ECTI Transactions on Computer and Information Technology
Journal homepage: https://ph01.tci-thaijo.org/index.php/ecticit/
Published by the ECTI Association, Thailand, ISSN: 2286-9131

A Hybrid GloVe-BERT Fusion Model with Multi-Level Attention-Based CNN-BiLSTM for Sentiment Analysis

Yashaswini Iyer[1] and Gnanaprasanambikai L[2]

ABSTRACT

Gauging public sentiment toward climate policy from information-rich news headlines remains challenging for conventional text classification approaches. Conventional sentiment analysis tools miss contextual subtleties in brief headlines, whereas deep learning models capture the public perception more accurately. The proposed work presents hybrid Convolutional Neural Network (CNN) and Bidirectional Long Short-Term Memory (BiLSTM) model that uses combination of GloVe and BERT embeddings with an attention layer for sentiment analysis of climate change news headlines. The novelty of this research lies in the use of GloVe and BERT embeddings through a multi-stage fusion strategy and an attention mechanism to enhance text classification performance in a hybrid model. The architecture employs a hierarchical layering approach to fuse static GloVe embeddings with dynamic, contextualized BERT representations through attention modules that enables the network to selectively focus on salient features. To further model complex semantic dependencies, the design incorporates parallel CNN-BiLSTM branches, structured with residual connections and bolstered with additional layers of attention. Evaluated on 1,023 climate-related headlines annotated on a three-point polarity scale, the proposed model achieves an accuracy of 80.47%, outperforming classical baselines (SVM, Naive Bayes, K-NN) and single branch deep networks (CNN:78.63%, BiLSTM:78.36%). The predictive accuracy of the hybrid model is evaluated using a paired t-test to determine whether the difference between models is statistically significant; this is confirmed by rejecting null hypothesis and accepting alternate hypothesis.This study demonstrates that compact, domain-adaptive deep learning models incorporating contextual embeddings and attention mechanisms that can effectively extract sentiment from news headlines, offering scalable, evidence based tools for tracking climate discourse and information policy decisions.

Article information:

Keywords: GLoVe, BERT, CNN, BiLSTM, Attention mechanism

Article history:
Received: September 4, 2025
Revised: October 23, 2025
Accepted: December 4, 2025
Published: December 27, 2025 (Online)

DOI: 10.37936/ecti-cit.2026201.263366

1. INTRODUCTION

Sentiment analysis is vital for understanding public perceptions of climate change, primarily through news headlines [5][13]. Social media specifically X (formerly Twitter), provide a unique opportunity to understand the public sentiments on societally relevant topics such as environmental challenges and climate change. Mi and Zhan (2023) show that sentiment analysis of X posts related to climate change reveals predominantly negative sentiment among opponents, while supporters primarily express fear concerning extreme weather. [5] found that news headlines exhibit positive correlation between sentiments related to "low carbon" and "electric cars," highlighting media's role in shaping public perceptions. Sentiment Analysis using traditional methods, such as lexicon-based systems, often misses contextual nuances [21]. In this research study, a dataset of over 1,000 climate-focused news headlines was taken from Kaggle is used to train a CNN-BiLSTM model for accurate sentiment classification, with the aim of informing environmental policy and enhance public awareness. The Dataset, consisting of climate-

[1,2] The authors are with the Department of Mathematics, Amrita School of Physical Sciences, Coimbatore, Amrita Vishwa Vidyapeetham, India, Email: yashaswiniiyer71@gmail.com and l_gnanaprasanambikai@cb.amrita.edu

focused news headlines with sentiments ranging from negative to positive, is well - suited for sentiment classification. This study aims to evaluate the proposed model's effectiveness in handling context-rich texts, uncovering media narratives, and interpreting public sentiment dynamics to support informed policymaking.

From an academic and methodological perspective, climate change sentiment analysis is used interdisciplinary research across environmental communication, sociology, psychology, and data science. The model leverages advanced computational methods such as BiLSTM networks, demonstrating methodological innovation and expanding research capabilities in data analysis and natural language processing (NLP). Ongoing sentiment monitoring provides longitudinal data, allowing researchers and policymakers to identify evolving public attitudes, emotional shifts, and responses to climatic events or international negotiations [13]. Consequently, sentiment analysis emerged as a comprehensive analytical framework that supports effective climate communication, strategic policymaking, and sustained academic inquiry.

Sentiment analysis has evolved from conventional machine learning approaches to highly develop deep learning techniques, each offering varying degrees of accuracy and contextual insight. In contrast, deep learning architectures, such as CNNs, LSTM, BiLSTM, and Transformers models (e.g., BERT) have recently been employed to capture varied complexities and subtle nuances of the human language [14]. BiLSTM models are particularly effective for sentiment analysis because they read process text in both forward and backward directions, thus making the comprehension of context richer and allowing better capture of finely-nuanced expressions of sentiment. The hybrid CNN-BiLSTM model can be enhanced by fusing contextual embeddings such as GLoVE and BERT, along with the introduction of attention mechanisms at multiple layers.

The research goal is to apply and compare machine learning models (SVM, NB, k-NN) and deep learning models (CNN, BiLSTM) for sentiment classification. These models will be compared based on standard performance measures to the proposed hybrid CNN-BiLSTM model, which is built with enhanced contextual accuracy.

2. RELATED LITERATURE

The evolution of sentiment analysis from rule-based lexicons to context-aware deep neural architectures, early systems would rely on curated dictionaries exemplified by SO-CAL, which uses an approach to polarity assignment based on manually annotated lexical entries. VADER extended this approach with valence-shifting heuristics which have proven effective across various forms of social media, including stock-market tweets [20]. Similarly Text Blob repackaged resources to support business intelligence [16]. Although lexicon–based methods are portable and require no training, they suffer from such well-known limitations such as sarcasm, negation scope, and domain-specific jargon, which have prompted a shift toward supervised approach. Comparative studies indicate that SVMs achieve higher accuracy than Random Forests on generic benchmarks datasets [23] .In domain-specific analysis, Naïve Bayes performs more superior than SVMs and k-NN baselines. Furthermore, methodological improvements to the like k-NN with advanced distance metrics, have even outperformed the classical configurations of SVM [19].

Climate change leads to rising temperature, melting of ice, extreme meteorological conditions, and sea-level rise [23]. Sentiment analysis of climate discourse provides a critical lens through which public emotion is sifted-and that emotion can vary from fear, scepticism, and optimism-to direct moral support or withdrawal from mitigation and adaptation policies. Analysts examine news articles and digital utterances to identify prevalent affective patterns [11], promote targeted messaging and behavior change [24] and to assess whether the collective psyche is ready, reluctant or resisting a set of interventions [4]. In addition to revealing media framing starting ranging from bias to sensationalism [26] sentiment analysis also designates how to carry out responsible journalism as well as advocacy approaches thus, the more they know about the media, the better equipped will be the ordinary citizens in recognizing disinformation and contributing positively to it [12]. From a methodological standpoint, climate-sentiment are grounded in communication, sociology, psychology, and data sciences, utilizing deep models such as BiLSTM to correctly contextualize and provide a longitudinal view of the shifting attitudes in relation to climatic events and negotiations [13]. Hence, sentiment analysis offers a multi-disciplinary framework for climate communication in real-time, the formulation of evidence-based policies, and ongoing academic research.

