Adaptive sentiment analysis using multioutput classification: a performance comparison

Taqwa Hariguna [Corresp., 1], **Athapol Ruangkanjanases** [Corresp. 2]

[1] Information Systems, Universitas Amikom Purwokerto, Purwokerto, Jawa Tengah, Indonesia

[2] Department of Commerce Chulalongkorn Business School, Chulalongkorn University, Bangkok, Thailand

Corresponding Authors: Taqwa Hariguna, Athapol Ruangkanjanases
Email address: taqwa@amikompurwokerto.ac.id, athapol@cbs.chula.ac.th

The primary objective of this research is to create a multi-output classification model for sentiment analysis through the combination of 10 algorithms: BernoulliNB, Decision Tree, K-nearest neighbor, Logistic Regression, LinearSVC, Bagging, Stacking, Random Forest, AdaBoost, and ExtraTrees. In doing so, we aim to identify the optimal algorithm performance and role within the model. The data utilized in this study is derived from customer reviews of cryptocurrencies in Indonesia. Our results indicate that LinearSVC and Stacking exhibit a high accuracy (90%) compared to the other eight algorithms. The resulting multi-output model demonstrates an average accuracy of 88%, which can be considered satisfactory. This research endeavors to innovate in adaptive sentiment analysis classification by developing a multi-output model that utilizes a combination of 10 classification algorithms

1 # Adaptive Sentiment Analysis Using Multioutput
2 # Classification: A Performance Comparison

4

6 ## Abstract

7 The primary objective of this study is to create a multi-output classification model for sentiment
8 analysis using a combination of 10 algorithms: BernoulliNB, Decision Tree, K-nearest neighbor,
9 Logistic Regression, LinearSVC, Bagging, Stacking, Random Forest, AdaBoost, and ExtraTrees.
10 In doing so, we aim to identify the optimal algorithm performance and role within the model. The
11 data used in this study were derived from customer reviews of cryptocurrencies in Indonesia. Our
12 results indicate that LinearSVC and Stacking exhibit high accuracy (90%) compared to the other
13 eight algorithms. The resulting multioutput model demonstrated an average accuracy of 88%,
14 which can be considered satisfactory. This research endeavors to innovate adaptive sentiment
15 analysis classification by developing a multioutput model that utilizes a combination of 10
16 classification algorithms.

18

19 ## Introduction

20 One method for increasing the accuracy of sentiment analysis is the development of multi-output
21 models (Wu et al., 2019; Alfazzi, 2022). Sentiment analysis, a subfield of natural language
22 processing (NLP), seeks to extract and interpret the emotions or attitudes conveyed in a text. It is
23 commonly utilized in marketing, public opinion, and social media (Kabir et al., 2021).

24 There are various methods for determining the sentiment of a text, including the use of multi-
25 output models. A multi-output model is capable of predicting multiple output classes
26 simultaneously. In the context of sentiment analysis, such models can be utilized to predict the
27 overall sentiment of a text and identify the specific type of sentiment present (e.g., positive,
28 negative, or neutral) (Kaur et al., 2021; Hayadi et al., 2023).

29 The significance of developing multi-output models for sentiment analysis lies in their ability to
30 enhance the accuracy of sentiment analysis. Single-output models, which can only predict a single
31 output class (e.g., only positive or negative), may struggle to detect more subtle forms of sentiment,
32 such as irony or sarcasm (Han et al., 2020). By contrast, multi-output models can more accurately
33 identify the specific type of sentiment in a text, enabling more informed decision-making based
34 on the sentiment conveyed (Sun et al., 2019; Saputra et al. 2023; Cao et al., 2023). In sentiment
35 analysis, multiple outputs refer to a situation where a model must output multiple decisions or
36 labels for each input. A combination of algorithms is often employed in various studies to improve
37 the accuracy of multi-output models in sentiment analysis.

38   The ensambling method is a commonly utilized approach in multiple-outcome sentiment analysis
39   research. This method involves the creation of multiple models using different algorithms, with
40   the decisions of each model serving as new features for a final model (model c) of the same type.
41   The final model can then learn from the decisions of the previous models and adapt accordingly,
42   leading to improved accuracy.

43   Han et al. (2020) utilized sentiment analysis and machine learning techniques in conjunction with
44   one another to predict stock market movements. The algorithms employed included support vector
45   machines (SVM) and Naive Bayes. Previous research has demonstrated that this combination of
46   algorithms exhibits higher accuracy than using a single algorithm. In a study conducted by Sun et
47   al. (2019), an ensemble algorithm consisting of multiple algorithms was utilized to perform
48   sentiment analysis on social media data. The algorithms employed included Random Forest,
49   Decision Tree, and Naive Bayes, and the results indicated that combining these algorithms was
50   more effective than using a single algorithm.

