EAI Endorsed Transactions
on Scalable Information Systems

Research Article EAI.EU

Multitask Sentiment Analysis and Topic Classification Using BERT

Parita Shah[1,*], Hiren Patel[2] and Priya Swaminarayan[3]

[1,2] Vidush Somany Institute of Technology and Research, Kadi Sarva Vishwavidyalaya, Sarva Vidyalaya Kelavani Mandal, Kadi, India
[3]Faculty of Information Technology and Computer Science, Parul University, Vadodara, India

Abstract

In this study, a multitask model is proposed to perform simultaneous news category and sentiment classification of a diverse dataset comprising 3263 news records spanning across eight categories, including environment, health, education, tech, sports, business, lifestyle, and science. Leveraging the power of Bidirectional Encoder Representations from Transformers (BERT), the algorithm demonstrates remarkable results in both tasks. For topic classification, it achieves an accuracy of 98% along with balanced precision and recall, substantiating its proficiency in categorizing news articles. For sentiment analysis, the model maintains strong accuracy at 94%, distinguishing positive from negative sentiment effectively. This multitask approach showcases the model's versatility and its potential to comprehensively understand and classify news articles based on content and sentiment. This multitask model not only enhances classification accuracy but also improves the efficiency of handling extensive news datasets. Consequently, it empowers news agencies, content recommendation systems, and information retrieval services to offer more personalized and pertinent content to their users.

Keywords: BERT; Analyzing Sentiments; Categorizing Topics; Multitasking in Learning; Processing Natural Language; Machine Learning Techniques; News Dataset; Retrieval of Information

Received on 04 March 2024, accepted on 31 May 2024, published on 11 July 2024

doi: 10.4108/eetsis.5287

1. Introduction

In today's information-driven world, news articles serve as a vital source of knowledge, offering insights, updates, and opinions on a vast array of topics. As the volume of news articles grows exponentially, the need for efficient and effective methods to organize and categorize this wealth of information becomes increasingly paramount. Accurate topic classification and sentiment analysis are pivotal components in achieving this goal, as they enable content providers, recommendation systems, and information retrieval services to better understand and respond to user preferences [1].

Traditional approaches in machine learning (ML) have historically addressed text classification tasks. However, their effectiveness has been somewhat limited when confronted with the complexities of human language. The intricacies, context, and dynamic trends in language present considerable challenges for conventional rule-based or statistical models. Additionally, the conventional ML method demands intricate feature engineering, which might not generalize effectively across various tasks.

Over the last decade, processing of Natural language has undergone a significant transformation with the emergence of deep learning models, notably Bidirectional Encoder Representations from Transformers. BERT, a transformer-based model pertained on language data, stands as a

[*]Corresponding author. Email: paritaponkiya@gmail.com

ground-breaking advancement in comprehending and processing human language [2]. Its capability to grasp bidirectional contextual information, considering the entire sentence or paragraph rather than relying solely on left-to-right or right-to-left context, has reshaped the landscape of numerous NLP tasks [3].

Our research embarks on an exploration of the potential of BERT in multitask learning, particularly in the domains of sentiment analysis and topic classification, while incorporating custom attention mechanisms. This novel approach seeks to leverage BERT's remarkable capacity for understanding the intricacies of natural language and enhance it with custom attention for more fine-grained control over its attention patterns. Our research objective is to design a unified architecture that can simultaneously address both topic classification and sentiment analysis in a single model to improve efficiency, contextual understanding, and reduce latency in text analysis.

Our research is motivated by the escalating demand for more sophisticated content curation and information retrieval systems. The need for advanced content curation systems is growing as a result of the enormous volume of news stories that are created every day. Traditional information retrieval algorithms find it difficult to handle the dynamic and complicated nature of human language, thus more sophisticated approaches are required. Current approaches frequently handle sentiment analysis and topic classification as independent tasks, which can lead to inefficiencies and worse accuracy. Furthermore, managing a variety of text data, including multilingual and user-generated information, is necessary due to the increase in online content. Traditional approaches are unable to modify attention mechanisms to enhance contextual comprehension and multitasking. Furthermore, models that improve text analysis latency and accuracy simultaneously must be designed to increase computing efficiency. In an era overwhelmed by the sheer volume of available information, there's a pressing need not only to categorize news articles into relevant topics but also to gauge their sentiments, providing readers with a more comprehensive context. This enhancement empowers content recommendation systems to offer tailored content and aids information retrieval services in presenting the most pertinent articles.

Additionally, the surge in online content and social media has led to a diverse array of text data, including user-generated content and multilingual text. BERT, with its multilingual capabilities and promise of understanding diverse linguistic nuances, emerges as an ideal solution to address these challenges. Incorporating custom attention mechanisms is anticipated to enable the model to adjust its attention patterns based on the task, thereby maximizing its multitasking capabilities.

Our research is fuelled by the escalating demand for more sophisticated content curation and information retrieval systems. In an era overwhelmed by the sheer volume of available information, there's a pressing need not only to categorize news articles into relevant topics but also to gauge their sentiments, providing readers with a more comprehensive context. This enhancement empowers content recommendation systems to offer tailored content and aids information retrieval services in presenting the most pertinent articles. Moreover, the surge in online content and social media has led to a diverse array of text data, including user-generated content and multilingual text. BERT, with its multilingual capabilities and promise of understanding diverse linguistic nuances, emerges as an ideal solution to address these challenges. Incorporating custom attention mechanisms is anticipated to enable the model to adjust its attention patterns based on the task, thereby maximizing its multitasking capabilities.

In terms of objectives, this research focuses on creating an integrated BERT-based model capable of multitasking, specifically for sentiment analysis and topic classification within news articles. It aims to evaluate the model's effectiveness considering both classification accuracy and computational efficiency. Additionally, it seeks to explore the potential application of this model for content curation, information retrieval, and personalized content recommendations.

Author's contributions to natural language processing and information retrieval are manifold. We introduce a novel approach to multitask learning using BERT and custom attention, expected to outperform traditional ML models in accuracy and efficiency. We demonstrate the potential of advanced transformer-based models like BERT in handling multiple text classification tasks, thereby enhancing content organization. Our research also offers insights into the advantages of custom attention mechanisms, allowing the model to adapt its attention patterns to the demands of sentiment analysis and topic classification. Finally, the findings and methodologies of this research have the potential to impact content providers, recommendation systems, and information retrieval services, leading to more personalized and relevant content delivery to users.

