JURNAL ECTIPE
Electronic, Control, Telecommunication, Information, and Power Engineering
Volume 12, Issue 2, October 2025, pp. 151-158
ISSN 2355-5068 ; e-ISSN 2622-4852
DOI: 10.33019/jurnalecotipe.v12i2.4561

Comparative Study of Sentiment Analysis for Interpreting the Customer Interest in Women Fashion Clothes

Eka Legya Frannita[1], Alifia Revan Prananda[2], Marwanto[3]
[1]Departement of Leather Product Processing Technology, Politeknik ATK Yogyakarta, Indonesia
[2]Departement of Information Technology, Faculty of Engineering, Universitas Tidar, Magelang, Indonesia
[3]Departement of Dance Education, Faculty of Language and Arts, Universitas Negeri Yogyakarta, Yogayakrta, Indonesia

ARTICLE INFO

Article historys:

Received : 28/04/2025
Revised : 04/06/2025
Accepted : 30/10/2025

Keywords:

Comparative Study; Customer Review; E-Commerce; Machine Learning; Sentiment Analysis

ABSTRACT

Public sentiment is widely recognized as a key factor influencing fluctuations in stock prices, product sales, and emerging trends. Since the user interest analysis played an essential role in representing the trend of the market and it is also extremely useful for generating the strategy and decision of trading, analyzing the public sentiment is quite important. In the contemporary era, virtual space has been perpetually evolving and used in plenteous applications including analyzing customer interest. This research work aimed to conduct comparative study in analyzing customer interest about women fashion clothes using sentiment analysis method-based machine learning approach. The proposed study was initiated by conducting data acquisition process. It was then continued with labelling and predicting the user sentiment by comparing eleven machine learning approaches. According to comparison result, Naïve Bayes successfully obtained the best performance with accuracy of 94%, precision of 87%, recall of 82%, f1-score of 84%. It can be inferred that Naïve Bayes was viable approach for predicting the user sentiment.

Corresponding Author:

Eka Legya Frannita
Department of Leather Product Processing Technology, Politeknik ATK Yogyakarta, Yogyakarta, Indonesia
Email: eka.legya@atk.ac.id

1. INTRODUCTION

The contemporary era has introduced how almost all human activities have been occurred in the virtual space [1]. In the current decade, it has been proven that online services, such as online discussion, e-commerce, e-learning, social media, etc., have become an effective way for almost all purposes [2–4]. It has accelerated significantly due to the frequent entrance of information that has giving effect to the development of innovation technology [5]. Thanks to the rapid growth of technology as a fresh wave of digitalization that has been fundamentally assisted human work and inspired to develop the improved technologies or even multifunction of technologies that offer wide range of uses [6].

E-commerce is one of the online services offering several facilities for encouraging and supporting customer needs [5]. In the last two years, the pace of e-commerce transaction has increased tremendously due to a high user population during the COVID-19 pandemic. In that situation, e-commerce has successfully assisted the customers for providing all needs such as foods, clothes, shoes, medicines, or other daily needs. It was then giving impact to the increasing of company profit [7].

Volume 12, Issue 2, October 2025, pp. 151-158
ISSN 2355-5068 ; e-ISSN 2622-4852
DOI: 10.33019/jurnalecotipe.v12i2.4561

Regardless that the e-commerce traffic was significantly increased in the last two years, a company who has employee the e-commerce still needs to analyse how does the market trend. It was needed to support the selling process [7]. In the new era, company also can use e-commerce to analysis the customers perspective about their products by using sentiment analysis. It was performed by analysing the customers opinion or review in market place or social media which was significantly useful for their selling process [8]. Sentiment analysis not only can give the companies customers perspective, but it also can give the companies some advises regarding to the customer needs or even somethings that make them feel satisfy or disappointed. This advice was extremely useful to create the selling strategies. Sentiment analysis results also can be used as a database for creating a prediction about what the future market trends [9–12].