51   Mukhtar et al. (2018) combined deep learning and feature selection techniques to perform
52   sentiment analysis, employing algorithms such as convolutional neural networks (CNN) and
53   logistic regression. The results indicated that the combination of these algorithms exhibited higher
54   accuracy than using a single algorithm. Similarly, in a study by Miao et al. (2021), feature selection
55   techniques and classification algorithms were combined to perform sentiment analysis, including
56   recursive feature elimination (RFE) and SVM. The results demonstrated that this combination of
57   algorithms had higher accuracy than using a single algorithm.

58   Salazar et al. (2022) combined deep learning techniques with rule-based approaches to perform
59   sentiment analysis on Twitter data, utilizing algorithms such as long short-term memory (LSTM)
60   and decision tree. The results indicated that the combination of these algorithms exhibited higher
61   accuracy than using a single algorithm.

62   The innovative aspect of this research lies in the combination of 10 algorithms in a multi-output
63   classification model, a practice that is uncommon in various studies, particularly in the field of
64   sentiment analysis. The primary objectives of this study are to demonstrate the effectiveness of the
65   combination of 10 algorithms and to compare the performance of the individual algorithms to
66   identify the optimal algorithm.

67   **Materials & Methods**
68   Figure 1 illustrates the process of developing the sentiment data classification model used in this
69   study:

70   [Figure 1 here]

71   Figure 1. Research Steps

72 *Identification of the Problem and Research Objectives*

73 The aim of this research is to develop a model capable of performing sentiment analysis through
74 the combination of 10 algorithms: BernoulliNB, Decision Tree, K-Nearest Neighbor, Logistic
75 Regression, LinearSVC (Support Vector Classification), Bagging, Stacking, Random Forest,
76 AdaBoost, and ExtraTrees. Once the model has been developed, a comparison will be conducted
77 to identify the algorithm with the highest level of performance in sentiment analysis. The findings
78 of this study will aid in the improvement of the model's ability to perform sentiment analysis and
79 provide insight into the strengths and limitations of each algorithm utilized.

80 *Data Collection*

81 The data collection method used in this study is data collection from platforms that provide
82 customer reviews on two cryptocurrency exchanges in Indonesia: Indodax and Binance. Data
83 collection was performed by accessing the customer reviews available on the platform. The
84 amount of data collected is 1599 data points that have been prepared in terms of characteristics
85 and normalization. The data collected is considered valid and relevant for this research. In addition,
86 the data also undergoes a normalization process to facilitate analysis and interpretation. Data
87 normalization is carried out with the aim of eliminating scale differences between data and
88 equalizing the magnitude of data values in one particular attribute. This is important because data
89 that has differences in scale can interfere with the accuracy of the analysis and interpretation
90 results. By using data that has been prepared in terms of characteristics and normalization, this
91 research can facilitate the analysis and interpretation process and produce more accurate results.
92 (Abo et al., 2018; Amalia et al., 2021; Yadav & Vishwakarma, 2020; Tseng et al., 2022; Ajiono
93 & Hariguna, 2023; Zeng & Nantida, 2023).

94 *Preprocessing*

95 The data in this study underwent several preprocessing steps to prepare them for further analysis.
96 The preprocessing phase included several procedures, including capitalization, punctuation and
97 conjunction removal, tokenization, stemming, and stop word elimination. Capitalization was used
98 to remove all distinctions between uppercase and lowercase letters irrelevant to the meaning of the
99 word. The next step was to remove punctuation and hyphens as they did not provide useful
100 information for text analysis. Tokenization was used to divide the text into tokens, which could
101 include words, phrases, or sentences with different meanings.

102 Stemming is a natural language processing technique used to extract the root or base form of a
103 word by removing any prefixes or suffixes attached to the word. The goal of stemming is to reduce
104 each word to a common base form, even if it is not a valid English word, to group words that have
105 the same meaning. By doing so, stemming can help to reduce the dimensionality of text data and
106 simplify analysis. Rather, converting a word to its base form is known as lemmatization (Del Arco
107 et al., 2022; Glimsdal & Granmo, 2021; Greco & Polli, 2020; Suhane & Kowshik, 2021; Jen and

108 Lin, 2021; Prayitno et al., 2021). Stemming, however, involves reducing the inflected form of a
109 word to its root. In addition, stop words, which are common English words such as "the," "a," "an,"
110 and "is," were removed during the preprocessing phase because they do not provide significant
111 meaning in text analysis. The purpose of these preprocessing steps was to prepare the data for
112 further analysis and facilitate data analysis.

113 Model Development

114 After the data is available, the next step is to select the algorithm to be incorporated into the model
115 (Budi & Yaniasih, 2022; Jafarian et al., 2021; Kaur et al., 2021). In general, machine learning
116 algorithms such as Naive Bayes, SVM, and neural networks are utilized in sentiment analysis.
117 However, depending on the purpose and characteristics of the data used, other algorithms such as
118 rule-based algorithms or hybrid algorithms may be employed.

