VAWKUM Transactions on Computer Sciences
http://vfast.org/journals/index.php/VTCS@ 2023 ISSN(e): 2308-8168, ISSN(p): 2411-6335
Volume 11, Number 1, January-June 2023 pp: 184-203

The Impact of COVID-19 on E-Learning: Context-Based Sentiment Analysis Discourse Using Text Mining

Aqsa Rehman[1*], Naeem Aslam[1], Kamran Abid[1], Muhammad Fuzail[1], Asif-Ur-Rehman[2]

[1]Department of Computer Science, NFC Institute of Engineering and Technology, Multan, Punjab, Pakistan; [2]ERMAKSAN, Bursa, Turkey

Keywords: *COVID-19, Distance learning, Twitter, Sentiment analysis, Word embedding, RNN, LSTM.*

Journal Info:
Submitted:
April 15, 2023
Accepted:
May 20, 2023 Published:
May 20, 2023

Abstract Finding the most dominant and pertinent user opinions on a certain topic is crucial to the sentiment analysis success factor. During the pandemic lockdowns around the world, the suspension of academic institutions leads to an exceptional increase in distance education. Academic institutions closed their campuses immediately to mitigate the effects of COVID-19 and prevent its pervasive spread, and educational activities were shifted to online platforms. The effectiveness of online education is a significant topic of interest for both students and their parents, especially in terms of how students and teachers perceive it and how technologically viable it is in a range of social circumstances. Before such a wide adoption of e-learning is possible, these issues must be analyzed from multiple perspectives. The present research aims to evaluate the efficacy of e-learning by examining individuals' perceptions of it. Opinions can be found on websites such as Instagram, Facebook, Twitter, etc. As social media has recently emerged as a significant means of communication. This study addresses factors connected to a significant change in the educational system. 200,000 tweets were gathered from Twitter to evaluate the opinions of Twitter users who were taking part in online learning. This study adopts VADER to analyze the subjectivity and polarity score of tweets, a topic model was also created using the LDA algorithm to determine the themes that were talked about on Twitter the most. The models have been constructed and evaluated using Word2Vec to capture the semantic relationships between words and LSTM and RNN sequential model for sentiment analysis. This study measured the efficiency of a sentiment analysis model using the accuracy metric, the conducted experiments reveal that the proposed hybrid model achieves an overall accuracy of 96.3%. The results also indicate a significant negative impact of the Covid-19 pandemic on individuals' emotions, with 64.4% of the analyzed tweets displaying negative sentiments. These findings provide valuable insights into the relationship between global events and individual emotions on social media platforms.

***Correspondence Author Email Address:**
2k20mscs121@nfciet.edu.pk; aqsa.rehman1995@gmail.com

1 Introduction

Twitter has become one of the most prominent social media platforms, allowing users to convey their views, opinions, and feelings regarding an event or incident through tweets. Not only has it provided all the necessary information, but it has also made it possible for people to speak freely about their perspectives. It has provided a forum for the average citizen to express his opinions and viewpoints. Thus, social media could be regarded as one of the finest methods for monitoring public opinion. The 280-character limit on tweets encourages users to convey themselves concisely. This brevity makes it simpler to extract and analyze tweets' sentiment, as they tend to include clear expressions of sentiments or thoughts [1]. Because of the widespread of COVID-19, it has become vital for us to make adjustments to our social lives to better protect ourselves from the virus [2]. while continuing our regular routines. Education is one of the most severely impacted sectors by the pandemic. Owing to the bad circumstances that transpired during the pandemic, governments throughout the globe halted face-to-face schooling, and for a while, online schooling took its place. People's reactions to such an unexpected event on such a wide scale were mixed. Some felt the new method would benefit students, while others condemned it and pointed out the limitations of remote education [3]. People largely voiced their ideas on internet channels due to social isolation during the outbreak. Throughout the pandemic, Twitter remained quite popular. Since each tweet contains diverse types of material, including limits, Twitter sentiment analysis is a complex method due to slang, hyperlinks, emoticons, and hashtags [4].

Collecting textual information about the Coronavirus (COVID-19) in the educational sector presents a difficult task. The current study explores the subjective aspects of e-learning-related textual data. It is crucial in today's difficult atmosphere to use textual data to analyze comments and people's emotions on such platforms. Routine educational activities were suspended as a result of the COVID-19 pandemic's influence on the world's educational system. Additionally, it is challenging to convert such massive amounts of unstructured textual input into structured data and extract pertinent information from it [5]. It is essential in today's difficult circumstances to evaluate comments and people's emotions on such platforms based on textual data. The COVID-19 pandemic affected the world's educational system and forced the suspension of regular educational activities. A tough process that necessitates cleaning, text transformation, and dimension reduction to choose meaningful and pertinent content is the conversion of such massive amounts of unstructured textual data into structured data and the extraction of valuable information from such data. Evaluation quality is the primary criterion for determining the validity of an evaluation. As a result, analyzing this data is essential for establishing the impact of the educational environment and for giving educational institutions better feedback so that they can raise the quality of their services [6].

The proliferation of COVID-19 altered people's daily routines, shifting them from living, traveling, and working to social connections. The education system, like many other sectors, is experiencing major implications that impact students, instructors, and institutions worldwide. Amid the global lockdowns due to the COVID-19 pandemic, traditional brick-and-mortar educational institutions have been shuttered, leading to a shift towards digitalization and an unparalleled increase in remote learning. E-learning, often known as online learning, is the process of learning in a synchronous or asynchronous environment while using a mobile device that has internet access, such as a laptop, phone, or tablet, among other things [7]. It is anticipated that the transition from conventional educational methods to online platforms would be slow and may provide some difficulties. Despite the advantages, obstacles associated with the shifting process may prevent online education from reaching its full potential. Numerous research endeavors have been conducted to assess the efficacy and benefits of online learning in comparison to traditional pedagogical approaches. The benefits encompass an increased scope of educational outreach, enhanced accessibility, and autonomy from temporal and spatial constraints, alongside accelerated learning velocity. In contrast, e-learning faces several major challenges that pose substantial risks in comparison to traditional classroom teaching methodologies. the lack of communication technology infrastructure, the high cost of tools and gadgets, the lack of technical expertise among teachers and students, and the necessary cultural transformation are one of the obstacles to effective and successful online education. The lack of communication technology infrastructure, the high cost of tools and gadgets, the lack of technical expertise among teachers and students, and the necessary cultural transformation are one of the obstacles to effective and successful online education. The COVID-19 pandemic had a significant impact on education worldwide [8], resulting in the suspension of regular instructional practices. A significant number of students enrolled in various educational and training programs were unable to participate in in-person classes. A vast number of educational and instructional establishments worldwide have transitioned their pedagogical practices to various electronic learning platforms and communication modalities. Online education offers substantial benefits in both pedagogy and learning, while also serving as a critical foundation for the worldwide education system amidst the ongoing COVID-19 pandemic. When transitioning from traditional in-person teaching to online learning [9].

It must be demonstrated that e-learning is at the very least an excellent alternative to traditional education, if not superior. Other research has concentrated on the COVID-19 issue based on comparable efforts [10]. Using Twitter data to raise community awareness of the COVID-19 epidemic [11]. Conducted a study that looked at the perspectives of the government, organizations, and individuals on Twitter during the epidemic. During the shutdown, the authors evaluated Twitter as a cross-language platform for European nations. [12] Given such complexity, Natural Language Processing (NLP) approaches are quite useful for developing various solutions [13]. One of the most prominent NLP research fields nowadays is sentiment analysis [14], which is used in a variety of fields, including recommender systems, data-driven systems, healthcare research, and others. Due to the popularity of social media, massive volumes of data are created on the Internet every day. People submit their opinions on a variety of social networking sites, including YouTube, Twitter, and Facebook. By examining these postings, one can draw interesting and beneficial conclusions as well as create goods. Because of the influence, these reviews have on both product deals and people's opinions [15] which have an impact on third-party vendors on shopping sites [16]. Unfortunately, creating such forecasts is a difficult task since channel oversaturation with the material makes it difficult to convert and connect. It becomes exceedingly challenging to conduct a full sentiment analysis with this type of unfiltered content, which comprises typos, links, emoticons, and grammatical errors.

The COVID-19 epidemic has had a tremendous influence on the global education industry, prompting schools and universities to use remote teaching and learning approaches. Although online education has been accessible for some time, the quick transition to online learning during the epidemic has brought new obstacles for both instructors and students. The usefulness of online teaching, particularly during the epidemic, is still being debated by scholars and educators [17].

