Journal of Soft Computing Exploration

Homepage: https://shmpublisher.com/index.php/joscex

p-ISSN: 2746-7686
e-ISSN: 2746-0991

.

A CoCoSo-based decision support framework for electric vehicle evaluation using deep learning-derived policy sentiment evidence

Nur Tulus Ujianto[1], Gunawan Gunawan[2], Gunawan Raharjo[3], Didiek Trisatya[4], Diajeng Tyas Purwa Hapsari[5]
[1,2,3,4,5]Department of Information System, Universitas Pancasakti Tegal, Indonesia

Article Info

Article history:

Received April 30, 2026
Revised May 23, 2026
Accepted July 4, 2026

Keywords:

Combined compromise solution (CoCoSo),
Decision support system,
Deep learning sentiment analysis,
Electric vehicle evaluation,
Multi-criteria decision making,

ABSTRACT

Electric vehicle adoption is shaped not only by technical and economic attributes but also by public perception of government incentive policies. Existing electric vehicle evaluation models commonly prioritize conventional criteria, while policy-related public sentiment is often treated only as descriptive context. This study proposes a hybrid decision support framework that integrates IndoBERT-based sentiment analysis with the Combined Compromise Solution (CoCoSo) method for electric vehicle evaluation in Indonesia. Public comments on electric vehicle incentives were classified into positive, neutral, and negative sentiment to construct a Policy Sentiment Evidence Index (PSEI). The index was incorporated into the CoCoSo decision matrix together with price, driving range, battery capacity, and charging time. The sentiment classifier achieved an accuracy of 0.6447 and a macro-F1 score of 0.5228, indicating moderate class-balanced performance under imbalanced sentiment distribution. The rescaled PSEI value of 0.3746 indicates relatively unfavorable public sentiment toward electric vehicle incentives. The CoCoSo results ranked Tesla Model 3 first, followed by Hyundai Ioniq 5, Nissan Leaf, and Wuling Air EV. Sensitivity analysis confirmed stable rankings across policy sentiment weighting scenarios, suggesting that sentiment evidence can enrich multi-criteria electric vehicle evaluation without destabilizing the decision outcome.

Corresponding Author:

Nur Tulus Ujianto,
Department of Information System,
Universitas Pancasakti Tegal,
Jl. Halmahera Km 1, Tegal Timur, Tegal, Central Java, Indonesia
Email: nurtulus@upstegal.ac.id
https://doi.org/10.52465/joscex.v7i3.80

1. INTRODUCTION

The transition to electric vehicles (EVs) has become an important part of low-carbon mobility strategies, particularly in countries seeking to reduce dependence on fossil fuels and improve transport sustainability. In Indonesia, EV adoption is not determined only by technological readiness, but also by affordability, infrastructure availability, consumer awareness, and the credibility of government incentive policies. Hakam and Jumayla show that EV adoption in Indonesia requires lessons from both developed and developing countries, especially because market readiness and public acceptance remain uneven [1]. Kresnanto and Putri further emphasize that EV subsidies can support green transportation, but their effectiveness depends on how the policy is perceived and implemented [2].

EV selection is also a multi-criteria decision problem. Consumers and policymakers must evaluate alternatives based on heterogeneous criteria, such as vehicle price, driving range, battery capacity, charging time, operating efficiency, and perceived value. Ecer demonstrates that MCDM methods can support battery electric vehicle assessment by structuring different performance indicators into a ranking framework [3]. Tian et al. also show that data-driven multi-criteria decision support can improve EV selection by combining product attributes with decision-relevant information [4]. More broadly, recent MCDM literature emphasizes that decision-support methods are useful when alternatives must be evaluated under multiple, conflicting, and uncertain criteria [5]–[7]. These studies indicate that MCDM-based decision support is suitable for EV evaluation, but most models still emphasize technical and economic attributes.

At the same time, public discourse has become an important source of evidence for understanding EV adoption. Online comments and social media discussions capture spontaneous public reactions to vehicle prices, infrastructure readiness, policy incentives, and trust in government programs. Nugroho and Widianto show that sentiment analysis can reveal public attitudes toward EV adoption in Indonesia, while Purnama et al. demonstrate that time-series social media analysis can support data-driven public policy for EV development [8], [9]. Similar evidence from EV-related public discourse also shows that social media and online review data can capture public attention, acceptance, concern, and preference toward electric vehicles [10], [11]. These studies confirm that sentiment evidence is useful for understanding the behavioral and policy context of EV adoption.

However, sentiment-based EV studies and MCDM-based EV evaluation studies are still often treated as separate research streams. Sentiment studies commonly describe public opinion, but they rarely transform sentiment outputs into structured decision criteria. Conversely, MCDM-based EV studies usually rank vehicle alternatives using technical and economic criteria, but they often treat public sentiment and policy perception as external contextual discussion. This separation limits the ability of decision support systems to represent the real adoption environment, especially in emerging markets where public trust in incentive policies can shape consumer acceptance.

Recent studies have started to combine online reviews, user preferences, and decision-support models for product evaluation. Yang et al. integrate customer reviews and product recommendation logic for new energy vehicle decision-making, while Wu et al. show that online reviews can support consumer preference analysis and product ranking [12], [13]. In EV-related sentiment research, Gong et al. demonstrate that online reviews can be transformed into analytical evidence for electric vehicle evaluation [14]. Recent studies on aspect-based sentiment analysis further show that fine-grained sentiment extraction can help transform unstructured opinions into more decision-relevant evidence [15], [16]. Nevertheless, limited attention has been given to converting deep learning-derived policy sentiment into an explicit criterion within a CoCoSo-based EV ranking framework.

The novelty of this study lies in the transformation of deep learning-derived policy sentiment evidence into a structured decision criterion within a CoCoSo-based electric vehicle evaluation framework. Unlike previous EV sentiment studies that mainly describe public opinion, and unlike conventional MCDM studies that primarily rely on technical and economic attributes, this study integrates probabilistic sentiment evidence from Indonesian public discourse into the multi-criteria ranking process.

Therefore, this study proposes a hybrid decision support framework that integrates IndoBERT-based sentiment classification with the Combined Compromise Solution (CoCoSo) method for evaluating electric vehicle alternatives in Indonesia. The objectives of this study are to extract policy sentiment from public comments on Indonesian EV incentives, construct a Policy Sentiment Evidence Index (PSEI), incorporate sentiment-derived evidence into a multi-criteria decision matrix, rank selected EV alternatives using CoCoSo, and evaluate the stability of the ranking through sensitivity analysis.

This study makes three main contributions. First, it extends EV evaluation beyond conventional specification-based ranking by incorporating policy sentiment as a measurable contextual criterion. Second, it integrates deep learning-based sentiment extraction with the CoCoSo multi-criteria ranking framework in a unified decision-support model. Third, it demonstrates a context-sensitive EV evaluation approach for emerging markets, where consumer trust, policy perception, and technical vehicle attributes jointly influence adoption decisions.

.

2. METHOD

Study design and analytical workflow

This study adopted a hybrid decision-support design that combined supervised deep learning-based sentiment classification with a CoCoSo-based multi-criteria ranking framework for electric vehicle (EV) evaluation. The logic of the framework was to treat public reactions to EV incentive policies not merely as contextual discussion, but as decision evidence that could be transformed into a structured input for downstream ranking. This design is consistent with recent work showing that EV adoption in Indonesia is strongly shaped by policy perception and public discourse, and that text-derived evidence can enrich decision-oriented EV analytics beyond purely technical specifications [1], [8], [17].

