COMPETENTIE JOURNAL
Vol 02, No. 02, 2025, pp. 23-28

https://journal.inacexter.com/index.php/competentie

Topic Modeling of Public Discourse on Electric Vehicles in Indonesia Using BERTopic

Muhammad Rizal*[1], Cici Suhaeni[2]

[1,2] IPB University, Bogor, Indonesia
*Corresponding Author: mmmdrizal02@gmail.com

Abstract*: The global shift toward sustainable transportation has increased public attention to electric vehicles (EVs) as an alternative to conventional mobility. This transition also supports the United Nations' Sustainable Development Goals (SDGs), particularly those related to affordable clean energy and climate action. This study examines how Indonesian users discuss EVs on X (formerly Twitter) by applying BERTopic, an embedding-based topic modeling framework that leverages multilingual sentence embeddings to identify latent themes in public discourse. The analysis reveals that online conversations are dominated by discussions of practical usage, affordability, technological readiness, and brand awareness, while environmental concerns appear less prominent. These insights contribute to a deeper understanding of public perceptions of electric mobility and highlight the social and economic factors influencing EV acceptance in Indonesia.*

Keywords: Electric vehicles; Topic Modeling; BERTopic

Introduction

The transition toward sustainable transportation has gained increasing global attention, driven by innovations in alternative energy technologies such as solar, hydrogen, and electric vehicles (EVs) (Tolani et al., 2025). Electric vehicles, in particular, are becoming increasingly popular due to their independence from fossil fuels and their potential to reduce greenhouse gas emissions (Alanazi, 2023). This growing interest aligns with the global agenda to achieve the United Nations' Sustainable Development Goals (SDGs), particularly Goal 7 (Affordable and Clean Energy) and Goal 13 (Climate Action) (United Nations, 2015).

EV adoption also supports several other SDGs: Goal 3 (Good Health and Well-being) by reducing air and noise pollution; Goal 9 (Industry, Innovation, and Infrastructure) by fostering technological advancement and sustainable industrial practices; Goal 11 (Sustainable Cities and Communities) by promoting cleaner and more efficient urban mobility; and Goal 12 (Responsible Consumption and Production) through the reduction of environmental impacts from vehicle manufacturing and operation. Collectively, the widespread utilization of EVs plays a pivotal role in advancing sustainable development by mitigating climate change, expanding clean energy use, and supporting resilient communities (Arefin et al., 2025).

However, the adoption of EVs varies across societies and is influenced by factors such as technological perception, trust, cost considerations, and environmental awareness. Social media has become a crucial platform where individuals express opinions, share experiences, and shape collective views about EVs. By sharing their experiences, expectations, and concerns, users collectively shape public perceptions and attitudes toward this emerging technology. Understanding these societal attitudes is essential for the future development and adoption of EVs. Consequently, analyzing user-generated content on social media offers valuable insights into public interest, concerns, and expectations regarding electric mobility (Özkara et al., 2025).

The vast and dynamic nature of discussions on social media presents challenges for understanding public discourse. These platforms contain spontaneous and diverse opinions, making them a valuable yet complex source of information that requires systematic analytical methods. To extract meaningful insights from such unstructured text, Natural Language Processing (NLP) techniques are often applied. Among these methods, topic modeling is widely used to uncover latent themes and identify key patterns in large-scale textual data (Suhaeni et al., 2025). Topic modeling algorithms work by identifying hidden patterns within word distributions across a collection of documents. The resulting topics consist of groups of words that frequently co-occur, representing distinct semantic structures or themes within the text (Aryani et al., 2024).

One of the most classical and widely adopted topic modeling approaches is Latent Dirichlet Allocation (LDA), which has been extensively applied across various domains, including news and social media analysis. However, LDA has well-documented limitations, particularly in handling short and unstructured text data, which often results in vague or overly general topics that provide limited analytical value. To address these challenges, recent advances have introduced embedding-based models such as BERTopic, which leverage transformer-based language representations to capture semantic context more effectively. By integrating BERT embeddings, clustering algorithms, and class-based TF–IDF (c-TF–IDF), BERTopic produces more coherent and interpretable topic representations, especially in social media and online discourse contexts where traditional models like LDA often fall short (Asnawi et al., 2024; Tamzila et al., 2025).