Traditional sentiment analysis methods struggle with contextual nuance and sequential meaning [27]. Lexicon-based methods rely on fixed dictionaries that fail to capture evolving language use, whereas classifiers such as SVM and Naïve Bayes are primarily based on shallow statistical features [28]. In climate-related tweet analysis, lexicon-based baselines such as SentiWordNet achieved only 36.1% accuracy on raw text, whereas hybrid approaches combining Text Blob with Logistic Regression attained an F1-score of 75.3% [29]. Because climate discourse embeds domain specific, context sensitive language, more expressive models are essential [20].

Recent advances in Deep Learning have considerably widened the performance gap. Convolutional Neural Networks (CNNs) capture local *n-gram* features, while Bidirectional Long Short-Term Memory (BiLSTM) networks model bidirectional contextual dependencies in restaurant review datasets. Integrating CNN and BiLSTM encoders results in enhanced representational capacity, as evidenced by a CNN–BiLSTM model evaluated on 2003 French newspaper articles, which consistently outperforms standalone CNN and BiLSTM baselines. The effectiveness of advanced embeddings and optimization strategies in sentiment analysis is evidenced by findings showing that Doc2Vec embeddings outperform Word2Vec for long-text representations, while the Adam optimizer achieves faster convergence than stochastic gradient descent.

Within climate change communication, headlines serve as interpretive frames shaping public perception of environmental risk and policy. Consequently, large-scale media analyses have adopted sentiment–topic modeling pipelines. [5] Processed global news to track attitudes toward a low-carbon economy, demonstrating sentiment shifts congruent with events such as forest fires and electric vehicle legislation. Using BERT topic and Latent Dirichlet Allocation, [15] charted climate discourse on social media blogs, whereas [18] integrated focus group insights with computational text mining to show German tabloids framing climate activists as security threats, thereby intensifying polarization. Cross-platform investigations corroborate the predominance of negative effect. One study analyzing a decade of Twitter, Reddit, and YouTube content using PMI-based polarity measures and NRCLex emotion detection identified fear, trust, and anticipation as dominant emotions surrounding discussions of climate change, air quality, and plastic pollution [3].

First, lexicon-based and classical machine learning approaches remain prevalent in multilingual climate datasets despite their limited capacity to apprehend nuanced or sarcastic language. Second, the temporal evolution of public sentiment vis-à-vis policy milestones and extreme weather events has not been comprehensively mapped. Third, the sentiment analysis is mainly in English, neglecting the inherently multilingual nature of climate discourse. Addressing these gaps will possibly necessitate hybrid architectures that integrate static lexical resources, domain-adapted embeddings, and contextual transformer models, augmented with attention mechanisms for interpretability. Such systems can leverage the bidirectional sensitivity of BiLSTMs, the local feature extraction of CNNs, and the semantic depth of transformers to decode short, information-dense headlines. Incorporating longitudinal and multilingual corpora into these advanced frameworks holds particular promise for enhancing the granularity, global reach, and policy relevance of climate sentiment monitoring.

BiLSTM networks overcome the uni-directional limits of early classifiers by reading text forward and backward, thereby encoding both preceding and succeeding context and yielding sharper semantic representation and higher accuracy [10].

Beyond architecture, three advances consistently lift performance. First, fusing local-pattern detectors with sequence encoders, CNN-BiLSTM hybrids capture both n-gram cues and long-range dependencies, improving sentiment predictions for short, information-dense texts such as headlines [8]. Second, attention layers direct the model to sentiment-laden tokens; an attention augmented BiLSTM on Chinese sticker comments outperformed plain BiLSTM, while an attention-assisted 1DCNN-BiLSTM reached R^2=0.9992 in a TMS-coil study. Third, the combination of static and contextual embeddings leads to richer semantic representations; BiLSTM–GloVe architectures improve headline classification performance, and BERT integrated with TCN, BiLSTM, and attention mechanisms achieves superior results in hotel-review sentiment analysis [6].

3. PROPOSED MODEL

The study proposes a hybrid GloVe-BERT + CNN BiLSTM + Attention architecture. It combines GloVe's static semantics with BERT's contextual embeddings, feeding this dual representation through parallel CNN and BiLSTM pathways, and applies an attention layer to capture the most salient tokens. By combining local lexical signals with long-range dependencies, the model delivers a better domain-specific sentiment classification. The present study extends the insights to climate headline sentiment, pairing GloVe and BERT embeddings with an attention enhanced CNN - BiLSTM to exploit domain vocabulary, highlight salient phrases, and achieve robust, context - sensitive classification. This model is planned to be utilized for climate - change headline analysis. The performance of the proposed hybrid model is compared with machine learning baselines and deep learning models.

4. METHODOLOGY

The study adopts a structured workflow in which climate change–related headlines are identified and collected from a publicly available Kaggle dataset, followed by data cleaning and tokenization to construct an analysis-ready dataset. Classical machine learning baselines (SVM, Naïve Bayes, k-NN) are trained next, followed by standalone deep learning models (CNN and BiLSTM). The final stage of the model integrates GloVe and BERT embeddings using an attention-based fusion strategy, followed by parallel CNN–BiLSTM branches augmented with residual connections and supplementary attention layers to model complex semantic patterns. All models are subsequently compared using performance metrics from the confusion matrix to determine the most effective approach for climate headline sentiment classification.

4.1 Phase 1: Data Preprocessing

An extract of the climate headline sentiment dataset is shown in Figure 1. It contains 1,024 climate change-related headlines from 15 high-GDP- based countries. The headlines have sentiment scores ranging from -1 to 1. The dataset also includes justification for the scores. Positive scores represent encouraging events (like advancements in climate initiatives), while negative scores point to troubling news (such as worsening crises). A score of zero indicates neutral, ambiguous, or mixed feelings.

