TRANSFORMER AND BILSTM-BASED SENTIMENT ANALYSIS FOR EMOTIONS IN SHORT STORIES

V. Kowsalya[1,*] [0000-0002-5109-7311], C. Divya[2][0000-0001-7364-0269]

[1] Ph.D. Research Scholar, Centre for Information Technology and Engineering, Manonmaniam Sundaranar University,Tirunelveli, TamilNadu, India.

[2]Assistant Professor, Centre for Information Technology and Engineering, Manonmaniam Sundaranar University, **Tirunelveli**, TamilNadu, India.

Abstract

Sentiment analysis of emotions in short stories is essential to determining how readers are affected by the story and how well it communicates its intended themes. However, because narrative texts are complex, current approaches frequently have difficulty adequately capturing the subtleties of mood. In this study, we offer a new model for sentiment analysis in short tales, based on Transformer with attention concept and BiLSTM. By exploiting the beneficial qualities of Transformer architecture for capturing global dependencies and BiLSTM for collecting local context, this model tackles the drawbacks of older methods. Combining these two architectures allows our model to provide an extensive analysis of sentiment variance while accurately capturing the story's emotional direction. A considerable improvement over previous methods was achieved with a valence score of around 76% and an arousal score of around 62% when our model's performance was assessed using common metrics like arousal and valence percentages. Our study offers insights into the emotional landscapes of short tales that may impact a variety of sectors, including education, movie directors, other script writers and mental health. It goes beyond academic research to contain significant practical consequences for society. For example, content producers might modify their narratives to elicit particular emotional reactions in their audience, and educators can use our approach to assess the emotional impact of instructional stories on students. The proposed Transformer and BiLSTM-based sentiment analysis methodology opens up new avenues for study on sentiment-aware narrative development and analysis, paving the way for future research in sentiment-aware narrative generation and analysis.

Keywords: Transformer, BiLSTM, Attention, Sentiment Analysis, Emotional Stories

1 Introduction

Sentiment analysis has emerged as a vital tool in understanding and interpreting human emotions expressed in textual narratives. This work explores sentiment analysis utilizing sophisticated Transformer and BiLSTM-based models on short texts. Although the main goal is not to create tales, the main goal is to correctly forecast the emotions that are present in the stories. The intricacies and subtleties of human emotions in text have long been a challenge for conventional approaches. But the emergence of deep learning models—Transformers and

BiLSTMs, in particular—has completely changed the area by providing improved capacities for identifying emotional nuances and contextual relationships. Because of its self-attention mechanism, the Transformer model is excellent at recognizing complex links between words in a tale and offers an extensive understanding of sentiment stream. Simultaneously, BiLSTMs refine sentiment predictions further by taking into account both past and future circumstances thanks to its bidirectional processing. The study intends to establish a new standard for the precision and dependability of sentiment analysis in short tales by utilizing these potent models, opening the door to more in-depth understanding of the emotional terrain of literature.

Sentiment Analysis on Stories

Sentiment analysis of stories involves evaluating and recognizing the emotional tone that stories express. Sentiment analysis may be used to automatically categorize the many emotions that are expressed in a text, including surprise, rage, grief, and joy. By using this method, content creators may better customize their tales to elicit desired emotional reactions by improving their understanding of audience engagement and emotional impact.

Benefits:

- Analyzes the emotions and reactions of the reader.
- Provides for the customization of content according to viewer mood.
- Helps to enhance stories' emotional appeal.

The research of sentiment analysis in tales has drawn a lot of scholarly attention. To tackle the challenging problem of identifying and forecasting emotions in narrative texts, a variety of approaches and datasets have been used. Using a Thai sentiment resource built from SenticNet2, a Sentiment Text Tagging System (STTS) [1] for Thai children's stories achieves improved accuracy with a support vector machine. The primary limitation of the system is its dependence on bi-directional translation, which can result in inaccuracies and decreased precision. used an ELECTRA model based on Transformers to introduce continuous valence and arousal annotations for a collection of children's stories [2]. Higher valence and arousal Concordance Correlation Coefficients (CCCs) were attained by their method. The intricacy of the model and the requirement for considerable fine-tuning, however, pose formidable obstacles.

The Story Cloze Task [3] employs lexical coherence and sentiment analysis in tandem. They illustrated the significance of sentiment in story comprehension, but they were unable to surpass baseline findings in accuracy. A sentiment analyzer and a sentimental generator are included in a framework for managing [4] fine-grained sentiment intensity in tale ending creation. Their approach faltered despite improvements due to the absence of explicit sentiment control during creation and fine-grained sentiment labeling. Applying explainable AI (XAI) to diachronic sentiment analysis and story production using GPT-4 [5]. Their greybox ensemble system coupled interpretability and efficiency of high-performance blackbox models with human-in-the-loop supervision that might be resource-intensive for successful implementation.

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

A creative method using sentiment analysis to raise the evolutionary suspense generator's fitness function [6]. They significantly improved suspense generation by using a computer method that modified narrative sequences depending on sentiment analysis results. Nevertheless, the study did not investigate more extensive uses of sentiment analysis in storytelling, instead concentrating on altering narrative situations. Sentiment analysis research in inflectional languages such as Turkish is limited. To address this, a technique for creating a sentiment dictionary and using dictionary-based sentiment analysis on Turkish texts is presented [7]. Although the project helped to build Turkish sentiment analysis tools, it was mainly concerned with creating dictionaries and did not go into more complex sentiment analysis methods or how they may be utilized to create narratives.

The bias in text generation models according to country [8], with special attention to the prejudice in GPT-2 against less internet-savvy nations. They proposed adversarial triggering as a technique for debiasing and provided evidence of its efficacy in reducing bias in created tales. Nevertheless, sentiment research methods and their usage in story creation were not included in the study; instead, it concentrated mostly on bias analysis and mitigation procedures. An in-depth examination of cutting-edge methods for sentiment-controlled text production was conducted [9], classifying approaches according to the qualities regulated and the control implementation. The study concentrated mostly on method categorization and comparison, with insufficient investigation of particular datasets or story production models, even though it offered insightful information about sentiment-controlled text generation strategies. Techniques for modeling story dynamics and structure include sentiment analysis and network-based modeling [10]. They illustrated the promise of network-based approaches in revealing the intricate nature of narrative dynamics by applying these methods to Victor Hugo's Les Misérables. However, the study did not thoroughly examine sentiment analysis techniques or their use in story development, instead concentrating exclusively on narrative modeling.