The key objectives of our study are to use Deep Learning models to assess the sentiments of tweets [18], to inquire about the influence of the epidemic on the school system, and to examine the social impacts of remote learning on people. By evaluating social media material connected to remote learning, the goal is to assist management companies in the development of educational systems and to allow people to correctly map emotions throughout the pandemic. To do this, sentiment analysis was performed using Twitter data [3]. The emotion of each statement was intent on, and polarity values were computed. Preprocessing, tweet categorization based on polarity values, content-based classification of data, training models with RNN, LSTM, Hybrid (LSTM-RNN) [19], and model optimization via hyperparameter modifications were used to collect a collection of English tweets on remote learning.

This study contributes to the existing body of knowledge by presenting a thorough analysis of the opinions and emotions of individuals towards e-learning during the Covid-19 pandemic. By leveraging the power of natural language processing techniques, this study offers a unique perspective on the efficacy and reception of online education in the context of a global crisis. The study has several notable contributions to the literature:

- The study utilized a huge dataset of 200,000 tweets obtained from Twitter, which represented a comprehensive discussion on e-learning during the outbreak. This study intended to give complete coverage of the whole pandemic duration, in contrast to existing datasets that are limited in size and cover only short periods. As a result, collecting Twitter data from scratch resulted in the creation of a unique dataset that provides a solid foundation for analyzing public opinions around this critical topic.
- Instead of depending on positive and negative word counting, sequential models (RNN, LSTM, and

a Hybrid LSTM-RNN) were used to capture context-based sentiment analysis in this study. The comparison of these models revealed that the hybrid LSTM-RNN strategy produced greater accuracy, allowing for a more sophisticated comprehension of emotions and viewpoints conveyed in social media content.
- The research utilized a combination of VADER for sentiment classification, LDA for topic modelling, and Word2Vec to extract semantic word associations. The hybrid methodology was become even more effective by the use of context-based sentiment analysis adopting sequential models.
- The suggested hybrid LSTM-RNN model attained an overall accuracy of 96.3% in sentiment analysis, illustrating the model's competence in capturing and understanding emotions and opinions stated in social media content.
- The analysis determined that 64.4% of the studied tweets contained negative emotions about e-learning during the Covid-19 outbreak. This study is important because it sheds light on the difficulties and potential pitfalls connected with the mass adoption of online education amid a worldwide crisis.
- The research findings can be helpful for academic institutions, policymakers, and educators in understanding the perceptions and insecurities of students and teachers, allowing them to develop strategies to deal with these problems and improve the overall effectiveness of e-learning.

Overall, this research paper makes a substantial contribution to the existing literature on e-learning sentiment analysis during the Covid-19 pandemic by shedding light on the relationship between global events and individual emotions on social media platforms. The custom dataset, context-based sentiment analysis approach, and innovative methodologies used in this study, including the comparison and selection of the hybrid LSTM-RNN model, establish a benchmark for future research in the domains of sentiment analysis and educational technology.

To begin, although multiple studies on sentiment analysis of tweets at COVID-19 [20], most of them concentrated on a specific region or nation. As a consequence, there is a need for research that covers a greater geographic region to provide a more comprehensive understanding of people's sentiments throughout the pandemic. Second, current sentiment analysis research has mostly relied on basic approaches of machine learning such as SVM and Naive Bayes [21]. Yet, considering recent developments in deep learning, more research into the usefulness of using deep learning models for sentiment analysis of Twitter data during the pandemic is required. Lastly, although most research has focused on smaller datasets for sentiment analysis, there is an urgent need for larger and more comprehensive datasets capable of capturing the full range of emotions expressed on Twitter. This research gap makes building trustworthy and robust sentiment analysis models challenging, limiting their potential effect in gauging public opinion and directing decision-making during the pandemic. Closing these study gaps will result in a better understanding of public sentiment during the COVID-19 outbreak [22], as well as insights for policymakers in developing effective communication approaches to address public concerns and anxiety.

2 Related Work

Depression in Bangla social media was analyzed using LSTM Deep Recurrent Network. This study used a small dataset of Bangla social media to show how hyper-parameter tuning might help depression research. With stratified datasets with repeated sampling, LSTM 5 layer of 128 units size with 25 batch sizes, 0.0001 learning rate over 20 epochs, the accuracy of depression detection is high. This discovery will help researchers and other psychologists identify sadness in virtual social interactions and reduce depression-related behaviors [23].

Three text features TF-IDF, Bag of Words, and Word2Vec used to vectorize the Turkish dataset for model application. The author compares machine learning and deep learning methods, results to see if social media can help disaster management. Deep learning was more accurate findings which showed that social media may help understand disasters and make decisions [24]. Recently, CNN ensembles are compared to LSTM RNNs. In addition, compare the Word2Vec and GloVe models to analyze such algorithms using SemEval data. Using their best score values, various tests and combinations are utilized to compare model

Table 1. Related Work

Ref	Year	Model	Dataset	Accuracy	Aim	Limitation
[23]	2019	LSTM	Bangla Tweets	86.3%	Depression analysis	Bangla language is used
[24]	2019	word2vec and MLP	3000 rows of Turkish tweets	81.86%	Opinion mining	Accuracy is not significant, small dataset used
[25]	2019	CNN and LSTM	Tweets from semantic evaluation	59.0%	Sentiment analysis	Albanian language tweets are used, Accuracy is not significant
[26]	2019	Naive Bayes and SVM	13,000 tweets of product reviews	93.54%	product reviews Sentiment analysis	The study is not about the sentiment analysis of online education
[27]	2020	LSTM and CNN	63,000 sentiment corpus for Arabic text	90.75%	Arabic sentiment analysis	Specific to the Arabic language, only 2.5% of internet-using language
[28]	2020	Multi-layer LSTM	27,357 tweets	82.4%	COVID-19 Sentiment analysis	Not significant Accuracy and tweets of COVID-19 general
[29]	2021	Naive-based classifier (model)	90,000 tweets were generated to discuss the findings	83.5%	Sentiment analysis impact of the pandemic on education	Topic modeling not performed and has no significant accuracy
[30]	2021	Dictionary-based approach	From Google, 154 articles were collected	90.0%	Opinion Mining on online education during COVID-19	Study used a machine learning approach

performance. The performance, advantages, and drawbacks of the aforementioned sentiment analysis methods are examined in this study utilizing a unified testing framework, identical datasets and computational environments [25].

Categorizing product reviews by using machine learning models give good results. For camera reviews, 98.17% accuracy of NB and 93.54% accuracy occurred SVM (Support Vector Machine) [26].Arabic sentiment analysis uses a deep learning model that expertly mixes a CNN one-layer architecture with a two-layer LSTM. This model well performed in recall, precision, F1-Score, and accuracy on a multi-domain corpus, scoring 89.10%, 92.14%, 92.44%, and 90.75%. Following a rigorous evaluation of emotion categorization embedding approaches in Arabic, the FastText model was shown to be better for learning semantic and syntactic information.

The suggested KNN and NB classifiers model is evaluated. The findings show that the best classifier is SVM, improving accuracy by 93.92%. Word embedding with WordNet's lexical database may increase quality [27].Deep learning models are used to analyze COVID-19 tweets to determine how from various cultures people reacted to the virus and the measures implemented by various nations [31]. The sentiment140 dataset has been trained to predict emotions and sentiment classification from retrieved tweets, deep long short-term memory models were applied with cutting-edge accuracy. Emoticons were used to assess Twitter tweet-based supervised deep learning models categorized sentiment polarity and emotions using a multi-layer LSTM model [28].

Opinion mining, or sentiment analysis, extracts people's ideas, feelings, and opinions about a subject from massive amounts of unstructured data. Sentiment analysis in many languages has been studied recently. NLTK and Naive Bayes Classifier analyzed 90,000 informative COVID-19 tweets for sentiment [29]. In another study to evaluate online learning sentiments, pandemic news and blogs were monitored. Web-scraping and to assess online learning contents. Dictionary-based with lexicon-based text mining detects 154 online learning articles on Google and other sites like reviews and blogs. Results showed positive but cautious opinions regarding online education in low-polarity public digital media. Even while blogs were more subjective than news pieces, the low overall subjectivity ratings suggest an evidence-based method for promoting public policy conversation through digital media [30].

3 Methodology

In this paper, we present a method for sentiment analysis. It is composed of the five primary components:

1. Snscrape data collection: We use the Snscrape library to collect Twitter data from scratch. With the help of this tool, tweets can be efficiently and completely extracted, creating a rich dataset for analysis.
2. Pre-processing: This part consists of core linguistic operations, creating a collection of custom features and a representation of word embedding. Through preprocessing, our dataset is well-structured and prepared for subsequent analysis.
3. Supervised Labeling with VADER: To categorize the collected tweets, we utilize the pre-trained model VADER for sentiment analysis, which offers accurate results for identifying sentiment in social media material.
4. LDA Topic Modeling: Our strategy uses LDA for topic modeling to discover underlying data themes.
5. LSTM-RNN and Word2Vec: For sentiment classification, our method takes advantage of the LSTM-RNN algorithm's capacity to represent long-range dependencies. Using Word2Vec embeddings as input, we successfully extract semantic information, hence enhancing the precision and efficacy of our sentiment analysis algorithm.

