SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15314
e-ISSN : 2541-2019
p-ISSN : 2541-044X

A Hybrid PULTS–SWARA–ELECTRE-I Model for Multi-Criteria Political Sentiment Classification on Indonesian Twitter Data

Taufik Yuandika Bahri [1)*], Chairini Aisyah [2)]
[1)] Department of Digital Business, Institut Modern Arsitektur dan Teknologi, Indonesia
[2)] Department of Visual Communication Design and Architecture, Institut Modern Arsitektur dan Teknologi, Indonesia
[1)] taufikyuandikabahri@gmail.com, [2)]chariniaisyah@hotmail.com

Submitted :Sep 6, 2025 | **Accepted** : Oct 10, 2025 | **Published** : Oct 15, 2025

Abstract: Social media platforms such as Twitter have become crucial for analyzing political sentiment, particularly in contexts where public opinion shifts rapidly. This study proposes a hybrid classification model that combines Probabilistic Uncertain Linguistic Term Set (PULTS), Stepwise Weight Assessment Ratio Analysis (SWARA), and ELimination Et Choice Translating REality (ELECTRE-I). Using a dataset of 7,800 tweets collected between January and July 2024 covering five major political parties in Indonesia, the model classifies tweets into positive, negative, and neutral sentiments. To address class imbalance, Easy Data Augmentation (EDA) was applied, while Term Frequency–Inverse Document Frequency (TF-IDF) was used for feature extraction. The results show that the proposed model achieves 90% accuracy and an F1-score of 85%, outperforming baseline methods such as SVM (86.7%), Naïve Bayes (83.3%), Decision Tree (88%), and K-Means (76.7%). These improvements demonstrate that the integration of linguistic uncertainty with expert-driven feature weighting provides measurable advantages in political sentiment classification. Beyond performance, the study contributes theoretically by extending multi-criteria decision-making methods into sentiment analysis and by offering a more interpretable alternative to opaque machine learning models. Together, these findings highlight the practical value of explainable decision frameworks for political communication while advancing methodological approaches for analyzing sentiment under uncertainty.

Keywords: Sentiment Classification; SWARA; ELECTRE; PULTS; Feature Selection

INTRODUCTION

The widespread use of social media has revolutionized the way political discourse is conducted, analyzed, and responded to(Gearhart et al., 2024). Platforms like Twitter (now known as X) allow individuals to express their political opinions in real-time, shaping public perception and influencing voter behavior(Alodat et al., 2023). In Indonesia, with its multiparty political system and high levels of digital engagement, Twitter has become a critical channel for expressing political support, criticism, and mobilization(Masduki & Wendratama, 2025).

This presents both an opportunity and a challenge for researchers seeking to extract meaningful sentiment from massive volumes of user-generated content that can be used as a consideration in choosing political party(N et al., 2025). Analyzing political sentiment from Twitter data involves significant challenges(Ansari et al., 2020). Tweets are typically short, informal, and noisy, containing slang, sarcasm, abbreviations, and multilingual expressions(Albladi et al., 2025). These characteristics make traditional sentiment analysis approaches difficult to apply directly(Mao et al., 2024). Lexicon-based methods, for example, rely on sentiment dictionaries to determine the polarity of words and phrases(Ojeda-Hernández et al., 2023). While they are interpretable and easy to implement, they struggle with context sensitivity, negation handling, and domain-specific expressions(Zaoui Seghroucheni et al., 2025).

This limitation often leads to reduced accuracy, especially in political contexts where implicit and ironic language is common(Puhacheuskaya & Järvikivi, 2022). Supervised machine learning models(Patrick et al., 2023)—such as Naïve Bayes(Fitri et al., 2019), Logistic Regression, and Support Vector Machines(Jaya Hidayat

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15314
e-ISSN : 2541-2019
p-ISSN : 2541-044X

et al., 2022)—represent another popular approach. These models use statistical learning from labeled datasets to perform classification(Wang et al., 2023). Although they outperform lexicon-based methods in many cases(Hill et al., 2025), they rely heavily on balanced, high-quality annotated data and require significant feature engineering(Mumuni & Mumuni, 2025). Furthermore, they treat sentiment classification as a crisp decision-making problem, failing to account for linguistic vagueness or uncertainty(Zaitseva et al., 2023).