The analytical workflow consisted of five sequential stages: (1) dataset preparation and quality control, (2) deep learning–based sentiment modeling, (3) construction of a policy sentiment evidence index, (4) integration of sentiment evidence into the CoCoSo-based decision support framework, and (5) validation through classification metrics, rank stability testing, and sensitivity analysis [17], [18].

The complete workflow is illustrated in Figure 1. This figure clarifies how unstructured policy-related public comments were transformed into structured sentiment evidence and subsequently integrated into the CoCoSo-based EV evaluation process. By presenting the workflow visually, the methodological connection between sentiment classification, PSEI construction, decision matrix formation, ranking, and sensitivity analysis becomes more transparent and reproducible.

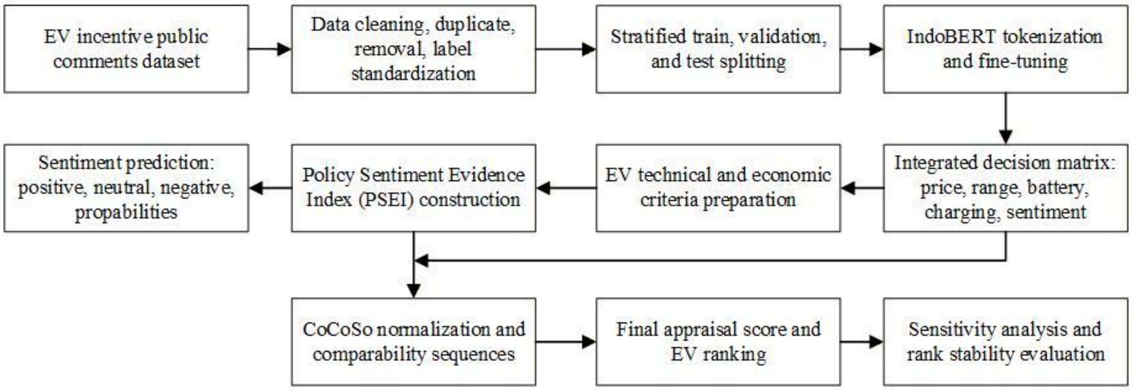

Figure 1. Research workflow integrating IndoBERT-based policy sentiment evidence and CoCoSo ranking.

Figure 1 presents the complete analytical workflow of the proposed decision-support framework. The process begins with the acquisition and preparation of public comments on electric vehicle incentive policies, followed by stratified data splitting and IndoBERT-based sentiment classification. The predicted sentiment probabilities are aggregated into the Policy Sentiment Evidence Index (PSEI), which is then integrated with technical and economic vehicle criteria in the CoCoSo decision matrix. The final stage produces the EV ranking and evaluates rank stability through sensitivity analysis.

Sentiment dataset and data preparation

The sentiment dataset used in this study was obtained from the Kaggle repository titled "Analisis Sentimen Terkait Intensif Mobil Listrik," available at https://www.kaggle.com/datasets/billycemerson/analisis-sentimen-terkait-intensif-mobil-listrik. Although the repository title uses the term "intensif," the dataset context refers to public sentiment toward electric vehicle incentive policies in Indonesia. The original file used in this study was "Sentimen Insentif Mobil Listrik Indonesia," which includes comment identifiers, account names, dates, cleaned Indonesian text, and sentiment labels. After initial inspection, the dataset contained 1,517 public comments and five fields: id_komentar, nama_akun, tanggal, text_cleaning, and sentimen. The target label consisted of three sentiment classes: positive, negative, and neutral. Because the file already included a cleaned text field, the preprocessing stage focused on quality assurance rather than full raw-text reconstruction. Records with missing cleaned text were removed, duplicate comments were identified and eliminated, labels were standardized, and the class distribution was preserved during data splitting using stratified sampling.

Table 1. Description of dataset fields.

Column name	Description
id_komentar	Unique identifier for each public comment.

nama_akun	Account name associated with the public comment.
tanggal	Date of the public comment.
text_cleaning	Cleaned Indonesian comment text used as the input for sentiment classification.
sentimen	Sentiment label consisting of positive, neutral, or negative.

To provide a clearer view of the input data, Table 2 presents representative examples of cleaned Indonesian comments and their corresponding sentiment labels. The examples show that the dataset contains informal expressions, shortened words, mixed evaluative cues, and policy-related terms, which make the sentiment classification task challenging.

Table 2. Representative examples of comments in the sentiment dataset.

Sentiment label	Cleaned Indonesian comment	English translation
Positive	saran sih bikin harga ionic sama kayak brio insya alloh laris manis	A suggestion: make the Ioniq price similar to the Brio; God willing, it will sell very well.
Positive	baik kualitas kembang dulu baik kualitas motor motor pabrikan jepang	It would be better to improve the quality first, similar to motorcycles from Japanese manufacturers.
Positive	kampung sekarang banyak banget bocil sama cewek pake motor	In villages, many teenagers and women are now using motorcycles.
Neutral	subsidi tepat sasar	The subsidy should be well targeted.
Neutral	iklan kudu kencang	Promotion should be stronger.
Neutral	hybrid	hybrid
Negative	problem subsidi kualitas diturunin harga dinaikin usaha gitu cari cuan subsidi sebab inflasi paling gede	The problem is that subsidies may reduce quality and increase prices; this seems like a way to profit from subsidies and may worsen inflation.
Negative	model jelek kwalitas buruk harga mahal	The model is unattractive, the quality is poor, and the price is expensive.
Negative	syarat ngaco woy anak muda blom punya rumah blom jd umkm bukan serta kur dapet ngaco	The requirements are confusing; young people who do not yet own a house or business may not qualify.

The preprocessing procedure was conducted in several steps to ensure reproducibility. First, records with missing values in the text_cleaning field were removed because they could not be used as input for the sentiment classifier. Second, duplicate comments were identified based on identical cleaned-text content and eliminated to avoid repeated observations. Third, sentiment labels were standardized into three classes, namely positive, neutral, and negative. Fourth, all text was converted into lowercase form to reduce lexical variation. Fifth, excessive whitespace, repeated punctuation, URLs, user mentions, and non-informative symbols were removed. Sixth, informal Indonesian expressions and common slang terms were normalized where possible based on their standard lexical equivalents. Seventh, emoticons and emojis were removed when they did not provide explicit textual meaning, while sentiment-bearing expressions already reflected in the cleaned text were retained. Eighth, comments were checked for extremely short or empty content after cleaning. Finally, tokenization was performed using the tokenizer associated with the selected IndoBERT model, so that the final input representation followed the subword tokenization scheme required by the transformer architecture. No aggressive stopword removal was applied because transformer-based language models can use function words and contextual cues to capture sentiment polarity [8], [14], [19].

Deep learning sentiment model

Because the dataset was written in Indonesian and the objective was three-class sentiment classification, this study employed a deep learning text classifier using a fine-tuned IndoBERT-style transformer architecture with a softmax output layer. A transformer-based approach was selected because modern sentiment analysis research consistently shows that contextual embeddings outperform shallow feature engineering when comments are short, semantically compressed, and sensitive to context. This is particularly relevant for policy sentiment, where expressions may be indirect, sarcastic, or dependent on local linguistic nuance. Recent studies on product reviews, consumer comments, and EV-related online discourse have shown that deep learning and transformer-based sentiment models are suitable for capturing this complexity in downstream decision applications.