The current study used pre-trained models to label the data, topic modeling to specify the most frequently used words to determine the themes that were talked about on Twitter the most and a deep learning-based hybrid model is used for sentiment analysis to investigate pandemic effects on the educational system. Figure.1 illustrates the proposed system architecture utilized. The workflow commences from dataset extraction from Twitter into a CSV file which is conducted based on the selected keywords string initial stage of the investigation involves ascertaining. The next phase is through several pre-processing steps such as removing duplicates, removing stop words, tokenization, and lemmatization dataset cleaning performed.

Further, The VADER tool [32] is utilized to annotate the data by assigning sentiment labels that correspond to the text's emotional tone negative and positive. The dataset that has been labeled is partitioned into distinct sets for training and testing deep learning models. These sets are utilized to train and evaluate the performance of the models. In this regard, LDA topic modeling [33] is performed to identify the topics that have been discussed on the social media platform. After that classification process, its outcomes, and the associated learning mechanisms RNN, LSTM, and LSTM-RNN architectures are evaluated. In the subsequent section, a concise description of each phase is provided.

3.1 Tweet Dataset Collection

In the study, the dataset was collected utilizing the Snscrape library [34], a Python program that is open-source and facilitates connection to the Twitter database and scrap tweets from Twitter. During the construction of the corpus, the following English keywords were utilized: "distance education"," covid-19", "e-learning", "online education", and "distance learning" We have gathered a total of 200,000 tweets about online education during Covid-19 and the tweets timeline is between '2019-01-01' and '2022-05-05'. Table.2 shows a collected tweets sample.

3.2 Data pre-processing

Data processing is required for data analysis which removes unnecessary acquisition of additional knowledge to enhance the efficacy of classification models and get accuracy. Tweets frequently contain extraneous content, surplus verbiage, and disorganized syntax and expressions, attributable to the inclusion

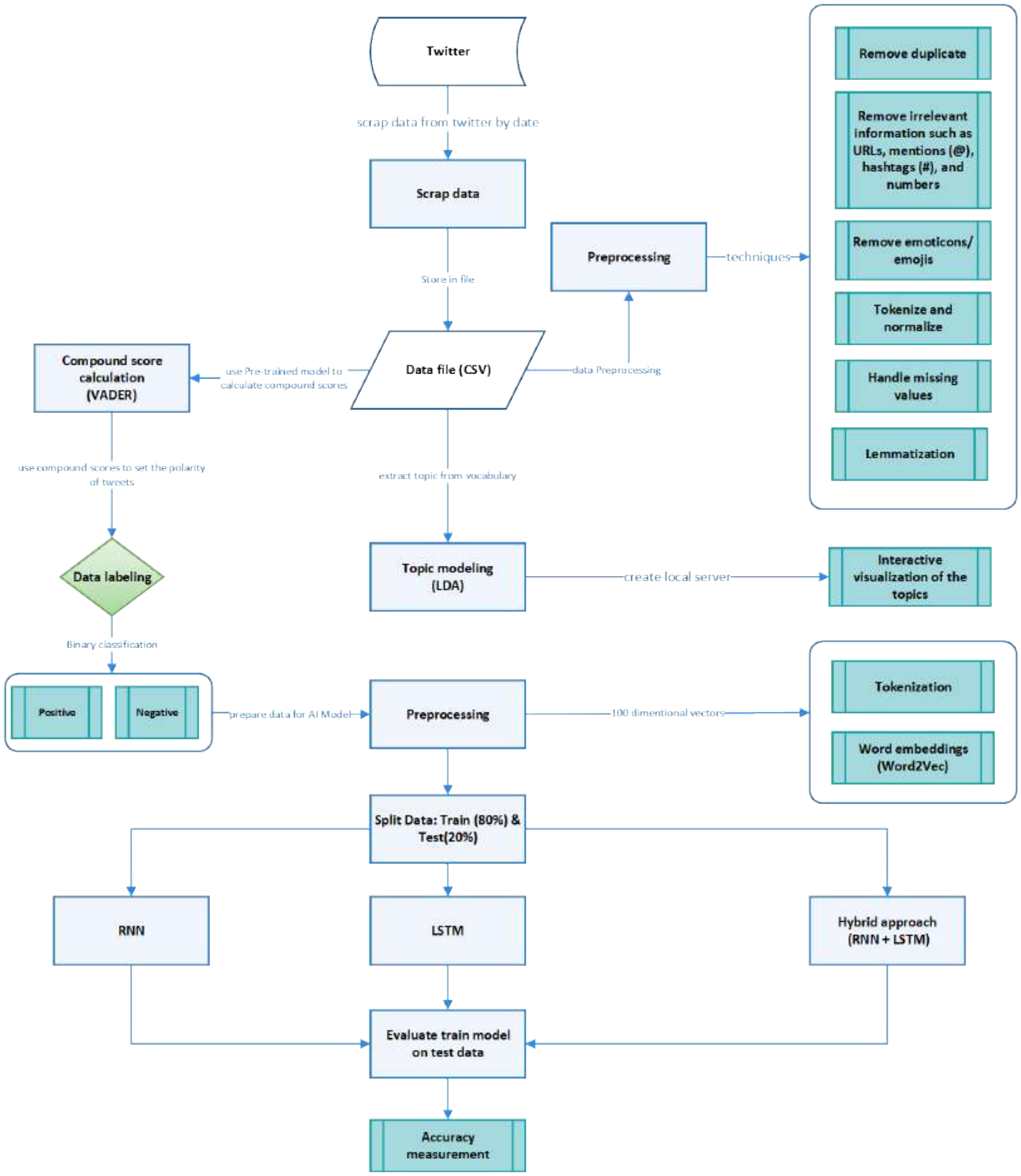

Figure 1. Proposed system architecture

Table 2. Tweet sample of online education

Raw tweets
@JamesDeanBoyd @mrcleancousin @mightyzo What? Quality of education, of health, of lives because of the idiocy and complacency in the pandemic responseâ€¦ what HASNâ€™T Kenneyâ€™s government negatively affected?
@realreporter A friend is just getting over Covid. She did some reading and figures she is good for 6 months and can travel all summer without taking precautions. Covid education for the general public is sadly lacking. Weâ€™re screwed!

of non-standard vocabulary, malformed lexemes, acronyms, and erroneous syntax. Superfluous information refers to data that either has no contribution or very little contribution to predicting the target class. Such data can introduce unnecessary computational complexity, which may degrade the performance of classification models if no preprocessing is carried out or if it is preprocessed improperly. This underscores the need for effective preprocessing to mitigate the impact of superfluous information and reduce the computational power required [35].

Thus, data cleaning or preprocessing [36] is required before analysis. For these reasons, it is necessary to make the text ready for analysis. Because whereas almost every tweet contains superfluous information such as emoticons, symbols, hashtags, and punctuation, it is necessary to filter them out. Several preprocessing techniques were applied to get a clean dataset. The first step is to clean up the noisy data by deleting any expressions that are unnecessary or useless for sentiment analysis. Many symbols have been eliminated using this technique, including emoticons, punctuation marks, usernames in Twitter texts that begin with the @sign, and links in text content that begin with "HTTP." After these steps, the text's entire content was converted to lowercase letters, and the beginning and end spaces were removed. Using Python's NLP package [37], tweets data were preprocessed for this investigation. Links, HTML (Hypertext Markup Language) components, and punctuation are removed once the text has been converted to lowercase. After applying lemmatization techniques to tidy up the text, stop words are removed at the end. After that, all tweets were removed which contain less than three words and sample tweets are shown in Table.3.

Table 3. Before and after pre-processing tweets

Raw tweets	After preprocessing
@JamesDeanBoyd @mrcleancousin @mightyzo What? Quality of education, of health, of lives because of the idiocy and complacency in the pandemic responseâ€¦ what HASNâ€™T Kenneyâ€™s government negatively affected?	quality education health life idiocy complacency pandemic response kenney government negatively affected
Post-Pandemic we have learned how online education and communication is an effective modality for learning. Keep it Innovative, Creative, and Accessible to all! @mdcpsdigiteam @SuptDotres #innovativeED #remotelearning #onlineeducation #onlinelearning	post-pandemic learned online education communication effective modality learning keep innovative creative accessible

3.3 Tokenization

Tokenization is the process of breaking up a large block of text into smaller tokens [38]. Tokens are the fundamental linguistic building elements that consist of words, subwords, and characters. As tokens are the fundamental units of Natural Language, token-based processing is the most common approach for processing unprocessed text. Three of the most prevalent deep learning architectures for NLP, RNN, GRU, and LSTM, also execute token-level analysis of the raw text. Using a tokenizer, unstructured data and text written in natural language are broken down into discrete informational components.