119 The next step after choosing an algorithm is to implement it in the model being developed. This
120 typically involves converting the data into a format that is suitable for the selected algorithm and
121 adjusting any necessary parameters (Xu et al., 2019; Liu, 2023; Wang et al., 2023). After
122 implementing the algorithm, the model must be validated by comparing the results it produces
123 with the actual results (ground truth). This allows for the accuracy of the model to be determined.
124 Upon validating the model, it is necessary to optimize it in order to obtain improved results. The
125 optimization can be achieved through a variety of methods, such as altering the model's
126 hyperparameters, modifying the number of layers or neurons, or selecting more effective features.

Results

128 In this results section, the research results are presented in accordance with the research objectives,
129 specifically the comparison of the performance of the algorithms used and the results of the
130 developed model.

131 *Algorithm Comparation*

132 The accuracy of an algorithm refers to the degree of accuracy it achieves in classifying data. A
133 higher accuracy indicates a better performance of the algorithm in classifying data.

134 The BernoulliNB algorithm, one of the naive Bayes algorithms used for binary classification,
135 achieved an accuracy of 0.87, indicating that it is able to classify data with 87% accuracy. The
136 DecisionTreeClassifier algorithm, a decision tree algorithm used to classify data based on its
137 attributes, achieved an accuracy of 0.86, indicating that it is able to classify data with 86%
138 accuracy.

139 The KNeighborsClassifier algorithm, a k-nearest neighbor algorithm used to classify data based
140 on the nearest data in the neighborhood, achieved an accuracy of 0.87, indicating that it is able to
141 classify data with 87% accuracy. The LogisticRegression algorithm, a logistic regression algorithm

142   used to classify binary data, achieved an accuracy of 0.89, indicating that it is able to classify data
143   with 89% accuracy. The LinearSVC algorithm, a support vector machine algorithm that utilizes a
144   linear kernel to classify data, achieved an accuracy of 0.90, indicating that it is able to classify data
145   with 90% accuracy.

146   The BaggingClassifier algorithm, a bagging algorithm that improves the accuracy of the algorithm
147   by using multiple models simultaneously, achieved an accuracy of 0.89, indicating that it is able
148   to classify data with an accuracy of 89%. The StackingClassifier model achieved an accuracy of
149   0.90, indicating that it is able to classify data into the appropriate class with an accuracy of 90%.

150   The RandomForestClassifier model achieved an accuracy of 0.88, indicating that it is able to
151   classify data into the correct class with an accuracy of 88%. The AdaBoostClassifier model also
152   achieved an accuracy of 0.89, indicating that it is able to classify data into the correct class with
153   an accuracy of 89%. Finally, the ExtraTreesClassifier model achieved an accuracy of 0.88,
154   indicating that it is able to classify data into the correct class with an accuracy of 88%. The
155   performance results of each algorithm can be seen in Figure 2.

156   [Figure 2 here]

157   Figure 2. Algorithm rankings

158   Linear Support Vector Classification (LinearSVC) and Stacking Classifier are two algorithms that
159   have demonstrated high accuracy in the classification of sentiment in text data. In particular, both
160   algorithms have shown to be effective in the detection and analysis of the sentiment (positive or
161   negative) present in a given text. In comparison to other algorithms, LinearSVC and Stacking
162   Classifier have demonstrated superior performance in the task of sentiment analysis.

163   Linear Support Vector Classification (LinearSVC) utilizes the Support Vector Machine method to
164   predict the class of an input data point through the identification of the optimal hyperplane that
165   separates two or more classes. This algorithm is highly accurate and has been shown to be effective
166   for both binary (two classes) and multi-class (more than two classes) classification tasks. Its
167   performance has been extensively demonstrated in various studies and has established itself as a
168   reliable and effective method for sentiment analysis and other classification tasks.

169   StackingClassifier is an ensambling method that combines multiple machine learning models into
170   a more powerful model. StackingClassifier predicts the class of an instance using multiple machine
171   learning models as base models, and then combines the prediction results of the base models with
172   other machine learning models as metamodels. In this way, StackingClassifier can improve
173   classification accuracy by using multiple machine learning models, each of which has different
174   advantages and disadvantages.

175   A multiple-output classification accuracy of over 80% can be deemed satisfactory, especially when
176   multiple algorithms are utilized. However, it is important to note that this benchmark may vary
177   depending on the specific context being analyzed and should not be treated as a definitive metric
178   of an algorithm's superiority. Additionally, it is crucial to consider the benefits and drawbacks of
179   each algorithm in relation to the problem at hand, as no algorithm is superior in all scenarios.

180   ***Model Simulation and Evaluation***

181   In order to utilize the model, the input data must be entered in the form of sentiments or ratings for
182   analysis. The data is then processed through the PreProcessed Review process, which displays the
183   analyzed data. Subsequently, the Predictions By Classifiers component utilizes 10 algorithms from
184   the multi-output model to analyze the data. The results of this analysis are presented on a scale
185   from 0 to 10, with 0 representing negative sentiment and 10 representing positive sentiment.
186   Finally, the TextBlob Sentiment Polarity (UnPreprocessed) display shows the number of words
187   that are not recognized by the model. The visualization of this model usage process is illustrated
188   in Figure 3.