	Headline	Link	Content	Sentiment	Justification
0	Australia's year ahead in climate and environment - ABC	https://www.abc.net.au/news/science/2024-01-23/what-2024-will-bring-for-australias-climate-and-environment/103362864	The year has barely started and extreme weather events are already in the headlines. Here are some more big environment issues to keep an eye on....	0	The headline is unclear about its direct impact on climate action.
1	Projections reveal the vulnerability of freshwater to climate change	https://news.griffith.edu.au/2024/01/09/projections-reveal-the-vulnerability-of-freshwater-to-climate-change/	â€œWater from groundwater, rivers and rainfall is undergoing disruption in its natural cycle due to climate and land use changes, which disrupt patterns and...	-0.5	The vulnerability of freshwater is concerning and highlights a negative impact, but it doesn't present a direct harmful action.
2	Record heat in 2023 worsened global droughts, floods and wildfires	https://www.anu.edu.au/news/all-news/record-heat-in-2023-worsened-global-droughts-floods-and-wildfires	2023 saw an increase in the frequency and intensity of rainfall events and river flooding, among other extreme weather events ...	-1	The headline describes worsening environmental impacts, which are harmful to the climate fight.
3	It's not just the total rainfall "â€œ why is eastern Australia ... - The Guardian	https://www.theguardian.com/australia-news/2024/jan/06/flash-flooding-cyclones-global-heating-australia-wild-weather	The number of storms in some regions is decreasing, but the intensity of rainfall is on the rise. Global heating seems to be playing a key role...	0	The headline is unclear about its stance on climate change or its environmental impact.
4	Expert Commentary: 2023 was the warmest year on record - CSIRO	https://www.csiro.au/en/news/all/news/2024/january/expert-commentary-2023-warmest-year-on-record	The European Union's Copernicus Climate Change Service (C3S) announced on 9 January 2024 that 2023 was officially the hottest year in recorded history...	-0.5	While highlighting the reality of climate change, it doesnâ€™t offer solutions or positive outcomes.
5	Climate migration "â€œ News, Research and Analysis - The Conversation	https://theconversation.com/topics/climate-migration-23371	Climate change is increasing the risks of extreme heat, floods and bushfires, meaning more people are having to consider moving home...	0	The headline is neutral, providing information without a clear positive or negative impact.
...					
1018	Russia strives to actively participate in global climate process	https://en.trend.az/business/green-economy/3960920.html	Putin noted that more than 20 percent of the world's forests are located on Russian territory, which, by absorbing greenhouse gases, helps slow down global	0.5	This suggests Russia is making an effort to participate in climate change efforts, which is a positive development.
1019	Boycott COP climate summit if Putin shows, Ukraine tells West	https://www.politico.eu/article/boycott-climate-cop-vladimir-putin-ukraine-vsevolod-chentsov/	The international community should shun the COP29 climate summit in Azerbaijan if Russian President Vladimir Putin attends, a top Ukrainian diplomat told...	-0.5	This is a politically charged statement and could harm the global climate dialogue by causing division and undermining
1020	How climate change will impact outdoor activities in the US | MIT News	https://news.mit.edu/2024/how-climate-change-will-impact-outdoor-activities-1022	Using the concept of â€œoutdoor days,"Â a study shows how global warming will affect people's ability to work or enjoy recreation outdoors. David Chandler | MIT...	0.5	This provides valuable insights into how climate change could affect daily life, contributing to public understanding and
1021	Climate and Environment - The New York Times	https://www.nytimes.com/section/climate	Earth's warming could trigger sweeping changes in the natural world that would be hard, if not impossible, to reverse....	0	This is a vague headline, and without more specific context, it is difficult to evaluate its impact on climate change.
1022	Climate Change News & Videos	https://abcnews.go.com/alerts/climatechange	Follow the latest Climate Change news stories and headlines. Get breaking news alerts when you download the ABC News App and subscribe to Climate Change...	0	Another general headline, offering information but no specific action or impact on climate change.
1023	Climate - BBC News	https://www.bbc.com/news/topics/cmj34zmwm1zt	Deadliest weather made worse by climate change - scientists. Human-caused climate change made recent extreme weather events more intense and more likely. new...	0	This is a general news update on climate change, without specific details or actions.
1024 rows x 6 columns					

Fig.1: *Extract of Dataset, Distribution of sentiments in the Data Set.*

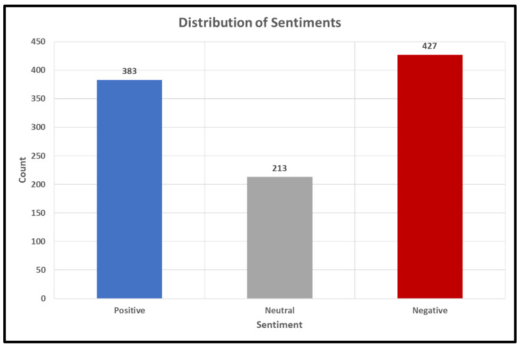

Fig.2: *Distribution of sentiments in the Dataset.*

From Figure 2, it is observed that most of the data is skewed toward Negative sentiment. A considerable number of entries are also positive, but fewer are Neutral. The figure that the issue of climate change tends to provoke stronger opinions rather than neutral ones.

For data preparation prior to applying Naïve Bayes, SVM, and KNN models, first the sentiment labels represented in numeric form were converted into textual categories. Numeric values representing negative sentiment were labeled as “Negative”, Numeric values representing positive sentiment were labeled as “Positive” and Numeric values representing zero were

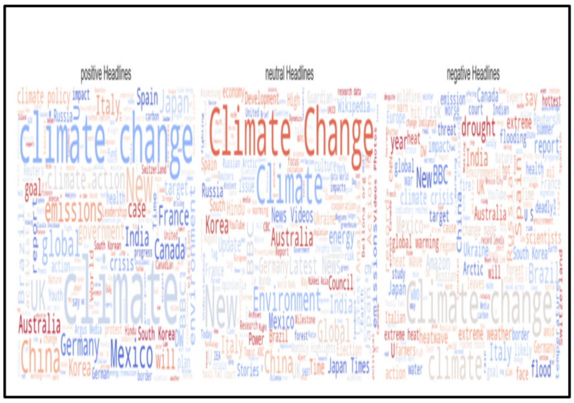

Fig.3: *Word cloud from the Dataset.*

labeled as “Neutral”. Second, a row having null sentiment has been deleted. Then, we have applied data cleaning, stop word removing, Part-of-Speech (POS) tagging, lemmatization and TF-IDF vectorization.

To prepare the input data for the CNN and BiLSTM models, we implemented a synthetic data generation strategy to augment our dataset. Sentiment labels were assigned as 2 for negative, 1 for positive, and 0 for neutral. Using synonym augmentation at a 30% rate, new headlines were created by replacing words with their synonyms, generating synthetic data up to 85% the size of the original dataset. These augmented headlines and labels were then combined with the original data, forming an expanded corpus. The final dataset was split 80/20 for training and testing. For model input, tokenization converted text into sequences of integers using a vocabulary size of 10,000, with out-of-vocabulary words replaced by a special token. Sequences were padded to a uniform length of 100 words to ensure compatibility with CNN and BiLSTM models. These preprocessing steps prepared the data for practical training on both original and augmented text.

4.2 Phase 2: Sentiment Analysis using Supervised Learning Algorithms

Supervised Learning is based on labeled training data. Sentiment analysis is a typical application of machine learning in NLP. It involves analyzing text data to determine the response expressed whether it’s positive, negative, neutral, or more varied emotions [30]. Deep learning is a part of machine learning which involves multi-layered neural networks to learn hierarchical feature representations of data. Deep learning has become one of the popular methods to carry out sentiment analysis, owing to its ability to identify complex patterns in the data of a textual nature [30]. For the improvement in model performance and combat overfitting, data augmentation has been implemented to increase dataset diversity without collecting new data [7], a learning rate scheduler has been employed to change the learning rate dynamically for better convergence and stability and early stopping has been used to stop the training once validation starts to decline in performance, hence ensuring good generalization.

