

CONSUMER ACCEPTANCE AND PERCEPTIONS OF ELECTRIC VEHICLES IN MALAYSIA USING SENTIMENT ANALYSIS (Penerimaan dan Persepsi Pengguna Terhadap Kenderaan Elektrik di Malaysia Menggunakan Analisis Sentimen)

NOR FARAWAHIDA ABDULLAH & NUR HAIZUM ABD RAHMAN*

ABSTRACT

Despite the excitement around electric automobiles, Electric Vehicles (EVs) in Malaysia are still relatively new compared to more advanced markets such as European countries. To understand Malaysian acceptance and perception of EVs, this study uses sentiment analysis to ascertain whether Malaysians view and utilise EVs positively or negatively. The dataset is scrapped data collected from comments on YouTube with Term Frequency-Inverse Document Frequency (TF-IDF) as the feature extraction method. The machine learning algorithm, support vector machine (SVM), is then created to automatically identify and classify comments about EV acceptance and perception. Key performance indicators, including accuracy, precision, and recall percentages, are used to assess the algorithm's performance. Results reveal that for the sentiment distribution, there is 55% positive and only 14% negative sentiment, while the remaining is neutral. The main concerns identified are battery, charging station, and price. The classification was further analysed, and the result suggests that SVM with TF-IDF feature extraction offered 85.18% accuracy in classifying the sentiment. The result of this study provides nuanced insights into public opinions regarding EVs in Malaysia by seamlessly integrating machine learning and sentiment analysis. The findings will have substantial implications for researchers in the sustainable transportation sector, policymakers, and industry stakeholders.

Keywords: electric vehicles; sentiment analysis; classification; machine learning; perceptions

ABSTRAK

Di sebalik keterujaan mengenai kereta elektrik, Kenderaan Elektrik (EV) di Malaysia masih agak baharu berbanding pasaran yang lebih maju seperti negara Eropah. Bagi memahami penerimaan dan persepsi Malaysia terhadap EV, kajian ini menggunakan analisis sentimen untuk memastikan sama ada rakyat Malaysia melihat dan menggunakan EV secara positif atau negatif. Set data adalah data yang dikikis daripada ulasan di YouTube dengan *Term Frequency-Inverse Document Frequency* (TF-IDF) sebagai kaedah pengekstrakan ciri. Algoritma pembelajaran mesin, Mesin Vektor Sokongan (SVM) kemudiannya dibangunkan untuk mengenal pasti dan mengklasifikasikan ulasan secara automatik tentang penerimaan dan persepsi EV. Penunjuk prestasi utama, termasuk kejituan, kepersisan dan peratusan imbas kembali, digunakan untuk menilai prestasi algoritma. Keputusan menunjukkan bahawa untuk taburan sentimen, terdapat 55% positif dan hanya 14% sentimen negatif, manakala selebihnya adalah neutral. Kebimbangan utama yang dikenal pasti ialah bateri, stesen pengecasan dan harga. Klasifikasi telah dianalisis selanjutnya, dan hasilnya menunjukkan bahawa SVM dengan pengekstrakan ciri TF-IDF menawarkan ketepatan 85.18% dalam mengklasifikasikan sentimen. Hasil kajian ini memberikan pandangan yang lebih mendalam tentang pendapat umum mengenai EV di Malaysia dengan mengintegrasikan pembelajaran mesin dan analisis sentimen dengan lancar. Penemuan ini akan mempunyai implikasi yang besar kepada penyelidik dalam sektor pengangkutan mampan, penggubal dasar dan pihak berkepentingan industri.

Kata kunci: kenderaan elektrik; analisis sentiment; klasifikasi; pembelajaran mesin; persepsi

1. Introduction

These few decades have seen a rapid increase in the global economy and technology, greatly benefiting human civilisation. However, with the fast pace of changes, where technological developments are increasing dramatically, technology is taking us farther down the path of disaster in the shape of resource constraints and climate change. Notably, resource constraints and climate change have become the main environmental concerns of the new millennium world. These issues endanger the health of our planet and present complex risks to international peace, socioeconomic institutions, and local livelihoods. As a result of numerous human activities that continuously and uncontrollably release hazardous air pollutants into the atmosphere, the world is also experiencing problems such as air pollution and greenhouse gas emissions. In particular, fossil fuel combustion is one of the leading causes of this issue (Asadi *et al.* 2021). Consequently, throughout the years, humanity has attempted to devise a number of solutions to the problems pertaining to preserving the environment and its sustainable growth (Ramachandaran *et al.* 2023).