More recently, deep learning models like Convolutional Neural Networks (CNN)(Meena et al., 2023), Long Short-Term Memory networks (LSTM)(Muhammad et al., 2021), and transformer-based architectures like BERT have achieved state-of-the-art results in sentiment analysis(Smairi et al., 2024). These models automatically learn hierarchical representations of text and can capture syntactic and semantic relationships(Onan, 2023). However, they also present critical drawbacks: they are data-hungry(Chia et al., 2022), computationally expensive(Oluwasakin et al., 2023), and often function as "black boxes"—offering little transparency in how predictions are made(Hassija et al., 2024). In high-stakes domains like political opinion mining, such lack of interpretability poses risks for trust, bias detection, and policy decision-making(Murikah et al., 2024).

To overcome the limitations of these traditional and modern approaches, this study adopts a multi-criteria decision-making (MCDM) framework that integrates linguistic modeling and expert-driven reasoning(Bueno et al., 2021). Specifically, we propose a hybrid classification method that combines the Stepwise Weight Assessment Ratio Analysis (SWARA), the ELimination Et Choice Translating REality (ELECTRE-I) method, and the Probabilistic Uncertain Linguistic Term Set (PULTS) environment. Each of these methods addresses a specific limitation in existing sentiment classification techniques. SWARA allows domain experts to determine the relative importance of different sentiment attributes—such as emotional tone, contextual strength, or topic relevance—through a structured, sequential weighting process(Sivageerthi et al., 2022). Unlike statistical models, SWARA enables the integration of expert judgment, which is particularly valuable when working with subjective or evolving concepts like political sentiment(Singer & Özşahin, 2025).

ELECTRE-I introduces an outranking mechanism for comparing and evaluating alternatives based on concordance and discordance indices(Fahmi et al., 2016). Rather than producing a single utility score, ELECTRE identifies which alternatives are preferable or incomparable based on multiple criteria(Ruan et al., 2024). This approach is more flexible and realistic in scenarios where trade-offs between attributes exist, and where some options cannot be clearly ranked(Vahdani et al., 2013). However, both SWARA and ELECTRE-I were originally developed for deterministic or crisp inputs(Akram & Ilyasa, 2025). This is where PULTS plays a critical role. The Probabilistic Uncertain Linguistic Term Set allows sentiment attributes to be represented as intervals of linguistic terms (e.g., “somewhat good”, “maybe poor”) with associated probabilities(Naz et al., 2025).

This framework captures the natural vagueness and hesitation often found in social media expressions and allows the model to process uncertain information in a structured and mathematically grounded way. The integration of these three methods—SWARA for weighting, ELECTRE-I for decision ranking, and PULTS for uncertainty modeling—offers a comprehensive solution that overcomes the shortcomings of other techniques. Compared to machine learning or deep learning approaches, our hybrid model is: More interpretable, since decisions are traceable through expert-defined weights and linguistic evaluations. More adaptable, because it does not require large labeled datasets or retraining when the political context changes. More aligned with real-world communication, especially in political discourse, where opinions are rarely absolute and often probabilistic.

The choice of the PULTS–SWARA–ELECTRE-I hybrid model was driven by the need to address specific limitations in traditional sentiment analysis approaches. PULTS (Probabilistic Uncertain Linguistic Term Set) allows for handling linguistic uncertainty, which is often present in political discourse. SWARA (Stepwise Weight Assessment Ratio Analysis) facilitates expert-driven feature weighting, enabling the model to incorporate domain-specific knowledge, particularly crucial in political contexts. ELECTRE-I (ELimination Et Choice Translating REality) enhances decision-making by evaluating alternatives based on multiple criteria, making it suitable for analyzing the multi-dimensional and complex nature of political sentiment. This combination of methods offers a more robust, interpretable, and adaptable solution than purely machine learning-based approaches, making it ideal for nuanced political sentiment analysis. The research gap in sentiment analysis for political discourse lies in the lack of models that can effectively capture the inherent uncertainty and context sensitivity of social media expressions. While machine learning models have achieved high accuracy, they often fail to provide transparent reasoning or account for the probabilistic nature of sentiment. This study fills this gap by proposing a hybrid approach that integrates uncertainty modeling, expert knowledge, and multi-criteria decision-making to create a more interpretable and adaptable solution for political sentiment classification.