4.2.1 Support Vector Machines (SVM) Classifier

For the proposed work, a SVM classifier has been used. It is a supervised learning method effective for tasks like sentiment analysis [1]. It identifies a hyper plane that maximally separates sentiment classes using the decision function:

$$f(x) = w \cdot x + b,$$

where w is the weight vector, x is the input feature vector, and b is the bias. The dataset was split into 80% training and 20% testing to evaluate performance. A linear kernel SVM was trained on TF-IDF-transformed text, leveraging its suitability for high-dimensional data to classify sentiments as positive, negative, or neutral.

4.2.2 Naïve Bayes (NB) Classifier

Another text classification model for our proposed work is NB. It is a probabilistic classifier which uses Bayes' theorem for text classification [1]. It estimates class probabilities using:

$$P(c|d) = P(d|c)P(c)/P(d),$$

Where P(c|d) is the probability of class c given document d. A Multinomial Naive Bayes (Multinomial NB) model [17] was trained on TF-IDF-transformed text and sentiment labels, leveraging word frequency distributions to effectively classify high-dimensional text data.

4.2.3 K-Nearest Neighbors (k-NN) Classifier

KNN is a simple, effective and lazy learner, requiring minimal training [1]. It classifies instances based on the majority sentiment of their k nearest neighbors using Euclidean distance:

$$EuclideanDistance = \sqrt{\sum (x_i - y_i)^2},$$

Where x and y are feature vectors. The dataset was split 80/20 for training and testing, and the model learned sentiment patterns by comparing feature similarities.

4.2.4 Convolutional Neural Networks (CNN)

The CNN is a specialized feed forward neural network which was inspired by the structure of the human visual cortex. The CNNs in NLP thus serve well to extract local features that are gathered through convolutional layers that pick spatially local patterns, such as the presence of sentiment cues which may occur at different locations inside the text.

The CNN architecture begins with an input layer accepting padded tokenized sequences, followed by a trainable embedding layer that maps terms onto a 300-dimensional space. The words are embedded into dense vectors from which the model can learn task-specific representations for each word during the training process. Multi-scale feature extraction is then applied using two parallel 1D convolutional layers, with kernel sizes of 3 and 5, and 128 filters each, equipped with ReLU activation. These two layers capture a wide variety of short and long n-gram patterns present in the input.

Max pooling layers are used on the outputs from convolutional layers to get significant features and diminish dimensionality [2]. We concatenate all the pooled outputs to a large feature vector that shows multi-scale patterns of the text. This vector is then sent to follow a dense layer with 64 units and ReLU activation with L2 regularization and dropout at a rate of 0.5 to avoid overfitting and promote generalization.

The output layer uses softmax activation to produce probability scores for classification. It is trained using the Adam optimizer with a learning rate of 0.001. A step decay scheduler reduces the learning rate by half every five epochs, and early stopping halts training if validation accuracy doesn't improve for five epochs. Training is conducted over a maximum of 30 epochs with a batch size of 32, ensuring robust evaluation after each epoch.

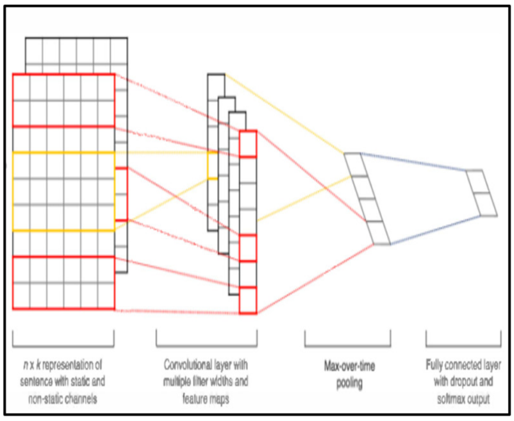

Fig.4: *Basic Architecture of CNN [21].*

4.2.5 Bidirectional Long Short-Term Memory (BiLSTM)

BiLSTM [25] extends an LSTM method where two LSTMs process the input, as illustrated in Figure 6. The working of the model starts with giving the sequence input to the first LSTM followed by the reverse of the input sequence given into the second LSTM. This model usually works for sentiment analysis as it grasps both past and future contexts in a sentence, rendering its understanding the sentiment expressed in the text. The model was developed through the following procedure.

In the proposed work, BiLSTM works begin with a trainable embedding layer converting each word of the headlines into a dense vector of 300 dimensions. This allows tweaking the word embeddings to improve sentiment classification during training. BiLSTM layer also contextual dependencies, mitigating overfitting, and setting both dropout and recurrent dropout set to 0.4. The dense layer is added after the dropout layer and has 32 units with ReLU activation. Its purpose is to pass the BiLSTM output and learn the more complex interactions of the features. Additional regularization is also gained from the dropout layer with rate 0.4. The final output layer uses softmax activation for classification.

With an initial learning rate of 0.001, BiLSTM is optimized by Adam optimizer and through a scheduler decreased to half every five epochs using a scheduler. Training is stopped early after seven epochs if no validation accuracy improvement is seen, making sentiment prediction testing more efficient. For the proposed work, the BiLSTM is trained for 30 epochs with a batch size of 32. All predictions on the test set is based the highest probability of the sentiment class.

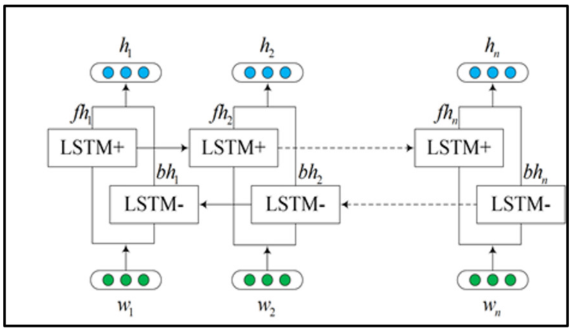

Fig.5: *Basic Structure of BiLSTM [31].*

4.3 Phase 3: Proposed Hybrid Model - Enhanced CNN-BiLSTM Algorithm

We have proposed a model of GloVe-BERT+CNN-BiLSTM+Attention, depicted in Figure 8, which includes various steps and layers. GloVe equips semantic representations of words using unsupervised learning, and BERT provides deep bidirectional contextual understanding, outperforming Word2Vec, GloVe, RNNs, and other contemporaneous models in multiple NLP tasks [1]. To augment model attention on specific parts of text, attention layers are included, which apply importance scores to input tokens in a dynamic manner. This architecture is tailored toward capturing local and global features for sentiment classification.