As the world embraces a green transportation system to reduce carbon emissions, it has introduced an alternative option, i.e., replacing traditional vehicles with new energy vehicles, such as Electric Vehicles (EVs). Thus, it is necessary to make the switch from traditional Internal Combustion Vehicles (ICVs), which run on fossil fuels such as petrol and diesel, to electric mobility, such as electric cars, EVs, in order to reduce the number of carbon emissions every day (Guo & Kontou 2021). Past research has mostly agreed that replacing gasoline-based automobiles with green transportation can significantly impact the ecosystem and shift the whole automobile industry to a better future. Attributable to the latest technology, the deployment of EVs has drawn attention and allowed EVs to become an essential part of future cars (Buekers *et al.* 2014).

During 2021, 6.5 million EVs were sold globally. Impressive sales patterns were observed, with a 100% rise from the previous year. About 50% of the market was held by mainland China, which has the largest share. Notably, government regulations, the Tesla effect, declining battery prices, 5G rollouts, and the introduction of the Battery-as-a-Service model are also the primary factors propelling the EV market (Bryła *et al.* 2023). Consumer interest in EVs has grown significantly as a result of their greater range, longer battery life, greater efficiency, and lower cost. From 2021 to 2027, the size of the global EV market is expected to expand at a Compound Annual Growth Rate (CAGR) of about 20%, or more than four times, to reach around USD 1.4 trillion.

Due to its reliance on energy sources, Malaysia is pressured to increase energy productivity and work towards lowering its carbon emissions (Adnan *et al.* 2018). Notably, Malaysia's market for EVs is still modest but expanding quickly. The number of EVs sold in 2021 was 2,717, as reported by the Malaysian Automotive Association (MAA), a notable increase from 1,642 units sold in 2020. With just 0.4% of all vehicle sales in Malaysia, the EV market share is relatively small compared to Norway, which is leading in the EV sector (Ritcher 2023). Nonetheless, the industry is anticipated to expand in the upcoming years due to the government's objective of having 125,000 EVs on the road by 2030 (International Trade Administration 2023).

As EVs are still increasing in Malaysia, public acceptability and spread of this topic are relatively new and vary. Due to that, minimal research has been conducted on this acceptance of contemporary and novel technology (Sang & Bekhet 2015; Adnan *et al.* 2016). The transition to electric mobility signifies a paradigm shift in the Malaysian automotive sector, shaped mainly by traditional internal combustion engine automobiles. In developing nations, especially Malaysia, where the government has recognised the advantages of adopting EVs, various steps

and measures have been taken to encourage their use (Adnan *et al.* 2017). Accordingly, it appears that customer choice and perceptions, which mostly stay in the domain of personal choice, are the driving force behind the acceptance of EVs in Malaysia (Liu *et al.* 2019).

Despite efforts to reduce the problem of environmental deterioration, the number of EVs on the road is still limited. The underdeveloped battery technology of EVs is one of the numerous barriers that must be overcome before EVs are widely used. Due to their high initial cost, long charging times, and restricted range, EVs are less desirable to the average consumer (Capuder *et al.* 2020). Furthermore, the limited number of available infrastructures for charging EVs is another significant barrier to widespread acceptance. As a result, the government has highlighted or stressed the various financial incentives and benefits it will provide to encourage the use of EVs (Muzir *et al.* 2022). Additionally, since EVs have a limited range of moves, some consumers could be concerned about running out of electricity. It has been proven that many drivers avoid EVs unless the infrastructure for charging them is stable and established. Thus, even though producing and using green transportation endorsements helps lower domestic petrol usage, an eco-friendly car's performance is marginally inferior to that of a regular car. Due to these drawbacks, consumers may think EVs are subpar.