189   [Figure 3 here]

190   Figure 3. Model result towards neutral data

191   The average accuracy of the developed model is 88%, which is considered to be quite accurate as
192   long as the input rating or sentiment data can be recognized by the model. However, if the rating
193   or sentiment data is dominated by words that are unfamiliar to the model or have not been captured
194   in its training data, most of the model's algorithms may classify the data as negative.

195   The results of the model are not simply binary (positive or negative) but rather are presented as a
196   score from 1 to 10 in order to more accurately reflect the sentiment of the input data. This is
197   because a given sentence may not contain exclusively positive or negative words, and presenting
198   the results on a more nuanced scale allows for greater flexibility and validity in different analysis
199   contexts. Furthermore, using a scale rather than a binary positive/negative categorization allows
200   for a more granular and potentially more actionable measurement. Additional simulation results
201   of the model can be found in Figures 4 and 5 below.

202   [Figure 4 here]

203   Figure 4. Model result towards positive data

204   [Figure 5 here]

205   Figure 5. Model result towards negative data

206

Discussion

In recent years, sentiment analysis has become a widely researched topic owing to its potential application in various fields. In the current work, the effectiveness of LinearSVC and StackingClassifier algorithms in sentiment analysis was investigated, and their superior performance compared to other algorithms was demonstrated. This result is consistent with previous studies, such as the work of Aini et al. (2021), which showed that SVM-based algorithms, such as LinearSVC, are effective in sentiment classification. Furthermore, the study highlights the importance of considering the specific context being analyzed when evaluating the performance of sentiment analysis algorithms. This finding is also supported by Hananto et al. (2022), who noted that the accuracy of sentiment analysis algorithms can vary depending on the type of data being analyzed and the specific domain or application. Therefore, it is important to evaluate the performance of sentiment analysis algorithms in relevant contexts to ensure their effectiveness.

The use of ensemble methods, such as the StackingClassifier, has also been explored in sentiment analysis. In a study by Hariguna & Rachmawati (2019), it was demonstrated that ensemble methods can improve sentiment classification accuracy by combining predictions from multiple base models. This finding is consistent with the results of the current study, which showed that the StackingClassifier can improve classification accuracy by leveraging the strengths and mitigating the weaknesses of multiple machine learning models. The current study also used a multi-output classification model to classify sentiments in tweets. This approach is similar to the work of Hariguna et al. (2019), who used a multilabel classification method to classify sentiments in movie reviews. The use of multiple algorithms in the multi-output classification model used in the current study is innovative and represents a potential improvement over previous methods that used only a single algorithm.

Finally, the high accuracy of the multi-output classification model used in this study suggests its potential for real-world applications. This finding is consistent with the work of Hung et al. (2021), who demonstrated the usefulness of sentiment analysis in predicting stock market trends. The high accuracy of sentiment analysis algorithms, such as the multi-output classification model used in this study, can enable more accurate predictions and decision-making in various fields. In conclusion, the current study provides valuable insights into the effectiveness of sentiment analysis algorithms, particularly LinearSVC and StackingClassifier, in relevant contexts. The use of multioutput classification models and ensemble methods is also innovative and represents potential improvements over previous methods. The results of this study are consistent with previous research in this area and suggest the potential for real-world applications of sentiment analysis. Based on the obtained research results, the followings can be pointed out:

First, the results of this study demonstrated that LinearSVC and StackingClassifier are two algorithms that can be effectively utilized in sentiment analysis, or the process of identifying and analyzing the sentiments (positive or negative) expressed in a text. Both algorithms exhibited higher accuracy in classifying sentiments compared to other algorithms. LinearSVC utilizes the

245   support vector machine (SVM) method to predict the class of an instance by finding the optimal
246   hyperplane to differentiate between two classes. This algorithm is recognized for its high accuracy
247   and strong performance in both binary and multi-class classification tasks.

248   In contrast, StackingClassifier is an ensemble method that combines multiple machine learning
249   models to form a more robust model. It predicts the class of an instance by using multiple machine
250   learning models as base models and then aggregating the prediction results of these base models
251   with the use of other machine learning models as meta-models. As a result, StackingClassifier can
252   improve classification accuracy by leveraging the strengths and mitigating the weaknesses of
253   multiple machine learning models. A classification accuracy of over 80% for tasks with multiple
254   output classes can be deemed satisfactory, especially when utilizing multiple algorithms. However,
255   it is important to note that this benchmark may vary depending on the specific context being
256   analyzed and should not be treated as a definitive measure of an algorithm's superiority.
257   Additionally, it should be remembered that no algorithm is superior in all scenarios, so it is crucial
258   to consider the benefits and drawbacks of each algorithm in relation to the problem at hand.

