one of a kind deep getting to know techniques to supply actual proof of their overall performance [35]. Deep gaining knowledge of strategies are turning into more and more popular. When assessing the overall performance of a single method on a single dataset in a precise area, the outcomes advise that CNN and RNN have fantastically appropriate accuracy.

3.2. Results and Discussion.

It shows the results of the deep learning approaches. The CNN-LSTM model achieved high accuracy (96%).

The confusion matrix is used to current the costs of TP, FP, TN, and FN of the sample. Based on these rates, the contrast metrics (specificity, accuracy, recall, precision, and F1-score) had been calculated to consider the CNN-LSTM mannequin the usage of unseen facts to predict the sentiment of customers. LSTM resulted in 82.24% TP, whilst CNN-LSTM resulted in 83.54% TP. As for misclassification, LSTM resulted in 6.39% FP and CNN-LSTM in 5.28% FP, indicating that the CNN-LSTM mannequin was once barely higher than the LSTM model.

The accuracy overall performance of LSTM for the education and validation datasets is introduced in Figure 6. The LSTM mannequin introduced growing accuracy for the duration of the education section (from 86% to 94%), whereas in the trying out phase, it carried out 91% accuracy with 10 epochs. The loss of the LSTM mannequin in the education section diminished from five to 0.35, whilst in the validation phase, the mannequin loss lowered from 0.3 to 0.27.

The accuracy overall performance of the CNN-LSTM throughout the coaching segment expanded from 87.50% to 97%. In the validation phase, the accuracy overall performance reached 94% (Figure 7(a)). The loss of the CNN-LSTM mannequin in the validation section used to be 0.20 (Figure 7(b)).

The dataset developed by using Rajkumar et al. [53] proposed SVM and Naive Bayes techniques to predict sentiment analysis. They accrued statistics from Amazon regarding cell phones, tablets, cameras, and televisions. They utilized the SVM approach to every dataset individually. Here, we utilized deep studying fashions to all the datasets combined. The empirical consequences of our device had been in contrast with the outcomes of [28] and are proven in Table four The CNN-LSTM mannequin accomplished an accuracy of 94%.

However, solely a few of them analyzed one of a kind deep getting to know techniques to supply actual proof of their overall performance [35]. Deep gaining knowledge of strategies are turning into more and more popular. When assessing the overall performance of a single method on a single dataset in a precise area, the outcomes advise that CNN and RNN have fantastically appropriate accuracy.. It makes use of n convolution filters that function over the factors of the enter sequence matrix to locate the convolutions for every sequence. We set the quantity of filters

Conclusion

Recently, sentiment evaluation has grow to be a precious device for the technology and contrast of distinct sorts of data, supporting the decision-making strategies that lead to the enchancement of groups and companies. Social networking creates a giant quantity of information that require processing and evaluation to gain applicable insights. In the current study, the experimental dataset was once amassed from the Amazon internet site and blanketed critiques of laptops, cellular phones, tablets, televisions, and video surveillance products. The lexicon-based strategy was once used for the calculation of the sentiment rating for every evaluation text.

