PCMJ

Progress in Computer and Mathematics Journal (PCMJ)
volume 3 [November, 2025]
e-ISSN: 3030-6728
Website: fskmjebat.uitm.edu.my/pcmj

SENTIMENT ANALYSIS AND VISUALIZATION OF PADU FROM MALAYSIAN X USERS USING BERT

AHMAD ISHRAF IMRAN MOHD HOSNI
Faculty of Computing and Mathematical Sciences, Univeristi Teknologi MARA
2023185765@student.uimt.edu.my

JAMALUDDIN JASMIS*
Faculty of Computing and Mathematical Sciences, Univeristi Teknologi MARA
jamaluddinjasmis@uitm.edu.my

Article Info

Received: March 2025
Accepted: September 2025
Available Online: November 2025

Abstract

On the surface, the introduction of PADU might be met with varying degrees of acceptance with Malaysians but knowing the actual sentiment without any biases is hard. Sentiment analysis of a certain topic, which in this study is PADU is a complex field that involves scraping datasets and classifying them with great accuracy where if one were to do it manually, would inevitably introduce some sort of bias to the results. The project provides a solution to the matter by developing a sentiment analysis model and appropriately visualising the data and results. The dataset used is scraped from X using Tweet Harvest which consists of 88 datapoints which were further augmented to 440 datapoints. The model is developed using bidirectional encoder representations from transformers that are trained with the dataset gathered. The model follows the software development methodology using waterfall and is released on a web platform. The result of the model that was trained with the combination of collected and augmented datasets showed 87% accuracy, 87% Precision, 87% Recall and F1-score of 87% compared with the model that was trained using only the collected dataset. In the future, further improvement to this project will be seen in the form of bigger language support for the model and the collection of data from a wide variety of social media platforms.

Keywords: PADU, Sentiment Analysis, BERT, Data Augmentation

INTRODUCTION (HEADING 1)

Pengkalan Data Utama more commonly known as PADU is an initiative in Malaysia that aims to enhance government services and decision-making. PADU serves as a centralised database hub that integrates information from various government departments and agencies. Its primary purpose is to provide a clearer picture of national household income. The government intends to use PADU for effective subsidy distribution and informed policy-making.

Given that PADU is a recently launched government initiative, there is a lack of an unbiased summarisation of the public perception towards PADU. This is because if one were to research the perception towards PADU on social media, the result would be skewed following the researcher's personal opinion on the matter. Furthermore, (Papakyriakopoulos et al., 2020) stated that hyperactive users can become opinion leaders and have an agenda-setting effect, thus creating an alternate picture of public opinion.

The issue of acquiring an unbiased view of the public perception towards PADU from social media such as X is a difficult task as any person who tries to find a conclusive verdict on the fact might find that their findings will differ compared to others. This is because social media may limit exposure to alternative ideas and favour the establishment of communities of like-minded individuals creating and reinforcing a common narrative as stated in (Cinelli et al., 2021). Through the use of sentiment analysis, the problem of acquiring an unbiased view towards PADU can be overcome.

The challenge inherent in sentiment analysis using social media posts lies in the intricate task of comprehending and contextualising data that often includes internet slang, abbreviations, and unconventional language. This complexity is amplified by the necessity to accurately detect and interpret the Malay language, especially when analysing posts made by Malaysians who predominantly express themselves online in Malay as found in (Nabiha et al., 2021). These inherent difficulties significantly impact the accuracy of sentiment analysis, as nuances play a pivotal role in understanding the true emotional tone behind each message.

LITERATURE REVIEW

Pengkalan Data Utama (PADU) is a comprehensive and secure central database containing individual and household profiles of Malaysian citizens and permanent residents. Developed by the Ministry of Economy in collaboration with the Department of Statistics Malaysia (DOSM) and the Malaysian Administrative Modernisation and Management

Planning Unit (MAMPU) and was launched on 2 Jan 2024 by Prime Minister Datuk Seri Anwar Ibrahim.

PADU aims to provide near real-time data for digitalization and periodic analytics, supporting data-driven policymaking and decision-making by integrating data from over 400 government agencies, offering a comprehensive overview of household income nationwide in order to increase the effectiveness of policy planning and resource allocation, particularly the distribution of government subsidies and support to those who need it. It also acts as a centralised and digitised reference source for the government's data-driven choices.

