Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 6s (2025)

# Hybrid Framework for Twitter Data Sentiment Classification using MLP-Deep Learning

***[1]L. Sudha Rani, [2]S. Zahoor-Ul-Huq, [3]C. Shoba Bindu**

*[1]Research Scholar, Department of CSE, Jawaharlal Nehru Technological University Anantapur (JNTUA), Anantapur 515002, India. Email: sudha10211@gmail.com

[2]Professor, Department of CSE, G. Pulla Reddy Engineering College, Kurnool, India. Email: szahoor@gmail.com

[3]Professor, Department of CSE, Jawaharlal Nehru Technological University Anantapur (JNTUA), Ananthapur, India. Email: shobabindhu.cse@jntua.ac.in

***Article History:***

***Received:*** *14-10-2024*

***Revised:*** *29-11-2024*

***Accepted:*** *10-12-2024*

**Abstract:**

Organizations collect data on user opinions from social media, enabling research into trends and behavior. Classifying text by identifying sentiments within content is complex due to varying expressions and contexts. Sentence-level analysis, often employed by traditional approaches, overlooks key aspects hidden in the data. A hybrid framework integrating Twitter-based feature selection with sentiment classification using a multi-layer perceptron (MLP) model is introduced in this paper. Techniques such as Support Vector classifiers and Random Forests are examined to enhance the accuracy of classification. Baseline models are compared with the proposed framework, which demonstrates reliable precision and improved classification in experiments across several datasets. The structured hybrid approach directly addresses challenges that arise when analyzing content from social media platforms.

**Keywords:** Twitter Data; Sentiment Analysis; Deep Learning; Word Embedding; Hybrid Framework

## 1. INTRODUCTION

Sentiment analysis identifies attitudes and opinions about objects or entities through text. Extracting precise sentiments can be difficult due to varying expressions across natural language. This process often operates at the document level to detect overall positive or negative polarity, but it may miss nuanced expressions and multiple opinions [1-2]. To address this limitation, sentence-level sentiment analysis examines individual sentences while considering all involved entities. Before performing this analysis, it distinguishes between objective and subjective sentences [3].

Though general polarity detection is useful, it can overlook detailed sentiments, resulting in incomplete analysis. Sentiments tied to specific aspects—such as product price, quality, or features—are identified through aspect-based sentiment analysis (ABSA) [4]. Some sentences may contain multiple positive or negative opinions, but standard models fail to capture this complexity. ABSA extracts explicit aspects directly from text and implicit aspects through advanced linguistic methods. Tools such as POS tagging and dependency parsing facilitate this process, although results vary when handling uncommon words or context-based sentiments.

Various hybrid approaches have been explored to enhance ABSA. Kim & Park introduced a framework combining machine learning with lexicons [5]. In another study, Wu et al. developed a hybrid method integrating rule-based and statistical strategies [6]. Other researchers, including Yan et al., applied node-ranking techniques to extract both implicit and explicit aspects [7]. Some studies employed point-wise mutual information to detect implicit aspects but only managed to obtain explicit features without context [8-10].

The paper is structured as follows: Section 2 reviews related work on sentiment analysis specific to Twitter data. Section 3 outlines the proposed framework, while Section 4 presents experimental comparisons with existing models. Section 5 concludes with insights from the research.

## 2. LITERATURE SURVEY

Researchers explored various strategies for text classification in aspect-based sentiment analysis (ABSA). These strategies fall into three primary categories: lexicon-based, machine learning, and hybrid approaches. Different challenges in processing opinions within text are addressed by each of these strategies.

### 2.1 Lexicon based approaches

Without needing a training dataset, lexicon-based methods assign sentiment scores to words or phrases. Typically, these scores range from -1 to 1, reflecting negative to positive sentiments. A study by Abadeh, Mowlaei, and Keshavarz [11] showed that performance drops when models struggle with unfamiliar lexicons. To resolve this issue, the study proposed two-generation procedures to enhance sentiment extraction. However, context-based opinions and a growing number of terms still posed challenges for these methods.