Transformer based Sentiment Prediction

Transformer-based models, like GPT or BERT, use the self-attention process to comprehend dependencies and context across the text. These models are very good at sentiment analysis because they are good at predicting sentiment by identifying subtle linguistic nuances and expressions.

Benefits:

- Excellent accuracy in identifying subtleties of complicated emotion.
- The capacity to analyze text with long-range dependencies.
- Due to parallelization, training and inference are done efficiently.

The Transformer architecture [11], which relied only on attention processes, transformed sequence transduction models. These models improve parallelizability and decrease training time by doing away with the requirement for recurrence and convolutions, while still producing state-

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

of-the-art outcomes in machine translation applications. However, the study's relevance to other areas is limited as it focuses solely on machine translation jobs. A hybrid model for sentiment analysis called RoBERTa-LSTM [12] integrates Transformer and Long Short-Term Memory (LSTM). The proposed model seeks to enhance sentiment analysis performance by utilizing data augmentation approaches and resolving sequence models' drawbacks. However, it can still run into issues with computational complexity and a diversity of data. An approach to generating tale endings that incorporates emotive trends to improve sentiment consistency in generated endings is the Gated Mechanism based Transformer Network (GMTF) [13]. The technique tries to enhance the quality of tale endings by fusing contextual information with sentiment analysis. It could, however, run into difficulties managing various story situations and conveying subtle emotional subtleties.

Technologies to convey cultural heritage using automated storytelling [14], with an emphasis on chatbots and big language models as models. While these technologies provide fresh approaches to interacting with audiences about cultural heritage, issues with story coherence and audience engagement in practical contexts may arise. About text-based sentiment analysis [15] in the financial industry, underlining the developments in natural language processing as well as the possibilities presented by transformer architecture and multimodal analysis. The work suggests several avenues for improving sentiment analysis in finance, including the use of transformer architecture and multimodal classifiers. However, more investigation is needed to address issues unique to the financial sector and data biases in sentiment analysis. Human attitude on social media data and the features of New York City's urban buildings are related [16]. To evaluate twitter messages on a social media network, they use transformer deep learning models, notably RoBERTa, and sentiment analysis algorithms. In order to document the change in building occupancy throughout the pandemic's first year, the research concentrates on the phrase "Stay-at-Home." Although the study clarifies the emotional differences associated with the features of urban buildings, socioeconomic elements influencing sentiment dynamics may not be fully covered.

A paradigm for identifying narrative structures in brief personal tales is called M-sense [17]. They provide the STORIES dataset, which has narrative components like climax and resolution manually annotated. Their computational model mixes contextual semantic embeddings with mental state information of the protagonist by using a pre-trained model based on social commonsense knowledge. Though the model takes a novel approach, its efficacy could be constrained by the intricacy of combining mental state representations with contextual semantic embeddings. Text transformers for sentiment classification in reviews based on Steam game reviews, utilizing TF-IDF, BERT, and SBERT [18] methods. Although the study provides valuable insights into the efficacy of various sentiment analysis techniques, its narrow focus on video game reviews may restrict its generalizability to other settings. Apply VADER and BERT models to analyze sentiment in Omicron vaccine discussions on Twitter [19]. Although the study's reliance on Twitter data may have overlooked opinions expressed through other media, it nonetheless offers insightful information about public opinion at a pivotal moment of the epidemic. Examine automated analysis [20] of family coaches' fragmented Dutch diaries using machine learning methods. With pre-trained language models such as BERTje, they get encouraging results in

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

sentiment analysis and essential social rights identification. The study's narrow emphasis on a particular dataset and language, however, can restrict how broadly its conclusions can be applied to other situations and tongues.

Predicting the sentiment using LSTM

An effective kind of recurrent neural network (RNN) for handling sequential data is the Long Short-Term Memory (LSTM) network. Because they can handle context across sentences and maintain long-term dependencies, LSTMs are resilient for sentiment analysis in narrative texts and can predict sentiment successfully.

Benefits:

- Ability to pick up knowledge from text's temporal and sequential patterns.
- Efficient at maintaining context across extended text passages.
- Performs better than conventional RNNs in handling vanishing gradient concerns.

RoBERTa-LSTM [21] is a hybrid model for sentiment analysis that combines the benefits of recurrent and Transformer neural network architectures. Their work tackles issues like lexical variety and unbalanced datasets and emphasizes the value of sentiment analysis across a range of fields. LSTM models' sequential structure, however, may result in slower computing speeds and longer execution periods. A divide-and-conquer strategy for sentiment analysis [22], with an emphasis on CNN and BiLSTM-CRF models for phrase type categorization. Their approach enhances sentiment analysis on many datasets by classifying phrases according to their complexity and sentiment aims. However, long-distance relationships in text may be difficult for conventional LSTM models to capture, which reduces their usefulness. Rhetorical structure theory is included into a hierarchical deep neural network [23] method to sentiment analysis in order to capture discourse-level sentiment. By taking into account the structural arrangement of texts, their methodology tackles the drawbacks of bag-of-words methods. However, training stability and convergence may be hampered by conventional LSTM models' propensity for problems like bursting and disappearing gradients. A hybrid binary classification framework for sentiment analysis of children's literature that combines the Random Forest algorithm and LSTM [24]. Their method reduces the computational cost and overfitting problems that come with using conventional LSTM models. However, LSTM models are prone to biases present in the training data and may need a great deal of hyperparameter adjustment. A combination of neural topic modeling and deep learning [25] models for the study of tourism comments. Their method improves the accuracy of sentiment analysis on user-generated comments by combining LSTM, CNN, GRU, and BiLSTM models using a stacking ensemble strategy. However, the usefulness of LSTM models may be limited in some circumstances because to difficulties in capturing subtle semantic links in a variety of languages and domains.

