Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2025 SCPE. Volume 26, Issues 4, pp. 1902–1918, DOI 10.12694/scpe.v26i4.4449

HYBRID DEEP LEARNING RECOMMENDATION SYSTEM FOR ACCURATE MOVIE AND PRODUCT REVIEW PREDICTIONS

SANJEEV DHAWAN*, KULVINDER SINGH† AND MANOJ YADAV‡

Abstract. This paper investigates the efficacy of deep learning models for sentiment analysis using two publicly available datasets: IMDb's movie review dataset and Amazon's product review dataset. The main objective was to evaluate the performance of various model architectures, particularly Long Short-Term Memory (LSTM) networks with dropout techniques, in emotion categorization under different settings. Key performance metrics, including accuracy, precision, recall, and F1 score, were used to train and validate several deep learning models: LSTM Spatial Dropout 1D, Bidirectional GRU-LSTM, Hybrid LSTM+GRU, and Bidirectional LSTM. The LSTM Spatial Dropout 1D model achieved remarkable results, with 93.00% accuracy and F1 score on the Amazon dataset, and an impressive 96.20% accuracy and 97.78% F1 score on the IMDb dataset. The Bidirectional GRU-LSTM model also performed exceptionally, achieving 98.69% accuracy, 96.16% precision, 94.62% recall, and 93.49% F1 score, outperforming many existing hybrid models in recommendation systems. By integrating forward and backward context, the Bidirectional GRU-LSTM model effectively captures complex temporal relationships, offering more accurate recommendations than traditional models that analyze data separately. This study underscores the robustness of LSTM-based architectures in sentiment analysis and highlights the potential of combining sentiment analysis with collaborative filtering to enhance precision and specificity in e-commerce recommendation systems.

Key words: Recommendation Systems, Collaborative filtering, Content-Based Filtering, Deep Learning, Bi-LSTM, Bi-LSTM-GRU.

1. Introduction. Introduction and examples. Users wanting to make wise decisions find it difficult to negotiate the enormous volume of data available online regarding service providers—such as hotels, restaurants, and merchandise. The abundant data might hinder decision-making procedures since people often struggle to sort through many possibilities to identify what really satisfies their needs. Recommendation systems (RSs) have been created to handle this problem by means of data processing simplification and tailored recommendations depending on user preferences. A restaurant recommendation system might, for instance, examine user dining behaviour and preferences to recommend nearby restaurants that fit them, therefore simplifying and accelerating the choosing process. Commonly used systems include content-based filtering (CBF), which suggests products based on their characteristics, and collaborative filtering (CF), which depends on the preferences of like users. Many recent RSs also apply hybrid approaches, combining CBF and CF to improve their recommendation accuracy. Online reviews have exploded as customer reviews progressively shape buying decisions. Many times, consumers base their decisions on the opinions and insights of others, such a review of a hotel stay or a purchase of a good or service. In this regard, sentiment analysis (SA) methods can be used to evaluate consumer perceptions of different services and products, therefore improving the quality of information retrieval by means of insightful analysis of the feedback. Using these technologies allows people to quickly obtain pertinent data and investigate hitherto unexplored fields, therefore improving the results of their decisions [22]). To solve this issue, recommendation systems have been created to sort data and give users individualized suggestions based on their preferencesBy harnessing these technologies, individuals can swiftly access information and delve into previously uncharted subjects, thereby enhancing the caliber of information retrieval services. Recommendation systems (RSs) frequently leverage collaborative filtering (CF), content-based filtering (CBF), and hybrid methodologies

*Faculty of Computer Science and Engineering, Department of Computer Science and Engineering, University Institute of Engineering and Technology (U.I.E.T), Kurukshetra University, Kurukshetra, Haryana, India (sdhawan2015@kuk.ac.in)

†Faculty of Computer Science and Engineering, Department of Computer Science and Engineering, University Institute of Engineering and Technology (U.I.E.T), Kurukshetra University, Kurukshetra, Haryana, India(ksingh2015@kuk.ac.in)

‡ PDepartment of Computer Science and Engineering, University Institute of Engineering and Technology (U.I.E.T), Kurukshetra University, Kurukshetra, Haryana, India (manoj200yadav@gmail.com)

that amalgamate both techniques to select items for recommendation [10], [2], [24]. In recent times, there has been a rise in the significance of customer reviews in the decision-making process over the use of services or purchases. Due to the fact that many buyers consider other people's opinions while making decisions, the quantity of customer evaluations that are placed online has significantly expanded. A customer's evaluation is based on their unique experiences using a certain service, such as renting a hotel, purchasing merchandise, or viewing a movie. In this situation, sentiment analysis (SA) techniques can be applied to gather information about customers' attitudes toward various concerns [19].

Sentiment Analysis seeks to identify the underlying sentiment of user-generated content about a given topic or entity [4]. To accomplish this, first determine whether overall tone of text is positive, negative, or neutral and then automatically extract relevant information about the entity being discussed. There are three tiers of data extraction at which Sentiment Analysis can be performed [26] a level, a sentence, and a whole document. The three primary strategies for addressing the SA issue are lexicon-based [28]. Machine learning-based [29] and hybrid approaches [25]. The first systems to be used for SA relied on dictionaries.

They fall into either the corpus-based or dictionary-based categories, depending on how they store and use lexicons and linguistic rules. Both classic and modern deep learning (DL) approaches are examples of Machine Learning (ML)-based techniques. Last but not least, a hybrid method integrates lexicons and machine learning tools [4]. It has been established that DL techniques applied to sentiment analysis are more effective than more traditional methods [14]. For sentiment categorization [30], can use a variety of Deep Learning models, including DNN stands for Deep Neural Network [1], [6], RNN (Recurrent Neural Network)[31], CNN (Convolutional Neural Network)[13],[5], LSTM (Long-Short Term Memory) [3], and GRU (Gated Recurrent Unit) [9]etc.

This work aims to enhance e-commerce recommendation systems by integrating collaborative filtering with sentiment analysis using deep learning. It introduces a novel Deep Learning based SA model tailored for this purpose, showing significant improvements in system efficiency and precision through empirical validation against baseline models.

2. Literature Review. Researchers are tackling data sparsity, cold start, and the gray sheep issue in collaborative filtering. Integrating sentiment analysis shows promise, with relevant studies and their contributions reviewed.

2.1. Recommendation System Based on Deep Sentiment Analysis. In [20] autoencoders and sentiment data enhance deep learning multi-criteria recommendations. LSTM and LSTM with Word2Vec excel in sentiment analysis. Their approach on the TripAdvisor dataset surpassed current methods, highlighting emotive data's crucial role.

In [9] the deep learning was employed for sentiment analysis to predict review sentiments for recommendations. LSTM and GRU models were evaluated using Amazon data, showing superior performance over traditional methods in the study.

In [22] the DSL-TR integrates CNN-BiLSTM for deep sentiment learning, incorporating review timestamps and factorization machine technique to predict ratings across Amazon datasets, outperforming alternatives.

In [21] is proposed new computer architecture enhances recommender systems with deep learning models integrating rating scores and emotional textual comments. Includes novel CFMDNN and MCNN models, showing strong performance across multilingual datasets (English and Arabic).

In [17] is presented a deep learning model-based recommendation system framework integrating textual comments and rating scores, overcoming existing limitations with models like MCNN and CFMDNN. Achieved strong performance across English and Arabic datasets.

In [32] are evaluated deep neural networks and collaborative filtering for enhanced recommendations, comparing LSTM, GRU, and a multilayer RNN to optimize system precision.

In [33] is proposed a deep learning framework that explores RNN for reviewer emotion analysis, scoring and recommending nearby locations based on social media reviews for accurate venue suggestions.

