

An Ensemble Classification of Mental Health in Malaysia related to the Covid-19 Pandemic using Social Media Sentiment Analysis

Nur 'Aisyah Binti Zakaria Adli¹, Muneer Ahmad^{2,*}, Norjihhan Abdul Ghani^{3*},
Sri Devi Ravana⁴, and Azah Anir Norman⁵

^{1,3,4,5}Department of Information Systems, Faculty of Computer Science and Information Technology, Universiti Malaya, 50603 Kuala Lumpur, Malaysia.

[e-mail: ¹aisyahzak@yahoo.com.my, ³norjihhan@um.edu.my, ⁴sdevi@um.edu.my, ⁵azahnorman@um.edu.my]

²Department of Computer Science, University of Roehampton, London SW15 5PH

[e-mail: ²muneer.ahmad@roehampton.ac.uk]

*Corresponding author: Muneer Ahmad, Norjihhan Abdul Ghani

*Received April 8, 2023; revised October 8, 2023; accepted October 26, 2023;
published February 29, 2024*

Abstract

COVID-19 was declared a pandemic by the World Health Organization (WHO) on 30 January 2020. The lifestyle of people all over the world has changed since. In most cases, the pandemic has appeared to create severe mental disorders, anxieties, and depression among people. Mostly, the researchers have been conducting surveys to identify the impacts of the pandemic on the mental health of people. Despite the better quality, tailored, and more specific data that can be generated by surveys, social media offers great insights into revealing the impact of the pandemic on mental health. Since people feel connected on social media, thus, this study aims to get the people's sentiments about the pandemic related to mental issues. Word Cloud was used to visualize and identify the most frequent keywords related to COVID-19 and mental health disorders. This study employs Majority Voting Ensemble (MVE) classification and individual classifiers such as Naïve Bayes (NB), Support Vector Machine (SVM), and Logistic Regression (LR) to classify the sentiment through tweets. The tweets were classified into either positive, neutral, or negative using the Valence Aware Dictionary or sEntiment Reasoner (VADER). Confusion matrix and classification reports bestow the precision, recall, and F1-score in identifying the best algorithm for classifying the sentiments.

Keywords: COVID-19, Ensemble Classification Method, Sentiment Analysis.

1. Introduction

In late 2019, the world grappled with the emergence of a novel coronavirus, famously known as SARS-CoV-2 [1]. Its origins are traced back to Wuhan City, Hubei Province, China, where it was suspected to have jumped from pangolins, snakes, and bats in the Wuhan wet markets [2], [3]. As the calendar flipped to January 30, 2020, the World Health Organization somberly declared the onset of a global pandemic, a declaration that set the stage for a relentless spread, with millions of cases and countless fatalities echoing across continents [4].

The far-reaching impact of this viral intruder was notably felt among high-risk individuals and seniors [6]. Those with pre-existing systemic health and social conditions found themselves at a heightened risk of falling critically ill and succumbing to the clutches of COVID-19, ushering in an era of profound physical and mental tribulations [6]. The novel nature of the virus added an air of uncertainty, casting a shadow over even the younger, healthier population as the true extent of potential long-term effects remained shrouded in mystery.

However, the ripples of the pandemic extended far beyond the infected. COVID-19, with its insidious spread, swiftly rewrote the norms of societies worldwide, prompting the implementation of rigorous Standard Operating Procedures (SOPs) [7]. Overburdened healthcare systems and financial losses became the unfortunate byproducts, laying the groundwork for a global stage where psychological distress and socio-economic strains played out in unison, transcending geographical boundaries [8].

Unpacking the layers of psychological impacts revealed a spectrum of struggles, from anxiety and depression to unhealthy coping mechanisms [9]. Vulnerable groups, including the elderly, homeless individuals, and ethnic minorities, found themselves navigating even more treacherous waters, as the pandemic exacerbated existing challenges [10], [11]. The fallout from these psychological reactions ranged from collective hysteria to more personal despair, potentially culminating in tragic outcomes like suicide.

Yet, the toll of COVID-19 transcended the realm of mental health, infiltrating the very fabric of non-COVID-19-related healthcare services globally [12]. Nations worldwide grappled with the implementation of restrictions and movement controls, disrupting the delicate balance of social and economic aspects. The resultant fallout, marked by skyrocketing unemployment rates and mental health issues, painted a bleak picture on the socio-economic canvas [13]. Domestic violence, a silent pandemic within the pandemic, emerged as a significant contributing factor, echoing the haunting echoes of similar psychological scars left by past epidemics [14].

As the economic recession triggered by the pandemic deepened, so did the shadows cast upon mental health [15]. Job loss, a harsh reality for many, especially in the adult demographic, brought with it a staggering 53 percent rate of depression and/or anxiety symptoms, eclipsing the mental well-being of those fortunate enough to retain their livelihoods. The younger generation, aged 18 to 24, bore the brunt with a reported 56 percent grappling with depression and/or anxiety, highlighting the disproportionate impact across age groups. Forecasts painted a grim picture, predicting a surge in self-harm, depression, substance use, and the heartbreaking rise in suicide rates [16].

This research study glimpses the following significant contributions,

- The study emphasizes the widespread impact of the pandemic on mental health, exploring severe disorders, anxieties, and depression.
- Social media is recognized as a valuable source for insights into the pandemic's mental health impact due to global connectivity.

- The study innovatively employs Word Cloud to visualize and identify frequent keywords associated with COVID-19 and mental health.
- Proposing a robust solution, the study integrates a sophisticated sentiment analysis framework, combining Majority Voting Ensemble and individual classifiers (Naïve Bayes, SVM, LR) for nuanced tweet sentiment classification.
- Utilizing the Valence Aware Dictionary or sEntiment Reasoner (VADER) enhances the precision of sentiment classification, categorizing tweets into positive, neutral, or negative sentiments.
- To validate the solution's effectiveness, the study adopts a thorough performance evaluation approach, using a confusion matrix and classification reports to quantify precision, recall, and F1-score metrics for sentiment classification algorithms.

The rest of the manuscript is organized as under, section two describes the review of relevant literature, section three presents the methodology, section four highlights the results and discussion, and section five concludes the discussion.

2. Literature Review

The research delves into the profound impact of the COVID-19 pandemic on global mental health, recognizing the surge in severe disorders, anxieties, and depression among people worldwide. With a notable shift in lifestyles since the WHO's pandemic declaration in January 2020, researchers have predominantly relied on surveys to assess mental health impacts. However, this study proposes a unique approach, leveraging social media as a rich source of insights into pandemic-related mental health sentiments. Employing innovative tools such as Word Clouds and sentiment analysis algorithms (Naïve Bayes, SVM, LR, Majority Voting Ensemble), the research aims to provide a nuanced understanding of sentiments expressed on platforms like Twitter, contributing valuable perspectives to the discourse on mental health challenges during the ongoing global crisis.

Generally, social media analytics (SMA) is the process of feeding social media data collected from social media platforms into analytics tools to extract information or insights that can be employed in decision-making tasks in any organization. SMA plays a significant role in both monodisciplinary and multidisciplinary fields because of how convenient it is for any researcher to get social media data online and from any desired platforms, despite the challenges in its technicality part such as running codes for web scraping or the ethical issues from extracting online users' details. Nevertheless, SMA is undoubtedly widely used in the healthcare industry as it is one of the most optimal ways to get closer to the public for various reasons such as proper information dissemination, raking in users' response towards the healthcare services provided, and others. However, the study is particularly inclined towards how SMA and mental health issues fared during the trying times of COVID-19. What better way than to get the answer from the direct source itself from platforms such as Twitter?

Based on the literature reviewed, several authors have concluded that COVID-19 did increase mental health issues among people. [17], [18] [19], [20] investigated the impacts of the COVID-19 outbreak on mental health in Malaysia and mentioned that a surge in anxiety, stress, obsessive-compulsive disorders, and effects on society due to the lockdown are to be expected. [21], [22] investigated international samples of Twitter content consisting of mental health and pregnancy during the COVID-19 pandemic mentioned that from the sentiment analysis, it was found that individuals and companies usually tweeted negatively, while

researchers and professionals tweeted neutrally. Additionally, from the thematic analysis, it was found that the most common themes discussed were stress and anxiety, followed by isolation, depression, and sleep difficulties.

Previous researchers observed to have used hashtags or location tags to identify some topics of certain countries. [23] collected their Twitter data using COVID-19 related keywords such as #COVID19Italia, #COVID19Italy, and #Italy to identify the COVID-19 spread in Italy. Similarly, [24] used the same method in collecting their dataset, where they used hashtags such as #IndiaLockdown and #IndiafightsCorona to understand the impact of the lockdown in India on their citizen during the COVID-19 outbreak. However, [25] observed to have used geographical filtering to filter only people based in India when analyzing the reaction of Indians towards the topic of stress, anxiety, and trauma that was built upon COVID-19 and the primary cause of them through Twitter.