The token occurrences of a document can be used to generate a vector representing the document. Instantaneously, an unstructured string (text document) is converted into a numerical data structure suitable for machine learning. The most fundamental step in processing textual data, tokenization involves identifying the words that comprise a series of characters. This is significant because the text's meaning can be easily determined by analyzing its terms (Webster & Kit, n.d.). Before removing stop words from the study, tokenization was performed [39].

3.4 Word embedding

The utilization of a vector space methodology for text representation has been primarily employed for document representation and has more recently been expanded to encompass word or term representation. This approach serves as the fundamental basis for achieving effective text representation. Neural language models were developed utilizing word embeddings and word vectors sourced from the Word2Vec toolkit [40]. Word embeddings are representations in a linguistic vector space that aim to preserve both semantic and grammatical relationships among words. Word embeddings are commonly referred to as word representations, as noted in reference [41]. The Word2vec tool offers two distinct architectures to represent words in a multi-dimensional space. These architectures are known as the continuous bag-of-words (CBOW) model and the skip-gram (SG) model. The Continuous Bag-of-Words (CBOW) model aims to learn embeddings by predicting the target word within a given text, in relation to the context word. The Skip Gram model employs a distinct approach from CBOW to predict the contextual words based on a given target word. The Word2Vec model employs the Continuous Bag-of-Words (CBOW) architecture to derive concept categorization through the vectorization of words. The entirety of the sanitized training data was utilized in generating word embeddings, with a designated vector size of 100. The minimum threshold for inclusion in the Word2Vec model is set at 5, thereby disregarding words that have a frequency of fewer than 5 occurrences. A window size of 5 indicates that the maximum allowable distance between the target word and any other contextual word within a sentence is 5. Four workers were chosen to facilitate faster training of the trained model through threading.

3.5 Sentence level sentiment polarity calculation

In the subsequent phase, a lexicon-based sentiment indicator is used to compute the polarity score of the tweets. To accomplish this, we fine-tuned tweet content on a newly generated dataset and used the pre-trained VADER model to determine whether each tweet is a positive or negative sample shown in Table.4. This allowed for the identification of tweet emotion classifications shown in Figure.2.

Table 4. Before and after pre-processing tweets

Tweet	**Sentiment**
quality education health life idiocy complacency pandemic response kenney government negatively affected	negative
post-pandemic learned online education communication effective modality learning keep innovative creative accessible	positive

1. Initially, each term present in the documents comprising the dataset was treated as a separate word.
2. Then, these words were matched against two distinct vocabularies of positive and negative English terms. When the number of positive words exceeded the number of negative words, a tweet was deemed positive. In contrast, it was labeled negative if the number of negative terms was greater. This method insured that each tweet's sentiment was accurately categorized.

In accordance with the purpose of our study, we divided the tweets into positive and negative categories and disregarded neutral comments, as previously stated. As shown in Table.5, this phase resulted in the classification of 200,000 tweets, of which 69,477 were classified as positive and 126,119 were classified as negative.

Each word in the input text is assigned a sentiment score based on its valence and intensity, and the text's overall sentiment is determined by aggregating these scores using a set of heuristics. The polarity score is calculated by dividing the sum of individual words' valence scores by a normalizing factor. A word's valence score indicates its emotional intensity, ranging from -4 (most negative) to +4 (most positive), with 0 representing neutrality. Each sentiment term is assigned a valence value in the sentiment lexicon used in our study. In the formula, the normalizing factor is a scaling factor that ensures the polarity score is

Table 5. Dataset sentiment class

Emotion	Count	Percentage (%)
Positive	69,477	35.520
Negative	126,119	64.479
Total	200,000	100

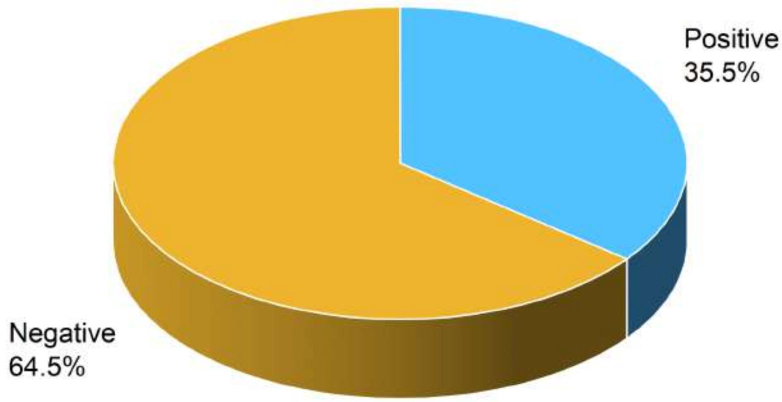

Figure 2. Dataset distribution

between -1 and +1, with -1 representing the most negative sentiment and +1 representing the most positive sentiment. The normalizing factor is calculated by taking the square root of the sum of the individual terms' squared valence scores.

$$\textit{polarity score} = \frac{\textit{sum of valence scores of individual words}}{\textit{normalizing factor}} \tag{1}$$

$$\textit{normalizing factor} = \sqrt{\left(\sum \textit{squared valence scores of individual words}\right)} \tag{2}$$

The dataset was then divided into training and testing sets, with 80% of the data used for training and 20% for testing. After the dataset was partitioned, text vectorization was performed. This procedure attempts to convert human-written data into a language that a deep-learning model can comprehend.

3.6 Sentiment analysis using LSTM-RNN and Word2Vec embeddings

The proposed method for developing the model involves converting textual data into numerical data. The input data is preprocessed using several techniques, after which it is passed through multiple layers, including LSTM, Time Distributed, and [26] RNN layers, with a classification activation function operating in the final layer. The architecture of our Hybrid model is shown in Figure.3. The model begins with an Embedding layer that utilizes Word2Vec embeddings, converting the input text into a dense representation suitable for the LSTM model [42]. The layer is configured with an input vocabulary size, an output dimension of 100, a pre-trained embedding matrix, an input length of 50, and is set to be non-trainable. The first LSTM layer is added, consisting of 128 units, with a return sequences option set to True, and dropout and recurrent dropout rates are both set at 20%. Two additional LSTM layers follow, the first containing 64 units and the second containing 32 units, both with return sequences set to True, and dropout and recurrent dropout rates of 20%. A Time Distributed layer is added, which applies a Dense layer with 128 units and a ReLU activation function [43] to each time step of the sequence, allowing the model to learn temporal relationships across different time steps. Next, two Simple RNN layers are incorporated, the first with 64 units and a ReLU activation function, and the second with 32 units and a ReLU activation function. The first Simple RNN layer has the return sequences option set to True, allowing the model to pass the sequence information to the subsequent layer. Finally, a Dense layer with a single neuron and a sigmoid activation function is added, which outputs the sentiment polarity prediction. The model is trained using the Adam optimizer with a learning rate of 0.001 over thirty epochs. This architecture leverages the power of Word2Vec embeddings, LSTM, Time Distributed, and RNN layers to effectively capture the temporal and contextual relationships within the text data, leading to more accurate sentiment analysis and predictions.

3.7 Latent Dirichlet allocation-based topic modeling

The text structure consists of a combination of topics that can be deduced using machine learning techniques to predict concealed variables. Multiple methods can therefore be used to extract information from a sequence of text, denoted as where xi corresponds to the ith tweet. Latent Dirichlet allocation (LDA) is one of the most widely employed methods for generating probabilistic models of sequences [44]. Probability-based modeling is a topic modeling technique. Let be the set of k topic models. A log-likelihood objective function is defined by [36] as follows:

$$log\, P(X\,|\,Z) = \sum_{x \in X} \log P(X\,|\,Z_{\theta}(X)) \tag{3}$$

The subject identity of the tweet is denoted by equation (3). A topic is documented as a procedure for compiling probabilities ranging from 0 to 1 across the entire corpus. In the LDA generator procedure [45], the document is regarded as a collection of diverse topics and words [18]. We utilized perplexity and topic coherence to assess the potential of the positive and negative topic models. Complexity and topic coherence are appropriate criteria for evaluating the quality of a topic model. The model is deemed advantageous when complexity is reduced. Topic coherence is used to determine the semantic similarity of a collection of content. Greater topic coherence correlates with the greater average pairwise similarity between terms.

4 Experimental results

Various preprocessing techniques were applied to the dataset, resulting in 200,000 tweets with positive or negative sentiments. To provide a comprehensive analysis, we performed topic modeling followed by classification based on sentiment analysis using deep learning algorithms. Figure.5 demonstrates that the topic two and topic three models have the highest coherence scores, which decrease considerably as the number of topics increases. As the number of topics increases, the coherence graph displays peaks and troughs. Based on the graphical representation, three appropriate topics were determined. Figure.4-a illustrates the topics extracted from all tweets.