259   Secondly, the multi-output classification model used in this study demonstrated superiority due to
260   its ability to classify data into multiple output classes. This is particularly useful for tasks that
261   involve classifying data with multiple output variables, as was the case in this study, which utilized
262   10 algorithms to classify sentiment in tweets. Additionally, the multi-output classification model
263   employed in this study represents a novel approach as it is the first multi-output sentiment
264   classification model to utilize 10 algorithms. This is likely to be highly beneficial in improving the
265   accuracy of the model, as the use of multiple algorithms provides a greater amount of information
266   that can be utilized in the classification process.

267   Furthermore, the accuracy of the multiple algorithm classification model used in this study was
268   found to be quite high at 88%. This indicates that the model is capable of accurately classifying
269   data, making it reliable for identifying the sentiment of tweets. It can therefore be concluded that
270   the multi-output classification model utilized in this study is a superior and innovative approach
271   for classifying data, particularly in the context of sentiment research.

272

Conclusions

274   This study successfully developed a sentiment classification model that combines 10 algorithms,
275   resulting in an average accuracy of 88%. This approach demonstrates the potential of combining
276   multiple classification algorithms to improve sentiment analysis accuracy. However, while the
277   accuracy achieved in this study represents a significant increase compared with previous studies
278   (Han et al., 2020; Miao et al., 2021; Salazar et al., 2022; Sun et al., 2019), there are still limitations
279   to consider. One limitation is the potential bias in the data used to train the model because
280   sentiment analysis is highly dependent on the context in which the text is produced. Therefore,
281   future studies should focus on creating more diverse and representative training datasets to

282 improve the accuracy of the model. Additionally, while LinearSVC and StackingClassifier showed
283 high accuracy in sentiment classification, it is important to note that no algorithm was superior in
284 all scenarios. Therefore, future research should focus on evaluating the advantages and
285 disadvantages of other algorithms and how they can be combined to improve the accuracy.

286 Furthermore, the use of multiple algorithms for sentiment classification requires significant
287 computational resources and time. Therefore, future research should focus on optimizing the
288 model to reduce computational cost while maintaining accuracy. Overall, although the sentiment
289 classification model developed in this study demonstrated significant improvements in accuracy
290 compared to previous studies, there are still limitations and opportunities for future work. By
291 addressing these limitations and exploring alternative approaches, the accuracy of sentiment
292 classification models can be further improved, ultimately leading to more accurate and reliable
293 sentiment analysis in various contexts.

294 Based on the conclusions of this study, there are several future perspectives and suggestions. First,
295 although the developed sentiment classification model has been shown to be accurate, further
296 studies should be conducted to improve its performance. This can be achieved by exploring the
297 potential of combining additional classification algorithms to improve the accuracy of the model.
298 Additionally, the model can be tested on a larger dataset to evaluate its scalability and robustness.
299 Second, future studies can explore the use of deep learning techniques such as convolutional neural
300 networks (CNNs) and recurrent neural networks (RNNs) for sentiment analysis. These techniques
301 have shown promising results in natural language processing tasks and may provide a more
302 accurate and efficient approach to sentiment analysis.

303 Third, future studies could investigate the applicability of the developed sentiment classification
304 model to other domains, such as product reviews, movie reviews, and news articles. This can
305 provide insight into the generalizability of the model and its potential use in other contexts. Finally,
306 future studies can explore the ethical implications of sentiment analysis and the use of machine
307 learning algorithms for decision-making. This can include examining the issues of bias, fairness,
308 and accountability in the development and use of sentiment analysis models. Finally, several
309 avenues for future research and suggestions can be made based on the findings of this study. These
310 include improving the accuracy of the developed sentiment classification model, exploring the use
311 of deep-learning techniques, evaluating the generalizability of the model, and examining the
312 ethical implications of sentiment analysis.

313

References

315 Abo, M. E. M., Shah, N. A. K., Balakrishnan, V., & Abdelaziz, A. (2018). Sentiment analysis
316   algorithms: evaluation performance of the Arabic and English language. *2018 International*
317   *Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE)*, 1–
318   5.

