A Hybrid Deep Learning Framework for Efficient Sentiment Analysis

Asish Karthikeya Gogineni, S Kiran Sai Reddy, Harika Kakarala, Yaswanth Chowdary Gavini,
M Pavana Venkat, Koduru Hajarathaiah, Murali Krishna Enduri
Algorithms and Complexity Theory Lab-Department of Computer Science and Engineering,
SRM University-AP, Amaravati, India

***Abstract*—In the era of Microblogging and the rapid growth of online platforms, an exponential rise is shown in the volume of data generated by internet users across various domains. Additionally, the creation of digital or textual data is expanding significantly. This is because consumers respond to comments made on social media platforms regarding events or products based on their personal experiences. Sentiment analysis is usually used to accomplish this kind of classification on a large scale. It is described as the process of going through all user reviews and comments that are discovered in product reviews, events, or similar sources in order to look for unstructured text comments. Our study examines how deep learning models like LSTM, GRU, CNN, and hybrid models (LSTM+CNN, LSTM+GRU, GRU+CNN) capture complex sentiment patterns in text data. Additionally, we study integrating BOW and TF-IDF as complementing features to improve model predictive power. CNN with RNNs consistently improves outcomes, demonstrating the synergy between convolutional and recurrent neural network architectures in recognizing nuanced emotion subtleties.In addition, TF-IDF typically outperforms BOW in enhancing deep learning model sentiment analysis accuracy.**

Keywords*—Sentiment analysis; LSTM; GRU; Convolutional Neural Networks (CNNs); BOW; TF-IDF

I. Introduction

Using accurate and reliable methods, sentiment analysis aims to automatically determine a text's sentiment and extract meaningful information. Various methodologies, tactics, algorithms, and approaches are investigated in sentiment analysis to understand textual emotions [1]. Its main purpose is to automatically classify text into neutral, joyful, furious, sad, and other emotions in addition to positive and negative ones. Sentiment analysis is essential for evaluating consumer evaluations of products and services since it reveals customer sentiment. The system uses advanced natural language processing algorithms to categorize customer evaluations, comments, and ratings as favorable, negative, or neutral (see Fig. 1). This enables organizations swiftly assess client opinion of their products.

Sentiment analysis also examines feedback to find traits that elicit positive or negative responses. Text-based movie reviews may be automatically analyzed and classified using sophisticated computational approaches for sentiment analysis [2]. Filmmakers and consumers may better communicate and enjoy film via movie reviews. Sentiment analysis algorithms evaluate reviews' language, tone, and context to determine positive, negative, or neutral sentiment [3]. This strategy helps filmmakers and spectators understand crowd responses. Sentiment analysis gives filmmakers vital audience feedback. Positive comments help filmmakers identify their talents by highlighting audience preferences. Conversely, negative feelings indicate places for progress. This helps improve storyline, character development, and cinematography in future projects. Sentiment analysis helps viewers choose movies [4]. Potential viewers may determine whether a film suits them by reading reviews.

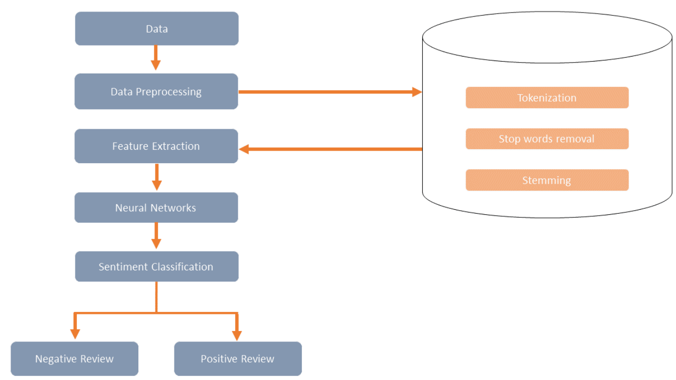

Fig. 1. Processing of DL algorithms.

An element of sentiment analysis employed in Coursera evaluations is the computerized evaluation of the emotion represented by students in their course remarks. Using ML or DL models trained on the labeled dataset, the sentiment analysis system learns to discover patterns that connect particular linguistic cues to particular sentiments [5]. The main motivation of this work is to analyze people's opinions by collecting reviews in the form of text and predicting the outcome, i.e., positive or negative reviews. Additionally, the goal is to assess the efficiency of deep learning models compared to machine learning techniques. When fresh Coursera reviews are fed into the algorithm, it forecasts the mood of each review, giving information about how students feel about the platform, instructors, course material, and more. This information can guide course improvements, influence instructional strategies, and offer valuable feedback to instructors, helping them refine their teaching methods based on learner sentiments.

The following sections offer a concise overview of pertinent details. In Section II, we examine prior research related to sentiment analysis, seeking to establish connections between existing knowledge and recent discoveries. Section III delves into the methodologies employed for this analysis, while Section IV is dedicated to detailing the dataset utilized in this paper. The conclusion of our work, along with the presentation of results, will bring this paper to a close.