Moreover, previous authors also employed certain keywords related to either both COVID-19 and mental health or either one of those keywords when gathering their dataset. [26] searched for keywords related to COVID-19, such as, "COVID", "corona", and "coronavirus" to identify, analyze, and visualize emotions in the tweets, whereas [27] gathered their tweets based on 20 keywords such as "coronavirus", "COVID-19", "pandemic", "quarantine", "virus", "lockdown", "stay home", "coronavirus", "social distancing", "new cases", "2019nCoV, and "coronaoutbreak". [28] did their study to investigate international samples of Twitter content consisting of mental health and pregnancy during the COVID-19 pandemic by gathering dataset that contained 14 hashtags related to COVID-19 such as "coronavirus", "covid-19", "virus", and "pandemicpregnancy", 21 hashtags related to pregnancy such as "birth", "laboranddelivery", "pregnancy", and "postpartum", and 11 hashtags related to mental health such as "anxiety", "worried", "depression to investigate international samples of Twitter content consisting of mental health and pregnancy during the COVID-19 pandemic. Some studies acquired their dataset by searching for the pair of both COVID-19 and mental health keywords, such as "COVID-19" and "loneliness" [29] to understand the loneliness topic during the COVID-19 crisis.

Additionally, previous studies were seen to gather data on the English language [30], non-retweet tweets, and tweets by individuals with no affiliation to any organization. In terms of visualization, some authors visualized their dataset using word cloud [24]. **Table 1** shows a list of reviews done on studies for sentiment analysis tools.

Table 1. Sentiment analysis

Citation	Concept	Sentiment Analysis tool	Best Model	Accuracy	F1-score
[25] (2021)	Analyzed the reaction of people in India towards the topic of stress, anxiety, and trauma that was built upon COVID-19 and the primary cause of them through Twitter.	Textblob	General ML/DL models	Shown improved than existing methods.	Presented enhanced as compared to other methods.
[28] (2021)	Investigated international samples of Twitter content consisting of mental health and pregnancy during the COVID-19 pandemic.	VADER	General ML/DL models	Shown improved than existing methods.	Presented enhanced than other methods.

[31] (2020)	Experimented to investigate VADER performance against NLTK analyzer for multi-classification system.	VADER	General ML/DL models	Shown improved than existing methods.	Presented enhanced than other methods.
[17] (2020)	Investigated the sentiment on US mental health during the COVID-19 pandemic using Twitter data.	VADER	General ML/DL models	Shown improved than existing methods.	Presented enhanced than other methods.
[32] (2021) , [33] (2021)	Investigated the sentiment on mental health from Twitter data to categorize the mental health of an individual compared to their well-being.	Textblob, MonkeyLearn	General ML/DL models	Shown improved than existing methods.	Presented enhanced than other methods.

We have highlighted the limitations of this study as follows,

- Social media data may introduce bias, as users may not represent the entire population.
- Findings might not apply universally, as sentiments on social media vary across demographics and cultures.
- Interpretation of Word Clouds is subjective and may not fully capture nuances in tweet sentiments.
- Sentiment classification algorithms (Naïve Bayes, SVM, LR) and Majority Voting Ensemble effectiveness can be influenced by training data and may not reflect human emotions accurately.
- The study may not capture evolving sentiments over time, missing the dynamic nature of the pandemic's impact on mental health.
- The study identifies correlations but doesn't establish causation between social media sentiment and mental health issues.
- Ethical concerns include privacy issues and the potential exploitation of personal information by using social media data.
- Limited focus on anxiety, depression, and severe disorders may overlook other nuanced mental health issues.
- The assumption that social media connectivity reflects genuine emotional expression may oversimplify human communication and emotions.

Table 2. Machine Learning on sentiment analysis on mental health

Citation	Concept	Sentiment Analysis tool	Best Model	Accuracy	F1-score
[34] (2018)	Implemented NLP and machine learning techniques on Twitter posts for sentiment analysis specifically for depression.	MNB, SVM	MNB	0.836	0.833
[35] (2020)	Understood the mental health of	LR	LR	0.947	0.947

	individuals through public tweets using sentiment analysis.				
[36] (2016)	Discussed Twitter sentiment analysis system using features extractions such as Unigram and Bigram, NLP, and machine learning techniques.	NB, SVM, MaxEnt	Not stated	Not stated	Not stated
[37] (2021)	Explored topics and discussions in tweets posted from Brazil and the USA related to COVID-19.	NB, MLP RF, AdaBoost, LR, LSVM	RF	Not stated	0.86
[38] (2020)	Identified any suicidal behavior in Twitter data using the WEKA tool.	NB, DT, MNB, LR, SVM	LR	0.955	Not stated
[39] (2019)	Explored discussion related to #MyTipsForMentalHealth during the World Mental Health Awareness Day on Twitter in 2017.	SVM, RF, LR	SVM	0.81	Not stated
[40] (2021)	Investigated the sentiment on social distancing in Canada using the Twitter approach.	SVM (and SentiStrength v2.3 tool)	SVM (and SentiStrength)	0.71	Not stated
[32] (2021)	Investigated the sentiment on mental health from Twitter data to categorize the mental health of an individual compared to their well-being.	SVM	SVM	0.67	Not Stated

Based on [Table 2](#), most of the sentiment analysis on mental health using Twitter data had SVM, LR, MNB, and RF as their best models. In summary,

- [34] Implemented NLP and machine learning techniques on Twitter posts for depression sentiment analysis, achieving an accuracy of 83.6% and an F1-score of 83.3%, utilizing models like MNB and SVM.
- [35] Examined mental health through public tweets, employing LR as the sentiment analysis tool, resulting in a high accuracy of 94.7% and an F1-score of 94.7%.
- [36] Explored Twitter sentiment analysis using features like Unigram, Bigram, NLP, and machine learning techniques, employing NB, SVM, and MaxEnt, with accuracy and F1-score not stated.
- [37] Investigated COVID-19-related discussions in tweets from Brazil and the USA, using multiple models (NB, MLP, RF, AdaBoost, LR, LSVM), with RF as the best model, achieving an F1-score of 86%.
- [38] Identified suicidal behavior in Twitter data using the WEKA tool, employing models like NB, DT, MNB, LR, and SVM, with LR as the best model, achieving an accuracy of 95.5%.
- [39] Explored discussions related to #MyTipsForMentalHealth on Twitter during World Mental Health Awareness Day in 2017, using SVM, RF, and LR, with SVM achieving an accuracy of 81%.
- [40] Investigated sentiment on social distancing in Canada using the Twitter approach, employing SVM and SentiStrength v2.3 tool, with SVM achieving an accuracy of 71%.
- [32] Examined sentiment on mental health from Twitter data, categorizing individuals' mental health compared to personal well-being, using SVM with an accuracy of 67% (F1-score not stated)

2.1 Ensembled Classification

Due to a problem identified when using an individual classifier, a hybrid or an ensemble classification has been used throughout many studies to create a new and improvised classifier that combines multiple individual classifiers as one classifier. However, there is very little approach to the ensemble classification of mental health sentiment on Twitter. Hence, [Table 3](#) shows the ensemble classification that exists in a different area of sentiment analysis using Twitter data.

Table 3. Ensembled classification in Twitter sentiment analysis

Citation	Machine Learning Model					Best Individual Classifier	Result
	NB	RF	SVM	LR	Hybrid Models		
[41] (2018)	✓	✓	✓	✓	NB + RF + SVM + LR	NB	The hybrid model outperformed the stand-alone classifiers.
[42] (2020)	✓	x	x	✓	MLP	NB	The hybrid model outperformed the individual classifiers.

[43] (2019)	✓	x	✓	x	SVM2AN2	SVM	The hybrid model outperformed the individual classifiers.
[44] (2018)	✓	✓	x	x	NB + RF + k-NN	k-NN	The hybrid model outperformed the individual classifiers.
[38] (2020)	✓	✓	✓	✓	AdaBoost, Voting Ensemble, Bagging Stacking	RF	RF outperformed the hybrid model.
[45] (2020)	✓	x	✓	✓	AdaBoost, SMO, HELM	SL	The hybrid model outperformed the individual classifiers.
[46] (2019)	✓	x	x	x	ExaTree	AdaBoost	The hybrid model outperformed the individual classifiers.
[47] (2018)	✓	x	✓	✓	MVE	Not stated	The hybrid model outperformed the individual classifiers.
[48] (2017)	✓	✓	x	✓	NB + RF + DT + LR	Not stated	The hybrid model outperformed the individual classifiers.
[49] (2020)	✓	x	✓	x	NB + NN + BN + MaxEnt + SVM	Not stated	The hybrid model outperformed the individual classifiers.

As observed in **Table 3**, hybrid models tend to outperform the individual classifiers. We can summarize it as follows,

- [41] Employed NB, RF, SVM, and LR individually, with a hybrid model combining all four. The hybrid model outperformed standalone classifiers, with NB being the best individual classifier.
- [42] Utilized NB, LR, and MLP individually, and a hybrid model combining them. The hybrid model demonstrated superior performance compared to individual classifiers, with NB as the best individual classifier.
- [43] Employed NB and SVM2AN2 individually, with a hybrid model combining them. The hybrid model surpassed individual classifiers, with SVM being the best individual classifier.
- [44] Employed NB, RF, and k-NN individually, and a hybrid model combining them. The hybrid model outperformed individual classifiers, with k-NN being the best individual classifier.
- [38] Implemented AdaBoost, Voting Ensemble, Bagging, and Stacking individually, with a hybrid model. RF outperformed the hybrid model.
- [45] Utilized NB, LR, AdaBoost, SMO, and HELM individually, and in a hybrid model. The hybrid model outperformed individual classifiers, with SL being the best individual classifier.
- [46] Employed ExaTree individually, and a hybrid model with AdaBoost. The hybrid model outperformed individual classifiers.
- [47] Utilized NB and MVE individually, with a hybrid model. The hybrid model outperformed individual classifiers, with classification tools not stated.
- [48] Employed NB, RF, DT, and LR individually, and a hybrid model combining them. The hybrid model outperformed individual classifiers, with classification tools not stated.
- [49] Utilized NB, NN, BN, MaxEnt, and SVM individually, with a hybrid model. The hybrid model outperformed individual classifiers, with classification tools not stated.