Several studies have explored public discourse on electric vehicles in Indonesia through various text analytics approaches. For instance, previous research has analyzed discussions using sentiment analysis of Indonesian-language data (Salsabila et al., 2023; Siregar et al., 2024), while others have implemented topic modeling methods such as BERTopic to examine YouTube comments related to EV adoption (Simanjuntak et al., 2024). However, these studies have remained limited in terms of modeling configuration and platform coverage, leaving broader patterns of public conversation unexplored.

This study addresses these gaps by applying BERTopic with a multilingual embedding model (paraphrase-multilingual-MiniLM-L12-v2) to analyze social media discussions about electric vehicles in X (formerly known as Twitter) in Indonesia. By utilizing default model parameters while modifying the embedding representation, this study aims to observe how multilingual embeddings perform in capturing thematic structures within Indonesian EV-related discourse. The findings are expected to offer a broader view of public narratives and perceptions surrounding electric mobility and support future research on sustainable transportation.

Methods

This study examines public discourse on electric vehicles (EVs) in Indonesia using social media text data collected from X (formerly Twitter). The dataset consists of posts containing EV related keywords published between 30 December 2024 and 2 September 2025. The analysis was conducted through several stages. First, tweets related to electric vehicles were retrieved using keyword-based filtering. The collected texts were then preprocessed, including lowercasing, removal of URLs and emojis, elimination of non-informative tokens, and text normalization to ensure data quality. Topic modeling was performed using BERTopic (Grootendorst, 2022), which includes multilingual sentence embeddings, dimensionality reduction using UMAP, clustering via HDBSCAN, and keyword extraction using class-based TF–IDF. Topic quality was assessed using topic coherence and topic diversity. To support further interpretation, hierarchical clustering was used to identify relationships between topics, which were subsequently summarized into broader thematic categories. The overall workflow is illustrated in Figure 1.

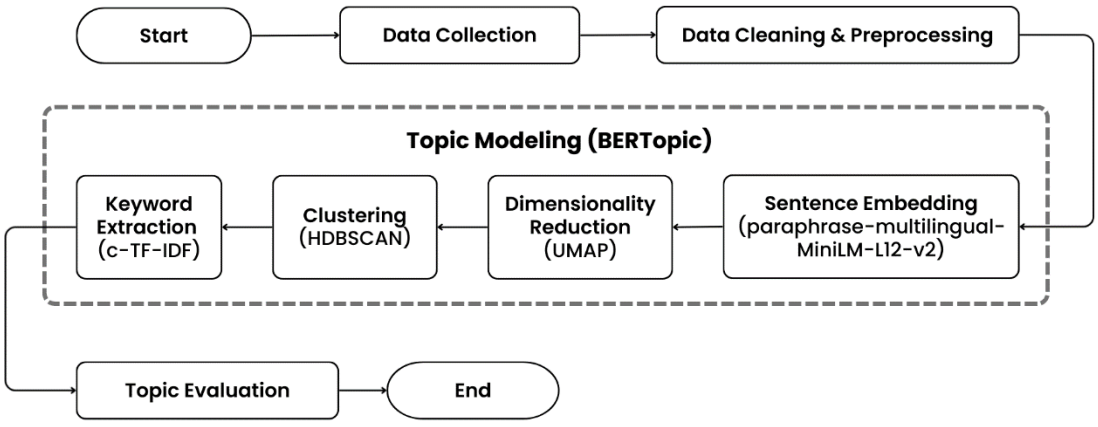

Figure 1. Research workflow for EV discourse analysis

Results and Discussion

In this study, the BERTopic model was applied to a corpus of tweets discussing electric vehicles (EVs) in Indonesia to uncover key themes in public discourse. Before modeling, the text data were cleaned, normalized, and transformed into sentence embeddings using the paraphrase-multilingual-MiniLM-L12-v2 model. The model identified 34 distinct topics representing diverse perspectives on EV adoption in Indonesia. To assess the quality of these topics, two intrinsic evaluation metrics, Topic Coherence and Topic Diversity, were calculated. The Topic Coherence Score was 0.5337, indicating that the identified topics were moderately coherent, while the Topic Diversity Score was 0.7853, reflecting a relatively high lexical variation across topics. These results suggest that the model produced a reasonable level of topic interpretability while maintaining diversity across themes. To further explore how these topics relate to one another, hierarchical clustering was conducted to visualize the structural organization of EV-related discussions in Indonesia.

