Frontiers in Health Informatics
ISSN-Online: 2676-7104
www.healthinformaticsjournal.com
2024; Vol 13: Issue 8
Open Access

Predicting E-Learning System Adoption Rates Using a Dense Convolutional Neural Network Integrated with BERT for Sentiment Analysis

B. Vivekanandam[1*]
[*] Email: vivekresearch2014@gmail.com

Mr. T. Manoj Praphakar[2]
[2]Assistant Professor, Sri Shakthi Institute of Engineering and Technology (Autonomous), Coimbatore, India
Email: Manojpsg.phd@gmail.com

Dr. M. Rajesh Babu[3]
[3]Professor & Head, Department of AI & DS, Kathir College of Engineering (Autonomous), Coimbatore, India
Email: drmrajeshbabu@gmail.com

Dr. T.Primya[4]
[4]Assistant Professor, Department of CSE,
KPR Institute of Engineering and Technology (Autonomous),
Coimbatore, India
Email: primyacse@gmail.com

Dr. T. Nandhini[5]
[5]Associate Professor, Department of MCA,
Moodlakatte Institute of Technology,
Kundapura, Karnataka, India
Email: nandhini.t13@gmail.com

Cite this paper as: B. Vivekanandam, Mr. T. Manoj Praphakar, Dr. M. Rajesh Babu, Dr. T.Primya, Dr. T. Nandhini (2024) Predicting E-Learning System Adoption Rates Using a Dense Convolutional Neural Network Integrated with BERT for Sentiment Analysis. *Frontiers in Health Informa 3978-3989*

Abstract

The exploration of sentiment extraction within the realm of e-learning systems through the utilization of social media networks has constituted an ongoing subject of investigation within the field of data mining. To discover the significant intention of the student to the e – learning system through sentimental analysis models stands to be inadequate and sparse as most of the data identified to be non actionable on employment of the existing sentiment prediction technique. The extracting the sentiment is alone not feasible in the forecasting the interest of the student to the e- learning platform. In addition to incorporation of the annotation, domain knowledge, Sensitiveness and subjective information related to forecasting and prediction analysis on the social network data extraction will eliminate the challenges in the traditional sentiment analysis approaches using machine learning algorithms. In this article, deep learning architecture entitled as Dense CNN +BERT model has been constructed on incorporation of annotation, sensitiveness, subjective information's, sentiments and domain knowledge.

Initially the preprocessing of the data is carried out for normalizing and annotating phrases and words in the data. Preprocessed work sequence is employed to BERT model as it considered as a natural language processing approach to extract the local and global features to obtain the opinion lexicon and it is represented as word embedding vector. Those embedding vector is passed to the proposed dense convolution neural network. Convolution layers composed of inbuilt constraints which compute the high level features from the word embeddings to eliminate the inconsistency related with the extracted features. The challenge of correlating dimensions with the max pooling layer has been addressed by implementing down sampling techniques. The utilization of the ReLU Activation Function has been investigated under specific conditions to gauge the effectiveness of actionable attributes' potency. Furthermore, domain knowledge has been applied for the dynamic refinement and validation of optimal features. Fully connected layer of containing softmax function determines the polarity of the feature and polarity of the feature predicts the adoption of the student to e learning in future. An experimental result has been carried in the twitter dataset to prove that proposed framework outperforms other conventional machine learning techniques with respect to efficiency, Fmeasure, parameter insensitiveness and accuracy.

Keywords: BERT Model, Dense Convolution Neural Network, Lexicon, Prediction, Sentiment analysis and Stock Analysis.

