DOI: https://doi.org/10.52756/ijerr.2024.v43spl.011 *Int. J. Exp. Res. Rev.*, Vol. 43: 146-159 (2024)

Original Article | Peer Reviewed | Open Access

International Journal of Experimental Research and Review (IJERR)
© Copyright by International Academic Publishing House (IAPH)
ISSN: 2455-4855 (Online)
www.iaph.in

Aspect-based sentiment analysis of Twitter mobile phone reviews using LSTM and Convolutional Neural Network

Nitendra Kumar[1], Ritu Talwar[2], Sadhana Tiwari[3]* and Priyanka Agarwal[4]

[1]Department of Decision Sciences, Amity Business School, Amity University, Noida-201303, Uttar Pradesh, India; [2]Department of Human Resource Management, NDIM, New Delhi -110062, India; [3]Department of Operations and Supply Chain Management, GNIOT Institute of Management Studies, Greater Noida-201310, Uttar Pradesh, India; [4]Department of Human Resource Management, Amity Business School, Amity University, Noida-201303, Uttar Pradesh, India

E-mail/Orcid Id:

NK, nkshukla.kumar@gmail.com, https://orcid.org/0000-0001-7834-7926; *RT*, ritu@ndimdelhi.org, https://orcid.org/0000-0002-8617-4479; *ST*, sadhanatiwari.sun@gmail.com, https://orcid.org/0000-0003-3786-3398; *PA*, pagarwal5@amity.edu, https://orcid.org/0000-0002-2943-2507

Article History:
Received: 11th Apr., 2024
Accepted: 21st Sep., 2024
Published: 30th Sep., 2024

Keywords:
LSTM, Recurrent neural network, Convolutional Neural Network, Sentiment analysis, Artificial intelligence, Fuzzy logic, Mobile phone reviews

How to cite this Article:
Nitendra Kumar, Ritu Talwar, Sadhana Tiwar and Priyanka Agarwa (2024). Aspect-based sentiment analysis of Twitter mobile phone reviews using LSTM and Convolutional Neural Network. *International Journal of Experimental Research and Review*, *43*, 146-159.
DOI: https://doi.org/10.52756/ijerr.2024.v43spl.011

Abstract: The proliferation of online shopping has led to a surge in product reviews, providing valuable information to consumers. However, the overwhelming volume and subjective nature of these reviews make it difficult to assess product performance accurately. We propose a machine learning-based system that extracts sentiment from online reviews to address this challenge. Our system effectively identifies positive, negative, and neutral sentiments and classifies sentiment for specific product aspects. By offering concise and clear information, our system empowers consumers to make informed purchasing decisions and assists manufacturers in improving their products. Our proposed LSTMCNN model, trained on a dataset of 62,563 reviews, achieved impressive results with an accuracy of 95.84%, precision of 95.6% and recall of 95.8%. This significantly outperforms existing models, demonstrating the effectiveness of our approach. Moving forward, we aim to further enhance the accuracy of our system, track sentiment changes over time, and develop personalized product recommendations. These advancements will continue to increase the value and utility of online reviews in e-commerce.

Introduction

The evaluation of products through customer feedback plays a critical role in enabling informed decision-making for other consumers (7 Reasons Why Customer Feedback Is Important To Your Business - Startquestion - Create Online Surveys and Forms, n.d.). Understanding the advantages and disadvantages of a product before making a purchase helps eliminate uncertainties and ensures a satisfying buying experience (*Understanding Customer Experience*, n.d.). However, conducting a comprehensive analysis of all aspects of a product can be time-consuming, given the wide range of brands, specifications, discounts, and attractive offers available in the market. Moreover, the abundance of opinionated content generated on social media platforms further complicates the decision-making process for customers, who often lack the time to read through all their opinions. Unfortunately, customers don't have the luxury of time to sift through numerous comments before making a purchase. Thus, customers seek a procedure that can effectively analyze a large number of comments on the same platform and consider the different aspects of the product, and present a brief summary of the pros and cons.

*Corresponding Author: sadhanatiwari.sun@gmail.com

A Aspect-based sentiment analysis (ABSA) is a very challenging task in the domain of natural language processing (NLP) aimed at determining the sentiment expressed towards various elements or features of a product or service ABSA is more challenging than traditional sentiment analysis, which is the typical classification of text under positive, negative, or neutral. As opposed to traditional sentiment analysis, here ABSA has to recognize the sentiment and, for every mention of an aspect, determine and analyze the associated sentiment. This is a much more difficult task, mainly because, in many cases, the aspects are implicit in the text and never mentioned directly. ABSA has been approached with various methods, such as rule-based systems, machine learning systems, and hybrid systems (Elahi et al., 2023; J and U, 2023). The rule-based system detects aspects and extracts sentiment by applying some predefined rules, while machine learning systems base the connection between aspects and sentiment on a statistical model. Hybrid systems combine rule-based and machine-learning approaches, enabling better quality in the analysis.

ABSA has applications across product recommendation (Elahi et al., 2023), social media monitoring (Asif et al., 2020; Lyu and Takikawa, 2022; Melton et al., 2021), and customer feedback analysis (Zou et al., 2022). Therefore, by understanding the sentiments of customers toward specific dimensions of their products or services, businesses can make informed improvements, enhance customer satisfaction, and improve market targeting. In this way, ABSA allows businesses to glean critical insights from customer opinions and align their offerings with customer preferences to drive growth and success.

Related Work

In modern world, aspect-based sentiment analysis is quite important in research. It helps us understand the emotions and thoughts associated with several items. With the help of AI techniques, consumers can make smarter choices while choosing the right products. In the next section, we explore the several AI techniques used in guided analysis to help consumers make informed decisions in their own best concern and the sentiment architecture is mentioned in figure 1.

Ananthajothi et al. (2021) presented a unique approach to aspect-based sentiment analysis (ABSA) for discussions of truth and uncertainty in data that includes demonetization-related keywords in India. The research considers the limitations in previous models using deep neural network-based classifiers and improves the selection of advantages with Adaptive Insect Swarm Optimization (SA-BSO) algorithm. Better performance metrics were achieved compared to the baseline innovation. The accuracy produced using the SA-BSO-RNN method increased significantly by 2.82% to 6.9% above the other methods such as FF -RNN, PSO-RNN, CSA-RNN, and BSO-RNN. The F1 score improved significantly ranging from 3.5% to 8.6%. Thus, findings demonstrate improvement reliability of the SA-BSO algorithm for sentiment classification in work and tweeting related specifically to demonetization discussions in India. Plays an important aspect. Pre-processing steps of removing abandoned messages, tokenization, low-code transformation, and stemming while ensuring quality and accuracy. As well, polarizing scores were calculated from the actual terms in the tweets using the Valence Awareness Dictionary and Affective Intensity Analyzer (Vader), where polarizing terms are scored, but the machine learning algorithms are not term reversed. Additionally, Word2Vec features are designed to determine semantic and syntactic resemblance between the words. Catelli et al. (2023) performed a thorough investigation into sentiments and emotions regarding COVID-19 in Italy using vaccination tweets where sentiments were categorized into four positions for analysis, participants as general users, advertising, medicine, or political related influencers. The findings indicate negative feelings along a timeline, particularly among the bulk of users. The analysis also identified distinct attitudes towards certain events (e.g., post-vaccine mortality) throughout the 14-month study period, possibly translatable into like sentiments. This informs ongoing dialogues and enables certain key actors in the dialogue to be identified via large datasets and specialized ledgers. This understanding is vital what the world needs from policy, practice, and communication experts to formulate intervention, action, and efficacy approach plan and solution. Thoughts. However, limitations of the study to remember about this study is timeframe and Twitter, recognizing the study may not fully represent the population on Twitter, limitation represent a priori. Future studies could begin to include other media and sometimes controlled for demographic factors. Cheng et al. (2022) developed new method, CFM-MHN method using absorptive analysis (ABSA). The CFM-MHN method introduced three modules: continuous content module, component focus module, and multi-head coordination module. Content extension employs the standard Bidirectional Encoder Representations from Transformers (BERT) to encode the sentences' content, including visual and