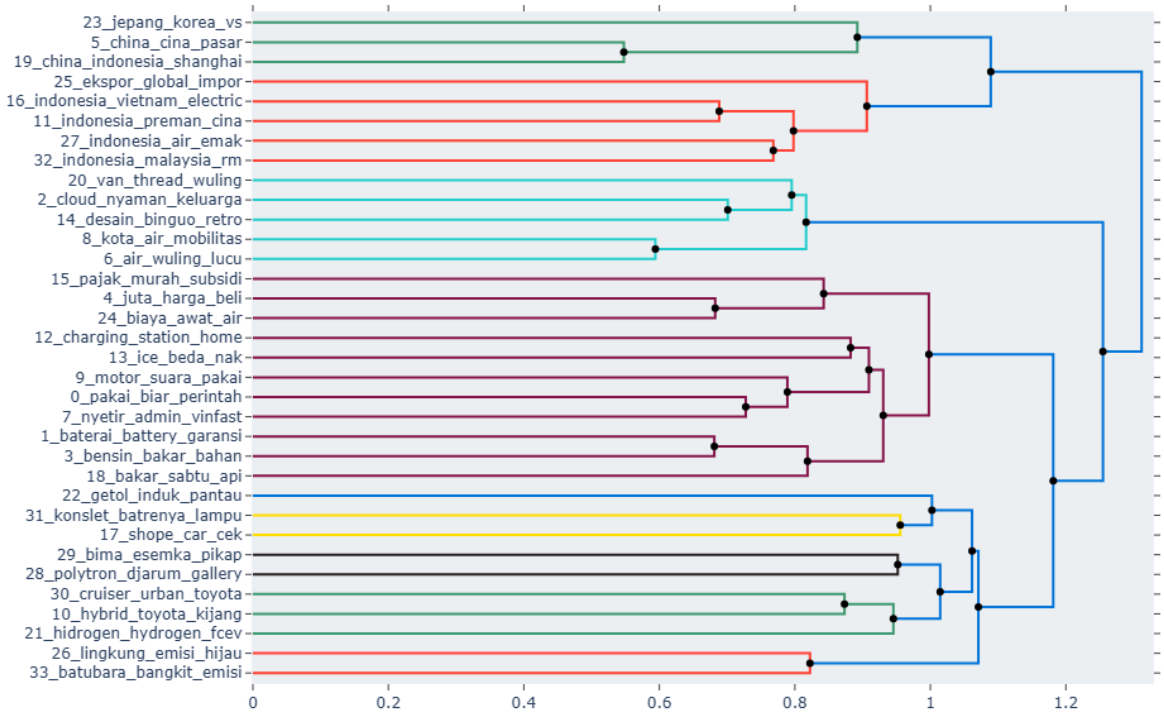

Figure 2. Hierarchical clustering of EV discourse topics in Indonesia

Figure 2 illustrates the hierarchical relationships among the 34 topics identified by BERTopic. Clusters with shorter linkages indicate stronger semantic relationships between topics. The figure shows that discussions related to pricing and consumer interest, technological aspects and infrastructure, and brand or product narratives tend to appear together in public conversations about electric vehicles. These patterns suggest that public discourse in Indonesia mainly revolves around economic and practical aspects of EV use, while environmental and policy-related discussions appear less prominent. Overall, the clustering structure highlights those conversations about electric vehicles in Indonesia are shaped primarily by affordability, technological readiness, and brand awareness rather than environmental considerations. From the hierarchical clustering, six general themes were synthesized to represent the major dimensions of public discourse, as summarized in Table 1.