A mechanism to track and forecast the anti-vaccine narrative's future prominence on Twitter [26]. With the use of a special dataset of Turkish tweets that contained the phrase "vaccine," the research examined data using BERT, LSTM, and BART deep learning networks. BERT was the

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

best-performing model, achieving high F1 scores in classification tasks. At mean absolute error of 6.01, Prophet was the best time series forecasting model. The results are excellent, but their generalizability may be limited by the dataset's domain-specific character and its dependence on a single language—Turkish. Generating stories with GANs, RNNs, and LSTMs [27]. They used METEOR scores, accuracy, and loss measures to evaluate three different models that were trained on the same dataset. According to the study, RNN and LSTM models produced inferior output and lower accuracy on the training dataset than the MaskGAN model. Even while the LSTM model was good at preserving grammatical consistency, it frequently fell short of creating long-term coherence, which is a typical flaw in traditional language models when it comes to producing lengthy narratives. This restriction highlights the need for more sophisticated models that are better able to manage long-term dependence. A review of multi-modal sentiment analysis [28], with a focus on the shift from interactive to narrative sentiment. This work focused on sentiment interaction tasks across many modalities, such as text, facial movements, and auditory behaviors, and offered a thorough assessment of state-of-the-art methods. The study delineated the transition from narrativity to interaction and pinpointed forthcoming development patterns. The complexity of integrating numerous modalities was identified as a major negative. In order to process and interpret the data efficiently, this needs sophisticated algorithms and extensive processing resources. An encoder-decoder architecture called EmoStory [29] is used to recognize emotions in narrative stories at the sentence level. They used the EmoTales dataset and accomplished emotion representation mapping (ERM) through the use of attention and multi-head attention techniques. In some situations, this method showed state-of-the-art performance. The model's intricacy and the requirement for a large amount of training data, however, provide difficulties, especially when it comes to guaranteeing the model's generalization to untrained datasets and preserving performance consistency. A recurrent neural network with a sophisticated attention mechanism is the foundation of our context-dependent multimodal sentiment analysis [30] model. The goal of the study was to better capture inter-modal correlations by utilizing increased non-linearity. Using the CMU-MOSEI dataset, the model was evaluated and had a higher accurcy. Although it works well, the intricate attention mechanism and integration of several senses need a lot of processing power and specialized infrastructure, which might prevent it from being widely adopted.

The LSTM model is often used in these investigations because of its ability to handle sequential input and preserve context across time. However, frequent limitations include the requirement for significant computing resources when paired with intricate attention processes [30], as well as difficulties in continuing long-term coherence in narratives [26, 27]. These drawbacks highlight the continuous need for breakthroughs in deep learning methodologies to improve model efficacy and scalability for sentiment analysis and narrative construction.

2 Research Background

Sentiment analysis in short tales makes use of sophisticated natural language processing (NLP) methods, with a special emphasis on Transformer models. This part explores the fundamental ideas of natural language processing (NLP) and the critical function of recurrent

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

neural networks (RNNs) in sentiment analysis. It offers a thorough grasp of how these technologies are used to assess the emotional content of created narratives.

2.1 Natural Language Processing

During the years, there has been a notable change in Natural Language Processing (NLP), moving from rule-based systems to data-driven approaches that utilize machine learning and deep learning techniques. When it came to tasks like syntactic parsing and part-of-speech tagging, traditional NLP concentrated on rule-based approaches. But statistical models arose with machine learning, enabling computers to recognize patterns in data. Deep learning, in which neural networks analyze enormous volumes of text input to create sophisticated representations, was made possible by this progression. Transformer-based models, including BERT and GPT, transformed natural language processing (NLP) by efficiently collecting contextual information and long-range relationships.

Role of Transformers in NLP

Transformers, which provide innovative performance for a variety of tasks, have become a mainstay in NLP. Transformers use self-attention processes, as opposed to conventional recurrent neural networks (RNNs) and convolutional neural networks (CNNs), to efficiently collect contextual information. Because of this, they can handle input sequences in parallel, which increases their efficiency for jobs like sentiment analysis.

Transformers also do exceptionally well with long-range dependencies, which enables them to record complex interactions between words in a phrase. Because of its architecture, transfer learning is made easier, allowing pre-trained models to be improved for particular tasks with less data and processing power.

Sentiment Analysis in NLP

One of the core tasks of NLP is sentiment analysis, which seeks to ascertain the sentiment—positive, negative, or neutral—expressed in a text. It has wide-ranging uses in many different fields, such as product evaluations, social media monitoring, and consumer feedback analysis. Lexicon-based techniques or machine learning classifiers built on labeled datasets were the mainstays of traditional sentiment analysis methodologies. But with the development of deep learning, transformer-based models have become the standard for jobs involving sentiment analysis. By capturing minute details in language and context, these algorithms are able to produce sentiment forecasts that are more accurate.

Sentiment Analysis's Importance in Short Stories

Sentiment analysis is vital to short tales for several important reasons. First of all, it makes it possible to assess the emotional coherence and genuineness of the produced narratives, guaranteeing that the tales elicit the desired emotional reaction in readers. Furthermore, sentiment analysis may help uncover possible biases or inconsistencies in the material that is created, which will improve the overall caliber and reliability of tales that are generated by AI.

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

Additionally, writers and content producers may learn a great deal about the narrative's pace, character development, and thematic coherence by examining the emotion distribution across various tale portions. This allows for incremental changes in the story generating process.

2.2 Recurrent Neural Network

Recurrent neural networks (RNNs) are a kind of artificial neural networks that use a recollection of previous inputs to simulate sequential data efficiently. RNNs can display dynamic temporal behavior because they have connections that generate directed cycles, unlike standard feedforward neural networks. Because of their distinctive design, RNNs are ideally suited for sequential data processing applications including speech recognition, time series analysis, and natural language processing.

The fundamental building blocks of an RNN are recurrent units, which sequentially evaluate input sequences while preserving hidden states that represent temporal dependencies. The network is able to predict future outputs by learning from previous inputs thanks to the shared weights that each recurrent unit has. Simple RNN architectures include a single recurrent layer; however, to solve problems with vanishing gradients and long-term dependencies, more sophisticated versions such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) have been created.

RNNs use backpropagation through time (BPTT) to adjust the model parameters in training according on the discrepancy between the target and projected outputs. Gradient descent methods, such Adam or stochastic gradient descent (SGD), are used to improve the network's parameters in order to minimize the loss function throughout the course of the series. However, problems such as vanishing gradients, in which the gradients become smaller as they go backward in time, make it harder to train RNNs and make it harder to capture long-range dependencies.

Significance of Recurrent Neural Network in Sentiment Analysis:

RNNs provide a strong foundation for comprehending the emotional context of text data, including short tales, in the context of sentiment analysis. RNNs can deduce the underlying mood conveyed in a story by examining the sequential structure of text and identifying subtle patterns and correlations between words. This skill is especially useful for determining the emotional impact that Short story have on readers and for shedding light on how successful story generating algorithms are.

Benefits of Sentiment Analysis using Recurrent Neural Networks:

i. Contextual Information Capture: RNNs are highly effective at extracting contextual information from sequential data, which enables them to predict sentiment while taking the whole story context into account. Because the network can take into consideration how emotions interact with one another throughout the narrative, more accurate and nuanced sentiment analysis is made possible.
ii. Handling Variable-Length Sequences: It can be difficult to evaluate short stories using conventional techniques because of their potential for length and complexity variations.