Sentiment analysis offers several benefits across various domains due to its ability to analyse and interpret the emotional tone and opinions expressed in text. Businesses can utilize customers sentiment for analysing customer responses, comments, and feedback. By understanding customer sentiments, companies can recognize areas needing enhancement, respond to customer feedback, and elevate overall service quality. Sentiment analysis helps companies monitor and manage their brand reputation online. By analysing social media mentions, reviews, and news articles, organizations can gain insights into how their brand is perceived and take corrective actions if needed. Understanding customer sentiments about products or services can guide product development and service improvements. Positive sentiments can be reinforced, and negative sentiments can be addressed to enhance the overall quality and satisfaction levels. Sentiment analysis is valuable in market research to gain public response about new products, advertising campaigns, or industry trends. This information aids companies in making informed business decisions and staying competitive in the market. Sentiment analysis is widely used for monitoring social media platforms to understand public opinions, trends, and discussions. This is crucial for social media managers, public relations teams, and marketing professionals to engage with the audience effectively [10]. In the financial industry, sentiment analysis is used to analyse news articles, financial reports, and social media discussions related to stocks and investments. This information can help investors make more informed decisions by considering market sentiment. Sentiment analysis can be integrated into customer support systems to automatically categorize and prioritize customer queries based on sentiment. This enables companies to provide timely and appropriate responses, enhancing customer support and engagement. Sentiment analysis is applied to analyse public opinions and sentiments regarding political candidates, parties, or policies. This can be useful for political campaigns and strategists to understand the public mood and tailor their messaging accordingly. In the healthcare industry, sentiment analysis can be applied to patient feedback, reviews of healthcare providers, and discussions on social media. This information can be valuable for improving healthcare services and patient experiences. Sentiment analysis can be used internally to analyse employee feedback, surveys, and sentiments expressed in communication channels. This information can be instrumental in improving workplace satisfaction and addressing employee concerns. Overall, sentiment analysis provides actionable insights by automating the analysis of vast amounts of text data, allowing organizations to make data-driven decisions, enhance user experiences, and respond effectively to changing sentiments in various contexts [13].

Sentiment analysis can be classified into different types based on the scope of analysis, the nature of sentiment, and the desired outcomes. Binary sentiment analysis aimed to determine whether the sentiment expressed in the text is positive or negative. Binary sentiment analysis is commonly used for applications where a simple positive/negative classification is sufficient, such as product reviews or social media sentiment. Multi-class sentiment analysis classified sentiment into multiple classes, such as positive, negative, neutral, and sometimes more fine-grained categories. It useful in scenarios where a more nuanced understanding of sentiment is required, such as customer feedback analysis with multiple sentiment categories. Aspect-based sentiment analysis analysed sentiment at a more granular level by identifying sentiments associated with specific aspects or features mentioned in the text. It was valuable for product reviews, where sentiment is analysed not only at an overall level but also for specific product features or aspects. Fine-grained sentiment analysis provided a more detailed sentiment analysis by categorizing sentiments into multiple fine-grained categories, such as very positive, positive, neutral, negative, very negative. It was useful when a more nuanced understanding of sentiment intensity

JURNAL ECOTIPE
Electronic, Control, Telecommunication, Information, and Power Engineering
Volume 12, Issue 2, October 2025, pp. 151-158
ISSN 2355-5068 ; e-ISSN 2622-4852
DOI: 10.33019/jurnalecotipe.v12i2.4561

is required, allowing for a more detailed analysis of the sentiment expressed. Emotion detection goes beyond basic sentiment analysis to identify specific emotions conveyed in the text, such as joy, anger, sadness, or surprise. It was applied in scenarios where understanding the emotional tone of the text is crucial, such as social media monitoring for brand sentiment during a marketing campaign. Intent analysis determined the intent behind a piece of text, such as whether it expresses a desire, request, complaint, or appreciation. It was helpful in understanding the underlying intentions of users, which can be valuable for customer support and interaction with chatbots. Temporal sentiment analysis analysed the changes in sentiment over time, tracking how sentiments evolve or fluctuate. It was useful for monitoring public opinion during events, product launches, or marketing campaigns to understand the impact over time. Cross-domain sentiment analysis involved training sentiment analysis models on data from one domain and applying them to another domain. It was enabled sentiment analysis models to adapt to different industries or domains, even when labelled data for the specific domain is limited. Domain-Specific Sentiment Analysis focused on sentiment analysis within a specific industry or domain, considering domain-specific language and context. It customized sentiment analysis models for industries like healthcare, finance, or technology where language and sentiments may be unique [14].