Introduction:

Nowadays E learning system is becoming vital technology across educational institutions, manufacturing and service-oriented industries. Especially, Covid -19 pandemic has increased the adoption rate of the peoples using the e learning applications across the world. Despite of its advancement in employing current technologies trends, it faces severe challenges in handling the user critics against the application services on daily basis. Thus, Sentiment analysis of the online reviews posted on social networks during covid -19 gains significant interest on the researchers to extract the opinion orientation with sentiment of the user or students on using e learning system. This process aids in acquiring insights and predicting the future adoption rate of students utilizing the e-learning system. Traditional machine learning models have been utilized to assess user sentiment based on their intentions. However, effectively identifying the meaningful intentions of students towards the e-learning system through sentiment analysis models remains limited due to the prevalence of non-actionable data. To address these challenges, it is necessary to incorporate annotations, domain knowledge, sensitivity, and subjective information. This study introduces a novel deep learning architecture called the Dense CNN + BERT model, which integrates the aforementioned functionalities.

Bidirectional Encoder Representation from Transformer known as BERT is a bidirectional attention mechanism to generate a language model which learns contextual relation between the words on the review or pre-processed reviews to generate the word embedding vector. Word embedding vector composed of the high level and low-level features. Embedding Vector of the BERT model is applied to the Dense Convolution Neural Network (DCNN). DCNN process the vector with convolution, Max pooling and fully connected layer to yields a optimal sentiment polarity to the opinion of the user. Those sentiment polarities predict the adoption rate of the user to e learning system.

. The subsequent sections of this paper are structured as follows: Section 2 provides an examination of relevant literature, while Section 3 presents the detailed description of the proposed deep learning model, referred to as Dense Convolutional Neural Network with Bidirectional Encoder Representation from Transformer architecture. Section 4 showcases the experimental results and performance metrics of the introduced model, utilizing a Twitter dataset. A comparative analysis of its performance against conventional methods is also emphasized in this section. Lastly, Section 5 concludes the article by summarizing its accomplishments.

Related Works:

Within this segment, a comprehensive analysis of sentiment analysis architectures employing machine learning models is undertaken. This examination delves into the intricacies of processing user opinions, encompassing aspects like feature extraction, feature selection, and class representation. Moreover, the prediction performance concerning sentiment-oriented opinion mining is assessed through the utilization of similarity measures, including cosine similarity and Euclidean distances. Among the array of machine learning techniques evaluated, those demonstrating noteworthy accuracy in the approach's evaluation are thoroughly expounded upon. Additionally, akin techniques yielding results akin to the proposed architecture are also elucidated below**.**

Sentiment Classification for Movie Reviews Using Improved Semantic Oriented Approach

This literature examines the realm of sentiment classification, which involves categorizing user reviews into either positive or negative polarity. Within this context, two prominent model types emerge: machine learning architecture and semantic orientation approaches. The process commences with the inclusion of a part-of-speech tagger, tasked with identifying phrases in the input data that encompass adjectives or adverbs. Subsequently, the computation of semantic orientation is achieved through the utilization of the mutual information technique [8].

Extraction of Opinions, Opinion categories, and Topics Expressed within Online News Media.

In this literature, semantic structures of the sentence have been exploited from the online news texts to identify the opinion of the user with topic generation. This method employs semantic role labelling as a constructive process to label an opinion and topic to the data on incorporating the three phases for opinion analysis: computing an opinion-bearing word, labelling semantic roles with respect to the word in the sentence, and then determining the user category and the topic of the opinion word among the labelled semantic roles. For a wide coverage, clustering approach predict the most probable class to a word which is not specified in previous works [9].

Proposed model:

Proposed architecture specifies a detailed specification of the forecasting and prediction of the user orientation and adoption rate using deep learning-based Sentiment analysis architecture. In order to achieve better results hyperparameter tuning of the deep learning layers has to be carried out to yield the prediction and forecasting results of the students review.

Problem Definition

The primary importance of the work is to analyse the review data to select sentiment to it. However traditionally sentiment analysis technique employs the rule-based approach and intension-based approach to detect the sentiment to the extracted feature vector. Those rule-based model and intension model faces numerous limitations on its implementation to streaming social media data. The limitation of the models is stated as follows

- The lengths of the review data will be changing dynamically. For example, few reviews contain thousands of words and some reviews contain only one sentence with two or three words.
- Extracting and classifying the context of the multilingual data on the polarity of the sentiment to the implicit and explicit word is highly complex [10].

