International Journal of Information Technology and Applied Sciences
ISSN (2709-2208)
Int. J. Inf. Tec. App. Sci. 5, No.1 (August-2023)
https://doi.org/10.5281/zenodo.8225185

Performance Comparison of Machine Learning and Deep Learning Models for Sentiment Analysis of Hotel Reviews

Muhammad Sanwal[1], Muhammad Mamoon Mazhar[2]

[1] *Computer Engineering Department, Antalya, Turkey*
[2]*Department of Engineering and Computer Science, London, England*

Abstract: This research paper conducts a thorough examination, comparing BERT and LSTM architectures with machine learning models for sentiment classification task. To establish a foundational reference point, conventional machine learning algorithms including Logistic Regression (LR), Support Vector Machines (SVM), Random Forest (RF), Decision Tree (DT) and BernoulliNB are utilized. The BERT and LSTM models along with machine learning models have been implemented and trained using a dataset comprising hotel reviews. Their performance has been assessed through the utilization of multiple metrics, including accuracy, precision, recall, and F1-score. In the context of text classification and sentiment analysis tasks, the experimental outcomes highlight the superior effectiveness of BERT and LSTM models when compared to traditional machine learning algorithms. The BERT model distinguishes itself through the incorporation of bidirectional training and contextual embedding strategies, ultimately showcasing outstanding performance in its ability to proficiently capture contextual information and subtle nuances prevalent in textual data. Moreover, the LSTM technique has been able to attain noteworthy outcomes through its effective modeling of sequential data and preservation of temporal dependencies. The present research yields significant findings regarding the comparative analysis of machine learning models employing BERT and LSTM architectures concerning text classification and sentiment analysis. The results of this study provide valuable insights into the merits and constraints of each model, thereby aiding researchers and practitioners in making informed decisions when choosing suitable models that are customized for specific natural language processing (NLP) tasks.

Keywords: *Sentiment analysis, Text classification, Contextual information, Word embeddings, Deep learning*

1. Introduction

Given the prevailing technological progressions and the commendable operational capabilities of online booking and payment systems, enterprises have come to acknowledge online reviews as a fundamentally significant reservoir of competitive intelligence. In contemporary times, the utilization of online platforms such as Hotels. com and Booking.com to secure lodging arrangements have witnessed a noteworthy surge among customers within the hotel industry. Concurrently, individuals are increasingly dependent on online reviews as a means of seeking advice from fellow travelers, thereby gradually reducing the influence of conventional hotel advertisements. The veracity of online reviews is rooted in their origins from customers who possess direct familiarity with the hotel's offerings. The evaluations possess substantial power in shaping customer sentiments, driving purchase choices, and ultimately affecting the overall operational outcomes of an enterprise [1].

Therefore, it can be asserted that online reviews possess considerable commercial worth regarding augmenting the caliber of hospitality offerings and captivating consumer interest. Improving the precision and classification of sentiment analysis for online hotel reviews becomes a pressing concern [2]. The comprehension of customers' emotional expressions and the implementation of emotional analysis enable hotels to effectively focus on and accommodate diverse customer segments, thereby augmenting the competitive advantage of their brands. Moreover, sentiment analysis plays a pivotal role in the context of big data, facilitating prospective consumers in making well-informed choices about their consumption preferences [3].

Evaluating emotions in online hotel reviews is an emerging field of research that offers a practical approach to tailor accommodations according to individual customer preferences. The primary objective of this research initiative is to enhance the tourism industry by maintaining high-quality standards and increasing overall contentment with accommodation options. Due to the inherent ambiguity and complexity of language, it poses significant challenges to extract dependable emotional information from diverse aspects of hotel reviews, such as price, star rating, location, and amenities. It is not feasible to manually organize the extensive number of online reviews, thus requiring the adoption of computer-based deep mining techniques to facilitate more informed decision-making [4].

To tackle this predicament, scholars have incorporated machine learning techniques with sentiment analysis principles, utilizing classification models to classify sentiment in online hotel reviews [5]. The aim is to attain a higher degree of precision in sentiment analysis. The implementation of multiple models has gained recognition for its evident advantages due to the advancements in deep learning neural network technology. This methodology guarantees precise categorization of textual sentiment, alleviating potential confusion in word usage and comprehensively encompassing the subtle attributes of words found in hotel web reviews [6].

