Journal of Digital Information Management

DLINE JOURNALS

ISSN: 0972-7272

JDIM 2026: 24 (3)

A Comparative Study of Multi-class Classification Models and Data Pre-Processing Methods for Analyzing Malaysian University Students' Views on Ai Chatbot

Ali Sameer Al-Abeej
Department of Physiology
College of Medicine, University of Basrah
Basrah, Iraq
ali.sameer@uobasrah.edu.iq

ABSTRACT

Social media contains opinions, ideas, and facts. Artificial intelligence (AI) has brought a mix of societal perspectives to social media events. This study used bilingual tweets (English and Malay). This study uses Malay stop words and number filters in data preprocessing, builds a machine learning classifier using mBERT, support vector machine, and neural network models, and measures performance using recall, precision, F-measure, and accuracy for each model. In the experimental results, Malay stop words and number filters caused variability in classifier performance. Malay stop words and number filters are important in preprocessing student sentiment analysis in this study, although they do not significantly affect the change in accuracy. This paper also discusses the KNIME workflow and the experimental results obtained. This study argues that the neural network is the best classifier, and the model's algorithm can be used to perform sentiment analysis on a new dataset.

Subject Categories and Descriptors: [**K.4.2 Social Issues]** [H.5.2 User Interfaces]: Natural language; [**H.1.2 User/Machine Systems**]; Human information processing: **[I.2.6 Learning]**

General Terms: Artificial Intelligence, Natural Language Processing, Human Information Processing,

Keywords: Sentiment Analysis, Artificial Intelligence, Machine Learning, Chatbot, Malay Stop Words, Number Filter, MBERT, Support Vector Machine, Neural Network

Received: 3 May 2026, Revised: 20 June 2026, Accepted: 23 July 2026

Review Metrics: Review Scale: 0-6, Review Score: 4.52, Inter-Reviewer Consistency: 84.5%

DOI: https://doi.org/10.6025/jdim/2026/24/3/149-157

1. Introduction

Nowadays, social media platforms have gained importance for readers because they allow people to share and express their opinions on certain topics and post messages simply across the world. Due to the rapid growth of social media, accurately analyzing the sentiment and classifying data is a challenging task. Most available data

is text based. Sentiment analysis uses natural language modelling, text interpretation, computational linguistics, and biometrics to define, research, isolate, and consistently perceive affective states.

Sentiment research is commonly applied to client product speech, with uses ranging from marketing and customer service to professional practice, such as feedback and audience results, internet and social media, and healthcare materials. Researchers often use sentiment analysis of social networks because it makes it easier to understand people's attitudes and behaviour.

In comparison, unlike previous sentiment analysis that targeted product reviews, users are free to write whatever they want [20].

So, the research focuses on approaches that can detect and mine sentiment opinions within large amounts of data. Sentiment analysis approaches are based on machine learning. Some testing strategies recommend combining lexicon based and machine learning approaches to optimise sentiment classification efficiency. Machine learning collects useful information and produces detailed outcomes from raw data [23].

This data helps us address dynamic, data rich issues. Machine learning algorithms learn from data and process it, and they can discover insights without explicit programming [3].

Machine learning also uses massive data from unlimited sources for processing and prediction, spam detection, increased security, and performance, etc. According to [22], machine learning approaches achieve high accuracy and recall. The data set space is defined in a hierarchical structure in which the data is partitioned by a condition on the attribute value, and this is a form called mBERT. [15]. stated that the advantages of a simple-to extract layout law, slightly lower measured size, ability to present substantial decision properties, and higher classification accuracy rate are commonly used in mBERT [10].

A group of supervised learning methods used for classification, regression, and outlier detection is support vector machines (SVM) [21]. claimed that there are layers of interconnected nodes in a neural network. Each node is a supervised learning algorithm equivalent to linear regression with multiple variables, and it feeds the generated signal into an activation function that may be nonlinear, such as multiple linear regression.

Words that are extracted before or during natural language data collection. It is a pre-processing technique. Although stop words typically refer to the most popular words or phrases in a language, not all natural language processing tools use a single universal list of stop words, and not all tools even use such a list [14].