Generally, the sentiment analysis process can be conducted using two famous methods which were lexicon approach and supervised learning based. Both of them had almost similar procedure to analyse the customer perspective. However, in the lexicon approach we have to provide the emotional words expressing the customers satisfaction about the products or the services which was difficult enough if we used the unstructured data like review or social media data. On the other hand, supervised learning approach worked by learning the customers expression and grouped the data that have similar meaning to analyse and interpret them. Hence, for the advanced analysis, supervised learning approach is more effective than lexicon approach [5].

This research work aims to develop sentiment analysis method for interpreting the women interest in fashion clothes. Customers opinion and reviews completed with rating score were employed and analysed to produce the customers sentiments. The analysis process had occurred by performing the following contributions.

1. Conducting comparison study of several classifier for interpreting the women interest in fashion clothes through their opinion, review or even the customer rate in the e-commerce.
2. Generating the rank of customers' interest in fashion clothes by considering the customers impression through their opinion, review or even the customer rate which was important for the companies to create a suitable selling strategy.

The rest of this article was structured as follow: section 2 defines the previous related research, section 3 describes about the data, methodology and evaluation method, section 4 illustrates the result and discussion, and section 5 presents the conclusion that was generated regarding to the results.

2. DATA AND RESEARCH METHOD

2.1. Brief Information about Dataset

This study used online e-commerce user review proposed by Agarap [15]. The dataset contained of 23,486 customer response completed with 10 features illustrated the data. The ten features were clothing ID, age, title, review text, rating (1 to 5), recommended IND, positive feedback count, division name, department name and class name. Example dataset is illustrated in following table.

Table 1. Example of data

No.	Review Text	…	Class Name
1	I had such high hopes for this dress and ….	…	Dress
2	Love this dress! it's sooo pretty. I happened...	…	Dress
3	I love, love, love this jumpsuit. it's fun, fl...	…	Pants
4	I love tracy reese dresses, but this one is not for the very petite. i am just under 5 feet tall and...	…	Dress
…	…	…	…
23486	This shirt is very flattering to all due…..	…	Blouse

JURNAL ECOTIPE
Electronic, Control, Telecommunication, Information, and Power Engineering
Volume 12, Issue 2, October 2025, pp. 151-158
ISSN 2355-5068 ; e-ISSN 2622-4852
DOI: 10.33019/jurnalecotipe.v12i2.4561

2.2. Scenario of Experiments

The experiment was started by preparing the dataset. In the beginning process, we conducted data acquisition aimed to remove unnecessary data. As illustrated in Table 1, there are several data that are not valid (illustrated by "Nan"). Those data existed in the title column. Furthermore, several features were not necessary in the analysis process, hence we have to remove those several features. The remained features that were used in the analysis process were review text, rating, class name, and age. Then, we conducted word count process to calculate the number of words in every review and to calculate how many those words were used. Hence, we got the new combination of data with the following features: review text, rating, class name, age and word count.