DOI: https://doi.org/10.52756/ijerr.2024.v43spl.011

environmental representation in content reconstruction. Objective analysis uses self-monitor methodology and Thiessen analysis to monitor key messages in terms of directionality and significance of content. Head integration employs multiple heads to evaluate the association between concept, concepts, and emotional polarization, ultimately analyzing the data here and will be presented. The evaluations here, the CFM-MHN method is a promising new method for employing ABSA. Huang et al. (2022) developed a novel ABSA model, LTNMR, which combines Logical Tensor Networks (LTN) and large-scale learning (Hu et al., 2021; Huang et al., 2022) semantic deeper learning. It surpasses state-of-the-art on Twitter, Lap14, and Rest14 datasets. Twitter data contains 1561 positive, 3127 neutral, and 1560 negative training tweets, with 692 for testing. Lap14 comprises 2328 training and 638 testing laptop reviews with three polarity classes. Rest14 is from SemEval-14 Task 4.

Chakraborty et al. (2020) employed a fuzzy rule-based approach for sentiment analysis of COVID-19 tweets. Using Doc2Vec (DBOW, DMC, DMM), they analyzed a dataset of approximately 23,000 retweets. Logistic Regression on DBOW+DMC yielded the best results. The model achieved up to 79% accuracy. The paper underscores the importance of monitoring social media for negative sentiment during crises like pandemics.

Sunitha et al. (2022) performed a sentiment analysis model for COVID-19 tweets, achieving 97.28% accuracy for Indian and 95.20% for European data. Huang et al. (2023) proposed AGSNP for aspect-level sentiment classification. Combining a modified GSNP with attention, AGSNP effectively processes content and aspect words. It employs two channels and modified GSNPs to capture dependencies. Evaluated on Restaurants, Laptops and Twitter datasets, AGSNP categorizes sentiment as positive, neutral, or negative.

AGSNP outperforms other models in classifying theories. Xiang et al. (2023) proposed a grammar recognition method using sentence block-level progression to capture concept connections. BDEP extracts concepts and analyzes similarities. The model surpasses state-of-the-art on four datasets. Sharifi and Shokouhyar (2022) analyzed 25,000 tweets on refurbished phones. Price, warranty, quality, and seller reputation are key factors affecting purchase decisions. ISM and MICMAC were used for analysis. Lv et al. (2021) introduced CAMN for aspect-level reasoning in short texts, outperforming baselines on SemEval2014 datasets. Zhao et al. (2022) proposed a GCN with multiple body weights for ABSA, improving performance by capturing relevant information. Han et al. (2023) introduced GMF-SKIA for hierarchical-focused opinion classification, achieving 91.56% accuracy on Rest16.

Žunić et al. (2021) suggested a graph for analyzing syntactic dependencies in drug reviews. Outperforming standard deep learning, the model effectively assesses sentiment in healthcare. Huang et al. (2022) introduced a hierarchical sentiment analysis model using asymmetric content weighting and bidirectional GRUs. The model surpasses other methods in accuracy and F1 score. You et al. (2022) presented ASK-RoBERTa for ABSC, capturing item-concept dependencies and emotional information. The model outperforms existing approaches on various datasets. Zhou et al. (2023) developed HD-GCN for ABSC, extracting syntactic and semantic information through GCNs. The model effectively removes noise and improves performance. Ahmed et al. (2023) proposed a decision-making method for selecting doctors based on patient preferences, combining theory-based analysis and TOPSIS. The approach effectively

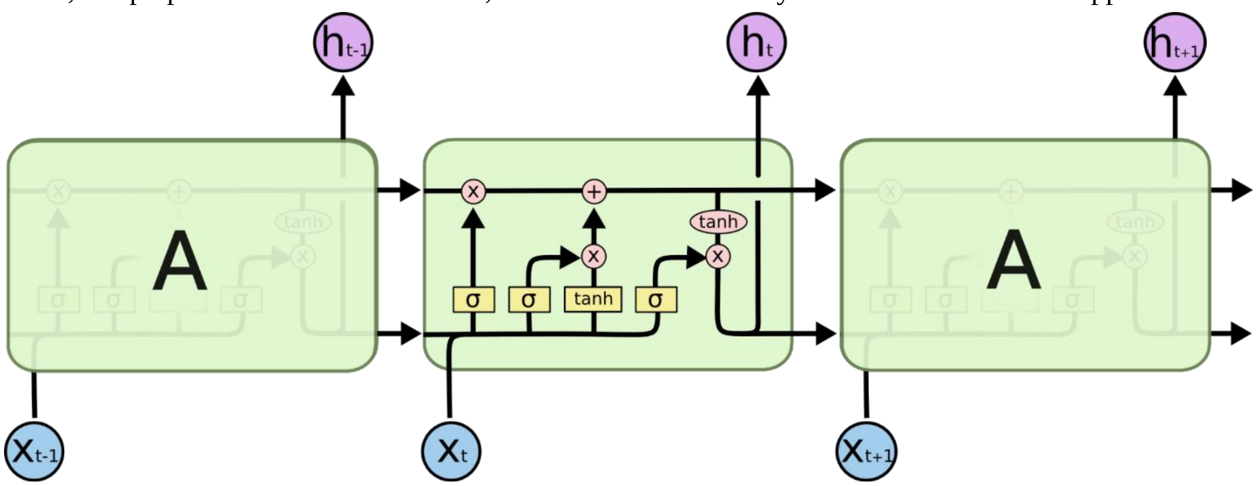

Figure 1. LSTM network architecture.

reduces data loss and outperforms existing methods. García-Méndez et al. (2023) introduced TABEA for identifying individual stock sentiments in tweets. Achieving over 90% accuracy, the model aids financial decision-making.

Dangi et al. (2023) propose an ARO-optimized RRVFLN model for emotional intelligence analysis, outperforming existing methods on three datasets. Wu et al. (2023) introduce KDGN for ABSA, combining knowledge base, text, and sentences. KDGN surpasses previous methods, highlighting the importance of domain knowledge in capturing sentiment.

Preliminaries

Deep Learning using LSTM network

LSTM networks capture text nuances, making them an effective tool for understanding sentiment towards specific aspects of products or services. The design for ABSA using LSTM networks usually consists of various components, comprising data preprocessing, word embedding, LSTM layers, attention mechanisms, and sentiment classification. In the first step, data preprocessing involves tokenising the text, removing stop words, punctuation, and special characters, and splitting the data into training, validation, and testing sets. Word embedding procedures, such as Word2Vec, GloVe, and FastText, are then used to represent words numerically, permitting the LSTM network to process them. The LSTM layer is built on the principles of structure and is designed to process data by capturing long-lived data and modeling the content of the text using storage devices that store and manipulate accurate information. Stacking multiple LSTM layers improves the network's ability to detect complex patterns. Attention systems are often included to highlight important features of the input text and the sentiments allied with them. Finally, sentiment classification predicts the sentiment polarity (positive, negative, or neutral) for respectively element, normally using a softmax layer to assign the highest possibility sentiment to each feature.

The general LSTM process can be defined mathematically as below:

Consider a time series input sequence of length T, which is denoted as $\{x_1, x_2, ..., x_t\}$.

The LSTM unit maintains a hidden state vector h_t and a memory cell vector c_t at each time step t, and these vectors are updated based on the preceding hidden state $h_{(t-1)}$, memory cell $c_{(t-1)}$, and the current input x_t.