### 2.2 Machine Learning Approaches

Unsupervised techniques analyze sentiments using unlabeled data, but the results have often been inconsistent [12]. Alternatively, supervised learning employs labeled datasets to train algorithms such as Naive Bayes (NB), Maximum Entropy (ME), and Support Vector Machine (SVM). Information Gain and Chi-Square techniques refine accuracy by identifying relevant features. In another study, a feedback framework helped improve the precision of text classification by focusing on multiple aspects [13]. Despite these efforts, SVM suffers from higher time complexity, making it less efficient than NB or ME.

### 2.3 Artificial Neural Network

In sentiment analysis, artificial neural networks (ANNs) have been widely applied. A dynamic convolutional neural network (CNN) model introduced in 2014 demonstrated potential for capturing complex sentiments [14]. Further research compared additional CNN models, which performed competitively against the original dynamic CNN [15]. Studies explored architectures like LSTM and deeper CNNs, discovering that multiple aspects of sentiment can be identified more accurately [16]. A neural network model was later proposed to handle noise in sentiment data, ensuring better sentiment classification [17].

### 2.4 Hybrid Approach for Sentiment Classification

In [18], the authors proposed a hybrid method that combines the attributes of POS tags with common-sense knowledge to analyze the intelligence of customers. In [19], the authors proposed a targeted-aspect-based method. In [20], the authors proposed a hybrid LSTM framework that combines the features of the Sentic LSTM and the recurrent addictive network. They were able to extract the implicit features by using five techniques.

While these methods are useful, they struggle with low-frequency words that carry context-specific meaning [21]. Despite progress in ABSA, existing models need further development to classify complex sentiments reliably [22-23]. This paper introduces a hybrid framework integrating Twitter-specific features with a multi-layer perceptron (MLP) model for text classification. By combining machine learning and deep learning, the framework performs multi-level sentiment detection and addresses challenges in analyzing opinions from social media platforms. In 2018, Ibrahim, Zahiruddin, and Salehmat [24] presented a hybrid method that combines the attributes of POS tags with the PCA+SVM framework. They were able to extract accuracy of 76.55%, 71.62%, and 74.24% for three datasets: the STS dataset, the STC dataset, and the Twitter attributes. The main objective of this research was to develop a method that can improve the performance of the proposed framework.

## 3. PROPOSED FRAMEWORK

We present a novel method for analyzing sentiment in a Twitter dataset. It takes into account the various levels of single and multi-word sentences and performs text pre-processing. Then, it discovers the aspects using the ARM method. The feature is powered by the Stanford DTT approach and the POS patterns. The hybrid method for sentiment detection employs a rule-based approach, which is followed by a feature ranking process, and a Principal Component Analysis framework. It then produces a classification of sentiments using a multi-level approach. In Figure 1, the framework shows a detailed representation of a multi-level approach for sentiment analysis in Twitter data.

### 3.1 Data Gathering

This study used a total of three datasets: the Stanford Twitter Sentiment dataset, the Twitter Airline Sentiment dataset, the Sanders Twitter Corpus dataset. The STS dataset was collected using the KNIME tool, which was used by Ibrahim, Zeinuddin, and Selamat in 2018 [24]. It is having seven categories and contains 177 negative and 180 positive tweets. The second dataset, which is called TAS, consists of five categories and contains 5,741 negative and 1,832 positive tweets. The STC dataset was composed of various categories such as Apple, Google, and Microsoft. For these experiments, we considered over 1,091 tweets.