Table 1. General Themes of Electric Vehicle Discourse in Indonesia

Theme	Interpretation
Everyday Use and User Perceptions	Covers users daily experiences with EVs, including driving feel, city mobility, compact-size benefits, convenience, parking ease, and social impressions. Also reflects public perceptions regarding brand presence, dealership visibility, and suitability of EVs for different lifestyles.
Affordability and Ownership Costs	Discusses financial considerations related to buying and maintaining EVs, including purchase price, subsidies, resale value, maintenance costs, and comparisons with ICE vehicles or EV pricing in other countries.
Technology Reliability and Charging Infrastructure	Encompasses technological concerns such as battery durability, degradation, warranties, fire incidents, and electrical faults. Also includes charging availability, charging speed, route planning issues, and infrastructure accessibility, highlighting the interplay of reliability and charging ecosystems.
Global Market Positioning and Industry Competition	Focuses on government incentives, tax reforms, industrial policies, import regulations, and EV adoption targets. Includes discourse on whether policies prioritize consumers, industry players, or broader national economic goals.
Policy Framework and National Strategy	Tweets reflect how EVs are used and discussed in daily life, emphasizing aspects of convenience, practicality, and integration into everyday routines. The theme represents how users share real-life experiences and perceptions of EV adoption in different social contexts.
Environmental Impact and Sustainability Perspectives	Highlights contrasting sustainability narratives enthusiasm for zero-emission mobility versus concerns about coal electricity, nickel mining, battery waste, industrial pollution, and upstream environmental burdens. Indicates polarized public sentiment on EVs' true environmental footprint.

Table 1 summarizes the six overarching themes derived from the hierarchical clustering of 34 BERTopic-generated topics. The first theme, Everyday Use and User Perceptions, describes how electric vehicles (EVs) are integrated into daily mobility routines in Indonesia. Public conversations commonly highlight the practical aspects of EV usage, including comfort, ease of driving, suitability for dense urban environments, and overall user impressions. These discussions also reflect how consumers perceive brand visibility and dealership presence as part of their everyday experience with EVs.

The second theme, Affordability and Ownership Costs, centers on economic considerations that strongly influence public attitudes toward EV adoption. Posts frequently refer to the purchase price of EVs, maintenance expenses, resale value, and comparisons with conventional vehicles or EV prices abroad. These concerns indicate that financial feasibility remains a major factor shaping consumer interest and decision-making.

The third theme, Technology Reliability and Charging Infrastructure, captures the public's attention to the technical performance and dependability of EVs. Many discussions relate to battery longevity, safety, charging accessibility, and the practical challenges of planning longer trips. These conversations highlight the extent to which perceptions of technological readiness and charging convenience affect confidence in EV ownership.

The fourth theme, Global Market Positioning and Industry Competition, captures the broader industry landscape within which EV adoption is taking place. Conversations include references to the dominance of certain manufacturers, emerging competition among global and regional brands, and Indonesia's efforts to strengthen its position within the global EV supply chain. Public interest also extends to product launches, brand strategies, and cross-border dynamics in EV production and distribution.

The fifth theme, Policy Framework and National Strategy, represents public reactions toward government-led initiatives aimed at accelerating EV adoption. Discussions often address incentives, tax policies, industrial strategies, and national targets, revealing mixed views on the effectiveness of these measures. The theme illustrates how policy direction influences both public expectations and perceptions of long-term EV deployment in Indonesia.

Finally, the theme Environmental Impact and Sustainability Perspectives captures a spectrum of views on the environmental implications of EVs. While some conversations highlight the potential of EVs to reduce emissions, others emphasize concerns related to electricity production sources, resource extraction, and waste management. These perspectives show that environmental considerations are present but not as dominant as economic and practical concerns in shaping public discourse. In summary, the themes show that EV-related conversations in Indonesia are primarily shaped by practical usage, affordability, and industry competition, while environmental issues receive comparatively less attention in public discussions.

Conclusion

This study examined public discourse on electric vehicles (EVs) in Indonesia using BERTopic, an embedding-based topic modeling framework that applies multilingual sentence embeddings to analyze social media text. The analysis revealed that public discussions about EVs are largely driven by practical usage, affordability, technological readiness, and brand awareness, while environmental issues receive comparatively less attention. These findings indicate that economic and functional factors remain central to how Indonesian users perceive and engage with electric mobility. Although the multilingual embedding captured relevant thematic structures, the results also suggest that language-specific embeddings such as those trained on Indonesian corpora could further improve topic coherence and contextual accuracy. Future research may extend this work by incorporating temporal or sentiment analyses to observe how public perceptions evolve alongside policy and technological developments.

