Research Journal of Engineering Technology and Medical Sciences (ISSN: 2582-6212), Volume 06, Issue 03, September-2023
Available at www.rjetm.in/

Sentiment Analysis from Social Media Contents using Attention based Machine Learning

Roli Shrivastava[1], Shatendra Kumar Dubey[2]
[1,2]Department of Information Technology, NRI Institute of Information Science and Technology, Bhopal (M.P), India
[1]roli.rlrl@gmail.com, [2]shatendradubey@gmail.com

* Corresponding Author: Roli Shrivastava

Abstract: *Sentiment analysis, also known as opinion mining, has evolved as a crucial field in natural language processing, with applications spanning from market trend forecasting to gauging public sentiment on critical issues. This paper presented an analyzes various machine learning techniques employed in sentiment analysis, encompassing both binary and multi-class sentiment classification. Additionally, the paper proposes a novel language-independent attention-based learning model (LI-ABLM) for sentiment analysis and presents an in-depth analysis of its results. The proposed LI-ABLM model demonstrates impressive accuracy, outperforming traditional algorithms and setting a new state-of-the-art in sentiment analysis.*

Keywords: *Sentiment analysis, attention mechanisms, deep learning, social media platforms, Classification.*

1. Introduction

The concept of "sentiment analysis" can be traced back to research from the early 2000s, which aimed to understand and forecast market trends based on textual evaluations. Over time, the term has become more versatile, often used interchangeably with "opinion mining," as it covers broader applications like detecting sentiment and subjectivity within textual content. In practice, sentiment analysis can be broken down into different granular levels: document-level, sentence-level, and feature-level analyses. While document and sentence-level analyses offer a broad-stroke picture of the writer's sentiment, they fall short in identifying specific subjects or topics that the sentiment is directed towards. This limitation paves the way for more nuanced, feature-level analyses. The growth of web 2.0 has led to an influx of user-generated content, providing a wealth of data for various fields to examine and gain insights from. However, challenges abound, including interpreting text laden with slang, emoticons, or ambiguous elements like sarcasm and irony.
Given the rise of social networking platforms, a growing body of research is devoted to sentiment analysis within the context of social media. In this domain, data scientists and researchers continuously explore the abundant user-generated content for various applications.For instance, one study utilized sentiment data from platforms like Amazon, TripAdvisor, Facebook, and Twitter to train classifiers that can identify different sentiment polarities. Their research demonstrated high accuracy rates in classifying reviews, and even showed promise in real-world applications like real-time traffic monitoring. Therefore, sentiment analysis has not only evolved as a field but also proven to be instrumental in leveraging big data for problem-solving and decision-making in the modern world. Sentiment analysis serves various applications beyond consumer goods and services. It has piqued interest as a tool for gauging public sentiment on issues like government policies, elections, and national events. However, the reliability of these methods remains a subject of debate, as illustrated by inaccurate predictions in some election forecasts. In the corporate world, companies are increasingly employing sentiment analysis to assess critical business metrics, such as brand perception, reputation, and customer engagement. For example, one study scrutinized customer feedback on ride-sharing apps by extracting comments from social media platforms.

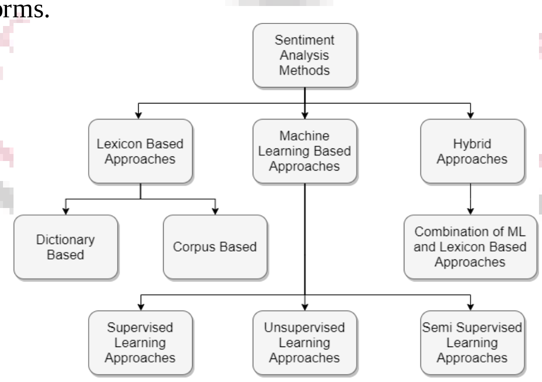

Figure 1: Sentiment analysis Methods

2. Literature Review

Chen et al. [1] introduced a deep sentiment analysis framework that leverages CNNs combined with LSTM. This model employs dual CNN layers to identify text's nuanced features. Post feature extraction, these are introduced into the LSTM to ascertain contextual data. By integrating this advanced deep learning method with a one-vs-rest training strategy, it's applied to multi-faceted sentiment classification. Testing revealed the model's accuracy to be 78.42% on dataset D1, surpassing the performance of existing models like SVMs, CNN, LSTM, and CNN-LSTM.