2.2. Collaborative and Content-Based Approaches. In [34] a Hybrid Neural Collaborative Filtering (HNCF) model integrates deep learning and interaction modeling, including Deep Multivariant Rating (DMR),

to enhance rating matrix-based recommender systems. It outperforms Yelp2014, Yelp2013, and IMDb in accuracy and reliability of top-n product suggestions.

In [7] it is proposed a genetic-based method identifies similar user preferences, creates suggestion spaces, and improves recommendation accuracy, recall, F-Measure, and Mean Absolute Error (MAE).

In [16] a hybrid recommendation strategy suggests tourist destinations using content analysis and collaborative filtering. It combines their strengths, using Cosine similarity and Singular Value Decomposition (SVD) for performance. Weighted hybridization of results outperforms singular CB and CF methods.

In [18] it is improved a vehicle-cargo matching on the VCM platform. The enhanced content-based VCM system recommends vehicles based on cargo owners' needs. A hybrid approach combining tag-based and collaborative filtering predicts shipper-driver ratings accurately, forming the PVCM model for personalized recommendations to shippers.

In [11] is proposed a streamlined method to recommend reading material for student courses, prioritizing information analysis in the administration recommendation system for integrated and rapid network asset growth.

In [10] is implemented a content-social filtering strategy introduces a novel document comparison method, distinguishing it from prior studies. Analysis validates its efficacy and practical implementation, outperforming Pure Collaborative Filtering and Singular Value Decomposition on real-world data.

2.3. Research Gaps. The literature underscores combining sentiment analysis and deep learning to enhance recommender systems. Practical application and scalability in real-world settings are critical due to reliance on simulated data. Further exploration across sectors like e-commerce, healthcare, and education is essential. Hybrid model investigations are needed beyond collaborative filtering and deep learning. User reception of these advanced systems remains pivotal for accuracy and satisfaction assessment.

3. Objecttives, motivation and research contributions.

3.1. Objectives.

- To analyze sentiment from Amazon and IMDb reviews to extract insights for developing a recommendation engine.
- To preprocess textual data to enhance accuracy and compatibility with sentiment analysis tools.
- To conduct exploratory data analysis (EDA) to understand the characteristics and patterns of the review datasets.
- To prepare data for training by encoding text information into numerical sequences and ensuring uniform input dimensions.
- To model with hybrid architectures combining LSTM, GRU, and other layers, optimizing hyperparameters for effective sentiment classification.

3.2. Motivation. This project addresses the challenge users face when choosing from numerous online options. Recommender Systems (RSs) like CF, CBF, and hybrids mitigate issues of sparsity and unreliable recommendations. Sentiment Analysis (SA) now influences buyer decisions by analyzing user-generated content using deep learning models like LSTM, GRU, CNN, and Bidirectional networks, improving classification results.

3.3. Research Contributions. The paper introduces a novel recommendation system (RS) integrating sentiment analysis (SA) techniques driven by deep learning to enhance the precision and uniqueness of e-commerce recommendations. By combining SA with collaborative filtering through hybrid deep learning methods like LSTM, GRU, and Bi-LSTM, the study significantly improves recommendation quality using e-commerce data.

4. Methodology. Here present the proposed deep learning methods for the recommender system based on sentiment analysis. Figure 4.1 depicts a schematic of the system's layout. The architecture makes it simple to set up the modules and how they interact, letting the app be pieced together using any of the available approaches. The structure consists of six individual elements (data collection, data pre-processing, Exploratory Data Analysis (EDA), Training, Modelling, and model evaluation). The reviews' raw data was applied to conduct experiments and train hybrid DL models on sentiment analysis.

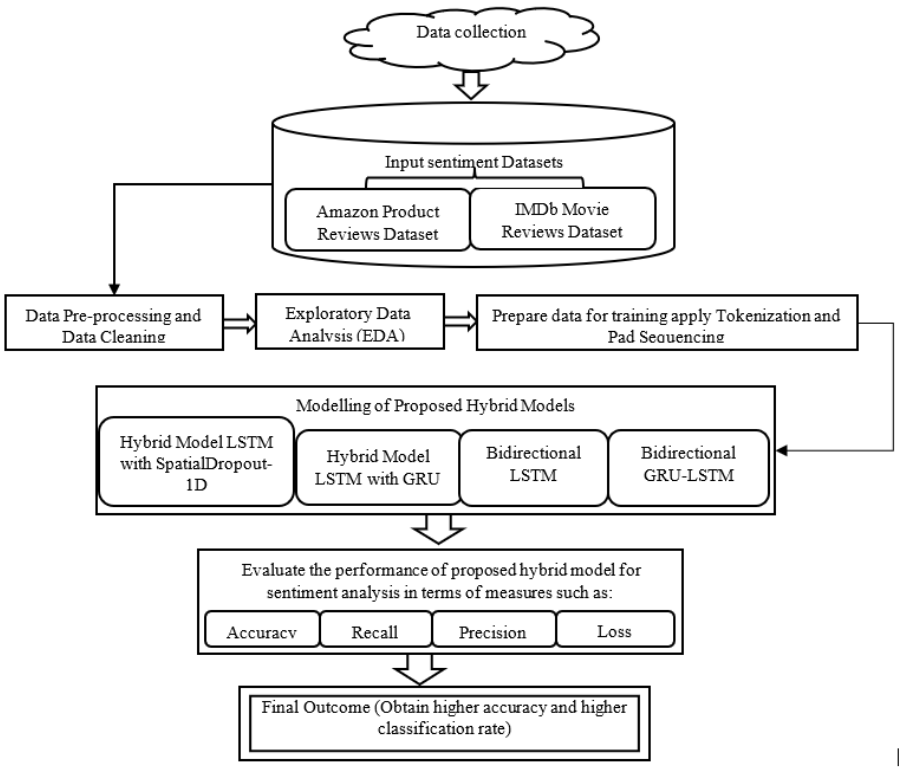

Fig. 4.1: The Proposed Methodology Architecture.

The recommended approach for the sentiment analysis-based recommendation system is depicted in Fig. 4.1. The various elements of this system, such as data collecting, data pre-processing, exploratory data analysis (EDA), classification models, and performance assessment, are elaborated upon below with thoroughness. The process is outlined in the following manner:

4.1. Data Collection. This work utilized two publicly accessible online sentiment datasets, namely Amazon's product review dataset and IMDb's movie review dataset. Together, the datasets used in this study—Amazon's product review dataset and IMDb's movie review dataset represent a great variety of samples vital for sentiment analysis. Comprising tens of thousands of evaluations, the Amazon product review dataset reflects a broad spectrum of consumer attitudes and rates from 1 to 5. This range helps to classify emotions into negative, neutral, and positive categories, so offering a whole picture of consumer satisfaction. Conversely, the IMDb movie review dataset comprises 50,000 reviews, therefore highlighting a range of viewpoints on a wide spectrum of films. Natural language processing (NLP) applications benefit especially from this dataset since it includes many genres and styles of writing, therefore strengthening the sentiment classification models' resilience. Training and validation of the hybrid models used in this research depend on the variety and volume of these datasets, which guarantees that they can generalise well across many contexts and fairly predict sentiments in both product and movie evaluations.

Amazon product review dataset. The research aims to analyze Amazon customer reviews to develop a product recommendation engine by extracting insights, including positive and negative sentiments. The process involves incorporating attribute-based reflective elements into the analysis. The dataset is a subset of Amazon.com reviews, randomly selected and numbering in the tens of thousands, with ratings from 1 to 5 indicating satisfaction levels. Ratings 1 and 2 are categorized as negative feedback, 4 and 5 as positive, and 3 as neutral

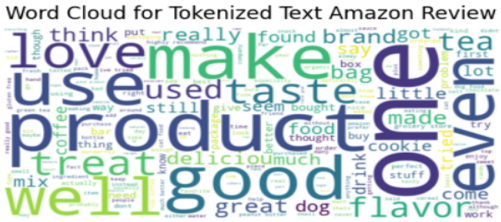

Fig. 4.2: Word cloud view of the text Amazon Movie review dataset

for sentiment classification experiments [35].