To improve research formulation, some tools explicitly avoid deleting these stop words [13]. Another pre-processing technique is number filtering, which filters numbers from the data. For filtering purposes, stop words are considered meaningless because they appear often in the language for which the indexing engine has been tuned. These terms are dropped at indexing time to save space and time, then skipped at search time. The study used both stop word and number filter techniques in pre processing [24].

At the personal and group level, social media has the potential to reveal useful insights into human emotions. Social media surveillance may be useful, as both the situation and individuals' responses change every moment during this volatile period [18]. Studying social media data could also play a key role in identifying students' actions and reactions to AI [17]. The data features must suit the next step in sentiment analysis and, in turn, help us understand people's insights on a given topic [9].

Healthcare sentiment analysis helps us better understand how individuals speak and feel about particular health issues or conditions. Sentiment analysis has the advantage of covering a larger population and providing immediate results [8].

According to Ahmad, Aftab, & Ali [2], using SVM (Support Vector Machine) is ideal for textual polarity detection for sentiment analysis using Weka. However, several other machine learning techniques can be used, such as

Naïve Bayes (NB), Random Forest (RF), mBERT, and many more (Soumeur, Mokdadi, Guessoum, & Daoud [26]). Researchers must choose a suitable classifier for their dataset features.

Therefore, this study aims to understand Malaysian students' insights and opinions about artificial intelligence in the country using sentiment analysis on social media [29].To measure performance accuracy, this study will use precision, recall, and F1, which are the best performance measures [28].

The main purpose of this study is to identify features suitable for sentiment analysis using the Facebook dataset. In addition, the study aims to build an ML classification model for sentiment analysis and evaluate its performance using precision, recall, and F1.

2. Related Work

One important reason is that social media has played a prominent role during the spread of AI among students and the general public, as these platforms, through responses, express real time public concerns about the spread of AI applications [1]. Social media often serves as a communication tool. Additionally, social media platforms can provide a wealth of information to raise community awareness and educate people about AI [11]. This study aims to analyse Malaysian students' topics and sentiments to better understand their opinions, ideas, and beliefs.

Social media has become a primary and preferred platform for public opinions and perceptions about various events and policies [19]. As a result, governments, businesses, and organizations have turned to social media as a vital communication tool to disseminate valuable and essential information to the public [5]. Table 1 illustrates the difference between previous studies and the author's research.

Author	Technique	Work on
Nahar, Jaradat, Atoum, and Ibrahim [17]	lexicon based approach	Facebook comments and spotting the polarity in the telecommunications sector
Jain and Dandannavar [13]	Machine learning algorithms (decision trees and naïve Bayes)	Proposed a Text analysis framework for Twitter data using Apache Spark and hence is more flexible, fast, and scalable
Fadel and Cemil [7]	Machine learning approaches and lexicon approaches	propose a model for automatically classifying users' reviews on Twitter. The performance of classification algorithms was measured using accuracy and F1 scores
Ersahin, Aktas, Kilinc, and Ersahin [6]	Lexicon based and machine learning (ML)-based approaches	Sentiment analysis in Turkish and test it with three different datasets
Iqbal, Hashmi, Fung, Batool, Khattak, Aleem, and Hung [12]	ML algorithms and lexical databases	Sentiment analysis employs state of the art to archives of online documents.
Al-Saffar, Awang, Tao, Omar, Al-Saiagh, and Al-Bared[4]	Lexicon method and machine learning approaches	classification performances based on the semantic orientate

Table 1. The Difference between the Previous Studies and the Author's Research

According to the researchers, sentiment analysis on social media is highly effective using machine learning methods [30]. However, SVM-based machine learning is underused as a classifier [27], which could help identify Malaysian students' opinions and insights on social media platforms [16]. This study used social media data to capture students' opinions about artificial intelligence. Sentiment classification categorized these sentiments as positive, negative, or neutral, and evaluated performance using precision, recall, F1 score, and accuracy [25]. The result of the classification was 97.8 for Support Vector Machine (SVM), 96.8 for K-Nearest Neighbor (K-NN), and 96.8 for Support Vector Machine (SVM).