Three main components of LSTM unit are as, input gate (i_t), forget gate (f_t), and the output gate (o_t). Each gate is responsible for managing the flow of information in and out of the memory cell.

The input gate (i_t) defines how much latest information should be stored in the memory cell and is computed as follows:

$i_t = \text{sigmoid} (W_i * [h_{(t-1)}, x_t] + b_i)$

Where, W_i represents the weight matrix for the input gate, $[h_{(t-1)}, x_t]$ is the concatenation of the previous hidden state $h_{(t-1)}$ and the current input x_t, and b_i is the bias term.

The forget gate (f_t) determines which information should be discarded from the memory cell. It is computed as:

$f_t = \text{sigmoid} (W_f * h_{(t-1)} , x_t] + b_f)$

Similarly, W_f denotes the weight matrix for the forget gate, and b_f is the bias term.

The memory cell (c_t) is updated by combining the previous memory cell value $c_{(t-1)}$ and the recently computed candidate value \tilde(c_t) and the candidate value can be computed as:

\tilde(c_t) = tanh ($W_c * h_{(t-1)}, x_t] + b_c$)

Here, W_c denotes the weight matrix for the candidate value, and b_c is the bias term.

The updated memory cell value can be calculated:

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

$o_t = \text{sigmoid} (W_o * h_{(t-1)}, x_t] + b_o)$

Where, W_o denotes the weight matrix for the output gate, and b_o is the bias term.

The hidden state (h_t), computed by applying the output gate to the memory cell value, i.e.,

$h_t = o_t * \tanh(c_t)$

TF-IDF (Term Frequency-Inverse Document Frequency)

TF-IDF is a technique used to measure the importance of terms in a collection of documents. Each examine paragraph is treated as a separate document, and a term-document matrix is created to represent the frequency of each term in each document. The matrix entry represents the frequency of a term in a given document. If the value is zero, it indicates that the term is not present in the review paragraph.

To assign numerical values to terms based on their dominance in the review, term weighting is applied. Term frequency (tf) is calculated as the ratio of a particular term's frequency (f) to the highest frequency value of any term in the document.

$tf = f / \max(f)$

The logarithmic ratio of the total number of review paragraphs or documents pertaining to a certain topic of interest to the total number of review documents in which the particular phrase appears is known as the inverse document frequency, or idf.

DOI: https://doi.org/10.52756/ijerr.2024.v43spl.011

log(N / n) = idf

where n is the number of review paragraphs that include the phrase and N is the total number of review paragraphs.

TF-IDF can provide a high-dimensional feature vector when the text corpus is huge, which could cause problems with overfitting in classification models. This may lower the model's accuracy, as TF-IDF ignores the text's semantic characteristics.

Materials and Methods

This algorithm takes a dataset of Twitter mobile phone reviews as input and outputs a trained hybrid LSTM-CNN. Here, the algorithm first preprocess the dataset by cleaning the text data, tokenizing the reviews into words, removing stop words, and performing stemming or lemmatization. The algorithm then performs aspect extraction to identify the aspects of the mobile phone being reviewed and extract aspect-specific sentences from the reviews. The algorithm then generates aspect-specific sentiment labels by labeling each aspect-specific sentence with sentiment polarity. The dataset is then split into training and testing sets. The text data is encoded using word embeddings. The hybrid LSTM-CNN model is built and trained on the training set. The model is evaluated on the testing set. The results of the model are visualized.

To retrieve mobile phone reviews from Twitter using the Twitter Application Programming Interface (API), we created a Twitter Developer Account and obtained the required API keys and access tokens. Next, we installed the tweepy library, which serves as a Python interface for interacting with the Twitter API. Once we had the API credentials, the application was authenticated by using the OAuthHandler class from tweepy. Once authenticated, we could leverage the api.search() method to search for tweets related to mobile phone reviews by specifying relevant keywords in the search query and indicating the desired number of tweets to be retrieved. The API provided a collection of tweets matching our search query. Finally, the retrieved tweets were processed for sentiment analysis.

Based on the membership function of the model, four class labels are generated: highly negative (HN), negative (N), positive (P), and highly positive (HP). Based on the specific features of the product, the suggested algorithm determines the mood associated with each statement. Customers may quickly weigh the benefits and drawbacks of a product before buying it because of this feature. Additionally, it helps producers enhance their products' qualities, boosting sales.

Pre-Processing

In the preprocessing stage, a dictionary marks subjective sentences and removes non-subjective sentences. Unmarked sentences are treated as context and discarded as empty sentences. Text normalization includes tasks such as statistical machine translation, dictionary, and spelling correction. This study used preprocessing steps such as tokenization, lowercase letters, orthographic correction, and lemmatization using dictionary maps for the analysis dataset. Tokenization breaks sentences into single words and then converts them to lowercase. Perform orthographic correction of the token, and lemmatization converts the token to its standard form (figure 2).

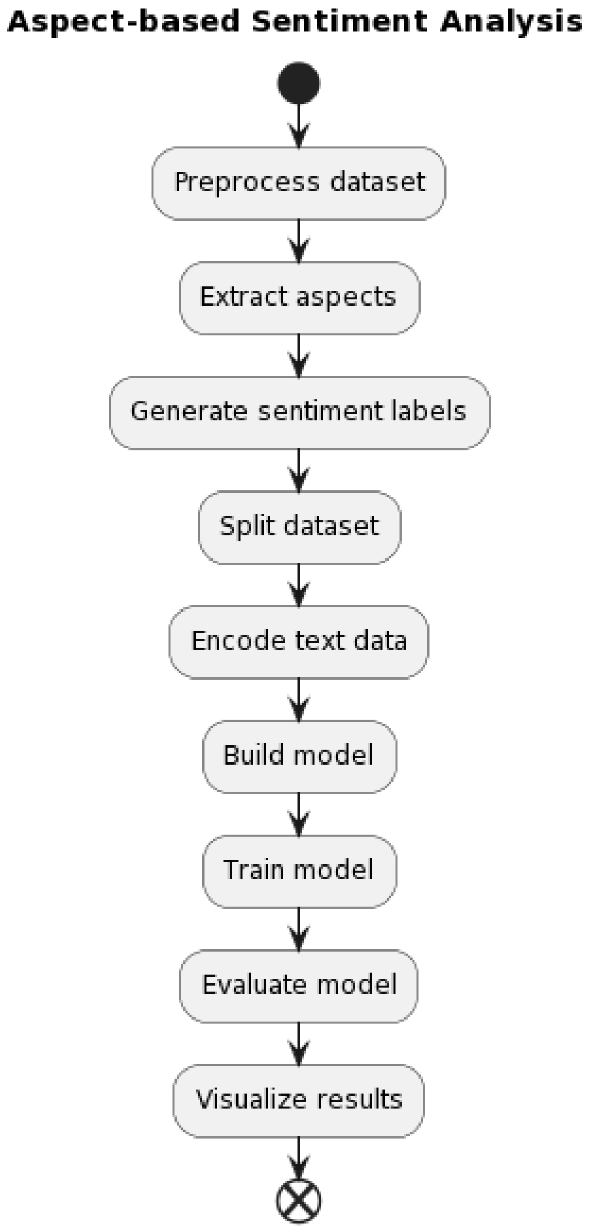

Figure 2. Aspect-based sentiment analysis of Twitter mobile phone.