A. Applications

1) Social media monitoring and brand management: Businesses use sentiment analysis to track mentions of their goods, providers of services, and brands on social media. By analyzing the sentiment of these mentions, you can gauge public perceptions, identify potential problems, and engage with your customers more effectively [6], [7].
2) Customer Opinion or Reviews: E-commerce platforms, restaurants, hotels, and other businesses use sentiment analysis to automatically process customer reviews. This helps us understand customer satisfaction, identify opportunities for improvement, and respond quickly to negative feedback [8], [9].
3) Movie and TV show review : Sentiment analysis is used to measure audience reactions to movies and TV shows, which aids in marketing and content creation [10], [11].
4) Product Development: Companies use sentiment analysis to analyze feedback and reviews of existing products, guiding improvements, and informing the development of new products that align with customer preferences [12], [13].

II. Related and Recent Work

In the recent past, the field of Machine Learning, Deep Learning, and Natural language processing (NLP) has witnessed a notable rise in sentiment analysis. With the huge expansion of textual information on the web, there is a growing need for automated tools that can accurately identify sentiment and emotional nuances within text. This increased focus on sentiment analysis has spurred the development of various novel methods and techniques that make use of deep learning models' capabilities to extract intricate sentiment insights from text data. These advances allow many applications, from social media sentiment analysis to corporate customer sentiment analysis. Here, we shall explain the recent contributions that have had a major influence on this subject and their relevance.

The work by Vateekul [14] used LSTM and DCNN deep learning algorithms on a Thai Twitter dataset. Deep learning outperformed several classic machine learning methods, according to their results. A reduced feature set and dimensionality reduction approaches were used in Akmal et al.'s clustering algorithm for emotional analysis [15]. The authors used deep learning to analyze sentiment in COVID-19 reviews [16], [17]. Their method uses an LSTM-RNN network and attention layers to improve features.For aspect extraction and emotion classification. A multitask learning technique was proposed by Akhtar et al. [18]. The extraction and sentiment categorization of aspects are seen to be separate activities. CNNs and BI-LSTMs execute the duties. The BI-LSTM layer infers Beginning and interior tags for each word using word embeddings to identify aspect keywords in a review sentence. Use a masking layer to remove non-aspect words from a text.

To extract sentiments from consumer evaluations about various characteristics of items, Pham and Le [19] used a multilayer architecture based on recurrent neural networks (RNNs). Their approach was used to examine 174,615 reviews relating to 1,768 hotels on tripadvisor.com. The results show that their technique has a lot of potential for sentiment analysis and hotel rating prediction. In this domain, Hassan and Mahmood used a hybrid technique that blends CNN and RNN architectures. They begin by training word embeddings with an unsupervised NLP model that has previously been fine-tuned on a large dataset. Following that, they use the capabilities of CNN for feature extraction and RNN for capturing interdependencies to detect sentiments in a variety of datasets [20].

We previously covered sentiment analysis using machine learning techniques [21] on similar datasets, as well as data cleaning and feature extraction, and this project serves as an expansion, digging into sentiment analysis using deep learning. Our findings reveal that machine learning models like Random Forest models demonstrate competitive performance, especially in scenarios with limited data. However, deep learning models, outperform Random Forest as the dataset size increases. We discuss the trade-offs between interpretability and performance, as Random Forest provides explicit feature importance, while deep learning models offer superior representation learning capabilities.

III. Methodology

This section offers a detailed overview of the approach employed for Sentimental Analysis using the neural networks.

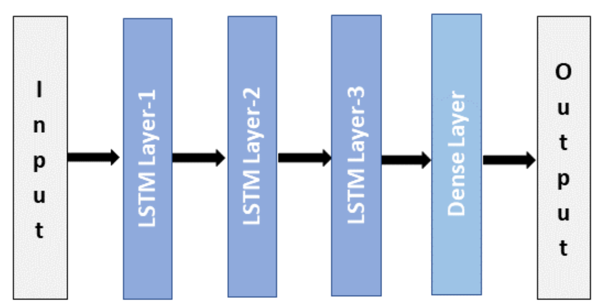

Fig. 2. LSTM Architecture.

A. LSTM

Long Short-Term Memory (LSTM) is a type of sophisticated recurrent neural network (RNN) crafted for proficient sequence modeling and prediction, mitigating the challenges of vanishing gradients in extended sequences [22]. In our model, first, we take input and pass it to LSTM Layers. Fig. 2 illustrates the three LSTM layers in our model. The first layer of the LSTM, Layer 1, has 32 units and returns sequences, passing on its output sequences to the subsequent layer. Layer 2, on the other hand, has 64 units and likewise returns sequences. Layer 3, on the other hand, has 128 units and does not return sequences, resulting in a fixed-size output. Data is eventually passed to a dense layer with 64 units and the ReLU activation function since it was an output layer for binary classification [23]. The output layer additionally has one neuron with a sigmoid activation function.