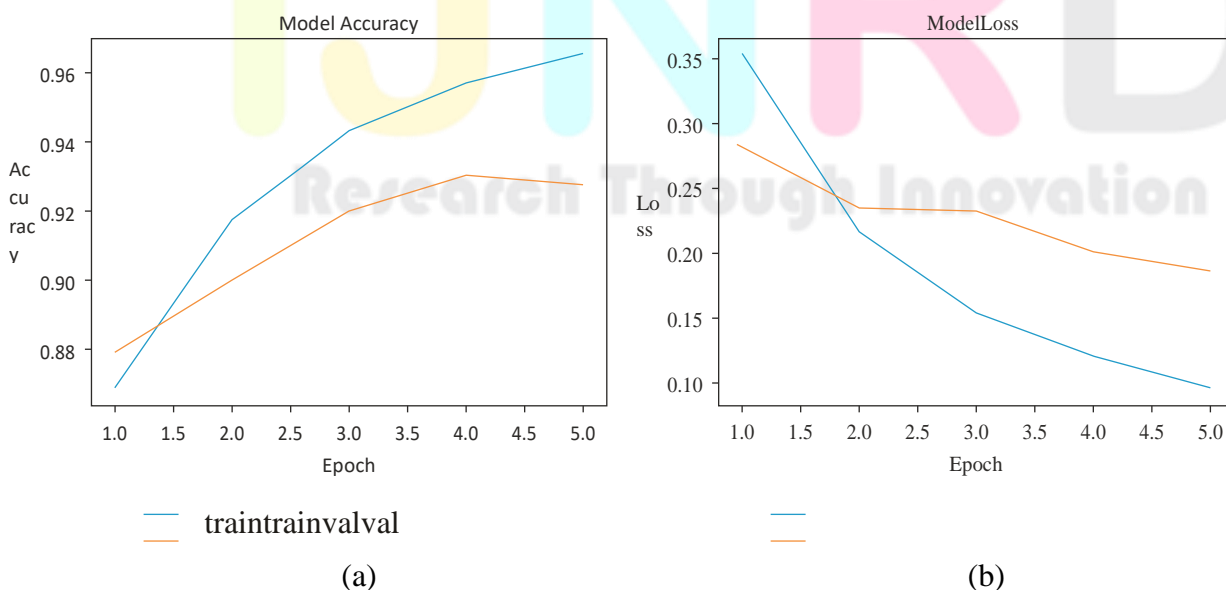


Figure 7: The performance of the CNN-LSTM model: (a) accuracy and (b) loss.

Table 4: Significant results of the CNN-LSTM model compared to the SVM method.

Models	Datasets	Accuracy (%)	References
Support vector machine [28]	Televisions, tablets, mobile phones, laptops, and video surveillance	88, 84, 92, 88, and 93	[1] Cambridge and
Proposed system (CNN-LSTM)	11 dataset	94	

D. S. A. Das, "Affective computing and sentiment analysis," *A Practical Guide to Sentiment Analysis*, vol. 31, pp. 1–10, 2017.

[2] O. Jagdale, V. Harmalkar, S. Chavan, and N. Sharma, "Twitter mining using R," *Int. J. Eng. Res. Adv. Tech.* vol. 3, pp. 252–256, 2017.

[3] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: a survey," *Ain Shams Engineering Journal*, vol. 5, no. 4, pp. 1093–1113, 2014.

[4] F. Sebastiani, "Machine learning in automated text categorization," *ACM Computing Surveys*, vol. 34, no. 1, pp. 1–47, 2002.

[5] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexicon-based methods for sentiment analysis," *Computational Linguistics*, vol. 37, no. 2, pp. 267–307, 2011.

[6] R. Prabowo and M. Thelwall, "Sentiment analysis: a combined approach," *Journal of Informetrics*, vol. 3, no. 2, pp. 143–157, 2009.

[7] Y. Dang, Y. Zhang, and H. Chen, "A lexicon-enhanced method for sentiment classification: an experiment on online product reviews," *IEEE Intelligent Systems*, vol. 25, no. 4, pp. 46–53, 2010.

[8] R. Jose and V. S. Chooralil, "Prediction of election result by enhanced sentiment analysis on twitter data using word sense disambiguation," in *Proceedings of the 2015 International Conference on Control Communication & Computing India (ICCC)*, pp. 638–641, Trivandrum, India, November 2015.

[9] A. Esuli and F. S. Sebastiani, *A High-Coverage Lexical Resource for Opinion Mining*, Institute of Information Science and Technologies (ISTI) of the Italian National Research Council (CNR), Pisa, Italy, 2006.

[10] G. A. Miller, "WordNet," *Communications of the ACM*, vol. 38, no. 11, pp. 39–41, 1995.

[11] R. Navigli, "Word sense disambiguation," *ACM Computing Surveys*, vol. 41, no. 2, pp. 1–69, 2009.

[12] M. Ibrahim, O. Abdillah, A. F. Wicaksono, and M. Adriani, "Buzzer detection and sentiment analysis for predicting presidential election results in a twitter nation," in *Proceedings of the 2015 IEEE International Conference on Data Mining Workshop (ICDMW)*, pp. 1348–1353, Atlantic City, NJ, USA, November 2015.

[13] N. Dokoohaki, F. Zikou, D. Gillblad, and M. Matskin, "Predicting Swedish elections with twitter: a case for stochastic link structure analysis," in *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pp. 1269–1276, Paris, France, August 2015.

[14] D. Liben-Nowell and J. Kleinberg, "The link-prediction problem for social networks,"