The size of each cluster correlates to the percentage of tweets in the corpus that are relevant to that issue, and each cluster represents a unique topic. In Figure.4-b, The percentage is highest for topic 1 and the distance between the bubbles represents the similarity between the topics, with bubbles far apart representing dissimilar topics. Figure.4 demonstrates the word collections associated with each topic. Using the corpus, the word lists are compiled, and their frequency is displayed in cyan. The red bars indicate the frequency with which each term is associated with the given topic. The terms with the longest red bars in the displayed data are the most common. The term "education" is most frequently associated with topic 1 in Figure.4, which corresponds to the English word "education." The term "learning" (which appears in English) is emphasized in topic 3, while the term "learning" (which appears in English) is highlighted in topic 2.

Figure.6 shows high coherence scores of negative tweets that belongs to number of topics three, five and nine. However, the lowest coherence scores observed with the number of topics one and four, so overall result shows highest coherence scores. Figure.7 depicts that the high coherence score was obtained for

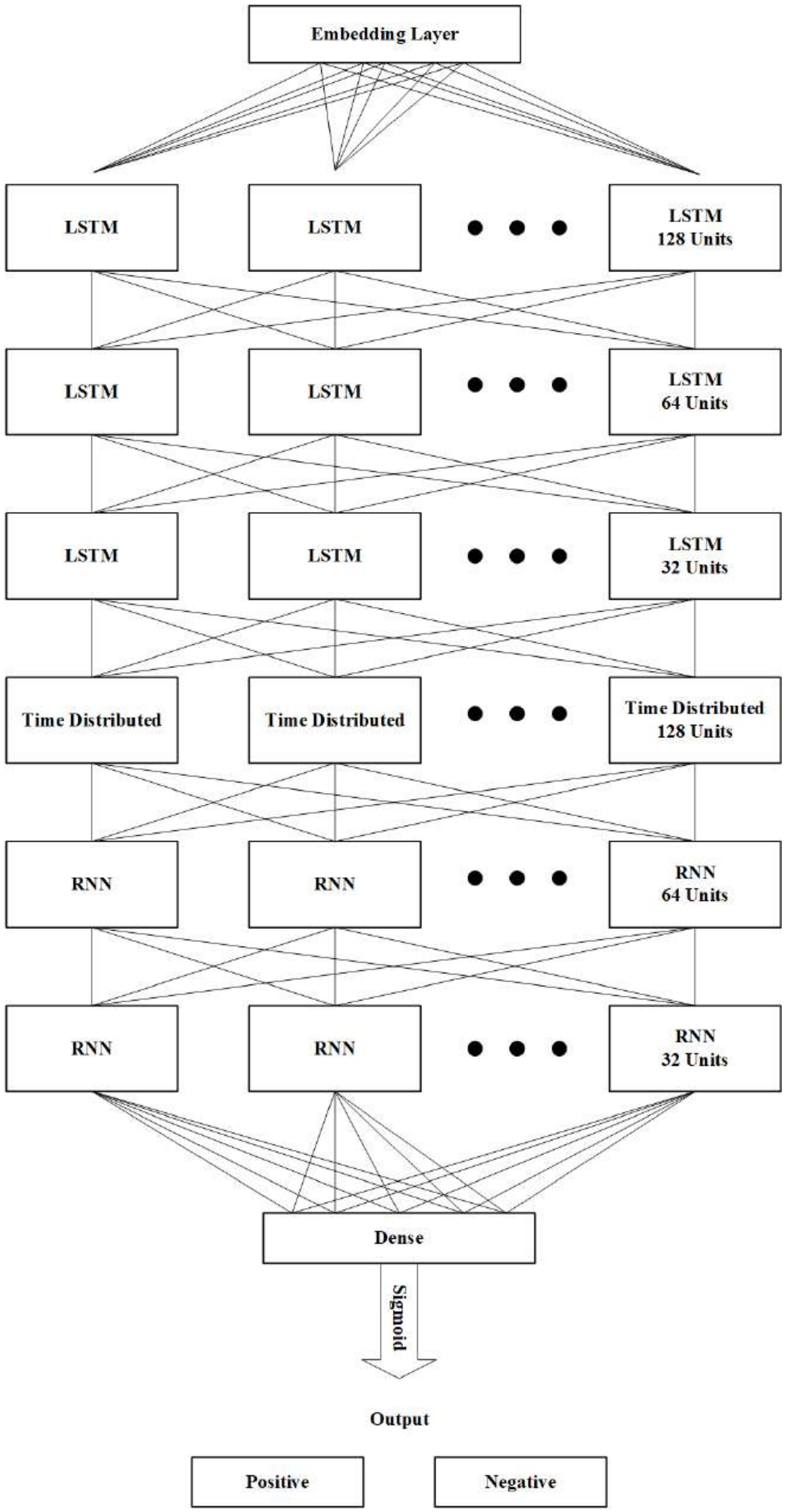

Figure 3. Illustration of proposed LSTM+RNN (Hybrid) model for Twitter sentiment

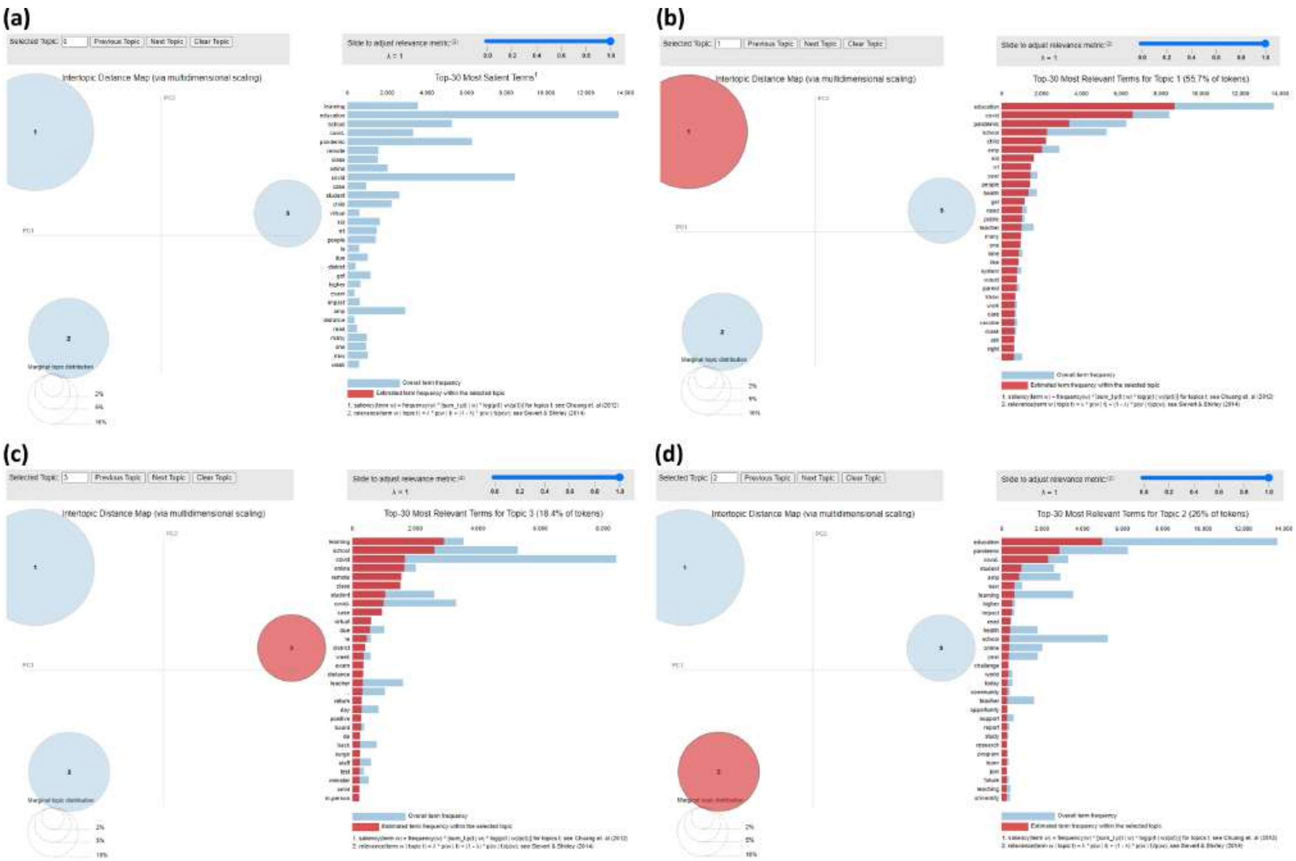

Figure 4. a) The sentiment analysis inter-topic distance map of tweets for distance learning Top-30 most relevant terms for b) subject 1 c) subject 2 d) subject 3