319   Aini, Q., Hammad, J., Taher, T., & Ikhlayel, M. (2021). Classification of Tweets Causing
320   Deadlocks in Jakarta Streets with the Help of Algorithm C4.5. *Journal of Applied Data*
321   *Sciences*, *2*(4), 143–156.
322   Ajiono, A., & Hariguna, T. (2023). Comparison of Three Time Series Forecasting Methods on
323   Linear Regression , Exponential Smoothing and Weighted Moving Average. *IJIIS:*
324   *International Journal of Informatics and Information Systems*, *6*(2), 89–102.
325   Alfazzi, F. (2022). A Knowledge Behavioral and Intelligence Management in Fostering
326   Entrepreneurship for Modern Industries. *International Journal for Applied Information*
327   *Management*, *2*(4), 95–105.
328   Amalia, A., Gunawan, D., & Nasution, K. (2021). Sentiment analysis of GO-JEK services quality
329   using Multi-Label Classification. *Journal of Physics: Conference Series*, *1830*(1), 12003.
330   Budi, I., & Yaniasih, Y. (2022). Understanding the meanings of citations using sentiment, role,
331   and citation function classifications. *Scientometrics*, 1–25.
332   Cao, Y., Chen, H., & Liu, Q. (2023). Utilizing Internet Analysis Technology to Conduct Data
333   Analysis of Student Attitude Survey. *IJIIS: International Journal of Informatics and*
334   *Information Systems*, *6*(2), 47–55.
335   Del Arco, F. M., Collado-Montañez, J., Ureña, L. A., & Martín-Valdivia, M.-T. (2022). Empathy
336   and Distress Prediction using Transformer Multi-output Regression and Emotion Analysis
337   with an Ensemble of Supervised and Zero-Shot Learning Models. *Proceedings of the 12th*
338   *Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis*,
339   239–244.
340   Glimsdal, S., & Granmo, O.-C. (2021). Coalesced Multi-Output Tsetlin Machines with Clause
341   Sharing. *ArXiv Preprint ArXiv:2108.07594*.
342   Greco, F., & Polli, A. (2020). Emotional Text Mining: Customer profiling in brand management.
343   *International Journal of Information Management*, *51*(December 2018), 1–8.
344   Han, Y., Liu, Y., & Jin, Z. (2020). Sentiment analysis via semi-supervised learning: a model based
345   on dynamic threshold and multi-classifiers. *Neural Computing and Applications*, *32*(9),
346   5117–5129.
347   Hananto, A., Rahayu, S., & Hariguna, T. (2022). COVID-19 Vaccination: A Retrospective
348   Observation and Sentiment Analysis of the Twitter Social Media Platform in Indonesia. *IJIIS:*
349   *International Journal of Informatics and Information Systems*, *5*(1), 56–69.
350   Hariguna, T., & Rachmawati, V. (2019). Community Opinion Sentiment Analysis on Social Media
351   Using Naive Bayes Algorithm Methods. *IJIIS: International Journal of Informatics and*
352   *Information Systems*, *2*(1), 33–38.
353   Hung, C.-W. (2021). Application of Quality Function Development Method to Establish
354   Application of New Product Development Information System. *International Journal for*
355   *Applied Information Management*, *1*(1 SE-Articles), 23–27.
356   Hayadi, B. H., Widawati, E., Bachtiar, M., & Tambunan, N. (2023). Certainty Factor Method
357   Analysis for Identification of Covid-19 Virus Accuracy. *IJIIS: International Journal of*
358   *Informatics and Information Systems*, *6*(1), 38–46.
359   Jafarian, H., Taghavi, A. H., Javaheri, A., & Rawassizadeh, R. (2021). Exploiting BERT to
360   improve aspect-based sentiment analysis performance on Persian language. *2021 7th*
361   *International Conference on Web Research (ICWR)*, 5–8.
362   Jen, L., & Lin, Y.-H. (2021). A Brief Overview of the Accuracy of Classification Algorithms for