This indicates that EVs in other countries are experiencing rapid acceptance and market expansion due to favourable conditions and supportive measures. In contrast, the perception and acceptance of EVs in Malaysia are progressing at a more gradual pace (Muzir *et al.* 2022). Hence, this paper examines Malaysian acceptance and perception of EVs using sentiment analysis as a measurement tool. By analysing YouTube comments, the study aims to determine whether Malaysians view EVs positively, negatively, or neutrally. The underlying hypothesis is that sentiment expressed in online discussions reflects public perception and acceptance of EVs. The sentiment classification results serve as a quantitative measure of public sentiment, providing insights into the factors influencing EV adoption in Malaysia.

2. Materials and Methods

2.1. Dataset

This study is based on primary data, which is scraped data that has been collected from comments in YouTube videos. The videos were selected based on the EVs-related keywords, which specifically had "Electric Vehicles," "EV Acceptance in Malaysia," "Tesla Malaysia," or "Electric Car" keywords in the title. Based on the keywords, the obtained videos include reviews, promotional videos, or discussions about EVs in Malaysia. The scrapping tool is the instant data scrapper, an extension of Google to obtain the data. Most of the video comments scraped are published about one year and below. The collected data consists of 389 observations with four variables named link, user, date, and comment after renaming to a readable name. It has 389 total observations in the dataset. After applying the Synthetic Minority Oversampling Technique (SMOTE) for resampling to balance the data, the total number of observations increased to 636. Consequently, the data is split into training and testing sets with an 80:20 ratio, which provides sufficient data for the training and reduces the variance in the model.

The sentiment classification process utilised Label Encoding, where categorical sentiment labels such as positive, neutral, and negative were converted into numerical values for model training. This process involved extracting the 'sentiment' column from the dataset, applying LabelEncoder from *sklearn.preprocessing* to transform the categorical labels into numeric values, and replacing the original sentiment labels with the encoded values. The mapping between the original labels and encoded values was verified to ensure correctness. The encoding followed a structured format where positive was assigned the value of 2, neutral was assigned

as 1, and negative was assigned as 0. This systematic approach ensured that sentiment classification was formatted appropriately for machine learning models, facilitating efficient training and evaluation.

2.2. Sentiment analysis approach

Sentiment analysis, also known as emotional intelligence or opinion mining, is the method or the process of collecting and extracting useful information based on unorganised or unstructured text. The process may include extraction from posts or comments from Twitter, online blogs, or other online platforms and sources. Sentiment analysis intends to help determine the text that the users generated and whether the opinion from that statement is positive, neutral, or negative (Taboada 2016). This method is utilised since online reviews of EVs usually include extensive details regarding particular characteristics or features, like appearance, performance, and interior features. In other words, aspect-based sentiment analysis breaks down the material to examine sentiments related to distinct parts or features of the product, in this case, EVs, rather than evaluating the entire review as a single sentiment.

Two sentiment analysis approaches are widely used for opinion mining and sentiment analysis: machine learning-based and lexicon-based. The machine learning-based method employs various supervised and unsupervised learning methods for sentiment classification. Conversely, the lexicon-based approaches employ a lexicon dictionary containing sentiment terms associated with a certain topic in order to classify sentiment. The polarity of every word, whether a term is positive or negative, is described in the dictionary. Note that the polarity of a word can be ascertained by matching it with its dictionary counterpart. Additionally, some studies have integrated methods based on lexicons and machine learning. Table 1 summarises the comparison of sentiment analysis techniques (Singh *et al.* 2020).

Table 1: Comparison of Sentiment Analysis Techniques

	Machine Learning-Based	Lexicon-Based
Approach	Supervised/Unsupervised	Unsupervised
Technique used	Naïve Bayes, Maximum Entropy, Support Vector Machine, Logistic Regression, Decision Tree, K-nearest Neighbours, Semi-supervised learning	PMI, Dictionary-based Approach
Data Preprocessing	Requires prior training in data	Does not require prior training in data
Factors Affecting Performance	Selected features Quality and quantity of training data Domain of the dataset	Opinion words included in the dictionary Polarity of many words in the domain context
Accuracy	70-90%	60-80%

For this paper, machine learning was chosen over a lexicon-based approach due to its superior performance and adaptability to the dataset (Machova *et al.* 2022). Since the dataset is the comments from YouTube, machine learning can learn the sentiment patterns from the labelled data, allowing them to capture context and nuances in user-generated text. In contrast, lexicon-based methods are unsupervised and depend solely on predefined dictionaries, which may fail to capture context, slang, or evolving language trends present in comments.