To enhance the robustness and generalization of the sentiment classification model, data augmentation was performed using three random transformations: synonym replacement (30% probability), word swapping, and word deletion (both at 10%). These were applied to generate synthetic headlines, expanding the training set by 85% of the original size—benefical for small or imbalanced datasets. The dataset was split 80:20 for training and testing. Three tokenization methods were used in parallel: Keras Tokenizer (top 15,000 words, sequences padded to 256), a separate tokenizer for GloVe (also padded to 256), and Hugging Face's BERT Tokenizer, producing input IDs and attention masks truncated/padded to 256 tokens. For embeddings, a GloVe embedding layer with pre-trained 300-dimensional vectors was applied with global max pooling for fixed-size outputs. BERT embeddings were also created using a trainable BERT-base model, extracting both [CLS] and mean-pooled representations, which were concatenated and passed through a GELU-activated dense layer to produce a 300-dimensional output.

The GloVe and BERT branch outputs are concatenated into a singular vector. This vector is subsequently restructured into a sequence and undergoes a self-attention mechanism where attention weights are learned through softmax. The representation with attention is then compressed to a form in which the prominent combined features stand out. This is succeeded by a multi-branch CNN-BiLSTM layer which extracts diverse features: an extremely detailed 1D convolutional layer with kernel size 1, and subsequently batch normalization, is used. Alongside it, a parallel BiLSTM with 192 units is used to capture the model's forward and backward dependencies. Global max pooling is performed on both branches to yield a multi-scale feature composition. The output received from both branches is subsequently merged, undergoing normalization followed by a residual connection, before passing through a second attention module that enhances the contextual representation further by focusing on the most informative learned features.

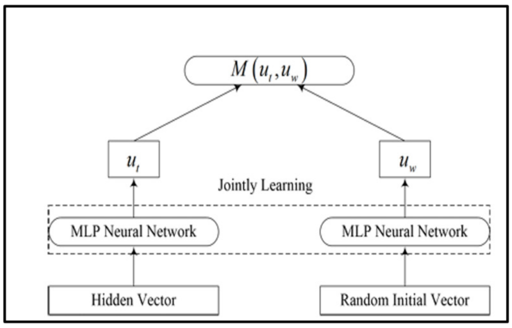

Fig.6: *Basic Self-Attention Mechanism [31].*

A Dropout layer with a rate of 0.3 is applied initially to prevent overfitting. Then a dense layer with GELU activation, L2 regularization for strong classification, and a Softmax output layer for sentiment class probability generation are implemented. The hybrid model is trained using the Adam optimizer, a Cosine Annealing Learning Rate Scheduler and Warm up, and Early Stopping for best convergence. The hybrid model is trained for 50 epochs with a batch size of 8.

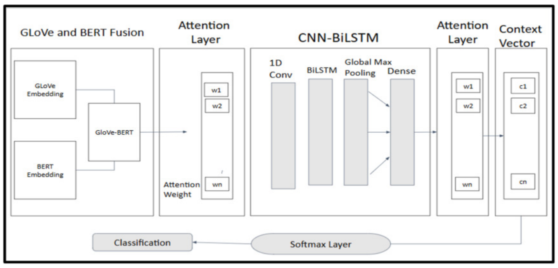

Fig.7: *Basic Framework of the proposed GLoVe-BERT+CNN-BiLSTM+Attention Model.*

5. EXPERIMENT AND RESULTS

5.1 Performance Metrics and Model Comparison

The study evaluates five supervised learning miodels against the proposed one using four confusion matrix metrics [2]: Accuracy, Precision, Recall, and F1-score, calculated based on True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). The sentiment classes are labelled as Neutral (0), Positive (1), and Negative (2). TP occurs when the predicted class matches the true class; FP when the model predicts an incorrect class; TN when the model correctly identifies that a sample does not belong to a class; and FN when the model fails to predict the actual class [25].

Accuracy is defined as the ratio of the correctly predicted sentiment class to the total number of predicted sentiment classes, whose equation is given by,

$$Accuracy = \frac{TP_0+TP_1+TP_2}{TP_0+TP_1+TP_2+FN_0+FN_1+FN_2+FP_0+FP_1+FP_2}$$

Precision is defined as the ratio of correctly predicted sentiment to the total number of expected sentiment in any class, whose equation is,

$$Precision_i = \frac{TP_i}{TP_i + FP_i}$$

Where i = 0, 1, 2

Recall is defined as the ratio correctly predicted sentiment to the total number of actual sentiment in any class, whose equation is given as,

$$Recall_i = \frac{TP_i}{TP_i + FN_i}$$

Where i = 0, 1, 2

F1 score is defined as the harmonic mean of precision and recall, whose equation is given as,

$$F1score_i = 2 * \frac{Precision_i * Recall_i}{Precision_i + Recall_i}$$

Where i = 0, 1, 2

The performance metrics obtained by each of the five supervised learning models are given in Table 1.

Table 1: *Comparison of Performance Metrics.*

Models	Accuracy	Precision	Recall	F1 Score
K-NN	58.05%	58.44%	58.05%	57.78%
SVM	62.44%	65.11%	62.44%	61.54%
NB	63.90%	64.59%	63.90%	63.25%
CNN	78.63%	79.92%	78.63%	78.65%
BiLSTM	78.36%	78.24%	78.36%	78.27%
GLoVe-BERT+ CNN-BiLSTM+ Attention	**80.47%**	**80.67%**	**80.47%**	**80.41%**

In Table 1, performance metrics comparison, the traditional models yielded relatively lower performance, with k-NN performing the weakest (Accuracy: 58.05%) and Naive Bayes slightly outperforming SVM. The reason for the NB classifier showing highest accuracy is its simplicity and feature independence which helps it perform well on text data (Munawaroh & Alamsyah, 2022). It has proven to be reliable due to its computational efficiency, achieving faster training and inference compared to other models, while also demonstrating strong performance on small datasets.

CNN and BiLSTM achieved significantly higher metrics, with CNN slightly surpassing BiLSTM across all evaluation criteria, as shown in Figure 9 and Figure 10. Figure 9 of the confusion matrix of CNN illustrates that the model achieves high accuracy overall, with most samples correctly classified, though some misclassifications occur primarily between classes 0 and 1. Figure 11 of the confusion matrix of BiLSTM shows that the model performs well overall, with most samples correctly classified, though some confusion is observed between classes 0 and 1, and between classes 1 and 2. CNNs excel at identifying key n-gram features, thus its accuracy reflects its ability to learn discriminative local features efficiently. BiLSTM can learn dependencies across distant words which helps maintain semantic coherence, improving performance where CNN may miss such relationships [25]. Although CNN slightly edges out BiLSTM in your results, both models complement each other well.