This paper is structured as follows: After providing the introduction in Section 1, Section 2 illustrates an overview of related work in the fields of sentiment analysis, topic classification, and the use of BERT in NLP tasks. The methodology used in our research, detailing the design of the BERT-based model for topic and sentiment classification are described in Section 3 followed by a comprehensive explanation of the experimental setup and data collection in Section 4. The results and discussion related to the model's performance are presented in Section 5 followed by a concluding remarks and outlines avenues for future research in Section 6. In summary, this research sets out to harness the capabilities of BERT, a cutting-edge NLP model, and augment its multitasking abilities through custom attention mechanisms. The findings are anticipated to provide a substantial leap forward in content curation, information retrieval, and user-specific content recommendations. Moreover, it underscores the promise of

BERT models in addressing complex and evolving NLP tasks and paves the way for further innovations in the field.

In recent years, the field of Natural Language Processing (NLP) has seen significant advancements driven by the emergence of deep learning models, particularly BERT (Bidirectional Encoder Representations from Transformers). BERT, introduced by Devlin et al. in 2018, is a pre-trained transformer-based model that has reshaped the landscape of NLP tasks through its capacity to capture contextual information bidirectional, enabling it to understand the intricacies of human language [1]. As the demand for more sophisticated content curation and information retrieval systems grows, research has shifted its focus toward leveraging the capabilities of BERT for multitask learning in sentiment analysis and topic classification. Research by Vaswani et al. in 2017 laid the foundation for the development of BERT by introducing the Transformer model, which revolutionized sequence-to-sequence tasks in NLP [2]. BERT has since demonstrated impressive results across various domains, including sentiment analysis. The work of Howard and Ruder in 2018 highlighted the significance of transfer learning and the potential of pre-trained models like BERT for improving sentiment analysis, emphasizing the role of pre-training and fine-tuning in achieving state-of-the-art results [3]. When it comes to multitask learning, the study by Liu et al. in 2019 explored the application of BERT for multitask learning in the medical domain, focusing on Named Entity Recognition (NER) and sentence classification tasks. The authors demonstrated that BERT's contextual embedding could be effectively adapted to multiple tasks, indicating its potential for multitasking [4]. Custom attention mechanisms have gained attention in recent research endeavours, primarily within the context of enhancing NLP models. Research by Vaswani et al. in 2017 introduced the attention mechanism as a foundational component of transformer models. These mechanisms allow models to focus on specific elements of input data and are integral to the multitasking capabilities of models like BERT [2]. Recent work by Xie et al. in 2021 explored custom attention mechanisms in transformer models, emphasizing their role in improving efficiency and interpretability [5]. In terms of application, the significance of multitask learning in sentiment analysis and topic classification has been recognized across various domains. Content recommendation systems and information retrieval services have benefited from research efforts in these areas. Personalized content recommendations, in particular, have been a focal point in recent work by Kang et al. in 2021, showcasing the potential of multitask learning in content curation and enhancing user experiences [6]. In conclusion, this literature review outlines the evolution of NLP research from the introduction of the Transformer model to the ground-breaking impact of BERT. The potential of BERT: bidirectional transformer representations for multitask learning in sentiment analysis and topic classification is underpinned by a strong foundation in transfer learning. Custom attention mechanisms have also

2. Related Work

Emerged as a powerful tool for fine-tuning models to specific tasks. Furthermore, the implications of this research extend to content providers, recommendation systems, and information retrieval services, promising more personalized and relevant content delivery to users. The work presented in this literature review underscores the role of BERT: bidirectional transformer representations with custom attention as a potent avenue for improving content organization and user-specific content recommendations [7]. To this end, we utilize the BERT: bidirectional transformer representations model as a text classifier for multitask sentiment analysis and topic classification in news articles. Several studies have demonstrated the effectiveness of BERT: bidirectional transformer representations in various text classification tasks, such as sentiment analysis and document classification. The work pointed out a new BERT model for sentiment classification of tweets, specifically focusing on identifying topic features for a social media bot [8]. Literature Review: Multitask Sentiment Analysis and Topic Classification Using BERT: bidirectional transformer representations with Custom Attention for Newson of the key applications of BERT: bidirectional transformer representations in text classification is sentiment analysis. BERT: bidirectional transformer representations has been widely used for sentiment analysis tasks, including aspect-based sentiment analysis and analysing the impact of topics such as the coronavirus on social life. For example, Xu et al. utilized BERT: bidirectional transformer representations for sentiment analysis on the impact of coronavirus in social life. They found that BERT: bidirectional transformer representations achieved strong performance in capturing sentiment from textual data. Similarly, in a study conducted on sentiment classification of tweets, researchers proposed a deep learning-based topic-level sentiment analysis model that utilized BERT: bidirectional transformer representations. The model aimed to independently identify topic features for a social media bot. The researchers found that BERT: bidirectional transformer representation’s pertaining to a large corpus allowed it to effectively capture sentiment information from the tweets, resulting in improved sentiment classification accuracy compared to other models. Furthermore, BERT: bidirectional transformer representations has also been applied to sentiment analysis in various languages, including Bengali. For instance, researchers discussed the systematic approach of using the BERT: bidirectional transformer representations model for sentiment analysis in Bengali documents [9]. They found that fine-tuning BERT: bidirectional transformer representations transformers and utilizing pre-training and fine-tuning during sequential text classification improved accuracy compared to other models. In addition to sentiment analysis, BERT: bidirectional transformer representations have also been utilized effectively for document classification in the context of news articles [10].