3. Machine Learning Algorithm Design

Figure 1 shows the methodology used in this study. We used a machine learning approach due to its adaptability and accuracy. It comprises six stages: Data Collection, Data Preprocessing, Feature Extraction, Training Data, Classification, and Results. After collecting the data, we performed preprocessing. This task eliminated numbers, punctuation, and stop words. The data were then ready for feature extraction. Next, we extracted features to select terms for training and classification. For the classification task, we used mBERT, Support Vector Machine (SVM), and Neural Network (NN). We used DT and SVM because many related works stated that these two classifiers were suitable for sentiment analysis classification. We also used NN as a classifier because it is popular today. The last task was evaluating the results. We used recall, precision and F-measure as the performance metrics.

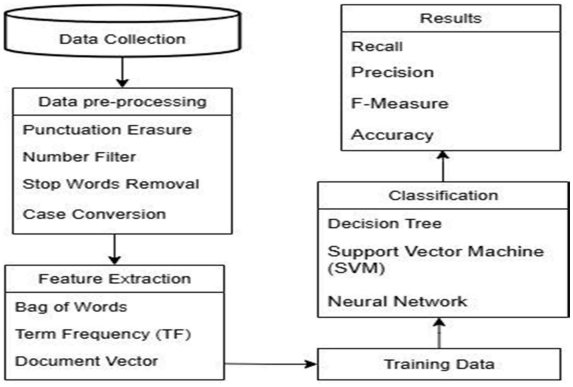

Figure 1. Methodology applied in this study

4. Experimental Design and Analysis

A. Data Collection

In this study, we conducted an experiment to gather Malaysian students' opinions on artificial intelligence. We then used machine learning technology to determine the polarity of the comments. After successfully identifying the polarity of the comments, we obtained 759 positive comments, 172 negative comments, and 83 neutral comments.

B. Data Pre-Processing

In this study, we conducted the research using Python. In this stage, we pre-processed the data through four activities: removing punctuation, numbers, and stop words, and converting text to lowercase. We used "remove punctuation" to remove all punctuation marks, and "remove numbers" to remove all numbers from comments. We also used both English and Malay stop words because the data obtained was bilingual.

C. Features Extraction

To represent the textual data numerically, we considered two feature extraction approaches: Bag-of-Words (BoW) and Term Frequency Inverse Document Frequency (TF-IDF). The BoW representation represents each comment according to the terms occurring in the document, while TF-IDF assigns a numerical weight to each term based on its frequency within an individual comment and its frequency across the complete corpus.

The TF-IDF weight of term (t) in document(d) is calculated as:

$$TFIDF\,(t, d) = TF\,(t, d)\ \text{x}\ IDF\,(t)$$

Where

$$IDF\,(t) = \log\left(\frac{N}{DF\,(t)}\right)$$

- t = term\word
- d =document\ comment
- N = total number of documents
- DF(t) =number of documents containing the term

TF(t, d) represents the term frequency of term (t) in document (d), and IDF(t) represents the inverse document frequency of the term across the corpus. In the experiments, we generated TF-IDF using the *TfidfVectorizer* implementation.

The vectorizer was fitted only on the training data, while the same transformation was subsequently applied to the test data to avoid data leakage. The resulting TF-IDF feature vectors were then provided as input to the mBERT, Support Vector Machine, and Multilayer Perceptron classifiers.

TF-IDF provides an additional feature representation for comparison with the original BoW based approach and allows evaluation of the effect of weighting informative terms in the sentiment classification task.

D. Training Data

The dataset consisted of three sentiment classes: positive, negative, and neutral. The class distribution was imbalanced, with 759 positive comments, 172 negative comments, and 83 neutral comments. To reduce potential bias toward the majority class, we applied a stratified train test split to preserve the relative proportion of each sentiment class in both the training and testing sets.