Ruz et al. [2] turned to the Bayes factor methodology, which led to the creation of more realistic networks. Their findings showcased the Bayes factor's potency and its superior predictive prowess in contrast to support vector machines and random forests, especially with an ample training data set. Additionally, the resulting networks highlighted word relationships, providing qualitative insights into event dynamics from both historical and social perspectives.

Arora et al. [3] embarked on the examination of health-related tweets, focusing on sentiments linked to Depression and Anxiety. They utilized both the Multinomial Naive Bayes and the Support Vector Regression (SVR) algorithm for classification.

Katchapakirin et al. [4] adopted NLP methodologies to devise a depression identification algorithm tailored for Thai language Facebook posts. With data from 35 Facebook participants, their findings indicated a strong correlation between Facebook activity and potential depression levels.

A unique model designed to predict anxious depression based on real-time tweets was put forth by Kumar et al. [5]. They trained this model using a trio of classifiers: multinomial naive bayes, gradient boosting, and random forest. The final verdict is determined via majority voting facilitated by an ensemble voting classifier. Initial tests on tweets from a sample of 100 users showcased the model's promising classification accuracy of 85.09%.

Cornn and colleagues [6] examined various machine learning techniques, such as logistic regression and support vector machines, a BERT-centered approach, and neural networks both with and without word embeddings (specifically CNN). The most efficient was the CNN model that didn't utilize word embeddings, achieving an impressive accuracy of 92.5% after just four epochs. The BERT-centric model closely followed, registering an accuracy of 85.7%.

Ahmad et al. [7] focused on devising a method to analyze content related to terrorism. Their primary objective was to categorize tweets into either extremist or non-extremist brackets. Utilizing user-generated content from Twitter, they crafted a tweet classification mechanism rooted in deep learning sentiment analysis techniques. This system's goal was to discern between extremist and non-extremist content. Their preliminary results were promising and pave the way for future academic pursuits in this realm.

Giakwad et at. [8] introduced an innovative technique that leverages a lexicon database to attribute a value termed as the 'Impact Factor' to each word within a text. This value signifies the influence of an individual word on the broader context of its containing sentence. Every word has its unique Impact Factor, indicating its sway over the sentence's overarching meaning. Words with a high Impact Factor exert greater influence. This method merges lexicon-centric techniques with machine learning. It employs the AFINN lexicon database for Impact Factor allocation and harnesses algorithms like SVM, KNN, and Naive Bayesian for model training and evaluation.

Jayakrishnan et al. [9] employed an SVM classifier to facilitate multi-emotional classification of Malayalam sentences. Their proposed methodology incorporates varied syntactic attributes like n-gram, POS markers, negation cues, level indicators, and more to enhance classification outcomes. This classifier can categorize Malayalam phrases into diverse emotional categories, such as happiness, sadness, anger, fear, and neutrality, while also providing intensity details (e.g., high or low). Additionally, the system can identify the sentence's nature – be it a dialogue, query, or other – ensuring an improved auditory experience when the content is relayed via a speech synthesizer.

Solakidis and his team [10] presented a comprehensive system capable of analyzing user-created texts to determine both their polarity (neutral, negative, or positive) and associated sentiment (joy, love, anger, or sadness). Their semi-supervised method resulted in two strategies for autonomously gathering training data, eliminating human oversight. Their methodology identifies self-descriptive features within the data, focusing on emoticons and emotionally charged keywords. When testing on a well-known forum using different classifiers and feature sets, the findings indicated various insights into the efficiency, benefits, and constraints of these methods when applied to Greek content. Keywords led to a 90% average accuracy in pinpointing subjectivity and 93% in determining polarity. In contrast, emoticons yielded average accuracies of 74% and 77% for the same measures, respectively.