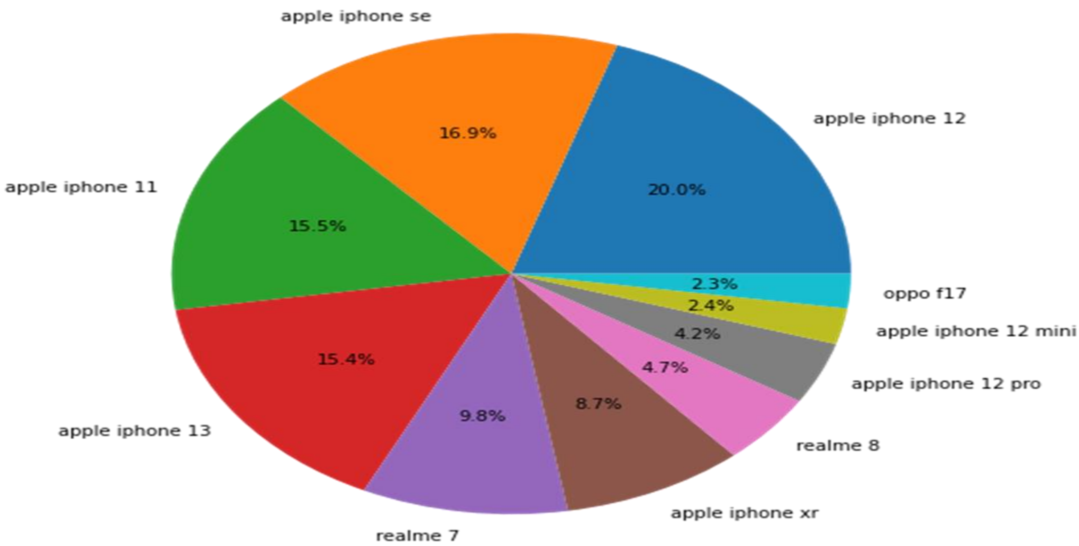

Figure 3. Twitter mobile phone data.

Sample Dataset

Realme has quietly introduced the C11 (2021) as its newest budget-friendly smartphone in India. To analyze user sentiment and preferences for this device, we employed a dataset of 62,563 online reviews, encompassing 66 unique phone names, 5,966 usernames, and 46,036 distinct tweets. Interestingly, the 'apple iphone 12' was the most frequently mentioned phone in the dataset, appearing 10,374 times. All reviews were written in English, providing a valuable resource for understanding consumer perceptions of the Realme C11 (2021) within the Indian market as shown in figure 3.

Tokenization

"Realme", "C11", "(", "2021", ")", "has", "silently", "been", "launched", "in", "India", "as", "the", "company's", "latest", "budget", "smartphone".

Convert tokens into lowercase

"realme", "c11", "(", "2021", ")", "has", "silently", "been", "launched", "in", "india", "as", "the", "company's", "latest", "budget", "smartphone".

Text normalization: spelling correction

"realme", "c11", "(", "2021", ")", "has", "silently", "been", "launched", "in", "india", "as", "the", "company's", "latest", "budget", "smartphone".

Text normalization: lemmatization (dictionary mapping)

We need to map each word to its base or dictionary form to perform lemmatization on the given tokens. Lemmatization helps in reducing inflected forms to their base or root form, which can improve the consistency and accuracy of the text.

Here's the result of lemmatization for the given tokens:

"realme", "c11", "(", "2021", ")", "have", "silently", "be", "launch", "in", "india", "as", "the", "company's", "latest", "budget", "smartphone".

As can be seen, some of the words have been transformed to their base forms, such as "has" changed to "have", "been" changed to "be", and "launched" changed to "launch". However, certain words like "realme", "c11", and "India" are not lemmatized as they are proper nouns or product names that do not have a standard base form.

Splitting long sentences

To split long sentences for aspect-based sentiment analysis using the ClausIE framework, one can leverage its ability to extract relationships and entities from sentences. “ClausIE” is a open-source Open Information Extraction (IE) tool that can classify subject-verb-object relationships and capture related information.

Here's an example of how “ClausIE” can be used to split a long sentence:

Original sentence: "The camera quality of the Realme C11 (2021) is excellent, but the battery life could be better".

ClausIE output:

01. Subject: The camera quality of the Realme C11 (2021)
    Relationship: is
    Object: excellent
02. Subject: the battery life
    Relationship: could be
    Object: better

By using “ClausIE”, one can extract the subject, relationship, and object information from the sentence, which aids in identifying the several features being discussed. In this case, we split the sentence into two parts based on the extracted information from the original

DOI: https://doi.org/10.52756/ijerr.2024.v43spl.011

sentence.

Split sentences:

01. "The camera quality of the Realme C11 (2021) is excellent."

02. "The battery life could be better."

Splitting the sentence this way allows us to analyze the sentiment associated with each aspect individually.

Feature Extraction

Extracting features from textual analysis is an important task in cognitive analysis. In the existing literature, popular feature extraction techniques include bag-of-words, TF-IDF, and word embedding. However, bag-of-words is not ideal for preserving sentence order and semantic features in aspect-based reasoning. It treats all words as unique regardless of their arrangement. Although TF-IDF outperforms Bag-of-Words, it may produce large feature vectors when processing large text files, which may cause overfitting in classification models. Moreover, TF-IDF ignores the nature of sentence analysis. Word embedding is a well-known technique that preserves the order and context of sentence analysis. It turns out to be a suitable method for extracting movies in appearance-based thinking. By utilizing word embeddings, we can capture the review's content and improve the classification model's accuracy while considering the review's features.

Word Embedding

One of the best word embedding techniques for dimension-based thinking in Twitter mobile language is Word2Vec. Word2Vec is a model based on shallow neural networks that represent a word as dense, continuous vectors in high-dimensional space. It captures semantic relationships and information content by learning word embeddings from large text corpora. The two main architectures of Word2Vec are Continuous Bag-of-Words (CBOW) and Cross-gram. Both CBOW and Skip-gram models use the softmax function to calculate the probability distribution of words for prediction purposes.

CBOW (Continuous Bag of Words) is generally faster to train and performs better when the target word appears many times in the dataset. It works well in cases where syntactic information and general context are important. We use CBOW technology to predict a word, topic or feature.

Sentiment score generation

Sentiment analysis is typically performed using rule-based methods or lexicon-based approaches. In this work, we used a lexicon-based approach to detect the polarity of a sentence. This approach involves creating a bag of words containing sentiment words and then checking for the presence of these words in a sentence using word embedding in an LSTM.

The sentences are given below:

"The camera quality of the Realme C11 (2021) is excellent."

"The battery life could be better."

Sentence: The camera quality of the Realme C11 (2021) is excellent. Polarity Score: 0
Sigmoid Score: 0.5
Sentence: The battery life could be better.
Polarity Score: 0
Sigmoid Score: 0.5
Epoch 1/10 2/2 [==============================] - 2s 17ms/step - loss: 0.6932 - accuracy: 0.0000e+00
Epoch 2/10 2/2 [==============================] - 0s 15ms/step - loss: 0.6933 - accuracy: 0.0000e+00
Epoch 3/10 2/2 [==============================] - 0s 14ms/step - loss: 0.6933 - accuracy: 0.0000e+00
Epoch 4/10 2/2 [==============================] - 0s 13ms/step - loss: 0.6932 - accuracy: 0.0000e+00
Epoch 5/10 2/2 [==============================] - 0s 12ms/step - loss: 0.6932 - accuracy: 0.0000e+00
Epoch 6/10 2/2 [==============================] - 0s 13ms/step - loss: 0.6932 - accuracy: 0.0000e+00
Epoch 7/10 2/2 [==============================] - 0s 13ms/step - loss: 0.6932 - accuracy: 0.0000e+00
Epoch 8/10 2/2 [==============================] - 0s 17ms/step - loss: 0.6932 - accuracy: 0.0000e+00
Epoch 9/10 2/2 [==============================] - 0s 13ms/step - loss: 0.6932 - accuracy: 0.0000e+00
Epoch 10/10 2/2 [==============================] - 0s 12ms/step - loss: 0.6932 - accuracy: 0.0000e+00
Sentiment Labeling using LSTM-CNN system

DOI: https://doi.org/10.52756/ijerr.2024.v43spl.011

For every input sentence, the LSTM produces an accurate polarity score. The learning process is continued until the predicted emotion score is produced. After that, the CNN is fed this score in order to classify sentences.