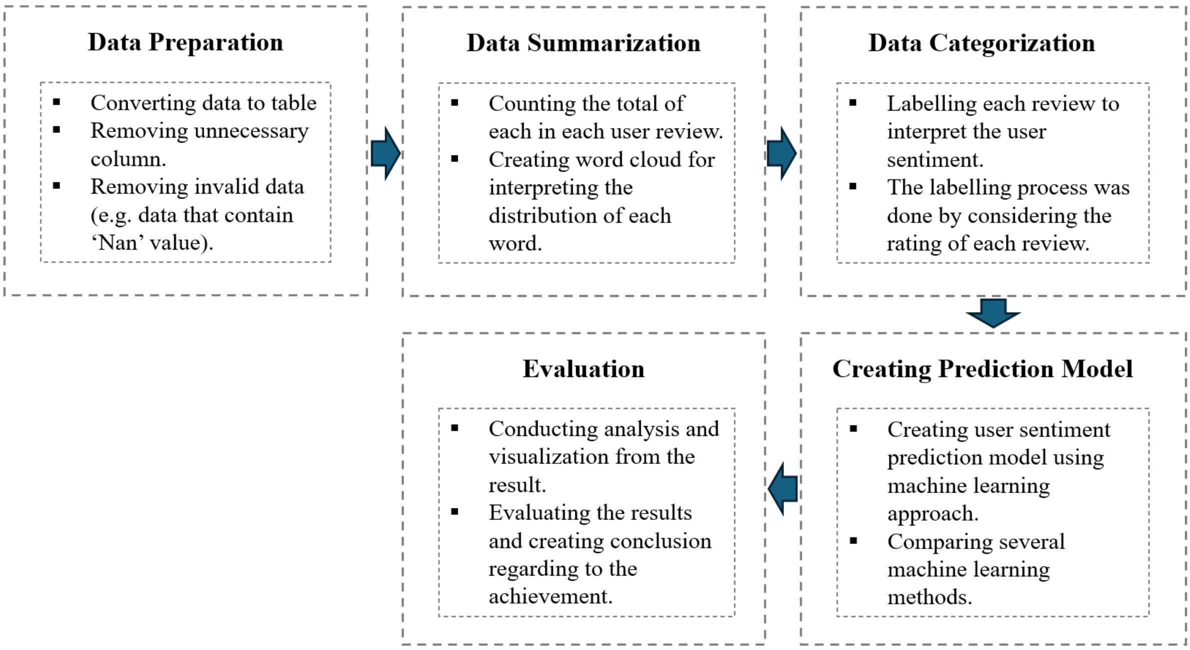


Figure 1. Block diagram of experiment

After preparing the data, we continued the process by labelling each review in three categories (positive, neutral and negative). The review which had 5 or 4 rating were labelled as positive sentiment. Review contained of 3 rating were labelled as neutral sentiment. While, review that contained 1 or 2 rating were labelled as negative sentiment. Then, the process was continued by conducting classification process. In this process, we used eleven classifiers in order to find the best machine learning method for predicting the user sentiment. Overall processes can be seen in the Figure 1.

3. RESULTS AND DISCUSSION

In the first process, we conducted data accusation process to remove some unnecessary features and created the word count to calculate the number of words. The result of this process can be defined in Table 2. In this process, we also calculated the number of each interesting word and created the word cloud to illustrate how the distribution of user interest. The illustration of word cloud and the calculated word are depicted in Figure 2.

Figure 2 illustrates the distribution of total user mention in every product. There are several women fashion clothes mentioned in the user review such as blouse, casual bottom, chemises, dress, jacket, jeans, knit, layering, legwear, outerwear, pant, short, skirt, sleepwear, sweater, swimming wear, and other fashion clothes. This distribution of total user mention in every product was then used to determine the most popular product and to create word cloud in order to visualize how the popularity of each product.

JURNAL ECOTIPE
Electronic, Control, Telecommunication, Information, and Power Engineering

Volume 12, Issue 2, October 2025, pp. 151-158
ISSN 2355-5068 ; e-ISSN 2622-4852
DOI: 10.33019/jurnalecotipe.v12i2.4561

Table 2. Result of word count process

ID.	Review Text	Rating	Class Name
1	I had such high hopes for this dress….	3	{'and': 3, 'be': 1, 'bottom': 1, 'but': 2, 'ch...
2	Love this dress! it's sooo pretty. I happened...	5	{'am': 1, 'and': 2, 'bc': 2, 'be': 1, 'below':...
3	I love, love, love this jumpsuit. it's fun, fl...	5	{'and': 1, 'but': 1, 'compliments': 1, 'every'...
…	…	…	…
23486	This shirt is very flattering to all due…..	5	{'adjustable': 1, 'all': 1, 'and': 1, 'any': 1...