**Table 1:** Twitter Datasets Properties

| Sl. No. | Dataset | Negative | Positive | Considered Tweets |
|---|---|---|---|---|
| 1 | Stanford Twitter Sentiment dataset (STS) | 177 | 180 | 357 |
| 2 | Twitter Airline Sentiment dataset (TAS) | 5741 | 1832 | 7573 |
| 3 | Sanders Twitter Corpus dataset (STC) | 572 | 519 | 1091 |

Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 6s (2025)

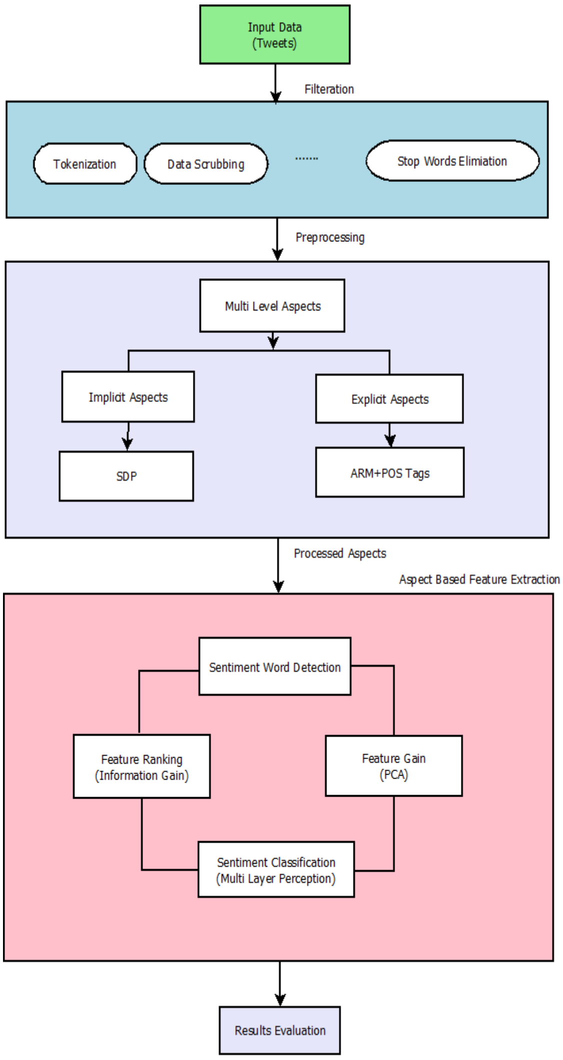


**Figure 1:** Multi-level Approach for Sentiment Analysis in Twitter Data

### 3.2 Data Transformation

Raw datasets often contain unnecessary elements such as usernames, links, emoticons, and smilies. Before analysis, these irrelevant elements are removed to avoid disrupting the classification process. Essential steps in pre-processing ensure smooth text processing by cleaning the data across multiple stages. The process involves eight key actions. It begins with eliminating web addresses and links, followed by tokenization, lemmatization, and case normalization. Word removal, stemming, and deletion of redundant data complete the cleanup, preparing the dataset for classification.

### 3.3 Aspect-Based Feature Extraction

Feature extraction simplifies large datasets by selecting relevant aspects. Two kinds of aspects—explicit and implicit—are extracted in this framework. POS tagging and rule mining methods handle explicit aspects, while implicit aspects are identified using the Stanford Dependency Parser (SDP). These methods work together to detect single and multi-word aspects effectively.

### 3.3.1 Stanford Dependency Parser for Implicit Aspects

Identifying implicit aspects can be more complex since they are not explicitly stated. Opinions and related aspects within sentences are connected using SDP, which detects their relationships. For instance, the phrase "the tooth pain got stale" reveals patterns such as "the DT," "pain JJ," and "stale NN." Paths like "nsubj-advmod" map adjectives and verbs (e.g., "bad RB," "had VBN") to implicit aspects. When the sentence reads "had really bad tooth pain," it identifies the aspect "pain tooth," associated with the dentist category in the STS dataset. By exploring patterns and dependencies, this approach discovers new implicit aspects efficiently.

### 3.3.2 Associate Rule Mining for Explicit Aspects

ARM identifies frequent patterns by analyzing relationships between elements in text. To ensure accuracy, the method uses Apriori algorithms with minimum support and confidence values. Items appearing frequently are detected and analyzed based on user-defined thresholds. Through ARM, single-word aspects such as "google," "united flight," and "nike" were extracted. The pattern NN-VBG identified "gop debate," while NN-JJ revealed "apple iphone." When generating multi-word aspects, phrases like "jet blue flight" followed the DT-JJ structure, demonstrating ARM’s capability to uncover complex textual patterns.