The proposed hybrid model, GloVe-BERT+CNN-BiLSTM+Attention outperformed all other models, achieving the highest Accuracy (80.47%) and F1 Score (80.41%), as shown in Figure 12. Figure 13 shows that the confusion matrix of the proposed model demonstrates strong overall performance, with most samples correctly classified, though minor misclassifications occur between classes 0 and 1, and be-

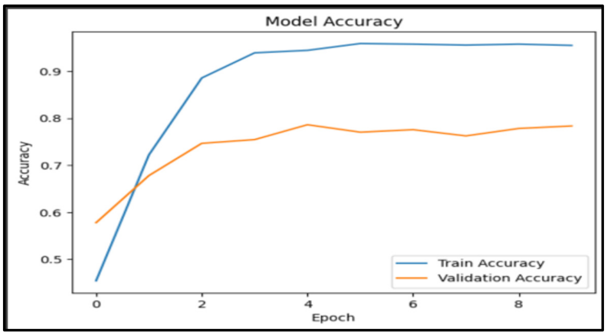

Fig.8: *Model Accuracy of CNN.*

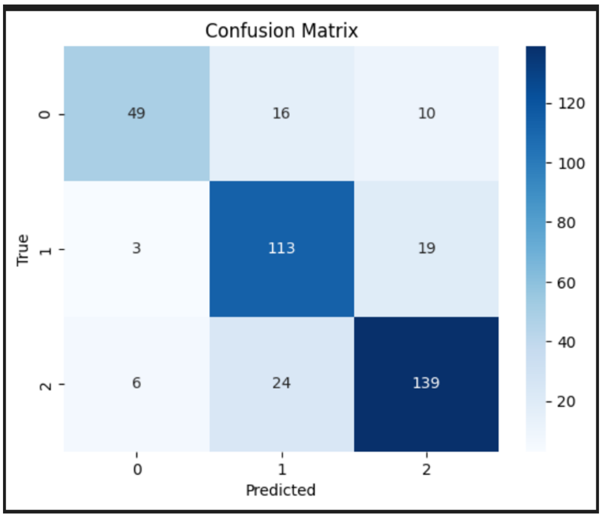

Fig.9: *Confusion Matrix of CNN.*

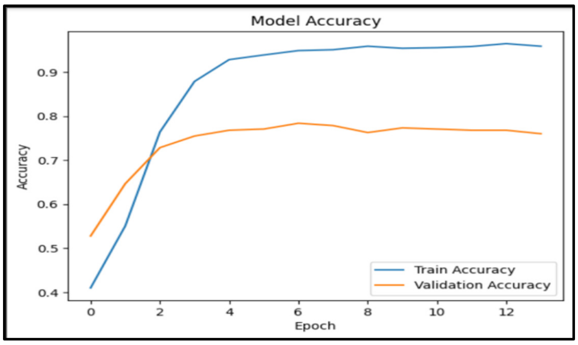

Fig.10: *Model Accuracy of BiLSTM.*

tween classes 1 and 2. This performance improvement highlights the strength of combining static and contextual word embeddings (GloVe and BERT) with deep sequence and feature extraction layers (CNN, BiLSTM) and an attention mechanism. The hybrid architecture effectively captures both local patterns and long-range dependencies, making it highly suitable for domain-specific text classification tasks such as climate change headline categorization.

The results presented in Table 1 show effectiveness of various models for sentiment classification of climate change-related headlines. We could find the traditional machine learning models show the accuracies

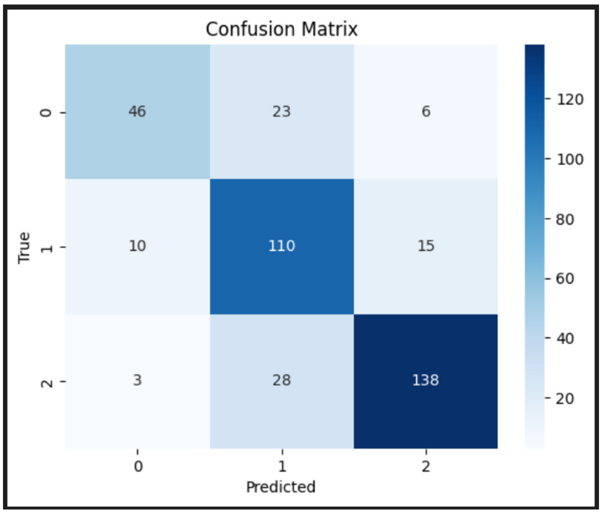

Fig.11: *Confusion Matrix of BiLSTM.*

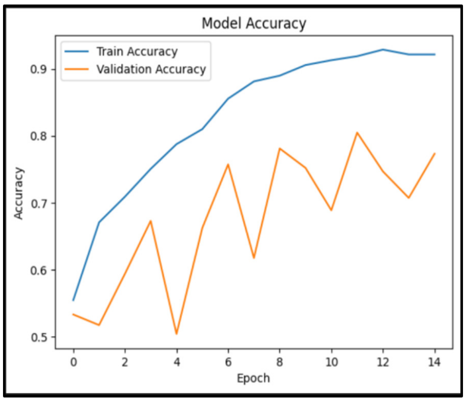

Fig.12: *Model Accuracy of GLoVe-BERT+CNN-BiLSTM+Attention.*

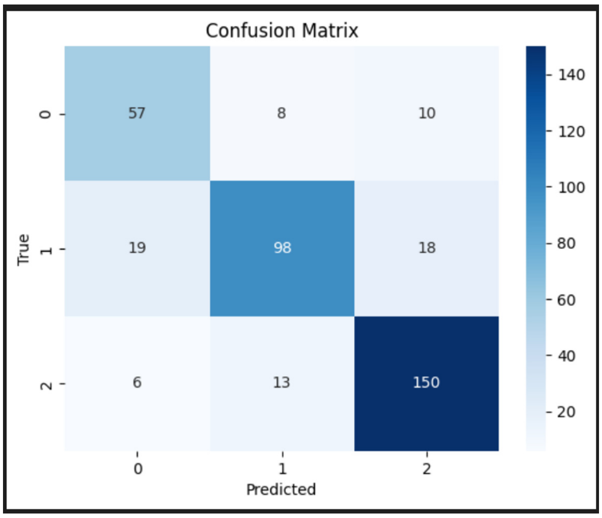

Fig.13: *Confusion Matrix of GLoVe-BERT+CNN-BiLSTM+Attention.*

range from 58.05% to 63.90%. Among them, Naive Bayes performed the best, likely due to its probabilistic nature and robustness in handling text data, at the same time k-NN lagged, possibly due to its sensitivity to high-dimensional sparse data typical in textual input.

In contrast, CNN and BiLSTM networks show significantly higher accuracies above 78%. CNN captured local n-gram features effectively, while BiLSTM leveraged contextual dependencies in both forward and backward directions. This demonstrates that deep neural architectures are better at modelling complex patterns in domain-specific textual data.

The proposed hybrid model, the combination GLoVe-BERT+CNN-BiLSTM+Attention, achieved the overall highest performance based on the metrics of accuracy of 80.47%, precision of 80.67%, recall of 80.47%, and F1 score of 80.41%. This result brings out the advantage of integrating both static (GloVe) and contextual (BERT) embeddings, which provide complementary semantic representations. The use of CNN and BiLSTM in the hybrid model captures both local and sequential features. The attention mechanism focuses on the most informative parts and gives the significance of the input.