Epoch 1/10 2/2 [==============================] - 2s 12ms/step - loss: 0.6878 - accuracy: 1.0000
Epoch 2/10 2/2 [==============================] - 0s 11ms/step - loss: 0.6625 - accuracy: 1.0000
Epoch 3/10 2/2 [==============================] - 0s 10ms/step - loss: 0.6354 - accuracy: 1.0000
Epoch 4/10 2/2 [==============================] - 0s 11ms/step - loss: 0.6018 - accuracy: 1.0000
Epoch 5/10 2/2 [==============================] - 0s 10ms/step - loss: 0.5611 - accuracy: 1.0000
Epoch 6/10 2/2 [==============================] - 0s 12ms/step - loss: 0.5073 - accuracy: 1.0000
Epoch 7/10 2/2 [==============================] - 0s 12ms/step - loss: 0.4364 - accuracy: 1.0000
Epoch 8/10 2/2 [==============================] - 0s 11ms/step - loss: 0.3460 - accuracy: 1.0000
Epoch 9/10 2/2 [==============================] - 0s 11ms/step - loss: 0.2395 - accuracy: 1.0000
Epoch 10/10 2/2 [==============================] - 0s 11ms/step - loss: 0.1373 - accuracy: 1.0000
1/1 [==============================] - 0s 428ms/step

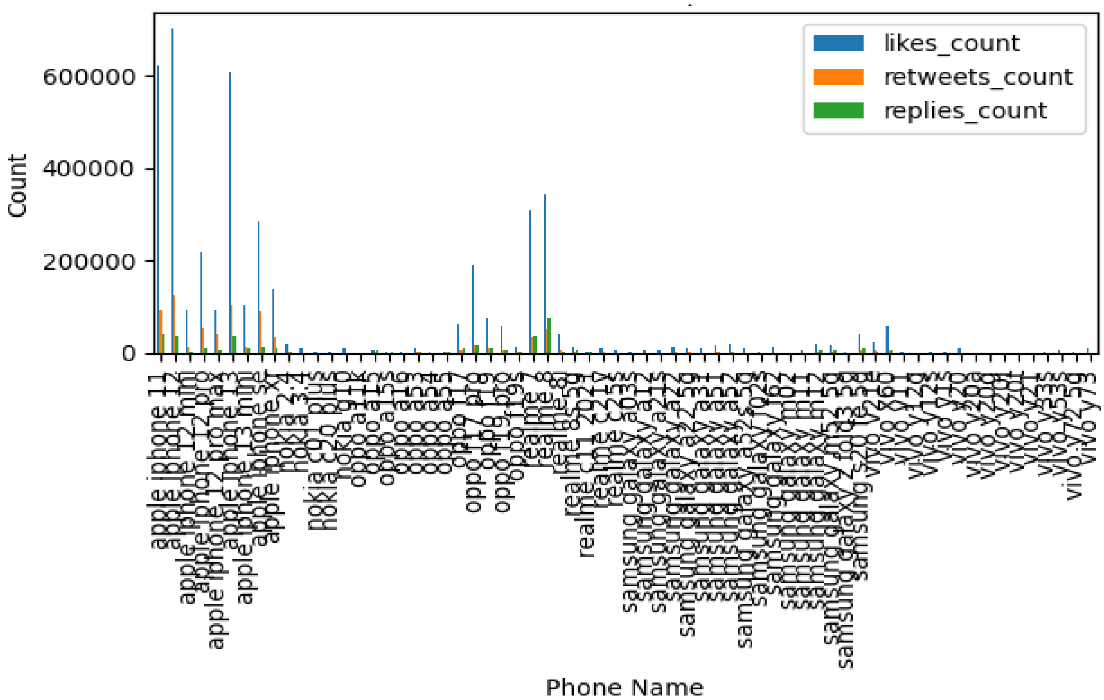

Figure 4. Number of likes, retweets and replied for each phone.

Experimental results and discussions

Dataset Description

The suggested LSTMCNN model was evaluated using the mobile phone reviews from Twitter.

The dataset contains 62,563 rows and several columns. Here's a brief summary:

The phone name column contains 66 unique phone names. The most frequently mentioned phone is the 'apple iphone 12', which appears 10,374 times. The date column contains 2,188 unique dates. The most frequent date is '2020-10-14', which appears 1,652 times. The username column contains 5,966 unique usernames. The

DOI: https://doi.org/10.52756/ijerr.2024.v43spl.011

most frequent username is 'realmecarein', which appears 4,738 times. The tweet column contains 46,036 unique tweets. The most frequent tweet is 'This MagSafe 3-in-1 Dock Costs 1/3rd Of The One Apple Sells, And Can Charge iPhone 12, Apple Watch, AirPods All At Once', which appears 167 times. The language column contains only one unique value, 'en', which means all tweets are in English. The replies_count, retweets_count, and likes_count columns contain numerical data about each tweet's number of replies, retweets, and likes.

We have also created a bar plot showing each phone name's total number of likes, retweets, and replies as mentioned in figure 4. Here's the plot:

Experimental Setup

The performance of the model was evaluated on an Intel Core i5-12400 12th Gen GPU processor with 16GB of RAM, running on the Windows 10 operating system. We used the Anaconda framework with Python 3.11 to develop and test the model. For preprocessing, we employed the WordNetLemmatizer, wordtokenize, and NumPy packages. Feature extraction was conducted using Word2Vec, pandas, and PCA libraries. The construction of the proposed LSTMCNN model was accomplished using the Keras library.

Proposed LSTMCNN model hyper parameters

The suggested LSTM CNN model performance is improved by adjusting various hyperparameters. The suggested model's learning, accuracy, and stability are affected by hyperparameters.

LSTM Layers

Input review text is converted to significant word vectors through the use of word embeddings. Experimentations during this phase were carried out to evaluate the performance of our suggested model for two sizes (64 and 128) regarding training input embeddings. The model outperformed a 64-length vector input embedding layer across three benchmark datasets. There used memory units in the model which able to memorise information from input review sentences better. In particular, the model employs 100 memory units in the LSTM component of the model that allows it to pick up on long review paragraphs. The model also had an output layer with 3 nodes generating sentiment scores for positive, negative and neutral classes. The proposed model exhibits promising performance in sentiment analysis tasks through these design choices and configurations.

Activation Function

ABSA establishes a significant challenge as it involves tackling a multilabel classification problem. To effectively address this challenge, the choice of an appropriate activation function is crucial. The sigmoid activation function shows to be well-suited for handling multilabel classification tasks, as it allows for the independent prediction of multiple sentiment labels. This does not apply for multilabel classification where a class is on or off but the sum of the predicted probabilities need not be one as in traditional classification problems. This is why the proposed model was intentionally designed to have a sigmoid activation function in the output layer. This allows the model to output individual sentiment scores for each aspect, thereby striving a better nuanced representation of what kind of sentiment needs to get extracted in terms that it should be aspect-based instead.

Number of epochs

During the training process of LSTM network, input dataset involved determining the number of iterations through the concept of epochs. The RMSE was carried out at the end of each epoch to evaluate the model's performance and track its progress. To identify the optimal training configuration, the proposed model was trained using different numbers of epochs, specifically 100, 200, 300, 400, and 500, with the corresponding RMSE values recorded. Throughout the whole analysis, it was observed that the model achieved commendable accuracy training at 500 epochs. Results show that a more extensive training duration contributes to the improved performance and efficacy of the proposed model.