The model input consisted of the text_cleaning field. Each text instance was tokenized using the tokenizer corresponding to the selected transformer model, padded or truncated to a fixed maximum sequence length,

.

and encoded into contextual token embeddings. The classification head produced posterior probabilities for the three sentiment labels. During fine-tuning, the categorical cross-entropy loss was optimized using class-aware weighting to reduce the effect of class imbalance, as the negative class was more prevalent than the neutral class. This choice is aligned with recent sentiment classification research showing that weighted optimization improves robustness in imbalanced multiclass text settings.

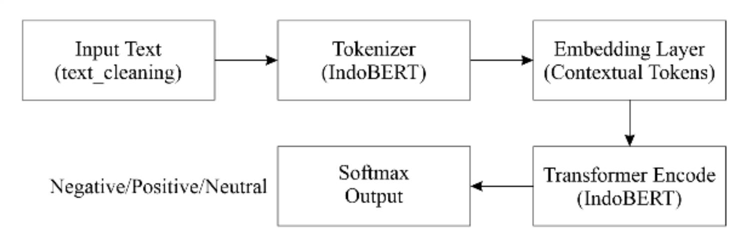

Figure 2. Layer-by-layer architecture of the IndoBERT-based sentiment classifier.

Figure 2 illustrates the layer-by-layer architecture of the IndoBERT-based sentiment classifier used in this study. The cleaned Indonesian comment is first converted into subword input tokens using the IndoBERT tokenizer. The tokenized input is then represented through token embeddings, segment embeddings, and positional embeddings. These embeddings are processed by 12 stacked transformer encoder layers with 12 self-attention heads and a hidden size of 768. The pooled representation of the [CLS] token is passed to a dropout layer and a fully connected classification head. Finally, the softmax layer produces probability scores for the three sentiment classes, namely negative, neutral, and positive.

IndoBERT fine-tuning configuration

The IndoBERT model was fine-tuned using the indobenchmark/indobert-base-p1 variant. This model follows the BERT-base architecture with 12 transformer encoder layers, 12 attention heads, and a hidden size of 768. The tokenizer used a vocabulary size of 50,000 and converted each input text into subword tokens. The maximum sequence length was set to 128 tokens.

Hyperparameter evaluation was conducted using validation macro-F1 as the model selection criterion because the dataset was imbalanced across sentiment classes. The best configuration used a learning rate of 3e-05, batch size of 16, 5 training epochs, dropout rate of 0.1, weight decay of 0.01, and warmup ratio of 0.1. The AdamW optimizer was used, and the random seed was fixed at 42 to support reproducibility.

To reduce the effect of class imbalance, balanced inverse-frequency class weights were applied in the cross-entropy loss function. The class weights were 0.5808 for negative, 3.5409 for neutral, and 1.0041 for positive. These weights gave greater penalty to errors in the minority neutral class while preventing the majority negative class from dominating the optimization process.

Table 3. Fine-tuning configuration of the IndoBERT sentiment classifier.

Hyperparameter	Value
Model variant	indobenchmark/indobert-base-p1
Model architecture	BertForSequenceClassification
Transformer encoder layers	12
Attention heads	12
Hidden size	768
Vocabulary size	50,000
Optimizer	AdamW
Learning rate	3e-05
Batch size	16
Training epochs	5
Maximum sequence length	128
Dropout rate	0.1
Weight decay	0.01
Warmup ratio	0.1
Class-weighting scheme	Balanced inverse-frequency class weights
Class weights	Negative = 0.5808; Neutral = 3.5409; Positive = 1.0041
Random seed	42
Model selection criterion	Highest validation macro-F1

The hyperparameter evaluation was performed using several candidate configurations. The evaluated configurations varied the learning rate, batch size, and maximum sequence length, while the number of epochs, dropout rate, weight decay, warmup ratio, optimizer, and class-weighting scheme were kept consistent. The candidate configurations included learning rates of 2e-05 and 3e-05, batch sizes of 8 and 16, and maximum sequence lengths of 128 and 256. The final configuration was selected based on the highest validation macro-F1 score, because macro-F1 is more appropriate than accuracy for imbalanced multiclass sentiment classification.

Table 4. Candidate hyperparameter configurations evaluated during model selection.

Run	Learning rate	Batch size	Epochs	Max sequence length	Validation macro-F1	Test macro-F1
1	2e-05	16	5	128	0.5467	0.5869
2	3e-05	16	5	128	0.5877	0.5228
3	2e-05	8	5	128	0.5438	0.5477
4	2e-05	16	5	256	0.5325	0.5419

Although Run 1 produced a higher test macro-F1 than Run 2, the final model was selected based on validation macro-F1 to avoid using the held-out test set for model selection. Therefore, Run 2 was retained as the final configuration because it achieved the highest validation macro-F1 score.

Data splitting, training protocol, and evaluation metrics

To ensure reliable model assessment, the dataset was divided into training, validation, and test subsets using stratified splitting. The main experimental protocol used an 80:10:10 partition for training, validation, and testing, respectively. The validation subset was used for model selection during hyperparameter evaluation, while the held-out test subset was used only for final performance reporting. The adoption of stratified splitting was motivated by the relatively limited dataset size and the imbalanced sentiment distribution. This strategy helped preserve the proportion of positive, neutral, and negative classes across the training, validation, and test subsets [8], [14], [19], [20].

Model performance was evaluated using accuracy, macro-precision, macro-recall, macro-F1, and the confusion matrix. Macro-averaged metrics were emphasized because they are more appropriate than raw accuracy in multiclass imbalanced settings. In addition, probability outputs were retained for post hoc aggregation into a sentiment evidence score. The evaluation framework follows current machine learning reporting practices that recommend class-balanced metrics, repeated validation, and transparent model comparison in applied decision-support studies [19], [21], [22].

Construction of deep learning–derived policy sentiment evidence

After the final sentiment model was trained, inference was performed on the full set of comments to obtain class probabilities for each observation. These outputs were then aggregated into a Policy Sentiment Evidence Index (PSEI). Let ppos, pneu, and pneg denote the predicted probabilities of positive, neutral, and negative sentiment for comment i, respectively. These probabilities were used to calculate the raw Policy Sentiment Evidence Index, as shown in Equation (1).

$$PSEI = \frac{1}{N}\sum_{i=1}^{N}(p_i^+ - p_i^-) \tag{1}$$

where n represents the total number of comments. The raw PSEI score was then linearly rescaled to the interval [0,1] so that it could be incorporated as a benefit-type criterion in the decision matrix. A higher PSEI value indicates more favorable public sentiment toward electric vehicle incentive policies, whereas a lower value indicates less favorable sentiment.

In this study, the PSEI was designed to represent the overall favorability of the public policy sentiment environment rather than to replace vehicle-specific technical attributes. The index summarizes public reactions to EV incentive policies and provides a contextual signal for downstream decision analysis. Therefore, the PSEI should be interpreted as policy-level sentiment evidence that complements engineering and economic criteria in the EV evaluation framework.

This formulation was chosen because it preserves uncertainty information from the classifier rather than collapsing all comments into hard labels. Such probability-aware aggregation is consistent with recent sentiment-informed DSS research, where text-derived evidence is transformed into structured decision signals rather than treated as purely descriptive output. It also provides a more defensible way to summarize public policy sentiment than simply counting positive and negative labels [17], [21], [23].

Integration with the CoCoSo-based decision support framework

.