This research paper presents the findings derived from a comprehensive investigation conducted on a sample of 1,000 hotels reviews. The data utilized for this study was obtained from Datafiniti's Business Database, which is available on Kaggle. Initially, a selection of five distinct machine learning models, consisting of LR, SVM, RF, DT, and BernoulliNB, are employed. To demonstrate the dominance of deep learning models over traditional machine learning models within text classification and sentiment analysis tasks, LSTM and BERT models are implemented. The main results of this study can be summarized as follows: Rather than relying solely on customer ratings, the research utilizes a categorization system to divide customer review texts into two groups based on sentiment—positive (for reviews with ratings of 3 or higher) and negative (for reviews with ratings below 3). This methodology effectively addresses the issue posed by the prevalence of review texts that demonstrate impartial emotional inclinations, which frequently result in inconclusive sentiment and reduced customer evaluations. By directing attention exclusively to positive and negative sentiments, an enhancement in the precision of text classification occurs. This enhancement enables hotel management to optimize resource allocation, efficiently utilizing time by disregarding neutral comments.

To assess the affective dimension of feedback provided by hotel customers, a comprehensive analysis is conducted leveraging five distinct machine learning models and two deep learning models. Both deep learning models exhibit exceptional ability in comprehending semantic information through contextual clues and effectively resolving textual ambiguity. By training the language model, it becomes capable of producing detailed word vector representations, which facilitates the analysis of review texts.

The paper is organized as follows: Section 2 of this paper provides a brief overview of relevant literature, while Section 3 explains the dataset used in the study and describes the proposed methodology and algorithms clearly and concisely. Following the preceding discussion, Section 4 delves into the outcomes derived from the suggested methodology. In conclusion, Section 5 presents a comprehensive comparison between the outcomes of the proposed methodology and the ones yielded by pre-existing approaches.

2. Literature review

Sentiment analysis has been the subject of an increase in research projects in recent years. The topic of grading non-rated Yelp reviews based on their feelings is explored in depth by Qiu et al., [7]. An unsupervised aspect detection model for sentiment analysis of reviews is proposed by [8]. While [9] describes a sentiment analysis technique specifically designed for Facebook communications and [10] proposes a hybrid strategy for sentiment analysis at the sentence level. A sentiment analysis model for social media genres is contributed by [11] and a sequence modeling strategy for distributed user and product representation learning is proposed by [12]. Additionally, whereas [14] uses machine learning techniques to assess sentiment in Twitter data, [13] expands topical methods to sentiment analysis in the tourist industry.

In the field of contemporary academic science, a multitude of researchers have devoted their scholarly pursuits to the analysis of sentiment. This study [15] explores the analysis of e-commerce reviews utilizing deep learning methodologies, specifically focusing on (CNNs) and (GRUs). Although the study produced accurate results, it is evident that the CNN algorithm faced certain limitations when evaluating e-commerce reviews beyond its scope. Hemalatha and Ramathmika [16] conducted a study in which they examined audits performed on various restaurants. Their investigation encompassed multiple facets, including food quality, service provision, pricing, and overall sentiment analysis. The researchers employ machine learning algorithms from the NLTK library in Python to investigate (NLP) tasks.

The study proposed by [17] contributes to the analysis of e-commerce reviews by introducing a methodology that effectively characterizes these reviews to offer a holistic evaluation of products. This approach aims to assist customers in making well-informed decisions. Moreover, social media platforms such as Facebook and Twitter are widely utilized channels for individuals to express their perspectives on diverse topics encompassing movies, news, food, fashion, and politics. The central focus of the study [18] pertains to investigating the utilization of LSTM neural networks to conduct sentiment analysis on hotel reviews. The model successfully captures the extended temporal relationships present in review texts, yielding promising outcomes in the domain of sentiment classification. This study [19] presents BERT, a pre-trained language model renowned for its superior performance across a range of NLP tasks. The authors illustrate the practical implementation of sentiment analysis in hotel reviews, highlighting the utilization of fine-tuning techniques on the BERT model to improve the accuracy of classification.