IMDb Movies Review dataset. The study applies sentiment analysis to IMDb's dataset of 50,000 movie reviews, showcasing insights from textual data. This extensive dataset offers a substantial resource for NLP and text analytics, providing a benchmark for binary sentiment classification using various methods including deep learning techniques.

4.2. Data Pre-processing and Cleaning. Data preparation involves transforming raw data into a suitable format. Initially, undesirable attributes are removed to enhance the quality of textual data through NLP text pre-processing. This includes eliminating stop words like "your," "I," and "it," which do not contribute to sentiment analysis. Emojis are cleaned by converting text to lowercase, removing carriage returns, and substituting non-UTF-8 characters. Hashtags are separated from words, and punctuation and non-alphabetic characters are filtered out from reviews. This rigorous process ensures consistency and improves the text's suitability for natural language processing applications, facilitating accurate sentiment analysis and model training.

To improve the quality of the input data, the data pretreatment processes comprised tokenising, text normalising, and stop word removal. Text normalising meant changing all text to lowercase and fixing spelling mistakes so guaranteeing consistency. Tokenising the book into separate words or phrases made analysis simpler. To cut noise and concentrate on meaningful words that support sentiment, stop words including "and," "the," and "is," were deleted. Furthermore used to rank significant features depending on their significance in the context of sentiment analysis were feature selection methods such TF-IDF (Term Frequency-Inverse Document Frequency), hence enhancing the model performance.

Methods were used to add variability in the current samples therefore improving the dataset. These methods assisted to vary the training data by synonym replacement, random insertion, and back-translation. This improved the robustness of the model so that, in sentiment analysis, it could effectively generalise and operate on unseen data.

4.3. Exploratory Data Analysis (EDA). Bar charts, histograms, and box plots are just some of the data visualization techniques that EDA uses. Create a heat map to visually represent the relationships (correlations) between different variables. See below for a summary and visual representation of both datasets.

Fig. 4.2 displays topical word clouds derived from datasets discussed in Section 4.1 word cloud visually represents text data, where the size of each word corresponds to its frequency. Common words, such as "product," "one," and "good," appear larger.

Fig. 4.3 shows a bar graph shows the top 20 most frequent words in Amazon review texts. The x-axis displays word frequency, with "like" being the most common and "time" the least, descending from left to right.

Fig. 4.4 horizontal box plot shows Amazon review lengths, with a 2000-character median, IQR, and whiskers highlighting outliers up to 7000 characters.

Fig. 4.5 shows histogram of text length for Amazon reviews, The x-axis signifies text length, and the y-axis signifies frequency. The histogram is blue in color. It has a peak at around 1000 text length. The histogram also has a long tail towards the right

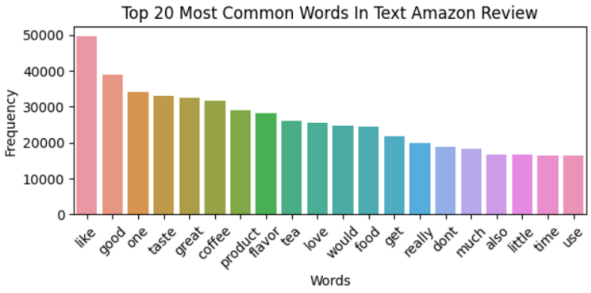

Fig. 4.3: Bar graph of top 20 most common words in text amazon review dataset.

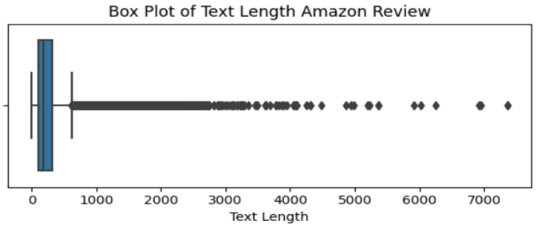

Fig. 4.4: Box Plot of text length amazon review dataset

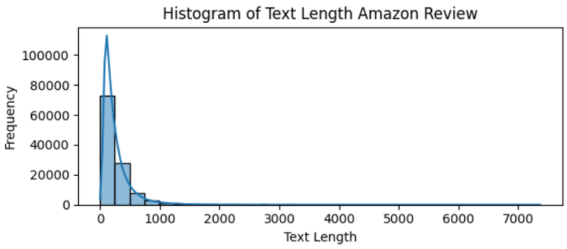

Fig. 4.5: Histogram of text length amazon review dataset.

Fig. 4.6 is a pie chart of the most frequent words in the Amazon review dataset. The chart is divided into 10 sections, each representing a different word. The words included are: “good”, “one”, “taste”, “great”, “coffee”, “product”, “flavour”, “tea”, “love”, and “like”. The largest section corresponds to the word “like” at 15.2%, followed by “good” at 11.7% and “love” at 9.6%. The smallest section represents the word “tea” at 8.01%.

Fig. 4.7 presents the IMDb movie review text dataset as a word cloud. Word cloud based on IMDb movie reviews’ most frequently used terms, including ”film,” ”movie,” ”one,” ”character,” ”scene,” ”good,” and so on. Furthermore, Fig. 4.9 presents a compilation of the top 20 words that occur most frequently within the IMDb Movie Reviews dataset.

Fig. 4.9 The Box Plot of IMDb Movie Review Text Length is displayed. The words in the dataset and their relative frequency are displayed as x and y axes, respectively, in the figure. Words associated with movies appear more often than others, such as ”make,” ”fire,” ”two,” ”character,” etc.

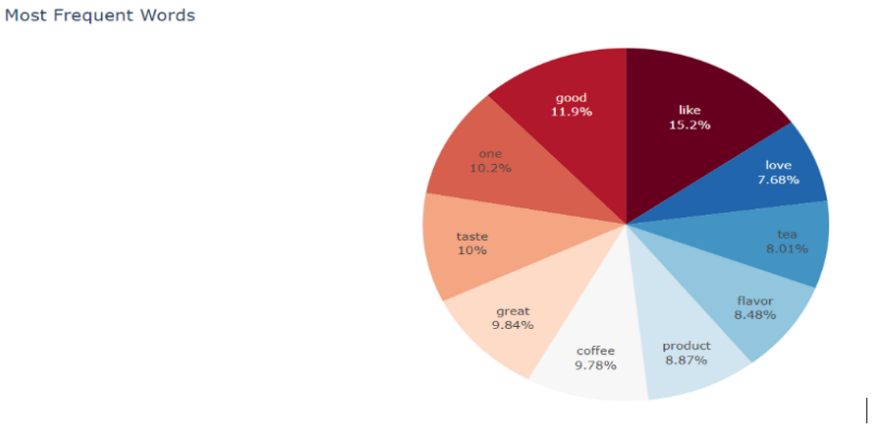

Fig. 4.6: Pie chart of most frequent words of amazon review dataset.

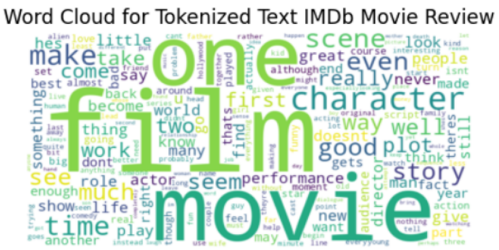

Fig. 4.7: Word cloud view of the text IMDb Movie review dataset

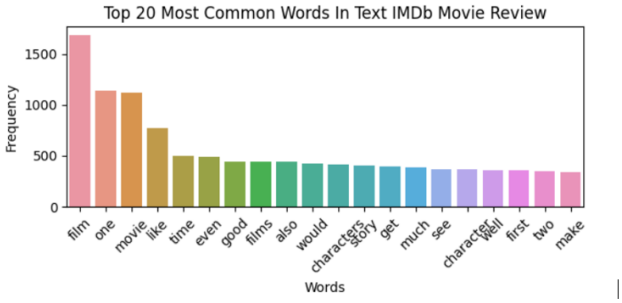

Fig. 4.8: Bar graph of top 20 most common words in text IMDb Movie review dataset.