B. GRU

One type of specialised recurrent neural network (RNN) is the gated recurrent unit (GRU). The unique way that a GRU manages the temporal flow of information sets it apart from other RNN designs, such the Long Short-Term Memory (LSTM) network. The two gates that GRU uses to regulate information flow within the network are an update gate and a reset gate. This architectural simplicity of the GRU contributes to its ease of training and reduced susceptibility to overfitting, particularly in scenarios characterized by limited data [24]. The model that we used, contains three GRU layers as illustrated in Fig. 3. The first and second layer contains 32 and 64 units as well as return sequences, whereas the third layer contains 128 units with no return sequences. Subsequently, it is forwarded to a dense layer comprising 64 units, employing the ReLU activation function. The final layer is made up of a single neuron that has sigmoid activity, serving as the output layer for binary classification.

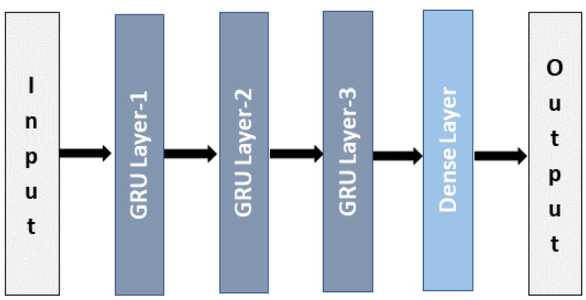

Fig. 3. GRU Architecture.

C. CNN

Convolutional neural networks, or CNNs for short, are a kind of Deep Learning neural network design that is frequently used in sentiment analysis, picture categorization, and other related applications. Convolutional neural networks, or CNNs, are known to be composed of several layers, including the input, convolutional, pooling, and dense layers. While the Pooling layer reduces the dimensions to lessen computational overhead, the Convolutional layer uses filters to extract features from the input. Ultimately, the fully connected layer is responsible for generating the final prediction [25]. In our model, the first layer is input layer which processes data with given shape and serves as entry point for review. Convolutional Layer 1, the next layer, has 32 filters, a 3×3 kernel, and ReLU activation. Using this, we then employed pooling. This pooling helps to reduce the size. Average and max poolings are the two main categories of pooling. Since max pooling performs better than average pooling and aids in the extraction of more significant features, we employed it with the default size. This process is repeated with Convolutional Layer 2, which has 64 filters, and then with an additional Convolutional Layer 3 using 128 filters. Convolutional layers extract various layers of information from the input. The main reason of using activation function as ReLU over other activation functions, i.e., Sigmoid, tanH and Softmax is that it does not activate all neurons at the same time. The output will be sent to the dense layer once the text has been processed via all of the pooling and convolutional layers. Since the dense layer requires input in the form of a 1-D array, we are unable to pass the convolutional layer's multi-dimensional output straight to it. To overcome this, we used the Flatten method between the dense layer and the convolutional layer. Finally, we used a dense layer to classify sentiment from convolutional layers as shown in Fig. 4. In our model, Since the output layer was intended for binary classification, it has one neuron with sigmoid activation function.

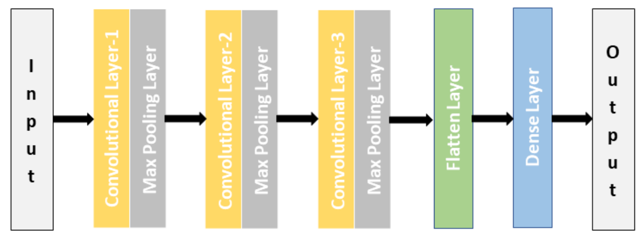

Fig. 4. CNN Architecture.

D. GRU+LSTM

As widely acknowledged, the Gated Recurrent Unit (GRU) is considered a highly promising algorithm within the realm of Recurrent Neural Networks (RNNs). In terms of functionality, both GRU and LSTM share similarities, However, GRU unifies the forget gate and input gate into a single update gate, using a single hidden state. Moreover, GRU creates a single state by merging the concealed state with the cell state. This efficient approach has led to GRU being recognized as a simplified variant of LSTM [26]. The model we have employed incorporates an initial layer with a single GRU layer containing 32 units and a return sequence. The subsequent layer features an LSTM layer with 64 units, also with a return sequence. We have added layer with a GRU and 128 filters the result then uses the ReLU activation algorithm to advance to a dense layer with 64 units. The final layer remains consistent with other hybrid models as depicted in Fig. 5.

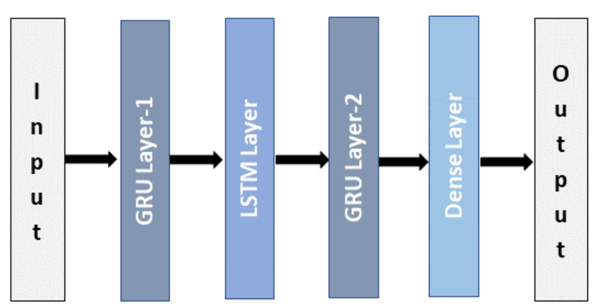

Fig. 5. GRU+LSTM Architecture.