6. HYPOTHESIS TESTING

Hypothesis testing is used to verify statistically whether one model is better than another one. From the accuracies of model in table1, the CNN and BiLSTM are chosen for hypothesis testing with the Hybrid model as they both performs best compare to classical machine learning model. CNN, BiLSTM, and the hybrid model are neural architectures designed to learn contextual and sequential patterns from the text. For performing paired t-test hypothesis testing, fivefold cross validation are performed on CNN vs Hybrid model and BiLSTM vs Hybrid model to determine statistical significance difference between these models.

6.1 CNN versus Hybrid Model Hypothesis Testing

To ensure a robust comparison, both CNN and hybrid models were evaluated using 5-fold cross - validation.

To determine whether the hybrid model provides a significant difference with the CNN model, a paired t-test hypothesis testing is performed, the hypothesis is stated as

Null Hypothesis:

$H_0 : \mu = 0$, (Both models are same)

Alternate Hypothesis:

$H_1 : \mu > 0$, (Hybrid model performs significantly better)

The paired t-test examines the mean difference accuracy between two models is statistically significant. The test statistic is given by

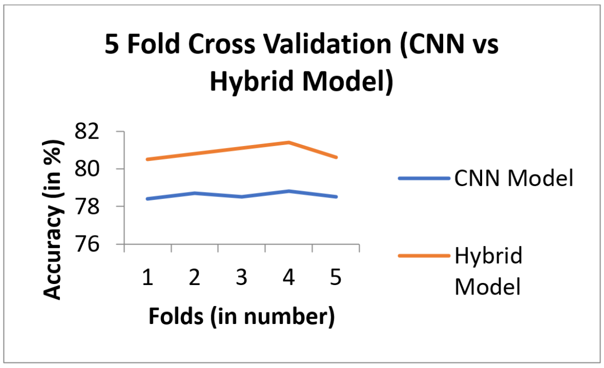

Fig.14: Five-fold Cross Validation(CNN vs Hybrid Model).

$$t = \frac{\bar{d}}{sd\sqrt{n}}$$

Based on the statistic, the hypothesis statement is evaluated.

The computed t-test statistic is 18.79 (2.98 conducted on the normalized accuracy values) for df=4 and the two tailed t-critical value at ($\alpha = .05$) is 2.776 where $18.07 > 2.776$, t stating the null hypothesis is rejected. Under the t distribution, the corresponding p value at 5% significance level ($\alpha = .05$) is .00003. Since the p value $< .05$, the null hypothesis is again rejected stating that the difference between the model is statistically significant.

6.2 BiLSTM versus Hybrid Model Hypothesis Testing

To perform hypothesis testing, paired t-test is examined. For the evaluation, the 5-fold cross validation

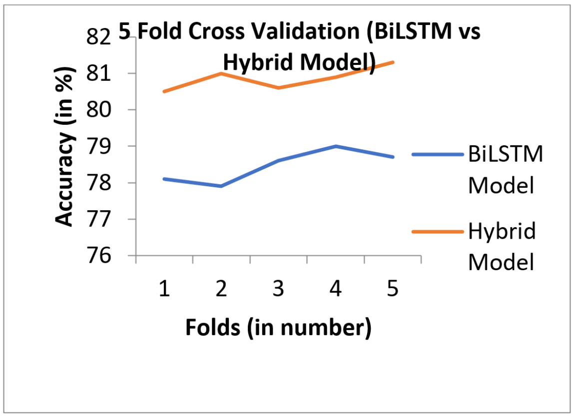

Fig.15: Five-fold Cross Validation (BiLSTM vs Hybrid Model).

The computed paired t-test statistic is 11.06 for df=4 and the two tailed t-critical value at ($\alpha = .05$) is 2.776 where $11.06 > 2.776$ stating the null hypothesis is rejected. Under the t distribution, the corresponding p value at 5% significance level ($\alpha = .05$), is .00003 < .05. Since the p value < .05, again the null hypothesis is rejected stating that the difference between the model is statistically significant. The hybrid model provides a substantial enhancement in sentiment prediction performance.

7. CONCLUSIONS

The proposed work GloVe-BERT fusion in CNN-BiLSTM-Attention architecture provides a robust solution for climate-headline sentiment analysis. By fusing Glove's static semantics with BERT's contextual embeddings, the model detects subtle emotional cues that elude conventional approaches. Parallel CNN-BiLSTM branches learn both local n-grams and long-range dependencies, and stacked attention layers focus the network on sentiment-carrying tokens and reveal the lexical triggers underlying each prediction, thus improving interpretability. Synonym-based data augmentation further broadens the model's coverage without distorting domain semantics, producing precision, recall, and F1-scores for positive, negative, and neutral classes. Tested on 1,023 climate-related headlines annotated on a three-point polarity scale, the hybrid system achieves 80.47% accuracy, outperforming classical baselines (SVM, Naïve Bayes, k-NN) and single-branch deep networks (CNN 78.63%, BiLSTM 78.36%). It provides a strong distinction between positive and neutral headlines, where vocabulary overlap typically impedes classification. Taken together, these results confirm that transformer-enhanced hybrid architectures, enriched with complementary embeddings and attention mechanisms, can deliver state-of-the-art accuracy and transparency for domain-specific sentiment tasks, providing a scalable tool for monitoring public climate discourse. Based on the Hypothesis testing, the hybrid model performs better than CNN and BiLSTM model is statistically verified.

Future work should broaden sentiment analysis of climate-change headlines beyond English to capture multilingual, cross-cultural perspectives; track sentiment over time to link shifts to major environmental events and policy milestones; test lightweight GRU-based models potentially surpassing CNN-BiLSTM accuracy when paired with fast optimisation methods such as swarm intelligence; and add interpretability layers that visualise token-level contributions, improving model transparency and trust.

AUTHOR CONTRIBUTIONS

Conceptualization, Y.I.; methodology, Y.I.; software, Y.I..; validation, G.L.; formal analysis G.L.; investigation, G.L.; data curation, Y.I.; writing—original draft preparation, Y.I.; writing—review and editing, G.L.; visualization, Y.I.; supervision, G.L.; All authors have read and agreed to the published version of the manuscript.

References

[1] S. Adhikari, R. Kaushik, A. J. Obaid, S. Jeyalaksshmi, D. Balaganesh and F. H. Hanoon, "YouTube sentimental analysis using a combined approach of KNN and K-means clustering algorithm," in *Proceedings of 3rd International Conference on Mathematical Modeling and Computational Science*, Singapore: Springer Nature Singapore, pp. 37–50, 2023.

[2] M. Alruily, "Sentiment analysis for predicting stress among workers and classification utilizing CNN: Unveiling the mechanism," *Alexandria Engineering Journal*, vol. 81, pp. 360–370, 2023.

[3] D. Amangeldi, A. Usmanova and P. Shamoi, "Understanding Environmental Posts: Sentiment and Emotion Analysis of Social Media Data," in *IEEE Access*, vol. 12, pp. 33504-33523, 2024.

