Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 45 No. 2 (2024)

Predicting Elections: Unveiling the Role of Social Media Data and Sentiment Analysis with ResBiLSTM Methodology

, Umanesan R[1],Sameer Yadav[2], Thiru Murugan R[3], Gagan Singh [4],.Nirmala Devi M[5], D Suganthi[6]

[1] Department of CSE, Saveetha School of Engineering, SIMATS, Thandalam, Chennai, Tamilnadu, India,

[2] Research Scholar, Department of Commerce and Business Administration, University of Allahabad, Prayagraj, Uttar Pradesh, India,

[3] Department of Business Administration, Kalasalingam Business School, Kalasalingam Academy of Research and Education, Tamilnadu, India,

[4] BBA Department, Meerut Institute of Technology, Meerut, Uttar Pradesh, India,

[5] Department of English, St.Martin's Engineering college, Secunderabad, Telangana, India,

[6]Department of Computer Science, Saveetha College of Liberal Arts and Sciences, SIMATS, Thandalam, Chennai, India,

Abstract. Several volumetric, sentiment, and social network methodologies are presented and assessed in this study with the aim of predicting significant decisions from social media platforms. The opinions of the people are quite essential when it comes to determining certain major decisions. People from all walks of life have had a global platform for public expression on social media for almost twenty years. When trying to gauge public opinion, sentiment analysis also called opinion mining can be useful. Each step of the process from selecting a method to preprocessing, feature extraction, and training the model requires exact sequentially. There were two distinct phases to the data preparation process. Separate from cleaning the Twitter data that was retrieved is data wrangling, this comprises changing the format of the data. A feature extraction process makes use of ngrams, which are collections of n words taken from a source text. Extracting features is the first step in training a ResBiLSTM model. Both Resnet and BiLSTM, two state-of-the-art algorithms, are surpassed by this novel approach. The results show that the accuracy reached 95.45%, which is a huge improvement.

Keywords: Election prediction · Categorical Proportional Difference (CPD) · Natural Language Processing (NLP).

1 Introduction

Predictions for the next presidential election have piqued the curiosity of both scholars and the broader public. The academic literature is dominated by two primary schools of thinking about election outcome prediction. There was an initial trend in political science. Political scientists have been working on models to predict elections since the 1980s. These models examine the interrelationships among economic development, a number of predictive variables, and the expected vote outcomes of a single presidential candidate, typically from the incumbent party. The second thread originates from the field of computer science. As big data on social media has grown in popularity in the 2010s, some academics have started to analyze Twitter sentiment as a predictor of future elections. Obtaining very accurate sentiment ratings from pertinent social media posts is usually the main goal. When voters see a candidate in a positive light, they are more likely to vote for that person. Although there are many scientific and practical advantages to both techniques, there are also some drawbacks. When using traditional models to forecast election outcomes, polling rates are essential. Poll surveys, however, can be expensive and time-consuming to conduct. The term "fake news" has grown in popularity as a blanket term for any deceptive material shared on the internet. Have examined the typology and distinguishing features of different types of incorrect information. Conversely, contend that further academic study is needed to examine where the prevalent disinformation comes from, how extensive it is, and what effects it has. The findings of this study add to the mounting evidence that the spread of politically motivated disinformation and false news can have serious consequences for society and government. The current conversation over the effects of the post-truth or fake news problem on society lacks evidence. Examples of such arguments include the claims made, who argue that the continued spread of disinformation poses a threat to democracy because it undermines public trust in government. The dissemination of misleading information is the most pressing social problem facing American adults today, according to a nationwide survey. The impact of such deceptive information on democratic processes has only been somewhat investigated, with only two studies addressing this significant issue are among the works that have focused on this topic previously. These works primarily involve experiments that examine how incorrect information influences people's opinions on the stories. An increasing amount of research has focused on electoral forecasting since its inception. Having trustworthy results projected is important for many people, including politicians, practitioners, and policymakers. This is true for both the present and the future. Such a feat would have an effect on party financing, political strategy, and strategic decisions. Big data, political markets, and polls have all played a role in creating complex models that try to make better predictions. More precise predictions were possible than with polls alone thanks to online data sources and the Internet in this case. People who utilize the Internet are more informed and more inclined to vote as a result. A number of studies have used search-volume indices such as Google Trends or tweet analysis to make predictions about online news, despite the fact that there are large datasets that contain this information found that there are a number of situations when looking at online news can help with financial, political, and economic predictions. An officially acknowledged group of people with shared political views and aims whose declared purpose is to influence government policy by electing party members to public office is called a political party. A political party takes over the government when it’s officially supported or approved candidates win the election. Also, they are in charge of political campaigns and making sure that people vote. To win elections and use those victories to shape public policy is the raison d'être of political parties. They need to rally a wide range of voters with shared values if they want to win. Whether their work is directly apparent to the public through the presentation of candidates or the electoral campaign, political parties perform multiple crucial functions that may influence the people who are eligible to vote. It's widely known that campaign appearances by candidates can also affect election outcomes. With the proliferation of social networking, microblogging, and blogging websites, individuals now have more methods than ever before to express themselves and generate vast quantities of data. Notable studies have consistently found little evidence to support the use of social media data for election prediction, and the criticism around this practice has been strong. You could begin to question these claims until you realize that most of these successful "predictions" happened after the election, when all the facts are known. Many famous studies that looked to have succeeded fell apart when evaluated objectively. Nevertheless, disregarding any technological worries, a comprehensive analysis of this topic uncovers some fascinating fundamental issues that occur in any attempt to understand and maybe predict human behavior using data from social media.