E. CNN+LSTM

A Convolutional Neural Network (CNN) operates on sequential data by employing sliding convolutional filters over the input, enabling it to capture features from both spatial and temporal dimensions. In contrast, an LSTM network processes sequential data by iterative processing time steps, capturing long-term dependencies between them. A CNN-LSTM network combines convolutional and LSTM layers to effectively extract insights from training data [27]. In our model, The initial layer serves as the input layer, processing data with a specified shape. Following this, The first Convolutional Layer is then presented; it has 32 filters and uses a 3×3 kernel with activation as ReLU. After this operation, we apply pooling. This sequence is then replicated with Convolutional Layer 2, which incorporates 64 filters, and subsequently with an additional Convolutional Layer 3, utilizing 128 filters. Following the passage through the pooling and convolutional layers, the data is directed to the LSTM Layer, housing 64 units with a specific activation function. Subsequently, the output moves on to a 64-unit dense layer, employing the ReLU activation function. The final layer is composed of a 1 neuron utilizing a sigmoid activation function as illustrated in Fig. 6, which is ideal for binary classification, serving as the output layer.

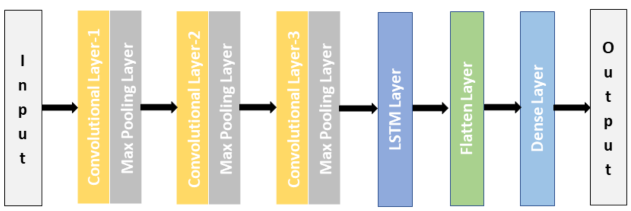

Fig. 6. CNN+LSTM Architecture.

F. GRU+CNN

The combination of CNN and Gated Recurrent Unit (GRU) components is our next hybrid model. The predictive performance of this dual-architecture method is enhanced. It also lessens the chance of overfitting and it is a novel work for the four different datasets in the context of sentimental analysis. The model we utilized, which is seen in Fig. 7, starts with a GRU (Gated Recurrent Unit) layer with 32 units and return sequences. Next, we add convolutional operations to the network by incorporating a Conv1D layer, which has 64 filters and a kernel size of 3. We next add another 128-unit GRU layer to the sequence, which is followed by a 64-unit dense layer with a ReLU activation function. Last but not least, we include a single-unit output layer and use a sigmoid activation function for binary classification.

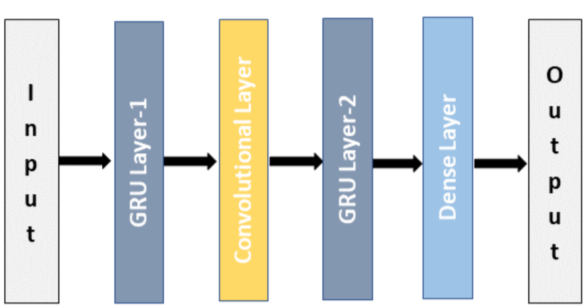

Fig. 7. GRU+CNN Architecture.

IV. Dataset Statistics

In our study, we harnessed a diverse array of four datasets to embark on comprehensive sentiment analysis. These datasets were meticulously selected for their relevance and diversity, thereby encompassing a broad spectrum of reviews from various domains. To elucidate further, the initial dataset centered around Movie Reviews and contained $34,000$ entries. Our second dataset, Coursera Reviews, proved to be expansive, comprising a total of $107,018$ reviews. Additionally, the Google Play Store Reviews dataset, our third one, provided valuable insights into the realm of mobile applications with its $12,485$ reviews. Lastly, the fourth dataset, Flipkart Reviews, encompassed a total of $9,976$ reviews, revealing a distinct polarity in emotional sentiment. This carefully curated assortment of datasets serves to ensure that our sentiment analysis model receives training from a diverse collection of linguistic expressions and contextual subtleties. There are $16,993$ instances of 0s and $17,007$ instances of 1s in the Movie Reviews data collection. There are $102,298$ occurrences of 0s and $4,720$ occurrences of 1s in the Coursera Reviews data set. There are $7,635$ instances of 0s and $4,850$ instances of 1s in Google Play Store Reviews. Finally, there are $8,975$ instances of 0s and $1,001$ instances of 1s in the Flipkart Reviews data set. Within their individual data sets, these binary values most likely indicate specific features or emotion labels connected with each review.

Formally, within a provided training dataset of reviews and their associated sentiment labels, a sentiment score of $'1'$ signifies a negative review, while a score of $'0'$ designates a positive one. We aim to characterize the feelings included within the supplied text collection appropriately.

A. Steps to Perform

In our research, we followed a structured process to prepare and preprocess our dataset [28].