The second stage of the framework used CoCoSo (Combined Compromise Solution) to evaluate EV alternatives based on a multi-criteria decision matrix [24]. In this study, the EV evaluation layer consisted of five criteria: price, driving range, battery capacity, charging time, and policy sentiment evidence. Price and charging time were treated as cost-type criteria, whereas driving range, battery capacity, and policy sentiment evidence were treated as benefit-type criteria. The inclusion of policy sentiment evidence was intended to represent the perceived favorability of the policy environment surrounding electric vehicle incentives. This treatment is consistent with recent MCDM and EV decision-support studies showing that EV evaluation can benefit from combining technical attributes, behavioral information, and intelligent decision-support mechanisms [7], [25], [26].

Because the sentiment dataset primarily reflects public reactions to EV incentive policies rather than direct reviews of each vehicle model, policy sentiment evidence was treated as a contextual criterion in the decision matrix. The PSEI captured the general policy sentiment environment, while alternative-level sentiment descriptors were used to differentiate how each EV alternative was positioned within that environment. This treatment allows policy sentiment to contribute to the CoCoSo ranking without assuming that all public comments directly evaluate each vehicle model. Thus, the proposed framework should be understood as a sentiment-informed decision support model rather than a purely product-review-based ranking system.

The initial decision matrix was defined as follows:

$$X = \left[x_{ij}\right]m \times n \tag{2}$$

where xij denotes the performance value of alternative i on criterion j, m represents the number of EV alternatives, and n represents the number of evaluation criteria.

To make all criteria comparable, the decision matrix was normalized. For benefit-type criteria, the normalized value was calculated as:

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{m} x_{ij}^2}} \tag{3}$$

For cost-type criteria, inverse-oriented normalization was applied as follows:

$$r_{ij} = \frac{1/x_{ij}}{\sqrt{\sum_{i=1}^{m} \left(1/x_{ij}\right)^2}} \tag{4}$$

where r_{ij} represents the normalized performance value of alternative i on criterion j.

After normalization, the additive comparability sequence S_i was calculated using the following equation:

$$S_i = \sum_{j=1}^{m} w_j r_{ij} \tag{5}$$

The multiplicative comparability sequence P_i was calculated as:

$$P_i = \sum_{j=1}^{n} \left(r_{ij}\right)^{w_j} \tag{6}$$

where w_j denotes the weight of criterion j, and the sum of all criterion weights is equal to one.

The CoCoSo method then calculates three compromise appraisal scores. The first appraisal score was calculated as:

$$k_{ia} = \frac{S_i + P_i}{\sum_{i=1}^{m} (S_i + P_i)} \tag{7}$$

The second appraisal score was calculated as:

$$k_{ib} = \frac{S_i}{min_i S_i} + \frac{P_i}{min_i P_i} \tag{8}$$

The third appraisal score was calculated as:

$$k_{ic} = \frac{\lambda S_i + (1-\lambda) P_i}{\lambda max_i S_i + (1-\lambda) max_i P_i} \tag{9}$$

where λ is a compromise coefficient. In this study, λ=0.5 was used to provide equal emphasis on the additive and multiplicative comparability sequences.

Finally, the overall CoCoSo appraisal score was calculated as:

$$k_i = (k_{ia} k_{ib} k_{ic})^{1/3} + \frac{1}{3}(k_{ia} k_{ib} k_{ic}) \tag{10}$$

The EV alternatives were then ranked in descending order based on the final k_i value. A higher k_i value indicates a more preferable electric vehicle alternative. Through this procedure, the proposed framework integrates technical, economic, and policy sentiment evidence into a unified multi-criteria decision support model.

Sensitivity analysis and ranking robustness

To assess the robustness of the ranking outcomes, a weight sensitivity analysis was performed by varying the weight assigned to the policy sentiment criterion across several scenarios. The purpose of this analysis was to examine whether the inclusion of policy sentiment evidence materially affected the stability of EV rankings. The policy sentiment weight was perturbed over a predefined range, while the remaining criteria weights were proportionally adjusted to preserve the total weight equal to one.

Rank consistency across scenarios was measured using Spearman's rho and Kendall's tau. These rank-correlation measures were used to compare the EV ranking under each sensitivity scenario with the baseline ranking. A stable ranking across different weight scenarios indicates that the proposed CoCoSo-based framework is not overly sensitive to reasonable changes in the importance assigned to policy sentiment evidence.

This procedure is important in MCDM-based decision support because hybrid models may appear methodologically sophisticated while contributing little to final ranking behavior if the added evidence is not structurally meaningful. Therefore, sensitivity analysis was used to evaluate whether policy sentiment functioned as a meaningful contextual modifier while maintaining ranking stability [9].

Explainability, reproducibility, and implementation

Although the sentiment model was deep learning-based, the framework was designed to remain interpretable at the decision-support level. Explainability was addressed in two ways. First, class-wise performance metrics and confusion patterns were reported for the sentiment classifier. Second, the effect of sentiment evidence on EV ranking was made transparent through explicit sensitivity scenarios and criterion-level reporting. Recent literature on trustworthy AI emphasizes that decision-support models should not be assessed only by predictive performance, but also by transparency, robustness, and usability for downstream decisions [27]–[29].

All analyses were implemented in Python. Data processing and numerical computation were performed using pandas and NumPy, while data splitting, class weighting, performance metrics, and confusion matrix evaluation were conducted using scikit-learn. IndoBERT fine-tuning was implemented using PyTorch and the Hugging Face Transformers library. The MCDM stage was implemented through matrix-based numerical computation. Random seeds were fixed for reproducibility, and all preprocessing rules, model settings, hyperparameter configurations, and evaluation steps were documented to reduce ambiguity in replication. Because the study used only publicly available comments and did not involve individual intervention or protected health information, formal ethical approval was not required.

The criterion weighting scheme used in the CoCoSo evaluation is visualized in Figure 3 to clarify the relative contribution of technical, economic, and contextual sentiment criteria in the decision-support framework.

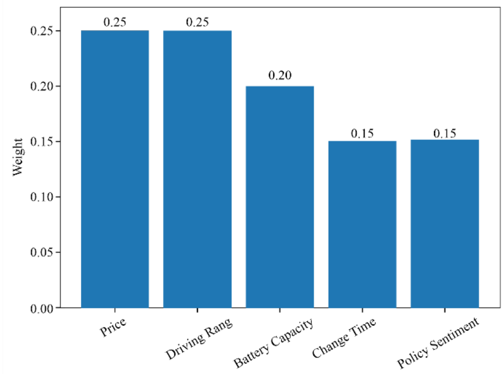

Figure 3. Criteria weight distribution used in the CoCoSo-based EV evaluation framework.

Figure 3 illustrates the relative importance assigned to the five criteria used in the CoCoSo-based electric vehicle evaluation framework, namely price, driving range, battery capacity, charging time, and policy sentiment. The distribution reflects the balance between technical, economic, and contextual dimensions in the proposed decision support model.

3. RESULTS AND DISCUSSIONS

Dataset characteristics

The sentiment dataset used in this study consisted of 1,514 comments related to electric vehicle (EV) incentive policies in Indonesia after data cleaning. The final dataset retained three sentiment classes, namely negative, positive, and neutral. The class distribution was imbalanced, with 869 negative comments (57.4%), 503 positive comments (33.2%), and 142 neutral comments (9.4%). This distribution indicates that public discourse on EV incentives was dominated by negative reactions during the observed period.

In addition to class imbalance, the dataset exhibited heterogeneous textual characteristics. The average comment length was 99.47 characters and 15.66 words, while the median values were 64 characters and 10 words, respectively. The longest comment contained 1,518 characters and 272 words, suggesting that the dataset included both short conversational responses and longer opinion-oriented statements. Only two duplicated cleaned-text records were identified, indicating that the corpus was generally diverse and suitable for supervised sentiment classification.