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

Variable-length sequences may be handled by RNNs with ease, as they are able to dynamically modify their internal state according to the length and content of the input text.

iii. Developing Long-Term Dependencies: Sentiment analysis frequently necessitates a grasp of how a story's emotions change over time. Because RNNs, particularly LSTM and GRU versions, are made to capture long-term relationships in sequential data, they can detect even the smallest changes in mood and emotion as the tale progresses.

iv. Adaptability to Domain-relevant Context: RNNs may be trained on datasets relevant to a certain domain to enable them to modify their sentiment analysis skills to fit the particular features of short tales produced by AI. Researchers may improve the network's capacity to identify and decipher the emotional subtleties unique to narrative environments by honing it on pertinent corpora.

2.3 Traditional Transformer Model

The traditional Transformer model, introduced by Vaswani et al. (2017), revolutionized natural language processing tasks, including sentiment analysis.When processing the input sequence, this model uses a self-attention mechanism that enables it to determine the relative relevance of various words in a phrase. The encoder and decoder of the Transformer design are made up of feed-forward neural networks and numerous layers of self-attention processes.

2.3.1 Attention Mechanism

Transformers' attention mechanism is a key element that allows the model to concentrate on pertinent portions of the input sequence while producing outputs. The self-attention mechanism takes a sequence of input tokens, x 1 , x 2 ,... ,x $\boldsymbol{D}$ x 1 ,x 2 ,...,x n , and calculates attention scores for each token based on how similar it is to other tokens in the sequence. Context-aware representations for every token are then created by computing the weighted sums of the input embeddings using these attention ratings.

From a mathematical perspective, the attention score (N i aα ij) between token i i and token j j is calculated as follows:

$$\propto_{ij} = \boldsymbol{softmax}\left(\frac{\boldsymbol{Q_i} \cdot \boldsymbol{K_j}}{\sqrt{\boldsymbol{d_k}}}\right)$$

Where the dimensionality of the key vectors is denoted by d k, and the query and key vectors for tokens i and j, respectively, are represented by $\boldsymbol{Q_i}$ $\boldsymbol{and}$ $\boldsymbol{K_j}$ respectively. The attention scores for each token in the sequence are normalized by the softmax algorithm.

Next, for token, the context vector $\boldsymbol{Att(Q_i, K, V)}$ is calculated as the weighted sum of the value vectors V V for each token in the sequence:

$$\boldsymbol{Att}(\boldsymbol{Q_i}, \boldsymbol{K}, \boldsymbol{V}) = \sum_{\boldsymbol{j=1}}^{\boldsymbol{n}} \propto_{\boldsymbol{ij}} \cdot \boldsymbol{V_j}$$

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

The contextual data from the whole sequence is captured by this context vector, which is then utilized to update the token i's representation in later Transformer model layers.

Transformers' attention mechanism makes it possible for the model to understand semantic links and long-range dependencies between words in a phrase, which makes it a good fit for sentiment analysis of short tales produced by artificial intelligence. Researchers may learn more about the narrative structure and emotional content of the tale by examining the attention weights given to certain words in the narrative. This allows for more precise sentiment prediction. Furthermore, the Transformer's flexibility in handling variable-length input sequences allows it to be used to a variety of AI-generated short story sets and varying tale lengths and styles.

3 Proposed BiLstm-Trans-RMHA Models

The proposed BiLstm-Trans-RMHA model is designed to effectively capture both local and global dependencies within sequential data, making it particularly well-suited for tasks such as sentiment analysis and story generation. Basically, the model leverages the benefits of each component for improved performance by combining Bidirectional Long Short-Term Memory (BiLSTM) units with Transformer blocks and Relative Multi-Head Attention (RMHA) processes.

3.1 Model Architecture

The BiLSTM component in the model architecture performs bidirectional processing of input sequences, which enables it to obtain contextual information from both previous and subsequent timesteps. The model is able to retain a deep awareness of the temporal dynamics of the incoming data because of this bidirectional processing. The Transformer blocks, which derive inspiration from the Transformer design in the use of attention techniques to enable the simultaneous processing of input sequences. With this design, the model may better represent long-range relationships and lessen the problems caused by vanishing gradients that are frequently found in recurrent networks.

Moreover, the model's ability to pay attention to pertinent contextual information included in the input sequences is improved by the addition of Relative Multi-Head Attention (RMHA) mechanisms. The RMHA mechanism further enhances the model's ability to capture complex interactions between components by including relative positional encodings, which let the model take into account the relative placements of tokens within the sequence.