- Data Cleaning: Special Character Removal: In this step, you remove special characters such as symbols, emojis, or any characters that don't contribute significantly to the text's meaning. This can be done using libraries like NLTK (Natural Language Toolkit) or regular expressions. Punctuation Removal: Similar to special characters, punctuation marks like periods, commas, and exclamation marks are often removed as they can add noise to the text.
- Tokenization: Dividing each sentence into individual words or tokens. For the majority of natural language processing (NLP) activities, this is a fundamental step.
- Stemming: The process of reducing words to their root or basic form is known as stemming. For instance, the terms "running," "ran," and "runs" may all have the same root, "run". This can increase processing efficiency and aid in lowering the dimensionality of the text data.
- Feature Extraction: After that, we used Bag of Words and TF-idf (Term Frequency-Inverse Document Frequency) feature extraction methods to quantitatively represent text data for deep learning or machine learning models. TF-idf weighted words by importance

in documents, while Bag of Words employed word counts to vectorize the text. Our data was better after these pretreatment stages and suitable for study analysis and modeling.

Bag-of-Words (BoW) characteristics were an important text analysis tool. BoW includes accumulating words to extract traits from text materials. For model training, a thorough vocabulary of different concepts from all training dataset records is needed. To facilitate text mining and information extraction, we used TF-IDF as a text data weight analysis approach. The significance of each word in a document is measured by TF-IDF. It normalizes word frequency according to text length, making it suitable for diverse document sizes. Additionally, TF-IDF rates each word's importance in a publication. Since it examines term frequency and normalizes it by text length, it works for a variety of document sizes. Furthermore, when determining the frequency with which a term appears in a particular text, TF (Term Frequency) accounts for changes in document lengths. On the other hand, IDF compares a word's meaning throughout the corpus to common, meaningless phrases like "that," "of," and "is". We may eliminate extraneous terms and concentrate on the most pertinent ones thanks to this strategy.

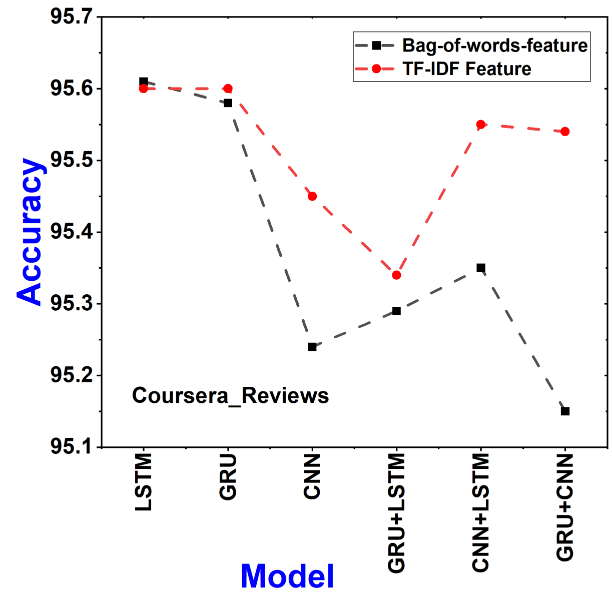

Fig. 9. Accuracy score for coursera reviews using six DL algorithms with TF-IDF and bag of words.

V. Results

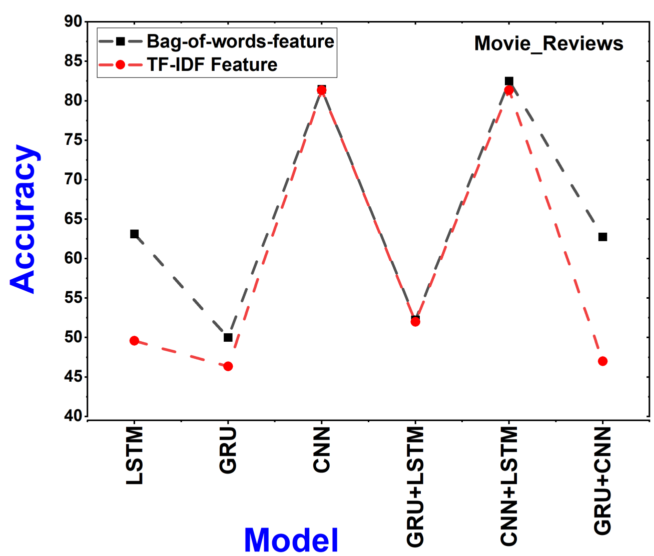

Fig. 8. Using six different DL algorithms with TF-IDF and Bag of Words, the accuracy score for movie reviews.