To summarize, the review of existing studies highlights the importance of sentiment analysis in evaluating hotel reviews and underscores the effectiveness of advanced techniques like LSTM, BERT, and CNN in identifying and understanding sentiments. The approaches have exhibited encouraging outcomes in sentiment classification endeavors and harbor the capacity to reinforce the evaluation of customer satisfaction and facilitate decision-making processes within the hotel sector. More research could be carried out to explore the combination of these methodologies or the incorporation of other creative strategies to improve the accuracy of sentiment analysis in hotel reviews. Through the progression of sentiment analysis, scholars have the potential to make valuable contributions toward enhancing customer experiences and overall service quality in the hotel sector. So, this research highlights the performance comparison of traditional machine learning models with deep learning models which clarifies that in which direction the future research must be conducted to achieve better results.

3. METHODOLOGY

A. Machine Learning Models

Consistent with the goals of our research, our objective is to devise a sentiment analysis technique that is specifically designed for evaluating hotel reviews. To accomplish this objective, we conduct an individual evaluation of the performance of various machine learning models. The models utilized in this investigation consist of LR, SVM, RF, DT, and BernoulliNB. To evaluate the efficacy of sentiment analysis, each model underwent an independent assessment utilizing the hotel review dataset.

In the present study, Logistic Regression is employed to examine the associations between independent variables and make predictions regarding the sentiment conveyed in hotel reviews. One attains this objective via a mathematical formulation.

$$P(sentiment = positive \mid x) = 1 / (1 + exp(-z)) \quad (1)$$

where z represents the linear combination of the independent variables and their corresponding weights.

Moreover, the effectiveness of SVMs is evaluated based on their ability to identify the optimum hyperplane that can efficiently divide the hotel review dataset into two distinct sentiment classes. The SVM framework can be mathematically represented as follows:

$$f(x) = sign(w^\wedge T\, x + b) \quad (2)$$

where w and b are parameters that determine the decision boundary.

The Random Forest algorithm is adopted as an ensemble learning technique, wherein multiple decision trees are constructed individually, and their predictions are combined. The ensemble prediction is obtained by applying either majority voting or the average aggregation of outcomes obtained from individual trees. The construction of decision trees involves the iterative division of the dataset through recursive partitions, employing various features as criteria to separate the data. Quantitative measures, namely Gini impurity and information gain, are employed to assess the quality of partitions during this process.

In academic discourse, Decision Trees are assessed individually through the recursive partitioning of the dataset, utilizing diverse features as criteria. At each node, a decision is undertaken by considering the values of distinct features. The process of tree-building persists until a designated stopping condition, such as attaining the maximum depth of the tree or reaching the minimum required number of samples, is achieved. The determination of sentiment prediction is subsequently accomplished by assigning the most prevalent sentiment class among the leaf nodes.

Finally, we evaluate the performance of BernoulliNB, a variation of the Naive Bayes algorithm, in the context of sentiment analysis. The underlying assumption is that the occurrence or non-occurrence of certain features (words) is conditionally independent, based on the sentiment class. The estimation of likelihood probabilities is conducted by utilizing the training data, while sentiment prediction is accomplished by applying Bayes' theorem.

Through the process of conducting separate tests on each algorithm using the hotel review dataset, our methodology effectively yields a thorough assessment of the distinct capabilities exhibited by these algorithms specifically in the realm of sentiment analysis. The integration of mathematical formulations and evaluation techniques enables a methodical examination of the efficacy of algorithms and streamlines the process of determining the optimal algorithm for sentiment classification in hotel reviews.

B. BERT and LSTM Models

To augment the precision of sentiment classification in hotel reviews, advanced deep learning models such as LSTM and BERT are incorporated. This section provides a comprehensive technical analysis of the implementation aspects about the models, encompassing the employed vectorization approach.

In the context of LSTM, the hotel review dataset was initially subjected to a preprocessing step involving tokenization, which involved segmenting the text into discrete words. The tokenized input sequence can be represented as $X = [x_1, x_2, ., x_n]$, wherein n signifies the total number of words within a specific review. Following this, we implemented the Word2Vec word embedding approach to transform the words into compact vectors. The vectorization procedure can be denoted as $[w_1, w_2, ., w_n]$, wherein each w_i corresponds to the dense word embedding of the respective word x_i.