Table 4. Summary of performance of ensembled classifiers and best individual classifiers

Citation	Type of Classifier	Classifier	Accuracy	Recall	F1-score
[41] (2018)	Hybrid	NB + RF + SVM + LR	Not stated	Not stated	0.767
	Individual	NB	Not stated	Not stated	0.756
[42] (2020)	Hybrid	MLP	0.82	Not stated	Not stated
	Individual	NB	0.801	Not stated	Not stated
[43] (2019)	Hybrid	SVM2AN2	0.974	0.963	0.964
	Individual	SVM	0.963	0.973	0.976
[44] (2018)	Hybrid	NB + RF + k-NN	0.787	Not stated	Not stated
	Individual	k-NN	0.636	Not stated	Not stated
[38] (2020)	Hybrid	Not stated	Not stated	Not stated	Not stated

	Individual	RF	0.985	0.982	Not stated
[45] (2020)	Hybrid	HELM	0.882	Not stated	Not stated
	Individual	SL (with IG+LSA feature selection)	0.882	Not stated	Not stated
[46] (2019)	Hybrid	ExtraTree	0.76	Not stated	Not stated
	Individual	AdaBoost	0.74	Not stated	Not stated
[47] (2018)	Hybrid	MVE	Not stated	Not stated	Not stated
	Individual	Not stated	Not stated	Not stated	Not Stated
[48] (2017)	Hybrid	NB + RF + DT + LR	0.916	Not stated	0.916
	Individual	Not stated	Not stated	Not stated	Not Stated

We can summarize **Table 4** as follows,

- [41] Hybrid model (NB + RF + SVM + LR) achieved an F1-score of 0.767; individual NB achieved an F1-score of 0.756, with accuracy and recall not stated.
- [42] Hybrid MLP achieved an accuracy of 0.82; individual NB achieved an accuracy of 0.801, with recall and F1-score not stated.
- [43] Hybrid SVM2AN2 achieved an accuracy of 0.974, recall of 0.963, and F1-score of 0.964; individual SVM achieved an accuracy of 0.963, recall of 0.973, and F1-score of 0.976.
- [44] The hybrid model (NB + RF + k-NN) achieved an accuracy of 0.787; individual k-NN achieved an accuracy of 0.636, with recall and F1-score not stated.
- [38] Hybrid model not stated; individual RF achieved an accuracy of 0.985 and recall of 0.982, with F1-score not stated.
- [45] Hybrid HELM achieved an accuracy of 0.882; individual SL (with IG+LSA feature selection) achieved an accuracy of 0.882, with recall and F1-score not stated.
- [46] Hybrid ExtraTree achieved an accuracy of 0.76; individual AdaBoost achieved an accuracy of 0.74, with recall and F1-score not stated.
- [47] Hybrid MVE, individual classifier details, accuracy, recall, and F1-score not stated.
- [48] Hybrid model (NB + RF + DT + LR) achieved an accuracy of 0.916 and F1-score of 0.916; individual classifier details, accuracy, recall, and F1-score not stated.

As observed in the related works, ensemble classifications showed improvements in terms of accuracy and F-measure compared to stand-alone classifiers. Additionally, in terms of the stand-alone classifiers adopted in previous studies as referred to in **Table 2** and **Table 4**, it was noted that NB, SVM, and LR seem to outperform other classifiers when compared against.

3. Methodology

This study deviates from traditional survey-based approaches in understanding people's views on the COVID-19 pandemic and mental health by developing an ensemble classification model using Twitter data. Leveraging the prevalence of sentiment analysis in the era of rapid social media development, the study aims to capture people's sentiments regarding the pandemic and mental health issues in Malaysia. The methodology involves three key analyses: sentiment analysis, word cloud visualization, and model evaluation. Raw data extraction is followed by preprocessing and tweet classification. Data collection is focused on keywords such as "COVID-19" and "mental health," with location-based tagging filtering for "Malaysia." Word cloud visualization is employed to identify prevalent keywords. The VADER lexicon model is used for sentiment extraction, categorizing tweets into positive, neutral, or negative. An ensemble classification approach incorporating SVM, NB, and LR—identified as high-performing classifiers—is utilized. The study concludes with a confusion matrix and classification reports, providing precision, recall, and F1-score metrics to determine the most effective algorithm for sentiment classification.

The proposed methodology is given below,

1. Load Twitter data into a dataset.
 2. Extract relevant columns such as "tweets" and "sentiments."
 3. Data Pre-processing:
 - a. Tokenization, stemming, and other text processing steps.
 - b. Split the data into training and testing sets: (D_{train} , D_{test}).
 4. Word Cloud Visualization:
 - a. Generate a word cloud to visualize and identify frequent keywords.
 5. Sentiment Analysis using VADER:
 - a. Apply the VADER lexicon model to extract sentiment features: $\text{VADER}(T_i)$ for each tweet (T_i).
 - b. Categorize tweets into positive ((P)), neutral ((N)), or negative ((Neg)) sentiments: $\text{Sentiment}(T_i)$.
 6. Ensemble Classification Model:
 - a. Choose multiple classifiers (e.g., SVM, NB, LR, RF).
 - b. Implement a voting ensemble model with selected classifiers:

$$\text{Ensemble}(T_i) = \text{Voting}(\text{SVM}(T_i), \text{NB}(T_i), \text{LR}(T_i), \text{RF}(T_i))$$
 7. Train the Ensemble Model:
 - a. Use the training set to train the ensemble classification model.
 8. Prediction:
 - a. Predict sentiments on the testing set using the trained ensemble model: $\hat{\text{Sentiment}}(T_i)$.
- Model Evaluation:
- a. Evaluate the model using:
 - Confusion Matrix: $\text{Confusion Matrix} = \begin{bmatrix} TP & FP \\ FN & TN \end{bmatrix}$
 - Classification Report (precision, recall, F1-score, etc.): $\text{Classification Report} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$
10. Output Results:
 - a. Display the results of the confusion matrix and classification report.

Previous research mostly used surveys or questionnaires to obtain people's views on the COVID-19 pandemic and associated mental health issues. This study develops an ensemble classification model to understand the sentiment between the COVID-19 pandemic and mental health in Malaysia through Twitter data. Due to the rapid development of social media, sentiment analysis has become one of the trendy approaches to be used today when it comes to understanding people's thoughts. Since people feel connected on social media, this study aims to get the people's sentiments about the pandemic related to mental issues based on Twitter data analysis.

The methodology of this study consists of three key analyses: sentiment analysis, word cloud, and model evaluation. **Fig. 1** demonstrates the workflow of the proposed approach. It begins with the extraction of raw data and continues to the data pre-processing section followed by the classification of tweets.

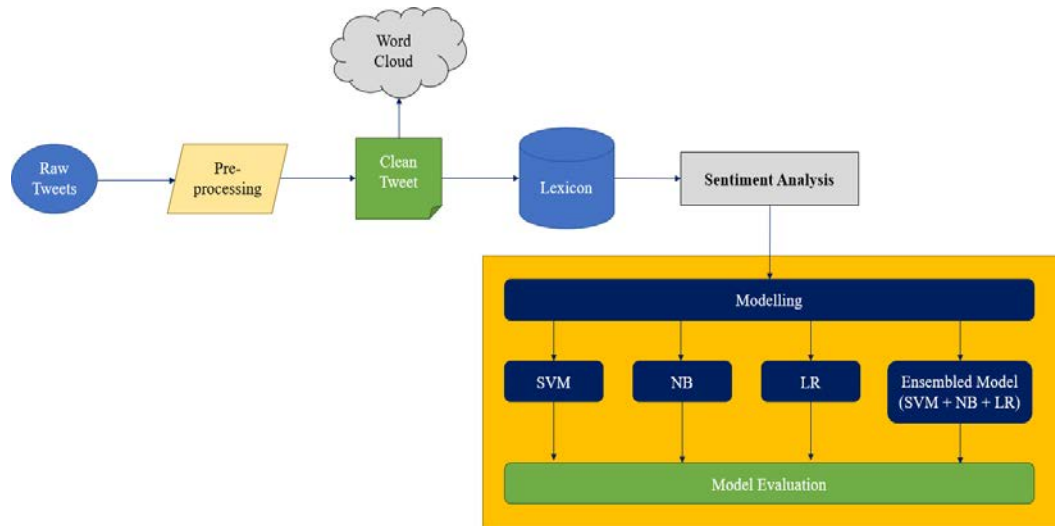


Fig. 1. Proposed methodology

Based on the previous studies, the respective authors searched data based on keywords, such as "COVID-19" to collect data related to COVID-19 from the tweets and/or "mental health" to get data related to mental health tweets. Additionally, since this paper is studying the sentiment of COVID-19 towards mental health in Malaysia, location-based tagging will be used to filter data related to "Malaysia" only. Since word cloud has been used a lot in previous studies [50]–[53], this study will employ word cloud to visualize and identify the most frequent keywords related to COVID-19 and mental health disorders. The lexicon model, VADER which is one of the well-known models used in natural language processing will be used to extract features from text to classify them into either positive, neutral, or negative. This study employs an ensemble classification of known classification algorithms to classify the sentiments through tweets. Machine learning techniques such as SVM, NB, and LR were selected after reviewing **Table 2** and **Table 4**, which showed these three classifiers as among high accuracy and working well together. The tweets will be classified into either positive, neutral, or negative. Confusion matrix and classification reports bestow the precision, recall, and F1-score in identifying the best algorithm for classifying the sentiments.