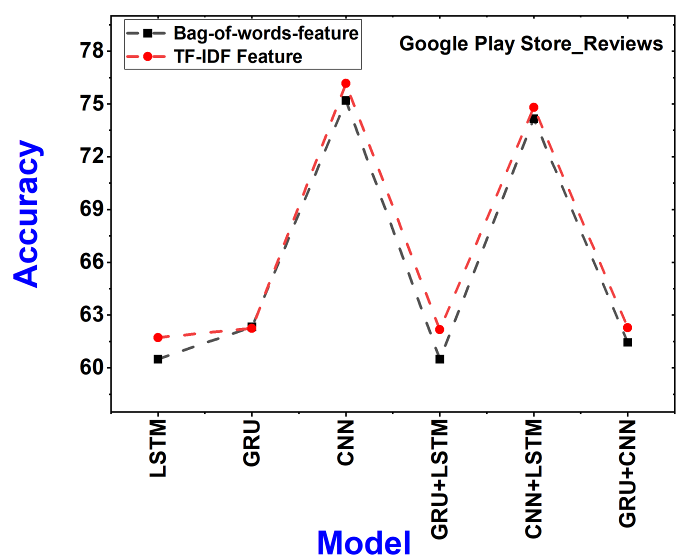

Fig. 10. Six DL algorithms with TF-IDF and bag of words were used to calculate the accuracy score for reviews on the google play store.

In this section, following the training of six distinct deep learning models (LSTM, GRU, CNN, LSTM+GRU, CNN+LSTM, and CNN+GRU), it became evident that the CNN+LSTM model consistently achieved higher accuracy scores across all the datasets, i.e., Movie, Coursera, Google Play Store, and Flipkart reviews compared to other deep learning models. The outcomes are displayed in Fig. 8, 9, 10, 11. Accuracy utilizing deep learning methods for different data sets is shown in Table I. We also observed that when the data set contains more reviews, deep learning models are more accurate in predicting compared to machine learning models. While CNNs on their own are typically not anticipated to surpass LSTM or GRU in terms of sentiment analysis performance due to their limited capacity to capture sequential dependencies, the combination of CNNs with RNNs, like CNN+LSTM often proves to be more effective than standalone models, shown in the Fig. 8, 9, 10, 11. Because these models can leverage the strength of CNNs in capturing local patterns and combine it with the sequential modeling capabilities of LSTMs. The

TABLE I. ACCURACY UTILIZING DEEP LEARNING METHODS FOR DIFFERENT DATA SETS

Model	Features	Movie	Coursera	Google Play Store	Flipkart
LSTM	Bag-of-words-feature	63.12	95.61	60.49	89.38
	TF-IDF Feature	49.59	95.60	61.72	89.94
GRU	Bag-of-words-feature	50.00	95.58	62.33	87.35
	TF-IDF Feature	46.34	95.60	62.25	88.91
CNN	Bag-of-words-feature	81.46	95.24	75.20	92.18
	TF-IDF Feature	81.34	95.45	76.16	93.08
GRU + LSTM	Bag-of-words-feature	52.25	94.29	60.49	82.42
	TF-IDF Feature	52.0	95.34	62.17	87.88
CNN + LSTM	Bag-of-words-feature	82.50	95.35	74.16	92.25
	TF-IDF Feature	81.36	95.55	74.80	93.45
GRU + CNN	Bag-of-words-feature	62.75	95.15	61.45	89.38
	TF-IDF Feature	47.00	95.54	62.28	89.97

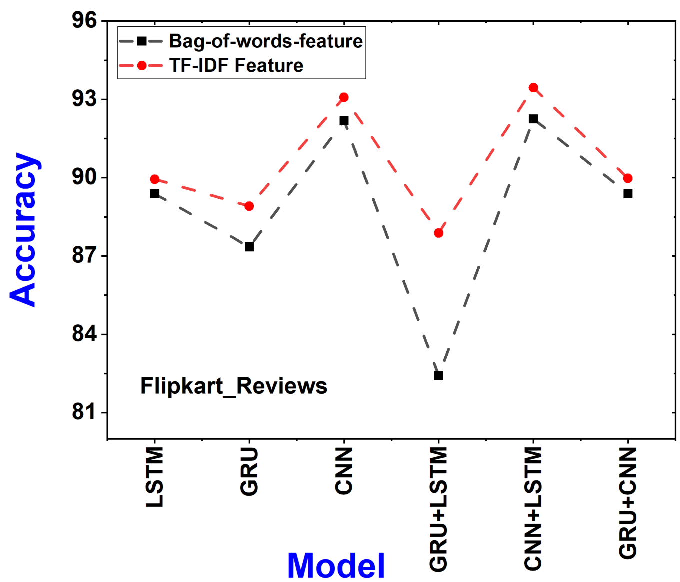

Fig. 11. The accuracy score for flipkart reviews is based on six DL algorithms, including bag of words and TF-IDF.

CNN component is capable of extracting high-level textual features, whereas the LSTM excels at capturing long-term dependencies. We also observed that the Combination of CNN with LSTM performs better than the combination of CNN with GRU. Also, we observed the combination of GRU+ LSTM will not lead good results compared to other deep learning models. This is due to a combination of both GRU and LSTM, which are similar in their functioning, and combining them does not introduce a significantly new perspective for sentiment analysis. Also, we can say that when CNNs are used for text-based tasks, they are typically employed in combination with RNNs or as part of more complex architectures, such as the Attention-based models, Transformers, or BERT-based models. These designs have performed admirably in a variety of NLP tasks, including sentiment analysis, by capturing both local and global patterns in text data. Ultimately, it was observed that Deep Learning models (LSTM, GRU, CNN, LSTM+GRU, CNN+LSTM, and CNN+GRU) outperform machine learning techniques (Logistic Regression, KNN Classifier, Bernoulli Naive Bayes, Multinomial Naive Bayes, XGBoost, Decision Tree, Random Forest Classifier) in the prediction of sentiment analysis. We also notice that these models yield improved results when utilizing the TF-IDF feature extraction technique.