The dataset was divided into training and testing subsets using a 75:25 ratio. In addition, class weighting was incorporated during model training to assign greater importance to the minority classes. The class weights were determined based on each class's frequency in the training data, assigning higher weights to the negative and neutral classes than to the positive class.

The class weighting strategy was applied only during model training and did not alter the test set's original distribution. This approach was used to ensure that model evaluation reflected the original dataset while reducing the tendency of the classifiers to favor the majority positive class.

E. Classification

This study used three machine learning classifiers: mBERT, Support Vector Machine (SVM), and Multilayer Perceptron (MLP). To address class imbalance in the dataset, we incorporated class weighting into the training process. The weighting strategy penalizes classification errors on minority classes more strongly than errors on the majority class. For the mBERT and Support Vector Machine models, class weights were assigned according to the class distribution in the training dataset.

For the Multilayer Perceptron model, we configured the training procedure to account for the unequal class distribution. We calculated the class weights from the training data rather than the full dataset to avoid

information leakage. Table 2 presents the class weights used to address the class imbalance in the dataset. We trained the models using the selected feature representations, including the original feature representation and the TF-IDF representation. We then evaluated model performance on the untouched test set.

Class	Samples	Percentage	Class weight
Positive	759	74.85%	0.445
Negative	172	16.96%	1.965
Neutral	83	8.19%	4.072

Table 2. Class weights used to address class imbalance

5. Evaluation and Results

In this section, we present the classifier results and our analysis. Using the methodology and software described in the previous section, we obtained the training sets for the three classifiers and measured each classifier's performance accuracy based on the preprocessing activities tested.

Model	Stop Words	Number Weight	Class	Accuracy (%)	Macro-F1 (%)
SVM	Retained	Retained	No	86.42	72.85
SVM	Removed	Removed	No	87.31	75.64
SVM	Removed	Removed	Yes	88.17	81.26
mBERT	Retained	Retained	No	89.74	77.91
mBERT	Removed	Removed	No	89.21	79.36
mBERT	Removed	Removed	Yes	88.92	83.48
NN	Retained	Retained	No	84.65	69.74
NN	Removed	Removed	No	85.83	72.91
NN	Removed	Removed	Yes	86.76	78.43

Table 3. Effect of Text Preprocessing and Class Weighting

Table 1 presents the effect of text preprocessing and class weighting on the performance of the three evaluated models: SVM, mBERT, and ML. The results show that removing stop words and numerical tokens generally improves the classification performance compared with retaining them, particularly for SVM and ML models. Furthermore, class weighting produces a more noticeable improvement in Macro-F1 than in accuracy. For SVM, the Macro-F1 score increases from 72.85% in the baseline configuration to 81.26% after applying preprocessing and class weighting. mBERT improves from 77.91% to 83.48% Macro-F1, while the ML model improves from 69.74% to 78.43%.

Although applying class weighting slightly decreases mBERT's overall accuracy from 89.74% to 88.92%, it substantially improves the balance of classification performance across sentiment classes. This improvement matters given the severe class imbalance in the dataset, where the Positive class represents the majority of samples, while the Neutral and Negative classes are considerably smaller. However, Macro-F1 provides a more informative measure of model performance than accuracy alone. Overall, the results indicate that combining appropriate text preprocessing with class weighting can reduce the model's bias toward the majority class and improve its ability to recognize minority sentiment classes.

Model	Class	Precision (%)	Recall (%)	F1-Score (%)
SVM + CW	Positive	92.10	91.80	91.95
	Negative	78.60	76.40	77.49
	Neutral	73.80	70.50	72.11
mBERT + CW	Positive	94.21	94.87	94.54
	Negative	80.36	78.49	79.41
	Neutral	79.71	73.87	76.67
NN + CW	Positive	90.83	90.15	90.49
	Negative	75.40	73.20	74.29
	Neutral	69.85	66.90	68.34

Table 4. Class-wise Performance of the Proposed Weighted Models

Table 3 presents the class wise precision, recall, and F1-score achieved by the weighted SVM, mBERT, and ML models. The results show that the Positive class achieves the highest performance across all three models, consistent with its dominance in the dataset. Class weighting improves the models' ability to identify the minority Negative and Neutral classes.