2. Literature Survey

The use of probability sampling and the expected form and solicitation of participant data are two of the most fundamental assumptions of conventional survey research that have been disrupted by social media predictions. [1] Starting off, probability-based sampling, a foundational method in scientific survey research, is "a possible compromise to measure the climate of public opinion" by assuming that all opinions should be correctly weighted, since they are all equally valid. A critical component in molding public opinion, the fact that individuals' interpersonal influence levels differ across the population is, however, ignored by this method [2]. We don't learn much about the leaders' views, which can influence the attitudes of many other citizens, from traditional surveys, although they can occasionally reveal whether the individuals who have a given perspective are a coherent unit. The predictive power of social media analysis is independent of user representativeness, as pointed out by [3] in their study. According to them, "if we assume that the politically active internet users act like opinion-makers who can influence (or to 'anticipate') the preference of a wider audience: consequently, it would be found that the preferences expressed through social media today would act (predict) the opinion of the entire population tomorrow" [4] 345 pages. To restate: just because social media data doesn't represent the population accurately doesn't mean it's useless for opinion mining. Also, it's possible to oversimplify public opinion by viewing it only through the lens of publicly expressed views on particular issues; [5]this approach misses the mark when it comes to the values and impact that are inherent in popular opinion. The ability to see into the future would be a godsend for many. [6] Several studies have proven that social media data is crucial for this type of prediction. Among the many fields that have made use of social media data for prediction are marketing, politics, and finance. [7] Our research in this area led us to concentrate on social media prediction and human behavior mining. Predicting the results of elections has thus been the writers' primary subject, and this has attracted a great deal of interest. The challenges of using social media data to forecast election outcomes, which is why the authors were inspired to perform this analysis. [8] Social media sentiment research can help us understand the public's mood and draw meaningful conclusions. Plus, there are a tonne of messages in the form of tweets, and technically speaking, collecting them is easy and straightforward. When it comes down to it, Twitter seems to be a fantastic spot to gauge public opinion on various political candidates and parties. [9] Another reason to perform this research is that most prominent figures, including politicians, celebrities, and actors, and normal users of both sexes and all ages (over thirteen), communicate their opinions on a variety of themes through tweets. Thus, reading their tweets can teach us about their emotions, which can help with data mining for human behavior, which can improve our ability to make predictions and educated decisions. [10] Another reason to look at social media sentiment is that it's more cost- and time-effective to acquire public opinion through social media platforms than to do a field survey. For example, polls are sent to a certain group of individuals to gauge their views on different political parties and the impending election in attempt to predict the results.[11] The phrase "social media" encompasses a wide range of online platforms, including but not limited to YouTube, Facebook, Instagram, LinkedIn, Twitter, WhatsApp, and Facebook. [12] Online forums like these bring people together from all over the globe to talk about societal, political, and economic issues. Politics being what it is, it's rare for a political event to go unnoticed on social media. Since elections are significant events in the lives of people all around the world, they naturally attract a lot of attention and discussion on social media. [13] Multiple studies have shown that the abundant usage of social media can drastically change the political scene. [14] The bulk of academics agree that social media facilitates extensive communication and influences election outcomes. Two well-known case studies that show how social media was crucial to the success of Barack Obama's 2008 US presidential campaign. [15] Since the general elections of 2014 in India, both politicians