4.4. Prepare Data for Training. Used Pad Sequencing and Tokenization on training data. The primary goal is to get text data ready for NLP work.

At first, take the text information from the DataFrame 'df' and put it into the variable 'review.' Subsequently, a tokenizer is used via the `Tokenizer` class, with a cap of 10,000 words as defined by 'num_words.' After the text is processed and a vocabulary is built, the tokenizer is "fit" on the "review" data. With the number of distinct words in the text data in mind, the variable 'vocab_size' is determined to be the length of

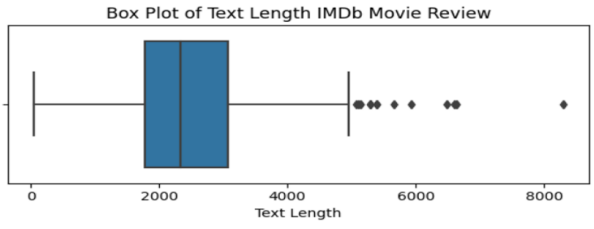

Fig. 4.9: Box Plot of text length IMDb Movie review dataset.

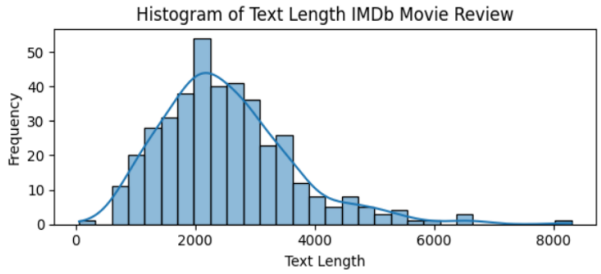

Fig. 4.10: Histogram of text length IMDb Movie review dataset.

the word index plus one.

Using the `tokenizer.texts_to_sequences` method, the text information is encoded into numerical sequences. The vocabulary is indexed according to each word in the 'review' data. The sequences have a maximum length of 200 and are padded using `pad_sequences` to guarantee uniform input dimensions for deep learning models. When dealing with sequences of varying lengths, this step is especially important because it guarantees that all data will be processed in the same way.

4.5. Modelling with Hybrid models. Hybrid Model LSTM with Spatial Dropout-1D; Hybrid Model LSTM with GRU; Hybrid Model LSTM with RNN; Hybrid Model RNN with LSTM; Hybrid Model RNN with LSTM; Hybrid-Training LSTM and Hybrid-Training GRU-LSTM. Key steps for putting into practice the models discussed below.

SpatialDropout1D. SpatialDropout1D is a specialized dropout layer used in deep neural networks for processing sequential data like time series or natural language. Unlike traditional dropout, SpatialDropout1D applies dropout independently to each feature along the time dimension (sequence length). This approach reduces the network's reliance on any single feature at a given time, fostering broader and more robust feature representations. During training, it generates a binary dropout mask based on a specified dropout rate, selectively zeroing out input features to introduce randomness and prevent overfitting. In inference, dropout is inactive, allowing the input to pass through unaltered. SpatialDropout1D effectively mitigates overfitting in models handling sequential data, particularly beneficial in tasks such as natural language processing where preserving word order is critical. Its adaptation enhances model performance by ensuring diverse feature utilization across sequences.

Hybrid Model LSTM with SpatialDropout1D. The implementation of the Hybrid model (LSTM with SpatialDropout1D) begins by initializing 'EmbeddingVectorLength' to 32, which defines the dimensionality of word embeddings used for NLP inputs. A Sequential model is then constructed, facilitating the sequential layering of neural network components. An embedding layer is added next to convert vocabulary words ('Vocab Size') into embedding vectors of specified length, with input sequences limited to 200 characters. Following this, a 0.25-dropout SpatialDropout1D layer is incorporated to randomly deactivate neurons during training, mitigating overfitting. A 50-unit LSTM layer follows, renowned for its effectiveness in processing sequential data, with

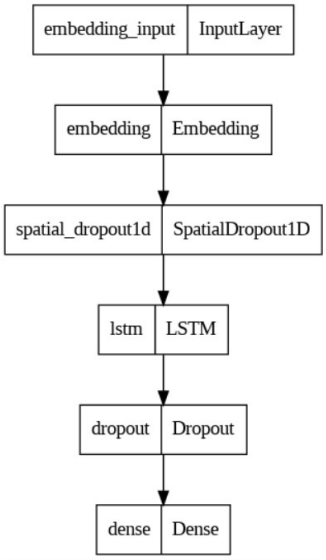

Fig. 4.11: The model architecture of hybrid model(LSTM with SpatialDropout-1D).

dropout rates set for both inputs and recurrent connections at 0.5. Lastly, a dense layer comprising a single unit with a sigmoid activation function is added for binary classification tasks, producing output probabilities ranging from 0 to 1. The model is compiled using binary cross-entropy loss, Adam optimizer, and accuracy metrics, ideal for binary classification tasks in NLP and other domains.

Fig. 4.11 architecture depicts a real-world application of a hybrid model consisting of and SpatialDropout1D. This schematic depicts how data moves through the model: Embedding the word embedding input layer is represented by the InputLayer class. The input tokens are subjected to embedding, a process that entails transforming them into dense vectors referred to as word embeddings. SpatialDropout1D is a dropout technique that eliminates complete 1D feature maps across channels, rather than individual elements. Dropout is a regularization method used to address the issue of overfitting. Densely interconnected layer. Dropping entire feature maps (along channels) instead of individual elements is what SpatialDropout1D is all about, and it makes for superior regularization of sequential data.

4.5.1. Hybrid Model LSTM with GRU. The construction of a Sequential model begins with an embedding layer that maps words to numerical vectors of specified dimensions ('EmbeddingVectorLength') and vocabulary size ('Vocab Size'). Input sequences are capped at 200 characters. A 0.25-percent-dropout SpatialDropout1D layer is added to prevent overfitting in sequence data. The model incorporates both GRU and LSTM layers, each with 25 units, to handle sequential input effectively. Dropout rates for inputs and recurrent connections are set to 0.4 to enhance model generalization. The LSTM layer also utilizes dropout regularization. Finally, an additional Dropout layer with a 0.2 dropout rate further mitigates overfitting concerns in the model. This architecture is tailored for NLP tasks requiring robust handling of sequential data.

Bidirectional GRU-LSTM models are computationally intensive yet crucial for capturing bidirectional dependencies in sequential data. They excel in sentiment analysis and recommendation systems, enhancing accuracy despite their complexity. To mitigate high variance, regularization methods like L2 regularization penalize large parameter values, preventing overfitting. Dropout further strengthens model robustness by randomly deactivating neurons during training. Ensembling techniques such as bagging or boosting combine predictions from multiple models to reduce variance. Cross-validation assesses model performance across diverse subsets, ensuring robustness. These strategies collectively enhance model generalization, enabling effective predictions on unseen data and improving overall performance in recommendation systems.