Among the evaluated models, mBERT with class weighting achieves the best overall class wise performance. It obtains an F1-score of 94.54% for the Positive class, 79.41% for the Negative class, and 76.67% for the Neutral class. In the comparison, the weighted SVM achieves F1-scores of 91.95%, 77.49%, and 72.11% for the Positive, Negative, and Neutral classes, respectively. where The ML model records comparatively lower performance, with F1-scores of 90.49%, 74.29%, and 68.34%, respectively.

Table 4 shows that our trained model can perform sentiment analysis on the new dataset.

Sentiment	Number
Positive	135
Negative	13
Neutral	7
Total	155

Table 4. Sentiment Analysis New Dataset Results

6. Conclusion

This study examines the sentiment analysis of Malaysian university students toward AI in social media, focusing on the Malay and English languages, using a machine learning approach based on DT, SVM, and NN. This study also aims to identify the most important preprocessing features using Malay stop words and a number filter. We use recall, precision, F-measure, and accuracy as classifier performance indicators. We use the best training model from the study to integrate new data, achieving strong results. The results of this study help people learn more about students' use of AI and their reactions to it. We also used the best classifier model with the best preprocessing features for sentiment analysis.

References

[1] Agarwal, A., Xie, B., Vovsha, I., Rambow, O., Passonneau, R. J. (2011, June). Sentiment analysis of Twitter data. In Proceedings of the Workshop on Language in Social Media (LSM 2011) (p. 30–38). Association for Computational Linguistics.

[2] Ahmad, M., Aftab, S., Ali, I. (2017). Sentiment analysis of tweets using SVM. International Journal of Computer Applications, 177(5), 25–29. https://doi.org/10.5120/ijca2017915758.

[3] Al Sallab, A., Hajj, H., Badaro, G., Baly, R., El-Hajj, W., Shaban, K. (2015, July). Deep learning models for sentiment analysis in Arabic. In Proceedings of the Second Workshop on Arabic Natural Language Processing (p. 9–17). *Association for Computational Linguistics*.

[4] Al-Saffar, A., Awang, S., Tao, H., Omar, N., Al-Saiagh, W., Albared, M. (2018). Malay sentiment analysis based on combined classification approaches and Senti-lexicon algorithm. PLoS ONE, 13(4), e0194852. https://doi.org/10.1371/journal.pone.0194852.

[5] Er, M. J., Liu, F., Wang, N., Zhang, Y., Pratama, M. (2016, July). User-level Twitter sentiment analysis with a hybrid approach. In International Symposium on Neural Networks (pp. 426–433). *Springer*.

[6] Ersahin, B., Aktas, Ö., Kilinc, D., Ersahin, M. (2019). A hybrid sentiment analysis method for Turkish. *Turkish Journal of Electrical Engineering Computer Sciences*, 27(3), 1780–1793.

[7] Fadel, I., Cemil, Ö. Z. (2020). A sentiment analysis model for terrorist attacks reviews on Twitter. Sakarya University Journal of Science, 24(6), 1294–1302.

[8] Gabarron, E., Dorronzoro, E., Rivera-Romero, O., Wynn, R. (2019). Diabetes on Twitter: A sentiment analysis. Journal of Diabetes Science and Technology, 13(3), 439–444. https://doi.org/10.1177/1932296818811679.

[9] Abdullah, A. S., Degan, M., Mohammad, A. J. M. (2024). Audit committee accounting background and audit fees: Evidence from Malaysia. *Journal of Accounting and Financial Studies* (JAFS), 19(68), 341–349.

[10] Ghaderi, R. (2015). Improving the retrieval of related questions in Stack Overflow (Doctoral dissertation, University of California, Irvine).

[11] Govindarajan, M. (2013). Sentiment analysis of movie reviews using hybrid method of Naive Bayes and genetic algorithm. *International Journal of Advanced Computer Research*, 3(4), 139–145.