Based on Multi class Sentiment Analysis

Tanna et al. [14] are working towards streamlining social media reactions and posts analysis with the introduction of a versatile social media platform. This platform enables users to post, like, comment, share, and even distribute content across other social platforms, essentially providing a one-stop solution encompassing the features of several social media platforms.

Arora et al. [15] delved into the analysis of tweets specifically focusing on mental health indicators like Depression and Anxiety. They leveraged the Multinomial Naive Bayes and Support Vector Regression (SVR) Algorithm for this classification task.

Asad et al. 16] have crafted a model grounded in data analytics to identify signs of depression in individuals. Through the utilization of machine learning, they processed data acquired from SNS users and employed Natural Language Processing (NLP) paired with Support Vector Machine (SVM) and Naïve Bayes algorithms to optimize the detection process.

Lyu's et al. [17] incorporated sentiment analysis with text-mining techniques to understand public sentiments towards child abuse incidents on the Weibo platform. The overarching aim was to discern public emotions and perceptions, potentially influencing the development and refinement of child protection policies in China.

Tariq et al. [18] introduced a unique method aimed at classifying individuals with chronic mental illnesses such as Anxiety, Depression, Bipolar, and ADHD using data from Reddit. Their co-training-based classification showcased improved performance by an average of 3% compared to contemporary techniques.

Jabreel's et al. [19] unveiled a pioneering deep learning system tailored for the intricate task of multi-label emotion classification on Twitter. They ingeniously transformed this into a binary classification challenge and applied deep learning techniques, surpassing previous benchmarks by reaching a notable accuracy score of 0.59 on the SemEval2018 Task 1.

Bouzazi et al. [20] put forth a groundbreaking technique that not only tackles binary and ternary sentiment classifications but also delves deeper, categorizing tweets into multiple sentiment classes. Even though the study was limited to seven sentiment classes, the method is scalable. The method showcased remarkable accuracy levels: 81.3% in binary classification (after excluding neutral tweets) and 70.1% in ternary classification.

Imran et al. [21] assessed the sentiments of individuals from diverse backgrounds regarding the novel Coronavirus and the measures implemented by nations. They employed deep LSTM models to gauge sentiment and emotional tone from tweets, achieving benchmark accuracy with the sentiment140 dataset. A notable aspect of their approach was the innovative use of emoticons as a validation tool for their supervised deep learning models applied to Twitter data.

3. Proposed Methodology

The suggested technique is comprised of three main components, which are outlined below (Figure 2).

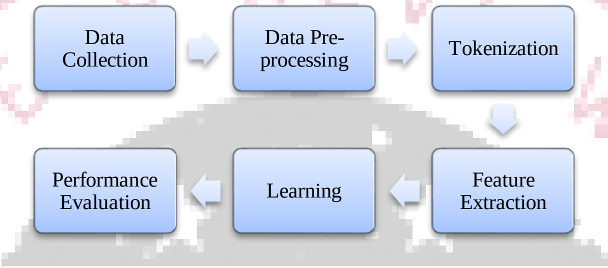

Figure 2: Proposed Methodology

3.1 Data Gathering

This stage prepares a dataset for various domains acquired from various domains. For further processing, the data is gathered from a variety of publicly accessible data sources.

3.2 Data Pre-processing

It is essential to filter the original information obtained from various sources. Data pre-processing is the phrase for this stage, as illustrated in figure 3. During pre-processing, any unneeded words in the review, such as commas and special symbols, are deleted since they don't add to any emotion values in the phrase or document. The acquired dataset is evaluated item by entity, and superfluous entities such as URLs, special symbols, commas, and other punctuation marks are deleted, leaving a clean dataset for future processing.