This study is the first to integrate MCDM methods (PULTS–SWARA–ELECTRE-I) into the political sentiment analysis of Indonesian Twitter data, addressing challenges of both uncertainty and interpretability. This study proposes a five-step framework: (1) data crawling of political tweets using keywords via the Twitter API, (2) preprocessing using the Sastrawi library for Indonesian-language normalization, (3) data augmentation using Easy Data Augmentation (EDA) and feature selection using TF-IDF, (4) classification using the SWARA–ELECTRE–PULTS hybrid model, and (5) evaluation using common performance metrics such as accuracy,

*name of corresponding author

precision, recall, and F1-score. This research contributes a novel, interpretable, and linguistically-grounded approach to sentiment classification for political discourse in Indonesia. It bridges the gap between traditional sentiment analysis and explainable decision support systems—advancing the field of political opinion mining under conditions of uncertainty.

LITERATURE REVIEW

Related Work

The following are several previous studies related to this research that have been conducted can be seen in Table 1.

Table 1. Related Work

Study	Method	Limitations
(Mishra et al., 2023)	Extended DNMA integrated with MEREC–SWARA for objective and subjective weighting under a Single-Valued Neutrosophic Set (SVNS) environment.	Focus on industrial location selection; does not address NLP or sentiment analysis; high complexity in criteria integration.
(Zhang & Song, 2024)	Text mining (LDA, TextRank, Word2Vec) combined with Large Group Decision Making (LGDM).	Big data not always representative; focus on product improvement, not political sentiment classification; limited interpretability.
(Huang et al., 2025)	Linking words-guided data augmentation combined with adversarial contrastive training (LWEDA-ACT).	Requires large labeled data; complex implementation; limited to ABSA domain, not multi-criteria decision-making.
(Naz et al., 2025)	Integration of SWARA–ELECTRE-I with Probabilistic Uncertain Linguistic Term Set (PULTS) for feature selection in image recognition.	Domain limited to image recognition; not tested on textual data; potential bias from expert weighting.
Proposed Method	Integration of PULTS for handling linguistic uncertainty, SWARA for expert weighting, and ELECTRE-I for outranking; dataset of 7,800 political tweets; preprocessing with Sastrawi + EDA.	Limited to Indonesian Twitter data from 2024; uses TF-IDF instead of contextual embeddings; subjectivity in expert weighting remains.

PULTS

The Probabilistic Uncertain Linguistic Term Set (PULTS) is a decision making tool that combines linguistic terms (such as ”high”, ”medium”, ”low”) with associated probabilities to represent uncertainty. It allows for the evaluation of alternatives under uncertainty by considering both qualitative linguistic terms and their associated likelihood. PULTS concepts are as follows(Xie et al., 2018).

1. PULTS uses a set of linguistic terms $S = \{\propto(\rho)|\rho = -\delta, \dots, \delta\}$, where each term $\alpha(\rho)$ represents a qulatitative description, suchas as “low”, “medium”, etc. These terms reflect a range of values in decision-making.
2. Each linguistic term is associated with a probability, $\rho(k)$, indicating the degree of certaiunty or uncertainty about that term.
3. If the sum of probabilities is less than 1, then normalize the probabilities to ensure consistency
4. Integrate the opinions of multiple decision-makers, PULTS uses weighted average operator. This method calculates a weighted average of the linguistic terms and their probabilities:
$$WA(Z_1(\rho), Z_2(\rho), \dots, Z_n(\rho) = \sum_{t=1}^{n} Z_t(\rho)\, w_t \quad (1)$$
Where w_t is the weight associated with each term $Z_t(p)$, and n is the total number of terms being considered.
5. The final decision is made based on the integrated result. This help decision-makers evaluate alternatives considering both the uncertainty in the linguistic terms and the probabilities associated with them.

ELECTRE

The ELECTRE method encompasses a group of outranking approaches that work by comparing alternatives in pairs and ranking the less favorable ones in order to identify the best option. ELECTRE concepts are as follows(Zahid et al., 2022).