In [11], the BERT SAN model emerges as a significant advancement in sentiment analysis, showcasing noteworthy enhancements compared to benchmarks. This improvement is particularly emphasized in its outcome measurement focused on sentiment evaluation. [12] highlights the Bert Base Chinese Model, surpassing LSTM in accuracy and excelling in capturing nuances in Emotional Attitudes in Texts and User Attitudes Towards Text Topics. [13] reports on the integration of topics with BERT, resulting in improved sentiment classification accuracy rates. The study zeros in on Sentiment Analysis on News Titles, underlining the model's effectiveness in this specific context. The sentiment analysis system developed in [14] identifies positive and negative indications in news titles using BERT. The study articulates its methodology with the BERT: Bidirectional Transformer Representations and sentiment analysis method, highlighting outcomes through Sentiment Analysis on News Titles. In [15], the BiLSTM-Attention model presents a fusion of bidirectional long-short-term memory networks and attention mechanisms, showcasing its application in understanding the broader sentiment landscape. [16] addresses challenges posed by stop-words within datasets, hindering classifier performance. The proposed model leverages BERT for sentiment analysis tasks, specifically gauging its success through Sentiment Analysis Task Performance. [17] positions BERT as a superior performer across eleven NLP tasks, with variations of the BERT base model showcasing Accuracy Enhancement in Aspect Level Sentiment Classification. The implementation of a pre-trained BERT model and RSS feed automation in [18] stands out for its impact on end-user productivity. Usability Rating, End User Productivity, and Business Costs collectively represent the measured outcomes, emphasizing the model's practical implications.[19] adopts top-performing models for sentiment and topic classification of news headlines. The study homes in on Sentiment Analysis of News Headlines and Topic Classification of News Headlines, illustrating its commitment to a nuanced understanding of news sentiment. In [20], the inadequacy of traditional word2vec models is addressed, shifting to BERT for feature extraction and classification using CNN. The outcome, Feature Extraction Using Word2Vec, signifies the model's departure from conventional methods. [21] introduces the BERT CBA model, showcasing anticipated results across various evaluation indicators, particularly in Emotional Probability Distribution in Text. This emphasizes the model's capacity to discern nuanced emotional tones in text.In [22], BERT's substantial success in sentence-level sentiment classification becomes evident, reflecting its prowess in Sentence Level Sentiment Classification and broader Sentiment Analysis. [23] presents a system designed for predicting sentiment classes in news articles, with a focus on the Classification Approach. This underscores the model's efficacy in categorizing sentiments in the complex domain of news articles. Academic interest surges in [24], with a specific focus on automatic extraction of expressions using BERT models. Positive Attitude Expressions and Negative Attitude Expressions emerge as the measured outcomes, indicating a nuanced analysis of sentiments. [25] combines BERT with CNN, RNN, and BiLSTM, offering higher accuracy rates in text classification tasks. The study delves into Classification Metrics, including Accuracy, Precision, Recall, and F1-score, providing a comprehensive assessment. The SA-BERT model in [26] demonstrates a positive impact in the field of sentiment analysis, specifically emphasizing Sentiment Classification. The study positions the model as a noteworthy contributor to sentiment analysis methodologies. In [27], the focus is on mining social information based on people's opinions through ML-based sentiment classification. The measured outcome revolves around Sentiment Classification, showcasing the model's capacity to distill sentiments from vast social datasets.[28] employs a multi-view learning framework for unified news representation in online articles. The measured outcome, Topic Categories of Online News Articles, highlights the model's ability to discern and categorize diverse topics within online news. In [29], web-based sentiment analysis holds significant value for corporations. The Outcome Measured includes Topics Correlated With Opinions: Positive, Negative, shedding light on the model's utility in understanding the alignment of opinions with specific topics. [30] modifies neural algorithms to capture five major insights in sentence-level documents. The Outcome Measured, Sentiment Classification, underlines the model's success in distilling sentiments from diverse documents. [31] emphasizes the superior performance of word embeddings combined with standard classifiers, particularly in Sentiment Analysis of Tweets. This novel approach indicates a departure from traditional sentiment analysis methods. BNgram method's superior performance on large Twitter datasets is highlighted in [32], focusing on Sentiment Classification Metrics. The study offers a meticulous analysis of sentiment classification outcomes. In [33], bert-base-cased emerges as the best-performing model for predicting sentiment in news headlines. The Outcome Measured, Sentiment Analysis of News Headlines, underscores the model's efficacy in analyzing sentiments within news contexts. [34] leverages topic distributions as features for sentiment analysis on documents, emphasizing Classification Accuracy. The study positions topic distributions as pivotal in enhancing sentiment analysis outcomes. BERT's capability in context embedding is illuminated in [35], focusing on Sentiment Analysis using BERT. The Outcome Measured, Classification Accuracy, signifies BERT's unique contribution to contextualizing sentiments within diverse texts. [36] reports that Bernoulli naive Bayes and Convolutional Neural Network achieve higher accuracy rates than other algorithms in headline classification. The Outcome Measured, Sentiment Polarity of News Headlines, underscores the model's success in distilling nuanced sentiments from headlines. BERT-based models demonstrate significant improvements in sentence classification across various platforms in [37], focusing on Sentence Classification. The study illustrates the versatility of BERT-based models in capturing and categorizing sentiments at the sentence level. [38] showcases multiple BERT-based models that improve text classification across different platforms, emphasizing Text Classification. The study highlights the adaptability of BERT-based models in diverse text analysis scenarios. In [39], deep bidirectional transformers enhance sentiment classification accuracy, recall, and F1 score. The Outcome Measured, Sentiment Classification, emphasizes the model's success in achieving nuanced sentiment analyses. [40] introduces a sentiment analyzer predicting news article sentiment at both sentence and document levels. The Outcome Measured includes Sentiment Analysis at Sentence Level and Sentiment Analysis at Document Level, reflecting the model's versatility. [41] introduces sentiment analysis software forecasting sentiment in news articles at both sentence and document levels based on topics. The Outcome Measured includes Sentiment Analysis at Sentence and Document Level, emphasizing the model's holistic approach. [42] identifies challenges in applying product review sentiment analysis to online news. The Outcome Measured includes Topics, Specific Sentiments, and Evolution of Topics and Sentiments, providing a comprehensive understanding of the challenges in cross-domain sentiment analysis. In [43], Bernoulli Naive Bayes technique exhibits higher accuracy in headline-based sentiment classification. The Outcome Measured includes Sentiment Orientation of News Headlines, signifying the model's proficiency in capturing nuanced sentiments from headlines. Similar challenges are identified in [44], focusing on Topics and Specific Sentiments. The study underscores the persistent challenges in applying product review sentiment. Predicting an object's associated emotion within a sentence is the aim of aspect-based sentiment categorization (ASC). While convolutional neural networks (CNN) and recurrent neural networks (RNN) are popular neural network-based techniques for ASC, recent work has started to incorporate syntactic features into graph neural networks (GNN) in order to address ASC challenges. Nevertheless, noise and ineffective utilization of syntactic dependency tree information frequently impede these methods. To overcome these drawbacks, a unique GNN-based deep learning model is presented in this study. The model reduces noise and makes better use of syntactic dependency tree information by building a part-of-speech (POS) guided syntactic dependency graph for a relational graph attention network (RGAT). Furthermore, to properly extract semantic relationships between surrounding words, a densely connected graph convolutional network (DCGCN) is equipped with a syntactic distance attention-guided layer. Tests conducted on three publicly available datasets show how successful the suggested model is; when compared to the current baselines, it reaches state-of-the-art performance. By constructing a novel POS-guided graph that eliminates unnecessary POS words, the new framework, called Dual Graph Neural Network (DGNN), accurately utilizes fine-grained syntactic information and extracts contextual feature representations, improving processing efficiency in graph neural networks [45]. For the vast majority of people in India, since it is the main source of revenue, agriculture is essential to the country's development. Timely and accurate predictions of crop output are vital for directing investments and forming agricultural policy. Crop rotation is one technique that improves soil fertility, but forecasting and predicting crop yields can greatly increase agricultural production. Crop yields, however, decline when farmers are ignorant of the nutrients and composition of the soil. The suggested approach creates an ensemble deep learning system to forecast rice crop yield based on soil nutrient levels in order to overcome these problems. The system makes use of databases on crop production statistics and soil nutrients, where crop production statistics show the yield in particular area. To fill up the gaps, pre-processing methods like mean attribute and normalization are used. The suggested approach makes use of Model Agnostic Meta-Learning (MAML), a stacking-based ensemble deep learning strategy that combines the outputs of three classifiers: Support Vector Machine (SVM), Deep Belief Network (DBN), and Deep Neural Network (DNN). Based on the specified soil, MAML then generates the final rice crop production projection. The technique predicts agricultural yields with a high accuracy of 89.5%, indicating its efficacy [46]. According to the DSM-5, antisocial behavior (ASB) is a type of personality disorder that is associated with disorders such as borderline, histrionic, and narcissistic personality disorders. It is characterized by a persistent disrespect for the rights of others. Since ASB can result in victims experiencing despair, anxiety, low self-esteem, and suicide ideation, it is a serious social and public health concern. Sites like Reddit and Twitter may unintentionally encourage this kind of conduct. This study suggests a proactive strategy to identify and address ASB on social media by utilizing deep learning and natural language processing, potentially averting serious consequences like suicide [47]. This work investigates the use of IoT big data and deep learning algorithms for the optimization of e-commerce logistics node design, and it suggests an enhanced plan. The study looks at five factors: node location and traffic conditions, information transmission accuracy, distribution efficiency, transportation costs, and assessments from e-commerce businesses. The better arrangement, according to the results, lowers transportation expenses for e-commerce companies, boosts information transmission accuracy by 4.34%, and improves logistics distribution efficiency by 3.69% [48].