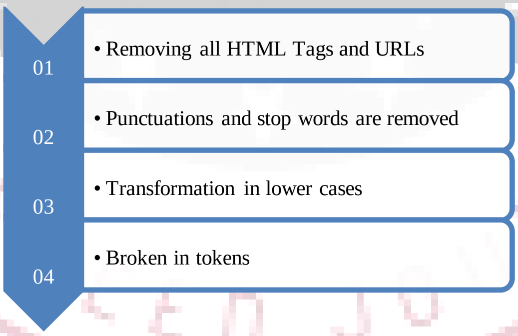

Figure 3: Data Pre-processing

3.3 Tokenization and Feature Extraction

In the context of Natural Language Processing (NLP), feature extraction and tokenization are critical preprocessing steps for tasks like sentiment analysis. First of all tokenization is performed. In this step, the proposed model use a weighted combined features using following, as presented in figure 3:

3.4 Classification

In this section, language independent attention-based learning model (LI-ABLM) is presented for analysing the tone of written content by discovering hidden patterns of coherence between words. Long short-term memory network (LSTM) was developed to address the problems, and it has been proven to be effective in learning sequence representations.

Proposed model architecture comprises an input layer that accepts sequences of integers, each representing a word token. These are transformed by a non-trainable Embedding layer into vectors of a predefined size. The model utilizes a combination of bi-directional GRU and LSTM layers, with 512 units for each type of layer, to capture both short-term and long-term dependencies in the sequence data. To mitigate overfitting, a Spatial Dropout layer with a dropout rate of 0.2 is applied after the Embedding layer. We further augment the model with an Attention layer, designed to weigh the significance of different parts of the input sequence. Following the attention mechanism, a Dropout layer with a rate of 0.2 is applied to further reduce overfitting. The sequence is then passed through a Dense layer of 64 units with ReLU activation, followed by another Dropout layer. The output layer consists of a Dense layer with three units employing a softmax activation function for classification. The model is compiled using the Adam optimizer and employs Binary Cross-Entropy as its loss function. The model's performance is gauged based on accuracy.

4. Result Analysis

In this section, paper presents the implementation details of sentiment analysis using social media content. For this a twitter dataset is prepared by collecting reviews or tweets from multiple sources. For training and testing samples are divided into 70:30 ratio. Testing was performed using random 30% samples taken from entire datasets. The whole implementation was done on python platform.

In this work, the proposed model is evaluated on twitter datasets collected from [14] and [15].

Text categorization and sentiment analysis tasks frequently employ these criteria. The following formula is used to compute these criteria:

Accuracy = Accuracy refers to the condition of a model efficiency being right or exact in all circumstances.

$$\frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

Precision = Precision refers to the correctness of the data being examined. It's a metric for determining if a positive occurrence is accurate or not. Presented by:

$$\frac{TP}{TP+FP} \tag{2}$$

Recall =: The frequency of positive experiences is measured by recalls. The Recall formula is as follows:

$$\frac{TP}{TP+FN} \tag{3}$$

F1 score = It is defined as the inverse of accuracy and recall multiplied by two, as provided by the equation.:

$$\frac{2}{1/precision+1/Recall} \tag{4}$$

Where TP, TN, FP, FN are true positive, true negative, false positive and false negative instances, respectively.

Table 1 Performance Evaluation

Parameters	Values
Accuracy	93.19
Precision	94
Recall	93
F1-Score	94
Execution Time	2.56s

Table 1 presents key performance efficiency of the model. A value of 93.19 suggests that the model correctly classifies approximately 93.19% of the instances, which indicates high accuracy.

Figure 4 shows the training and validation accuracy graph of the model and it shows that approx. 90% of accuracy was achieved by the model during training and validation. Whereas figure 5 shows the training and validation loss of the model and it shows approx. 0.15.