### 3.4 Hybrid Text Classification

This section presents a hybrid classification model that integrates rule-based methods and machine learning algorithms. It identifies sentiment words while extracting opinions from the given datasets.

### 3.4.1 Sentiment word detection using rule-based method

The rule-based method captures aspects overlooked by SDP. It identifies relevant sentiment words and extracts their polarity values and relationships with aspects. For instance, the STS dataset contains a tweet: “Worked harder, now only one exam left, and I feel so happy, will have fun soon, anxiously waiting.” Using relationships such as "aux" and "xcomp," the SDP method extracts "will fun" and "feel happy." Algorithm 1 identifies sentiment words based on their position relative to aspects, evaluating them within a defined distance. This method ensures that each sentiment word contributes accurately to classification outcomes.

**Algorithm 1: Rule Based Method**

For each tweet:

1. Analyze the aspect in the sentence.
2. If the aspect matches:
   - Retrieve the sentence word.
   - If the sentiment word’s distance $\leq 4$:
     - Add it to the results.
   - Else:
     - Exclude the word.

3. For each result:
   - o Compute the sentiment word value.
   - o Display the result.

---

### 3.4.2 Feature selection of sentiment words

The framework uses information gain to rank features. PCA was selected for feature reduction after comparing it with other techniques, such as Latent Semantic Analysis and Genetic Algorithms. The feature selection process starts with transforming the dataset into a statistical model, followed by calculating covariance and extracting eigenvalues. Analysis is conducted by training the top vector to represent relevant features effectively.

### 3.4.3 Classification algorithm

Following feature selection, the text is classified using a multi-layer perceptron (MLP) model. This model includes an input layer, an output layer, and at least one hidden layer. Data flows between layers through feed-forward connections. Each predictor variable connects with neurons in the next layer, with the final hidden layer linking to the output. The model includes three hidden layers, each with a unique number of neurons and activation functions, to ensure accurate classification.

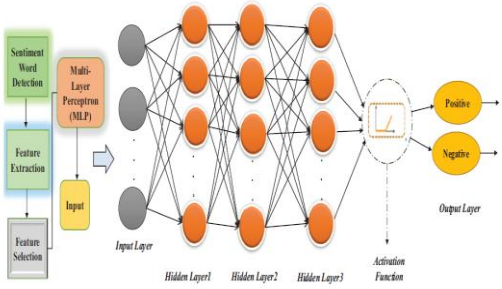


**Figure 2:** Multi-layer perception of the proposed framework

### 3.4.4 Activation functions

Activation functions help convert input into outputs through hidden layers. Differences in these functions' speed and efficiency determine their usage. Hyperbolic tangent functions range from -1 to 1, while sigmoid functions have an output between 0 and 1. The identity function outputs linearly, while ReLU addresses the vanishing gradient issue, making it a popular choice in deep learning models. Multiple configurations are tested using these functions to ensure optimized classification results.

$$f(x) = x \quad //identify \tag{1}$$

$$f(x) = 1/1 + \exp(-x) \quad //Logistic\,(or)\,Sigmoid \tag{2}$$

$$f(x) = 1-\exp(-2x)/1+\exp(-2x) \quad //Hyperbolic\,\mathrm{Tan}\,gent\ Function \tag{3}$$

$$R(x)=\max(o,x)\, i.e\, R(x)=\begin{cases}0 & if\ R<0\\ x & if\ R\geq 0\end{cases}\quad //\mathrm{Re}LU \tag{4}$$

## 4. RESULT ANALYSIS

The ASAMLP model was tested on three datasets: STS, TAS, and STC, as detailed in Table 1. In the first experiment, Association Rule Mining (ARM) extracted explicit aspects using nouns and verbs. For the second experiment, POS tags were combined with ARM to identify multi-word aspects. Different categories, including “north,” “yankees,” “cheney,” “jquery,” and “startrek,” were successfully classified by this approach. Improved results were obtained by integrating POS tagging with ARM.