3. News Topic and sentiment classification using BERT

Proposed research entails a comprehensive approach for the classification of news articles using BERT-based models. The process initiates with data loading and splitting into training and testing sets. First, a Topic Classification model is established, which involves BERT model initialization, text tokenization, label conversion, and the construction of TensorFlow datasets. This model is trained iteratively, evaluating performance and generating classification reports after each epoch. Following the topic classification, a Sentiment Analysis model is set up with similar steps. The sentiment model is trained and evaluated in a parallel manner. Both models generate results and classification reports that are stored and can be visualized for insights. This methodology enables the effective analysis of news articles, categorizing them into topics and assessing sentiment, providing valuable insights into news content.

3.1 Data collection and Pre-processing

In the pursuit of conducting comprehensive sentiment analysis and topic classification on Indian news articles, we encountered a significant challenge - the unavailability of a ready-made dataset that encompassed both news topic categories and sentiment scores. To address this gap, we embarked on a multi-stage data collection and pre-processing journey, meticulously crafting our dataset from the ground up. This process consisted of four fundamental phases: web scraping, data collection, data pre-processing, and sentiment analysis.

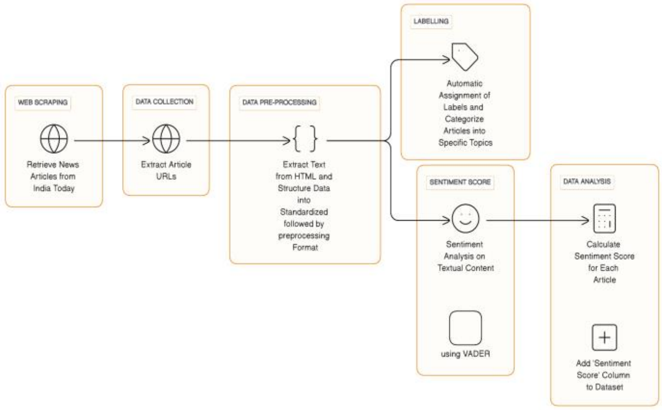

Figure 1. Steps to perform web scraping to create news dataset

Following steps are performed to create a news dataset with topic labelling and sentiment score as shown in figure 1. Web Scraping**:** Initially, our researchers undertook web scraping efforts to retrieve news articles directly from India Today's website, and in some cases, other news sources pertaining to India. This entailed the systematic extraction of article URLs, the downloading of article content, and the amalgamation of this data into a structured format.

Data Collection: Subsequent to web scraping, the data collection phase commenced. It involved the careful curation of articles, where raw HTML content was harvested, bearing in mind the importance of preserving the original structure and metadata. This wealth of data was then aggregated into a cohesive dataset.

Data Pre-processing: To ensure that our dataset was suitable for analysis, a pivotal pre-processing step was executed. This entailed cleaning the raw data, extracting pertinent text content from the HTML, and structuring it into a standardized format. This phase is vital in ensuring the quality and consistency of the dataset.

Labelling: To further enrich the dataset, labels were automatically assigned to each news article based on the categories available on the website. This allowed us to categorize news articles into specific topics for subsequent analysis.

Sentiment Score: Beyond these stages, we aimed to delve deeper into the dataset by conducting sentiment analysis on the textual content of the news articles. Sentiment analysis is a potent natural language processing technique employed to quantify the emotional tone or sentiment expressed within a given text. For this purpose, we harnessed the capabilities of VADER (Valence Aware Dictionary and sentiment Reasoned), a sentiment analysis tool adept at interpreting sentiment in text, particularly in the context of news articles. The sentiment analysis process involved the computation of a sentiment score for each article within the dataset. The sentiment score represents a single numerical value that encapsulates the overall sentiment within the text, considering both positive and negative sentiments. The authors specifically made use of VADER's compound score for their work, which amalgamates various sentiment scores into one comprehensive value. This compound score ranges between -1 and 1, with -1 denoting highly negative sentiment, 1 signifying highly positive sentiment, and 0 indicating a neutral sentiment. To execute the sentiment analysis, we designed a Python script to load the dataset from a CSV file, initialized the SentimentIntensityAnalyzer from NLTK's VADER package, and subsequently applied a sentiment analysis function to each article's text. This function computed the sentiment score for each article and stored the outcomes in a dedicated column labelled 'sentiment score' within our dataset.

3.2 Topic and sentiment classification of news using BERT

The proposed model efficiently handles two classification tasks, Topic Classification and Sentiment Analysis, with a clear separation between the two as shown in figure 2. It leverages BERT's powerful language modelling capabilities and TensorFlow for building, training, and evaluating the models. The results obtained from this workflow can be valuable for understanding and categorizing news articles based on both their topics and sentiment.