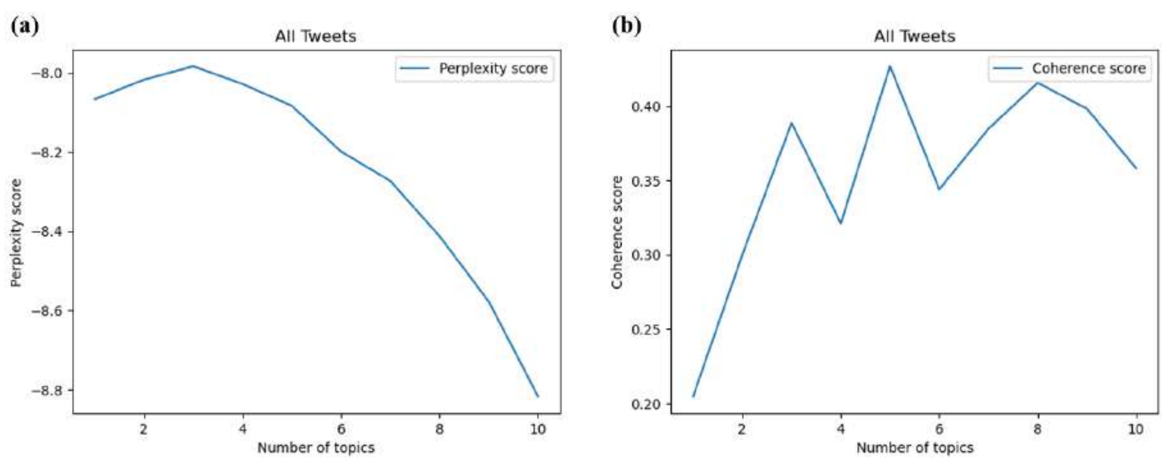

Figure 5. Distinct topics while evaluating all tweets: a) Coherence Score b) Perplexity Score

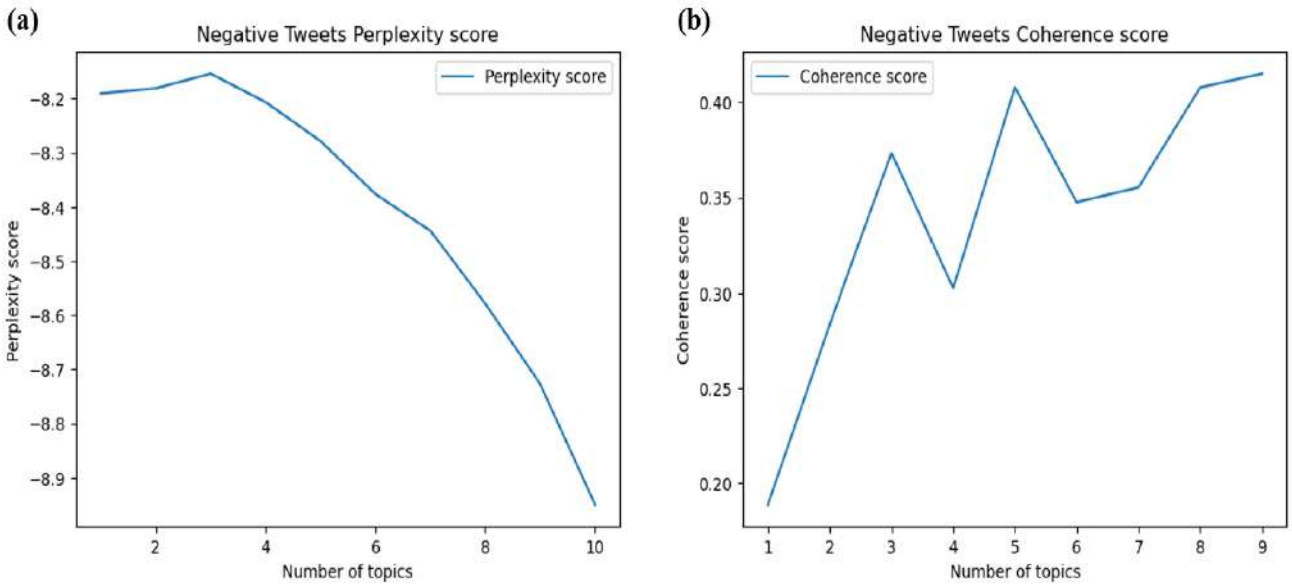

Figure 6. Distinct topics while evaluating negative tweets: a) Coherence Score b) Perplexity Score

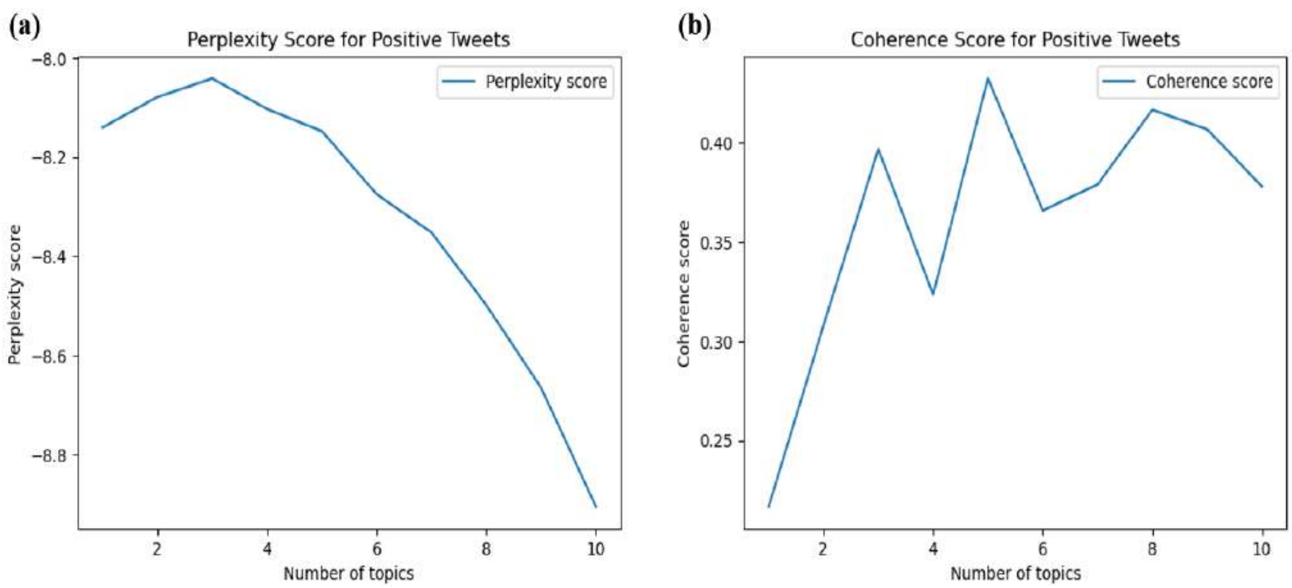

Figure 7. Distinct topics while evaluating positive tweets: a) Coherence Score b) Perplexity Score

the number of topics three, five and eight when the positive tweets were evaluated and number of topics one, four and ten belongs to the low coherence score. When the number of the topics were evaluated low coherence scores were not considered but the high coherence score was examined because the high coherence refers to the high level of consistency, similarity and semantically connected text.

Furthermore, in 8 a word cloud of the pre-processed data is produced for all tweets, negative tweets, and positive tweets separately. Word clouds exhibit a positive relationship between the size of a word and its significance, as the former is determined by the frequency of occurrence.

Figure 8. Most frequently used words in all tweets of 2020, b negative tweets, c positive tweets

After identifying the underlying semantic structure through topic modeling, the research project moved forward with the development of Natural Language Processing (NLP) classification models for sentiment analysis. Long-Short Term Memory (LSTM) [46] and Recurrent Neural Networks (RNN) were used to train two deep learning classifiers on the Twitter dataset pertaining to distance learning (RNN). Combining long-short-term memory (LSTM) [47] and recurrent neural networks (RNN) will increase the complexity of the models and increase training and validation accuracy. However, as the complexity of the models increased, training accuracy continued to rise while validation accuracy remained constant. To address this issue, regularization, and dropout layers were implemented to ensure that the training and validation sets had comparable accuracy and loss scores. This modification enabled the models to learn from the dataset and generalize it, which was essential for the classification task involving sentiment analysis. The final models were evaluated using accuracy to determine their efficacy at classifying tweets as positive or negative sentiment.

In addition to the aforementioned modifications, we also experimented with different epoch sizes, optimizers, and learning rates to further optimize the performance of our sentiment classification models [48]. After evaluating various optimizers, we found that Adam worked best for the Hybrid and LSTM model, while Stochastic gradient descent was the most accurate for the RNN model [49]. We also tried different learning rates and achieved the highest accuracies with 0.001 for the Hybrid model Figure.11 and 0.001 for the RNN Figure.9 and LSTM model Figure.10.