Binary values—positive or negative—were used for sentiment classification. Precision and recall measured the accuracy of the classification system. A confusion matrix further analyzed performance using false positives, false negatives, and true negatives. Evaluating sentiment accuracy involved these detailed metrics.

The method was compared against models proposed by Kim [25] and Zhang et al. [26]. Figures 3, 4, and 5 display precision, recall, and F-score results for the STS, TAS, and STC datasets, respectively. Higher values were achieved by the proposed approach than those obtained by Kim [25] and Zhang et al. [26]. This outcome shows that combining ARM with POS tags enhances classification accuracy by capturing details that single methods might overlook.

**Table 2:** Analysis of the proposed Framework on STS Dataset

| Sl. No. | Name of the Approach | Features | Precision | Recall | F-Score | Accuracy |
|---|---|---|---|---|---|---|
| 1 | Proposed Method | POS tags + unigram | 0.851 | 0.847 | 0.848 | 84.72 |
| 2 | ASA+SVC | POS tags + unigram | 0.812 | 0.805 | 0.808 | 80.55 |
| 3 | ASA+RF | POS tags + unigram | 0.802 | 0.791 | 0.796 | 79.16 |
| 4 | ASA+KN | POS tags + unigram | 0.722 | 0.761 | 0.740 | 76.16 |
| 5 | ASA+AB | POS tags + unigram | 0.812 | 0.805 | 0.808 | 80.55 |
| 6 | Kim’s [25] | POS tags + unigram | 0.824 | 0.781 | 0.801 | 80.85 |
| 7 | Zhang et al.’s [26] | POS tags + unigram | 0.742 | 0.723 | 0.732 | 77.28 |

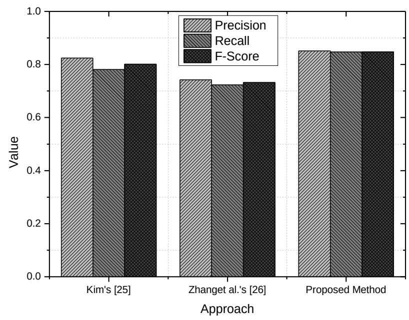


**Figure 3:** Precision, Recall and F-Score Value of STS Dataset

**Table 3:** Analysis of the proposed Framework on TAS Dataset

| Sl. No. | Name of the Approach | Features | Precision | Recall | F-Score | Accuracy |
|---|---|---|---|---|---|---|
| 1 | Proposed Method | POS tags + unigram | 0.833 | 0.84 | 0.836 | 84.09 |
| 2 | ASA+SVC | POS tags + unigram | 0.773 | 0.764 | 0.768 | 78.43 |
| 3 | ASA+RF | POS tags + unigram | 0.792 | 0.803 | 0.797 | 80.33 |
| 4 | ASA+KN | POS tags + unigram | 0.815 | 0.827 | 0.820 | 82.12 |
| 5 | ASA+AB | POS tags + unigram | 0.815 | 0.818 | 0.816 | 80.16 |
| 6 | Kim’s [25] | POS tags + unigram | 0.813 | 0.742 | 0.775 | 79.25 |
| 7 | Zhang et al.’s [26] | POS tags + unigram | 0.726 | 0.714 | 0.719 | 76.34 |