Several hybrid models are used in this work to improve the performance of sequential data processing tasks

Table 4.1: Hyperparameters Details

Parameter	**Details**
Model	Sequential
Layers	LSTM, GRU, SpatialDrop1D, Bi-LSTM, Bi-GRU-LSTM
Dropout	20%
Activation	Sigmoid
Loss	Binary Cross Entropy
Epochs	20
Validation Split	25%
Batch Size	1024

including sentiment analysis and recommendations by especially combining Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), and 1D Convolutional Neural Networks (1D CNN). Both LSTM and GRU, both variations of recurrent neural networks (RNN), shine in sequential data in capturing long-term dependencies. While GRU, a condensed form, performs quicker with less parameters, LSTM with its memory cells efficiently addresses vanishing gradient problems. Conversely, the 1D CNN is incorporated to detect local patterns in the sequence data, thereby detecting important elements before forward sending them to the recurrent layers. Usually serving as a feature extractor in the hybrid models, 1D CNN analyses input sequences to gather pertinent local information. LSTM or GRU layers then receive this output to record temporal dependencies, therefore offering a complete knowledge of both local and sequential properties. Combining 1D CNN with LSTM and GRU helps the model to leverage the strengths of both convolutional and recurrent networks, hence enhancing performance in applications like sentiment analysis, where knowledge of local features and long-term dependencies is absolutely vital. It is imperative to establish the originality of the suggested design and can be achieved by means of a comparative study with current recommendation systems. This study will show how the suggested deep learning models—such as the Bidirectional GRU-LSTM—outperform conventional approaches in terms of accuracy and efficiency, so stressing their special contributions to the discipline of e-commerce suggestions. The justifications for using particular techniques, such hybrid models including bidirectional GRU-LSTM and LSTM with Spatial Dropout. This covers talking about how these techniques guarantee correct content recommendations by efficiently tackling the difficulties of user profile and sentiment analysis. Furthermore underlined should be issues of the resilience of these models in managing various datasets, such as Amazon and IMDb. By clearly stating these decisions, you lay a strong basis for your research and show how each approach improves the general analysis and fits the aims of the study.

LSTM, GRU, and 1D CNN's hybrid technique was selected above conventional models because of its better capacity to detect intricate patterns and dependencies in sequential data, including user reviews and feelings. Long-term dependencies are handled effectively by LSTM and GRU, therefore reducing the vanishing gradient issue typical in conventional recurrent networks. Combining these with 1D CNN improves feature extraction by spotting local patterns within the data, so producing better input sequence representation. More subtle sentiment analysis made possible by this integrated architecture produces more accurate and successful suggestions in e-commerce platforms.

i. Activation function. Activation functions like sigmoid, ReLU, and softmax determine neuron activation in neural networks by processing inputs to produce output values. They control a neuron's behavior based on its input and activity state. The sigmoid function, chosen here, limits outputs to a range between 0 and 1, making it ideal for binary classification tasks. It computes the likelihood of an input belonging to a category, crucial for the final layer in such models[26]:

$$Sigmoid = \frac{1}{1 + e^{-x}} \dots (1) \tag{4.1}$$

ii. Loss function. The loss function, also known as the error function, evaluates model performance by comparing predicted outputs to actual labels using training data. It's critical to choose a metric that accurately

measures model effectiveness. For deep neural networks (DNNs), tailored loss functions address specific objectives, such as minimizing classification errors. In this case, binary cross-entropy was used for straightforward binary classification, producing a probability value between 0 and 1 as output. The formula for binary cross entropy is $-1 \times \log(\ textprobability)$ [36].

$$Loss = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - p(y_i)) \tag{4.2}$$

Here, Yi is the real class and log(p(yi) is the probability that Yi belongs to that class.

- p(yi) is the probability of one
- 1-p(yi) is the probability of zero

iii. Batch size. This accounts for the training dataset that is specific to each batch. The process of iterating is dependent on the dataset, which necessitates dividing it into successive batches due to the impracticality of feeding a single epoch into the computer in its entirety.

iv. Epoch. An epoch is considered complete when all the data has been both fed into and outputted from the neural network simultaneously. The parameters of a neural network can be adjusted by an iterative process of supplying it with training data. After each iteration, the parameter is revised. Generally, extending the number of epochs results in enhanced accuracy and less loss.

4.6. Performance Evaluation Metrics. In evaluating a machine learning model's effectiveness for sentiment analysis using deep learning on Amazon product and IMDb movie review datasets, performance metrics like accuracy, precision, recall, and loss were assessed. Models were selected based on minimal validation loss and tested using a confusion matrix to compare classification results, including True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN).

Accuracy. The accuracy of a classifier is evaluated by calculating the ratio of correct classifications to the total number of predictions. In this work model gets 98% and 100% accuracy on both datasets. The formula for precision is given by the following equation:

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \tag{4.3}$$

Precission. Precision in classification refers to the proportion of correctly identified true positives. In this work model gets 98% and 100% precision on both datasets. The equation can be represented accurately as:

$$Precision = \frac{TP}{TP + FP} \tag{4.4}$$

Recall. Recall is calculated by dividing the number for true positives by the sum of true positives & false negatives. Precisely measuring the number of positive designations is essential in this situation. In this work model gets 94% and 100% recall on both datasets. The formula for memory is as stated:

$$Recall = \frac{TP}{TP + FN} \tag{4.5}$$

F1 score. The F1 score is a statistical measure that balances precision and recall, providing a single metric for model performance evaluation.

$$F1 = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} \tag{4.6}$$

Epoch and loss curve. These curves are graphed over time and help interpret the model-training process by showing how the parameters of epoch and loss change over time. In this work model gets 0.04% and 0.012% loss on both datasets.

Table 5.1: Comparison of deep learning models on Amazon Product Review Dataset

Deep Learning Model	Accuracy	Precision	Recall	F1 Score	Loss
LSTM Spatial Dropout1D	98.01	94.00	92.04	93.00	0.05
Hybrid LSTM+GRU	98.04	93.60	92.70	91.72	0.05
Bidirectional LSTM	97.77	93.28	91.16	91.66	0.06
Bidirectional GRU-LSTM	98.69	96.16	94.62	93.49	0.04

Table 5.2: Comparison of deep learning models on IMDb Movie Review Dataset

Deep Learning Model	Accuracy	Precision	Recall	F1 Score	Loss
LSTM Spatial Dropout1D	96.20	97.31	98.26	97.78	0.012
Hybrid LSTM+GRU	93.72	81.67	72.80	76.98	0.022
Bidirectional LSTM	96.39	90.02	84.37	87.10	0.024
Bidirectional GRU-LSTM	93.47	81.02	70.80	75.57	0.012

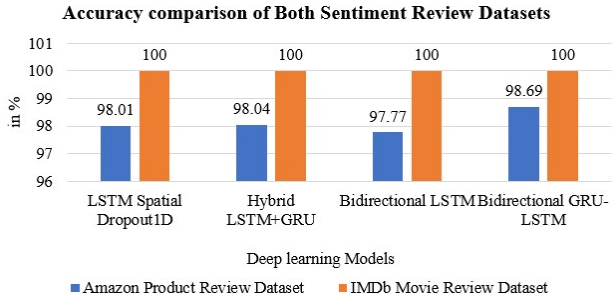

Fig. 5.1: Bar graph of Accuracy comparison between both datasets with deep learning models

5. Comparative analysis with Discussion. In this section, the empirical results used to assess the efficacy of the proposed approach for constructing recommender systems are presented. The findings of the investigation are documented in detail. Empirical data was acquired using Google Colab, leveraging Python, which is widely recognized as a leading programming language for machine learning and data analysis due to its extensive library support. For this study, Python was chosen for its accessibility and robust libraries, including NumPy, pandas, seaborn, scikit-learn, matplotlib, Keras, and TensorFlow.

The datasets used in this study included Amazon product reviews from Amazon website and movie reviews from IMDb. These datasets were partitioned into training and test sets. Sentiment data was classified and analyzed using various deep learning models, including Hybrid Model LSTM with SpatialDropout-1D, Hybrid Model LSTM with GRU, Bidirectional LSTM, and Bidirectional GRU-LSTM.