363   Data Prediction in Machine Learning Applications. *Journal of Applied Data Sciences*, *2*(3),
364   84–92.

365 Kabir, M., Kabir, M. M. J., Xu, S., & Badhon, B. (2021). An empirical research on sentiment
366   analysis using machine learning approaches. *International Journal of Computers and*
367   *Applications*, *43*(10), 1011–1019.
368 Kaur, H., Ahsaan, S. U., Alankar, B., & Chang, V. (2021). A proposed sentiment analysis deep
369   learning algorithm for analyzing COVID-19 tweets. *Information Systems Frontiers*, *23*(6),
370   1417–1429.
371 Liu, Y. (2023). Research on Deep Learning-Based Algorithm and Model for Personalized
372   Recommendation of Resources. *Journal of Applied Data Sciences*, *4*(2), 68–75.
373 Miao, H., Zhang, Y., Wang, D., & Feng, S. (2021). Multi-Output Learning Based on Multimodal
374   GCN and Co-Attention for Image Aesthetics and Emotion Analysis. *Mathematics*, *9*(12),
375   1437.
376 Mukhtar, N., Khan, M. A., & Chiragh, N. (2018). Lexicon-based approach outperforms Supervised
377   Machine Learning approach for Urdu Sentiment Analysis in multiple domains. *Telematics*
378   *and Informatics*, *35*(8), 2173–2183.
379 Prayitno, J., Saputra, B., & Waluyo, R. (2021). Data Mining Implementation with Algorithm C4.5
380   for Predicting Graduation Rate College student. *Journal of Applied Data Sciences*, *2*(3), 74–
381   83.
382 Salazar, C., Montoya-Múnera, E., & Aguilar, J. (2022). Sentiment analysis in learning resources.
383   *Journal of Computers in Education*, 1–26.
384 Saputra, J. P. B., Rahayu, S. A., & Hariguna, T. (2023). Market Basket Analysis Using FP-Growth
385   Algorithm to Design Marketing Strategy by Determining Consumer Purchasing Patterns.
386   *Journal of Applied Data Sciences*, *4*(1), 38–49. https://doi.org/10.47738/jads.v4i1.83
387 Suhane, A., & Kowshik, S. (2021). Multi output learning using task wise attention for predicting
388   binary properties of tweets: Shared-task-on-fighting the COVID-19 infodemic. *Proceedings*
389   *of the Fourth Workshop on NLP for Internet Freedom: Censorship, Disinformation, and*
390   *Propaganda*, 110–114.
391 Sun, Z., Sarma, P. K., Sethares, W., & Bucy, E. P. (2019). Multi-modal sentiment analysis using
392   deep canonical correlation analysis. *ArXiv Preprint ArXiv:1907.08696*.
393 Tseng, S., Liang, C., & Tsai, H. (2022). A Study on the Relationships Among Personality Traits ,
394   Gender and Customer Knowledge Preferences. *International Journal for Applied Information*
395   *Management*, *2*(3), 1–17.
396 Wang, F., Sungho, P., & Cattareeya, S. (2023). Software Defect Fault Intelligent Location and
397   Identification Method Based on Data Mining. *Journal of Applied Data Sciences*, *4*(2), 84–92.
398 Wu, J., Lu, K., Su, S., & Wang, S. (2019). Chinese micro-blog sentiment analysis based on
399   multiple sentiment dictionaries and semantic rule sets. *IEEE Access*, *7*, 183924–183939.
400 Xu, D., Shi, Y., Tsang, I. W., Ong, Y.-S., Gong, C., & Shen, X. (2019). Survey on multi-output
401   learning. *IEEE Transactions on Neural Networks and Learning Systems*, *31*(7), 2409–2429.
402 Yadav, A., & Vishwakarma, D. K. (2020). A comparative study on bio-inspired algorithms for
403   sentiment analysis. *Cluster Computing*, *23*(4), 2969–2989.
404 Zeng, J., & Nantida, S. (2023). Study on the Ideological and Political Practice Teaching of College
405   Students Based on the Internet + Technology. *IJIIS: International Journal of Informatics and*
406   *Information Systems*, *6*(1), 24–30. https://doi.org/10.47738/ijiis.v6i1.148
407