Data Pre-processing

Data preprocessing of opinion data represented as user generated text contains noisy and redundancy data which is eliminated using normalized on stop word removal process, stemming process, tokenization process and weighting scheme.

- **Stop Work Removal Process**

It is removing the meaningless words in the opinion word which is considered as common words in the natural language such as articles, preposition, pronouns and conjunctions etc. It is to remove the meaningless words which resides in the opinion sentences abundantly. It is also considered aslow-level information containing common words in the natural language such as articles, preposition, pronouns and conjunctions etc. Removing the common words will not produce any negative consequence to the data analytics model instead it reduces the dataset size and training time of model.

- **Stemming process**

It is the process of reducing the word from its base word which is affixed to suffixes and prefixes or to the roots of words termed as lemmas. It is also considered as word normalization technique. Stemming process is vital process as it produces the resultant token containing reduced character length.

- **Tokenization Process – word-based tokenization**

It is to process which splits opinion sentence into separate meaningful tokens. Token can be easily compiled by the learning models and its machine. Token can be character, word and sub word. In this word-based tokenization, splitting process of the sentences eliminates only punctuation, whitespaces, delimiters from the sentence from further processing as it is considered as meaningless in sentiment analysis.

- **Weighting scheme**

Token contains more information for contextual learning instead it is required to extract the best words or token.

$$\mathrm{Tf}\,u\,,wi \;=\; \frac{TF'}{\sum wi \,\xi\, W\,(TF')}$$

Here, TFu,wi represents the Term Frequency value of wi in document f. W denotes the set of keywords. Subject and Object Extraction: The extraction of subjects and objects is primarily accomplished through context mining. This process aims to efficiently yield essential background knowledge for subsequent stages. Another approach for extracting subjects and objects involves text analysis, where opinion-oriented information is extracted directly from the textual content. The extraction of properties relies on sentiment, modifiers, and rule bases. Furthermore, the concept of Maximum Likelihood Estimates is explored. It simply use the frequencies in the data

$$\hat{P}(k_i) = \frac{N(k = k_{i=0})}{N}$$

Designing and incorporating the Sentimental analysis based on the data polarity

We suggest creating sentiment, modifier, object, and rule bases using BERT, a rule-based approach, based on the polarity of the data. The procedure is outlined as follows."

Semantic Orientation –BERT Analysis

- Sentiment Base: The sentiment analysis process comprises two closely intertwined modules: the sentiment lexicon and the sentiment polarity associated with terms in the lexicon. The lexicon encompasses subjectivity cues, such as verbs, adjectives, and nouns, annotated with their polarity (positive, negative, or neutral) and strength (strong or weak). However, this lexicon solely defines a word's original polarity, whereas a word's actual polarity can be influenced by its context within a sentence.
- Various methods that consider word context have been proposed to compute the sentiment orientation of words. For instance, the sentiment polarity of a word might be established by its morphemes. If the morphemes of a word are more prevalent in the positive lexicon compared to the negative lexicon, the word is considered positive; otherwise, it is categorized as negative.
- Modifier Base: Building on the earlier assumptions, a sentence's initial sentiment is determined by sentiment words. However, this sentiment can be modified by adverbs. Adverbs of negation can reverse sentiment polarity, indicating the opposite sentiment (e.g., "fast" is positive, but becomes negative if preceded by "not"). Similarly, degree adverbs that amplify or mitigate sentiment intensity must also be taken into account. Furthermore, the sentence structure contributes to the sentiment polarity value of a sentence.
- Semantic Rule Base: The creation of semantic rules holds paramount importance within the rule-based approach, particularly regarding student sentiment and interest in adopting e-learning. Each sentiment word has distinct modifiers. The core relationships between a sentiment word and its modifiers hinge on their placements and categories within the sentence [14].