The accuracy metric, representing the percentage of correctly classified instances in the validation set, was used to evaluate the models. The experiments, summarized in tables 5.1,5.2, present the accuracy of several deep learning methods. The results demonstrate the effectiveness of the proposed models in accurately categorizing sentiment data, validating their potential for improving recommendation systems.

The evaluation will be conducted on two separate datasets. Reviews of products on Amazon or reviews of movies on IMDb. In addition, the comparison results are shown in bar graph form in Fig. 5.1,5.2,5.3,and 5.4.

The table 5.1,5.2,and bar charts depict the divergent levels of effectiveness between deep learning models that were trained on either the Amazon Product Review Dataset or the IMDb Movie Review Dataset.

The examination of deep learning models on the IMDb Movie Review Dataset and the Amazon Product Review Dataset offers interesting insights about the efficiency of several architectures for sentiment analysis

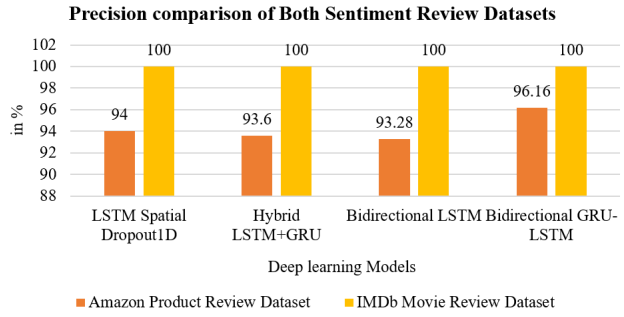

Fig. 5.2: Bar graph of Precision comparison between both datasets with deep learning models

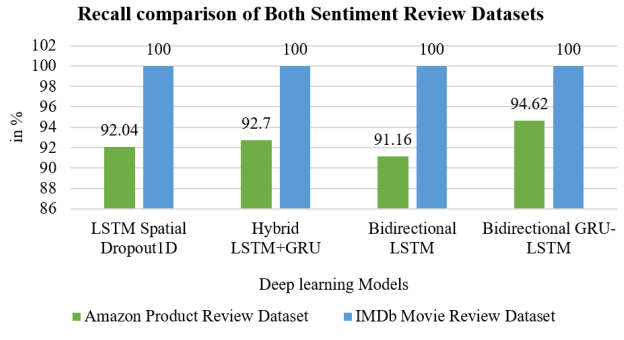

Fig. 5.3: Bar graph of Recall comparison between both datasets with deep learning models

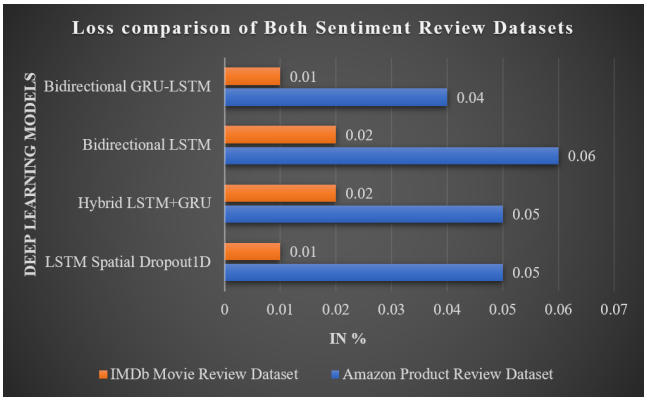

Fig. 5.4: Bar graph of Loss comparison between both datasets with deep learning models

problems. With an astounding accuracy of 93.00%, the LSTM Spatial Dropout1D model ranked first in the first table, which summarises the performance on the Amazon dataset. With a precision score of 90.73% and a recall of 95.38%, this model showed an amazing capacity to strike a balance between accuracy and recall, producing an F1 score of 93.00%. With an accuracy of 93.50% and a greater precision of 92.57%, the Bidirectional GRU-LSTM model quite faithfully followed. Its recall of 94.44% and F1 score of 93.49% show how well it reduces false positives while preserving a great capability for spotting true positive cases. Though they did not outperform

the top models in general efficacy, the other models—including Hybrid LSTM+GRU and Bidirectional LSTM showed satisfactory performance. On the other hand, IMDb dataset results showed much better performance measures. Once more leading with an accuracy of 96.20%, accompanied by an amazing precision of 97.31% and recall of 98.26%, the LSTM Spatial Dropout1D model produced an F1 score of 97.78%. This shows that the model not only fairly projected the results but also performed exceptionally well in spotting genuine favourable attitudes among the evaluations. Though it showed a clear decline in precision and recall when compared to the LSTM Spatial Dropout 1D model, the Bidirectional LSTM model also performed admisitely, with an accuracy of 96.39%. The Hybrid LSTM+GRU model suffered notably, especially on the IMDb dataset, with a precision of only 81.67% and a recall of 72.80%, so highlighting its shortcomings in efficiently capturing sentiment nuances in challenging text data.

Indicating its resilience in sentiment analysis tasks, the study highlights generally the advantage of the LSTM Spatial Dropout1D model across both datasets. Its capacity to control overfitting and efficiently capture contextual information is shown by the always high performance measures over several contexts. These results imply that deep learning models—especially those using LSTM architectures with dropout techniques—are suited for sentiment analysis and future study could investigate optimisations and advanced architectures to improve performance even further in varied textual datasets.

Deep learning models on the IMDb Movie Review Dataset and the Amazon Product Review Dataset show diverse patterns and efficacy across several architectures in performance comparison. With an accuracy of 93.00%, the LSTM Spatial Dropout1D model stands out in the Amazon dataset showing a solid balance between precision (90.73%) and recall (95.38%), hence producing an F1 score of 93.00%. With an accuracy of 93.50% and underlining its capacity to reduce false positives while keeping a high true positive rate, the Bidirectional GRU-LSTM model rather closely follows. On the other hand, the IMDb dataset shows even better overall performance; the LSTM Spatial Dropout1D model once again leads at 96.20% accuracy and remarkable precision of 97.31% and recall of 98.26%, therefore producing an F1 score of 97.78%*. Though the Hybrid LSTM+GRU model underperforms greatly, especially in precision and recall, suggesting its limits in catching emotion nuances, the Bidirectional LSTM model also performs well.

Consistent performance across both datasets shows that the LSTM Spatial Dropout1D model is resilient for sentiment analysis tasks overall.

Unexpected outcomes can offer insightful analysis and point out areas of interest for more investigation in both scholarly fields and useful applications. If a model meant to improve sentiment classification, for example, shows lower-than-expected performance on a particular dataset, it could point to underlying data complexity—such as confusing language or varied contexts not sufficiently recorded. This insight implies that to improve model resilience, more complex feature extraction techniques or the inclusion of more data sources is necessary. Such results can inspire academics to look at alternate approaches such using ensemble techniques or transfer learning or challenge the limits of current methods. Practically speaking, knowing these surprising results will help companies rethink their user engagement plans or product recommendations so they better fit user attitude. Overall, acknowledging and evaluating surprising outcomes helps one to have a constant improvement attitude, so opening the path for developments in academic study as well as practical applications.

Presenting the results with confidence intervals or error bars will help to improve the dependability of the conclusions. Error bars can graphically show the variability and uncertainty in the measurements for every performance metric—accuracy, precision, recall, F1 score—that model uses. Usually, statistical techniques allow one to determine a 95% confidence interval, so clarifying the range within which the actual model performance is like to fall. This method not only helps to find statistically significant variations between models but also guarantees a more complete interpretation of the data by strengthening the general conclusions derived from the investigation.