1. Let $A = \{a_1, \dots, a_m\}$ be the set of alternatives and $C = \{1, \dots, n\}$ the index set of criteria. Each criterion $i \in C$ has weight $w_i \geq 0$ with $\sum_{i=1}^{n} w_i = 1$. Th performance of a_j on criterion i is $g_i(a_j)$ (larger is better)

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15314

e-ISSN : 2541-2019
p-ISSN : 2541-044X

2. For each ordered pair (a_j, a_k), define the concordance-citeria set
3. Let $v_i > 0$ be the veto threshold for criterion i. Define per-criterion and aggregate discordance
4. Outranking relation and exploitation
5. Given the indifference and preference thresholds $q_1 \geq 0$ and $p_i \geq q_i$ for each criterion i, define the partial support factor

$$\emptyset_I^{JK} = \begin{cases} 1, \; g_i(a_j) + q_i \geq g_i(a_k) \\ \frac{g_i(a_j)+p_i-g_i(a_k)}{p_i-q_i}, g_i(a_j) + q_i < g_i(a_k) \leq g_i(a_j) + p_i \\ 0, g_i(a_k) > g_i(a_j) + p_i \end{cases} \quad (2)$$

Then the concordance becomes

$$C(a_j, a_k) = \sum_{i=1}^{n} \emptyset_i^{jk} w_i \quad (3)$$

While the concordance test can still be enforced with the same $D(a_j, a_k)$ criterion

SWARA

In the SWARA method, the relative significance of each factor is determined based on the input provided by decision makers. By evaluating these factors, the method helps in identifying and selecting the most important one. SWARA concepts are as follows(Sivageerthi et al., 2022).

1. Initial prioritization of factors
   The evaluation factors (criteria) are collected from experts or decision makers. These factors are ranked according to their relative importance, starting from the most important to the least important.
2. Determination of the coefficient (G_j)
3. The initial weights are calculated iteratively using:

$$B_j = \begin{cases} 1, j = 1 \\ \frac{B_{j-1}}{G_j}, j > 1 \end{cases} \quad (4)$$

4. The relative weights are obtained by normalizing the initial weights:

$$V_j = \frac{B_j}{\sum_{k=1}^{m} B_k}, j = 1,2, \dots, m \quad (5)$$

5. Finally, the factors are ranked according to their relative weights V_j. The factor with the highest V_j is considered the most influential, and the rest follow in descending order.

METHOD

An outline of the research methodology can be seen in Figure 1.

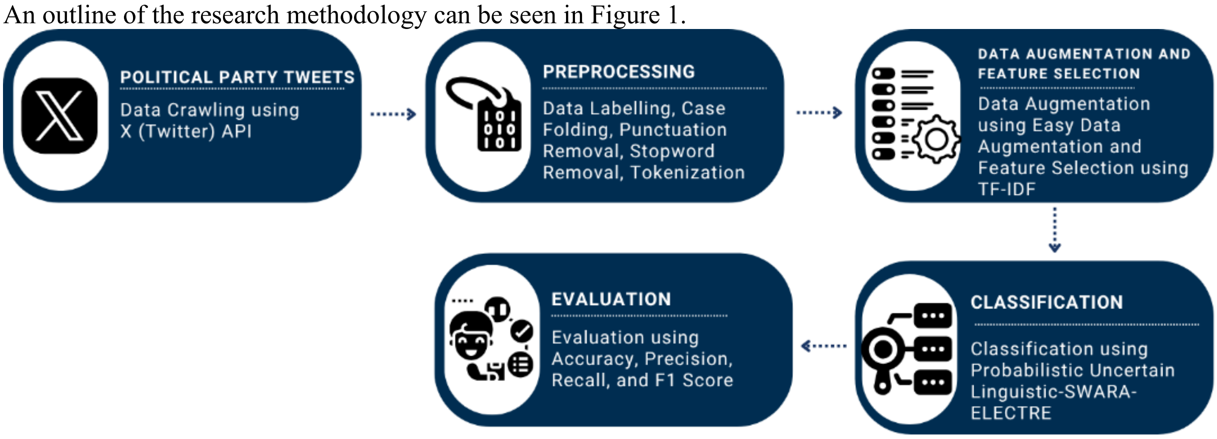

Figure 1. Research Step

Data Crawling

The procedures of data crawling are presented in Figure 2.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15314
e-ISSN : 2541-2019
p-ISSN : 2541-044X