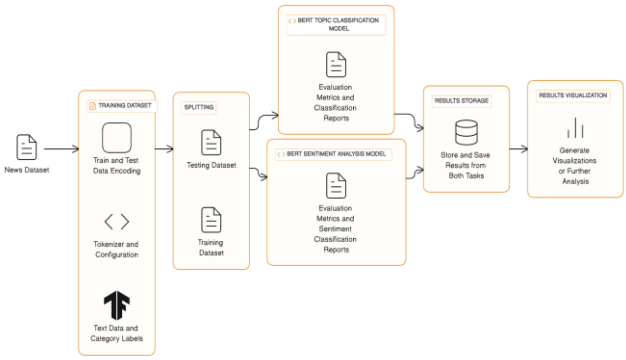

Figure 2. Topic and sentiment classification of news using BERT

3.2.1 Data Loading and Splitting:

The code starts by loading a news dataset from a CSV file, which is assumed to have 'Text', 'Category', and 'Sentiment' columns. It splits the dataset into training and testing sets with an 80-20 split ratio.

3.2.2 Topic Classification Model Setup:

In this segment of the code, the setup for the Topic Classification model is established to categorize news articles into different topics. The steps involved are as follows:

The first step involves initializing the BERT model tailored for Topic Classification, utilizing the "bert-base-uncased" variant. Simultaneously, the code initializes the BERT tokenizer to preprocess the text data. The number of labels for the classification task aligns with the count of unique categories present in the dataset.

Following the model initialization, the text undergoes tokenization and encoding using the BERT tokenizer. This process ensures that the text data is segmented into words or subwords comprehensible to BERT. Additionally, the encoded data includes special tokens for padding purposes, with a predetermined maximum sequence length set at 128.

Subsequently, the category labels, representing different topics, are converted into numpy arrays to facilitate their utilization within the model architecture.

To facilitate the training process, TensorFlow datasets are constructed. These datasets encompass the encoded text data and their corresponding topic labels, organizing the information required for model training.

Finally, the model undergoes compilation, which involves configuring an optimizer, defining a suitable loss function, and selecting an evaluation metric crucial for the model's training and assessment.

3.2.3 Topic Classification Training Loop:

The loop overseeing the training of the Topic Classification model operates for a designated number of epochs. Throughout this process: Training involves shuffling and batching the training dataset for model training. The optimizer is utilized to minimize the cross-entropy loss, and model accuracy is continuously monitored during the training phase. Following each training epoch, the model's performance undergoes evaluation on the test dataset. Essential evaluation metrics, such as loss and accuracy, are systematically recorded to gauge the model's effectiveness. Results and Reports are integral to this process, with the code generating classification reports. These reports provide precision, recall, and F1-score metrics for each category involved in the topic classification task. The outcomes are systematically stored in a list for further analysis and reference.

3.2.4 Sentiment Analysis Model Setup:

This section lays the groundwork for the Sentiment Analysis model designed to assess the sentiment of news articles. To initiate the BERT model for sentiment analysis, a process akin to setting up the topic classification model is employed. Sentiment analysis, in this context, involves a binary classification task, distinguishing between positive and negative sentiments. Following the model initialization, the text data undergoes tokenization and encoding using the BERT tokenizer. This process involves incorporating special tokens and padding to facilitate subsequent analysis. The sentiment labels, representing positive or negative sentiments, are then transformed into numpy arrays, preparing them for integration into the model's input. Subsequently, TensorFlow datasets are constructed, incorporating the encoded text data and the converted sentiment labels. These datasets serve as the foundational elements for the model's training and evaluation processes.

3.2.5 Sentiment Analysis Training Loop:

The code generates classification reports for sentiment analysis, including precision, recall, and F1-score for positive and negative sentiment. The Sentiment Analysis model undergoes training for a specified number of epochs, akin to the topic classification loop. During this process: Training involves the optimization of the model using the sentiment analysis dataset, where the optimizer works to minimize the cross-entropy loss. The model's accuracy is continually monitored and recorded throughout the training phase. Following each training epoch, the model's performance undergoes evaluation using the test dataset. Metrics such as loss and accuracy are computed and documented for assessment. Subsequently, the code generates comprehensive classification reports for sentiment analysis. These reports encompass precision, recall, and F1-score metrics specifically tailored for both positive and negative sentiment classifications.

4. Performance Evaluation

The evolution of performance throughout this process signifies a series of systematic improvements in the Topic Classification model's setup. Initiated by the precise initialization of the BERT model, specifically tailored for Topic Classification using the "bert-base-uncased" variant, the evolution progresses to encompass efficient text tokenization and encoding via the BERT tokenizer. This transformation ensures the segmentation of text into BERT-comprehensible units, integrating special tokens for padding while maintaining a concise maximum sequence length of 128. Further, the conversion of category labels into numpy arrays streamlines their integration within the model architecture. Ultimately, the culmination of these meticulous steps, including the construction of TensorFlow datasets and the comprehensive model compilation, denotes a holistic evolution aimed at refining the model's performance, optimizing its efficiency, and facilitating its adeptness in classifying news articles into distinct topics.

4.1 Experiment Setup

Our dataset, constructed from a careful and rigorous process of web scraping, data collection, and pre-processing, serves as the cornerstone for our research, enabling sentiment analysis and topic classification on Indian news articles. It comprises a diverse array of news articles obtained from India Today's website, as well as other reputable sources dedicated to Indian news. Each article is meticulously structured and labelled according to its specific news category.

In this study, author organized the dataset into training and testing subsets using a stratified split. Around 80% of the articles comprised the training dataset, employed to train our sentiment analysis and topic classification models. The remaining 20% formed the testing dataset, serving as an independent evaluation set to gauge the model's performance. This stratified method aimed to maintain a proportional representation of diverse news categories in both training and testing sets, ensuring dataset diversity. We adjusted two hyper parameters: the first, epochs, determined the number of training iterations the model underwent through the entire dataset. The second, batch size, defined the quantity of training examples processed together in each training round, impacting training efficiency and hardware requirements.

To gauge the effectiveness of our sentiment analysis and topic classification models, we rely on a set of comprehensive evaluation metrics. For sentiment analysis, the performance of our models is assessed using standard metrics such as accuracy, precision, recall, F1-score, and the receiver operating characteristic (ROC) curve. These metrics enable us to measure the accuracy of sentiment predictions, identify false positives and false negatives, and assess the overall model performance. Additionally, for topic classification, we employ metrics including categorical cross-entropy, accuracy, and confusion matrices. These metrics provide insights into the model's ability to categorize news articles correctly and identify potential misclassifications.