4. Experiment and Results

The raw dataset was scraped using scrape modules for Twitter. For this study, only relevant data to the scope of this study were scraped, such as:

- date: Creation date and time of the tweet, UTC format. (Column: *Datetime*)
- id: A unique identifier of the tweet. (Column: *Tweet Id*)
- content: Content of the tweet. Keywords can be used to search for relevant tweets, non-sensitive. (Column: *Tweet*)
- user.username: Username of the user. (Column: *Username*)
- user.verified: An identifier of whether the Twitter account is verified (or an official account) or not. If verified, then "True", else "False". (Column: *isVerified*)
- user.location: The location of the user, can appear as the name of a city, or a country. (Column: *UserLocation*)
- lang: Predicted language of the tweet, auto-generated. (Column: *TweetLanguage*)
- retweetedTweet: If the tweet is a retweeted tweet, then a tweet id will be displayed, else it will appear as "None". (Column: *isRetweeted*)

There are two types of keywords listed, which were COVID19-related keywords, such as "covid19", "2019-ncov", "coronavirus", and "ncov", and mental health-related keywords, such as "depression", "anxiety", "disorder", and "trauma". These keywords are then nested together and appended into one list to appear such as "covid19 depression", "2019-ncov depression", and "ncov trauma". The complete details of the keywords used are in [Table 5](#) below:

Table 5. Topics and keywords related to COVID-19 tweets

Topic	Keywords
Keywords related to COVID-19	'corona', 'coronavirus', 'covid19', 'covid', 'covid-19', 'sarscov2', 'sars', 'cov2', 'cov 2', 'covid_19', 'ncov', 'ncov2019', '2019-ncov', 'pandemic', '2019ncov', 'chinese virus', 'chinesevirus', 'wuhan virus', and 'wuhanvirus'
Keywords related to mental health	'mental', 'depression', 'anxiety', 'disorder', 'insomnia', 'bipolar', 'stress', 'psycho', 'fear', 'suicide', 'dissociation', 'dissociative', 'paranoia', 'ptsd', 'trauma', 'giving up', 'fatality', 'psychiatrist', 'loneliness', 'uncertainties', and 'worried'

Afterward, scrape used these keywords to search relevant content of the tweets that contain these keywords and produce the results. Besides the listed keywords, the author also included keywords where "COVID-19" and "Malaysia" were in a tweet. Additionally, only tweets where the "language" column has an "en" value were scraped.

4.1 Data Pre-Processing

We present here the algorithm of data preprocessing we performed on our data.

Algorithm: Twitter Data Pre-processing

Input: Raw Twitter dataset

Part 1: Initial Cleaning and Filtering

(a) Filter tweets from Malaysia based on user location.

(b) Exclude retweeted data where "isRetweeted" is "None".

- (c) Exclude verified (official) accounts based on "isVerified" and username criteria.
 - (d) Convert "Datetime" column to proper datetime format, adjust for MYT (Malaysia Time).
 - (e) Add new columns for month and hour based on the "Datetime" column.
 - (f) Clean the "location" column for consistency.
 - (g) Remove unnecessary columns: "Datetime", "isVerified", "TweetLanguage", "Username", "isRetweeted".
 - (h) Remove duplicates based on the "Tweet" column.
 - (i) Remove the "Tweet Id" column.
- # Save cleaned data into a CSV file

Part 2: Further Text Cleaning for Analysis and Modeling

- (a) Manually clean the CSV file to ensure data integrity.
 - (b) Remove any duplicated tweets.
 - (c) Remove stop words, punctuations, emojis, mentions, URLs, and hashtags.
 - (d) Trim and clean any white spaces.
 - (e) Decode ASCII codes.
 - (f) Perform lemmatization.
 - (g) Perform stemming.
 - (h) Tokenization.
 - (i) Vectorize the text for analysis and modeling.
- # Output: Cleaned and pre-processed Twitter dataset ready for analysis and modeling.

We further highlight here the working of the preprocessing algorithm. In this stage, there are two parts of data pre-processing. The **first part** involved the steps below before saving the cleaned dataset into a csv format:

- (a) Included tweets from Malaysia based on the location of the user and removed other countries and unidentified locations.
- (b) Excluded retweeted data, where "isRetweeted" column is "None".
- (c) Excluded any verified (or official) accounts from the dataset, where "isVerified" column is "None", and the username does not contain any of these words: reporter, bot, yahoo, humanity, tv, fake, online, india, bhd, berhad, asia, _my, global, radio, officialupnm, group, fmtoday, express, malaymail, cinema, mall, gsc, tgv, world, fm, kl, news, post, radio, or malaysia.
- (d) Change the format of "Datetime" column to a proper datetime format in python and added 8 hours delta to ensure the "Datetime" column was in MYT.
- (e) Added new columns for month and hour based on the "datetime" column.
- (f) Cleaned the "location" column to ensure the values were consistent, for example, instead of "Kota Kinabalu, Sabah", this value was transformed to "Sabah".
- (g) Removed the "Datetime", "isVerified", "TweetLanguage", "Username", and "isRetweeted" columns.
- (h) Removed duplicates using "Tweet" as a reference column.
- (i) Removed "Tweet Id" column.

After saving the data into a csv format, another round of cleaning was done manually through the csv to ensure the data was included as per scope. The second part of the data pre-processing stage was more tailored to the data analysis and data modeling part. The reason for having the first part of data pre-processing was to avoid re-scraping the whole data from the beginning.

The **second part** of the data pre-processing stage focused on cleaning the text in tweets, such as:

- (a) Removing any duplicated tweets.
- (b) Removing stop words, punctuations, emojis, mentions, urls, and hashtags.
- (c) Trimming and cleaning any white spaces.
- (d) Decoding ascii codes.
- (e) Lemmatization.
- (f) Stemming.
- (g) Tokenization.
- (h) Vectorizing

4.2 Exploratory Data Analysis

4.1.1 Data Attributes

During data exploration, it was noted that the dataset contained 4783 rows and four columns, which were the month, hour, tweet, and location.

4.1.1.1 Bar Plot

Three bar plots were built to observe the number of tweets posted by hour, month, and location. The purpose of this analysis is to investigate when the users usually posted about mental health and/or COVID-19, as well as to observe the trend of tweets over the past eight months on topics related to COVID-19 and mental health issues in Malaysia. Hence, for this analysis. The analysis on location was done to investigate the common location of users who posted such tweets.

For the number of tweets by hour, the x-axis represented the hour, where the values were from 0 to 23, whereas the y-axis represented the number of tweets, where each of the tweets was counted. **Fig. 4.1** shows the number of tweets by hour.

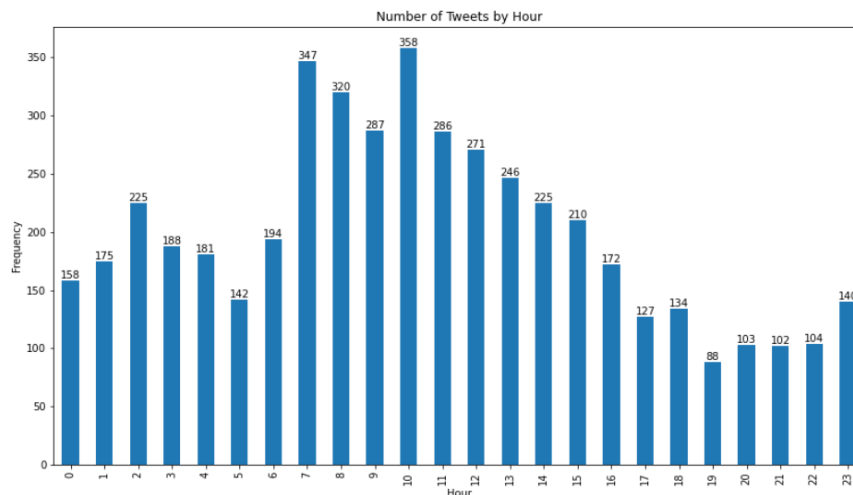


Fig. 2. Number of tweets by hour

As observed in **Fig. 2**, tweets were mostly posted in the morning between seven to ten. However, tweets were less likely to be posted around the evening between seven to ten. In addition, another bar plot for the number of tweets by month was created, where the x-axis

showed the month and the value started from one to eight, while the y-axis showed the number of tweets, where each of the tweets was counted. **Fig. 3** shows the number of tweets by month.