The model integrating RoBERTa, Long Short-Term Memory (LSTM), Bidirectional LSTM (BiLSTM), and Gated Recurrent Unit (GRU), achieves an accuracy of 94.9%, with precision, recall, and F1 score of 95 each. This hybrid technique combines the strengths of its components: RoBERTa's Transformer-based architecture for enhanced context awareness, LSTM and BiLSTM's sequence modeling and long-term dependency capture, and GRU's efficiency in training while maintaining high performance in sequential tasks. However, the proposed Bidirectional GRU-LSTM model demonstrates a significant improvement, achieving 98.69% accuracy, 96.16%

Table 5.3: Comparative Analysis of Existing Model and Proposed Highest Performance Model

Model	Accuracy	Precision	Recall	F1 Score
Hybrid (RoBERTa) [37]	94.9	95	95	95
Bidirectional GRU-LSTM	98.69	96.16	94.62	93.49

precision, 94.62% recall, and an F1 score of 93.49%. By processing input bidirectionally, this model captures both forward and backward dependencies, leveraging the strengths of GRU and LSTM architectures to better handle intricate patterns and temporal linkages in data. The superior accuracy indicates the Bidirectional GRU-LSTM's ability to handle nuanced dataset subtleties, making it the best-performing model in this study. Moreover, it surpasses many existing hybrid recommendation systems, which often separately analyze user behavior, item characteristics, and sentiment data. The Bidirectional GRU-LSTM's integration of bidirectional context processing enables richer, dynamic insights, enhancing recommendation precision by accounting for complex temporal interactions.

6. Discussion. This work shows the great promise of deep learning models—especially LSTM-based architectures—in sentiment analysis over several datasets. Showcasing their resilience in precisely categorising emotions, the LSTM Spatial Dropout 1D and Bidirectional GRU-LSTM models fared quite well. Particularly the Bidirectional GRU-LSTM model improved recommendation precision by capturing both forward and backward contextual dependencies, hence outperforming conventional hybrid models applied in recommendation systems. This work also emphasises the interesting way sentiment analysis may be combined with collaborative filtering in e-commerce systems to improve the relevance and accuracy of suggestions in dynamic, real-world environments.

7. Conclusion. Through a focus towards enhancing content recommendations in streaming services including movies and merchandise, this paper discusses the difficulties of building reliable user profiles for recommendation systems. The work shows the ability of deep learning models such as LSTM with Spatial Dropout-1D, LSTM+GRU, bidirectional LSTM, and bidirectional GRU-LSTM in improving sentiment analysis and recommendation accuracy by means of their respective application in Particularly the bidirectional GRU-LSTM model shines in capturing bidirectional relationships in data, hence enhancing the quality of recommendations by considering both forward and backward temporal dependencies. Preventing overfitting and guaranteeing the resilience of the model depend on methods including regularisation and cross-valuation. Understanding complex emotions, such irony, which emphasises the need of constant model improvement, still presents difficulties, nevertheless. The results of this work highlight, especially via sentiment analysis, the possibility of including deep learning techniques into recommendation systems. With its capacity to capture intricate temporal correlations and so reflect the high performance of the bidirectional GRU-LSTM model, above conventional hybrid models it demonstrates its great advantage. Future studies might look at using feature selection methods and ensemble learning to raise accuracy and flexibility even more. Expanding sentiment analysis into other fields, such banking and healthcare, might offer insightful information impacting decision-making procedures. Incorporating multimodal data sources—such as audio and visual materials—may also improve sentiment classification sensitivity. Furthermore vital for enhancing model interpretability, fostering user confidence, and guaranteeing recommendation system transparency is investigating explainable artificial intelligence (XAI). The limits of the study, which depend on publically accessible datasets, draw attention to the possible differences in real-world user behaviour and the necessity of honing models for different cultural environments to guarantee worldwide applicability. Notwithstanding these constraints, the work provides insightful analysis on how to improve recommendation system performance by means of sophisticated deep learning methods.

Future Coverage. The results of this work open various directions for next investigations. Extending the sentiment analysis paradigm to fields like banking and healthcare, where knowledge of user sentiment can greatly influence decision-making procedures, seems to be one bright future path. Including multimodal data sources—such visuals and audio—alongside text could further improve sentiment categorisation sensitivity. Improving model interpretability and user confidence also depends on researching cutting-edge technologies

such explainable artificial intelligence (XAI). At last, analysing the performance of the model in many cultural settings helps to highlight the subtleties of sentiment expression, hence guiding the creation of more flexible and successful recommendation systems.

Limitations. The fact that this study depends on publicly accessible data limits it in that it might not fairly depict actual user behaviour. Furthermore, the emphasis on sentiment analysis could cause one to ignore other elements impacting advice. The performance of the model may also change depending on the cultural setting, therefore influencing its generalisability and efficiency. Strength & weakness. Indicating strong sentiment analysis resilience, the LSTM Spatial Dropout 1D model shines in accuracy, precision, and recall over both datasets. With much lower precision and recall, the hybrid LSTM+GRU model highlights difficulties in precisely capturing emotional nuances in difficult texts.

REFERENCES

[1] A. Seilsepour, R. Ravanmehr, and R. Nassiri, *Topic sentiment analysis based on deep neural network using document embedding technique*, J. Supercomput., vol. 79, no. 17, pp. 19809–19847, 2023, doi: 10.1007/s11227-023-04514-y.

[2] M. Nilashi, O. Ibrahim, and K. Bagherifard, *A recommender system based on collaborative filtering using ontology and dimensionality reduction techniques*, Expert Syst. Appl., vol. 92, pp. 507–520, 2018, doi: 10.1016/j.eswa.2017.10.058.

[3] S. D. Gogula, M. Rahouti, S. K. Gogula, A. Jalamuri, and S. K. Jagatheesaperumal, *An emotion-based rating system for books using sentiment analysis and machine learning in the cloud*, Appl. Sci., vol. 13, no. 2, p. 773, 2023, doi: 10.3390/app13020773.

[4] M. Birjali, M. Kasri, and A. Beni-Hssane, *A comprehensive survey on sentiment analysis: Approaches, challenges and trends*, Knowl.-Based Syst., vol. 226, p. 107134, 2021, doi: 10.1016/j.knosys.2021.107134.

[5] S. Manikandan, P. Dhanalakshmi, K. C. Rajeswari, and A. D. C. Rani, *Deep sentiment learning for measuring similarity recommendations in twitter data*, Intell. Autom. Soft Comput., vol. 34, no. 1, pp. 183–192, 2022, doi: 10.32604/iasc.2022.015452.

[6] P. Sugumaran and A. B. B. K. Uma, *Real-time twitter data analytics of mental illness in COVID-19: sentiment analysis using deep neural network*, Indones. J. Electr. Eng. Comput. Sci., vol. 26, no. 1, pp. 560–567, 2022, doi: 10.11591/ijeecs.v26.i1.pp560-567.

[7] I. Bouacha and S. Bekhouche, *An Evolutionary Based Recommendation Approach*, in 2021 International Conference on Theoretical and Applicative Aspects of Computer Science (ICTAACS), pp. 1–9, 2021, doi: 10.1109/ICTAACS53038.2021.9594028.

[8] I. Karabila, N. Darraz, A. El-Ansari, N. Alami, and M. El Mallahi, *Enhancing collaborative filtering-based recommender system using sentiment analysis*, Future Internet, vol. 15, no. 7, p. 235, 2023, doi: 10.3390/fi15070235.

[9] H. Zarzour, M. Al-Ayyoub, Y. Jararweh, et al., *Sentiment analysis based on deep learning methods for explainable recommendations with reviews*, in 2021 12th International Conference on Information and Communication Systems (ICICS), pp. 452–456, 2021, doi: 10.1109/ICICS52457.2021.9464597.

[10] A. Pal, P. Parhi, and M. Aggarwal, *An improved content based collaborative filtering algorithm for movie recommendations*, in 2017 Tenth International Conference on Contemporary Computing (IC3), pp. 1–3, 2017, doi: 10.1109/IC3.2017.8284338.