Batch size

The batch size plays an important role during the training of LSTM network, and it determines the number of input review sentences processed at each iteration. In the setting up of this proposed model, experiments were conducted by training the LSTM with varying batch sizes, specifically 4, 8, 16, and 32, using three benchmark datasets. Results were observed that the proposed model achieved notable accuracy when trained with a batch size of 4. This results highlights the significance of carefully selecting the batch size, as it can substantially impact the model's overall performance.

Optimizer

Optimizing the LSTM model with respect to gradient descent using a better optimizer. Adam optimizer is proposed in the considered model and fine-tune the learning rate at default values. The choice was a very intelligent one as it boosted the model accuracy greatly. The model presented has been able to improve the performance by taking advantage of the Adam optimizer. Adam optimizer was used because it proved to be very efficient and effective in training the LSTM network, hence leading to a significant improvement in the accuracy of the sentiment analysis task.

Dropout

The release itself is a gradual task used to reduce the imported classes for sentence analysis — doing this will help mitigate potential overfitting. We thus also developed dual-output method for the LSTM network to tackle this issue. As a continuous last operation, this first output layer is added between the embedding layer and the LSTM layer, dropping some units of input up to each training session. This prevents the network from becoming dependent on a few features and therefore leads to a much better input. Additionally, the second output layer on top of thick+mxd layers was beneficial in improving constructability and decreasing delamination. The proposed model effectively solves the overfitting problem by adding layers of key points in the network, resulting in better performance and better accuracy.

Evaluating the Classification Performance of the Proposed LSTMCNN Model

Statistical methods Precision and recall were calculated using equations (1) and (2), respectively for the evaluation of effectiveness of the proposed model. These methods provide a good idea of the model's ability to identify different event classes. Also, the simple mehods accuracy in the division of labor is calculated using Equation (3). Using these methods, an overall assessment of the performance of the proposed model can be obtained, which can lead to a better assessment and value of its distribution.

Precision Equation

Precision = True Positives/(True Positives+False Positives)

Recall

Recall, also known as sensitivity or true positive value, measures the proportion of correct predictions of each positive event. It is calculated using the following equation:

Precision = True Positives/(True Positives+False Positives)

Accuracy

Accuracy represents the proportion of correctly

Table 1. Sentiment classification precision comparison of algorithms review wise.

Precision Comparison of Algorithms			
Reviews	LSTM	BI-LSTM-FUZZY	LSTM-CNN
Positive	0.936	0.952	0.956
Negative	0.938	0.952	0.960
Neutral	0.938	0.947	0.961

Table 2. Sentiment classification recall comparison of algorithms review-wise.

Recall Comparison of Algorithms			
Reviews	LSTM	BI-LSTM-FUZZY	LSTM-CNN
Positive	0.922	0.946	0.958
Negative	0.924	0.946	0.957
Neutral	0.921	0.948	0.955

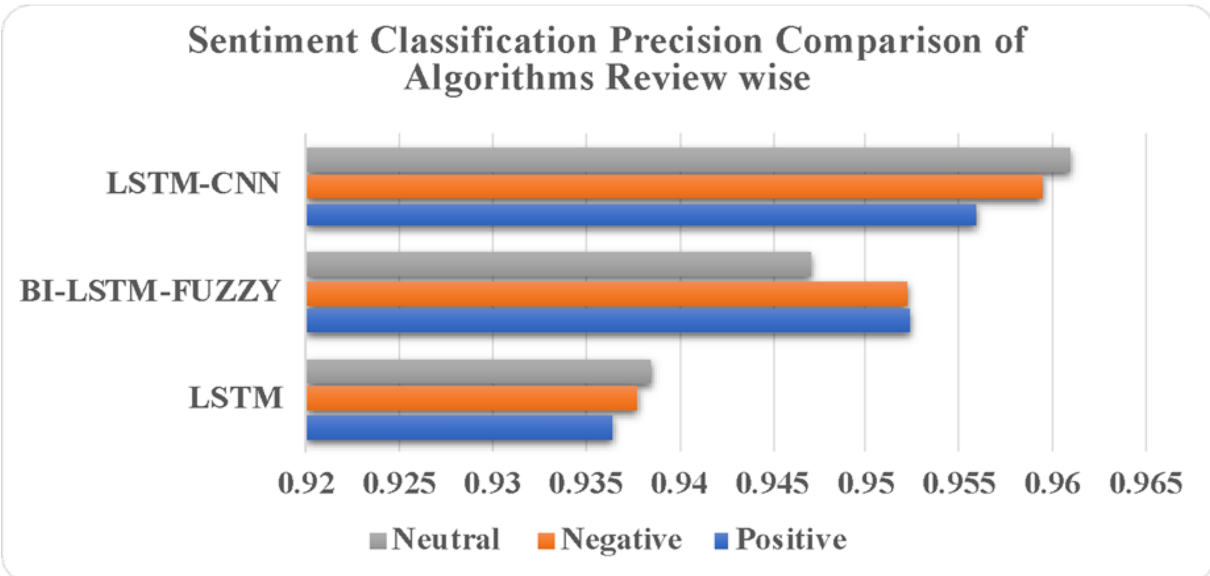

Figure 5. Sentiment classification precision comparison of algorithms review-wise.

DOI: https://doi.org/10.52756/ijerr.2024.v43spl.011

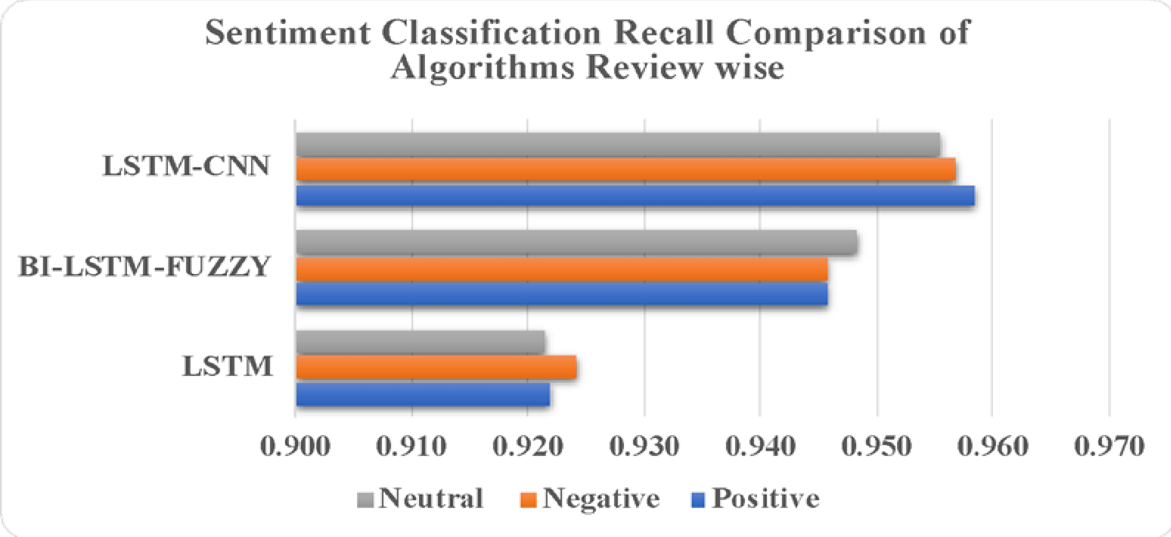

Figure 6. Sentiment classification recall comparison of algorithms review wise.