4.2 Results and discussion

Multitask learning using BERT has proven to be effective for both topic classification and sentiment analysis of news articles. By jointly training the model on two related tasks, we harness the ability of BERT to understand contextual information and relationships in text, which has led to significant improvements in both tasks. The tasks of topic classification and sentiment analysis share some common linguistic features. Leveraging the pre-trained BERT model allows the network to transfer knowledge between these related tasks. This transfer learning has resulted in a model that performs well on both tasks, even though they are distinct. Multitask learning has enhanced the model's ability to generalize across different categories and sentiments. This is particularly important in a real-world scenario where news articles may cover various topics and exhibit a range of sentiments. The training process demonstrates that the model is making substantial progress, with training accuracy consistently improving. Overall, the model's increasing accuracy and decreasing loss indicate that it is learning the topic classification task effectively. Continuing the training process, monitoring validation results, and potentially fine-tuning hyper parameters will likely result in an even more robust and accurate model for topic classification.

The results obtained from the multitask learning model, which combines topic classification and sentiment analysis using BERT, are highly promising. In topic classification, the model showcases an impressive accuracy of 98.4%, indicating its ability to accurately categorize news articles into eight distinct categories, including environment, health, education, tech, sports, business, lifestyle, and science. Moreover, the model's macro average F1-Score of 92.9% and weighted average F1-Score of 94.04% reflect a well-balanced trade-off between precision and recall across all categories, indicating its reliability and effectiveness in this task. In sentiment analysis, the model maintains a strong accuracy of 94%, enabling it to effectively differentiate between positive and negative sentiment in news articles. The model's macro average F1-Score of 81% and weighted average F1-Score of 85% underscore its proficiency in sentiment classification. Overall, these results highlight the model's robustness in understanding and classifying news articles, considering both their content and emotional tone. Insights gained from multitask learning.

The model's performance in multitask learning is noteworthy for several reasons. First, the joint training on related tasks, topic classification, and sentiment analysis, demonstrates the power of leveraging pre-trained BERT representations to learn from textual data effectively. The ability to understand context and relationships in text enables the model to excel in both tasks. Furthermore, the high accuracy achieved in topic classification is indicative of the model's capacity to categorize news articles accurately, a crucial requirement for information retrieval and recommendation systems. In sentiment analysis, the model performs well in distinguishing between positive and negative sentiment, which is valuable for assessing the overall sentiment conveyed in news articles. However, further improvements may be explored, potentially by incorporating more granular sentiment categories or through fine-tuning of hyper parameters. Overall, the results demonstrate that multitask learning with BERT is a promising approach for understanding and classifying news articles across categories and sentiments, offering significant potential for real-world applications in the field of news analysis and recommendation.

Table 1. represents the results of News Topic Classification model, employing BERT architecture, demonstrates a promising performance across epochs. Initially, in the early epochs (1 and 2), the precision and recall values hover around 75%-84%, indicating moderate performance. However, from epoch 3 onwards, substantial improvements are evident, notably in the F1-Score which steadily increases. Epoch 5 stands out with an impressive F1-Score of 94%, showcasing the model's ability to effectively balance precision and recall, particularly notable in handling diverse classes. Although epoch 4 exhibits a drop in recall, the subsequent epochs showcase consistent and high precision-recall balance, notably in epochs 6 to 10, suggesting robust learning and convergence. Overall, the model demonstrates exceptional accuracy (98%), signalling its proficiency in correctly classifying topics, particularly after fine-tuning in later epochs. However, attention is warranted in instances, like epoch 4, where recall drops significantly, possibly indicating specific challenges in certain class distinctions that merit further investigation or targeted improvements.

Table 1. Result summary of news topic classification using Bert

Epoch	Precision(%)	Recall(%)	F1-Score(%)
1	75	72	74
2	84	72	78
3	87	92	90
4	91	59	71
5	91	97	94
6	96	89	92
7	97	87	92
8	97	87	92
9	97	87	92
10	97	87	92
Accuracy	-	-	98

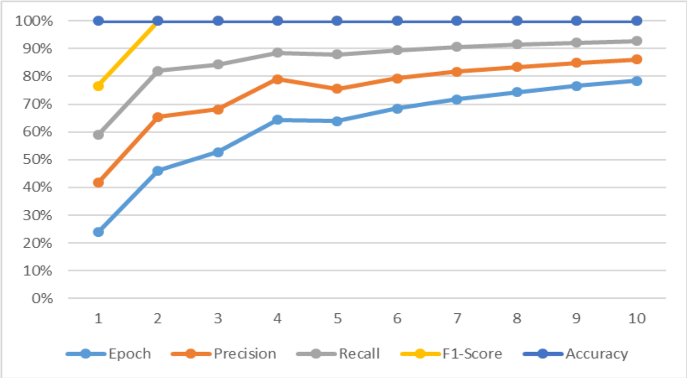

Figure 3. Result summary of news topic classification using Bert

Figure 3 represents the plotted graph demonstrates the evolution of precision, recall, and F1-score metrics across multiple epochs in news topic classification. Initially, precision starts at 75% in epoch 1 and gradually increases to reach a stable high of 97% from epoch 6 onwards, showcasing consistent and significant improvements. In contrast, recall shows more variability, peaking at 97% in epoch 5 but fluctuating notably between epochs 3 to 7 before stabilizing. F1-score exhibits a similar pattern to precision, showing a consistent increase from 74% in epoch 1 to approximately 92% in epoch 6, followed by a plateau. The support remains relatively constant throughout the epochs, indicating consistent representation across classes. The convergence of precision, recall, and F1-score towards higher values, coupled with steady support, Indicates the model's capacity to generalize and accurately categorize news topics as the training advances.

The sentiment analysis model employing BERT architecture showcases an overall promising performance across multiple epochs. Initially, as shown in table 2, in the early stages, the model demonstrated moderate precision around 51% in epoch 1, progressively improving to 72% by epoch 6. A similar trend is observed in recall, starting at 85% in the initial epochs and stabilizing around 88-90% in later epochs. The F1-score, which balances precision and recall, depicts a consistent upward trajectory, starting at 64% and reaching a stable 90% by epoch 7, remaining consistent through epochs 8 to 10. This suggests the model's robustness in effectively classifying sentiments within the dataset. The accuracy stands at an impressive 94%, indicating the model's high proficiency in accurately categorizing sentiments across all epochs. Overall, the incremental improvements in precision, recall, and F1-score culminate in a notably high accuracy, underlining the model's efficacy in sentiment classification while maintaining stable performance across multiple epochs.