2.3. Related works

In recent years, sentiment analysis has become a pivotal tool for understanding consumer acceptance and perception across various fields. Numerous studies have utilised advanced

computational techniques to assess and interpret sentiment in textual data, providing valuable insights into consumer behaviour, social trends, and market dynamics. A study focusing on Arabic social media texts utilised TF-IDF and MRMR and subsequently implemented k-nearest neighbours (KNN), support vector machines (SVM), decision trees, and random forests algorithms. The findings revealed that SVM, when paired with TF-IDF, achieved an accuracy of 89%, outperforming other classifiers in the context of customer reviews for restaurants and cafes (Alsemaree *et al.* 2024).

Similarly, research analysing Amazon product reviews conducted a comparative study between machine learning models and lexicon-based approaches. The paper compared six sentiment classification methods: three supervised machine learning models (SVM, gradient boosting, and logistic regression) and three lexicon-based techniques (VADER, Pattern, and SentiWordNet). The study was conducted on Amazon product reviews, and NLP techniques such as stopword removal, word lemmatisation, and TF-IDF vectorisation were applied. The models are evaluated based on accuracy, precision, recall, and F1-score. The results show that all models effectively classified positive and negative reviews, with machine learning models outperforming lexicon-based methods in all metrics. This is likely because lexicon-based approaches rely on predefined word lists, which may not fully capture sentiment nuances. Among the machine learning models, the study reported that SVM, utilising TF-IDF vectorisation, attained an accuracy of 89%, surpassing other models in effectively classifying consumer sentiments (Nguyen *et al.* 2018).

Another investigation conducted a comprehensive sentiment analysis of app reviews to understand user opinions and behaviours. The primary objective was to classify sentiments expressed in these reviews and to explore university students' interactions with the app market.

The study employed various machine learning algorithms, including support vector machines (SVM), logistic regression (LR), and naïve Bayes (NB), utilising text representation schemes such as Term Presence (TP), Term Frequency (TF), and Term Frequency-Inverse Document Frequency (TF-IDF). The study revealed that the SVM classifier, combined with a tri-gram TF-IDF scheme, achieved the highest accuracy of 93.37% and an F1-score of 0.88 (Ranjan & Mishra 2020).

These studies collectively underscore the effectiveness of combining SVM classifiers with TF-IDF feature extraction in accurately assessing consumer sentiments. The consistent success of this methodology across diverse datasets and domains suggests its robustness and applicability in understanding consumer acceptance and perception.

2.3.1. Term Frequency-Inverse Document Frequency (TF-IDF)

In sentiment analysis, appropriate feature selection methods are crucial for finding pertinent characteristics and raising classification accuracy. Accordingly, five primary categories can be used to classify feature selection techniques: statistical, hybrid, clustering-based, Natural Language Processing (NLP), or heuristic-based.

This paper employs the Term Frequency-Inverse Document Frequency (TF-IDF). It is proven effective in capturing the semantic meaning of words in documents. The TF-IDF is a numerical statistic that indicates a word's significance in a document in relation to a set of documents. It is computed by taking the inverse document frequency (the rarity of a term throughout the collection of documents) and multiplying it by the term frequency (the number of times a term appears in a document) (Liang & Niu 2022).

$$TF(t, d) = \frac{\text{Number of times term } t \text{ appears in document } d}{\text{Total number of terms in document } d}, \quad (1)$$

where t is the term (word), and d is the specific document.

$$IDF(t, D) = \log \left(\frac{\text{Total number of documents in the corpus } D}{\text{Number of documents containing term } t} \right), \quad (2)$$

where D is the set of all documents in the corpus.

$$TD - IDF(t, d, D) = TF(t, d) \times IDF(t, D). \quad (3)$$

In information retrieval and text mining applications, TF-IDF is frequently used to determine the importance of words in texts. This method takes into account the frequency of words, ignoring their order.

In NLP and machine learning, TF-IDF is a widely used method for text analysis tasks like information retrieval, document classification, and clustering. This is attributable to the fact that they are effective tools for extracting relevant information from text data and aid in capturing the underlying semantic meaning of words.