Sentiment Analysis

According to (Pang & Lee, 2008) "Sentiment analysis, also called opinion mining, is the field of study that analyzes people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes." It involves using natural language processing, text analysis and computational linguistics to systematically identify, extract, quantify, and study subjective information. Sentiment analysis works by evaluating the tone of a piece of text. The goal is to classify the text as either positive, negative, or neutral. The process begins with data collection, where large volumes of text are gathered from different sources. Then, through machine learning algorithms and linguistic rules, the system identifies patterns and markers that signify sentiment.

Bidirectional Encoder Representations from Transformers

Bidirectional Encoder Representations from Transformers or BERT is a text embedding model published by Google in 2018. As stated by (Ravichandiran, 2021), the major difference in BERT compared to other classification algorithms is that it understands the meaning of a word based on its surrounding context. This results in a more nuanced understanding of language, significantly improving the performance of various NLP tasks such as sentiment analysis.

As found in (Kumar & Gawade, 2023) BERT was able to outperform other sentiment analysis algorithms such as VADER, LSTM and Textblobwith a substantially higher accuracy of 93.4%. This goes to show that BERT is a powerful algorithm for sentiment analysis.

PCMJ

Progress in Computer and Mathematics Journal (PCMJ)
volume 3 [November, 2025]
e-ISSN: 3030-6728
Website: fskmjebat.uitm.edu.my/pcmj

METHODOLOGY

The PADU Insight website is developed using the Waterfall methodology along with BERT as it machine learning algorithm for sentiment analysis.

Waterfall

The methodology that will be used in this project is the Waterfall model. The Waterfall model is a sequential software development methodology that follows a linear progression through distinct phases. Figure 1 shown below shows the phases that is in Waterfall model. Due to the nature of this project, this model is adjusted to exclude the maintenance phase. This method is chosen due to its clear and well-defined requirements, predictable timelines and thorough documentation.

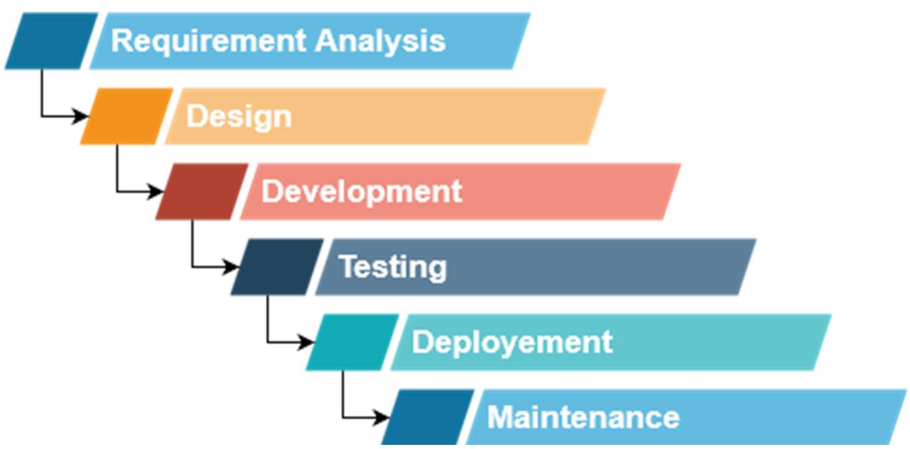

Figure 1: Illustration of Waterfall model phases

This model is particularly beneficial for projects with clear, well-defined requirements. It operates on the premise that each phase: requirements gathering, system design, implementation, and testing must be completed before the next phase can begin. This structured approach ensures that each phase's goals are met before moving forward, which can be particularly useful when there is little to no ambiguity in the project's requirements. Furthermore, due to its orderly nature, the waterfall model allows the establishment of a clear schedule based on the estimated duration of each phase thus making it easier to forecast timelines, costs and deliverables. To build on this, the model emphasizes the thorough documentation of each phase which provides a clear, step-by-step guide to project completion.

Bidirectional Encoder Representations from Transformers

The methodology of employing BERT for natural language understanding and sentiment analysis, demonstrating its effectiveness in interpreting the nuances and context of social media communication will be examined. The aim is to gain valuable insights into public opinion and sentiment trends as reflected in PADU's Twitter interactions.