From a decision-support perspective, these characteristics are relevant for two main reasons. First, the dominance of negative sentiment suggests that public reactions to EV policies may influence downstream evaluation when sentiment evidence is incorporated into the decision model. Second, the class imbalance presents a realistic challenge for deep learning classification, particularly when assessing macro-level performance across minority classes.

The descriptive characteristics of the final sentiment dataset are summarized in Table 5, including the number of comments, class distribution, text-length statistics, and duplicate records after cleaning.

Table 5. Descriptive characteristics of the EV incentive sentiment dataset.

Characteristic	Value
Number of comments	1,514
Negative comments	869 (57.4%)
Positive comments	503 (33.2%)
Neutral comments	142 (9.4%)
Mean text length (characters)	99.47
Median text length (characters)	64
Mean text length (words)	15.66
Median text length (words)	10
Maximum text length (characters)	1,518
Maximum text length (words)	272
Duplicated cleaned-text records	2

Sentiment model performance

The IndoBERT-based sentiment classifier achieved a test accuracy of 0.6447, indicating that the model correctly classified approximately 64.47% of the test instances. Because the dataset was imbalanced and the neutral class contained substantially fewer observations than the negative and positive classes, macro-averaged metrics provide a more informative evaluation of model behavior across all sentiment categories. The model obtained a macro-precision of 0.5410, macro-recall of 0.5138, macro-F1 score of 0.5228, and weighted-F1 score of 0.6364.

These results suggest that the classifier was able to capture the dominant polarity structure of the dataset, particularly for the negative and positive classes, but its balanced performance across all three classes remained moderate. This pattern is consistent with the class distribution of the dataset, in which negative comments dominated the corpus while neutral comments represented the smallest class. The relatively lower macro-F1 score indicates that minority-class recognition, especially neutral sentiment, remained challenging.

From a methodological standpoint, the sentiment classifier was sufficient for constructing a probabilistic policy sentiment signal, but the results should not be interpreted as evidence of a fully optimized sentiment classification system. The main role of the deep learning module in this study was to transform public policy discourse into structured sentiment evidence for downstream decision support. Future studies may further improve classification performance by using larger datasets, richer annotation schemes, more extensive hyperparameter search, and comparative transformer models.

The classification performance of the IndoBERT-based sentiment model is reported in Table 6 using accuracy, macro-precision, macro-recall, macro-F1, and weighted-F1. These metrics provide both overall and class-balanced evaluation of the model performance on the test set.

Table 6. Performance of the deep learning sentiment classifier.

Metric	Value
Accuracy	0.6447
Macro-Precision	0.5410
Macro-Recall	0.5138
Macro-F1	0.5228
Weighted-F1	0.6364

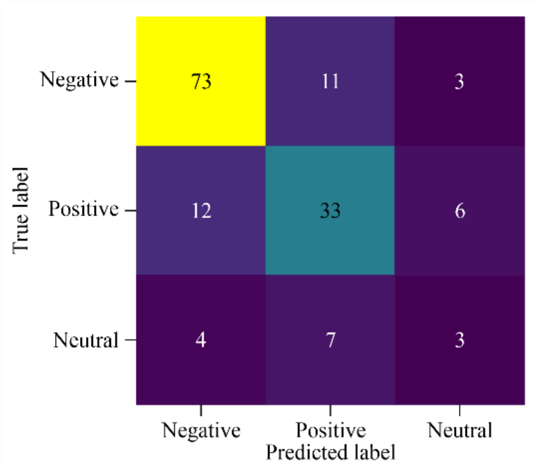

Figure 4. Confusion matrix of the three-class sentiment classifier.

Figure 4 presents the confusion matrix of the three-class sentiment classifier on the test dataset. The model correctly classified 66 negative comments, 3 neutral comments, and 29 positive comments. The strongest classification performance was observed in the negative class, with 66 out of 87 negative comments correctly predicted. The positive class also showed moderate recognition, with 29 out of 51 positive comments correctly classified. By contrast, the neutral class remained the most difficult category, with only 3 out of 14 neutral comments correctly predicted.

Misclassification mainly occurred between neutral and polarity-bearing comments, as well as between positive and negative comments with mixed evaluative cues. Several negative comments were predicted as positive, while several positive comments were predicted as negative, reflecting the semantic overlap between criticism, conditional support, and informal expressions in public discussions on EV incentive policies. Overall, the confusion pattern suggests that the model captured the dominant polarity structure of the dataset, but its ability to distinguish the minority neutral class remained limited.

Policy sentiment evidence

The outputs of the sentiment classifier were next aggregated to construct the Policy Sentiment Evidence Index (PSEI). Rather than relying on hard-label counts alone, the framework used the mean predicted probabilities of the three sentiment classes to preserve uncertainty information from the deep learning model. The average class probabilities across all comments were 0.3305 for positive sentiment, 0.0882 for neutral sentiment, and 0.5813 for negative sentiment. These values confirm that negative sentiment dominated the overall discourse captured in the dataset. Using the probability-based aggregation scheme, the raw PSEI value was -0.2508, which after rescaling to the interval [0, 1] yielded a final PSEI score of 0.3746. Because this value lies below the neutral midpoint of 0.5000, the result indicates that public sentiment toward EV incentive policies in Indonesia was generally unfavorable during the period covered by the dataset.

.

This finding is important for the decision-support perspective of the study. In conventional EV evaluation models, the decision environment is usually represented only by technical and economic criteria, such as price, range, charging time, and energy efficiency. However, the present results suggest that the broader policy context, as captured through public sentiment, may not be uniformly supportive and may therefore influence how EVs are perceived and prioritized in real-world decision settings. The policy sentiment evidence thus provides an additional layer of contextual intelligence that can complement the engineering-oriented attributes typically used in EV selection.

The aggregated probability outputs used to construct the Policy Sentiment Evidence Index are presented in Table 7, including the mean class probabilities, raw PSEI, and rescaled PSEI value.

Table 7. Aggregated policy sentiment evidence derived from the deep learning model.

Indicator	Value
Mean positive probability	0.3305
Mean neutral probability	0.0882
Mean negative probability	0.5813
Raw PSEI	-0.2508
Rescaled PSEI	0.3746

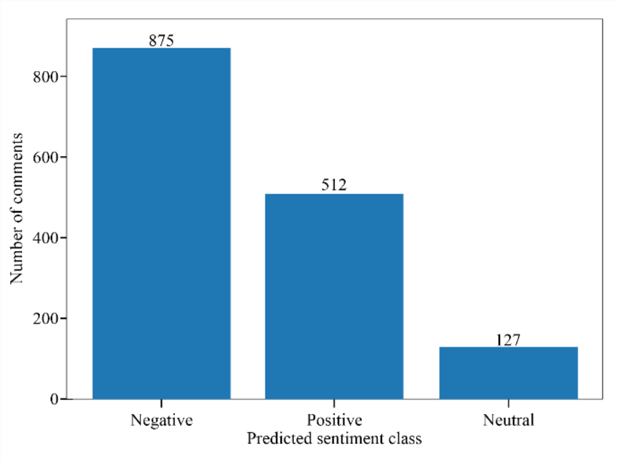

Figure 5. Distribution of predicted sentiment across the three classes.

The relatively low PSEI value implies that, during the observed period, public support for EV incentive policies was weaker than expected from a purely policy-driven transition narrative. This may reflect concerns related to affordability, unequal access to incentives, infrastructure limitations, or skepticism regarding implementation effectiveness. However, this interpretation should be limited to the public comments contained in the dataset and should not be generalized to the entire Indonesian population. From a policy analytics standpoint, this result strengthens the argument that EV evaluation in emerging markets should not rely solely on technical vehicle performance, but should also account for contextual acceptance signals derived from public discourse.