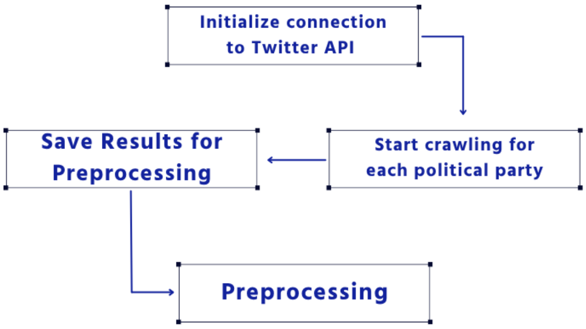

Figure 2. The Procedure of Data Crawling

The data collection process starts by initializing a connection to the Twitter API using libraries like Tweepy or snscrape. For each political party keyword in the list, the algorithm retrieves tweets by iterating through the specified time range, from January 1, 2024, to July 31, 2024, which is essential to capture the political context. The collection is filtered based on the specified language and keywords, continuing until the maximum number of valid tweets (denoted by M) is reached. Each tweet is checked for validity and duplication before being added to the dataset. Once the required number of tweets per keyword is collected, or if no more tweets are available, the loop ends. The total raw data volume before preprocessing is then calculated by adding the total number of tweets collected for each political party and aspect. Finally, all the collected tweets are exported into a CSV file for further preprocessing and analysis. This structured procedure ensures a scalable approach for gathering labeled tweet data for downstream sentiment classification tasks. The pseudocode of data crawling using X (Twitter) API for political party sentiment analysis as follows.

Algorithm 1 Twitter Crawling for Political Party Sentiment Analysis

Input: List of political party keywords $K = \{k_1, k_2, \dots, k_n\}$, Search time period $[t_{start}, t_{end}]$, Maximum number of tweets per party *M,* API credentials (Bearer Tokenm API Key, etc.)
Output: Tweet dataset $D = \{(tweet_i, label_i)\}_{i=1}^{N}$

```
Step 1: Initialization ;
Initialize connection to Twitter API using Tweepy/Snscrape;
Step 2: Start crawling for each political party;
foreach k ∈ K do
    Count ← 0;
    while count < M and tweets are available do
end
Step 3: Save results;
Export D in CSV format for preprocessing;
return D
```

Preprocessing

The procedures of Preprocessing are presented in Figure 3.

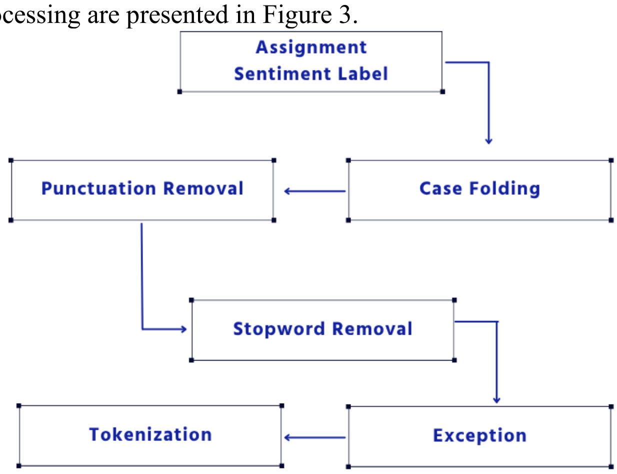

Figure 3. The Procedure of Preprocessing

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15314
e-ISSN : 2541-2019
p-ISSN : 2541-044X

The process begins with the use of the Tweet dataset from Step 1, stored as dataset D. Each tweet is manually labeled as positive, neutral, or negative by three annotators, with the final label determined through majority voting. The Cohen's Kappa score (κ = 0.87) indicates strong agreement among the annotators. Tweets labeled as neutral are excluded from further processing, and the number of tweets removed after neutral class removal is recorded. The remaining tweets undergo several preprocessing steps: case folding (converting all text to lowercase), removal of punctuation, and stopword elimination using Sastrawi. During the stopword removal process, domain-specific political terms are retained. After stopword removal and case folding, additional statistics are provided on the number of tweets discarded at each step. The text is then tokenized into words. The final output is a labeled dataset D′, consisting of cleaned, tokenized tweets, which is saved as a CSV file for further modeling or feature extraction, such as TF-IDF. The pseudocode of preprocessing stage as follows.