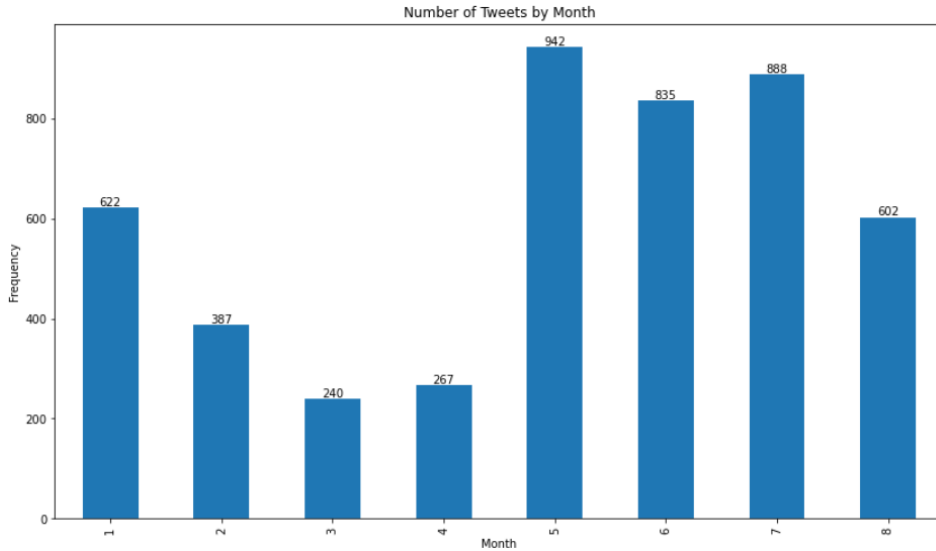


Fig. 3. Number of tweets by month

From **Fig. 3**, it was noted that out of the eight months, May showed the highest number of tweets, while March showed the lowest. Interestingly, the trend in this graph was similar to the trend of COVID-19 cases in Malaysia. Lastly, for the number of tweets by location, the x-axis represents the location, where the values represent all the states in Malaysia and Malaysia as a country, whereas the y-axis represents the number of tweets, where each of the tweets was counted. **Fig. 4** shows the number of tweets by location.

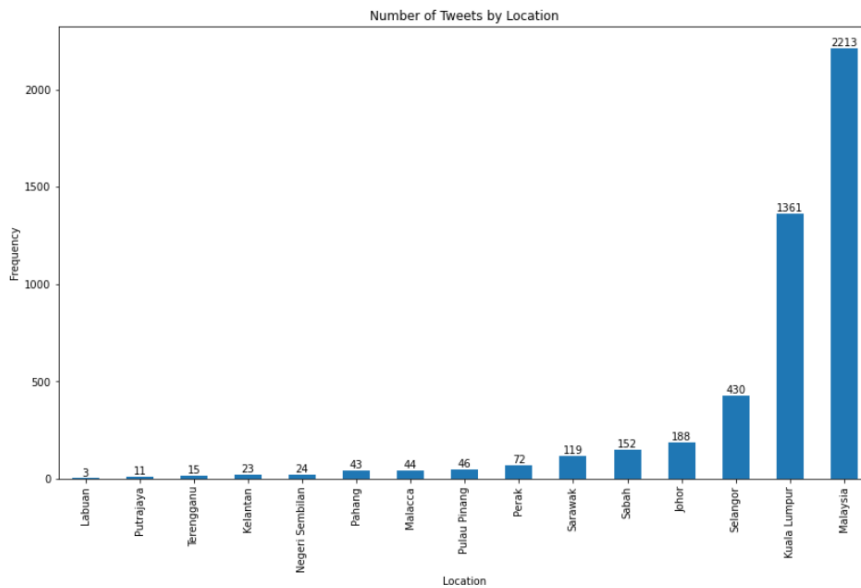


Fig. 4. Number of tweets by location

It was observed that users based in "Kuala Lumpur" and "Malaysia" often posted about COVID-19 and/or mental health issues, whereas users from "Labuan" were less likely to post on these topics. "Malaysia" value was input when the user only input "Malaysia" or if they did not put their location containing states in Malaysia.

4.1.1.2 Word Cloud

The word cloud analysis was used to visualize the common words used in all tweets. As observed in Fig. 5, the size of each word represented the counts of each word in the document, in this case, the tweets. The bigger the word in the word cloud, the higher the count of the word existing in the tweets, whereas the smaller the word showed that the word was less used in the tweets. Fig. 5 shows a word cloud where COVID-19 and mental health-related keywords were included.



Fig. 5. Word Cloud (with COVID-19 and mental health keywords)

As observed in Fig. 5, the commonly used words in the dataset were "corona", "mental", "impact", "life", "kill", "case", "health", "life", "safe", "physical", and "normal". However, when excluding COVID-19 and mental health-related keywords, the word occurrence can be observed in Fig. 6.

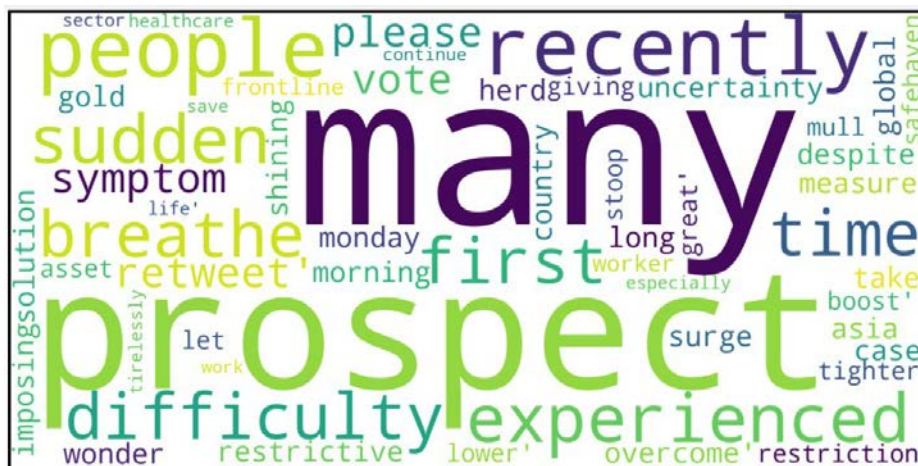


Fig. 6. Word Cloud (without COVID-19 and mental health keywords)

In **Fig. 6**, the common words observed were "prospect", "many", "difficulty", "experienced", "imposingsolution", "overcome", "recently", "people", "sudden", "breathe", "time", "first", "restriction", and "tighter". This could infer that when there were topics related to COVID-19 and/or mental health, these keywords were commonly used in the topic.

4.3 Sentiment Analysis

In this section, a sentiment analysis model implementing the Lexicon sentiment tool was used. The sentiment analysis labeled each data into either positive, neutral, or negative sentiment. During the process, only the "Tweet" column was selected. An object called *SentimentIntensityAnalyzer* which is one of the objects available in VADER was used to apply the polarity scores to the text. If the polarity score is above 0 then it can be considered as a positive sentiment, whereas if the polarity score was below 0, then it was considered as a negative sentiment, else neutral. Some examples of data after *SentimentIntensityAnalyzer* was implemented can be observed in **Table 6**.

Table 6. Sentiment polarity table

Tweet	Polarity Value	Negative Value	Neutral Value	Positive Value	Sentiment Type
COVID created increased anxiety depression insomnia	-0.2004	0.418	0.222	0.36	Negative
blessed supportive friends always bring positive vibes mental support covid fighters covid real guys	0.9451	0	0.317	0.683	Positive
worrying vaccinated	0	0	1	0	Neutral

Moreover, from the overall tweets, the polarity score seems higher towards the positive as seen in **Fig. 7**.

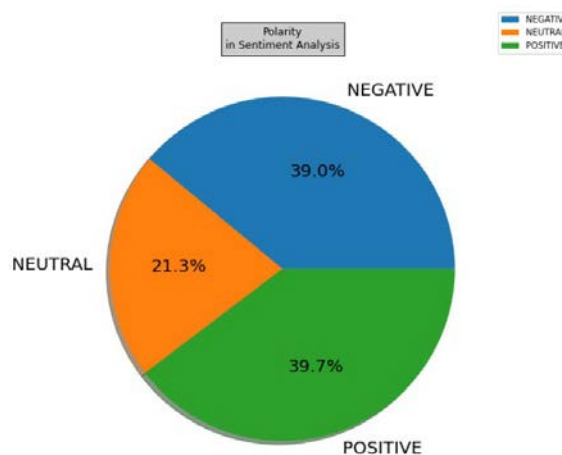


Fig. 7. Polarity pie chart

As observed, positive sentiment was slightly higher than negative sentiment, by 0.7 percent. The total count of tweets for each polar can be seen in **Table 7**.