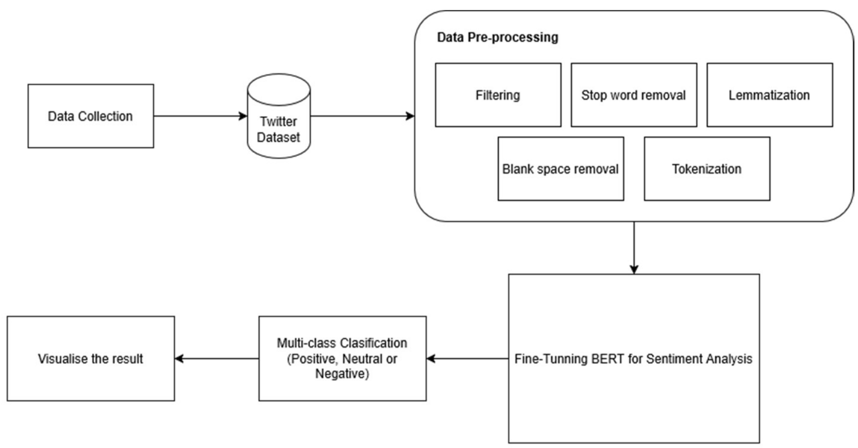

Figure 2: Illustration of the model diagram used for sentiment analysis

Based on Figure 2 above, in the training phase, tweets will be scraped from X using Selenium to serve as the training dataset. This dataset will undergo pre-processing and embeddings to prepare it to be sent to the BERT model to pre-train the model to understand the language. This pre-trained BERT model will then be used for sentiment analysis by fine-tuning it for sentiment analysis. This fine-tuned BERT model then will receive the testing dataset which is also scraped from X that is related to PADU which also undergoes pre-processing. The output of these processes will then be visualised and summarised.

RESULT AND DISCUSSION

The model and system was tested in terms of its functionality, usability and its accuracy. This was done through functionality and usability testing as well as a comparative analysis of the accuracy achieved through the different dataset.

Functionality Result

Functionality testing is a type of software testing that verifies whether a software application performs and functions according to its defined requirements and specifications. The goal is to ensure that every feature and function of the application works correctly and meets the end-user's expectations. Table 1 below shows the result of the testing.

Table 1 Functionality testing result

Test Case	Expected Results	Test Result
To test whether a user can view the dashboard	The user can view the dashboard and it is shown dynamically	The user successfully viewed the dashboard
To test whether a user can view the graph page	The user can view the graph page and it is shown dynamically	The user successfully viewed the graph page
To test whether a user can view the dataset download page	The user can view the dataset download page and it is shown dynamically	The user successfully viewed the dataset download page
To test whether a user can download the dataset	The user can download the desired dataset by clicking the respective button	The user successfully downloads the desired dataset

Usability Test Result

Usability testing is a type of software testing that evaluates how easy and user-friendly an application is for its intended users. It focuses on the overall user experience (UX), including navigation, accessibility, design, and satisfaction, ensuring that the application meets the needs of its target audience. In this study, the System Usability Scale (SUS) is used which was done through a ten-question survey as shown on Table 2 below.

Table 2: System Usability Scale Questions

No.	SUS Questionnaire
1.	I think that I would like to use this system frequently.
2.	I found the system unnecessarily complex.
3.	I thought the system was easy to use.
4.	I think that I would need the support of a technical person to be able to use this system.
5.	I found the various functions in this system were well integrated.
6.	I thought there was too much inconsistency in this system.
7.	I would imagine that most people would learn to use this system very quickly.
8.	I found the system very cumbersome to use.
9.	I felt very confident using the system.
10.	I needed to learn a lot of things before I could get going with this system.

The System Usability Scale (SUS) is calculated by first adjusting the scores from each of the 10 questions. For odd-numbered questions, subtract 1 from the user's rating, and for even-numbered questions, subtract the rating from 5. These adjusted scores are then summed and multiplied by 2.5, resulting in a final usability score out of 100.

Once all individual SUS scores are collected from multiple participants, the mean SUS score is calculated to determine the overall usability of the system. A higher score indicates better usability, while a lower score suggests usability issues.