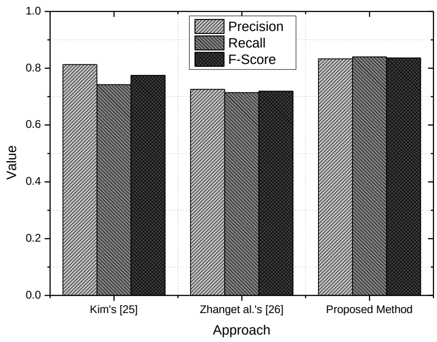


**Figure 5:** Precision, Recall and F-Score Value of TAS Dataset

**Table 4:** Analysis of the proposed Framework on STC Dataset

| Sl. No. | Name of the Approach | Features | Precision | Recall | F-Score | Accuracy |
|---|---|---|---|---|---|---|
| 1 | Proposed Method | POS tags + unigram | 0.79 | 0.789 | 0.789 | 78.99 |
| 2 | ASA+SVC | POS tags + unigram | 0.758 | 0.757 | 0.757 | 75.79 |
| 3 | ASA+RF | POS tags + unigram | 0.712 | 0.722 | 0.716 | 72.08 |
| 4 | ASA+KN | POS tags + unigram | 0.714 | 0.721 | 0.717 | 70.16 |
| 5 | ASA+AB | POS tags + unigram | 0.755 | 0.753 | 0.752 | 75.34 |
| 6 | Kim’s [25] | POS tags + unigram | 0.743 | 0.718 | 0.730 | 72.35 |
| 7 | Zhang et al.’s [26] | POS tags + unigram | 0.684 | 0.657 | 0.670 | 68.32 |

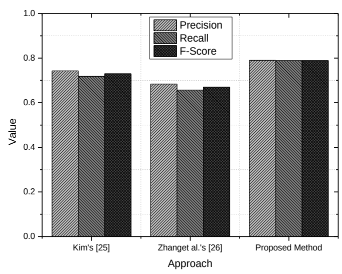


**Figure 6:** Precision, Recall and F-Score Value of STC Dataset

The proposed model achieved 84.72% accuracy in the STS dataset, with precision, recall, and F-score recorded as 0.851, 0.847, and 0.848, respectively (see Table 2). For the TAS dataset, the best accuracy observed was 84.09%, along with precision of 0.833, recall of 0.840, and an F-score of

0.836 (see Table 3). Using POS tags and unigram features, the model in the STC dataset attained an average accuracy of 78.99%. Precision, recall, and F-score for this dataset were 0.790, 0.789, and 0.789, respectively (see Table 4).

## 5. CONCLUSION

This study introduces a hybrid framework that merges aspect-based sentiment analysis (ABSA) with multi-level sentiment detection using Association Rule Mining (ARM). POS tagging is combined with heuristic rules to improve the extraction of sentiment-rich aspects from the data. Implicit aspects are detected using the Stanford Dependency Parser (SDP) by analyzing relationships between dependencies within sentences.

Feature selection is integrated into the framework to identify and rank relevant elements, improving classification outcomes. Deep learning techniques embedded in the model support multi-layer sentiment classification. This hybrid approach fills gaps in previous models by balancing linguistic patterns and statistical methods, resulting in detailed extraction of sentiment from complex datasets.

## REFERENCES

[1] Jang, Hyeju, Emily Rempel, David Roth, Giuseppe Carenini, and Naveed Zafar Janjua. "Tracking COVID-19 discourse on twitter in North America: Infodemiology study using topic modeling and aspect-based sentiment analysis." *Journal of medical Internet research* 23, no. 2 (2021): e25431.

[2] Ishaq, Adnan, Sohail Asghar, and SairaAndleeb Gillani. "Aspect-based sentiment analysis using a hybridized approach based on CNN and GA." *IEEE Access* 8 (2020): 135499-135512.

[3] Ay Karakuş B, Talo M, Hallaç İ.R, Aydin G. 2018. Evaluating deep learning models for sentimentclassification. Concurrency and Computation: Practice and Experience 30(21):e4783DOI 10.1002/cpe.4783.

[4] Bansal B, Srivastava S. 2019. Hybrid attribute based sentiment classification of online reviews forconsumer intelligence. Applied Intelligence 49(1):137–149 DOI 10.1007/s10489-018-1299-7.

[5] Tripathi, Milan. "Sentiment analysis of nepali covid19 tweets using nbsvm and lstm." *Journal of Artificial Intelligence* 3, no. 03 (2021): 151-168.