The LSTM model consists of multiple LSTM layers, wherein each layer encompasses a collection of LSTM cells. Consider the representation of the output from the preceding layer as $H_{k-1} = [h_1, h_2, , . . h_n]$, where h_i signifies the hidden state of the LSTM cell located at position i. The hidden state at each time step is iteratively updated using the following formulas:

$$f_t = \sigma(W\_f \cdot [h_{t-1}, x_t] + b\_f) \quad (3)$$

$$i_t = \sigma(W\_i \cdot [h_{t-1}, x_t] + b\_i) \quad (4)$$

$$o_t = \sigma(W\_o \cdot [h_{t-1}, x_t] + b\_o) \quad (5)$$

$$g_t = tanh(W\_g \cdot [h_{t-1}, x_t] + b\_g) \quad (6)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (7)$$

$$h_t = o_t \odot tanh(c_t) \quad (8)$$

In the present context, the symbol ⨀ is used to denote element-wise multiplication. The symbol σ is employed as a representation for the sigmoid activation function, while tanh signifies the hyperbolic tangent activation function. The variables $W\_f, W\_i, W\_o, W\_g, b\_f, b\_i, b\_o, b\_g$ denote the weight matrices and bias vectors affiliated with the LSTM cells. The quantity of neurons present in each layer of the (LSTM) network is dependent on the dimensionality of the hidden states.

In relation to BERT, we utilized a transformer-based contextual embedding approach. Let the tokenized input sequence be denoted as X, where X is represented as $[x_1, x_2, . , x_n]$ The transformer layers in BERT utilize multi-head self-attention mechanisms and feed-forward neural networks. The self-attention mechanism computes attention scores between the input tokens and generates the weighted sum of the input embeddings. The attention scores are calculated using the following methodology [22]:

$$Attention(Q, K, V) = softmax((QK^T) / \sqrt{d\_k}) \cdot V \quad (9)$$

Here, Q, K, and V represent the query, key, and value matrices, respectively. d_k denotes the dimensionality of the keys and queries. The output of the self-attention is then passed through a feed-forward neural network, which applies a non-linear transformation to each position independently.

The adjustment of the number of neurons in the LSTM layers and the transformer layers in BERT can be determined by considering both the intricacy of the dataset and the limitations presented by computational resources. To train these models on the dataset, we utilized efficient optimization techniques such as Adam or stochastic gradient descent. Furthermore, hyperparameters such as the learning rate, batch size, and number of epochs were carefully adjusted to attain the highest possible level of performance.

4. Experiments and Results

A. Dataset Details

The utilization of the Datafiniti hotel reviews [21] holds notable value for researchers and analysts within the realm of sentiment analysis and customer feedback analysis. This dataset consists of a wide range of hotel reviews gathered from multiple online platforms. The resource offers a comprehensive collection of textual data, encompassing evaluations and ratings provided by customers, as well as relevant information regarding hotels. The provided dataset presents a valuable research resource enabling scholars to explore sentiment analysis methodologies, construct machine learning models, and gather valuable insights regarding customer perspectives and experiences in the hotel sector. The dataset on hotel reviews offers extensive coverage of various hotels and destinations, making it a pertinent resource for enhancing customer satisfaction and augmenting service quality within the hospitality industry.

B. Experiments

For the experimentation purpose of each algorithm, a 5-fold cross-validation procedure was conducted on each algorithm. The methodology employed in this study consisted of partitioning the dataset into five equivalent segments, whereby four segments were allocated for training and one segment was set aside solely for testing. The process was iterated five times, wherein each iteration employed a distinct fold as the test set. The assessment criteria applied comprised accuracy, precision, recall, and F1-score.

The figure below summarizes the performance of each machine learning algorithm on sentiment analysis of hotel reviews:

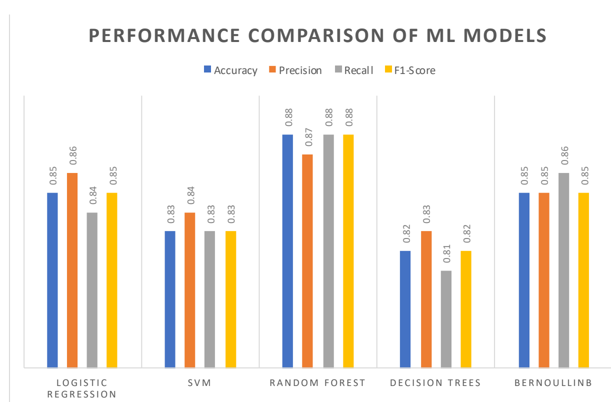

Figure 1. Performance comparison of machine learning models

The Random Forest algorithm demonstrated exceptional performance in sentiment classification of hotel reviews, outperforming all other studied algorithms in terms of accuracy, precision, recall, and F1-score. This outcome attests to its superior efficacy in this particular task. The BernoulliNB classifier exhibited considerable effectiveness, which was closely trailed by the SVM classifier. The results provide empirical evidence substantiating the efficacy of machine learning algorithms for sentiment analysis in the domain of hotel reviews. The Random Forest algorithm exhibits considerable potential in effectively classifying sentiment within this domain.