In addition to the previously mentioned findings, it is noteworthy to highlight the quantitative performance of the proposed CNN+LSTM and CNN+GRU models in comparison to other state-of-the-art deep learning algorithms across diverse datasets, namely Movie, Coursera, Google Play Store, and Flipkart. The assessment was conducted using both bag-of-words and TF-IDF features to comprehensively evaluate the models' capabilities.

The CNN+LSTM model, leveraging bag-of-words features, exhibited commendable accuracy scores of 82.50, 95.35, 74.16, and 92.25 across the respective datasets. Employing TF-IDF features, the CNN+LSTM model demonstrated consistently high accuracy scores of 81.36, 95.55, 74.80, and 93.45, except in the Movie review dataset, where TF-IDF slightly outperformed bag-of-words. Notably, in the Movie review dataset, bag-of-words not only outperformed TF-IDF for the CNN+LSTM model but also surpassed the other five algorithms in accuracy.

In stark contrast, the GRU+LSTM model, when utilizing bag-of-words features, showed lower accuracy scores of 52.25, 94.29, 60.49, and 82.42 across the Movie, Coursera, Google Play Store, and Flipkart datasets, respectively. Similarly, with TF-IDF features, the GRU+LSTM model exhibited accuracy scores of 52.0, 95.34, 62.17, and 87.88. It is noteworthy that, compared to the other 5 DL models, the performance of the GRU+LSTM model is relatively poor.

VI. CONCLUSIONS

This article compares and contrasts sentimental categorization and analysis using several Deep-learning techniques that are currently available. Finally, several significant insights are shown by Table I of model accuracies based on different features and datasets. TF-IDF features are superior to Bag-of-Words features in capturing the nuances of the text data, as demonstrated by the consistently superior performance of the models that used them. The GRU with TF-IDF features and the CNN with TF-IDF features demonstrated the highest accuracies across the separate models across several datasets, demonstrating their efficacy in sentiment analysis tasks. In the realm of sentiment analysis, deep learning models exhibit superior performance when contrasted with traditional machine learning models across various domains, including reviews on Coursera, Flipkart, movies, and the Google Play Store. It's important to remember, though, that the model's performance varied greatly based on the dataset that was employed.

The majority of the models in the data from the Coursera and Google Play Store had consistently excellent accuracy scores, whereas the Movie dataset had lower accuracy ratings. Additionally, Flipkart data showed that different models performed differently, with some doing remarkably well and others having difficulty. The performance of the combination models, like CNN + LSTM and GRU + CNN, was frequently good, demonstrating the advantages of integrating several neural network architectures for improved sentiment analysis. In actuality, the model and feature representation used should be based on the needs of the particular dataset and application. Even if TF-IDF features work well most of the time, they might not be the ideal option every time. Deep learning excels in sentiment analysis by automatically learning intricate patterns, capturing contextual dependencies, and enabling end-to-end feature extraction. Hierarchical feature learning and adaptability to diverse data types make deep learning models like CNNs and LSTMs effective, the unique properties of the data should be taken into account while deciding between combination models and individual models. Future work should explore enhancing the interpretability of deep learning models in sentiment analysis, addressing its nature through methods like attention mechanisms. Investigating efficient training strategies, and data-efficient approaches, and applying novel architectures, including transformer-based models, will contribute to advancing the field and making deep learning more accessible for sentiment analysis tasks.

REFERENCES

[1] 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.

[2] B. Liu *et al.*, "Sentiment analysis and subjectivity." *Handbook of natural language processing*, vol. 2, no. 2010, pp. 627–666, 2010.

[3] A. Tripathy, "Sentiment analysis using machine learning techniques," Ph.D. dissertation, 2017.

[4] M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," *Artificial Intelligence Review*, vol. 55, no. 7, pp. 5731–5780, 2022.

[5] A. Abbasi and H. Chen, "Writeprints: A stylometric approach to identity-level identification and similarity detection in cyberspace," *ACM Trans. Inf. Syst.*, vol. 26, no. 2, apr 2008. [Online]. Available: https://doi.org/10.1145/1344411.1344413

[6] M. N. Sadiku, T. J. Ashaolu, A. Ajayi-Majebi, and S. M. Musa, "Artificial intelligence in social media," *International Journal of Scientific Advances*, vol. 2, no. 1, pp. 15–20, 2021.

[7] T. Balaji, C. S. R. Annavarapu, and A. Bablani, "Machine learning algorithms for social media analysis: A survey," *Computer Science Review*, vol. 40, p. 100395, 2021.