and ordinary residents have made heavy use of social media, especially Twitter. [16]Instagram didn't launch until 2010, Facebook and Twitter didn't launch until 2006, and other modern SM platforms are still in their early stages. Social media began to be used in modern political activities and was even considered as a method to forecast elections soon after it was released. It is possible that Tilton was among the first to attempt to predict election outcomes using SM data [17]."Could Facebook be used to estimate the results of a student election?" was the question he set out to address. It took place barely two years after Facebook became publicly accessible. [18] In this case, the linked society in issue was a school. His model has a 21 percent success rate in predicting the candidates' final total out of 27 elections. Despite the fact that other researchers in the field use Tilton's work less frequently, we think it provides a very useful foundational examination of the subject (maybe because it is unrelated to formal political scenarios). Two were pivotal in utilizing SM data to predict political elections, and nearly all subsequent studies credit them. In 2010, [19] documented the results of the German federal election that took place. They were able to correlate the amount of tweets with the results of the election by collecting all tweets that mentioned any of the six parliamentary parties in Germany or prominent politicians from these parties. The researchers claimed that "the mere number of tweets mentioning a political party can be considered a plausible reflection of the vote share and its predictive power even comes close to traditional election polls" as a result. [20]replicated consumer confidence and presidential job approval polls from the same year by using an emotion detector based on Twitter data. Study participants' levels of positive and negative messaging were intended to be related to the outcomes of the election. [21] They gathered all tweets on a political party or candidate and classified them as positive, negative, or neutral using sentiment analysis. Data collection through sentiment analysis and an open search on Twitter proved to be the most challenging aspects of our investigations. These online applications, often referred to as social media, allow individuals from all over the globe to engage with one another. These platforms make it easier for people to communicate, share knowledge, and work together. [22]state that it can be seen as a tool for education that promotes information sharing and discovery, which in turn facilitates connections with like-minded individuals. Contrary to other instances, several businesses, such as hospitality and tourism, use websites and mobile applications for brand management [23]. These characteristics are largely responsible for the meteoric rise of social media during the past two decades. [24]notes that in 2003, a singular social network called MySpace enabled users to make free profiles, interact with each other, and share material, which led to the network's subsequent stratospheric ascent. The intended audience loved these new features. In less than a year, we went from zero to one million users and briefly overtook Google in US traffic. Roughly one-third of the world's population is linked to some kind of social network. Years after MySpace went live, other popular social networks began to draw users; one such platform is Facebook [25]. Since social media platforms enable communication regardless of physical location, a growing number of people utilize them regularly. No matter where they were in the world, users could virtually meet others who shared their passions and perspectives. The rise of these platforms brought to a new phenomenon: the online polarization effect on users. Topics abound on these social media sites because to the proximity of individuals who have same interests and personalization algorithms that show users what they've indicated a preference. The rest of this paper is structured as follows. The methods section covers the study area, data acquisition, processing, and analysis, while the materials part details the study methods, model evaluation criteria, and training parameters for ResNet, BiLSTM, and ResBiLSTM models. Section II provides an introduction to the methods and materials. Section III details the results of the ResBiLSTM model's predicted accuracy for various growth stages, day delays, and model comparisons, while Section IV describes the additional study. Lastly, Section IV presents the conclusions.