Computation of Score Value through CNN Approach - Positive or Negative

Significant thematic sentences typically occupy prominent positions such as titles, opening sentences, and concluding sentences to enhance emphasis. Consequently, when determining the overall polarity of a document, the sentiment sentence's location should be taken into account. Practically, a sentence's significance within a document can be indicated by its weight during polarity calculation. Thematic sentences should carry more weight compared to other sentences in the document.

The Convolutional Neural Network (CNN)[15] is a deep learning architecture employed to predict student adoption of online e-learning systems. In this segment, the activation function's parametric tuning for student adoption prediction is explored within the context of objective functions, based on the latent factors derived from opinion extraction. The training model utilizes max pooling to transform the extracted latent features into pairwise feature representations. These pair-wise representations are then processed within a convolutional function, representing student preferences. This process aims to reconstruct features into an opinion [16].

- Max pooling layer

Within this layer, latent features in the form of subsets are organized into pairs of features, representing various preference pairs for predicting student intentions, considering bias coefficients. This is achieved through the construction of a student preference matrix, established via matrix factorization. This model operates as a latent factor model, focused on predicting sentiment for e-learning opinions expressed by students.

The max pooling layer within the neural network is utilized to generate latent features for the embedding layer. These feature representations aim to improve student adoption of the e-learning system, particularly during Pearson correlation similarity computation regarding student preferences. The results provide insights into features that enhance student adoption.

$$\mathbf{x} = \sum_{i=1}^{N} x_i\, v_i = x_1 v_1 + x_2 v_2 + \ldots + x_N v_N$$

The components for tuning the prediction of the Fully Connected Neural Network are outlined in Table 1, encompassing scheduling details.

Table 1: Hyperparameters for Fine-Tuning Fully Connected Neural Networks to Achieve Desired Results.

Scheduling Component's Hyper Parameters	Values of Parameter
Matrix Batch Size	118
Model Learning Rate	0.08
Dimension Size	25
Number of Epoch	30
Error function	Cross entropy

- **Layer of Embedding**

The Embedding layers encapsulate the preference feature subset hierarchically, employing their inherent mechanism to distil abstract and discriminative features with minimal hyperparameters from the fully connected network. The result of this process is the latent dimension of the evolving stock characteristics. Within this layer, numerous evolving traits of the student preference variable are embedded into the activation function for discerning student adoption. The architecture of the proposed model is as follows:

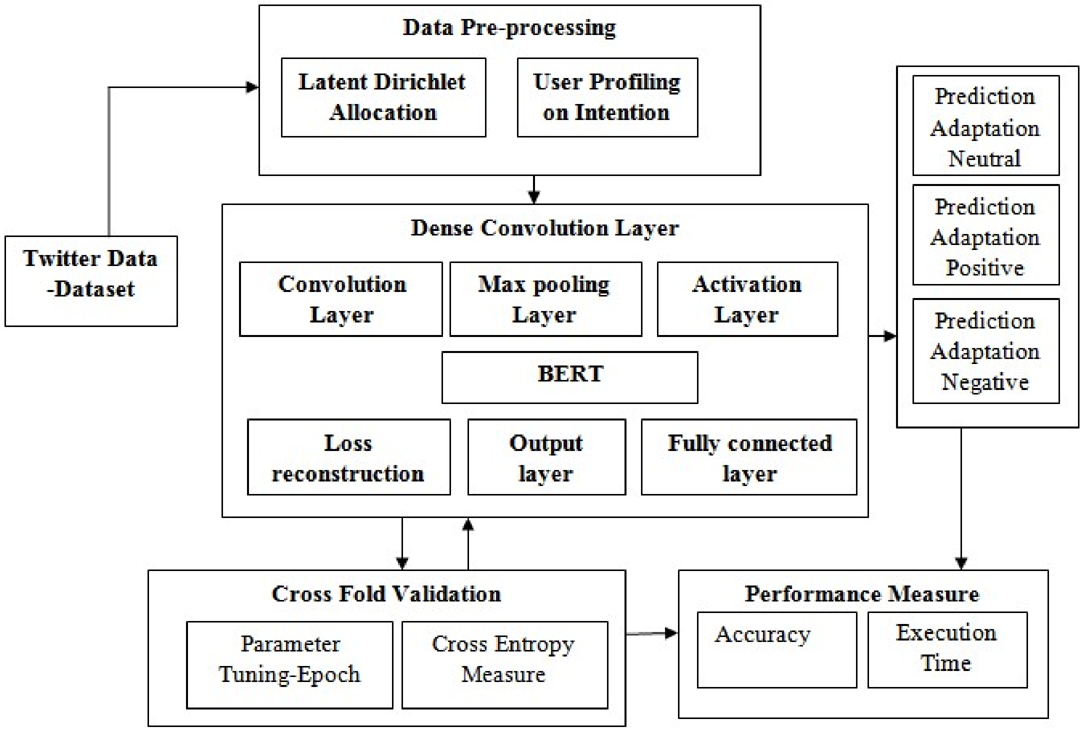

Figure1: Architecture of the proposed model

- **Function of Activation**

The suggested framework employs the rectified linear units (ReLU) activation function [17] in the embedded layer for stock recommendation within the evolving student pool, which includes the preference vector and pairwise vector of e-learning data. This activation function generates the relevant student intention for each element within the embedding matrix. The embedded vector is further manipulated using parameterized values to produce accurate predictions for the chosen student vector and for each epoch's parameter updates.

- **Output Layer**

The output layer of the deep neural network comprises the prediction results, encompassing student suggestions for the e-learning system. Additionally, in this layer, a parametric flattening of predictions is executed using similarity measures to facilitate student recommendations to the e-learning system. This process involves the utilization of Softmax optimization [18] on the resulting set, while the cross-entropy mechanism is employed to assess the efficacy of student adoption towards e-learning applications.

- **Loss Layer**

This layer is designed to enhance prediction accuracy during the fine-tuning process by refining parameters across various layers of the deep neural network. Its objective is to minimize the reconstruction error between the feature max pooling layer and the ReLU activation layer. Moreover, the utilization of the cross-entropy loss function has been employed to effectively manage prediction outcomes [19].

Algorithm 1: CNN learning for forecasting the student adoption
Input: Discriminative student opinion data to e learning applications
Output: forecasting student adoption
Process
Pre-processing ()

Stopword removal()
Stemming ()
Tokenization()
BERT Investigation ()
Calculate the latent feature set $F_{\cdot s}$
Deep Learning Application:
Convolution learning ()
Set Kernel () to Feature vector
Embedded Feature Map containing sentiment association to token
Max Pooling ()
Compute weighted student characteristics to e learning system
Layer of Embedded ()
Estimation of features incorporating sentiment.
Activation Layer Utilization ()
Parameterized Tuning of ReLU Function.
Output Layer Integration ()
Loss layer ()
Cross Entropy (confusion Matrix)
Softmax() ---forecasting student adoption rate

This process promotes the formation of embedded latent feature points on a representative map to facilitate forecasting student adoption in the e-learning system.

Results and Discussion:

Amid the COVID-19 pandemic period, online student opinions related to the e-learning system were gathered from various categories of students worldwide through Online Social Networks, including platforms like Twitter. Subsequently, this data was structured into a corpus of records. The online student opinions, gathered as reviews, revolved around student intentions, sensitiveness, and subjectiveness.

To process this information, the data was tokenized and further refined by eliminating stop words and undergoing stemming processes within the dataset [20].