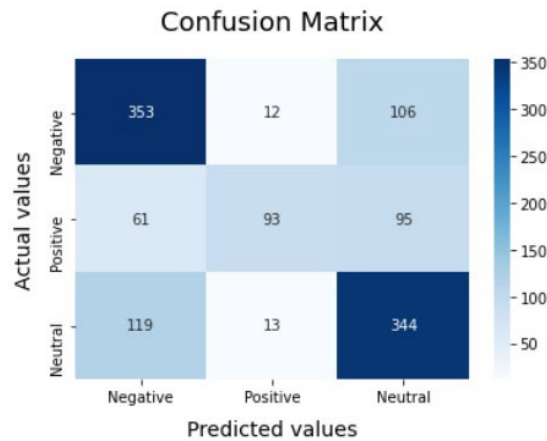
Table 7. Sentiment Polarity (Tweet Count)

Sentiment Type	Tweet Count
Positive	1906
Negative	1856
Neutral	1021

4.4 Data Modelling Evaluation

As mentioned in the methodology, the modeling techniques chosen in this study were NB, LR, SVM, and ensembled classifiers of the three models using Majority Voting. In this section, the performance of each technique will be observed and compared using the confusion matrix and classification results.

4.4.1 Naïve Bayes Classifier

**Fig. 8.** Confusion matrix for MNB**Table 8.** Classification report for MNB

	Precision	Recall	F1-score	Support
Negative	0.66	0.75	0.70	471
Neutral	0.79	0.37	0.51	249
Positive	0.63	0.72	0.66	476
Macro avg	0.69	0.62	0.63	1196
Weighted avg	0.68	0.66	0.65	1196

Fig. 8 shows the performance of the confusion matrix for the MNB model. This model has successfully predicted 344 values of words with neutral sentiment, 93 values of words with positive sentiment, and 353 of words with negative sentiments. In terms of precision, recall, and F1-score, the result can be observed in **Table 8**.

4.4.2 Support Vector Machine

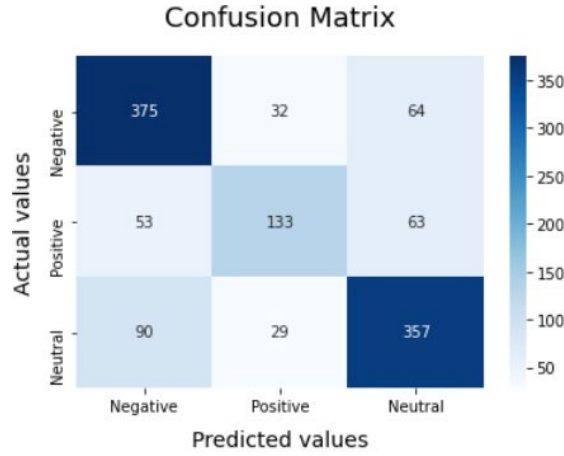


Fig. 9. Confusion matrix for SVM

Table 9. Classification report for SVM

	Precision	Recall	F1-score	Support
Negative	0.72	0.80	0.76	471
Neutral	0.69	0.53	0.60	249
Positive	0.74	0.75	0.74	476
Macro avg	0.72	0.69	0.70	1196
Weighted avg	0.72	0.72	0.72	1196

Fig. 9 shows the performance of the confusion matrix for the SVM model. This model has successfully predicted 357 values of words with neutral sentiment, 133 values of words with positive sentiment, and 375 words with negative sentiments. In terms of precision, recall, and F1-score, the result can be observed in **Table 9**.

4.4.3 Logistic Regression

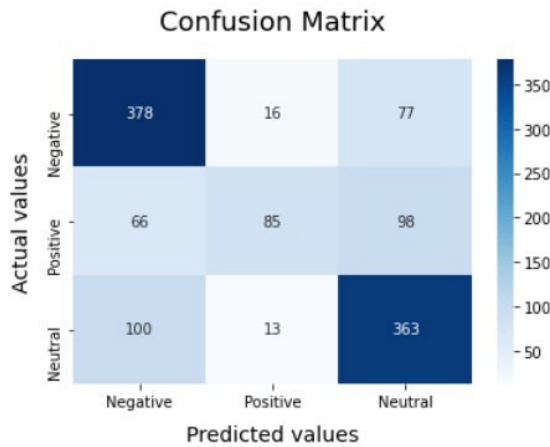


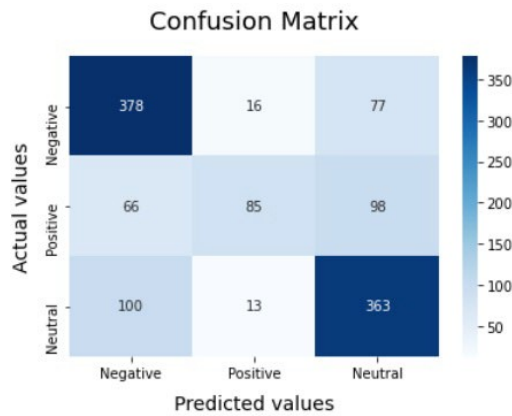
Fig. 10. Confusion matrix for LR

Table 10. Classification report for LR

	Precision	Recall	F1-score	Support
Negative	0.69	0.80	0.74	471
Neutral	0.75	0.34	0.47	249
Positive	0.67	0.76	0.72	476
Macro avg	0.71	0.64	0.64	1196
Weighted avg	0.70	0.69	0.68	1196

Fig. 10 shows the performance of the confusion matrix for the LR model. This model has successfully predicted 363 values of words with neutral sentiment, 85 values of words with positive sentiment, and 378 words with negative sentiments. In terms of precision, recall, and F1-score, the result can be observed in **Table 10**.

4.4.4 Majority Voting

**Fig. 11.** Confusion matrix for Ensembled Classification**Table 11.** Classification Report for Ensembled Classification

	Precision	Recall	F1-score	Support
Negative	0.74	0.80	0.77	471
Neutral	0.70	0.58	0.63	260
Positive	0.72	0.74	0.73	465
Macro avg	0.72	0.70	0.71	1196
Weighted avg	0.73	0.73	0.72	1196

Fig. 11 shows the performance of the confusion matrix for the ensembled classification model using the Majority Voting Classifier. This model has successfully predicted 339 values of words with neutral sentiment, 153 values of words with positive sentiment, and 373 of words with negative sentiments. In terms of precision, recall, and F1-score, the result can be observed in **Table 11**.

4.5 Comparison

Fig. 12 shows the summary of all classifiers in a bar plot. As observed, the ensembled classifier showed a slightly higher accuracy compared to SVM, which had the best performance among stand-alone classifiers.

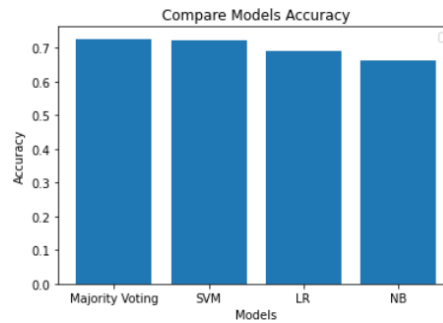


Fig. 12. Accuracy comparison among classifiers (bar plot)

Table 12. Accuracy comparison among classifiers (tabular)

	Majority Voting	SVM	LR	NB
Accuracy	72.66%	72.23%	69.06%	66.39%

As presented in **Table 12**, In the context of a majority voting ensemble model using three different classifiers (SVM, LR, NB), the results indicate the following accuracies:

- Majority Voting: 72.66%
- SVM (Support Vector Machine): 72.23%
- LR (Logistic Regression): 69.06%
- NB (Naive Bayes): 66.39%

Here's a brief interpretation of the results:

- **Majority Voting (72.66%):** The majority voting approach combines the predictions of SVM, LR, and NB. It achieved the highest accuracy among the individual classifiers, suggesting that the ensemble benefits from the diverse perspectives of these models. This is a common observation in ensemble learning where combining multiple models often results in improved performance.
- **SVM (72.23%):** SVM achieved a slightly lower accuracy than the majority voting ensemble. SVM is known for its effectiveness in high-dimensional spaces and works well when there's a clear margin of separation between classes. The accuracy suggests that SVM contributes significantly to the ensemble's overall performance.
- **LR (69.06%):** Logistic Regression, a linear model, achieved a slightly lower accuracy compared to SVM. LR is commonly used for binary classification problems but can be extended to multiclass problems. In this context, its performance is outperformed by the more complex SVM and the combined majority voting.
- **NB (66.39%):** Naive Bayes, often efficient for text classification tasks, achieved the lowest accuracy among the individual classifiers. Despite its simplicity and assumption of feature independence, NB can still be valuable, especially in scenarios where its assumptions align well with the data.

In summary, the majority voting ensemble outperformed individual classifiers, showcasing the strength of combining different models. It's essential to consider not only accuracy but also other metrics like precision, recall, and F1-score to have a comprehensive understanding of the model's performance. Additionally, the choice of the ensemble approach can have a significant impact, and further exploration into hyperparameter tuning or different ensemble techniques may provide insights for potential improvements.

5. Conclusion

This research aims to raise awareness of the mental health issues that are surging now. There are many issues with suicidal attempts, even suicides that have shown an increase due to many reasons, such as lockdown, unemployment, isolation, stress, and so on. Social media is used to talk about many issues, including this topic. Many users are more aware of the adverse effects of mental health that could lead them. Additionally, there are few studies on sentiment analysis on mental health issues during COVID-19 in Malaysia. Hence, this study is hopefully becoming a stepping stone for more studies to come.