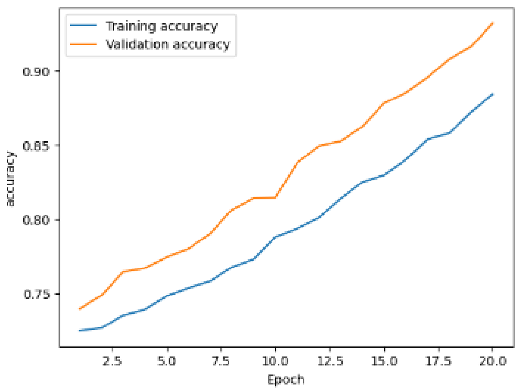

Figure 4: Training and Validation Accuracy

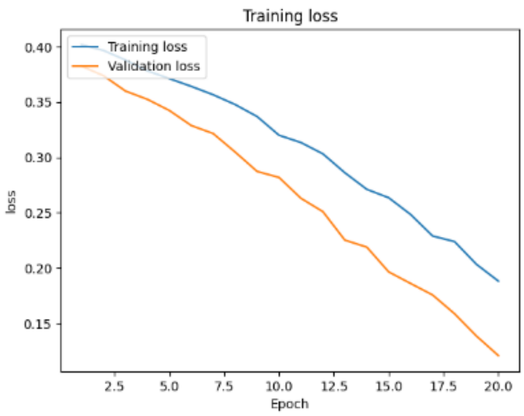

Figure 5: Training and Validation Loss

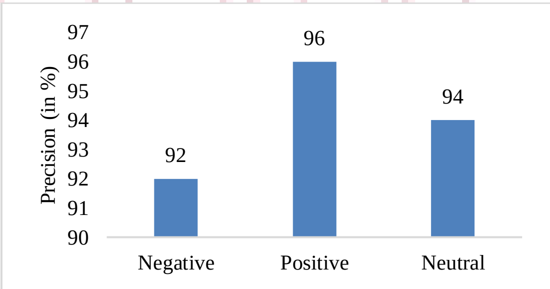

Figure 6: Precision Evaluation

The figure 6 focuses on the precision values for a classification model that classifies tweets into Negative, Positive, and Neutral categories. This means that when the model predicts a tweet to be negative, it is correct 92% of the time. For positive tweets, the model excels with a precision score of 96. This indicates that the model is highly accurate when classifying tweets as positive, being correct 96% of the time.

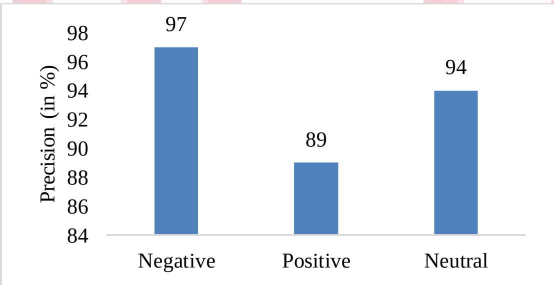

Figure 7: Recall Evaluation

The figure 7 presents the recall values for a classification model that categorizes tweets into three classes: Negative, Positive, and Neutral. With a recall value of 97, the model is extremely effective at correctly identifying negative tweets.

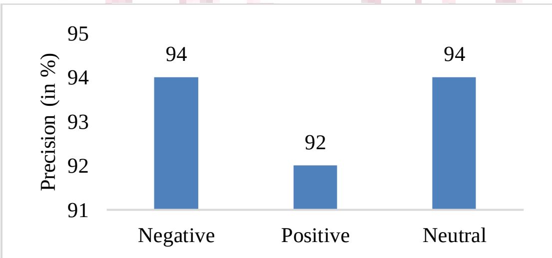

Figure 8: F1-Score Evaluation

The figure 8 presents the f1-score values for a classification model that categorizes tweets into three classes: Negative, Positive, and Neutral. With a f1-score value of 94, the model is extremely effective at correctly identifying negative tweets.

In summary, the proposed LI-ABLM model sets a new state-of-the-art in terms of accuracy, followed closely by Multinomial Naive Bayes [10]. Traditional algorithms like Logistic Regression and Random Forest also perform competitively, but KNN appears to be less suitable for this task.