Table 2. Result summary of new sentiment classification using Bert

Epoch	Precision(%)	Recall(%)	F1-Score(%)
1	51	85	64
2	57	85	68
3	67	76	71
4	69	76	72
5	71	69	70
6	72	70	71
7	90	90	90
8	91	88	90
9	91	88	90
10	91	88	90
Accuracy			94

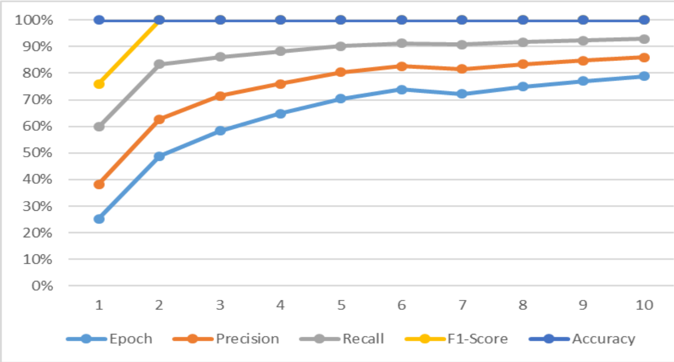

Figure 4. Result summary of news topic classification using Bert

Figure 4 illustrates the performance of a sentiment analysis model over various epochs using BERT for news sentiment classification. Initially, the model demonstrates moderate precision at 51% in epoch 1, steadily progressing to 72% by epoch 6. Concurrently, recall starts at 85% and stabilizes around 88-90% from epochs 7 to 10. The F1-score, a balance between precision and recall, exhibits a consistent upward trend, starting at 64% and culminating at a stable 90% from epochs 7 onwards. This graph signifies the model's incremental enhancement in precision, recall, and ultimately, F1-score, suggesting its improved ability to correctly classify sentiments as training progresses. The stability and convergence of these metrics in the later epochs indicate the model's robustness and effectiveness in sentiment classification tasks, resulting in a high accuracy of 94% across all epochs.

5. Conclusion

This research presents a comprehensive methodology for news article classification using BERT-based models, simultaneously addressing topic classification and sentiment analysis. The research significantly enriches the domains of natural language processing and news analysis, offering a comprehensive comprehension of news content and its conveyed emotions. The applications of this work extend to content recommendation, trend analysis, and monitoring public sentiment, offering valuable insights to journalists, media houses, and researchers. In this endeavour, Researchers made use of a multitask model that was BERT-powered to address the dual challenges of news category and sentiment classification. With a rich dataset comprising 3263 records across eight diverse categories, this model exhibits exceptional performance. Its precision and recall balance in topic classification, resulting in an accuracy of 98%, highlights its proficiency in classifying news articles, while sentiment analysis with 94% accuracy

underscores its ability to differentiate emotional nuances within these articles. This study elucidates the significance of multitask learning, enabling us to glean holistic insights from news data, catering to both the categorization and sentiment aspects, thus positioning it as a valuable tool in news analysis and recommendation systems. While this methodology showcases robust performance, future work may focus on fine-tuning models, exploring diverse pre-trained BERT variants, and enhancing dataset quality. Further research avenues could include real-time data integration and user feedback to create an adaptive, dynamic news analysis system. This research lays a strong foundation for advancing news analysis techniques in the future, making it an exciting area of study with continued potential for growth and innovation.

References

[1]. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Bidirectional Encoder Representations from Transformers. arXiv preprint arXiv:1810.04805.

[2]. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. In Advances in neural information processing systems (pp. 30-48).

[3]. Howard, J., & Ruder, S. (2018). Universal language model fine-tuning for text classification. arXiv preprint arXiv:1801.06146.

[4]. Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Zettlemoyer, L. (2019). Roberta: A robustly optimized bert pretraining approach. arXiv preprint arXiv:1907.11692.

[5]. Xie, Z., Xiao, Y., Wang, T., Zhou, B., Lin, Z., & An, L. (2021). Towards Transparent and Controllable Attention Mechanisms in NLP. arXiv preprint arXiv:2102.11941.

[6]. Kang, J. Y., Lee, S. H., & Jang, S. (2021). A personalized news recommendation model based on

multi-task learning. Expert Systems with Applications, 165, 114118

[7]. R. Khandelwal, A. Nayak, H. Harkous and K. Fawaz. "CookieEnforcer: Automated Cookie Notice Analysis and Enforcement". Jan. 2022.

[8]. H. Zankadi, A. Idrissi, N. Daoudi and I. Hilal. "Identifying learners' topical interests from social media content to enrich their course preferences in MOOCs using topic modeling and NLP techniques". Nov. 2022.

[9]. M. F. Mridha, M. A. H. Wadud, A. Hamid, M. M. Monowar, M. Abdullah-Al-Wadud and A. Alamri. "L-Boost: Identifying Offensive Texts From Social Media Post in Bengali". Jan. 2021.

[10]. M. Schirmer, U. Kruschwitz and G. Donabauer. "A New Dataset for Topic-Based Paragraph Classification in Genocide-Related Court Transcripts". Jan. 2022.

[11]. Y. Wang, Q. Chen, and W. Wang, 'Multi-task BERT for Aspect-based Sentiment Analysis', in 2021 IEEE International Conference on Smart Computing (SMARTCOMP), 8 2021.

[12]. Shiwang Huang, Xiaoyu Wang, Xiaohan, Ji, Jing Xie, and Qin Tang, 'Network News Sentiment Analysis Based on BERT'.

[13]. Sarojadevi Palani, P. Rajagopal, and Sidharth Pancholi, 'T-BERT - Model for Sentiment Analysis of Micro-blogs Integrating Topic Model and BERT'.

[14]. M. F. Abdussalam, D. Richasdy, and M. A. Bijaksana, 'BERT Implementation on News Sentiment Analysis and Analysis Benefits on Branding', JURNAL MEDIA INFORMATIKA BUDIDARMA, vol. 6, no. 4, p. 2064, Oct. 2022.

[15]. G. Li et al., 'A BERT-based Text Sentiment Classification Algorithm through Web Data', in 2022 International Conference on Computer Engineering and Artificial Intelligence (ICCEAI), 7 2022.

[16]. Samir, S. M. Elkaffas, and M. M. Madbouly, 'Twitter Sentiment Analysis Using BERT', in 2021 31st International Conference on Computer Theory and Applications (ICCTA), 2021.

[17]. Z. Gao, A. Feng, X. Song, and X. Wu, 'Target-Dependent Sentiment Classification With BERT', IEEE Access, vol. 7, pp. 154290–154299, 2019.

[18]. 'Online News Monitoring and Sentiment Analysis using BERT Approach', International Journal of

[31]. Notes in Computer Science, Springer International Publishing, 2019, pp. 181–193.