There are several limitations identified during the experiment conduction. In Twitter, the location of the user can be empty, and there might be some users who might have come from Malaysia but were not able to be identified in the dataset due to the filtering criteria only including "Malaysia" and states in Malaysia. Similarly, if the user puts the location as a different name, such as "Hogwarts", then this user will be excluded as well. Moreover, Malaysia is a multilingual country, which means some of the tweets might be in Bahasa Malaysia, Mandarin, or Tamil. Since the author excluded other language than English, their tweets, which could improve the sentiment in the dataset, were filtered out. However, some tweets mixed between English and Bahasa Malaysia were found in the dataset.

For future research, a similar study can be done on another language than English, for example, focusing more on the national language of Malaysia. Malaysia is not just a multilingual country but also has many slang. Hence, in the future, a similar study that includes slang can also be done. Additionally, more models can be included in the hybrid model to improve overall accuracy. Moreover, the researcher could conduct more data cleaning on the location of the user to identify users who came from Malaysia. Lastly, since Prominent Features Set, Lexicon-based (Lex) and PoS features provided better support to Bag-of-Words (BoW), future research could also include this.

Funding statement

- This work was supported by the Universiti Malaya (Research University Grant - Program Faculty), under Project No: GPF098B-2020.

References

- [1] L. Ma, K. Song, and Y. Huang, "Coronavirus Disease-2019 (COVID-19) and Cardiovascular Complications," *Journal of Cardiothoracic and Vascular Anesthesia*, 35(6), 1860-1865, 2021. [Article \(CrossRef Link\)](#).
- [2] X. Dong et al., "Eleven faces of coronavirus disease 2019," *Allergy Eur. J. Allergy Clin. Immunol.*, 75(7), 1699-1709, 2020. [Article \(CrossRef Link\)](#).
- [3] A. S. Dousari, M. T. Moghadam, and N. Satarzadeh, "COVID-19 (Coronavirus disease 2019): A new coronavirus disease," *Infection and Drug Resistance*, 13, 2819-2828, 2020. [Article \(CrossRef Link\)](#).
- [4] "WHO Coronavirus Disease (COVID-19) Dashboard," *Bangladesh Physiother. J.*, 2020. [Article \(CrossRef Link\)](#).
- [5] WHO and W. (PRC) Aylward, Bruce (WHO); Liang, "Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19)," *WHO-China Jt. Mission Coronavirus Dis. 2019*, 2020.

- [6] World Health Organization, "WHO Coronavirus Disease 2019 (COVID-19) Dashboard," *WHO.int*, 2021.
- [7] F. Wu et al., "A new coronavirus associated with human respiratory disease in China," *Nature*, 579, 265-269, 2020. [Article \(CrossRef Link\)](#).
- [8] H. Ouassou et al., "The Pathogenesis of Coronavirus Disease 2019 (COVID-19): Evaluation and Prevention," *Journal of Immunology Research*, 2020. [Article \(CrossRef Link\)](#).
- [9] G. De Girolamo et al., "Mental health in the coronavirus disease 2019 emergency - The Italian response," *JAMA Psychiatry*, 77(9), 974-976, 2020. [Article \(CrossRef Link\)](#).
- [10] A. Vergara-Buenaventura, M. Chavez-Tuñón, and C. Castro-Ruiz, "The Mental Health Consequences of Coronavirus Disease 2019 Pandemic in Dentistry," *Disaster Med. Public Health Prep.*, 2020. [Article \(CrossRef Link\)](#).
- [11] G. H. Bahn, "Coronavirus disease 2019, school closures, and children's mental health," *J. Korean Acad. Child Adolesc. Psychiatry*, 14(6), e31 - e34, 2020. [Article \(CrossRef Link\)](#).
- [12] A. Beckstein, B. Rathakrishnan, P. B. Hutchings, and N. H. Mohamed, "THE COVID-19 PANDEMIC AND MENTAL HEALTH IN MALAYSIA: CURRENT TREATMENT AND FUTURE RECOMMENDATIONS," *Malaysian J. Public Heal. Med.*, 21(1), 260-267, 2021. [Article \(CrossRef Link\)](#).
- [13] L. P. Wong et al., "Escalating progression of mental health disorders during the COVID-19 pandemic: Evidence from a nationwide survey," *PLoS ONE*, 16(3), e0248916, 2021. [Article \(CrossRef Link\)](#).
- [14] L. Rosida, I. M. Putri, K. Komarudin, N. Fajarini, and E. K. Suryaningsih, "The domestic violence during the COVID-19 pandemic: Scoping review," *Open Access Maced. J. Med. Sci.*, 9(F), 660-667, 2021. [Article \(CrossRef Link\)](#).
- [15] N. Panchal et al., "The Implications of COVID-19 for Mental Health and Substance Use," *Kaiser Fam. Found.*, 2021.
- [16] A. Roberts et al., "Alcohol and other substance use during the COVID-19 pandemic: A systematic review," *Drug and Alcohol Dependence*, 229, 109150, 2021. [Article \(CrossRef Link\)](#).
- [17] D. Valdez, M. ten Thij, K. Bathina, L. A. Rutter, and J. Bollen, "Social media insights into US mental health during the COVID-19 pandemic: Longitudinal analysis of twitter data," *J. Med. Internet Res.*, 22(12), e21418, 2020. [Article \(CrossRef Link\)](#).
- [18] A. Amerio et al., "Covid-19 lockdown: Housing built environment's effects on mental health," *Int. J. Environ. Res. Public Health*, 17(16), 5973, 2020. [Article \(CrossRef Link\)](#).
- [19] A. N. Mat Ruzlin, X. W. Chen, R. M. Yunus, E. Z. Samsudin, M. I. Selamat, and Z. Ismail, "Promoting Mental Health During the COVID-19 Pandemic: A Hybrid, Innovative Approach in Malaysia," *Front. Public Heal.*, 9, 2021. [Article \(CrossRef Link\)](#).
- [20] A. K. Tay and S. Balasundaram, "Mental health services for refugees in Malaysia during the COVID-19 pandemic," *The Lancet Psychiatry*, 8(2), 2021. [Article \(CrossRef Link\)](#).
- [21] A. Salman, U. Kamekar, M. Jaafar, and D. Mohamad, "Empirical analysis of COVID-19 induced socio cognitive factors and its impact on residents of Penang Island," *Int. J. Tour. Cities*, 8(1), 210-222, 2022. [Article \(CrossRef Link\)](#).
- [22] P. L. Chui et al., "The covid-19 global pandemic and its impact on the mental health of nurses in malaysia," *Healthc*, 9(10), 1259, 2021. [Article \(CrossRef Link\)](#).
- [23] S. Andreadis et al., "A social media analytics platform visualising the spread of COVID-19 in Italy via exploitation of automatically geotagged tweets," *Online Soc. Networks Media*, 23, 2021. [Article \(CrossRef Link\)](#).
- [24] G. Barkur, Vibha, and G. B. Kamath, "Sentiment analysis of nationwide lockdown due to COVID 19 outbreak: Evidence from India," *Asian Journal of Psychiatry*, 51, 2020. [Article \(CrossRef Link\)](#).
- [25] S. V. Praveen, R. Ittamalla, and G. Deepak, "Analyzing Indian general public's perspective on anxiety, stress and trauma during Covid-19 - A machine learning study of 840,000 tweets," *Diabetes Metab. Syndr. Clin. Res. Rev.*, 15(3), 667-671, 2021. [Article \(CrossRef Link\)](#).
- [26] M. Y. Kabir and S. Madria, "EMOCOV: Machine learning for emotion detection, analysis and visualization using COVID-19 tweets," *Online Soc. Networks Media*, 23, 2021. [Article \(CrossRef Link\)](#).