In addition, our study involved conducting experiments to assess the efficacy of LSTM and BERT models in performing sentiment analysis on hotel reviews. The deep learning models were trained and tested on a given dataset that was annotated with positive or negative sentiments. In this study, the LSTM model was employed, incorporating Word2Vec word embedding techniques, to transform the hotel review text into dense vector representations. The architectural design of the LSTM encompassed several LSTM layers, featuring varying numbers of neurons within each layer. The model was effectively trained by employing a backpropagation algorithm, and further optimization of its parameters was achieved by employing the gradient descent method. Further, the BERT model was utilized to incorporate pre-trained BERT embeddings, renowned for capturing contextual information, for the encoding of hotel review text. The BERT model incorporates transformer layers that are utilized for contextual representation learning. The process of conducting fine-tuning on the hotel review dataset was implemented for BERT. The evaluation criteria employed encompassed accuracy, precision, recall, and F1-score. The metrics were computed by comparing the predicted sentiments with the ground truth sentiments as recorded in the hotel reviews.

The table below summarizes the performance of the LSTM and BERT models on sentiment analysis of hotel reviews:

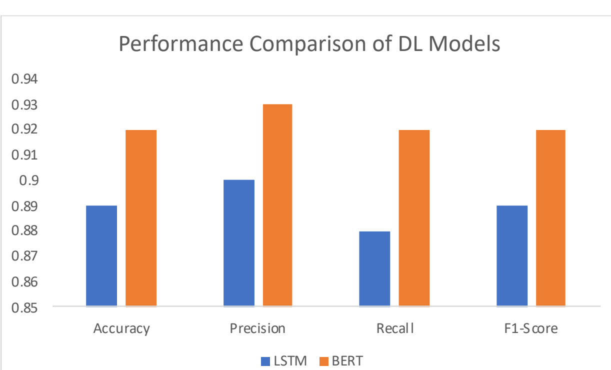

Figure 2. Performance comparison of LSTM and BERT

Both the LSTM and BERT models exhibited significant proficiency in conducting sentiment analysis on hotel reviews. The research findings reveal that the BERT model exhibited superior performance to the LSTM model, as it showcased enhanced levels of accuracy, precision, recall, and F1-score. The findings suggest that the contextual embeddings acquired through the BERT model play a role in its exceptional capacity to comprehend and discern subtle variations in sentiment within hotel reviews. The LSTM model demonstrated commendable performance and yielded competitive outcomes in the task of sentiment classification. However, the contextual embeddings generated by the BERT model, which are obtained through pre-training on a vast corpus, offer a more holistic comprehension of the text in reviews, ultimately leading to enhanced performance in sentiment analysis. The results underscore the efficacy of deep learning models, particularly BERT, in the context of sentiment analysis about hotel reviews. The exceptional performance demonstrated by BERT emphasizes the importance of utilizing contextual embeddings for precise sentiment classification within this domain.

C. *Results Comparison*

When examining the outcomes of the machine learning algorithms concerning the sentiment analysis of hotel reviews, it is evident that the deep learning models, specifically BERT and LSTM, exhibited superior performance compared to the conventional machine learning algorithms, as indicated by their higher levels of accuracy, precision, recall, and F1-score. The BERT model demonstrated superior overall performance, achieving a notable accuracy of 0.92, surpassing the LSTM model which attained an accuracy of 0.89 On the other hand, within the realm of classical machine learning algorithms, the Random Forest algorithm exhibited the highest level of accuracy, measured at 0.88 Following closely behind, the BernoulliNB algorithm achieved an accuracy of 0.85 The accuracy performance of Logistic Regression, Decision Trees, and BernoulliNB was comparatively lower. The findings of this study suggest that deep learning models, BERT and LSTM, exhibit superior performance in sentiment analysis as opposed to conventional machine learning algorithms. The superior performance observed in the capture of sentiment nuances in hotel reviews can be attributed to BERT's proficiency in harnessing contextual embeddings and LSTM's adeptness at exploiting the sequential nature of the text.