Table 2: Comparative State-of-Art

Models	Accuracy (%)
SVM [1]	74.18
Naive Bayes [2]	74.2
Multinomial Naive Bayes [3]	78.0
Random Forest [5]	81.04
Logistic Regression [6]	84.8
KNN [7]	72.0
LI-ABLM (Proposed)	93

5. Conclusion

Sentiment analysis has emerged as a versatile tool with applications in diverse domains, including social media, mental health assessment, and public opinion tracking. This paper has provided a comprehensive overview of various machine learning techniques employed in sentiment analysis, highlighting their strengths and weaknesses. The proposed LI-ABLM model, designed for sentiment analysis, has showcased exceptional accuracy, achieving a 93.19% accuracy rate in classifying tweets into Negative, Positive, and Neutral categories. This performance surpasses traditional algorithms such as SVM, Naive Bayes, and Logistic Regression, making LI-ABLM a promising advancement in the field of sentiment analysis. While the LI-ABLM model demonstrates significant promise, there is room for further research and improvement. Future work could focus on expanding the model's capabilities to handle sentiment analysis in multiple languages, enhancing its generalizability, and exploring applications beyond social media sentiment analysis. Overall, sentiment analysis continues to evolve, driven by innovative models like LI-ABLM, and holds substantial potential for practical applications in various industries and fields.

References

[1] B. Chen, Q. Huang, Y. Chen, L. Cheng, and R. Chen, "Deep Neural Networks for Multi-class Sentiment Classification," Proc. - 20th Int. Conf. High Perform. Comput. Commun. 16th Int. Conf. Smart City 4th Int. Conf. Data Sci. Syst. HPCC/SmartCity/DSS 2018, pp. 854–859, Jun. 2018, doi: 10.1109/HPCC/SMARTCITY/DSS.2018.00142.

[2] P. Arora and P. Arora, "Mining Twitter Data for Depression Detection," 2019 Int. Conf. Signal Process. Commun. ICSC 2019, pp. 186–189, Mar. 2019, doi: 10.1109/ICSC45622.2019.8938353.

[3] G. A. Ruz, P. A. Henríquez, and A. Mascareño, "Sentiment analysis of Twitter data during critical events through Bayesian networks classifiers," Futur. Gener. Comput. Syst., vol. 106, pp. 92–104, 2020, doi: https://doi.org/10.1016/j.future.2020.01.005.

[4] K. Katchapakirin, K. Wongpatikaseree, P. Yomaboot, and Y. Kaewpitakkun, "Facebook Social Media for Depression Detection in the Thai Community," 2018 15th Int. Jt. Conf. Comput. Sci. Softw. Eng., 2018, doi: 10.29007/TSCC.

[5] A. Kumar, A. Sharma, and A. Arora, "Anxious Depression Prediction in Real-time Social Data A R T I C L E I N F O", Accessed: Aug. 04, 2023. [Online]. Available: http://ssrn.com/link/ICAESMT-2019.html=xyz

[6] K. Cornn, "Identifying Depression on Social Media".

[7] S. Ahmad, M. Z. Asghar, F. M. Alotaibi, and I. Awan, "Detection and classification of social media-based extremist affiliations using sentiment analysis techniques," Human-centric Comput. Inf. Sci., vol. 9, no. 1, pp. 1–23, Dec. 2019, doi: 10.1186/S13673-019-0185-6/TABLES/12.

[8] G. Gaikwad and D. J. Joshi, "Multiclass Mood classification on twitter using lexicon dictionary and machine learning algorithms," Proc. Int. Conf. Inven. Comput. Technol. ICICT 2016, vol. 1, Jul. 2016, doi: 10.1109/INVENTIVE.2016.7823247.

[9] G. S. Solakidis, K. N. Vavliakis, and P. A. Mitkas, "Multilingual Sentiment Analysis Using Emoticons and Keywords," 2014 IEEE/WIC/ACM Int. Jt. Conf. Web Intell. Intell. Agent Technol., vol. 2, pp. 102–109, Oct. 2014, doi: 10.1109/WI-IAT.2014.86.