2.3.2. Support vector machine

The support vector machine (SVM) is a supervised machine learning technique used for regression and classification problems. Its main goal is to determine the best hyperplane to divide data points into distinct classes in feature space. The technique maximises the margin between these support vectors, improving the model's generalisation to new, unseen data. Notably, "support vectors" are the data points closest to the decision boundary. With the introduction of kernel functions, SVM can handle non-linear decision boundaries and performs well in high-dimensional spaces. It is especially helpful in situations when the data exhibits distinct margins of separation. The equation for the hyperplane is represented as (Evgeniou & Pontil 2001):

$$f(x) = \text{sign}(w \cdot x + b), \quad (4)$$

where w is the weight vector, x is the feature vector derived from the text data, b is the bias term, and the $\text{sign}()$ is the sign function that determines the class label.

Optimising hyperparameters is essential to enhancing an SVM model's sentiment analysis performance. Possible hyperparameters for SVM are C (Regularisation Parameter), Kernel Type, Gamma (Kernel Coefficient), and Class Weight. The regularisation parameter, C , manages the trade-off between obtaining low testing and training errors. A bigger C seeks to classify all training data accurately, whereas a smaller C permits a more straightforward decision boundary. Moreover, SVM can employ various kernel functions such as 'linear', 'rbf' or custom kernels. Gamma is relevant when utilising 'rbf' or other non-linear kernels. Hence, higher values indicate the influence of a single training example, and more complex decision limits result from using it. Lastly, the 'class_weight' parameter is relevant for imbalanced datasets. Options include 'None' and 'Balanced.'. The classification task is a multi-class classification problem since the study divides sentiments into three categories: positive, neutral, and negative. The SVM model was trained using a one-on-one classification approach, ensuring it accurately assigns a sentiment category comments.

2.4. Evaluation classifier

A complete view of categorisation outcomes is provided by a confusion matrix that further divides down true positives, true negatives, false positives, and false negatives. The particular objectives of the work will determine which metrics are used, taking into account the implications of various mistake types within the application's environment. In the end, a classifier's evaluation entails a sophisticated interpretation of these measures, which directs advancements and well-informed model development decisions.

Metrics like recall, accuracy, precision, and F1-score thoroughly reveal how well the classifier can categorise occurrences. While precision and recall concentrate on particular classification elements, such as the accuracy of positive predictions and the model's capacity to capture all real positive cases, accuracy measures the overall correctness of predictions. The F1-score is a balanced statistic that considers both false positives and false negatives by combining precision and recall (Ting 2011).

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (7)$$

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \quad (8)$$

where TP is True Positive, FP is False Positive, TN is True Negative, and FN is False Negative.

In addition, the Receiver Operating Characteristic (ROC) curve and Area Under the Curve (AUC-ROC) demonstrate the classifier's performance at various threshold settings. An ideal classifier would hug the top-left corner of the ROC space, which denotes high sensitivity and specificity. Meanwhile, an AUC of 0.5 for a random guess might be represented by a diagonal line that runs from bottom left to top right. The model performs better when the curve is closer to the upper-left corner.

A classification model's overall performance is measured by its AUC. It suggests that the likelihood that a randomly selected positive instance would be ranked higher by the model than a randomly selected negative case. In other words, it assesses how well the model can discriminate between positive and negative classes. AUC is a number between 0.5 and 1, where 1 denotes a perfect classifier, and 0.5 denotes a random categorisation.

3. Results and Discussion

This paper aims to thoroughly grasp the attitudes currently held about EVs in the Malaysian environment. To achieve this objective, manual or automated labelling of the feedback and reviews collected will be conducted. Each of the comments in the dataset will be assigned a label, either positive, negative, or neutral. Based on the data, there are 209 positive, 119 neutral, and 52 negative labels. The class was mapped as Negative: 0, Neutral: 1, and Positive: 2. As a

result, positive sentiment scores and good classification accuracy suggest that Malaysian respondents are typically supportive and prepared to move forward with the topic or activity under consideration, according to the sentiment analysis classifier results. This good sentiment implies a positive perspective or agreement among the responders, suggesting a favourable attitude or willingness to move forward.