3. Proposed System

Since the advent of social media, users have had a strong voice to express themselves. Institutions like businesses can't make better decisions without first understanding the public's stance on important subjects. Political entities utilize public opinion data to inform their campaign strategy; hence the arena of politics is one such use. A lot of individuals believe that sentiment analysis on data from social media is a fantastic method to monitor people's preferences.

3.1 Preprocessing

The data preparation phase was divided into two parts. Data wrangling, which entails altering the data's format, and cleaning the retrieved Twitter data are two separate processes.

3.1.1 Wrangling of data

For the aim of analysis or human inspection, data wrangling is employed to examine complicated and, in theory, massive data sets. Data wrangling can be approached in various ways. Some of these alternatives include graphing, schema matching, converting data repair value formats, entity settlement and merging, and data extraction from sources such as deep networks or online tables. The study used data wrangling to convert the raw data obtained from Twitter into a more understandable manner. Data modification occurs when multiple tables' worth of information is combined into a single processing component via a join query. It is possible to collect information about both the user and their tweets by combining data from the Tweet table with the User database in a specific method.

3.1.2 Cleaning of Data

In order to make data analysis easier, data cleaning involves stripping tweets of unnecessary characters such as spaces, numbers, punctuation, and capitalization. Data cleansing entails three stages: cleaning, stop-wording, and stemming [22]. To clean it up, you have to change all the capital letters to lowercase and remove any numbers, spaces, or punctuation. The term "stop-word" describes words that are not as important as other symbols. As it stands, no NLP system employs a universal, language-specific blacklist of words [26]. When it comes to twitter data, terms like "on" and "on" in English are treated as stop words. Although these terms don't really signify much, they're essential for making filtering, indexing, crawling, and final tweeting more efficient. Some types of words are excluded from stemming in order to create a standard morphological definition called the stem. When two words share a root, stemming makes them simpler by removing the inflective and derivational affixes from them. On sometimes, only the suffixes that are attached to the right side of the root are left out. Following stemming, there will be more material phrases and a better chance that they will meet the context notion.

3.2 Feature Extraction

In this case, ngrams provide the foundation for all feature extraction methods. A set of n words taken from a text is called a ngram. This sequence is known as a unigram when only one word is chosen at a time. Another option would be to select pairs of words from the text. For each of these chosen sequences, the word "bigram" is most suited. Here, it ran with the bag of words method, which uses a ngram's document-specific frequency as its foundation. Each word in a set of words, called a vocabulary, which consists of unigrams (one word), is assigned an index number r. h_r is the word with index number r, and v is the number of words in the vocabulary.

$$n = \{h_1, h_2, h_3, \dots, h_r, \dots, h_v\}, r, v \in \mathbb{V} \quad (1)$$

It is possible to represent a document as a vector f with n lexical dimensions. The document-wide counting frequency of h_r from vocabulary n is stored in every element g_r of vector f, in reference to (2).

$$f = [g_1, g_2, g_3, \dots, g_r, \dots, g_v], r, v \in \mathbb{V} \quad (2)$$

Selecting appropriate vocabulary has an immediate impact on the vector representation of documents. As a rule, vocabulary consists merely of a compilation of all unigrams extracted from corpora. It can reduce the original vocabulary to a more manageable set of ngrams that are better suited for class separation by employing feature selection techniques. The Chi-Square test can be used to determine the level of relationship between ngram and classes [23]. When ngrams occur frequently in many classes, the Chi-Square value will be low; when they occur in few classes, the value will be high. Use Categorical Proportional Difference (CPD) to find out how much of an impact a term has on class discrimination. It was originally designed to be used for text classification tasks. For each class, a word's CPD can take on a value between -1 and 1, where 1 indicates that the term never appears in any class and -1 indicates that it appears solely in that class. Finding ngrams and phrases with CPDs greater than a certain level allows us to focus our search. Categorical Probability Proportional Difference (CPPD) is an optimization of the prior method [27]. With CPPD, we may measure how much an ngram or sentence fits into a given class. Simply put, CPD merely assesses the distribution of classes, but CPPD rates them based on the likelihood of the same ngram or phrase occurring in that class and how those classes are distributed. Our search is then limited to ngrams and phrases with CPD and probability values greater than certain thresholds. For this investigation, can determine the optimal probability ranking terms by applying the following restriction to CPPD. Notation for the necessary feature count is n, and the critical path distance (CPD) is 1.

3.3 Model Training

3.3.1 ResNet

In local receptive fields, ResNet's foundational arithmetic integrates channel-wise and spatial-wise input to uncover significant characteristics. The ResNet architecture revolves around the residual block. This block's components include an activation layer called ReLU, a batch normalization layer called BN, a convolution layer called the weight layer, an output called $W(t)$, and a residual mapping function called $D(t)$. It is not an easy process to adapt the model so it matches the actual mapping $W(t)$. The restructured ResNet residue building components in a hierarchical format. In order to solve the problem of performance loss caused by network stacking, ResNet implements a "skip connection" that allows it to directly fit the residual mapping $D(t)$. The model can approach the actual mapping, which is actually$W(t) = D(t) + t$, by decreasing the residual function $D(t) = W(t) - t$.

$$t^{(y+1)} = t^y + \sum_{r=1}^{T-1} D\left[t^{(r)} + H^{(r)}\right] \tag{3}$$

Internal covariate shift occurs when there is a change in the data distribution of internal nodes during training the model. The BN layer is used to rectify issues with improper or uneven data distribution in deep neural networks. With the BN layer, training deep neural networks becomes much faster [28]. Data is normalized after the activation function transforms the input data from the preceding layer in a non-linear way. This retains the network's characterization ability, speeds up convergence, makes sure the network can be trained, and reduces the influence of big changes on the network's internal distribution. With its sparse activation characteristic, the Rectified Linear Unit (ReLU) activation function avoided overfitting. Mathematically, it seems like

$$\frac{\partial loss}{\partial t^{(y)}} = \frac{\partial loss}{\partial t^{(Y)}} \frac{\partial t^{(Y)}}{\partial t^{(y)}}$$
$$= \frac{\partial loss}{\partial t^{(Y)}} \left(1 + \frac{\partial}{\partial t(y)} \sum_{r=1}^{Y-1} D\left(t^{(r)} + H^{(r)}\right)\right) \tag{4}$$