References

Alanazi, F. (2023). Electric Vehicles: Benefits Challenges and Potential Solutions. *Journal of Applied Scienc*, *13*, 1–23.

Arefin, A. A., Meraj, S. T., Lipu, M. S. H., Rahman, M. S., Rahman, T., Hasan, K., Sarker, M. R., & Muttaqi, K. M. (2025). Societal, environmental, and economic impacts of electric vehicles towards achieving sustainable development goals. *Results in Engineering*, *27*(September). https://doi.org/10.1016/j.rineng.2025.107060

Aryani, D., Lucia Kharisma, I., Sujjada, A., & Kamdan, K. (2024). Topic Modeling of the 2024 Election Using the BERTopic Method on Detik.com News Articles. *Inform : Jurnal Ilmiah Bidang Teknologi Informasi Dan Komunikasi*, *9*(2), 171–180. https://doi.org/10.25139/inform.v9i2.8429

Asnawi, M. F., Hanafi, M., Kurniawan, N. F., Suwondo, A., Nasrullah, A., & Setyawan, C. (2024). Topic Modelling Analysis on Indonesian News Using BERT Topic Model. *2024 6th International Conference on Cybernetics and Intelligent System (ICORIS)*, 1–6. https://doi.org/10.1109/ICORIS63540.2024.10903779

Grootendorst, M. (2022). *BERTopic: Neural topic modeling with a class-based TF-IDF procedure*. http://arxiv.org/abs/2203.05794

Özkara, Y., Bilişli, Y., Yildirim, F. S., Kayan, F., Başdeğirmen, A., Kayakuş, M., & Yiğit Açıkgöz, F. (2025). Analysing Social Media Discourse on Electric Vehicles with Machine Learning. *Applied Sciences (Switzerland)*, *15*(8), 1–20. https://doi.org/10.3390/app15084395

Salsabila, H., Habibi, R., & Harani, N. H. (2023). Social Media-Based Sentiment Analysis: Electric Vehicle Usage in Indonesia. *Indonesian Journal of Computer Science*, *12*(3), 1132–1146. https://doi.org/doi.org/10.33022/ijcs.v12i3.3250

Simanjuntak, K. A., Koyimatu, M., & Ervanisari, Y. P. (2024). Analisis Perubahan Opini Publik Terhadap Kendaraan Listrik di Indonesia Melalui Komentar YouTube: Pendekatan Topic Modeling BERTopic. *Jurnal Inovasi Kewirausahaan*, *1*(3), 1–9.

Siregar, A. M., Faisal, S., Fauzi, A., Indra, J., Masruriyah, A. F. N., & Pratama, A. R. (2024). Model machine learning for sentiment analysis of the presence of electric vehicle in Indonesia. *BIS Information Technology and Computer Science*, *1*, V124022. https://doi.org/10.31603/bistycs.140

Suhaeni, C., Nissa, L., Mualifah, A., & Wijayanto, H. (2025). LDA Topic Modeling Analysis of Public Discourse on Indonesia's Free Nutritious Meals Program (MBG). *International Journal on Informatics for Development*, *14*(1), 587–600. https://doi.org/10.14421/ijid.2025.5211

Tamzila, A. S., Sulistya, A., & Lidiawaty, B. R. (2025). A Comparative Evaluation of LDA and BERTopic for Topic Modeling of Traffic Complaints in Short Social Media Texts. *2025 International Electronics Symposium (IES)*, 567–572. https://doi.org/10.1109/ies67184.2025.11161719

Tolani, K., Manohar, S., & Rao, S. (2025). Measuring sustainable mobility of electric vehicles: determining critical factors with policy mix support for developing economies and user convenience. In *Journal of Innovation and Entrepreneurship* (Vol. 14, Issue 1). https://doi.org/10.1186/s13731-025-00560-2

United Nations. (2015). Arsenic and the 2030 Agenda for sustainable development. In *Transforming our world: The 2030 Agenda for Sustainable Development*.