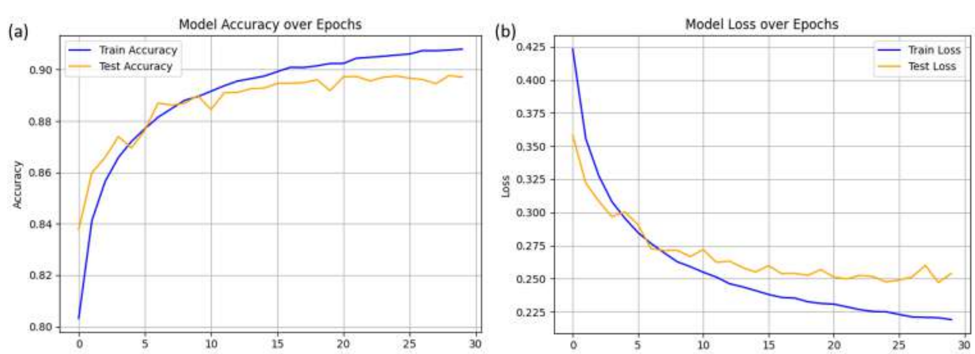

Figure 9. The RNN model's accuracy and loss trends

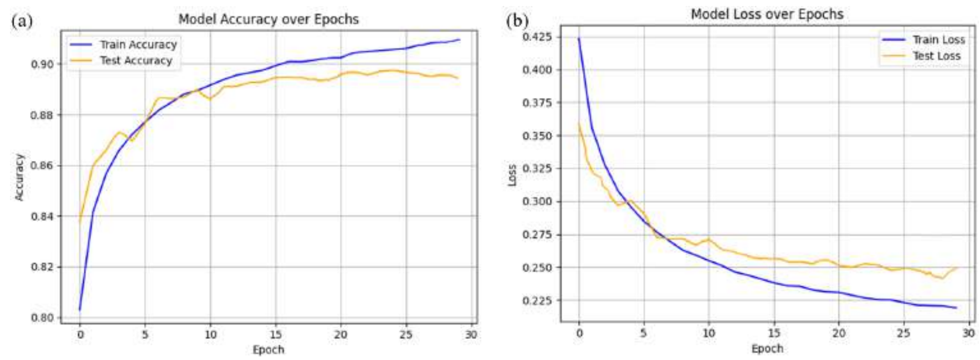

Figure 10. Loss trends and accuracy of the Simple LSTM model

Figure.11 demonstrates how the learning trajectory of the composite model, which combines Long Short-Term Memory (LSTM) and Recurrent Neural Networks (RNN), became unstable during the second epoch. To address this issue, we decided to modify the model by replacing its Embedding layer with a Word2Vec word embedding that had been trained. Figure.11 demonstrates that this modification significantly enhanced the learning process of the model.

In a comparison between the LSTM model without word embedding and the RNN model with Word2Vec embedding, the former achieved a higher accuracy score of 83% than the latter, which reached 76%. Notably, the RNN model demonstrated consistent learning throughout the training process, but it still performed less well than the LSTM model with Word2Vec. Further investigation revealed that the pre-trained Word2Vec word embedding played a crucial role in enhancing the LSTM model's learning capacity. Using Word2Vec enabled the model to incorporate semantic relationships between words, which contributed to its improved performance [50].

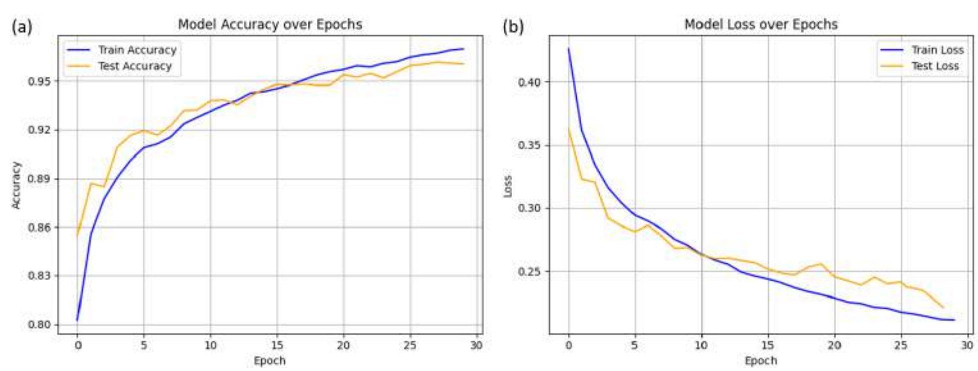

Figure 11. Accuracy and loss trends of the Word2Vec LSTM + RNN hybrid model

Our findings indicate that incorporating pre-trained word embeddings, such as Word2Vec, can substantially enhance the learning process of deep learning models, especially LSTM and RNN hybrid models. Table.6 provides an exhaustive summary of these outcomes. In addition, future research could investigate the effect of other pre-trained embeddings and architectures on the efficacy of such hybrid models.

5 Conclusion

In response to the COVID-19 pandemic, measures were implemented by governing bodies to safeguard public health while fulfilling societal obligations. As a result, conventional brick-and-mortar educational establishments were compelled to suspend operations, and remote learning emerged as a prevalent alternative. Although certain countries were more agile in their endeavors to improve the standard of education,

Table 6. Results of sequential models-based sentiment analysis

Classifier	Val. Acc.	Val. Loss	Train Acc.	Train loss (%)
RNN	0.774	0.473	0.764	0.498
LSTM	0.897	0.254	0.908	0.219
LSTM+RNN (Hybrid)	0.963	0.223	0.981	0.212

several nations encountered difficulties in effecting this transformation owing to various impediments. The experiment conducted was of a significant magnitude and unfortunately resulted in a reduction of societal resilience and belief. As a consequence, the population experienced psychological difficulties and a tendency towards introversion. The significance of understanding the substance of this page stems from the convenience that social media provides for individuals to express themselves.

The objective of this investigation is to ascertain the predominant themes and attitudes surrounding COVID-19 remote learning as expressed on social media. The efficacy of distance education for students remains a topic of debate. Therefore, we aim to examine the changes in awareness and attitudes towards this subject to gain a deeper understanding of the overall trend. The principal objective of the initial classification of contributions was to ascertain the overall attitude or emotional tone of individuals. In order to accomplish this objective, social media posts spanning the years 2019 to 2021 were gathered. Various text preprocessing methods, such as tokenization, normalization, stop-word elimination, lemmatization, and others, were employed to cleanse the content. Subsequently, the utilization of word embedding methodologies was employed to create neural language models through the implementation of Word2Vec's word vectors. Subsequently, a number of sentiment models were created utilizing deep learning and topic modeling techniques with the aid of the LDA algorithm. The augmentation of the Hybrid model's efficacy has been noted through the utilization of word embedding. The Word2Vec-based Hybrid model demonstrated the highest level of performance, achieving an optimal outcome with a validation accuracy of 96%. Approximately 64.5% of the populace holds unfavorable perceptions regarding online education. The principal cause of this circumstance is the shift in education from a communal activity to an individualized practice at home, which has resulted in the need for new skill acquisition and a reduction in social activities due to restricted campus access.

Author Contributions

Aqsa Rehman: Conceptualization, Methodology, Software, Writing-Original draft preparation. **Dr. Naeem Aslam**: Supervision. **Kamran Abid**: Validation. **Muhammad Fuzail**: Writing- Reviewing and Editing. **Dr. Asif-Ur-Rehman**: Visualization, Investigation.

Compliance with Ethical Standards

It is declare that all authors don't have any conflict of interest. It is also declare that this article does not contain any studies with human participants or animals performed by any of the authors. Furthermore, informed consent was obtained from all individual participants included in the study.

Author Information

ORCID:
Aqsa Rehman: 0000-0002-7824-7703

References

[1] S. Edosomwan *et al.*, "The history of social media and its impact on business," *Journal of Applied Management and entrepreneurship*, vol. 16, no. 3, pp. 79–91, 2011.

[2] Y. Martínez-Díaz, H. Méndez-Vázquez, L. S. Luevano, M. Nicolás-Díaz, L. Chang, and M. González-Mendoza, "Towards accurate and lightweight masked face recognition: An experimental evaluation," *IEEE Access*, vol. 10, pp. 7341–7353, 2022.

[3] Y. Bengio, H. Schwenk, J. Sencal, F. Morin, and J. Gauvain, "Neural probabilistic language models," *Journal of Machine Learning Research*, vol. 3, no. 6, pp. 1137–1155, 2003.

[4] 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, Dec 2014.

[5] A. Al-Rawi, K. Grepin, X. Li, R. Morgan, C. Wenham, and J. Smith, "Investigating public discourses around gender and covid-19: a social media analysis of twitter data," *J Healthc Inform Res*, vol. 5, no. 3, pp. 249–269, Sep 2021.