Figure 1 illustrates the word clouds for the positive and negative sentiments. The larger the word appears, the more frequent the words are in the comments. Some keywords include Tesla, battery, price, and charging. The word Tesla itself exhibits its impact on EVs in Malaysia, as Tesla is one of the most popular EV companies in Malaysia, and it is often talked about in Malaysia. Charging stations often appear from the user's perspective based on Figure 1. This is because the charging stations are rarely found in Malaysia, and due to that, potential users are concerned that there will be no changes to the number of charging areas. Thus, it will be inconvenient for them if no further action is taken.

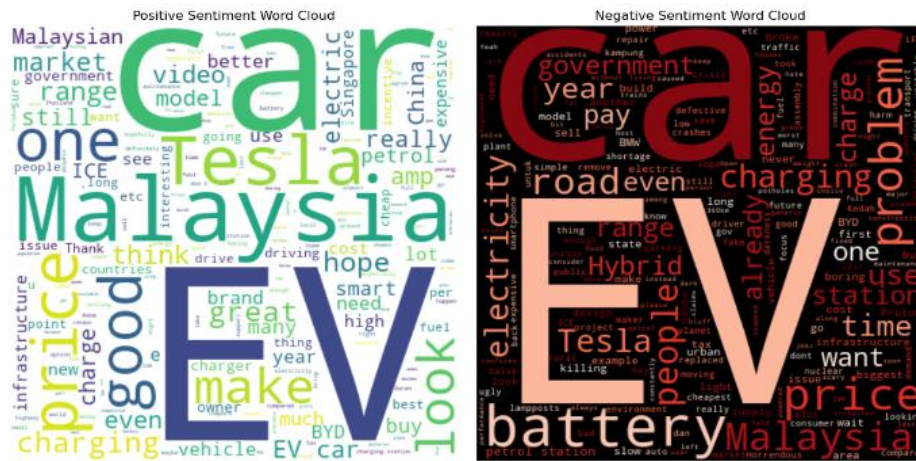


Figure 1: Word Clouds for The Positive and Negative Words

Figure 1 displays two-word clouds representing positive and negative sentiments associated with EVs. The left side illustrates the positive sentiment word cloud, highlighting terms favourably associated with EVs. Prominent words like "car," "Malaysia," "EV," "Tesla," "good," "make," "charging," and "price" indicate positive perceptions about EVs, their impact in Malaysia, and the brand Tesla. Words such as "range," "model," and "market" also suggest that people are optimistic about the driving range, variety, and market prospects of EVs. On the right side, the negative sentiment word cloud emphasises the challenges and concerns related to EVs. Key terms include "EV," "price," "battery," "problem," "road," "government," "charging," "electricity," "range," and "station." These words reflect issues such as high costs, battery problems, infrastructure shortcomings, and governmental policies. Overall, the image illustrates a balanced view of EVs, highlighting both enthusiasm for their potential and recognition of the obstacles that must be addressed for wider acceptance.

For sentiment prediction, a thorough assessment of the created machine learning algorithms was evaluated using performance indicators, including recall percentages, accuracy, and precision. The assessment procedure aims to confirm the algorithm's dependability and appropriateness for identifying complex viewpoints on the acceptance and perception of EVs in the Malaysian setting. SVM demonstrated a good response and accuracy when using TF-IDF to capture the semantic meaning of words.

SVM Classifier Performance:
Accuracy: 0.8518095238095238
Classification Report:

	precision	recall	f1-score	support
0	0.95	0.97	0.96	36
1	0.78	0.93	0.85	42
2	0.95	0.77	0.85	48
accuracy			0.88	126
macro avg	0.89	0.89	0.89	126
weighted avg	0.89	0.88	0.88	126

Figure 2: Classifier Performance for SVM based on TF-IDF

Based on Figure 2, the SVM has an accuracy of 85.18% based on the TF-IDF method. In addition, in Figure 2, the precision on classification based on TF-IDF is in the range of 78% to 95%, recall of 77% to 97%, and F1-score of 85% to 96% for SVM. The results were obtained after performing K-Fold Cross-Validation, which reduces the variability in the performance estimate, leading to a more reliable evaluation. Overall, SVM has higher accuracy and F1-scores for each class, concluding that SVM with TF-IDF is able to predict the sentiment of EV in Malaysia.