[8] S. Ramaswamy and N. DeClerck, "Customer perception analysis using deep learning and nlp," *Procedia Computer Science*, vol. 140, pp. 170–178, 2018.

[9] A. Iqbal, R. Amin, J. Iqbal, R. Alroobaea, A. Binmahfoudh, and M. Hussain, "Sentiment analysis of consumer reviews using deep learning," *Sustainability*, vol. 14, no. 17, p. 10844, 2022.

[10] M. Soleymani, D. Garcia, B. Jou, B. Schuller, S.-F. Chang, and M. Pantic, "A survey of multimodal sentiment analysis," *Image and Vision Computing*, vol. 65, pp. 3–14, 2017, multimodal Sentiment Analysis and Mining in the Wild Image and Vision Computing. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0262885617301191

[11] H. Zhao, Z. Liu, X. Yao, and Q. Yang, "A machine learning-based sentiment analysis of online product reviews with a novel term weighting and feature selection approach," *Information Processing & Management*, vol. 58, no. 5, p. 102656, 2021.

[12] L. Yang, Y. Li, J. Wang, and R. S. Sherratt, "Sentiment analysis for e-commerce product reviews in chinese based on sentiment lexicon and deep learning," *IEEE access*, vol. 8, pp. 23 522–23 530, 2020.

[13] M. E. Alzahrani, T. H. Aldhyani, S. N. Alsubari, M. M. Althobaiti, and A. Fahad, "Developing an intelligent system with deep learning algorithms for sentiment analysis of e-commerce product reviews," *Computational Intelligence and Neuroscience*, vol. 2022, 2022.

[14] C. Udomcharoenchaikit, P. Boonkwan, and P. Vateekul, "Adversarial evaluation of robust neural sequential tagging methods for thai language," *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, vol. 19, no. 4, pp. 1–25, 2020.

[15] H. Akmal, F. Hardalaç, and K. Ayturan, "A fetal well-being diagnostic method based on cardiotocographic morphological pattern utilizing autoencoder and recursive feature elimination," *Diagnostics*, vol. 13, no. 11, p. 1931, 2023.

[16] C. Singh, T. Imam, S. Wibowo, and S. Grandhi, "A deep learning approach for sentiment analysis of covid-19 reviews," *Applied Sciences*, vol. 12, no. 8, p. 3709, 2022.

[17] H. Kaur, S. U. Ahsaan, B. Alankar, and V. Chang, "A proposed sentiment analysis deep learning algorithm for analyzing covid-19 tweets," *Information Systems Frontiers*, pp. 1–13, 2021.

[18] M. S. Akhtar, T. Garg, and A. Ekbal, "Multi-task learning for aspect term extraction and aspect sentiment classification," *Neurocomputing*, vol. 398, pp. 247–256, 2020.

[19] D.-H. Pham and A.-C. Le, "Learning multiple layers of knowledge representation for aspect based sentiment analysis," *Data & Knowledge Engineering*, vol. 114, pp. 26–39, 2018.

[20] A. Hassan and A. Mahmood, "Convolutional recurrent deep learning model for sentence classification," *Ieee Access*, vol. 6, pp. 13 949–13 957, 2018.

[21] M. K. Enduri, A. R. Sangi, S. Anamalamudi, R. C. B. Manikanta, K. Y. Reddy, P. L. Yeswanth, S. K. S. Reddy, and A. Karthikeya, "Comparative study on sentimental analysis using machine learning techniques," *Mehran University Research Journal Of Engineering & Technology*, vol. 42, no. 1, pp. 207–215, 2023.

[22] M. S. Divate, "Sentiment analysis of marathi news using lstm," *International journal of Information technology*, vol. 13, no. 5, pp. 2069–2074, 2021.

[23] Q. Lu, X. Sun, Y. Long, Z. Gao, J. Feng, and T. Sun, "Sentiment analysis: Comprehensive reviews, recent advances, and open challenges," *IEEE Transactions on Neural Networks and Learning Systems*, 2023.

[24] S. Sachin, A. Tripathi, N. Mahajan, S. Aggarwal, and P. Nagrath, "Sentiment analysis using gated recurrent neural networks," *SN Computer Science*, vol. 1, pp. 1–13, 2020.

[25] X. Ouyang, P. Zhou, C. H. Li, and L. Liu, "Sentiment analysis using convolutional neural network," in *2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*, 2015, pp. 2359–2364.

[26] R. Ni and H. Cao, "Sentiment analysis based on glove and lstm-gru," in *2020 39th Chinese Control Conference (CCC)*, 2020, pp. 7492–7497.

[27] A. U. Rehman, A. K. Malik, B. Raza, and W. Ali, "A hybrid cnn-lstm model for improving accuracy of movie reviews sentiment analysis," *Multimedia Tools and Applications*, vol. 78, pp. 26 597–26 613, 2019.

[28] S. Elzeheiry, W. A. Gab-Allah, N. Mekky, and M. Elmogy, "Sentiment analysis for e-commerce product reviews: Current trends and future directions," 2023.