4.1. *Performance Evaluation*

The assessment of the forecasting's efficacy involves the utilization of the subsequent performance metrics:

- **Precision:** Also known as positive predictive value, precision quantifies the portion of retrieved instances that hold relevance within the cluster of student data.
- **Recall:** Often referred to as sensitivity, recall gauges the proportion of pertinent instances that have been successfully retrieved. Both precision and recall are predicated on an appreciation of relevance and serve as measures thereof.

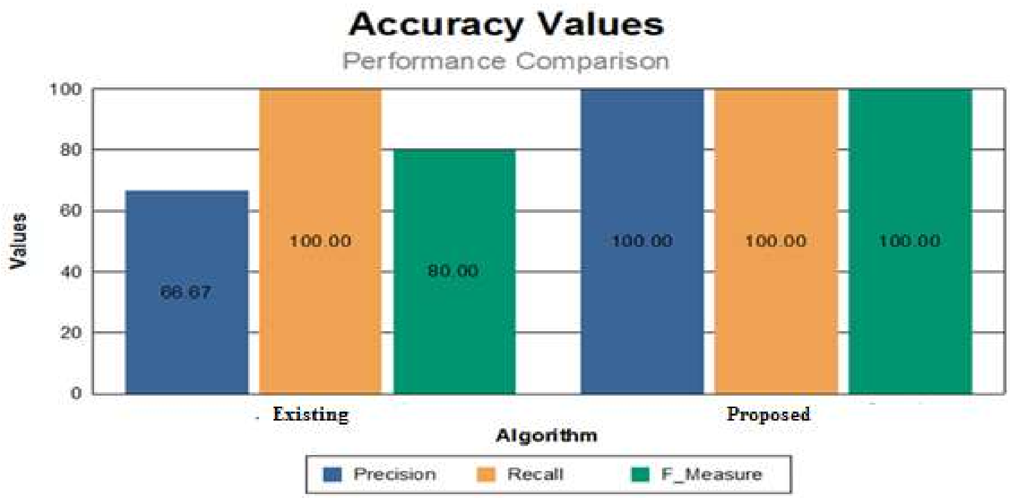

Figure 2: Assessment of the prediction outcomes' performance

The model is evaluated on the performance of sentiment analysis in the student data for forecasting against the precision, recall and accuracy. The classification accuracy is the degree of closeness of computed results to actual (true) value. Figure 2 represents the performance evaluation of the proposed model in terms of accuracy. Table 2 represents the performance evaluation of the proposed model against the existing model.

Table 2: Evaluation Performance

Technique	Precision	Recall	Fmeasure
Dense CNN +BERT–Proposed	100	99	99.9
RNN-LTSM- Existing	99	88	93

Since the statistical approach demonstrated greater accuracy in comparison to the semantic approach, we chose the statistical approach to label the reviews.

5 Conclusion:

The Sentiment analysis in student opinion data analysis is carried in this work to forecast the student intension to e-learning system. In this work, a new architecture is proposed and implemented using Dense CNN architecture with BERT model as it is rule based approach for sentimental analysis on the web based data. The opinion data has been structured as a predictive analysis, while the proposed solution has been devised to scrutinize student opinion data for predicting future student adoption of the e-learning system. This process involves utilizing the latent semantic analysis algorithm, which exhibits improved performance in terms of scaling and accuracy compared to data classification. The sentiment analysis is conducted by computing polarity scores, namely Positive, Negative, and Neutral, based on word occurrences.

References

[1] H. Li, L. Dombrowski, and E. Brady, “Working toward empowering a community: How immigrant-focused nonprofit organizations use twitter during political conflicts,” in Proceedings of the 2018 ACM Conference on Supporting Groupwork. ACM, 2018, pp. 335–346.

[2] D. Tang, B. Qin, F. Wei, L. Dong, T. Liu, and M. Zhou, “A joint segmentation and classification framework for sentence level sentiment classification,” IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 23, no. 11, pp. 1750–1761, 2015.