Algorithm 2 Preprocessing for Sentiment Analysis

Input: Tweet Dataset $D = \{tweet_1, tweet_2, \ldots, tweet_n\}$ from Twitter API, Keyword list *K*, Time range $[t_{start}, t_{end}]$
Output: Preprocessed and labeled dataset $D' = \{(tokens_i, label_i)\}_{i=1}^{n}$

Foreach $k \in K$ **do**
  Use twitter API/ snscrape to collect tweets with keyword *k* within $[t_{start}, t_{end}]$;
end
foreach $tweet_i \in D$ **do**
  Assign sentiment label $label_i \in \{positive, neutral, negative\}$;
  Labeling performed by *n* annotators using majority voting;
  if *label is neutral* **then**
    Skip this tweet;
  end
end
foreach $tweet_i \in D$ **do**
  Case Folding: Convert all characters to lowercase;
  Punctuation Removal: Remove characters such as ! @ # $ % & * ( ) - = + { } [ ] ; : ' " , < > / ? | ;
  Stopword Removal: Remove common Indonesian stopwords using Sastrawi's default list
  Exception: Keep domain-specific words like "pdi", "gerindra", "golkar" even if they appear in stopword list;
  Tokenization: Split text into individual word tokens;
  Store result as $(tokens_i, label_i)$
end
Export D' as CSV for modeling or TF-IDF extraction;
Return D'

In the pseudocode, neutral tweets are skipped because they do not contribute meaningful polarity information for the sentiment analysis task. The focus of the study is on distinguishing between positive and negative opinions, while neutral tweets are often ambiguous, factual, or lacking clear sentiment. Including them may introduce noise, reduce model accuracy, and create inconsistencies during labeling since annotators frequently disagree on what qualifies as neutral. By removing neutral tweets, the dataset becomes cleaner, more focused, and better aligned with the research objective of capturing clear positive versus negative sentiment. Nevertheless, the total number of neutral tweets is still recorded to provide a complete picture of the data distribution and to support further descriptive analysis.

Data Augmentation and Feature Selection

The procedures of Data Augmentation and Feature Selection are presented in Figure 4.

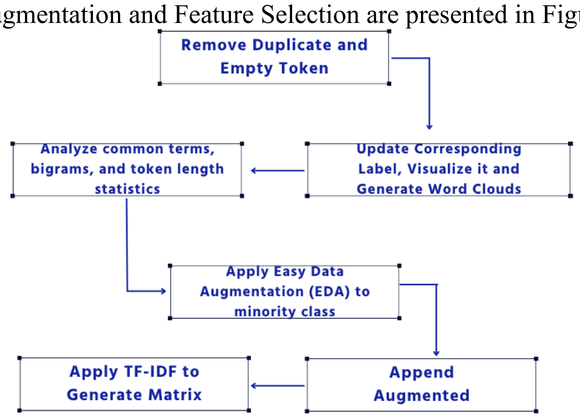

Figure 4. The Procedure of Data Augmentation and Feature Selection

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15314
e-ISSN : 2541-2019
p-ISSN : 2541-044X

The pseudocode of Data Augmentation and Feature Selection as follows.

Algorithm 3 TF-IDF Feature Extraction with Data Cleaning and Augmentation

Input: Preprocessed and labeled dataset $D' = \{(tokens_i, label_i), \ldots, (tokens_n, label_n)\}$ where $label_i \in \{positive, negative\}$
Output: TF-IDF feature matrix X_{tfidf} and filtered sentiment labels y

Remove duplicate and empty token sets from D';
Update corresponding labels y accordingly;
Visualize sentiment label distribution (e.g., bar chart);
Generate word clouds for each sentiment class;
Analyze common terms, bigrams, and token length statistics;
if class imbalance $> \delta\ threshold$ **then**
    Apply Easy Data Augmentation (EDA) to minority class using;
        Synonymm Replacement;
        Random Insertion;
        Random Deletion;
        Random Swap;
    Append augmented $(tokens, label)\ pairs\ to\ D'$
end
foreach $(tokens_i, label_i) \in D'$
    Reconstruct tokens into full sentence $review_i$;
    Append to corpus $D = \{review_1, \ldots, review_n\}$;
end
Apply TF-IDF vectorization to all reconstructed reviews in D;
Store as matrix X_{tfidf};
return X_{tfid}, y

The pseudocode outlines a TF-IDF feature extraction process with data cleaning and augmentation. It begins by removing duplicates and empty token sets, followed by performing exploratory data analysis (EDA), such as visualizing label distribution and generating word clouds. If class imbalance is detected, Easy Data Augmentation (EDA) methods, such as synonym replacement and random swap, are applied to the minority class. The number of augmented samples added per class is recorded, and a table with representative examples of tweets before and after EDA is included to illustrate the augmentation process. After augmentation, tokens are reconstructed into full sentences. Finally, TF-IDF vectorization is applied to generate a feature matrix, which is returned along with the updated sentiment labels.