[12] Iqbal, F., Hashmi, J. M., Fung, B. C., Batool, R., Khattak, A. M., Aleem, S., Hung, P. C. (2019). A hybrid framework for sentiment analysis using genetic algorithm-based feature reduction. *IEEE Access*, 7, 14637–14652.

[13] Jain, A. P., Dandannavar, P. (2016). Application of machine learning techniques to sentiment analysis. *In 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology* (iCATccT) (p. 628–632). *IEEE*.

[14] Jansen, B. J. (2010). Gen X and Y's attitudes on using social media platforms for opinion sharing. In CHI '10 Extended Abstracts on Human Factors in Computing Systems (p. 3853–3858). *ACM*.

[15] Kamel, M., Selim, S. Z. (1994). The application of induction learning algorithm. Pattern Recognition, 27(3), 421–428.

[16] Medhat, W., Hassan, A., Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), 1093–1113.

[17] Nahar, K. M., Jaradat, A., Atoum, M. S., Ibrahim, F. (2020). Sentiment analysis and classification of Arab Jordanian Facebook comments for Jordanian telecom companies using lexicon based approach and machine learning. International Journal on Emerging Technologies, 6(3), 52–71.

[18] Khan, S. I., Abdullah, N. N., Degan, M., Mohammed, A. J., Shexa, N. S. (2023). Consequences of job insecurity on job performance among bank employees: Exploring the work engagement as mediating role. *Mitanni Journal of Humanitarian Sciences*, 4(2), 528–534. https://doi.org/10.25156/ptjhss.v4n2y2023.p528-534.

[19] Pang, X., Zhou, Y., Wang, P., Lin, W., Chang, V. (2020). An innovative neural network approach for stock market prediction. The Journal of Supercomputing, 76(3), 2098–2118. https://doi.org/10.1007/s11227-019-02754.

[20] Rajaraman, A., Ullman, J. D. (2011). Data mining. In Mining of Massive Datasets (p. 1–17). Cambridge University Press. https://doi.org/10.1017/CBO9781139058452.002.

[21] Ren, Y. W. (2016). A topic enhanced word embedding for Twitter sentiment classification. Information Sciences, 369, 188–198.

[22] Roberts, K. R. (2012). EmpaTweet: Annotating and detecting emotions on Twitter. In Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12) (p. 3806–3813).

[23] Singhal, T. (2020). A review of coronavirus disease 2019 (COVID-19). The Indian Journal of Pediatrics, 87, 281–286. https://doi.org/10.1007/s12098-020-03263-6.

[24] Alfawareh, F. S., Degan, M., Mohammad, A. J. (2025). The influence of FinTech and gender diversity on bank financial stability: Experience from Jordan. Discover Sustainability, 6, Article 511. https://doi.org/10.1007/s43621-025-01382-8.

[25] Soliman, T. H., Elmasry, M. A., Hedar, A., Doss, M. M. (2014). Sentiment analysis of Arabic slang comments on Facebook. *International Journal of Computers Technology*, 12(5), 3470–3478.

[26] Soumeur, A., Mokdadi, M., Guessoum, A., Daoud, A. (2018). Sentiment analysis of users on social networks: Overcoming the challenge of the loose usages of the Algerian dialect. *Procedia Computer Science,* 142, 26–37.

[27] Degan, M., Al-Abeej, A. S., Obaid, A. H. (2025). The effect of artificial intelligence on accounting and finance. *International Journal of Scientific Conferences*, 26, 764–777.

[28] Verma, B., Thakur, R. S. (2018). Sentiment analysis using lexicon and machine learning-based approaches: A survey. *In Proceedings of International Conference on Recent Advancement on Computer and Communication* (p. 441–447). *Springer*.

[29] Xhymshiti, M. (2020). Domain independence of machine learning and lexicon based methods in sentiment analysis (Bachelor's thesis, University of Twente).

[30] Yin, M., Wortman Vaughan, J., Wallach, H. (2019, May). Understanding the effect of accuracy on trust in machine learning models. *In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (p. 1–12). *ACM*.