[3] H. Wang, D. Can, A. Kazemzadeh, F. Bar, and S. Narayanan, “A system for real-time twitter sentiment analysis of 2012 us presidential election cycle,” in Proceedings of the ACL 2012 System Demonstrations. Association for Computational Linguistics, 2012, pp. 115–120.

[4] D. Paul, F. Li, M. K. Teja, X. Yu, and R. Frost, “Compass: Spatio temporal sentiment analysis of us election what twitter says!” in Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2017, pp. 1585–1594.

[5] F. Bravo-Marquez, E. Frank, and B. Pfahringer, “Annotate-sampleaverage (asa): A new distant supervision approach for twitter sentiment analysis,” in 22nd European Conference on Artificial

Intelligence (ECAI), vol. 285. IOS Press, 2016, pp. 498–506.
[6] B. Pang, L. Lee et al., "Opinion mining and sentiment analysis," Foundations and Trends R in Information Retrieval, vol. 2, no. 1–2, pp. 1–135, 2008.
[7] K.-L. Liu, W.-J. Li, and M. Guo, "Emoticon smoothed language models for twitter sentiment analysis," in Proceedings of the Twenty- Sixth AAAI Conference on Artificial Intelligence, 2012.
[8] J. Zhao, L. Dong, J.Wu, and K. Xu, "Moodlens: an emoticon-based sentiment analysis system for chinese tweets," in Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2012, pp. 1528–1531.
[9] A. Go, R. Bhayani, and L. Huang, "Twitter sentiment classification using distant supervision," CS224N Project Report, Stanford, vol. 1, no. 12, 2009.
[10] D.-T. Vo and Y. Zhang, "Target-dependent twitter sentiment classification with rich automatic features." in IJCAI, 2015, pp. 1347– 1353.
[11] E. Cambria, "Affective computing and sentiment analysis," IEEE Intelligent Systems, vol. 31, no. 2, pp. 102–107, 2016.
[12] D. Tang, F. Wei, N. Yang, M. Zhou, T. Liu, and B. Qin, "Learning sentiment-specific word embedding for twitter sentiment classification," in Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, 2014, pp. 1555–1565.
[13] K. Schouten and F. Frasincar, "Survey on aspect-level sentiment analysis," IEEE Transactions on Knowledge and Data Engineering, vol. 28, no. 3, pp. 813–830, 2016.
[14] J. Zhao and X. Gui, "Comparison research on text pre-processing methods on twitter sentiment analysis," IEEE Access, vol. 5, pp. 2870–2879, 2017.
[15] S. Symeonidis, D. Effrosynidis, and A. Arampatzis, "A comparative evaluation of pre-processing techniques and their interactions for twitter sentiment analysis," Expert Systems with Applications, 2018.
[16] M. Thelwall, K. Buckley, and G. Paltoglou, "Sentiment in twitter events," Journal of the Association for Information Science and Technology, vol. 62, no. 2, pp. 406–418, 2011.
[17] N. Du, Y. Liang, M. Balcan, and L. Song, "Influence function learning in information diffusion networks," in International Conference on Machine Learning, 2014, pp. 2016–2024.
[18] M. Tsytsarau, T. Palpanas, and M. Castellanos, "Dynamics of news events and social media reaction," in Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2014, pp. 901–910.
[19] S. Stieglitz and L. Dang-Xuan, "Emotions and information diffusion in social mediasentiment of microblogs and sharing behavior," Journal of Management Information Systems, vol. 29, no. 4, pp. 217–248, 2013.
[20] Y. Fu, Y. Ge, Y. Zheng, Z. Yao, Y. Liu, H. Xiong, and J. Yuan, "Sparse real estate ranking with online user reviews and offline moving behaviors," in 2014 IEEE International Conference on Data Mining (ICDM). IEEE, 2014, pp. 120–129.