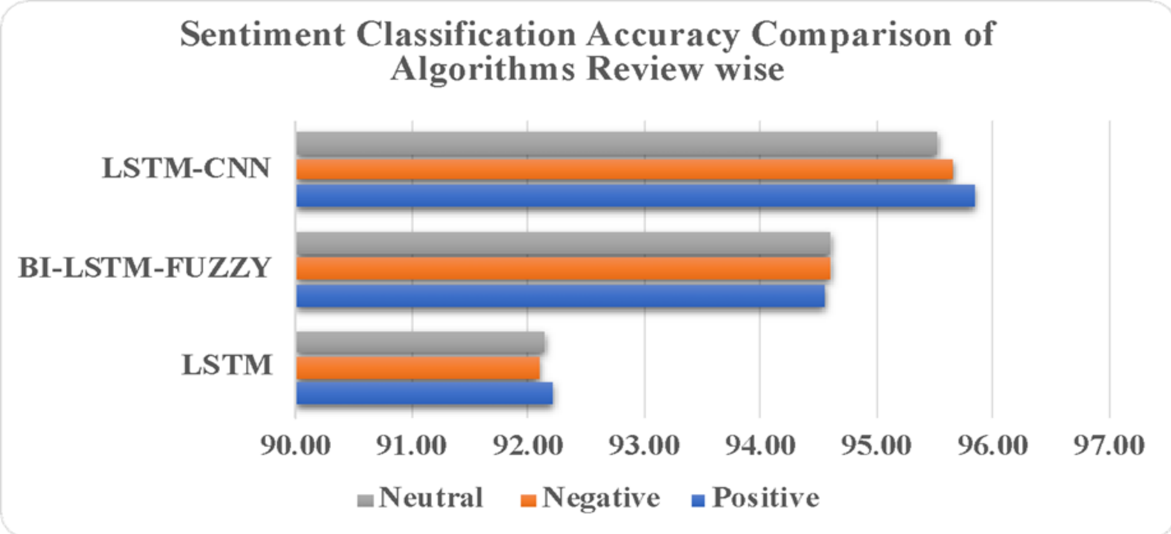

Figure 7. Sentiment classification accuracy comparison of algorithms review wise.

Table 3. Sentiment classification accuracy comparison of algorithms review wise.

Accuracy Comparison of Algorithms			
Reviews	**LSTM (%)**	**BI-LSTM-FUZZY (%)**	**LSTM-CNN (%)**
Positive	92.20	94.55	95.84
Negative	92.08	94.60	95.65
Neutral	92.15	94.61	95.51

Table 4. Experimental results comparison of proposed model with state-of-the-art methods.

Method	**Twitter cell phone reviews**		
	Precision	Recall	Accuracy
LSTM	93.6	92.2	92.2
Bi-LSTM-Fuzzy	95.2	94.6	94.55
LSTMCNN	95.6	95.8	95.84

predicted instances out of the total number of occurrences and can be calculated by using the following below equation:

Accuracy = (True Positives + True Negatives) / (True Positives + True Negatives + False Positives + False Negatives).

The evaluation of the suggested model requires some critical parameters to evaluate its efficiency in sentence recognition. These indices include True Positives (TP), which represent that the number of articles was indeed proper and positive, and negative (TN), which represent

DOI: https://doi.org/10.52756/ijerr.2024.v43spl.011

that the number of negative reviews too be classified as negative. Also, those articles that do not belong to the corresponding class are mentioned, such as False Positives (FP), which are articles that are not classified as positive but are negative ones, and False Negatives (FN), which are articles incorrectly classified as negative while, they are positive. All of them are the kinds of juxtaposition combinators. From synthesizing these kinds of examples, the whole model may be assessed on the grammatical role it seems to have. The proposed "LSTMCNN" model is evaluated using messages on Twitter. Considering the measurement model, the file having 62,563 rows and columns. The experimental outcome are compared with the well-known techniques (LSTM and BiLSTM-Fuzzy), as shown in Table 4. The proposed model reported 95.6% precision and 95.8% recall. Compared with existing models such as LSTM and BiLSTM-Fuzzy, the proposed model has more accuracy, recall and precision. The predominant values have been highlighted in Table 4 and figures 5, 6 and 7.

Conclusion

This study introduces a new approach to aspect-based sentiment analysis using LSTMCNN. Our model incorporates the features of the ClausIE framework, which effectively breaks down lengthy sentences into smaller, meaningful segments. We experimented with and without word embedding techniques for feature extraction to evaluate its performance. Interestingly, our results revealed that the word embedding technique proved highly beneficial for aspect-based sentiment analysis. The proposed LSTMCNN model was put to the test using Twitter mobile phone reviews as the dataset, which comprised 62,563 rows and several columns. The outcomes were promising, as our model achieved an accuracy of 95.84%, with precision and recall reaching 95.6% and 95.8%, respectively. Notably, when compared to existing models like LSTM and BiLSTM-Fuzzy, our proposed model exhibited higher accuracy, recall, and precision. These encouraging findings open up exciting possibilities for the future. Our model can be extended to tackle aspect-based sentiment analysis on more intricate aspect corpuses gathered from various online shopping portals in real-time scenarios. This advancement could lead to even more accurate and reliable sentiment analysis in a dynamic, ever-changing online shopping landscape.

Future Scope

Despite the promising results achieved in this study, several challenges remain to be addressed in future research. One significant limitation is the reliance on a single data set, which may not fully capture the diversity of consumer opinions. Expanding the dataset to include reviews from multiple sources and regions can enhance the generalizability of the findings. Additionally, exploring techniques to handle sentiment analysis in languages other than English would broaden the applicability of the proposed system. Furthermore, investigating methods to address the issue of subjectivity and bias in online reviews can improve the accuracy and reliability of sentiment extraction. Finally, developing a real-time sentiment analysis system using artificial intelligence methods that can continuously monitor and analyze emerging trends in consumer sentiment would provide valuable insights for businesses.

Conflict of Interest

The authors declare no known conflict of interest to publish the article.

References

Ahmed, Z., & Wang, J. (2023). A fine-grained deep learning model using embedded-CNN with BiLSTM for exploiting product sentiments. *Alexandria Engineering Journal*, *65*, 731–747. https://doi.org/10.1016/j.aej.2022.10.037

Ananthajothi, K., Karthikayani, K., & Prabha, R. (2022). Explicit and implicit oriented Aspect-Based Sentiment Analysis with optimal feature selection and deep learning for demonetization in India. *Data & Knowledge Engineering*, *142*, 102092. https://doi.org/10.1016/j.datak.2022.102092

Asif, M., Ishtiaq, A., Ahmad, H., Aljuaid, H., & Shah, J. (2020). Sentiment analysis of extremism in social media from textual information. *Telematics and Informatics*, *48*, 101345. https://doi.org/10.1016/j.tele.2020.101345

Catelli, R., Pelosi, S., Comito, C., Pizzuti, C., & Esposito, M. (2023). Lexicon-based sentiment analysis to detect opinions and attitude towards COVID-19 vaccines on Twitter in Italy. *Computers in Biology and Medicine*, *158*, 106876. https://doi.org/10.1016/j.compbiomed.2023.106876

DOI: https://doi.org/10.52756/ijerr.2024.v43spl.011

Chakraborty, K., Bhatia, S., Bhattacharyya, S., Platos, J., Bag, R., & Hassanien, A. E. (2020). Sentiment Analysis of COVID-19 tweets by Deep Learning Classifiers—A study to show how popularity is affecting accuracy in social media. *Applied Soft Computing*, *97*, 106754. https://doi.org/10.1016/j.asoc.2020.106754

Cheng, L.C., Chen, Y.L., & Liao, Y.Y. (2022). Aspect-based sentiment analysis with component focusing multi-head co-attention networks. *Neurocomputing*, *489*, 9–17. https://doi.org/10.1016/j.neucom.2022.03.027

Dangi, D., Telang Chandel, S., Kumar Dixit, D., Sharma, S., & Bhagat, A. (2023). An efficient model for sentiment analysis using artificial rabbits optimized vector functional link network. *Expert Systems with Applications*, *225*, 119849. https://doi.org/10.1016/j.eswa.2023.119849