This study employed an improved residual block structure that addressed the issue of gradients disappearing or exploding during deeper network training. The modified structure differed from the original version in that gradients in the modified block could be directly connected to any earlier layer via shortcuts. When the data distribution of internal nodes changes during model training, it causes internal covariate shift. To address the issue of inconsistent and uneven data distribution in deep neural networks, the Batch Normalization (BN) layer is employed. Deep neural networks can be trained more quickly with the help of the BN layer. It normalizes the data after the activation function does a non-linear transformation on the input data from the previous layer. This way, the neural network can keep the input data distribution consistent, be trained, have its internal distribution less affected by large changes, converge faster, and keep its characterization ability. To prevent overfitting, Rectified Linear Unit (ReLU) exhibited sparse activation as an activation function. This is the mathematical formula:

$$d(y) = \max(0, y) \tag{5}$$

3.3.2 BiLSTM

An adaptation of the RNN architecture, the Memory Unit is the fundamental building block of the Long Short-Term Memory (LSTM) structure; it is able to forget, remember, and output via its gate structure. You can control the information state by calculating the forgetting gate d_s, memory gate r_s, and output gate m_s using the prior hidden state w_{s-1} and the current input y_s. With the Memory Unit, RNN no longer experiences gradient explosion or gradient disappearance since it remembers crucial data while discarding irrelevant data.

3.3.3 ResBiLSTM

The ResBiLSTM model integrates two sub-models to extract spatial-temporal information from the dataset. The meta-learner enhances its data fitting capabilities by gaining a deeper understanding of the derived spatio-temporal features. In order to combine and understand the gridded SWC data from the ResNet branch and the meteorological time series data from the BiLSTM branch, a meta-learner is used. The model is trained using Adam, an iterative optimization strategy, with MSE serving as the loss function. Combining continuous multi-depth SWC with weather data can transform several time series datasets into a two-dimensional matrix. Using a top-down arrangement of time series, the grid data matrix shows meteorological and multi-depth SWC variables from left to right. This gridded time series data can be described in a few different ways:

$$T_{m,x} = \begin{bmatrix} T_{1,x-v} & T_{2,x-v} & \cdots & T_{m,x-v} \\ T_{1,x-v+1} & T_{2,x-v+1} & \cdots & T_{m,x-v+1} \\ \vdots & \vdots & \ddots & \vdots \\ T_{1,x} & T_{2,x} & \cdots & T_{m,x} \end{bmatrix} \quad (6)$$

the input variables' count (m), the historical time step x, and the input items' matrix $T_{m,x}$. Grid data used by the ResNet subset undergoes a progression that includes a 2D convolution layer, two subsequent residual blocks, output flattening, and a full connection layer containing 62 filters. In order to keep the model lightweight and make sure enough information is recovered, considering how small the input pieces are in comparison to the images, it is constructed with two layers of leftover blocks. A 2 x 2 convolution kernel size and 62 feature mappings make up the 2D CNN layer [24]. A 2D convolutional neural network (CNN) layer with 62 filters and a 4 x 4 convolution kernel size is set up in the same way for both residual blocks. The BiLSTM branch flattens the input items before feeding them into two successive layers. Lastly, they output to a fully linked layer of 62 filters. We can improve our ability to capture time domain characteristics by stacking the two BiLSTM layers and superimposing a trend prediction on top of each time step's forecast. Using fully linked layers, we can train a metalearner to learn more feature parameters from the ResNet and BiLSTM branches. The initial step in creating a dense layer of 252 neurons is to splice 130 neurons from each branch. All feature layers in the ResNet and BiLSTM branches use the ReLU function in their activation functions to prevent overfitting. The layer responsible for learning meta-output makes use of the linear activation function.

4. Result and Discussion

The "sentiment analysis" practice of attempting to assess public opinion on a certain problem or event has piqued the interest of natural language processing experts since the emergence of social media in the previous decade. Finding out how someone feels about a topic is what sentiment analysis is all about. The public's perception of a candidate impacts their likelihood of becoming the leader of the nation. A large, diverse data collection collected from Twitter represents the current public opinion on the candidates. It uses a lexicon-based approach to analyze the collected tweets and determine the public's sentiment.