[10] R. Jayakrishnan, G. N. Gopal, and M. S. Santhikrishna, "Multi-Class Emotion Detection and Annotation in Malayalam Novels," 2018 Int. Conf. Comput. Commun. Informatics, ICCCI 2018, Aug. 2018, doi: 10.1109/ICCCI.2018.8441492.

[11] M. Deshpande and V. Rao, "Depression detection using emotion artificial intelligence," in 2017 International Conference on Intelligent Sustainable Systems (ICISS), 2017, pp. 858–862. doi: 10.1109/ISS1.2017.8389299.

[12] R. Chatterjee, R. K. Gupta, and B. Gupta, "Depression Detection from Social Media Posts Using Multinomial Naive Theorem," IOP Conf. Ser. Mater. Sci. Eng., vol. 1022, no. 1, Jan. 2021, doi: 10.1088/1757-899X/1022/1/012095.

[13] R. Jose and V. Chooralil, "Prediction of election result by enhanced sentiment analysis on twitter data using classifier ensemble Approach," 2016, pp. 64–67. doi: 10.1109/SAPIENCE.2016.7684133.

[14] S. Tariq et al., "A Novel Co-Training-Based Approach for the Classification of Mental Illnesses Using Social Media Posts," IEEE Access, vol. 7, pp. 166165–166172, 2019, doi: 10.1109/ACCESS.2019.2953087.

[15] D. Tanna, M. Dudhane, A. Sardar, K. Deshpande, and N. Deshmukh, "Sentiment Analysis on Social Media for Emotion Classification," 2020 4th Int. Conf. Intell. Comput. Control Syst., pp. 911–915, May 2020, doi: 10.1109/ICICCS48265.2020.9121057.

[16] P. Arora and P. Arora, "Mining Twitter Data for Depression Detection," 2019 Int. Conf. Signal Process. Commun., pp. 186–189, Mar. 2019, doi: 10.1109/ICSC45622.2019.8938353.

[17] N. Al Asad, M. A. Mahmud Pranto, S. Afreen, and M. M. Islam, "Depression Detection by Analyzing Social Media Posts of User," 2019 IEEE Int. Conf. Signal Process. Information, Commun. Syst. SPICSCON 2019, pp. 13–17, Nov. 2019, doi: 10.1109/SPICSCON48833.2019.9065101.

[18] Y. W. Lyu, J. C. C. Chow, and J. J. Hwang, "Exploring public attitudes of child abuse in mainland China: A sentiment analysis of China's social media Weibo," Child. Youth Serv. Rev., vol. 116, Sep. 2020, doi: 10.1016/J.CHILDYOUTH.2020.105250.

[19] M. Jabreel and A. Moreno, "A Deep Learning-Based Approach for Multi-Label Emotion Classification in Tweets," Appl. Sci. 2019, Vol. 9, Page 1123, vol. 9, no. 6, p. 1123, Mar. 2019, doi: 10.3390/APP9061123.

[20] M. Bouazizi and T. Ohtsuki, "A Pattern-Based Approach for Multi-Class Sentiment Analysis in Twitter," IEEE Access, vol. 5, pp. 20617–20639, Aug. 2017, doi: 10.1109/ACCESS.2017.2740982.

[21] A. S. Imran, S. M. Daudpota, Z. Kastrati, and R. Batra, "Cross-Cultural Polarity and Emotion Detection Using Sentiment Analysis and Deep Learning on COVID-19 Related Tweets," Ieee Access, vol. 8, p. 181074, 2020, doi: 10.1109/ACCESS.2020.3027350.

[22] F. Yao and Y. Wang, "Domain-specific sentiment analysis for tweets during hurricanes (DSSA-H): A domain-adversarial neural-network-based approach," Comput. Environ. Urban Syst., vol. 83, p. 101522, Sep. 2020, doi: 10.1016/J.COMPENVURBSYS.2020.101522.