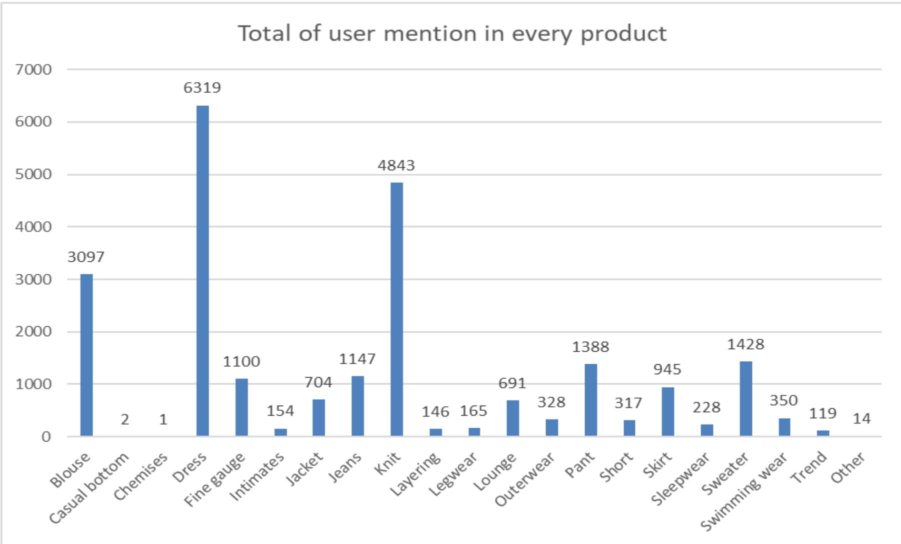


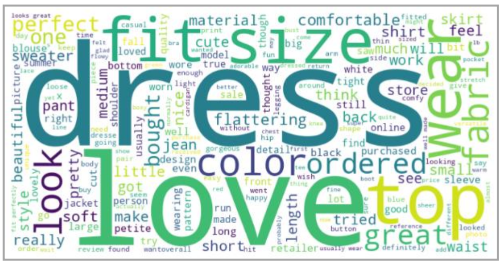
Figure 2. Distribution of user mention in each women fashion clothes

In this step, we not only counted the total of user mention in every product, but we also counted the total of each term that illustrated the satisfaction of the customers. The example words were 'love'. 'pretty', 'great', 'super', 'happy', 'disappointed', etc. This process was used to evaluate the satisfaction of the customer. Result of this step is depicted in Figure 3(a). This figure illustrates the top mentioned words which illustrates how was the user satisfaction about each product.

```
Selected Words
love      8951
great     6117
super     1726
happy      705
glad       614
dtype: int64

Class Names
Dresses       6319
Knits         4843
Blouses       3097
Sweaters      1428
Pants         1388
Name: Class Name, dtype: int64
```

(a)



(b)

Figure 3. Data acquisition results: (a) the number of each word and (b) word cloud

Result in Table 3 illustrates how every word in every user review was counted. This process was conducted to analyze the most interesting women fashion clothes and how did the satisfaction of each user. After conducting word counting in each user review. We evaluated the counted result as depicted in Figure 3. In the Figure 3(a), we can analyze the most popular women fashion clothes and how did their opinion about the products. According to Figure 3(a), we found that there are five the most type of

JURNAL ECOTIPE
Electronic, Control, Telecommunication, Information, and Power Engineering
Volume 12, Issue 2, October 2025, pp. 151-158
ISSN 2355-5068 ; e-ISSN 2622-4852
DOI: 10.33019/jurnalecotipe.v12i2.4561

women fashion clothes which are dresses, knits, blouses, sweaters, and pants. Dresses became the most mentioned women fashion clothes with the total mention of this type of clothes of 6319 mentions. Knits wear was in the second position with the total mention of 4843 mentions. Blouses was in the third position with the total mention of 3097 mentions. Sweaters was in the fourth position with the total mention of 1428 mentions. The last type of clothes, pants, was in the fifth position with the total mention of 1388 mentions. Figure 3(a) also illustrates the total of each word mentioned in the user review. In this process, we found five the most words that have been mentioned in the user review. The term 'love' became the most mentioned word that was mentioning by the user with the total mention of 8951 words. The term 'great' was in the second position with the mention of 6117 words. The term 'super' was in the third position with the mention of 1726 words. The term 'happy' was in the fourth position with the mention of 705 words. The term 'glad' was in the fifth position with the mention of 614 words. According to Figure 3(a) we can concluded that the users were satisfy about almost all product that were reviewed. It was indicated by the word counting result which the five most mentioned word were referred to the positive review. Based on the Figure 3(a) we also can concluded that most loved women fashion clothes were dresses. It was indicated by the total mention in the review data in which the term 'dress' was mentioned more than 6000 times. In the second position, there was knit wear that became the most mentioned clothes type after dresses. In the third position, there was blouses followed by sweaters and pants.