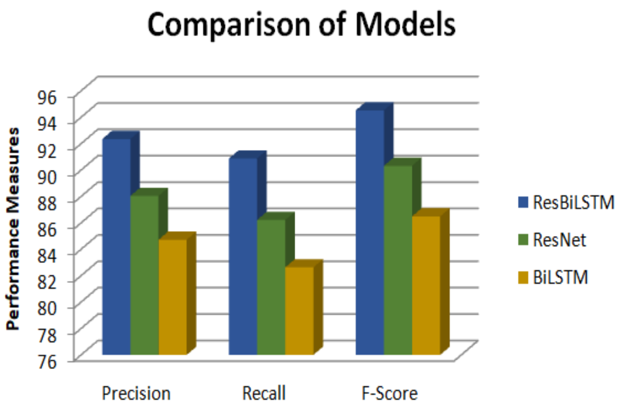

Fig. 1. Comparison of Precision, Recall and F-measure for all CNN Models

When comparing the three models' performance on the dataset, as shown in Figure 1, the results demonstrated significantly improved f-measure, precision, and recall. Resnet is well-known for feature selection, and Bi-LSTM empowers the model to incorporate context by supplying past and future sequences; together, they form ResBiLSTM, a dataset classification algorithm.

Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 45 No. 2 (2024)

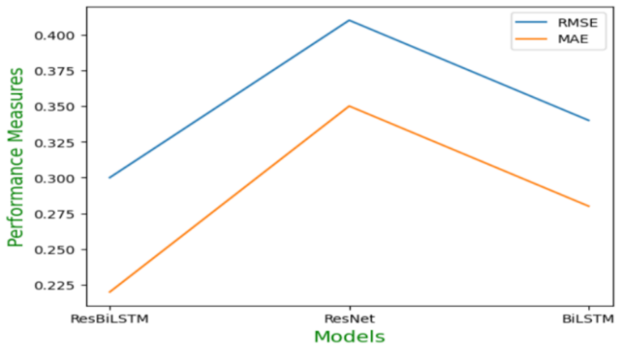

Fig. 2. Comparison of Error Rate

Figure 2 shows a comparison of the ResBiLSTM models' error rates as measured by MAE and RMSE. The ResBiLSTM models with various parameter settings are shown on the X-axis of the bar chart, while the performance measurements for each ResBiLSTM model are shown on the Y-axis.

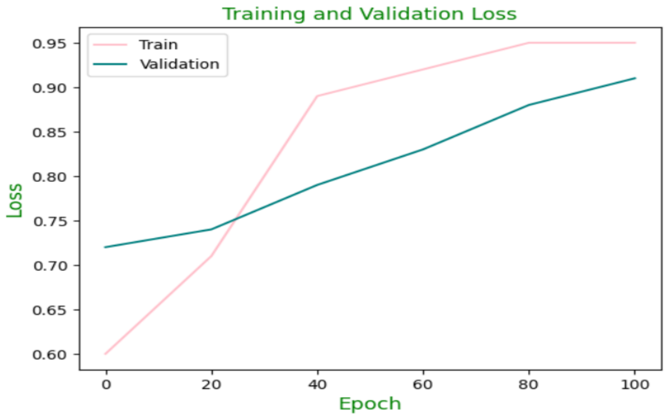

Fig. 3. Training and Validation Accuracy of ResBiLSTM

The models' ResBiLSTM training and validation accuracy is displayed in Figure 3. Different performances and forecast times are displayed at epoch 20. Therefore, ResBiLSTM provides us with the highest validation accuracy and the shortest execution time.

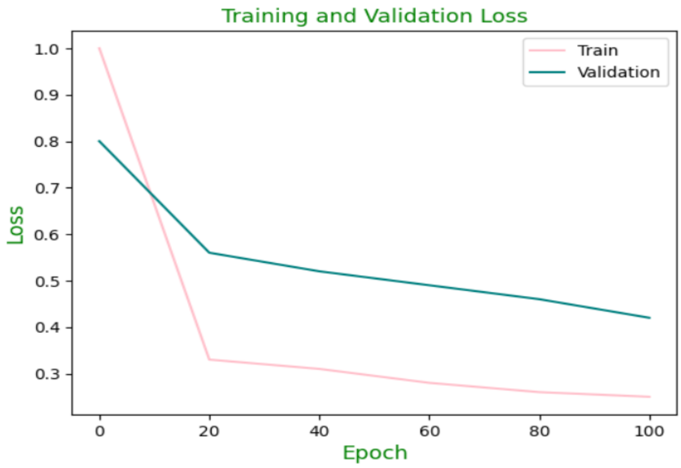

Fig. 4. Training and Validation Loss of ResBiLSTM

Model loss behavior is seen in Fig. 4. Figure 4 illustrates that there is overfitting since the training loss is nearly zero and the validation set loss is rising. Training loss of 0.25% for the proposed model.