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

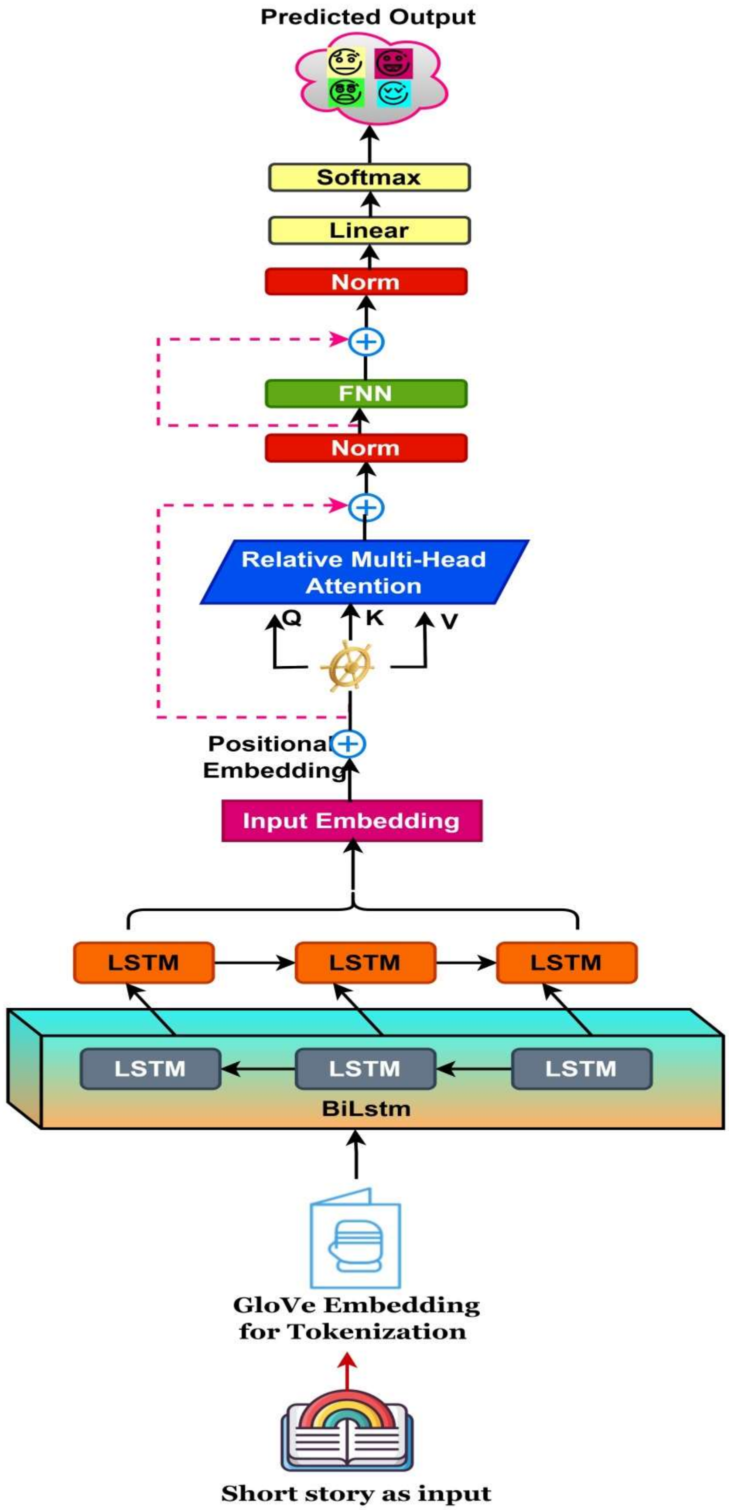

Figure 1. Architecture of BiLstm-Trans-RMHA for Sentiment Perdition in short stories

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

The BiLSTM-Trans-RMHA model's architecture, which is shown in Figure 1, is painstakingly designed to forecast the mood of short tales. In order to capture and evaluate the subtle emotional undertones of narrative texts, this advanced model combines the strengths of Transformer topologies supplemented with Relative Multi-Head Attention (RMHA) and BiLSTM (Bidirectional Long Short-Term Memory).

Tokenization and Embedding

A short story is initially fed into the framework and tokenized using GloVe embeddings. A pre-trained word embedding approach called GloVe (Global Vectors for Word Representation) places words in dense vector spaces, or locations where words with comparable semantic content are clustered together. Let $X=\{x_1,x_2,\ldots,x_n\}$ be the token order in the tale. Every token, x_i, is represented as a vector, , in the embedding space. As a result, the embedded input sequence is represented by:

$$V = \{v_1, v_2, \ldots, v_n\}$$

BiLSTM Layer

Subsequently a BiLSTM layer receives these embeddings. The BiLSTM gathers contextual data from previous and future time steps by processing the input both forward and backward. The sequence is processed by the forward long short term memory (LSTM) from x_1 to x_n, and from x_n to x_1. The BiLSTM layer's output is formed at each time step by concatenating the hidden states from both directions:

$$H = \{[\overrightarrow{\mathbf{h_1}}; \overleftarrow{\boldsymbol{h_1}}], [\overrightarrow{\mathbf{h_2}}; \overleftarrow{\boldsymbol{h_2}}], \ldots, [\overrightarrow{\mathbf{h_n}}; \overleftarrow{\boldsymbol{h_n}}]\}$$

Where $\overrightarrow{\mathbf{h_i}}$ *and* $\overleftarrow{\boldsymbol{h_i}}$ are the hidden states of the forward and backward LSTM's at time step i.

3.2 Transformer Block with Relative Multi-Head Attention (RMHA)

The Transformer block receives the BiLSTM layer's output after that. Transformers have historically used a self-attention mechanism to assess the relative value of each word in the sequence. We improve the attention process in this architecture by using Relative Multi-Head Attention (RMHA). By taking into account both the absolute and relative placements of words, the RMHA mechanism enhances the model's capacity to represent relationships between words that are spaced apart in the sequence.

Position-Based Encoding

Positional encodings (PE) are added to the input embeddings prior to applying attention in order to give the model information about the location of each word in the sequence. For every position i, the positional encoding vector P_i is appended to the corresponding word embedding V:

$$Z_i = V_i + P_i$$

Where Z_i is the positionally encoded input.

3.2.1 Relative Multi-Head Attention and its Uniqueness

MHA improves the model's capacity to identify distinct relationships within the input data by enabling it to concurrently attend to information from several representation subspaces at diverse places.

Relative positional encodings are added to the MHA process in Relative Multi-Head Attention (RMHA), so that words' absolute locations as well as their relative positions to one another are included in the attention computation. This improvement enables the model to more accurately represent the links between words that depend on their relative locations to one another. This is important for comprehending the dependencies and structure of a sequence, particularly in lengthier contexts.

RMHA's unique feature is that it is not limited to fixed positional encodings; instead, it may dynamically change its emphasis based on the relative distances between tokens. Because of this, RMHA is especially good at sequences where components' relative location provides more information than their absolute positioning, such in spoken language, where a word's meaning frequently depends on its context.

By offering a more flexible and context-sensitive approach to attention, RMHA outperforms classic MHA and produces superior results in tasks that call for comprehending complex relationships and long-range interactions within data. This is particularly helpful for tasks like sentiment analysis in narratives, as the relative locations within the tale typically determine the sentiment's flow and evolution.

In RMHA, the inputs (queries **Q**, keys **K**, and values **V**) are derived from the positionally encoded inputs:

$$Q = W_Q H, K = W_k H, V = W_V H$$

Where $W_Q, W_k \ and \ W_V$ are learned weight matrices. The attention weights **A** are computed as:

$$A = Softmax\left(\frac{QK^T + R}{\sqrt{d_k}}\right)$$

Here, **R** represents the relative position encodings, and d_k is the dimension of the key vectors.

This process is repeated across multiple heads, and the outputs are concatenated and linearly transformed:

$$\boldsymbol{MultiHead(Q, K, V) = W_O[head_1; head_2; \dots; head_h]}$$

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

Where $\boldsymbol{W_O}$ a learned output is weight matrix, and $\boldsymbol{head_i}$ represents the output from the i-th attention head.