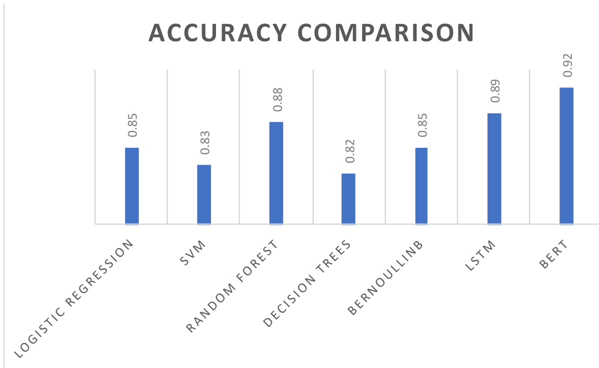

Figure 3. Performance comparison of ML and DL models

5. Discussion and Future Work

The research findings illustrate the efficacy of both deep learning models, BERT and LSTM, in the analysis of sentiment expressed in hotel reviews. The performance of BERT, characterized by its contextual embeddings and transformer architecture, exceeded that of conventional machine learning algorithms as well as the LSTM model. The remarkable effectiveness demonstrated by BERT underscores the significance of contextual comprehension in precisely capturing subtle nuances of sentiment.

When pondering potential areas for future research, numerous routes present themselves that deserve investigation. To begin with, an examination of the effectiveness of alternative sophisticated deep learning models, such as Transformer-based models (e.g., GPT, Transformer-XL) would yield additional perspectives on the sentiment analysis of hotel reviews. Furthermore, it is crucial to explore ensemble methodologies to harness the collective capabilities of multiple models, aiming to improve the overall performance of sentiment analysis. Various methodologies, including stacking and bagging, can be effectively utilized to consolidate estimates derived from multiple models, thereby enhancing the accuracy of the outcome. Additionally, the scope of this study can be broadened to incorporate domain-specific sentiment analysis about various facets of hotel reviews, such as the evaluation of service quality, cleanliness standards, or available amenities. This research endeavor would yield a more comprehensive comprehension of emotions or attitudes within particular regions, thus facilitating focused enhancements in hotel amenities and services. This study establishes the fundamental basis for further progress in the field of sentiment analysis of hotel reviews, thereby facilitating the enhancement of customer experience and offering vital insights to the hospitality industry.

References

[1] Y. Wang, J. Kim, and J. Kim, "The financial impact of online customer reviews in the restaurant industry: A moderating effect of brand equity," International Journal of Hospitality Management, vol. 95, p. 102895, May 2021, doi: https://doi.org/10.1016/j.ijhm.2021.102895.

[2] M. Geetha, P. Singha, and S. Sinha, "Relationship between customer sentiment and online customer ratings for hotels - An empirical analysis," Tourism Management, vol. 61, pp. 43–54, Aug. 2017, doi: https://doi.org/10.1016/j.tourman.2016.12.022.

[3] S. Shayaa et al., "Sentiment Analysis of Big Data: Methods, Applications, and Open Challenges," IEEE Access, vol. 6, pp. 37807–37827, 2018, doi: https://doi.org/10.1109/access.2018.2851311.

[4] W. van Atteveldt, M. A. C. G. van der Velden, and M. Boukes, "The Validity of Sentiment Analysis: Comparing Manual Annotation, Crowd-Coding, Dictionary Approaches, and Machine Learning Algorithms," Communication Methods and Measures, vol. 15, no. 2, pp. 121–140, Jan. 2021, doi: https://doi.org/10.1080/19312458.2020.1869198.

[5] P. Yang and Y. Chen, "A survey on sentiment analysis by using machine learning methods," IEEE Xplore, Dec. 01, 2017. https://ieeexplore.ieee.org/abstract/document/8284920 (accessed Mar. 13, 2022).

[6] R. Man and K. Lin, "Sentiment Analysis Algorithm Based on BERT and Convolutional Neural Network," IEEE Xplore, Apr. 01, 2021. https://ieeexplore.ieee.org/abstract/document/9421110 (accessed Jul. 21, 2023).