Figure 3(b) illustrates the word cloud as the visualization of results in Figure 3(a). Word cloud was used to illustrate the distribution of total mention of each word. For example, in the term of 'dress', in Figure 3(b) the term of 'dress' becomes the biggest word compared with other which means that the term 'dress' obtained the most mentioned word. Finally, according to Figure 3(a) and 3(b), we can summarize that the most interesting women fashion clothes based on the customer review is dress wear.

In the next process, we labelled the data into three categories which were positive, neutral and negative. Then, we predicted the user sentiment in several classifier. In this study, we performed 11 classifiers. This process was used to predict how are the customer sentiment regarding to their review. By classifying each customer review in three categories (positive, neutral and negative) we can understand what the customer perspective to the product. According to labelled process, we found that 18,208 review indicated positive sentiment, 2,871 reviews indicated neutral sentiment, and 2,407 reviews indicated negative sentiment. Those data were then trained to create prediction model. The following table illustrates the comparison result of all classifier.

Table 3. Result of word count process

Method	Accuracy	Precision	Recall	F1-score
Logistic regression	94%	87%	81%	83%
Naïve Bayes	**94%**	**87%**	**82%**	**84%**
SVM	93%	88%	72%	77%
Neural Network	93%	85%	80%	82%
Stochastic Gradient Descent	92%	79%	80%	79%
K-Nearest Neighbors	89%	81%	51%	49%
Decision Tree	90%	76%	62%	66%
Random Forest	90%	76%	61%	64%
SVM (Linear Kernel)	94%	88%	78%	81%
SVM (Polynomial Kernel)	89%	44%	50%	47%
SVM (RBF Kernel)	89%	44%	50%	47%

Table 3 illustrates the comparison results between eleven classifiers in predicting the user sentiment. According to Table 2, Naïve Bayes achieved the best performance with accuracy of 94%, precision of 87%, recall of 82%, f1-score of 84%. Hence, it can be concluded that Naïve Bayes was suitable for classifying customer review in three classes which are positive sentiment, neutral sentiment and negative sentiment. This result indicated that Naïve Bayes was significant for interpreting and predicting the user interest in the cases of women fashion clothes reviews.

Volume 12, Issue 2, October 2025, pp. 151-158
ISSN 2355-5068 ; e-ISSN 2622-4852
DOI: 10.33019/jurnalecotipe.v12i2.4561

4. CONCLUSION

This research work aimed to conduct comparison study in development of sentiment analysis method-based machine learning approach which was performed in the online user review of women fashion clothes. The proposed study was started by conducting data acquisition process. It was then continued with labelling and predicting the user sentiment by comparing eleven machine learning approach. According to comparison result, Naïve Bayes successfully obtained the best performance with accuracy of 94%, precision of 87%, recall of 82%, f1-score of 84%. It can be inferred that Naïve Bayes was viable approach for predicting the user sentiment.

REFERENCES

[1] J. An and W. M. N. Wan Zainon, "Integrating color cues to improve multimodal sentiment analysis in social media," *Eng. Appl. Artif. Intell.*, vol. 126, p. 106874, 2023, doi: https://doi.org/10.1016/j.engappai.2023.106874.

[2] H. Wang and M. Hou, "Quantum-like implicit sentiment analysis with sememes knowledge," *Expert Syst. Appl.*, vol. 232, p. 120720, 2023, doi: https://doi.org/10.1016/j.eswa.2023.120720.