5. Conclusion

Sentiment analysis is a part of computer science that studies how people express themselves emotionally and subjectively in written form. When used in this sense, "polarity" denotes a problem with classifying words, more especially with identifying those words as positive or negative. Assessing public opinion of candidates and political parties has never been easier than with Twitter sentiment analysis. The procedure of preparing the data consisted of two separate steps. Separate from cleaning the obtained Twitter data is data wrangling, which involves altering the data's format. In a feature extraction procedure, ngrams groups of n words extracted from a source text are utilized. All of the accessible parameters are considered by the ResBiLSTM algorithm during training. The proposed strategy routinely achieves better results than both the ResNet and BiLSTM models, which achieve an average accuracy of 95.45 percent.

References

[1] J. Ratkiewicz, M. D. Conover, M. Meiss, B. Gonçalves, A. Flammini, and F. Menczer, “Detecting and Tracking Political Abuse in Social Media,” *Proc. 5th Int. AAAI Conf. Weblogs Soc. Media, ICWSM 2011*, pp. 297–304, 2011, doi: 10.1609/icwsm.v5i1.14127.

[2] B. Metin, M. Atasoyu, E. Arslan, N. Herencsar, and O. Cicekoglu, “A meta-analysis of state-of-the-art electoral prediction from Twitter data,” *Midwest Symp. Circuits Syst.*, vol. 2017-Augus, pp. 739–742, 2017, doi: 10.1109/MWSCAS.2017.8053029.

[3] A. Ceron, L. Curini, S. M. Iacus, and G. Porro, “Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens’ political preferences with an application to Italy and France,” *New Media Soc.*, vol. 16, no. 2, pp. 340–358, 2014, doi: 10.1177/1461444813480466.

[4] W. N. Espeland and M. Sauder, “Rankings and reactivity: How public measures recreate social worlds,” *Am. J. Sociol.*, vol. 113, no. 1, pp. 1–40, 2007, doi: 10.1086/517897.

[5] K. H. Kwon, S. Il Moon, and M. A. Stefanone, “Unspeaking on Facebook? Testing network effects on self-censorship of political expressions in social network sites,” *Qual. Quant.*, vol. 49, no. 4, pp. 1417–1435, 2015, doi: 10.1007/s11135-014-0078-8.

[6] A. Bermingham and A. F. Smeaton, “On Using Twitter to Monitor Political Sentiment and Predict Election Results,” *Psychology*, pp. 2–10, 2011.

[7] Saravanakumar, S., & Thangaraj, P. (2019). A computer aided diagnosis system for identifying Alzheimer’s from MRI scan using improved Adaboost. Journal of medical systems, 43(3), 76.

[8] A. Boutet, H. Kim, and E. Yoneki, “What’s in your tweets? I know who you supported in the UK 2010 general election,” *ICWSM 2012 - Proc. 6th Int. AAAI Conf. Weblogs Soc. Media*, pp. 411–414, 2012, doi: 10.1609/icwsm.v6i1.14283.

[9] Saravanakumar, S., & Saravanan, T. (2023). Secure personal authentication in fog devices via multimodal rank-level fusion. Concurrency and Computation: Practice and Experience, 35(10), e7673.

[10] W. Budiharto and M. Meiliana, “Prediction and analysis of Indonesia Presidential election from Twitter using sentiment analysis,” *J. Big Data*, vol. 5, no. 1, pp. 1–10, 2018, doi: 10.1186/s40537-018-0164-1.

[11] Thangavel, S., & Selvaraj, S. (2023). Machine Learning Model and Cuckoo Search in a modular system to identify Alzheimer’s disease from MRI scan images. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 11(5), 1753-1761.

[12] S. Ahmed, K. Jaidka, and J. Cho, “The 2014 Indian elections on Twitter: A comparison of campaign strategies of political parties,” *Telemat. Informatics*, vol. 33, no. 4, pp. 1071–1087, 2016, doi: 10.1016/j.tele.2016.03.002.