Implications for CoCoSo-based decision support

The CoCoSo-based decision support framework was applied to four EV alternatives, namely Tesla Model 3, Hyundai Ioniq 5, Nissan Leaf, and Wuling Air EV, using five evaluation criteria: price, driving range, battery capacity, charging time, and policy sentiment. To enable comparability across alternatives, interval-based vehicle specifications reported in the source comparison were converted into midpoint values. Meanwhile, the policy sentiment criterion was represented using alternative-level sentiment descriptors that were encoded into ordinal scores. This approach allowed policy sentiment evidence to vary across alternatives while remaining consistent with the overall public sentiment environment represented by the PSEI.

Before applying the CoCoSo procedure, the five evaluation criteria were organized into a normalized decision matrix consisting of technical, economic, and contextual policy sentiment indicators. The policy sentiment score was included as a benefit-type contextual criterion to reflect the perceived favorability of each EV alternative within the broader public policy sentiment environment. The normalized decision matrix used in the CoCoSo evaluation is presented in Table 8.

Table 8. Normalized decision matrix of EV alternatives used in the CoCoSo evaluation.

Alternative	Price	Driving Range	Battery Capacity	Charging Time	Policy Sentiment
Tesla Model 3	0.424	0.633	0.653	0.488	0.492
Hyundai Ioniq 5	0.391	0.557	0.632	0.558	0.492
Nissan Leaf	0.475	0.423	0.348	0.488	0.369
Wuling Air EV	0.665	0.330	0.232	0.460	0.615

As shown in Table 8, Tesla Model 3 achieved the highest normalized values for driving range and battery capacity, indicating strong technical performance. Hyundai Ioniq 5 also showed competitive performance, particularly in battery capacity and charging time. Wuling Air EV obtained the highest normalized values for price and policy sentiment, reflecting its affordability and favorable contextual perception. However, its lower driving range and battery capacity reduced its overall position in the CoCoSo aggregation. Nissan Leaf showed moderate performance across most criteria but did not dominate any criterion. These normalized values served as the basis for calculating the additive comparability sequence, multiplicative comparability sequence, and final CoCoSo appraisal score.

The final CoCoSo ranking results are presented in Table 9, including the additive comparability score, multiplicative comparability score, final appraisal score, and rank position for each electric vehicle alternative. Among the four evaluated alternatives, Tesla Model 3 achieved the highest final CoCoSo score (ki = 2.8172), followed by Hyundai Ioniq 5 (ki = 2.5702), Nissan Leaf (ki = 2.2668), and Wuling Air EV (ki = 1.3823). This ranking indicates that Tesla Model 3 offers the most balanced compromise across the selected technical, economic, and contextual criteria within the scope of the evaluated alternatives. Therefore, the ranking should be interpreted as a comparative decision-support result based on the selected criteria and alternatives, rather than as a universal market ranking of all electric vehicles.

Table 9. CoCoSo-based ranking results for EV alternatives incorporating policy sentiment evidence.

Alternative	S_i	P_i	k_i	Rank
Tesla Model 3	0.6623	4.5252	2.8172	1
Hyundai Ioniq 5	0.6940	3.8644	2.5702	2
Nissan Leaf	0.5327	3.6419	2.2668	3
Wuling Air EV	0.4000	2.0000	1.3823	4

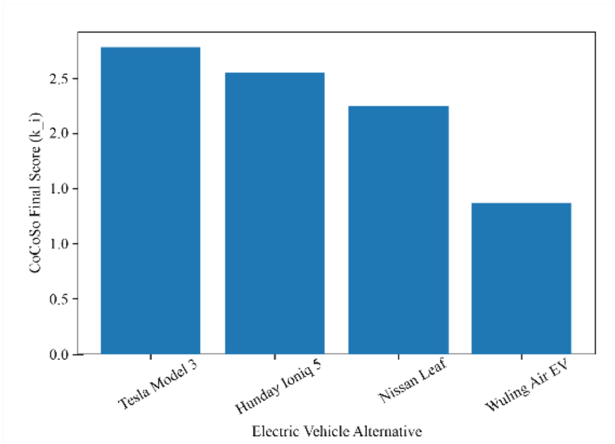

Figure 6. Final CoCoSo appraisal scores of electric vehicle alternatives.

The final CoCoSo appraisal scores are illustrated in Figure 6. The results show that Tesla Model 3 obtained the highest compromise score among the evaluated alternatives. Hyundai Ioniq 5 ranked second due to its strong charging performance and competitive driving range. Nissan Leaf placed third, while Wuling Air EV ranked fourth because of its relatively limited battery capacity and driving range despite its affordability.

The leading position of Tesla Model 3 can be attributed to its consistently strong performance across multiple criteria, particularly battery capacity, driving range, and overall technical efficiency. Although Hyundai Ioniq 5 achieved the best value in charging speed and remained highly competitive in driving range, its higher price reduced its final compromise advantage relative to Tesla Model 3. Nissan Leaf occupied the third position because it offered moderate technical performance and lower price than Tesla and Hyundai, but its weaker policy sentiment score and charging characteristics limited its overall ranking. Wuling Air EV ranked fourth, despite its affordability and very high policy sentiment score, because its substantially lower battery capacity, shorter driving range, and longer charging time reduced its overall performance under the multi-criteria framework.

These findings demonstrate that the proposed CoCoSo-based DSS does not reward a single dominant criterion in isolation. Instead, it identifies alternatives that perform consistently across a mix of cost, benefit, and contextual criteria. This is especially relevant in EV evaluation, where policy support alone may improve market attractiveness, but cannot fully compensate for substantial technical disadvantages.

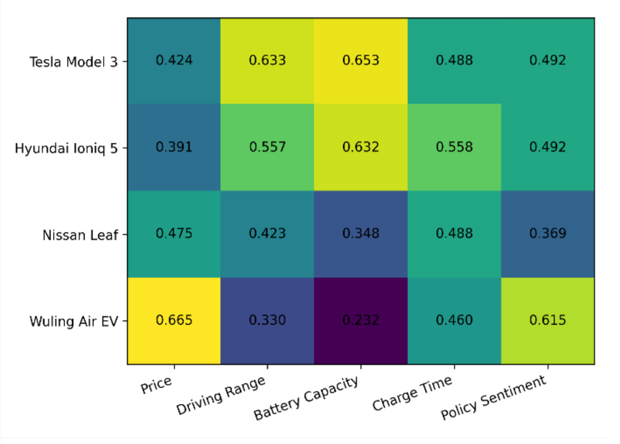

Figure 7. Normalized decision matrix visualization.

Figure 7 presents the normalized values of the EV decision matrix across the five evaluation criteria. Cost criteria were transformed using inverse-oriented normalization, while benefit criteria were normalized using vector normalization. The visualization highlights the relative performance distribution of Tesla Model 3, Hyundai Ioniq 5, Nissan Leaf, and Wuling Air EV under the proposed CoCoSo-based decision support framework.

Sensitivity and managerial implications

Recent EV decision-support studies also emphasize that EV-related decisions should account for local characteristics, infrastructure readiness, and contextual implementation conditions [30]. To assess the robustness of the ranking outcomes, this study performed a sensitivity analysis by varying the weight assigned to the policy sentiment criterion from 0.10 to 0.40. As shown in Table 10, the resulting rankings remained unchanged across all tested scenarios. Both Spearman's rho and Kendall's tau were equal to 1.000 for all comparisons relative to the baseline scenario, and no rank changes were observed.