[4] R. J. Brulle, J. Carmichael and J. C. Jenkins, "Shifting public opinion on climate change: An empirical assessment of factors influencing concern over climate change in the U.S., 2002–2010," *Climatic Change*, vol. 114, no. 2, pp. 169–188, Sep. 2012.

[5] C. Bucur, B. Tudorica, J. V. Andrei, D. Dusmanescu, D. Paraschiv and C. Teodor, "Sentiment analysis of global news on environmental issues: Insights into public perception and its impact on low-carbon economy transition," *Frontiers in Environmental Science*, vol. 12, p. 1360304, Jun. 2024.

[6] D. Chi, T. Huang, Z. Jia and S. Zhang, "Research on sentiment analysis of hotel review text based on BERT-TCN-BiLSTM-attention model," *Array*, vol. 25, p. 100378, Mar. 2025.

[7] E. D. Cubuk, B. Zoph, D. Mané, V. Vasudevan and Q. V. Le, "AutoAugment: Learning Augmentation Strategies From Data," *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, pp. 113-123, 2019.

[8] M. K. Daradkeh, "A hybrid data analytics framework with sentiment convergence and multi-feature fusion for stock trend prediction," *Electronics*, vol. 11, no. 2, p. 11020250, Jan. 2022.

[9] J. Devlin, M.-W. Chang, K. Lee and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proceedings of NAACL-HLT 2019*, pp. 4171–4186, 2019.

[10] S. Divya, S. Kiruthika, A. N. Anton and S. Padmavathi, "Segmentation, tracking and feature extraction for Indian sign language recognition," *International Journal on Computational Science & Applications*, vol. 4, no. 2, pp. 57–72, 2014.

[11] D. N. Dwivedi and G. Mahanty, "Mental health in messages: Unravelling emotional patterns

through advanced text analysis," in *Using Machine Learning to Detect Emotions and Predict Human Psychology*, Hershey, PA, USA: IGI Global, pp. 187–208, 2024.

[12] L. Edwards *et al.*, "Rapid evidence assessment on online misinformation and media literacy," Final Report, 2021.

[13] M. El Barachi, M. AlKhatib, S. Mathew and F. Oroumchian, "A novel sentiment analysis framework for monitoring the evolving public opinion in real-time: Case study on climate change," *Journal of Cleaner Production*, vol. 312, p. 127820, Aug. 2021.

[14] M. Geetha, N. Aloysius, D. A. Somasundaran, A. Raghunath and P. Nedungadi, "Toward Real-Time Recognition of Continuous Indian Sign Language: A Multi-Modal Approach Using RGB and Pose," in *IEEE Access*, vol. 13, pp. 60270-60283, 2025.

[15] T. Gokcimen and B. Das, "Exploring climate change discourse on social media and blogs using a topic modeling analysis," *Heliyon*, vol. 10, no. 11, p. e32464, 2024.

[16] J. P. Gujjar and H. P. Kumar, "Sentiment analysis: Text blob for decision making," *International Journal of Scientific Research & Engineering Trends*, vol. 7, no. 2, pp. 1097–1099, 2021.

[17] N. S. Harzevili and S. H. Alizadeh, "Mixture of latent multinomial naive Bayes classifier," *Applied Soft Computing*, vol. 69, pp. 516–527, 2018.

[18] J. Heide, "Polarizing social figures? Climate activists in German media and popular discourse," *European Societies*, vol. 27, no. 5, pp. 959–991, 2025.

[19] S. Hota and S. Pathak, "KNN classifier based approach for multi-class sentiment analysis of Twitter data," *International Journal of Engineering & Technology*, vol. 7, no. 3, pp. 1372–1375, 2018.

[20] C. Hutto and E. Gilbert, "VADER: A parsimonious rule-based model for sentiment analysis of social media text," *Proc. Int. Conf. The Eighth International AAAI Conference on Weblogs and Social Media (ICWSM-14)*, vol. 8, no. 1, pp. 216–225, 2014.

[21] T. Islam *et al.*, "Lexicon and deep learning-based approaches in sentiment analysis on short texts," *Journal of Computer and Communications*, vol. 12, no. 1, pp. 11–34, 2024.

[22] T. R. Karl and K. E. Trenberth, "Modern global climate change," *Science*, vol. 302, no. 5651, pp. 1719–1723, 2003.

[23] T. A. Khan, R. Sadiq, Z. Shahid, M. M. Alam and M. B. M. Su'ud, "Sentiment analysis using support vector machine and random forest," *Journal of Informatics and Web Engineering*, vol. 3, no. 1, pp. 67–75, 2024.

[24] F. S. Khatibi, A. Dedekorkut-Howes, M. Howes and E. Torabi, "Can public awareness, knowledge and engagement improve climate change adaptation policies?," *Discover Sustainability*, vol. 2, no. 18, pp. 1–24, 2021.

[25] S. Khruahong, O. Surinta and S. C. Lam, "Sentiment analysis of local tourism in Thailand from YouTube comments using BiLSTM," in *Proc. Int. Conf. Multi-disciplinary Trends in Artificial Intelligence*, Switzerland: Springer, pp. 169–177, 2022.

[26] S. O'Neill *et al.*, "Dominant frames in legacy and social media coverage of the IPCC Fifth Assessment Report," *Nature Climate Change*, vol. 5, no. 4, pp. 380–385, 2015.

[27] A. Razaque *et al.*, "State-of-art review of information diffusion models and their impact on social network vulnerabilities," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 1, pp. 1275–1294, 2022.

[28] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Foundations and Trends in Information Retrieval*, vol. 2, no. 1–2, pp. 1–135, 2008.

[29] N. M. Sham and A. Mohamed, "Climate change sentiment analysis using lexicon, machine learning and hybrid approaches," *Sustainability*, vol. 14, no. 8, p. su14084723, 2022.

[30] K. L. Tan, C. P. Lee, and K. M. Lim, "A survey of sentiment analysis: Approaches, datasets, and future research," *Applied Sciences*, vol. 13, no. 7, p. app13074550, 2023.

[31] J. Xie, B. Chen, X. Gu, F. Liang and X. Xu, "Self-Attention-Based BiLSTM Model for Short Text Fine-Grained Sentiment Classification," in *IEEE Access*, vol. 7, pp. 180558-180570, 2019.

Yashaswini Iyer earned her Master's degree in Applied Statistics and Data Analytics from Amrita Vishwa Vidyapeetham, Coimbatore. She has research experience in the area of hybrid deep learning models She has professional experience as a Data Analyst in the e-commerce sector where she primarily focused on ERP coordination and analytical workflows. Her research interests include application-based analysis in healthcare and environment, applied machine learning, and data-centric problem solving.

Gnanaprasanambikai L working as Assistant Professor, Department of Mathematics in the School of Physical Sciences, Amrita Vishwa Vidyapeetham, Coimbatore. Her research interests includes Machine learning, Soft computing, Deep learning. She has published research findings in reputed international journals and conferences and is actively involved in teaching undergraduate and post graduate courses.