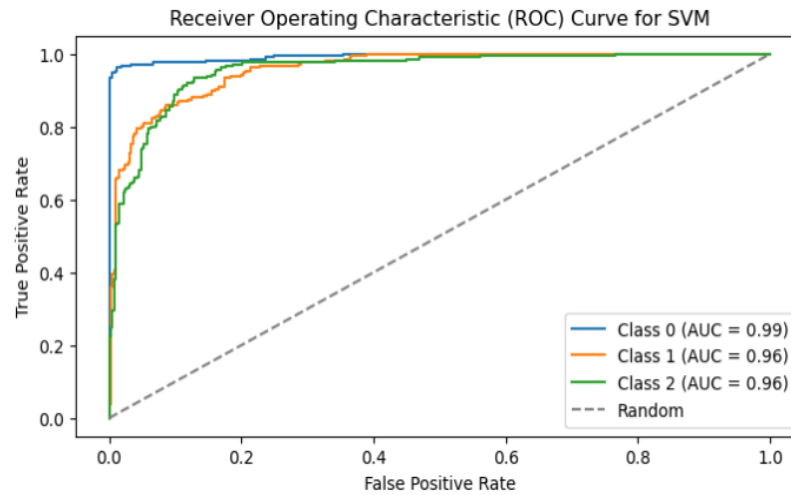


Figure 3: ROC Curve and AUC Score for TF-IDF in SVM

Since the accuracy of SVM using TF-IDF is good, the ROC curve and AUC also produced the same results as displayed in Figure 3. This indicates that the model performs better since the curve is closer to the upper-left corner, and the AUC score exhibits a very good result. This proves that the TF-IDF SVM model is consistent across the three classes, demonstrating robustness to overfitting and reliability in multiple scenarios. The implication for further analysis is that SVM using TF-IDF can be used for text classification prediction.

4. Conclusion

This study aims to identify how electric vehicles (EVs) are accepted and perceived in Malaysia by conducting sentiment analysis and employing the machine learning algorithm known as

support vector machines (SVM) for sentiment prediction. Data reflecting Malaysian consumer attitudes towards EVs was collected from YouTube comments. The results show that 55% of the sentiments are positive, 14% are negative, while the rest are neutral. The SVM model achieved high accuracy in all assessments, confirmed by the AUC score, proving the effectiveness of SVM for text classification and sentiment prediction in this dataset. While some dissenting opinions exist, Malaysians appear open to embracing the transition to electric vehicles. The increasing popularity of EVs in Malaysia is evident, particularly as the government promotes their acceptance by enhancing the number of charging stations at petrol stations to facilitate environmentally friendly transportation. The Malaysian government has also set ambitious goals to boost EV adoption, targeting 100,000 electric cars, 100,000 electric motorcycles, 2,000 electric buses, and 125,000 charging stations by 2030. With such government initiatives and growing public acceptance, Malaysia stands ready to enter the era of electric mobility and support a greener future. However, challenges remain, including battery costs, limited charging infrastructure, and the overall expense of EVs in Malaysia. Many potential buyers express concerns over the high costs of electric vehicles, making them seem less worthwhile. Therefore, government intervention may be crucial in encouraging EV acceptance by offering financial incentives such as tax credits, subsidies, or reduced registration fees for prospective buyers use.

Acknowledgments

The authors would like to thank the anonymous reviewers for their thorough assessment and constructive comments. Their expertise and valuable insights have greatly enhanced the quality and clarity of the content.