The unchanged ranking across all sensitivity scenarios indicates that the proposed framework is stable under reasonable changes in the policy sentiment weight. However, this result also suggests that technical and economic criteria remained the dominant determinants of the final ranking. In this sense, policy sentiment evidence functioned as a contextual modifier rather than a disruptive ranking factor. This is substantively meaningful because public sentiment can enrich the decision model without making the ranking overly dependent on volatile online discourse.

Compared with previous EV evaluation studies that mainly relied on technical and economic criteria, the present findings show that policy sentiment can be incorporated as contextual decision evidence without destabilizing the final ranking [3], [4], [25], [26].

The main substantive contribution of this study lies in the transformation of unstructured policy discourse into structured evidence for decision support. While previous EV sentiment studies have mainly focused on describing public opinion, and previous MCDM-based EV studies have mainly emphasized technical and economic attributes, the proposed framework connects these two perspectives. This integration allows EV evaluation to reflect not only what a vehicle offers technically, but also how the surrounding policy environment may shape public acceptance and decision relevance.

From a managerial perspective, the findings imply that technical and economic criteria remain the primary drivers of EV ranking, while policy sentiment functions as a meaningful but non-disruptive contextual modifier. For manufacturers, this means that positive policy sentiment may strengthen market attractiveness, but it cannot substitute for weak technical performance, limited driving range, or poor charging convenience. For policymakers, the results suggest that improving public sentiment toward EV incentives may enhance adoption conditions, but such efforts should be accompanied by clear policy communication, better affordability mechanisms, and infrastructure readiness. Therefore, policy sentiment should be understood as a complementary decision signal rather than a standalone determinant of EV competitiveness.

More broadly, the results support the value of integrating deep learning sentiment evidence into multi-criteria EV decision support. Rather than treating public discourse as an external narrative, the proposed framework formalizes policy sentiment as quantifiable decision evidence. This contributes to a more realistic evaluation environment, particularly in emerging markets where consumer trust in policy design may influence adoption alongside vehicle specifications.

The sensitivity analysis results are summarized in Table 10, showing how changes in the policy sentiment weight affect Spearman's rho, Kendall's tau, and the number of rank changes relative to the baseline ranking.

Table 10. Sensitivity analysis under different policy sentiment weighting scenarios.

Scenario	Policy sentiment weight	Spearman's rho	Kendall's tau	Number of rank changes
Low sentiment weight	0.10	1.000	1.000	0
Baseline scenario	0.20	1.000	1.000	0
Medium-high sentiment weight	0.30	1.000	1.000	0
High sentiment weight	0.40	1.000	1.000	0

4. CONCLUSION

This study proposed a hybrid decision-support framework that integrates deep learning–based sentiment analysis with the CoCoSo multi-criteria decision-making method for evaluating electric vehicle alternatives. By transforming public policy sentiment into a quantitative decision criterion, the framework offers a more context-aware approach for incorporating societal perspectives into technology evaluation.

The results show that the IndoBERT-based sentiment classifier was able to capture the main polarity structure of public comments on electric vehicle incentive policies, although the class-balanced classification performance remained moderate. The model achieved an accuracy of 0.6447 and a macro-F1 score of 0.5228 on the test set. The probability-based aggregation produced a rescaled PSEI value of 0.3746, indicating that public sentiment toward EV incentives in the observed dataset was relatively unfavorable. The CoCoSo evaluation ranked Tesla Model 3 as the most preferable alternative among the four evaluated EVs, followed by Hyundai Ioniq 5, Nissan Leaf, and Wuling Air EV. Sensitivity analysis further confirmed that the ranking remained stable under different policy sentiment weighting scenarios.

These findings suggest that sentiment-driven evidence can enrich multi-criteria decision support by adding behavioral and policy-sensitive dimensions to conventional technical evaluation. The study contributes to EV decision-support research by demonstrating how public policy sentiment can be converted into structured decision evidence rather than being treated only as descriptive context. Nevertheless, this study is limited by the size and scope of the sentiment dataset, the moderate performance of the sentiment classifier, the use of selected EV alternatives, and the contextual nature of the policy sentiment score. Future research may extend the framework using larger multi-source public discourse datasets, alternative-specific sentiment extraction, comparative deep learning models, additional technical and economic criteria, and other MCDM methods to further validate the robustness and generalizability of the proposed approach.

ACKNOWLEDGMENTS

The authors would like to thank Universitas Pancasakti Tegal for academic support during the preparation of this research. The authors have reviewed, validated, and take full responsibility for the accuracy, integrity, and final content of the manuscript.

AUTHOR CONTRIBUTIONS

Nur Tulus Ujianto: Conceptualization, methodology, sentiment modeling, data curation, formal analysis, writing – original draft preparation. Gunawan Gunawan: Conceptualization, supervision, validation, project administration, writing – review and editing. Gunawan Raharjo: CoCoSo modeling, decision-support framework development, validation, writing – review and editing. Didiek Trisatya: Data preprocessing, visualization, software implementation, formal analysis. Diajeng Tyas Purwa Hapsari: Literature review, interpretation of results, manuscript refinement, writing – review and editing.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The sentiment dataset used in this study was obtained from the publicly available Kaggle repository titled “Analisis Sentimen Terkait Intensif Mobil Listrik.” The original file used in the analysis was “Sentimen Insentif Mobil Listrik Indonesia.” The processed dataset and analysis scripts can be made available by the corresponding author upon reasonable request.

REFERENCES

[1] D. F. Hakam dan S. Jumayla, “Electric vehicle adoption in Indonesia: Lesson learned from developed and developing countries,” *Sustain. Futur.*, vol. 8, hal. 100348, 2024, doi: https://doi.org/10.1016/j.sftr.2024.100348.

[2] N. C. Kresnanto dan W. H. Putri, “Subsidies for electric vehicles as a form of green transportation: Evidence from Indonesia,” *Transp. Res. Interdiscip. Perspect.*, vol. 27, hal. 101230, 2024, doi: https://doi.org/10.1016/j.trip.2024.101230.

[3] F. Ecer, “A consolidated MCDM framework for performance assessment of battery electric vehicles based on ranking strategies,” *Renew. Sustain. Energy Rev.*, vol. 143, hal. 110916, 2021, doi: https://doi.org/10.1016/j.rser.2021.110916.

[4] Z. Tian, H. Liang, R. Nie, X. Wang, dan J. Wang, “Data-driven multi-criteria decision support method for electric vehicle selection,” *Comput. Ind. Eng.*, vol. 177, hal. 109061, 2023, doi: https://doi.org/10.1016/j.cie.2023.109061.

[5] P. A. Alvarez, A. Ishizaka, dan L. Martínez, “Multiple-criteria decision-making sorting methods: A survey,” *Expert Syst. Appl.*, vol. 183, hal. 115368, 2021, doi: https://doi.org/10.1016/j.eswa.2021.115368.

[6] A. S. Yalcin, H. S. Kilic, dan D. Delen, “The use of multi-criteria decision-making methods in business analytics: A comprehensive literature review,” *Technol. Forecast. Soc. Change*, vol. 174, hal. 121193, 2022, doi: https://doi.org/10.1016/j.techfore.2021.121193.

[7] Y. Li, X. He, L. Martínez, J. Zhang, D. Wang, dan X. A. Liu, “Comparative analysis of three categories of multi-criteria decision-making methods,” *Expert Syst. Appl.*, vol. 238, hal. 121824, 2024, doi: https://doi.org/10.1016/j.eswa.2023.121824.