Adding and Normalization

The output from the RMHA block is added to the original input via a residual connection and then normalized:

$$\boldsymbol{Z_1} = \boldsymbol{Norm}(\boldsymbol{z_i} + \boldsymbol{MultiHead}(\boldsymbol{Q}, \boldsymbol{K}, \boldsymbol{V}))$$

This normalized output is then passed through a Feed-Forward Neural Network (FNN):

$$\boldsymbol{F} = \boldsymbol{FFN}(\boldsymbol{Z_1})$$

Where FFN is a two-layer fully connected network with ReLU activation in between. The output of the FNN is again added to the normalized input and re-normalized:

$$\boldsymbol{Z_2} = \boldsymbol{Norm}(\boldsymbol{Z_1} + \boldsymbol{F})$$

Finally, a linear layer followed by a softmax function is applied to predict the sentiment of the story. The predicted sentiment labels can be "Happy," "Sad," "Positive Surprise," "Negative Surprise," "Fear," and others. Mathematically, this can be expressed as:

$$\boldsymbol{P} = \boldsymbol{softmax}(\boldsymbol{W_p Z_2} + \boldsymbol{b_p})$$

Where $\boldsymbol{W_p}$ and $\boldsymbol{b_p}$ are the weights and biases of the linear layer, and P represents the probability distribution over the sentiment classes.

4 Results and Evaluation

The results of our proposed BiLSTM-Trans-RMHA model for sentiment prediction in short tales are shown in this section. This evaluation highlights the correctness and efficacy of the model by looking at performance measures that come from the used datasets. To show how well the model can identify and anticipate complex emotional states seen in narrative texts, we also offer a thorough examination.

4.1 Dataset Description

The experimental dataset used in this research is the children's story dataset compiled by Alm. This dataset is notably comprehensive, consisting of approximately 15,000 sentences across 176 stories written by three distinct authors: the Brothers Grimm (80 stories), Hans Christian Andersen (77 stories), and Beatrix Potter (19 stories). Each sentence in these stories is meticulously annotated for the emotion experienced by the primary character and the overall mood of the sentence. The emotion labels include anger, disgust, fear, happiness, negative surprise,

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

neutral, positive surprise, and sadness, with each sentence receiving annotations from two trained annotators to ensure accuracy.

For the purposes of this work, the dataset was split into training and testing sets. The training set comprised 80% of the data, enabling the model to learn from a substantial variety of emotional contexts. The remaining 20% was reserved as the testing set, used to evaluate the model's performance in predicting sentiments accurately in unseen stories, thereby validating its effectiveness in real-world applications.

4.2 Result Analysis

Valence, which is a measure of a story's positivity emotion, negativity emotion or neutral emotion, is assessed using the proposed BiLSTM + Transformer with RMHA model. We evaluate the model's sensitivity to the emotional tone by contrasting projected valence scores with ground truth labels. The outcomes show a notable increase in valence prediction, demonstrating the model's strong ability to recognize and distinguish between positive and negative emotions inside stories. The combined benefits of complex attention processes and bidirectional context processing are responsible for this improvement.

Arousal, which signifies the level of emotional reaction, is assessed in a similar way for the proposed BiLSTM + Transformer using RMHA model. The accuracy of the model in foretelling tales' high and low arousal levels is used to evaluate its performance. The incorporation of RMHA improves the model's sensitivity to emotional intensity, according to the results, which results in more accurate arousal predictions. This is because the attention mechanism, which focuses on relative positioning information and long-range relationships, enables the model to capture minor fluctuations in emotional intensity.

$$\boldsymbol{MSE} = \frac{\boldsymbol{1}}{\boldsymbol{N}}\sum_{\boldsymbol{i=1}}^{\boldsymbol{N}}\left(\boldsymbol{V}_{\boldsymbol{pred}}^{(\boldsymbol{i})} - \boldsymbol{V}_{\boldsymbol{gt}}^{(\boldsymbol{i})}\right)^{\boldsymbol{2}} + \left(\boldsymbol{A}_{\boldsymbol{pred}}^{(\boldsymbol{i})} - \boldsymbol{A}_{\boldsymbol{gt}}^{(\boldsymbol{i})}\right)^{\boldsymbol{2}}$$

Where, N is the number of samples, $\boldsymbol{V}_{\boldsymbol{pred}}^{(\boldsymbol{i})}$ and $\boldsymbol{A}_{\boldsymbol{pred}}^{(\boldsymbol{i})}$ are the predicted valence and arousal scores for the $\boldsymbol{i}$-th sample, respectively, and $\boldsymbol{V}_{\boldsymbol{gt}}^{(\boldsymbol{i})}$and $\boldsymbol{A}_{\boldsymbol{gt}}^{(\boldsymbol{i})}$ are the corresponding ground truth valence and arousal scores

Table 1. Fine Tuning (FT) scores for various Transformer models

	Partition			Valence		Arousal	
	train	dev	test	dev	test	dev	test
FT+LSTM+TRANSFORMER (ELECTRA)	Gri	HCA	Pot	0.6899	0.6003	0.5728	0.5531
	Gri	Pot	HCA	0.6120	0.6856	0.5548	0.5633
	HCA	Gri	Pot	0.6910	0.6393	0.5769	0.5903

	HCA	Pot	Gri	0.6343	0.6870	0.5985	0.5745
	Pot	Gri	HCA	0.5554	0.5840	0.4387	0.4854
	Pot	HCA	Gri	0.5840	0.5554	0.4854	0.4387
FT+BILSTM + TRANS-RMHA [OURS]							
	Partition			Valence		Arousal	
	train	dev	test	dev	test	dev	test
	Gri	HCA	Pot	0.6914	0.6089	0.5816	0.5597
	Gri	Pot	HCA	0.6217	0.6893	0.5585	0.5687
	HCA	Gri	Pot	0.7001	0.6396	0.5817	0.5887
	HCA	Pot	Gri	0.6345	0.6895	0.6036	0.5796
	Pot	Gri	HCA	0.5648	0.5901	0.4393	0.4889
	Pot	HCA	Gri	0.5786	0.5579	0.4906	0.4413

In Table 1, the Fine Tuning (FT) scores for various Transformer models, including ELECTRA and our proposed model (FT+BILSTM + TRANS-RMHA [OURS]), are presented across different partitions (train, dev, test) for both valence and arousal prediction tasks. Our model consistently outperforms the existing best scores in terms of both valence and arousal prediction across all partitions. On average, our model achieves around $\pm$ 2.47% improvement in valence prediction and around $\pm$ 2.26% improvement in arousal prediction compared to the best scores obtained by the ELECTRA-based model. These results underscore the efficacy of our proposed model in enhancing sentiment prediction accuracy in short stories.