Figure 1

Figure_1._Research_steps
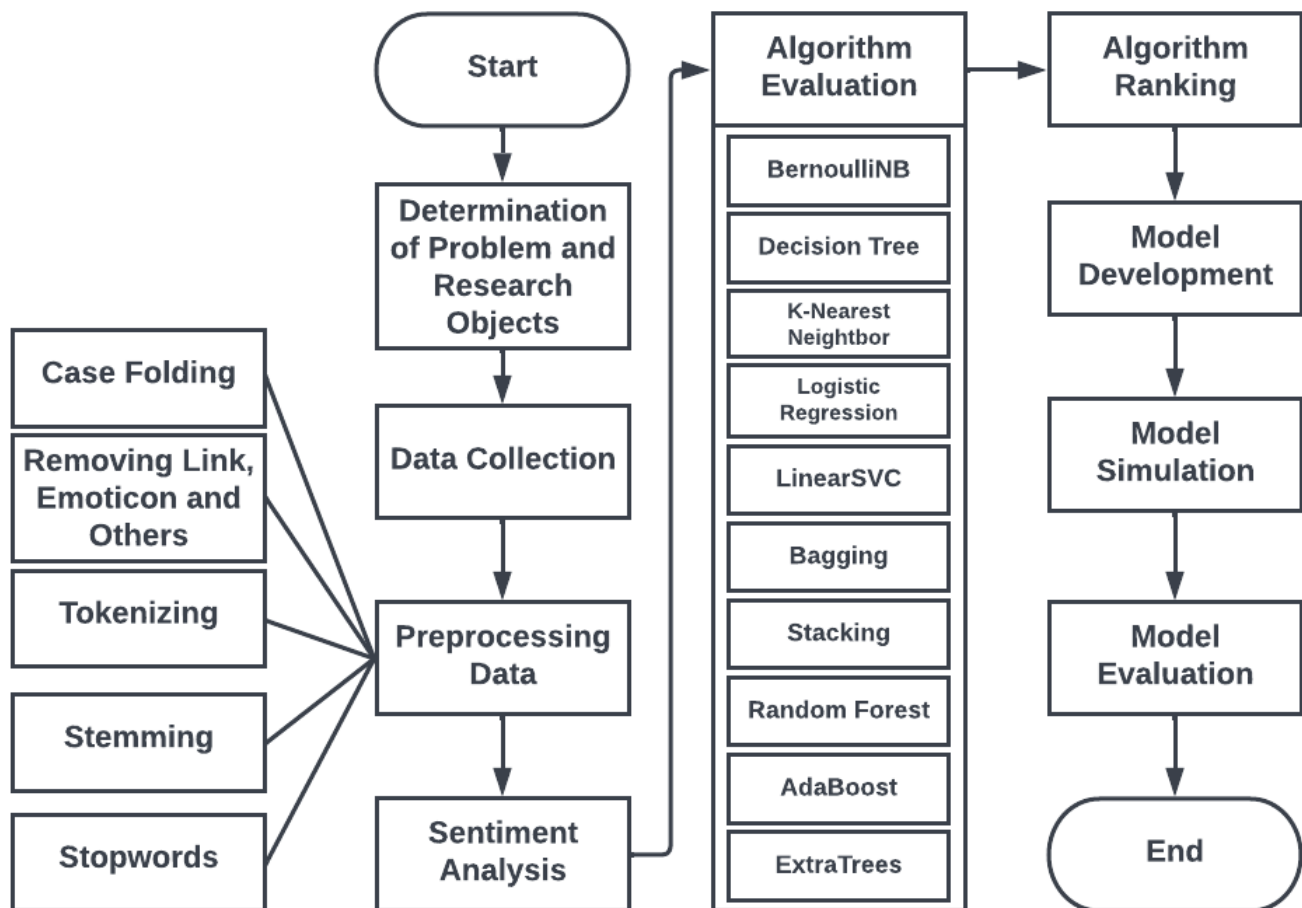
Figure 2

Figure_2._Algorithm_rankings

Model	Accuracy	Precision	Recall	ROC Score
BernoulliNB	0.87	0.72	0.68	0.53
DecisionTreeClassifier	0.86	0.88	0.70	0.60
KNeighborsClassifier	0.87	0.64	0.71	0.61
LogisticRegression	0.89	0.70	0.61	0.60
LinearSVC	0.90	0.72	0.88	0.63
BaggingClassifier	0.89	0.65	0.76	0.62
StackingClassifier	0.90	0.68	0.77	0.69
RandomForestClassifier	0.88	1.00	0.83	0.52
AdaBoostClassifier	0.89	0.57	0.45	0.70
ExtraTreesClassifier	0.88	0.66	0.76	0.50

Figure 3

Figure_3._Model_result_towards_neutral_data

```
In [933]:  #Simulasi
           custom_review = "lumayan sih transaksi di triv tapi kyk ga ada yang pake ini website yah"
           custom_review

Out[933]:  'lumayan sih transaksi di triv tapi kyk ga ada yang pake ini website yah'

In [934]:  generate_rating(ml_models_list,custom_review)

           PreProcessed Review:
           ['lumayan sih transaksi di triv tapi kyk ga ada yang pake ini website yah']

           Predictions By Classifiers:

           BernoulliNB : [0]
           DecisionTreeClassifier : [1]
           KNeighborsClassifier : [1]
           LogisticRegression : [1]
           LinearSVC : [0]
           BaggingClassifier : [0]
           StackingClassifier : [1}
           RandomForestClassifier : [1]
           AdaBoostClassifier : [0]
           ExtraTreesClassifier : [1]

           Dampak Komentar (0: Negative | 10: Positive): 5

           TextBlob Sentiment Polarity(UnPreprocessed): 0.5
```

Figure 4

Figure_4._Model_result_towards_positive_data

```
In [922]:  custom_review = "Keren nih transaksi di indodax, bagus banget pelayanannya apalagi kalo ada member banyak deh promonya"
           custom_review

Out[922]:  'Keren nih transaksi di indodax, bagus banget pelayanannya apalagi kalo ada member banyak deh promonya'

In [923]:  generate_rating(ml_models_list,custom_review)

           PreProcessed Review:
           ['Keren nih transaksi di indodax, bagus banget pelayanannya apalagi kalo ada member banyak deh promonya']

           Predictions By Classifiers:

           BernoulliNB : [1]
           DecisionTreeClassifier : [1]
           KNeighborsClassifier : [0]
           LogisticRegression : [1]
           LinearSVC : [1]
           BaggingClassifier : [1]
           StackingClassifier : [1]
           RandomForestClassifier : [1]
           AdaBoostClassifier : [1]
           ExtraTreesClassifier : [1]

           Dampak Komentar (0: Negative | 10: Positive): 9

           TextBlob Sentiment Polarity(UnPreprocessed): 0.0
```

Figure 5

Figure_5._Model_result_towards_negative_data

```
In [931]: custom_review = "Pengalaman terburuk yang saya dapat adalah ketika proses pengiriman dengan binance sangat mahal dan lama"
          custom_review

Out[931]: 'Pengalaman terburuk yang saya dapat adalah ketika proses pengiriman dengan binance sangat mahal dan lama'

In [932]: generate_rating(ml_models_list,custom_review)

          PreProcessed Review:
          ['Pengalaman terburuk yang saya dapat adalah ketika proses pengiriman dengan binance sangat mahal dan lama']

          Predictions By Classifiers:

          BernoulliNB : [0]
          DecisionTreeClassifier : [0]
          KNeighborsClassifier : [0]
          LogisticRegression : [0]
          LinearSVC : [0]
          BaggingClassifier : [0]
          StackingClassifier : [0]
          RandomForestClassifier : [0]
          AdaBoostClassifier : [0]
          ExtraTreesClassifier : [0]

          Dampak Komentar (0: Negative | 10: Positive): 0

          TextBlob Sentiment Polarity(UnPreprocessed): 0.0
```