- [27] J. Xue et al., "Twitter discussions and emotions about the COVID-19 pandemic: Machine learning approach," *J. Med. Internet Res.*, 22(11), 2020. [Article \(CrossRef Link\)](#).
- [28] J. Talbot, V. Charron, and A. T. Konkle, "Feeling the void: Lack of support for isolation and sleep difficulties in pregnant women during the covid-19 pandemic revealed by twitter data analysis," *Int. J. Environ. Res. Public Health*, 18(2), 2021. [Article \(CrossRef Link\)](#).
- [29] J. X. Koh and T. M. Liew, "How loneliness is talked about in social media during COVID-19 pandemic: Text mining of 4,492 Twitter feeds," *J. Psychiatr. Res.*, 145, 317-324, 2022. [Article \(CrossRef Link\)](#).
- [30] D. Gruda and S. Hasan, "Feeling anxious? Perceiving anxiety in tweets using machine learning," *Comput. Human Behav.*, 98, 245-255, 2019. [Article \(CrossRef Link\)](#).
- [31] S. Elbagir and J. Yang, "Sentiment Analysis on Twitter with Python's Natural Language Toolkit and VADER Sentiment Analyzer," *IAENG Transactions on Engineering Sciences*, pp. 63-80, 2020. [Article \(CrossRef Link\)](#).
- [32] L. Alsudias and P. Rayson, "Social media monitoring of the COVID-19 pandemic and influenza epidemic with adaptation for informal language in Arabic twitter data: Qualitative study," *JMIR Med. Informatics*, 9(9), 2021. [Article \(CrossRef Link\)](#).
- [33] C. R. Machuca, C. Gallardo, and R. M. Toasa, "Twitter sentiment analysis on coronavirus: Machine learning approach," *J. Phys.: Conf. Ser.*, 1828, 2021. [Article \(CrossRef Link\)](#).
- [34] M. Deshpande and V. Rao, "Depression detection using emotion artificial intelligence," in *Proc. of 2017 International Conference on Intelligent Sustainable Systems (ICISS)*, 2017. [Article \(CrossRef Link\)](#).
- [35] Imamah and F. H. Rachman, "Twitter sentiment analysis of Covid-19 using term weighting TF-IDF and logistic regresion," in *Proc. of 2020 6th Information Technology International Seminar (ITIS)*, 2020. [Article \(CrossRef Link\)](#).
- [36] O. Bharti and M. M. Malhotra, "International Journal of Computer Science and Mobile Computing SENTIMENT ANALYSIS ON TWITTER DATA," 2016.
- [37] K. Garcia and L. Berton, "Topic detection and sentiment analysis in Twitter content related to COVID-19 from Brazil and the USA," *Appl. Soft Comput.*, 101, 2021. [Article \(CrossRef Link\)](#).
- [38] S. T. Rabani, Q. R. Khan, and A. M. Ud Din Khanday, "Detection of suicidal ideation on Twitter using machine learning & ensemble approaches," *Baghdad Sci. J.*, 17(4), 2020. [Article \(CrossRef Link\)](#).
- [39] K. Saha, J. Torous, S. K. Ernala, C. Rizuto, A. Stafford, and M. De Choudhury, "A computational study of mental health awareness campaigns on social media," *Transl. Behav. Med.*, 9(6), 1197-1207, 2019. [Article \(CrossRef Link\)](#).
- [40] C. Shofiya and S. Abidi, "Sentiment analysis on covid-19-related social distancing in Canada using twitter data," *Int. J. Environ. Res. Public Health*, 18(11), 2021. [Article \(CrossRef Link\)](#).
- [41] Ankit and N. Saleena, "An Ensemble Classification System for Twitter Sentiment Analysis," *Procedia Computer Science*, 132, 937-946, 2018. [Article \(CrossRef Link\)](#).
- [42] A. K. Abbas, A. K. Salih, H. A. Hussein, Q. M. Hussein, and S. A. Abdulwahhab, "Twitter Sentiment Analysis Using an Ensemble Majority Vote Classifier," *J. Southwest Jiaotong Univ.*, 2020. [Article \(CrossRef Link\)](#).
- [43] S. Sangam and S. Shinde, "Sentiment classification of social media reviews using an ensemble classifier," *Indones. J. Electr. Eng. Comput. Sci.*, 16(1), 355-363, 2019. [Article \(CrossRef Link\)](#).
- [44] E. P. - and K. S. -, "New Ensemble Approach to Analyze User Sentiments from Social Media Twitter Data," *SIJ Trans. Ind. Financ. Bus. Manag.*, 2018. [Article \(CrossRef Link\)](#).
- [45] S. Sharma and A. Jain, "Hybrid Ensemble Learning With Feature Selection for Sentiment Classification in Social Media," *Int. J. Inf. Retr. Res.*, 10(2), 2020. [Article \(CrossRef Link\)](#).
- [46] D. Tiwari and N. Singh, "Ensemble Approach for Twitter Sentiment Analysis," *Int. J. Inf. Technol. Comput. Sci.*, 11(8), 20-26, 2019. [Article \(CrossRef Link\)](#).
- [47] M. M. Fouad, T. F. Gharib, and A. S. Mashat, "Efficient Twitter Sentiment Analysis System with Feature Selection and Classifier Ensemble," in *Proc. of The International Conference on Advanced Machine Learning Technologies and Applications (AMLT2018)*, 516-527, 2018. [Article \(CrossRef Link\)](#).

- [48] R. Wijayanti and A. Arisal, "Ensemble approach for sentiment polarity analysis in user-generated Indonesian text," in *Proc. of 2017 International Conference on Computer, Control, Informatics and its Applications (IC3INA)*, 2017. [Article \(CrossRef Link\)](#).
- [49] L. Singh, P. Gupta, R. Katarya, and P. Jayvant, "Twitter data in emotional analysis - A study," in *Proc. of 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, 2020. [Article \(CrossRef Link\)](#).
- [50] F. M. Moy and Y. H. Ng, "Perception towards E-learning and COVID-19 on the mental health status of university students in Malaysia," *Sci. Prog.*, 104(3), 1-18, 2021. [Article \(CrossRef Link\)](#).
- [51] E. A. Othman, M. Mohamad, and V. Giampietro, "Patient Satisfaction With Teleconsultation During Covid-19 Pandemic: A Descriptive Study For Mental Health Care In Malaysia," *Malaysian J. Public Heal. Med.*, 21(2), 243-251, 2021. [Article \(CrossRef Link\)](#).
- [52] A. Azman, P. S. J. Singh, J. Parker, and S. Ashencaen Crabtree, "Addressing competency requirements of social work students during the COVID-19 pandemic in Malaysia," *Soc. Work Educ.*, 39(8), 1058-1065, 2020. [Article \(CrossRef Link\)](#).
- [53] A. Perveen, S. Motevalli, H. Hamzah, F. Ramlee, S. M. Olagoke, and A. Othman, "The Comparison of Depression, Anxiety, Stress, and Coping Strategies among Malaysian Male and Female During COVID-19 Movement Control Period," *Int. J. Acad. Res. Bus. Soc. Sci.*, 487-496, 2020. [Article \(CrossRef Link\)](#).



Nur Aisyah Zakaria Adli received the bachelor's degree in business information system from Universiti Teknologi PETRONAS, Malaysia in 2017 and the master's degree in Data Science from University of Malaya (UM), Malaysia in 2021. She is currently working as a Senior Consultant at an Australian start-up firm and was previously a Consultant and Data Analytics Specialist at one of the top Technology Consulting firms worldwide.



Norjihan Abdul Ghani received the Bachelor's Degree in Information Technology from Universiti Utara Malaysia, in 2000, the Master's Degree in Information Technology from Universiti Kebangsaan Malaysia, in 2000, and the Ph.D. Degree in Computer Science from Universiti Teknologi Malaysia, in 2013. She is currently an Associate Professor with the Department of Information Systems, FCSIT. She has managed to secure various research grants from internal and external funds as a principal investigator and a co-researcher. She has established collaboration with national and international collaborators and also published in many ISI- and Scopus-indexed journals. Her research interests include the areas of information security where she focuses more on information privacy, data security, data protection, and access control which now she extended it in smart-based applications.



Muneer Ahmad completed his PhD in Computer Science from Universiti Teknologi PETRONAS, Malaysia in 2014. He has 18 years of teaching, research and administrative experience internationally. Dr. Muneer Ahmad has authored numerous research papers in refereed research journals, international conferences and books. Further, he successfully completed several funded research projects. His areas of interest include Data science, big data analysis, machine learning, bioinformatics and medical informatics.



Azah Anir Norman received the bachelor's degree in computer science from Universiti Kebangsaan Malaysia (UKM), the master's degree in electronic commerce security from the Royal Holloway University of London, U.K., in 2004, and the Ph.D. degree from the University of Malaya (UM), Malaysia, in 2014. Before joining the academic field, she worked with MSC Trustgate.com (an MDEC subsidiary in collaboration with VeriSign Inc.) and was a consultant to Top-500 companies for internet security implementation for their businesses. She is currently an Associate Professor with the Department of Information Systems, Faculty of Computer Science and Information Technology, UM. She is also active in the working group WG/G/5- 1 Information Security Management System (ISMS), Standard Department Malaysia, and in the International Standard Organization (ISO) as an Expert. Her works were published in respectable ISI-cited journals and Scopus-cited journals. Her research interest includes information systems focusing on information systems security. She is involved in various research, leading to the publication of many academic papers in the areas of information security governance, information security management, information security systems, information security and trust, information security and privacy, information security education awareness, and information security assurance, and information security policy and governance. Dr. Norman is also a member of the Association of Information Systems (AIS). She won a prestigious award from the Royal Academy of Engineering of U.K. (RAENG) and the Leaders of Innovation, in 2018.



Sri Devi Ravana received the bachelor's degree in information technology from the National University of Malaysia, in 2000, the master's degree in software engineering from the University of Malaya (UM), Malaysia, in 2001, and the Ph.D. degree from The University of Melbourne, Australia, in 2012. She is currently an Associate Professor with the Department of Information Systems, UM. She has actively collaborated with researchers from many countries, including The University of Auckland, The University of Melbourne, University of Leeds, The University of Hong Kong, and the National University of Singapore. Her research interests include information retrieval heuristics, text indexing, and data analytics. She received a couple of best paper awards in international conferences within the area of information retrieval. She was a recipient of the Competitive AUA Scholars Award, in 2019 and 2020. She also serves as an editorial board member for Scopus/WoS indexed journals.