[13] D. Gayo-Avello, “A Balanced Survey on Election Prediction using Twitter Data,” pp. 1–13, 2012, [Online]. Available: http://arxiv.org/abs/1204.6441.

[14] Saravanakumar, S. (2020). Certain analysis of authentic user behavioral and opinion pattern mining using classification techniques. Solid State Technology, 63(6), 9220-9234..

[15] A. Herdağdelen, W. Zuo, A. Gard-Murray, and Y. Bar-Yam, "An exploration of social identity: The geography and politics of news-sharing communities in twitter," *Complexity*, vol. 19, no. 2, pp. 10–20, 2013, doi: 10.1002/cplx.21457.

[16] Kumaresan, T., Saravanakumar, S., & Balamurugan, R. (2019). Visual and textual features based email spam classification using S-Cuckoo search and hybrid kernel support vector machine. Cluster Computing, 22(Suppl 1), 33-46.

[17] S. Tilton, "Virtual polling data: A social network analysis on a student government election," *Webology*, vol. 5, no. 4, 2008.

[18] M. Cerezo, "Las nociones de Sachverhalt, Tatsache y Sachlage en el Tractatus de Wittgenstein," *Anu. Filos.*, vol. 37, no. 2, pp. 455–479, 2004, doi: 10.15581/009.37.29387.

[19] E. Sang and J. Bos, "Predicting the 2011 dutch senate election results with twitter," *Proc. Work. Semant. Anal.* ..., no. 53, pp. 53–60, 2012, [Online]. Available: http://dl.acm.org/citation.cfm?id=2389969.2389976%5Cnhttp://dl.acm.org/citation.cfm?id=2389976.

[20] J. K. Mandal, S. C. Satapathy, M. K. Sanyal, P. P. Sarkar, and A. Mukhopadhyay, "Modeling Indian General Elections: Sentiment Analysis of Political Twitter Data," *Adv. Intell. Syst. Comput.*, vol. 339, no. October, 2015, doi: 10.1007/978-81-322-2250-7.

[21] N. D. Prasetyo and C. Hauff, "Twitter-based election prediction in the developing world," *HT 2015 - Proc. 26th ACM Conf. Hypertext Soc. Media*, pp. 149–158, 2015, doi: 10.1145/2700171.2791033.

[22] N. Anstead and B. O'Loughlin, "Social media analysis and public opinion: The 2010 UK general election," *J. Comput. Commun.*, vol. 20, no. 2, pp. 204–220, 2015, doi: 10.1111/jcc4.12102.

[23] C. H. Chan and K. W. Fu, "The Relationship Between Cyberbalkanization and Opinion Polarization: Time-Series Analysis on Facebook Pages and Opinion Polls During the Hong Kong Occupy Movement and the Associated Debate on Political Reform," *J. Comput. Commun.*, vol. 22, no. 5, pp. 266–283, 2017, doi: 10.1111/jcc4.12192.

[24] J. Chung and E. Mustafaraj, "Can Collective Sentiment Expressed on Twitter Predict Political Elections?," *Proc. 25th AAAI Conf. Artif. Intell. AAAI 2011*, pp. 1770–1771, 2011, doi: 10.1609/aaai.v25i1.8065.

[25] T. Diehl, B. E. Weeks, and H. Gil de Zúñiga, "Political persuasion on social media: Tracing direct and indirect effects of news use and social interaction," *New Media Soc.*, vol. 18, no. 9, pp. 1875–1895, 2016, doi: 10.1177/1461444815616224.

[26] L. A. Bokhoeva, A. B. Baldanov, V. E. Rogov, A. S. Chermoshentseva, and T. Ameen, "the Effect of the Addition of Nanopowders on the Strength of Multilayer Composite Materials," *Ind. Lab. Mater. Diagnostics*, vol. 87, no. 8, pp. 42–50, 2021, doi: 10.26896/1028-6861-2021-87-8-42-50.

[27] C. M. A. Carvalho, H. Nagano, and A. K. Barros, "A Comparative Study for Sentiment Analysis on Election Brazilian News," *Proc. 11th Brazilian Symp. Inf. Hum. Lang. Technol.*, pp. 103–111, 2017, [Online]. Available: https://aclanthology.org/W17-6613.

[28] J. Yu *et al.*, "A deep learning approach for multi-depth soil water content prediction in summer maize growth period," *IEEE Access*, vol. 8, pp. 199097–199110, 2020, doi: 10.1109/ACCESS.2020.3034984.