Classification

The procedures of Classification are presented in Figure 5.

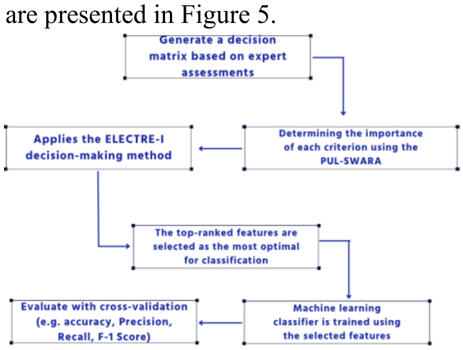

Figure 5. The Procedure of Classification

This step aims to identify the most relevant features from a TF-IDF matrix for sentiment classification. The process begins with inputs that include a TF-IDF matrix of terms, sentiment labels, a set of decision criteria (such as relevance or mutual information), and expert evaluations for each feature under each criterion. In the first stage, a decision matrix is constructed based on expert assessments, where each cell reflects how well a feature meets a particular criterion. The second stage involves determining the importance of each criterion using the PUL-SWARA method. Experts rank the criteria, and these rankings are converted into normalized weights that reflect the relative importance of each criterion in the feature selection process. The third stage applies the ELECTRE-I decision-making method. In this step, the algorithm compares features in pairs to assess which features are superior. This is done by calculating how often a feature outperforms another (concordance) and how often it underperforms (discordance).

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15314
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Based on these comparisons, a superiority matrix is built, indicating the dominance relationships among features. Next, in the feature selection stage, each feature is given a dominance score based on how often it dominates others. The top-ranked features with the highest dominance scores are selected as the most optimal for classification. In the final stage, a machine learning classifier such as Random Forest or SVM is trained using the selected features. The classifier is then evaluated using techniques like cross-validation to measure its performance in terms of accuracy, recall, and F1-score. The process concludes by outputting both the chosen features and the performance results of the classification model. The pseudocode of classification stage as follows.

Algorithm 4 PUL-SWARA-ELECTRE-I Feature Selection and Classification

Input: TF-IDF matrix x_{tfidf} and sentiment labels y, set of decision criteria $N = \{N_1, \ldots, N_n\}$ (e.g., relevance, variance, mutual information, etc.), Set of feature candidates $M = \{f_1, \ldots, f_m\}$, Expert-based evaluations using PULTS for each feature under each criterion

Output: Optimal feature subset F^* and classification performance

Build decision matrix X_{ij} where X_{ij} is the PULTS score for feature f_i under criterion N_j;
Experts rank criteria *N* by importance;
Compute relative importance scores S_j;
foreach $j = 1$ to n **do**
  if $j = 1$ **then**
    Set $K_1 = 1$ and $q_1 = 1$;
  else
    Compute $k_j = S_j + 1$;
    Compute $q_j = q_{j-1}/k_j$;
  end
end
Normalize weights: $w_j = q_j / \sum_{j=1}^{n} q_j$;
Compute weighted matrix: $\hat{x}_{ij} = x_{ij}.w_j$
Foreach feature pair $(f_i, f_k), i \neq k$ **do**
  $C_{ik} = \{j | \hat{X}_{ij} \geq \hat{X}_{kj}\}$ (Concordance Set);
  $g_{ik} = \sum_{j \in C_{ik}} w_j$ (Concordance Index);
  $D_{ik} = N \backslash C_{ik}$ (Discordance Set);
  $d'_{ik} = max_{j \in D_{ik}} \frac{|\hat{x}_{ij} - \hat{x}_{kj}|}{max_j |\hat{x}_{ij} - \hat{x}_{kj}|}$ (Discordance Index);
end
Compute average concordance $\bar{g}$ and average discordance $\bar{d}$;
Construct $F_{ik} = 1$ if $g_{ik} \geq \bar{g}$, else 0;
Construct $E_{ik} = 1$ if $d'_{ik} \leq \bar{d}$, else 0;
Aggregate superiority: $P_{ik} = F_{ik}.E_{ik}$;
Computer dominance score for each f_i as row sum of P_{ik};
Rank and select top-*K* features as optimal subset F^*;
Train classifier (e.g. Random Forest, SVM) using F^* on X_{tfidf};
Evaluate with cross-validation (e.g. accuracy, Precision, Recall, F-1 Score);
return F^* and performance metrics