[3] P. Hajek, L. Hikkerova, and J.-M. Sahut, "Fake review detection in e-Commerce platforms using aspect-based sentiment analysis," *J. Bus. Res.*, vol. 167, p. 114143, 2023, doi: https://doi.org/10.1016/j.jbusres.2023.114143.

[4] P. Savci and B. Das, "Prediction of the customers' interests using sentiment analysis in e-commerce data for comparison of Arabic, English, and Turkish languages," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 35, no. 3, pp. 227–237, 2023, doi: https://doi.org/10.1016/j.jksuci.2023.02.017.

[5] M. Demircan, A. Seller, F. Abut, and M. F. Akay, "Developing Turkish sentiment analysis models using machine learning and e-commerce data," *Int. J. Cogn. Comput. Eng.*, vol. 2, pp. 202–207, 2021, doi: https://doi.org/10.1016/j.ijcce.2021.11.003.

[6] J. Wang and J.-J. Lee, "Predicting and analyzing technology convergence for exploring technological opportunities in the smart health industry," *Comput. Ind. Eng.*, vol. 182, p. 109352, 2023, doi: https://doi.org/10.1016/j.cie.2023.109352.

[7] R. K. Das, M. Islam, M. M. Hasan, S. Razia, M. Hassan, and S. A. Khushbu, "Sentiment analysis in multilingual context: Comparative analysis of machine learning and hybrid deep learning models," *Heliyon*, vol. 9, no. 9, p. e20281, 2023, doi: https://doi.org/10.1016/j.heliyon.2023.e20281.

[8] R. V Karthik and S. Ganapathy, "A fuzzy recommendation system for predicting the customers interests using sentiment analysis and ontology in e-commerce," *Appl. Soft Comput.*, vol. 108, p. 107396, 2021, doi: https://doi.org/10.1016/j.asoc.2021.107396.

[9] F. Xu, Z. Pan, and R. Xia, "E-commerce product review sentiment classification based on a naïve Bayes continuous learning framework," *Inf. Process. Manag.*, vol. 57, no. 5, p. 102221, 2020, doi: https://doi.org/10.1016/j.ipm.2020.102221.

[10] P. Atandoh, F. Zhang, D. Adu-Gyamfi, P. H. Atandoh, and R. E. Nuhoho, "Integrated deep learning paradigm for document-based sentiment analysis," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 35, no. 7, p. 101578, 2023, doi: https://doi.org/10.1016/j.jksuci.2023.101578.

[11] M. Nakayama and Y. Wan, "The cultural impact on social commerce: A sentiment analysis on Yelp ethnic restaurant reviews," *Inf. Manag.*, vol. 56, no. 2, pp. 271–279, 2019, doi: https://doi.org/10.1016/j.im.2018.09.004.

[12] N. Ramshankar and J. P. P.M., "Automated sentimental analysis using heuristic-based CNN-BiLSTM for E-commerce dataset," *Data Knowl. Eng.*, vol. 146, p. 102194, 2023, doi: https://doi.org/10.1016/j.datak.2023.102194.

Volume 12, Issue 2, October 2025, pp. 151-158
ISSN 2355-5068 ; e-ISSN 2622-4852
DOI: 10.33019/jurnalecotipe.v12i2.4561

[13] P. Kathuria, P. Sethi, and R. Negi, "Sentiment analysis on E-commerce reviews and ratings using ML & NLP models to understand consumer behavior," in *2022 International Conference on Recent Trends in Microelectronics, Automation, Computing and Communications Systems (ICMACC)*, 2022, pp. 1–5. doi: 10.1109/ICMACC54824.2022.10093674.

[14] S. S. Sohail, J. Siddiqui, and R. Ali, "Feature extraction and analysis of online reviews for the recommendation of books using opinion mining technique," *Perspect. Sci.*, vol. 8, pp. 754–756, 2016, doi: https://doi.org/10.1016/j.pisc.2016.06.079.

[15] A. F. Agarap, "Women's E-Commerce Clothing Reviews," *Kaggle Dataset*, 2017. https://www.kaggle.com/datasets/nicapotato/womens-ecommerce-clothing-reviews