# Literature Review

Sentiment analysis is a fundamental task of natural language processing used to determine whether the sentiments portrayed in data are positive, negative, or neutral. Sentiment analysis is often performed to help businesses to gather customer feedback on their products and service or whether comprehend customer needs. It has been proved to be a quite efficient text mining technique while aspect-level sentiment analysis is an even more improved technique often implemented on textual data to perform a detailed analysis on customer feedback that helps businesses to produce products and services that are up to customer satisfaction and needs.

Aspect-level sentiment analysis is a text mining technique that groups data by aspect and categorizes the sentiment to each aspect of data. Aspect-level sentiment analyses are performed to evaluate customer feedback by integrating certain opinions on several aspects of any product or service. Aspects are the components or attributes of a product such as a user's review after experiencing a specific product or service, any feature of a product, or customer service provided by an organization. Aspect level sentiment analysis is specifically the positive or negative opinion on a particular aspect that can be any feature or category of a product or service.

Recently, Xiaodi Wang et al., declared that the position features can improve the performance of the model on four different datasets. Therefore, the study proposes a classification model by combining attention mechanism and position features named Multi-Level Interactive Bidirectional Gated Recurrent Unit (MI-biGRU). Word embeddings are developed by position features of a word in a sentence so that the used approach can extract context and features of the targeted term by utilizing a well-constructed model, at last, to pay more attention to the words that are significant for sentiment classification an attention mechanism is used. The correlation between position features and the multilevel interactive attention network has been shown through experimental results [1].

A study focuses on retrieving aspect terms from each record extracted from Amazon customer review data by identifying the Parts-of-Speech and applying classification models including Naïve Bayes and Support vector machines to classify the sentiments from data. while performing the aspect level sentiment analysis feature terms of a product are the key target which depends on the product attributes. In the study, the aspect level sentiment analysis is categorized in three steps; identification, classification, and aggregation. Experiments are performed through applying preprocessing mainly Part-of-speech tagging to each term in a sentence and frequently used words are extracted. Later frequently used aspects are extracted from data through the Apriori feature extraction algorithm and opinion words such as a set of adjectives that best describe the aspect of the product are extracted after applying feature pruning. At last classification algorithms are applied to classify the sentiments into labels including positive, negative, and neutral [2].

In a study by Puspita Kencana Sari et al. Tokopedia review data is used for aspect-level sentiment analysis. The user reviews are classified based on five e-Servqual dimensions including; personalization, reliability, responsiveness, trust, web design, and sentiment analysis positive and negative. The naïve Bayes classification method is applied for the classification of the data because of its support to process large data and high-level accuracy [3].

Pratik P. Patile et al. focused on polarity prediction using a large data set extracted from Amazon customer reviews labeled into three classes such as positive, negative, or neutral. The classification task is performed by logistic regression, naïve Bayes, and support vector machines. Before performing classification, the data has been cleansed by applying some preprocessing methods including tokenization, stop word removal, and stemming. LDA and k-means algorithms are applied to extract and cluster the aspects or topics. The experimental results show that logistic regression has outperformed the other classifiers applied [4].

Ashwath et al. performed a study on aspect-level sentiment analysis by using different notion examination techniques on client audit data. data has been preprocessed by applying stemming, stop word removal techniques and POS tagging is used to extract mainly used terms from the data, that has been further classified by two states of the art machine learning models including, naïve Bayes and support vector machines where Naïve Bayes has the high accuracy score than the other model used [5].

A study proposed a pre-trained language model based on transformer bidirectional encoder representation named SA-BERT. Semantic information of the context of data is encoded into a word vector by Bert. While the text features were extracted by using the attention mechanism on a deeper level to comprehend the semantics of the text information and to accomplish the sentiment analysis task of e-commerce data. the experiments are performed on a JD mobile review dataset that shows the efficiency of the SA-BERT model in aspect-level sentiment analysis [6].

A study focused on developing a hybrid method for aspect-level sentiment analysis with the combination of machine learning approaches and lexicons. It has been proved that the word-level analysis by word cloud visualization provides primary results regarding the customer opinion on a product or service. For review classification, two lexicons named Syuzhet and Sentimentr are compared at the sentence level to classify the reviews. Syuzhet package has provided better performance and is used to train labeled text provided by it. The experimental results show that the naïve Bayes has a better accuracy score while using the Syuzhet lexicon package among other applied models [7].

Qiang Lu et al. proposed a model for aspect-level sentiment analysis named interactive rule attention network (IRAN). IRAN is proposed to designs a grammar rule encoder by standardizing the output of adjacent positions to simulates the grammatical functions in the sentence. It also learns attention information from context and target by constructing an interaction attention network. The experiments have been performed on the ACL 2014 Twitter dataset and SemEval 2014 dataset. It has been shown that the IRAN can learn effective features and has shown better performance as compared to the traditional models [8].

Another study presents a continuous learning framework based on naïve Bayes for sentiment classification of a large-scale and multi-domain e-commerce product review. The parameter estimation mechanism has been extended in naïve Bayes to support continuous learning. The experiments are performed in two different domains including; cross-domain sentiment classification and Domain-specific sentiment classification. In cross-domain classification, reviews are taken from different domains, and in domain-specific classification, reviews are taken from the same domain [9].