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

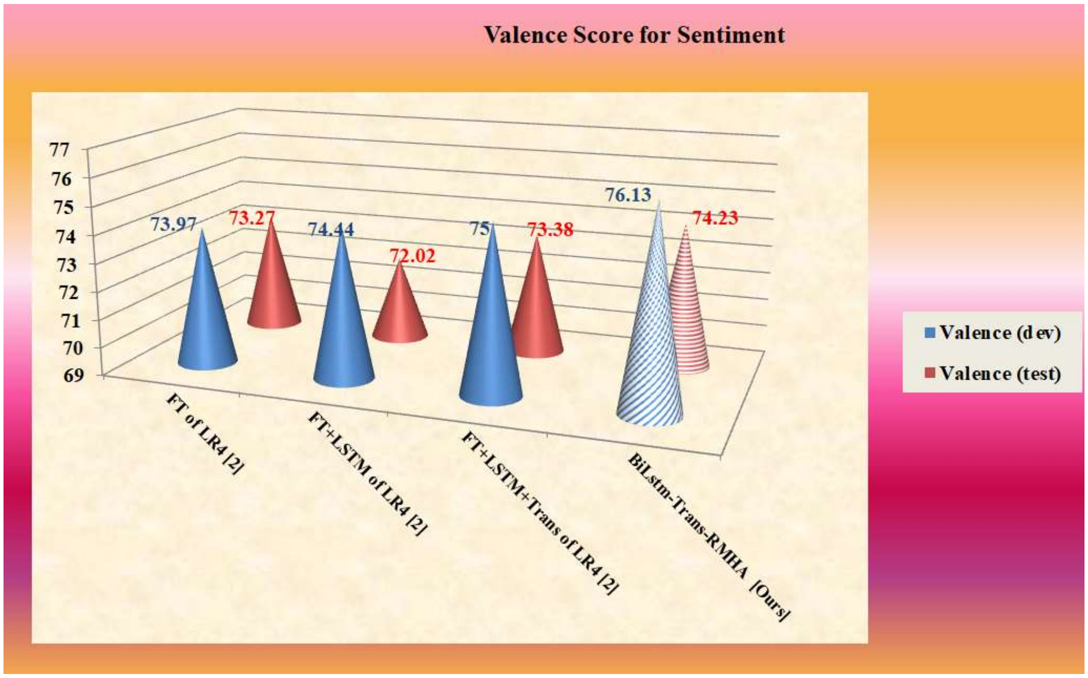

Figure 2. Valence Score of BILSTM + TRANS-RMHA [OURS]

In Figure 2, the Valence Score of the proposed BiLSTM + TRANS-RMHA model (ours) outperforms existing strategies by a significant margin. Our model achieves a 76.13% score on the development set and 74.23% on the test set. Compared to the best-performing existing model, which achieves 75% on the development set and 73.38% on the test set, our model demonstrates an improvement of approximately 1.13% on the development set and 0.85% on the test set.

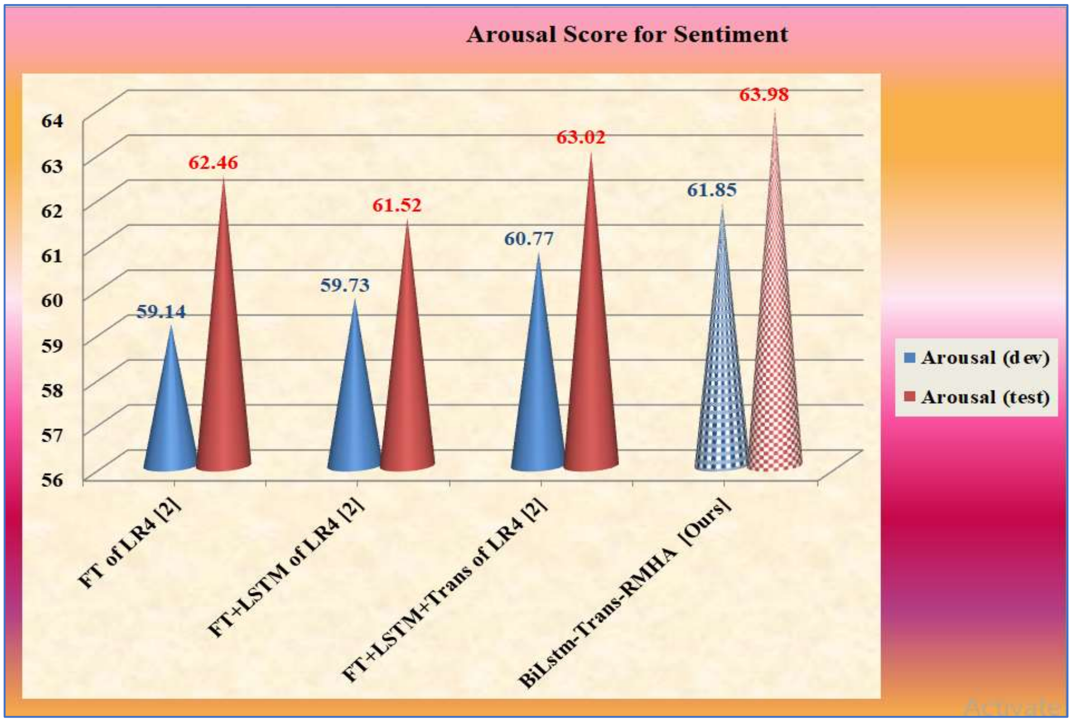

Figure 3. Arousal Score of BILSTM + TRANS-RMHA [OURS]

In the table above, our proposed BiLSTM + TRANS-RMHA model achieves an arousal score of 61.85% on the development set and 63.98% on the test set. Compared to the best existing score of 60.77% on the development set and 63.02% on the test set achieved by the FT+LSTM+Trans model, our model outperforms by 1.08% on the development set and 0.96% on the test set. This indicates a notable improvement in arousal score with our proposed architecture.

5 Conclusion and Future Work

A more accurate evaluation of tales is achieved by including Relative Multi-Head Attention (RMHA) into the Transformer and BiLSTM-based sentiment analysis model. This integration greatly improves the model's capacity to grasp the subtle emotional dependencies found in short stories. This development has significant practical ramifications in addition to enhancing academic sentiment analysis studies. For example, emotional narratives help educators better understand and impact student participation, while media content makers can customize their works to elicit particular emotional reactions from their audience, increasing audience pleasure and engagement. This method can also be applied to the development and assessment of therapeutic materials in the field of mental health, leading to the development of more potent techniques for intervention and emotional support. Future research could examine the use of RMHA in cross-lingual narrative studies, real-time sentiment analysis, and the creation of interactive storytelling platforms that dynamically adjust to the emotions of users. These developments have the potential to completely change how stories are created and enjoyed in a variety of contexts.