[7] J. Qiu, C. Liu, Y. Li, and Z. Lin, "Leveraging sentiment analysis at the aspects level to predict ratings of reviews," Information Sciences, vol. 451–452, pp. 295–309, Jul. 2018, doi: https://doi.org/10.1016/j.ins.2018.04.009.

[8] A. Bagheri, M. Saraee, and F. de Jong, "An Unsupervised Aspect Detection Model for Sentiment Analysis of Reviews," Natural Language Processing and Information Systems, pp. 140–151, 2013, doi: https://doi.org/10.1007/978-3-642-38824-8_12.

[9] A. Ortigosa, J. M. Martín, and R. M. Carro, "Sentiment analysis in Facebook and its application to e-learning," Computers in Human Behavior, vol. 31, pp. 527–541, Feb. 2014, doi: https://doi.org/10.1016/j.chb.2013.05.024.

[10] O. Appel, F. Chiclana, J. Carter, and H. Fujita, "A hybrid approach to the sentiment analysis problem at the sentence level," Knowledge-Based Systems, vol. 108, pp. 110–124, Sep. 2016, doi: https://doi.org/10.1016/j.knosys.2016.05.040.

[11] A. Muhammad, N. Wiratunga, and R. Lothian, "Contextual sentiment analysis for social media genres," Knowledge-Based Systems, vol. 108, pp. 92–101, Sep. 2016, doi: https://doi.org/10.1016/j.knosys.2016.05.032.

[12] T. Chen, R. Xu, Y. He, Y. Xia, and X. Wang, "Learning User and Product Distributed Representations Using a Sequence Model for Sentiment Analysis," IEEE Computational Intelligence Magazine, vol. 11, no. 3, pp. 34–44, Aug. 2016, doi: https://doi.org/10.1109/mci.2016.2572539.

[13] M. Rossetti, F. Stella, and M. Zanker, "Analyzing user reviews in tourism with topic models," Information Technology & Tourism, vol. 16, no. 1, pp. 5–21, Nov. 2015, doi: https://doi.org/10.1007/s40558-015-0035-y.

[14] A. P. Jain and P. Dandannavar, "Application of machine learning techniques to sentiment analysis," IEEE Xplore, Jul. 01, 2016. https://ieeexplore.ieee.org/abstract/document/7912076 (accessed Mar. 15, 2021).

[15] L. Yang, Y. Li, J. Wang, and R. S. Sherratt, "Sentiment Analysis for E-Commerce Product Reviews in Chinese Based on Sentiment Lexicon and Deep Learning," IEEE Access, vol. 8, pp. 23522–23530, 2020, doi https://doi.org/10.1109/access.2020.2969854.

[16] H. S and R. Ramathmika, "Sentiment Analysis of Yelp Reviews by Machine Learning," IEEE Xplore, May 01, 2019. https://ieeexplore.ieee.org/document/9065812 (accessed Dec. 20, 2020).

[17] Z. Singla, S. Randhawa, and S. Jain, “Statistical and sentiment analysis of consumer product reviews,” IEEE Xplore, Jul. 01, 2017. https://ieeexplore.ieee.org/document/8203960 (accessed May 14, 2020).

[18] A. U. Rehman, A. K. Malik, B. Raza, and W. Ali, “A Hybrid CNN-LSTM Model for Improving Accuracy of Movie Reviews Sentiment Analysis,” Multimedia Tools and Applications, vol. 78, no. 18, pp. 26597–26613, Jun. 2019, doi: https://doi.org/10.1007/s11042-019-07788-7.

[19] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” arXiv.org, Oct. 11, 2018. https://arxiv.org/abs/1810.04805

[20] J. G. J. S. Raja and S. Juliet, “Deep Learning-based Sentiment Analysis of Trip Advisor Reviews,” IEEE Xplore, May 01, 2023. https://ieeexplore.ieee.org/abstract/document/10140848 (accessed Jul. 28, 2023).

[21] “Hotel Reviews,” www.kaggle.com. https://www.kaggle.com/datasets/datafiniti/hotel-reviews (accessed Jul. 28, 2023).

[22] D. Bahdanau, K. Cho, and Y. Bengio, “Neural Machine Translation by Jointly Learning to Align and Translate,” arXiv.org, 2014. https://arxiv.org/abs/1409.0473