Elahi, M., Khosh Kholgh, D., Kiarostami, M. S., Oussalah, M., & Saghari, S. (2023). Hybrid recommendation by incorporating the sentiment of product reviews. *Information Sciences*, *625*, 738–756. https://doi.org/10.1016/j.ins.2023.01.051

García-Méndez, S., De Arriba-Pérez, F., Barros-Vila, A., & González-Castaño, F. J. (2023). Targeted aspect-based emotion analysis to detect opportunities and precaution in financial Twitter messages. *Expert Systems with Applications*, *218*, 119611. https://doi.org/10.1016/j.eswa.2023.119611

Han, Y., Zhou, X., Wang, G., Feng, Y., Zhao, H., & Wang, J. (2023). Fusing sentiment knowledge and inter-aspect dependency based on gated mechanism for aspect-level sentiment classification. *Neurocomputing*, *551*, 126462. https://doi.org/10.1016/j.neucom.2023.126462

Huang, B., Guo, R., Zhu, Y., Fang, Z., Zeng, G., Liu, J., Wang, Y., Fujita, H., & Shi, Z. (2022). Aspect-level sentiment analysis with aspect-specific context position information. *Knowledge-Based Systems*, *243*, 108473. https://doi.org/10.1016/j.knosys.2022.108473

Huang, H., Zhang, B., Jing, L., Fu, X., Chen, X., & Shi, J. (2022). Logic tensor network with massive learned knowledge for aspect-based sentiment analysis. *Knowledge-Based Systems*, *257*, 109943. https://doi.org/10.1016/j.knosys.2022.109943

Huang, Y., Peng, H., Liu, Q., Yang, Q., Wang, J., Orellana-Martín, D., & Pérez-Jiménez, M. J. (2023). Attention-enabled gated spiking neural P model for aspect-level sentiment classification. *Neural Networks*, *157*, 437–443. https://doi.org/10.1016/j.neunet.2022.11.006

J, S., & U, K. (2023). Sentiment analysis of amazon user reviews using a hybrid approach. *Measurement: Sensors*, *27*, 100790. https://doi.org/10.1016/j.measen.2023.100790

Kumar, N., Agarwal, P., Bansal, S., Kumar Yadav, V., & Bhowmik, D. (2024). Sentiment analysis using novel deep learning methods. Proceedings on Engineering Sciences, 6(2), 853–862. https://doi.org/10.24874/PES06.02A.012

Leelawat, N., Jariyapongpaiboon, S., Promjun, A., Boonyarak, S., Saengtabtim, K., Laosunthara, A., Yudha, A. K., & Tang, J. (2022). Twitter data sentiment analysis of tourism in Thailand during the COVID-19 pandemic using machine learning. *Heliyon*, *8*(10), e10894. https://doi.org/10.1016/j.heliyon.2022.e10894

Lv, Y., Wei, F., Cao, L., Peng, S., Niu, J., Yu, S., & Wang, C. (2021). Aspect-level sentiment analysis using context and aspect memory network. *Neurocomputing*, *428*, 195–205. https://doi.org/10.1016/j.neucom.2020.11.049

Lyu, Z., & Takikawa, H. (2022). Media framing and expression of anti-China sentiment in COVID-19-related news discourse: An analysis using deep learning methods. *Heliyon*, *8*(8), e10419. https://doi.org/10.1016/j.heliyon.2022.e10419Mehra, P. (2023). Unexpected surprise: Emotion analysis and aspect based sentiment analysis (Absa) of user generated comments to study behavioral intentions of tourists. *Tourism Management Perspectives*, *45*, 101063. https://doi.org/10.1016/j.tmp.2022.101063

Melton, C. A., Olusanya, O. A., Ammar, N., & Shaban-Nejad, A. (2021). Public sentiment analysis and topic modeling regarding COVID-19 vaccines on the Reddit social media platform: A call to action for strengthening vaccine confidence. *Journal of Infection and Public Health*, *14*(10), 1505–1512. https://doi.org/10.1016/j.jiph.2021.08.010

Piryani, R., Gupta, V., Singh, V. K., & Ghose, U. (2017). A linguistic rule-based approach for aspect-level sentiment analysis of movie reviews. In S. K. Bhatia, K. K. Mishra, S. Tiwari, & V. K. Singh (Eds.), *Advances in Computer and Computational Sciences*, *553*, 201–209. Springer Singapore. https://doi.org/10.1007/978-981-10-3770-2_19

'7 Reasons Why Customer Feedback Is Important To Your Business - Startquestion - create online surveys and forms'. [Online]. Available: https://www.startquestion.com/blog/7-reasons-

'Understanding Customer Experience'. [Online]. Available: https://hbr.org/2007/02/understanding-customer-experience.

Sharifi, Z., & Shokouhyar, S. (2021). Promoting consumer's attitude toward refurbished mobile phones: A social media analytics approach. *Resources, Conservation and Recycling*, *167*, 105398. https://doi.org/10.1016/j.resconrec.2021.105398

Sunitha, D., Patra, R. K., Babu, N. V., Suresh, A., & Gupta, S. C. (2022). Twitter sentiment analysis using ensemble based deep learning model towards COVID-19 in India and European countries. *Pattern Recognition Letters*, *158*, 164–170. https://doi.org/10.1016/j.patrec.2022.04.027

why-customer-feedback-is-important-to-your-business/.

Wu, H., Huang, C., & Deng, S. (2023). Improving aspect-based sentiment analysis with Knowledge-aware Dependency Graph Network. *Information Fusion*, *92*, 289–299. https://doi.org/10.1016/j.inffus.2022.12.004

Xiang, Y., Zhang, J., & Guo, J. (2023). Block-level dependency syntax based model for end-to-end aspect-based sentiment analysis. *Neural Networks*, *166*, 225–235. https://doi.org/10.1016/j.neunet.2023.05.008

You, L., Han, F., Peng, J., Jin, H., & Claramunt, C. (2022). ASK-RoBERTa: A pretraining model for aspect-based sentiment classification via sentiment knowledge mining. *Knowledge-Based Systems*, *253*, 109511. https://doi.org/10.1016/j.knosys.2022.109511

Zhao, Z., Tang, M., Tang, W., Wang, C., & Chen, X. (2022). Graph convolutional network with multiple weight mechanisms for aspect-based sentiment analysis. *Neurocomputing*, *500*, 124–134. https://doi.org/10.1016/j.neucom.2022.05.045

Zhou, T., Shen, Y., Chen, K., & Cao, Q. (2023). Hierarchical dual graph convolutional network for aspect-based sentiment analysis. *Knowledge-Based Systems*, *276*, 110740. https://doi.org/10.1016/j.knosys.2023.110740

Zou, F., Zhu, X., Qian, Y., & Chang, D. (2022). A sustainable product design approach based on data mining of dynamic user demands: A case study on Huawei mate 40 series. *IFAC-Papers OnLine*, *55*(10), 1056–1061. https://doi.org/10.1016/j.ifacol.2022.09.529

Žunić, A., Corcoran, P., & Spasić, I. (2021). Aspect-based sentiment analysis with graph convolution over syntactic dependencies. *Artificial Intelligence in Medicine*, *119*, 102138. https://doi.org/10.1016/j.artmed.2021.102138

How to cite this Article:

Nitendra Kumar, Ritu Talwar, Sadhana Tiwari and Priyanka Agarwal (2024). Aspect-based sentiment analysis of Twitter mobile phone reviews using LSTM and Convolutional Neural Network. *International Journal of Experimental Research and Review*, *43*, 146-159.

DOI : https://doi.org/10.52756/ijerr.2024.v43spl.011