[32]. K. Cai, S. Spangler, Y. Chen, and L. Zhang, 'Leveraging sentiment analysis for topic detection', Web Intelligence and Agent Systems: An International Journal, vol. 8, no. 3, pp. 291–302, 2010.

[33]. D. Rajput and S. Verma, 'An Attention Arousal Space for Mapping Twitter Data', in Lecture Notes in Electrical Engineering, Springer Singapore, 2020, pp. 381–395.

[34]. S. D. Tembhurnikar and N. N. Patil, 'Topic detection using BNgram method and sentiment analysis on twitter dataset', in 2015 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions), 9 2015.

[35]. Singh and G. Jain, 'Sentiment Analysis of News Headlines Using Simple Transformers', in 2021 Asian Conference on Innovation in Technology (ASIANCON), 2021.

Advanced Research in Technology and Innovation, Jan. 2023.

[19]. S. R. Pingili and L. Li, 'Target-Based Sentiment Analysis using a BERT Embedded Model', in 2020 IEEE 32nd International Conference on Tools with Artificial Intelligence (ICTAI), 11 2020.

[20]. J. Lei, Q. Zhang, J. Wang, and H. Luo, 'BERT Based Hierarchical Sequence Classification for Context-Aware Microblog Sentiment Analysis', in Neural Information Processing, Springer International Publishing, 2019, pp. 376–386.

[21]. V. Yadav and S. Shakya, 'Sentiment Analysis and Topic Modeling on News Headlines', Journal of Ubiquitous Computing and Communication Technologies, vol. 4, no. 3, pp. 204–218, Sep. 2022.

[22]. R. Man and K. Lin, 'Sentiment Analysis Algorithm Based on BERT and Convolutional Neural Network', in 2021 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC), 2021.

[23]. G. Wei, 'Research on Internet Text Sentiment Classification Based on BERT and CNN-BiGRU', in 2022 11th International Conference on Communications, Circuits and Systems (ICCCAS), 2022.

[24]. Jinbin Cai, Fei Chen, and Si-xuan Chen, 'Sentiment Classification Based On BERT'.

[25]. Florian Bütow, Florian Schultze, and Leopold Strauch, 'Semantic Search : Sentiment Analysis with Machine Learning Algorithms on German News Articles'. .

[26]. M. Agarwal, P. K. Chaudhary, S. K. Singh, and C. Vij, 'Sentiment Analysis Dashboard for Socia Media comments using BERT', in 2023 International Conference on Advancement in Computation & Computer Technologies (InCACCT), 2023.

[27]. Bello, S.-C. Ng, and M.-F. Leung, 'A BERT Framework to Sentiment Analysis of Tweets', Sensors, vol. 23, no. 1, p. 506, Jan. 2023.

[28]. Xiaohong Cai, Hui Cao, and Jin-gang Ma, 'Sentiment Analysis of E-commerce Comments Based on BERT'.

[29]. 'Performance based Machine Learning Algorithm for Topic Oriented Text Categorization', International Journal of Recent Technology and Engineering, vol. 8, no. 2S11, pp. 3501–3506, Nov. 2019.

[30]. C. Wu, F. Wu, T. Qi, Y. Huang, and X. Xie, 'Title-Aware Neural News Topic Prediction', in Lecture

[36]. D.Deepa, 'Bidirectional Encoder Representations from Transformers (BERT) Language Model for Sentiment Analysis task: Review'.

[37]. S. S. Hossain, Y. Arafat, and M. E. Hossain, 'Context-Based News Headlines Analysis: A Comparative Study of Machine Learning and Deep Learning Algorithms', Vietnam Journal of Computer Science, vol. 08, no. 04, pp. 513–527, Apr. 2021.

[38]. H. Batra, N. S. Punn, S. K. Sonbhadra, and S. Agarwal, 'BERT-Based Sentiment Analysis: A Software Engineering Perspective', in Lecture Notes in Computer Science, Springer International Publishing, 2021, pp. 138–148.

[39]. J. Zheng, X. Chen, Y. Du, X. Li, and J. Zhang, 'Short Text Sentiment Analysis of Micro-blog Based on BERT', in Lecture Notes in Electrical Engineering, Springer Singapore, 2019, pp. 390–396.

[40]. S. Kaman, 'News Sentiment Analysis By Using Deep Learning Framework', May 2020.

[41]. P. Liu, J. A. Gulla, and L. Zhang, 'Dynamic Topic-Based Sentiment Analysis of Large-Scale Online

News', in Web Information Systems Engineering -- WISE 2016, Springer International Publishing, 2016, pp. 3–18.
[42]. S. Rahman, S. S. Hossain, S. Islam, M. I. Chowdhury, F. B. Rafiq, and K. B. M. Badruzzaman, 'Context-Based News Headlines Analysis Using Machine Learning Approach', in Computational Collective Intelligence, Springer International Publishing, 2019, pp. 167–178.
[43]. P. Liu, J. A. Gulla, and L. Zhang, 'RETRACTED ARTICLE: A joint model for analyzing topic and sentiment dynamics from large-scale online news', World Wide Web, vol. 21, no. 4, pp. 1117–1139, Jul. 2017.
[44]. X. Zhang, Z. Wu, K. Liu, Z. Zhao, J. Wang, and C. Wu, 'Text Sentiment Classification Based on BERT Embedding and Sliced Multi-Head Self-Attention Bi-GRU', Sensors, vol. 23, no. 3, p. 1481, Jan. 2023.
[45]. Xiao, L., Xue, Y., Wang, H., Hu, X., Gu, D., & Zhu, Y. (2022). Exploring fine-grained syntactic information for aspect-based sentiment classification with dual graph neural networks. Neurocomputing, 471, 48-59.
[46]. Chandraprabha, M., and Rajesh Kumar Dhanraj. "Ensemble Deep Learning Algorithm for Forecasting of Rice Crop Yield based on Soil Nutrition Levels." EAI Endorsed Transactions on Scalable Information Systems 10, no. 4 (2023).
[47]. Singh, R., Subramani, S., Du, J., Zhang, Y., Wang, H., Miao, Y., & Ahmed, K. (2023). Antisocial Behavior Identification from Twitter Feeds Using Traditional Machine Learning Algorithms and Deep Learning. EAI Endorsed Transactions on Scalable Information Systems, 10(4).
[48]. Li, L. (2023). Deep Learning Algorithm Aided E-Commerce Logistics Node Layout Optimization Based on Internet of Things Network. EAI Endorsed Transactions on Scalable Information Systems, 10(4), e16-e16.