Yue Han et al. proposed a Pretraining and Multi-task learning model in their study based on Double BiGRU namely PM-DBiGRU. In the proposed model short text-level drug review are used to learn pretrained weight to initialize related weight for the model for sentiment classification task model. After that two BiGRU networks are executed to produce the bidirectional semantic representations of the drug and target review, and target-specific representation is obtained attention mechanism for aspect-level sentiment classification of drug review. To transfer helpful domain knowledge from the corpus multi-task learning is further utilized [10].

A study used a mobile product review dataset for aspect-level sentiment classification. The data has been initially preprocessed by using tokenization and POS tagging. POS tagging has been used to extract the most frequent terms from the data. later the classification has been performed by using four different machine learning models including Bernoulli Naïve bayesian, multinomial naïve Bayesian, k nearest neighbor, and Support vector machines. The experimental results show that the KNN has the highest accuracy among all the other models on APPLE products and has performed comparatively better than the other classifiers [11].

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Ref No.** | **Year of Publication** | **Dataset** | **Methods** | **Accuracy** |
| **1** | 2020 | SemEvals dataset series; SemEvals 2014 laptop sales, and SemEvals 2014, 2015, 2016 restaurants feedback. | Multi-Level Interactive Bidirectional Gated Recurrent Unit (MI-biGRU) | 82.54 |
| **2** | 2018 | Amazon customer review data | POS tagging, Apriori feature extraction algorithm, SVM and NB classifiers | 90.42 |
| **3** | 2018 | Tokopedia customer review data | five e-Servqual dimensions and naïve bayes classifier | 90 |
| **4** | 2019 | Amazon review data | LDA and k-means algorithms used to extract topics, NB, SVM and LR are used for classification | 81.5 |
| **5** | 2019 | Client audit data | POS tagging, NB, SVM | 81.30 |
| **6** | 2020 | JingDong mobile review data | SA-Bert | 91.35 |
| **7** | 2020 | Amazon review data | Syuzhet and Senti-mentr lexicons with NB, KNN, SVM classifiers | 86.3 |
| **8** | 2020 | SemEval 2014 and ACL 2014 Twitter Dataset datasets | Interactive rule attention network (IRAN) | 81.96 |
| **9** | 2020 | Amazon Product and movie review dataset | Naïve Bayes using domain-specific and cross-domain sentiment classification | 75.68 |
| **10** | 2020 | SentiDrugs dataset | PM-DBiGRU | 78.26 |
| **11** | 2020 | Mobile review dataset | Bernoulli Naïve bayesian, multinomial naïve Bayesian, k nearest neighbor, Support vector machines | 91.5 |

# References

1. Wang, X., Chen, X., Tang, M., Yang, T. and Wang, Z., 2020. Aspect-Level Sentiment Analysis Based on Position Features Using Multilevel Interactive Bidirectional GRU and Attention Mechanism. *Discrete Dynamics in Nature and Society*, *2020*.
2. Vanaja, S. and Belwal, M., 2018, July. Aspect-level sentiment analysis on e-commerce data. In *2018 International Conference on Inventive Research in Computing Applications (ICIRCA)* (pp. 1275-1279). IEEE.
3. Sari, P.K., Alamsyah, A. and Wibowo, S., 2018, March. Measuring e-Commerce service quality from online customer review using sentiment analysis. In *Journal of Physics: Conference Series* (Vol. 971, No. 1, p. 012053). IOP Publishing.
4. Patil, P.P., Phansalkar, S. and Kryssanov, V.V., 2019. Topic modelling for aspect-level sentiment analysis. In *Proceedings of the 2nd International Conference on Data Engineering and Communication Technology* (pp. 221-229). Springer, Singapore.
5. ASHWATH, K. and VINOTH, J., Aspect Based Sentiment Analysis for E-Commerce Using Classification Techniques.
6. Cai, X., Cao, H. and Ma, J., Sentiment Analysis of E-commerce Comments Based on BERT.
7. Marshan, A., Kansouzidou, G. and Ioannou, A., 2020, November. Sentiment Analysis to Support Marketing Decision Making Process: A Hybrid Model. In *Proceedings of the Future Technologies Conference* (pp. 614-626). Springer, Cham.
8. Lu, Q., Zhu, Z., Zhang, D., Wu, W. and Guo, Q., 2020. Interactive rule attention network for aspect-level sentiment analysis. *IEEE Access*, *8*, pp.52505-52516.
9. Xu, F., Pan, Z. and Xia, R., 2020. E-commerce product review sentiment classification based on a naïve Bayes continuous learning framework. *Information Processing & Management*, *57*(5), p.102221.
10. Han, Y., Liu, M. and Jing, W., 2020. Aspect-level drug reviews sentiment analysis based on double BiGRU and knowledge transfer. *IEEE Access*, *8*, pp.21314-21325.
11. Abraham, M.P. and KR, U.K.R., 2020. Feature Based Sentiment Analysis of Mobile Product Reviews using Machine Learning Techniques. *International Journal*, *9*(2).