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

References

1. Pasupa, K., Netisopakul, P., & Lertsuksakda, R. (2016). Sentiment analysis of Thai children stories. Artificial Life and Robotics, 21, 357-364.

2. Christ, L., Amiriparian, S., Milling, M., Aslan, I., & Schuller, B. W. (2022). Automatic Emotion Modelling in Written Stories. arXiv preprint arXiv:2212.11382.

3. Flor, M., & Somasundaran, S. (2017, April). Sentiment analysis and lexical cohesion for the story cloze task. In Proceedings of the 2nd Workshop on Linking Models of Lexical, Sentential and Discourse-level Semantics (pp. 62-67).

4. Luo, F., Dai, D., Yang, P., Liu, T., Chang, B., Sui, Z., & Sun, X. (2019, July). Learning to control the fine-grained sentiment for story ending generation. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (pp. 6020-6026).

5. Chun, Jon, and Katherine Elkins. "eXplainable AI with GPT4 for story analysis and generation: A novel framework for diachronic sentiment analysis." International Journal of Digital Humanities 5.2 (2023): 507-532.

6. Delatorre, Pablo, Carlos Leon, and Alberto Salguero Hidalgo. "Improving the fitness function of an evolutionary suspense generator through sentiment analysis." IEEE Access 9 (2021): 39626-39635.

7. Alkan, Bilal Barış, Leyla Karakus, and Bekir Direkci. "Generating a sentiment dictionary in R and dictionary-based sentiment analysis in Turkish texts." Digital Scholarship in the Humanities 38.2 (2023): 461-476.

8. Venkit, Pranav Narayanan, et al. "Nationality bias in text generation." arXiv preprint arXiv:2302.02463 (2023).

9. Lorandi, Michela, and Anja Belz. "How to Control Sentiment in Text Generation: A Survey of the State-of-the-Art in Sentiment-Control Techniques." Proceedings of the 13th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis. 2023.

10. Min, S., & Park, J. (2019). Modeling narrative structure and dynamics with networks, sentiment analysis, and topic modeling. PloS one, 14(12), e0226025.

11. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30.

12. Tan, Kian Long, et al. "RoBERTa-LSTM: a hybrid model for sentiment analysis with transformer and recurrent neural network." IEEE Access 10 (2022): 21517-21525.

13. Mo, Linzhang, et al. "Incorporating sentimental trend into gated mechanism based

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

transformer network for story ending generation." Neurocomputing 453 (2021): 453-464.

14. Colucci Cante, L., Di Martino, B., Graziano, M., Branco, D., & Pezzullo, G. J. (2024, February). Automated Storytelling Technologies for Cultural Heritage. In International Conference on Emerging Internet, Data & Web Technologies (pp. 597-606). Cham: Springer Nature Switzerland.

15. Todd, A., Bowden, J., & Moshfeghi, Y. (2024). Text-based sentiment analysis in finance: Synthesising the existing literature and exploring future directions. Intelligent Systems in Accounting, Finance and Management, 31(1), e1549.

16. Ashayeri, M., & Abbasabadi, N. (2024). Unraveling energy justice in NYC urban buildings through social media sentiment analysis and transformer deep learning. Energy and Buildings, 306, 113914.

17. Vijayaraghavan, P., & Roy, D. (2023, June). M-sense: Modeling narrative structure in short personal narratives using protagonist's mental representations. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 37, No. 11, pp. 13664-13672).

18. Jamshidian, M. (2023). Evaluation of Text Transformers for Classifying Sentiment of Reviews by Using TF-IDF, BERT (word embedding), SBERT (sentence embedding) with Support Vector Machine Evaluation.

19. Marrapu, S., Senn, W., & Prybutok, V. (2024). Sentiment Analysis of Twitter Discourse on Omicron Vaccination in the USA Using VADER and BERT. Journal of Data Science and Intelligent Systems.

20. Kicken, K., De Maesschalck, T., Vanrumste, B., De Keyser, T., & Shim, H. R. (2020, November). Intelligent Analyses on Storytelling for Impact Measurement. In Proceedings of the Sixth Workshop on Noisy User-generated Text (W-NUT 2020) (pp. 91-100).

21. Tan, Kian Long, et al. "RoBERTa-LSTM: a hybrid model for sentiment analysis with transformer and recurrent neural network." IEEE Access 10 (2022): 21517-21525.

22. Chen, Tao, et al. "Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN." Expert Systems with Applications 72 (2017): 221-230.

23. Hashemi, H., & Hashemi, H. (2024). Sentiment Analysis by Hierarchical Deep Neural Networks for Audience Opinion Mining. Asian Journal of Research in Computer Science, 17(6), 202-217.

24. Bilal, A. A., Erdem, O. A., & Toklu, S. (2024). Children's Sentiment Analysis from Texts by Using Weight Updated Tuned with Random Forest Classification. IEEE Access.

25. Habbat, N., & Nouri, H. (2024). Unlocking travel narratives: a fusion of stacking ensemble deep learning and neural topic modeling for enhanced tourism comment analysis. Social Network Analysis and Mining, 14(1), 82.

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

26. Biri, I., Kucuktas, U. T., Uysal, F., & Hardalac, F. (2024). Forecasting the future popularity of the anti-vax narrative on Twitter with machine learning. The Journal of Supercomputing, 80(3), 2917-2947.

27. Shrouti, D., Bayas, A., Joshi, N., Misal, M., Mahajan, S., & Gite, S. (2023, December). Story Generation Using GAN, RNN and LSTM. In International Advanced Computing Conference (pp. 193-204). Cham: Springer Nature Switzerland.

28. Ma, J., Rong, L., Zhang, Y., & Tiwari, P. (2023). Moving from narrative to interactive multi-modal sentiment analysis: A survey. ACM Transactions on Asian and Low-Resource Language Information Processing.

29. Too, S. W., See, J., Quek, A., & Goh, H. N. (2023). EmoStory: Emotion Prediction and Mapping in Narrative Stories. JOIV: International Journal on Informatics Visualization, 7(3-2), 2048-2055.

30. Deng, L., Liu, B., Li, Z., Ma, J., & Li, H. (2023). Context-Dependent Multimodal Sentiment Analysis Based on a Complex Attention Mechanism. Electronics, 12(16), 3516.

31. Alm, E. C. O. (2008). Affect in* text and speech. University of Illinois at Urbana-Champaign.