[11] M. Kommineni, P. Alekhya, T. M. Vyshnavi, V. Aparna, K. Swetha, and V. Mounika, *Machine learning based efficient recommendation system for book selection using user based collaborative filtering algorithm*, in 2020 Fourth International Conference on Inventive Systems and Control (ICISC), pp. 66–71, 2020, doi: 10.1109/ICISC47916.2020.9171208.

[12] C. N. Dang, M. N. Moreno-García, and F. De la Prieta, *An approach to integrating sentiment analysis into recommender systems*, Sensors, vol. 21, no. 16, p. 5666, 2021, doi: 10.3390/s21165666.

[13] H. An and N. Moon, *Design of recommendation system for tourist spot using sentiment analysis based on CNN-LSTM*, J. Ambient Intell. Humaniz. Comput., vol. 13, no. 3, pp. 1653–1663, 2022, doi: 10.1007/s12652-021-03474-1.

[14] W. Zhao, Z. Guan, L. Chen, X. He, D. Cai, B. Wang, and Q. Wang, *Weakly-supervised deep embedding for product review sentiment analysis*, IEEE Trans. Knowl. Data Eng., vol. 30, no. 1, pp. 185–197, 2017, doi: 10.1109/TKDE.2017.2727504.

[15] I. Ahmed, *Comparative study of Sentiment Analysis on Amazon Product Reviews using Recurrent Neural Network (RNN)*, Int. J., vol. 11, no. 3, 2022, doi: 10.3991/ijes.v11i3.23915.

[16] V. Garipelly, P. T. Adusumalli, and P. Singh, *Travel Recommendation System Using Content and Collaborative Filtering-A Hybrid Approach*, in 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), pp. 1–4, 2021, doi: 10.1109/ICCCNT51525.2021.9546022.

[17] R. L. Rosa, G. M. Schwartz, W. V. Ruggiero, and D. Z. Rodríguez, *A knowledge-based recommendation system that includes sentiment analysis and deep learning*, IEEE Trans. Ind. Informat., vol. 15, no. 4, pp. 2124–2135, 2018, doi: 10.1109/TII.2018.2868974.

[18] J. Zhong, Z. Yang, and J. Sun, *A Hybrid Approach with Joint Use of Tag and Rating for Vehicle and Cargo Matching*, in 2021 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), pp. 1397–1401, 2021, doi: 10.1109/IEEM50564.2021.9673072.

[19] Q. Wang, W. Zhang, J. Li, F. Mai, and Z. Ma, *Effect of online review sentiment on product sales: The moderating role of review credibility perception*, Comput. Hum. Behav., vol. 133, p. 107272, 2022, doi: 10.1016/j.chb.2022.107272.

[20] L. Berkani, M. Zaouidi, and R. Brahimi, *Sentiment deep learning algorithm for multi-criteria recommendation*, in 2022 First International Conference on Big Data, IoT, Web Intelligence and Applications (BIWA), pp. 77–82, 2022, doi: 10.1109/BIWA54433.2022.9755212.

[21] M. Almaghrabi and G. Chetty, *Multilingual sentiment recommendation system based on multilayer convolutional neural networks (mcnn) and collaborative filtering based multistage deep neural network models (cfmdnn)*, in 2020 IEEE/ACS 17th International Conference on Computer Systems and Applications (AICCSA), pp. 1–6, 2020, doi: 10.1109/AICCSA50499.2020.9316505.

[22] S.-W. Lee, G. Jiang, H.-Y. Kong, and C. Liu, *A difference of opinion in online review based on sentiment analysis and expert collaborative filtering*, in 2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA), pp. 595–600, 2021, doi: 10.1109/ICMLA52953.2021.00103.

[23] N. Vedavathi and A. K. M. Kumar, *E-learning course recommendation based on sentiment analysis using hybrid Elman similarity*, Knowl.-Based Syst., vol. 259, p. 110086, 2023, doi: 10.1016/j.knosys.2022.110086.

[24] U. Thakker, R. Patel, and M. Shah, *A comprehensive analysis on movie recommendation system employing collaborative filtering*, Multimedia Tools Appl., vol. 80, no. 19, pp. 28647–28672, 2021.

[25] G. K. Shinde, V. N. Lokhande, R. T. Kalyane, V. B. Gore, and U. M. Raut, *Sentiment analysis using hybrid approach*, Int. J. Res. Appl. Sci. Eng. Technol. (IJRASET), vol. 9, pp. 282–285, 2021.

[26] D. Sharma and A. Kumar, *Levels and classification techniques for sentiment analysis: A review*, in Proc. Int. Conf. Adv. Commun. Comput. Technol., Springer, 2019, pp. 333–345.

[27] S. Sharma, S. Sharma, and A. Athaiya, *Activation functions in neural networks*, Towards Data Sci., vol. 6, no. 12, pp. 310–316, 2017.

[28] S. Siswanto, Z. Mar'ah, A. S. D. Sabir, T. Hidayat, F. A. Adhel, and W. S. Amni, *The Sentiment Analysis Using Naïve Bayes with Lexicon-Based Feature on TikTok Application*, J. Varian, vol. 6, no. 1, pp. 89–96, 2022.

[29] A. Ziani, N. Azizi, D. Schwab, M. Aldwairi, N. Chekkai, D. Zenakhra, and S. Cheriguene, *Recommender system through sentiment analysis*, in Proc. 2nd Int. Conf. Autom. Control, Telecommun. Signals, 2017.

[30] J. Joseph, S. Vineetha, and N. V. Sobhana, *A survey on deep learning based sentiment analysis*, Mater. Today: Proc., vol. 58, pp. 456–460, 2022.

[31] R. Alroobaea, *Sentiment analysis on Amazon product reviews using the recurrent neural network (RNN)*, Int. J. Adv. Comput. Sci. Appl., vol. 13, no. 4, 2022.

[32] D. Anil, A. Vembar, S. Hiriyannaiah, G. M. Siddesh, and K. G. Srinivasa, *Performance analysis of deep learning architectures for recommendation systems*, in Proc. 2018 IEEE 25th Int. Conf. High Perform. Comput. Workshops (HiPCW), IEEE, 2018, pp. 129–136.

[33] G. Preethi, P. Venkata Krishna, M. S. Obaidat, V. Saritha, and S. Yenduri, *Application of deep learning to sentiment analysis for recommender system on cloud*, in Proc. 2017 Int. Conf. Comput. Inf. Telecommun. Syst. (CITS), IEEE, 2017, pp. 93–97.

[34] M. Ibrahim, I. S. Bajwa, N. Sarwar, F. Hajjej, and H. A. Sakr, *An intelligent hybrid neural collaborative filtering approach for true recommendations*, IEEE Access, 2023.

[35] R. C. Patil and N. S. Chandrashekar, *Sentimental Analysis on Amazon Reviews Using Machine Learning*, in Proc. Int. Conf. Ubiquitous Comput. Intell. Inf. Syst., Springer, 2022, pp. 467–477.

[36] B. Jain, M. Huber, and R. Elmasri, *Increasing Fairness in Predictions Using Bias Parity Score Based Loss Function Regularization*, arXiv preprint arXiv:2111.03638, 2021.

[37] K. L. Tan, C. P. Lee, K. M. Lim, and K. S. M. Anbananthen, *Sentiment Analysis With Ensemble Hybrid Deep Learning Model*, IEEE Access, vol. 10, no. September, pp. 103694–103704, 2022, doi: 10.1109/ACCESS.2022.3210182.

Edited by: Manish Gupta
Special issue on: Recent Advancements in Machine Intelligence and Smart Systems
Received: Jul 29, 2024
Accepted: Nov 9, 2024