[6] Mowlaei, Mohammad Erfan, Mohammad SanieeAbadeh, and Hamidreza Keshavarz. "Aspect-based sentiment analysis using adaptive aspect-based lexicons." *Expert Systems with Applications* 148 (2020): 113234.

[7] Liu, Ning, and Bo Shen. "Aspect-based sentiment analysis with gated alternate neural network." *Knowledge-Based Systems* 188 (2020): 105010.

[8] Neelakandan, S., and D. Paulraj. "A gradient boosted decision tree-based sentiment classification of twitter data." *International Journal of Wavelets, Multiresolution and Information Processing* 18, no. 04 (2020): 2050027.

[9] Dang, NhanCach, María N. Moreno-García, and Fernando De la Prieta. "Sentiment analysis based on deep learning: A comparative study." *Electronics* 9, no. 3 (2020): 483.

[10] Peng, Haiyun, Lu Xu, Lidong Bing, Fei Huang, Wei Lu, and Luo Si. "Knowing what, how and why: A near complete solution for aspect-based sentiment analysis." In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 05, pp. 8600-8607. 2020.

[11] Mowlaei ME, Abadeh MS, Keshavarz H. 2020. Aspect-based sentiment analysis using adaptiveaspect-based lexicons. Expert Systems with Applications 148(C–1):113234

[12] Do HH, Prasad P, Maag A, Alsadoon A. 2019. Deep learning for aspect-based sentiment analysis: a comparative review. Expert Systems with Applications 118(6):272–299 DOI 10.1016/j.eswa.2018.10.003.

[13] Feldman R. 2013. Techniques and applications for sentiment analysis. Communications of the ACM 56(4):82–89 DOI 10.1145/2436256.2436274.

[14] Go A, Bhayani R, Huang L. 2009. Twitter sentiment classification using distant supervision. CS224N Project Report, Stanford 1(12):2009.

[15] HaCohen-Kerner Y, Miller D, Yigal Y. 2020. The influence of preprocessing on text classification using a bag-of-words representation. PLOS ONE 15(5):e0232525 DOI 10.1371/journal.pone.0232525.
[16] Hong J, Fang M. 2015. Sentiment analysis with deeply learned distributed representations of variable length texts. Stanford University Report. 1–9.
[17] Kalchbrenner N, Grefenstette E, Blunsom P. 2014. A convolutional neural network for modelling sentences. Available at https://arxiv.org/abs/1404.2188.
[18] Kastrati Z, Imran AS, Kurti A. 2020. Weakly supervised framework for aspect-based sentiment analysis on students2019 reviews of moocs. IEEE Access 8:106799–106810.
[19] Kaur D. 2017. Sentimental analysis on apple tweets with machine learning technique. International Journal of Science, Engineering and Computer Technology 7(9):76.
[20] Liu B. 2012. Sentiment analysis and opinion mining. Synthesis Lectures on Human Language Technologies 5(1):1–167.
[21] Liu B, Zhang L. 2012. A survey of opinion mining and sentiment analysis. In: Aggarwal C, Zhai C, eds. Mining Text Data. Boston: Springer, 415–463.
[22] Liu H, Dougherty ER, Dy JG, Torkkola K, Tuv E, Peng H, Ding C, Long F, Berens M, Parsons L, Yu L, Zhao Z, Forman G. 2005. Evolving feature selection. IEEE Intelligent Systems 20(6):64–76
[23] Liu N, Shen B. 2020. Aspect-based sentiment analysis with gated alternate neural network. Knowledge-Based Systems 188:105010.
[24] Zainuddin N, Selamat A, Ibrahim R. 2018. Hybrid sentiment classification on twitter aspectbasedsentiment analysis. Applied Intelligence 48(5):1218–1232.
[25] Kim, Y. Convolutional neural networks for sentence classification. arXiv 2014, arXiv:1408.5882
[26] Zhang, Y.; Zhang, Z.; Miao, D.; Wang, J. Three-way enhanced convolutional neural networks forsentence-level sentiment classification. Inf. Sci. 2019, 477, 55–64.