References

- Adnan N., Nordin S.M., Amini M.H. & Langove N. 2018. What make consumer sign up to PHEVs? Predicting Malaysian consumer behavior in adoption of PHEVs. *Transportation Research Part A: Policy and Practice* **113**: 259–278.
- Adnan N., Vasant P.M., Rahman I. & Noor A. 2016. Adoption of plug-in hybrid electric vehicle among Malaysian consumers. *Industrial Engineering & Management* **5**(2): 185.
- Adnan N., Nordin S.M., Rahman I. & Rasli A.M. 2017. A new era of sustainable transport: An experimental examination on forecasting adoption behavior of EVs among Malaysian consumer. *Transportation Research Part A: Policy and Practice* **103**: 279–295.
- Alsemaree O., Alam A.S., Gill S.S. & Uhlig S. 2024. Sentiment analysis of Arabic social media texts: A machine learning approach to deciphering customer perceptions. *Heliyon* **10**(9): e27863.
- Asadi S., Nilashi M., Samad S., Abdullah R., Mahmoud M., Alkinani M.H. & Yadegaridehkordi E. 2021. Factors impacting consumers' intention toward adoption of electric vehicles in Malaysia. *Journal of Cleaner Production* **282**: 124474.
- Bryła P., Chatterjee S. & Ciabiada-Bryła B. 2023. Consumer adoption of electric vehicles: A systematic literature review. *Energies* **16**(1): 205.
- Bueckers J., Van Holderbeke M., Bierkens J. & Panis L.I. 2014. Health and environmental benefits related to electric vehicle introduction in EU countries. *Transportation Research Part D: Transport and Environment* **33**: 26–38.
- Capuder T., Sprčić D.M., Zoričić D. & Pandžić H. 2020. Review of challenges and assessment of electric vehicles integration policy goals: Integrated risk analysis approach. *International Journal of Electrical Power & Energy Systems* **119**: 105894.
- Evgeniou T., & Pontil M. 2001. Support vector machines: Theory and applications. In Paliouras G., Karkaletsis V. & Spyropoulos C.D. (eds.). *Machine Learning and Its Applications: Advanced Lectures*: 249–257. Berlin, Heidelberg: Springer.
- Guo S. & Kontou E. 2021. Disparities and equity issues in electric vehicles rebate allocation. *Energy Policy* **154**: 112291.
- Liang M. & Niu T. 2022. Research on text classification techniques based on improved TF-IDF algorithm and LSTM inputs. *Procedia Computer Science* **208**: 460–470.

- Liu Y., Ouyang Z. & Cheng P. 2019. Predicting consumers' adoption of electric vehicles during the city smog crisis: An application of the protective action decision model. *Journal of Environmental Psychology* **64**: 30–38.
- Machova K., Mach M. & Vasilko M. 2022. Comparison of machine learning and sentiment analysis in detection of suspicious online reviewers on different type of data. *Sensors* **22**(1): 155.
- Muzir N.A.Q., Mojumder M.R.H., Hasanuzzaman M. & Selvaraj J. 2022. Challenges of electric vehicles and their prospects in Malaysia: A Comprehensive review. *Sustainability* **14**(14): 8320.
- International Trade Administration. 2023. Malaysia electric vehicles. 2023. <https://www.trade.gov/market-intelligence/malaysia-electric-vehicles> (3 January 2025).
- Nguyen H., Veluchamy A., Diop M. & Iqbal R. 2018. Comparative study of sentiment analysis with product reviews using machine learning and lexicon-based approaches. *SMU Data Science Review* **1**(4): 7.
- Ramachandran S.D., Ng H., Rajermani R. & Raman A. 2023. Factor influencing consumer's adoption of electric cars in Malaysia. *TEM Journal* **12**(4): 2603-2612.
- Ranjan S. & Mishra S. 2020. Comparative sentiment analysis of app reviews. *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* in Kharagpur, India, pp. 1–7.
- Richter F. 2023. Electric mobility: Norway leads the charge. <https://www.statista.com/chart/17344/electric-vehicle-share/> (15 February 2024).
- Sang Y.N. & Bekhet H.A. 2015. Modelling electric vehicle usage intentions: An empirical study in Malaysia. *Journal of Cleaner Production* **92**: 75–83.
- Singh N.K., Tomar D.S. & Sangaiah A.K. 2020. Sentiment analysis: a review and comparative analysis over social media. *Journal of Ambient Intelligence and Humanized Computing* **11**: 97-117.
- Taboada M. 2016. Sentiment analysis: An overview from linguistics. *Annual Review of Linguistics* **2**(1): 325–347.
- Ting K.M. 2011. Confusion matrix. In Sammut C. & Webb G.I. (eds.). *Encyclopedia of Machine Learning*: 209-209. Boston, MA: Springer.

Centre for Mathematical Sciences
Universiti Malaysia Pahang Al-Sultan Abdullah
26300 Kuantan
Pahang, MALAYSIA
*E-mail: farawahida1519.abdullah@gmail.com, haizum@umpsa.edu.my**

Received: 7 June 2024
Accepted: 6 March 2025

*Corresponding author