[6] K. K. Bhagat, S. Mishra, A. Dixit, and C. Y. Chang, "Public opinions about online learning during covid-19: A sentiment analysis approach," *Sustainability (Switzerland)*, vol. 13, no. 6, Mar 2021.

[7] X. Zhu and J. Liu, "Education in and after covid-19: Immediate responses and long-term visions," *Postdigital Science and Education*, vol. 2, no. 3, pp. 695–699, Oct 2020.

[8] S. Tadesse and W. Muluye, "The impact of covid-19 pandemic on education system in developing countries: a review," *Open Journal of Social Sciences*, vol. 8, no. 10, pp. 159–170, 2020.

[9] C. B. Mpungose, "Emergent transition from face-to-face to online learning in a south african university in the context of the coronavirus pandemic," *Humanit Soc Sci Commun*, vol. 7, no. 1, Dec 2020.

[10] A. D. Dubey. (2020) Twitter sentiment analysis during covid19 outbreak. [Online]. Available: https://ssrn.com/abstract=3572023

[11] B. P. Pokharel. (2020) Twitter sentiment analysis during covid-19 outbreak in nepal. [Online]. Available: https://ssrn.com/abstract=3624719

[12] U. Gunduz, "Stayhome hashtag: Sentiment analysis on twitter during the covid-19 pandemic," *European Scientific Journal ESJ*, vol. 16, no. 34, Dec 2020.

[13] G. G. Chowdhury, "Natural language processing," *Annual Review of Information Science and Technology*, vol. 37, no. 1, pp. 51–89, Jan 2005.

[14] Z. Feng, H. Zhou, Z. Zhu, and K. Mao, "Tailored text augmentation for sentiment analysis," *Expert Syst Appl*, vol. 205, Nov 2022.

[15] X. Li, C. Wu, and F. Mai, "The effect of online reviews on product sales: A joint sentiment-topic analysis," *Information and Management*, vol. 56, no. 2, pp. 172–184, Mar 2019.

[16] W. Song, W. Li, and S. Geng, "Effect of online product reviews on third parties' selling on retail platforms," *Electron Commer Res Appl*, vol. 39, Jan 2020.

[17] M. E. Osman, "Global impact of covid-19 on education systems: the emergency remote teaching at sultan qaboos university," *Journal of Education for Teaching*, vol. 46, no. 4, pp. 463–471, Aug 2020.

[18] K. N. Alam, M. S. Khan, A. R. Dhruba, M. M. Khan, J. F. Al-Amri, M. Masud, and M. Rawashdeh, "Deep learning-based sentiment analysis of covid-19 vaccination responses from twitter data," *Computational and Mathematical Methods in Medicine*, 2021.

[19] D. Wei, B. Wang, G. Lin, D. Liu, Z. Dong, H. Liu, and Y. Liu, "Research on unstructured text data mining and fault classification based on rnn-lstm with malfunction inspection report," *Energies*, vol. 10, no. 3, p. 406, 2017.

[20] R. Lamsal, "Design and analysis of a large-scale covid-19 tweets dataset," *Applied Intelligence*, vol. 51, pp. 2790–2804, 2021.

[21] A. M. Rahat, A. Kahir, and A. K. M. Masum, "Comparison of naive bayes and svm algorithm based on sentiment analysis using review dataset," in *2019 8th International Conference System Modeling and Advancement in Research Trends (SMART)*. IEEE, 2019, pp. 266–270.

[22] K. M. Ridhwan and C. A. Hargreaves, "Leveraging twitter data to understand public sentiment for the covid-19 outbreak in singapore," *International Journal of Information Management Data Insights*, vol. 1, no. 2, p. 100021, 2021.

[23] A. H. Uddin, D. Bapery, and A. S. M. Arif, "Depression analysis from social media data in bangla language using long short term memory (lstm) recurrent neural network technique," in *2019 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2)*. IEEE, Jul 2019, pp. 1–4.

[24] C. Baru, *Sentiment Analysis in Turkish with Deep Learning*.

[25] D. Goularas and S. Kamis, "Evaluation of deep learning techniques in sentiment analysis from twitter data," in *2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML)*. IEEE, Aug 2019, pp. 12–17.

[26] R. S. Jagdale, V. S. Shirsat, and S. N. Deshmukh, *Sentiment analysis on product reviews using machine learning techniques*. Springer Verlag, 2019, pp. 639–647.

[27] A. H. Ombabi, W. Ouarda, and A. M. Alimi, "Deep learning cnn–lstm framework for arabic sentiment analysis using textual information shared in social networks," *Soc Netw Anal Min*, vol. 10, no. 1, Dec 2020.

[28] A. S. Imran, S. M. Daudpota, Z. Kastrati, and R. Batra, "Cross-cultural polarity and emotion detection using sentiment analysis and deep learning on covid-19 related tweets," *IEEE Access*, vol. 8, pp. 181 074–181 090, 2020.

[29] S. S. Cheeti, Y. Li, and A. Hadaegh, "Twitter based sentiment analysis of impact of covid-19 on education globaly," *International Journal of Artificial Intelligence & Applications*, vol. 12, no. 03, pp. 15–24, May 2021.

[30] K. K. Bhagat, S. Mishra, A. Dixit, and C. Y. Chang, "Public opinions about online learning during covid-19: A sentiment analysis approach," *Sustainability (Switzerland)*, vol. 13, no. 6, Mar 2021.

[31] N. Chintalapudi, G. Battineni, and F. Amenta, "Sentimental analysis of covid-19 tweets using deep learning models," *Infect Dis Rep*, vol. 13, no. 2, Apr 2021.

[32] C. Hutto and E. Gilbert, "Vader: A parsimonious rule-based model for sentiment analysis of social media text," in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 8, no. 1, 2014, pp. 216–225.

[33] H. Jelodar, Y. Wang, C. Yuan, X. Feng, X. Jiang, Y. Li, and L. Zhao, "Latent dirichlet allocation (lda) and topic modeling: models, applications, a survey," *Multimedia Tools and Applications*, vol. 78, pp. 15 169–15 211, 2019.

[34] T. Sarkar and N. Rajadhyaksha, "Tla: Twitter linguistic analysis," *arXiv preprint arXiv:2107.09710*, 2021.

[35] "Research on home based online learning of remote rural students during the epidemic," *Curriculum and Teaching Methodology*, vol. 5, no. 9, 2022.

[36] V. Agarwal, "Research on data preprocessing and categorization technique for smartphone review analysis," *International Journal of Computer Applications*, vol. 131, no. 4, pp. 30–36, 2015.

[37] Y. Kang, Z. Cai, C. W. Tan, Q. Huang, and H. Liu, "Natural language processing (nlp) in management research: A literature review," *Journal of Management Analytics*, vol. 7, no. 2, pp. 139–172, 2020.

[38] H. Saleem, A. Naeem, K. Abid, and N. Aslam, "Sarcasm detection on twitter using deep handcrafted features," *Journal of Computing & Biomedical Informatics*, vol. 4, no. 02, pp. 117–127, 2023.

[39] J. J. Webster and C. Kit, "Tokenization as the initial phase in nlp," 2023.

[40] V. K. Ayyadevara, *Word2vec*. Apress, 2018, pp. 167–178.

[41] S. Ghannay, B. Favre, Y. Estève, and N. Camelin. Word embeddings evaluation and combination. [Online]. Available: https://code.google.com/p/word2vec/

[42] K. S. Tai, R. Socher, and C. D. Manning, "Improved semantic representations from tree-structured long short-term memory networks," *arXiv preprint arXiv:1503.0 0 075*, 2015.

[43] A. F. Agarap, "Deep learning using rectified linear units (relu)," *arXiv preprint arXiv:1803.08375*, 2018.

[44] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *Journal of Machine Learning Research*, vol. 3, no. Jan, pp. 993–1022, 2003.

[45] C. Roque, J. L. Cardoso, T. Connell, G. Schermers, and R. Weber, "Topic analysis of road safety inspections using latent dirichlet allocation: A case study of roadside safety in irish main roads," *Accident Analysis & Prevention*, vol. 131, pp. 336–349, 2019.

[46] A. Graves, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, p. 1735, 1997.

[47] J. Xu, D. Chen, X. Qiu, and X. Huang, "Cached long short-term memory neural networks for document-level sentiment classification," in *Proceedings of the EMNLP*, 2016, pp. 1660–1669.

[48] T. Pang, B. Pang, and L. Lee, "Thumbs up? sentiment classification using machine learning," in *Proceedings of the EMNLP*, 2002, pp. 79–86.

[49] T. Mikolov, M. Karafit, L. Burget, J. Cernocky, and S. Khudanpur, "Recurrent neural network based language model," in *Proceedings of the INTERSPEECH Conference of the International Speech Communication Association*, 2010, pp. 1045–1048.

[50] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv preprint arXiv:1301.3781*, 2013.