[8] S. A. Nugroho dan S. Widianto, “Exploring electric vehicle adoption in Indonesia using zero-shot aspect-based sentiment analysis,” *Sustain. Oper. Comput.*, vol. 5, hal. 191–205, 2024, doi: https://doi.org/10.1016/j.susoc.2024.08.002.

[9] D. A. Purnama *et al.*, “Data-driven public policy for electric vehicles (EV) through open innovation and dynamic consumer preferences: A time-series social media analysis using integrated IPA-product improvability model,” *J. Open Innov. Technol. Mark. Complex.*, vol. 11, no. 3, hal. 100583, 2025, doi: https://doi.org/10.1016/j.joitmc.2025.100583.

[10] Q. Qin, Z. Zhou, J. Zhou, Z. Huang, X. Zeng, dan B. Fan, “Sentiment and attention of the Chinese public toward electric vehicles: A big data analytics approach,” *Eng. Appl. Artif. Intell.*, vol. 127, hal. 107216, 2024, doi: https://doi.org/10.1016/j.engappai.2023.107216.

[11] H. Wang, Y.-J. Xin, M. Deveci, W. Pedrycz, Z. Wang, dan Z.-S. Chen, “Leveraging online reviews and expert opinions for electric vehicle type prioritization,” *Comput. Ind. Eng.*, vol. 197, hal. 110579, 2024, doi: https://doi.org/10.1016/j.cie.2024.110579.

[12] Z. Yang, Q. Li, V. Charles, B. Xu, dan S. Gupta, “Supporting personalized new energy vehicle purchase decision-making: Customer reviews and product recommendation platform,” *Int. J. Prod. Econ.*, vol. 265, hal. 109003, 2023, doi: https://doi.org/10.1016/j.ijpe.2023.109003.

[13] P. Wu, T. Tang, L. Zhou, dan L. Martínez, “A decision-support model through online reviews: Consumer preference analysis and product ranking,” *Inf. Process. Manag.*, vol. 61, no. 4, hal. 103728, 2024, doi: https://doi.org/10.1016/j.ipm.2024.103728.

[14] B. Gong, R. Liu, X. Zhang, C.-T. Chang, dan Z. Liu, “Sentiment analysis of online reviews for electric vehicles using the SMAA-2 method and interval type-2 fuzzy sets,” *Appl. Soft Comput.*, vol. 147, hal. 110745, 2023, doi: https://doi.org/10.1016/j.asoc.2023.110745.

[15] A. Dey dan M. Jenamani, “Aspect based sentiment analysis of consumer reviews using unsupervised attention neural framework,” *Appl. Soft Comput.*, vol. 167, hal. 112259, 2024, doi: https://doi.org/10.1016/j.asoc.2024.112259.

[16] Y. C. Hua, P. Denny, J. Wicker, dan K. Taskova, “A systematic review of aspect-based sentiment analysis: domains, methods, and trends,” *Artif. Intell. Rev.*, vol. 57, no. 11, hal. 296, 2024, doi: 10.1007/s10462-024-10906-z.

[17] A. Abdulla, G. Baryannis, dan I. Badi, “An integrated machine learning and MARCOS method for supplier evaluation and selection,” *Decis. Anal. J.*, vol. 9, hal. 100342, 2023, doi: https://doi.org/10.1016/j.dajour.2023.100342.

[18] S. K. Jauhar, S. Sethi, S. S. Kamble, S. Mathew, dan A. Belhadi, “Artificial intelligence and machine learning-based decision support system for forecasting electric vehicles’ power requirement,” *Technol. Forecast. Soc. Change*, vol. 204, hal. 123396, 2024, doi: https://doi.org/10.1016/j.techfore.2024.123396.

[19] A. Daza, N. D. González Rueda, M. S. Aguilar Sánchez, W. F. Robles Espíritu, dan M. E. Chauca Quiñones, “Sentiment Analysis on E-Commerce Product Reviews Using Machine Learning and Deep Learning Algorithms: A Bibliometric Analysis, Systematic Literature Review, Challenges and Future Works,” *Int. J. Inf. Manag. Data Insights*, vol. 4, no. 2, hal. 100267, 2024, doi: https://doi.org/10.1016/j.jjimei.2024.100267.

[20] B. Selvakumar dan B. Lakshmanan, “Sentimental analysis on user’s reviews using BERT,” *Mater. Today Proc.*, vol. 62, hal. 4931–4935, 2022, doi: https://doi.org/10.1016/j.matpr.2022.03.678.

[21] N. Darraz, I. Karabila, A. El-Ansari, N. Alami, dan M. El Mallahi, “Integrated sentiment analysis with BERT for enhanced hybrid recommendation systems,” *Expert Syst. Appl.*, vol. 261, hal. 125533, 2025, doi: https://doi.org/10.1016/j.eswa.2024.125533.

[22] I. Tougui, A. Jilbab, dan J. El Mhamdi, “Impact of the Choice of Cross-Validation Techniques on the Results of Machine Learning- Based Diagnostic Applications,” *Healthc. Inform. Res.*, vol. 27, no. 3, hal. 189–199, 2021, doi: https://doi.org/10.4258/hir.2021.27.3.189.

[23] X. Jia, M. Xu, C. Xu, dan P. Zhu, “An online review-driven decision-making method for electric vehicle selection based on satisfaction analysis,” *Expert Syst. Appl.*, vol. 300, hal. 130518, 2026, doi: https://doi.org/10.1016/j.eswa.2025.130518.

[24] M. Yazdani, P. Zarate, E. Kazimieras Zavadskas, dan Z. Turskis, “A combined compromise solution (CoCoSo) method for multi-criteria decision-making problems,” *Manag. Decis.*, vol. 57, no. 9, hal. 2501–2519, Nov 2018, doi: 10.1108/MD-05-2017-0458.

[25] Y. Song, Y. Li, H. Zhu, dan G. Li, “A decision support model for buying battery electric vehicles considering consumer learning

and psychological behavior," *J. Retail. Consum. Serv.*, vol. 73, hal. 103303, 2023, doi: https://doi.org/10.1016/j.jretconser.2023.103303.
[26] A. İ. Şimşek, Y. E. Gür, dan E. Ünal, "Innovative MCDM-ML algorithms-based decision-support system for electric vehicle selection," *Environ. Dev. Sustain.*, 2025, doi: 10.1007/s10668-025-06476-x.
[27] M. Kim, H. Sohn, S. Choi, dan S. Kim, "Requirements for trustworthy artificial intelligence and its application in healthcare," *Healthc. Inform. Res.*, vol. 29, no. 4, hal. 315–322, 2023, doi: https://doi.org/10.4258/hir.2023.29.4.315.
[28] S. S Band *et al.*, "Application of explainable artificial intelligence in medical health: A systematic review of interpretability methods," *Informatics Med. Unlocked*, vol. 40, hal. 101286, 2023, doi: https://doi.org/10.1016/j.imu.2023.101286.
[29] I. D. Mienye *et al.*, "A survey of explainable artificial intelligence in healthcare: Concepts, applications, and challenges," *Informatics Med. Unlocked*, vol. 51, hal. 101587, 2024, doi: https://doi.org/10.1016/j.imu.2024.101587.
[30] N. Kunicina *et al.*, "Decision-Making System for Electric Vehicle Management by Integrating Smart Technologies and Local Characteristics," *Applied Sciences*, vol. 14, no. 23. hal. 11150, 2024, doi: https://doi.org/10.3390/app142311150.

.