PCMJ

Progress in Computer and Mathematics Journal (PCMJ)
volume 3 [November, 2025]
e-ISSN: 3030-6728
Website: fskmjebat.uitm.edu.my/pcmj

Table 3: System Usability Scale Score Result

Participant	Age	Gender	Experience with similar websites	Question										Raw SUS Score	Final SUS Score
				1	2	3	4	5	6	7	8	9	10		
1	24	Male	Some	4	5	4	2	4	2	5	2	4	2	28	70
2	24	Male	Some	4	1	4	1	4	1	5	1	4	3	34	85
3	20	Male	Expert	5	5	5	1	5	1	5	1	5	1	36	90
4	23	Male	None	5	1	5	1	5	1	5	1	5	1	40	100
5	23	Male	None	5	1	5	1	5	1	5	1	5	1	40	100
6	23	Prefer Not To Say	None	5	1	5	1	5	1	5	1	4	1	39	97.5
7	18	Male	None	5	1	5	1	5	1	5	1	5	1	40	100
8	23	Male	Some	4	1	5	1	5	2	5	1	4	1	36	90
9	22	Male	Some	4	2	4	4	4	1	5	1	4	1	32	80
10	23	Male	None	5	1	4	2	5	1	5	2	5	2	36	90
Average															90.25

Based on Table 3 which shows the SUS score result of each participant that ranges from 70 to 100. The background of these participants are predominantly Male with them being in their early 20s. The final SUS score obtained was 90.25 which is an A grade based.

Accuracy Result

The accuracy of the model will be tested by looking at how accurately the model can predict the sentiment of each sentence. The accuracy result that will be outline in this section will include the result from different datasets which are the original dataset, the augmented dataset and the original dataset combined with the augmented dataset.

The model will be evaluated through a couple of metrics such as Precision, Recall and F1-Score. These metrics help in measuring how well the model distinguishes between classes. Precision measures how many of the predicted positive cases were actually correct which is useful when false positives are costly. Recall measures how many actual positive cases were correctly identified. It is crucial when false negatives are costly.F1-Score balances precision and recall which is useful when both false positives and false negatives are important.

Table 4: The overall F1 and Accuracy score of each dataset

Dataset	F1-Score	Accuracy Score
Original	0.44	0.44
Augmented	0.84	0.84
Augmented + Original	0.87	0.87

Based on the Table 4 shown above, among all the dataset that were used to train the model, the Augmented in combination with the Original dataset produce the highest F1 and Accuracy score among all the other dataset with 87%. This result is satisfactory considering the limitation encountered during the development of the model such as a limited dataset and an over reliant on augmented data.

CONCLUSION

To conlcude the research developed a robust sentiment analysis model that leverages the powerful capabilities of the bert-base-multilingual-uncased model. This pre-trained model was meticulously fine-tuned using a carefully curated dataset of tweets related to PADU, ultimately achieving an impressive accuracy of 87%. The comprehensive system development process involved multiple critical stages including data collection, rigorous pre-processing, detailed annotation, thorough analysis, strategic dataset splitting, methodical model training and extensive evaluation.

Progress in Computer and Mathematics Journal (PCMJ)
volume 3 [November, 2025]
e-ISSN: 3030-6728
Website: fskmjebat.uitm.edu.my/pcmj

REFERENCES

Papakyriakopoulos, O., Serrano, J. C. M., & Hegelich, S. (2019). Political communication on social media: A tale of hyperactive users and bias in recommender systems. *Online Social Networks and Media*, *15*, 100058. https://doi.org/10.1016/j.osnem.2019.100058

Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, *118*(9). https://doi.org/10.1073/pnas.2023301118

Nabiha, A., Mutalib, S., & Malik, A. M. A. (2021). Sentiment Analysis for Informal Malay Text in Social Commerce. *2021 2nd International Conference on Artificial Intelligence and Data Sciences, AiDAS 2021.*, *9*, 1–6. https://doi.org/10.1109/aidas53897.2021.9574436

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, *2*(1–2), 1–135. https://doi.org/10.1561/1500000011

Ravichandiran, S. (2021). Getting Started with Google BERT: Build and train state-of-the-art natural language processing models using BERT. In *Packt Publishing*.

Kumar, S., & Gawade, S. (2023). Sentiment Analysis of opinions about Electric Vehicles Using Twitter data. *2022 IEEE 6th Conference on Information and Communication Technology (CICT)*, 1–5. https://doi.org/10.1109/cict59886.2023.10455654