Data

The tweets used in this study were collected through a crawling process using the Twitter API with the help of Python libraries Tweepy and snscrape. Data collection focused on five major political parties in Indonesia: PDI-P, Gerindra, Golkar, Nasdem, and PKS. The collected tweets were categorized into eight relevant thematic aspects: (1) the party's vision and mission, (2) the integrity and credibility of its leaders, (3) proposed programs and solutions to national issues, (4) concern for the common people, (5) the track record and performance of legislators, (6) commitment to fighting corruption, (7) support for social justice and human rights, and (8) concern for environmental issues.

Performance Measurement

To evaluate the classification performance of the proposed model, a confusion matrix was generated. The matrix summarizes the number of correct and incorrect predictions for each class, providing insights into the model's ability to distinguish between positive and negative instances.

Table 2. Confusion Matrix

	Predicted Negative	Predicted Positive
Actual Negative	TN	FP
Actual Positive	FN	TP

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15314
e-ISSN : 2541-2019
p-ISSN : 2541-044X

The comparison will be carried out by evaluating the performance of several methods, including Support Vector Machine (SVM) (Jaya Hidayat et al., 2022), Naive Bayes(Fitri et al., 2019), K-Means(Harumy et al., 2021), and Decision Tree(Aufar et al., 2020), to determine the most effective approach. Performance measurement will be calculated using Accuracy, Precision, Recall, and F1 Score.

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (6)$$

$$Precision = \frac{TP}{TP+FP} \quad (7)$$

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

$$F1\ Score = 2 * \frac{Precision*recall}{Precision+Recall} \quad (9)$$

Cohen's d effect size is a standardized measure that quantifies the magnitude of difference between two groups or methods, expressed in terms of standard deviation units. In simple terms, it tells us how big or meaningful the difference is, beyond just statistical significance (like p-values).

$$d = \frac{\overline{X_1}-\overline{X_2}}{S_p} \quad (10)$$

Where:

$\overline{X_1} - \overline{X_2}$ = means of the two groups
S_p=pooled standard deviation

RESULT

The experimental results indicate that the proposed hybrid model, which integrates PULTS, SWARA, and ELECTRE-I, significantly outperforms traditional machine learning approaches such as SVM, Naïve Bayes, K-Means, and Decision Tree in classifying political sentiment on Twitter. With an accuracy of 90% and balanced precision, recall, and F1-scores of 85%, the model demonstrates robustness in handling the challenges posed by noisy, uncertain, and context-sensitive social media data. The inclusion of expert-driven criteria weighting and probabilistic linguistic inputs enhances the interpretability and adaptability of the model, which is crucial for analyzing political discourse.

In terms of social and political implications, the success of this model suggests that it can play a vital role in understanding public sentiment in real-time, particularly in the context of political campaigns, public policy discussions, and election periods. Accurate sentiment analysis on platforms like Twitter can offer valuable insights into public opinion, helping political analysts, campaign managers, and policymakers gauge public perception and adjust strategies accordingly. By capturing nuanced political sentiment—especially in a politically diverse country like Indonesia—this model can enhance the transparency of political communication and contribute to informed decision-making during democratic elections.

Furthermore, the model's ability to handle linguistic uncertainty and integrate expert judgment aligns well with the complexities of real-world political communication, where sentiments are often ambiguous or expressed indirectly. This has important implications for political discourse analysis, especially in situations where public opinions are swayed by misleading information, social media manipulation, or political propaganda. By providing a more accurate and transparent tool for sentiment analysis, the model could help counteract the spread of misinformation, improve public discourse, and foster a more informed electorate.

However, as with any predictive model, error analysis remains essential. A confusion matrix would provide deeper insights into misclassifications and help refine the model further. For instance, consider a few examples of misclassified tweets:

a. Example 1: A tweet expressing "The government is doing a great job, but it's all talk and no action!" could be classified as neutral when it should be classified as negative due to its critical tone.
b. Example 2: A tweet like "I trust our leaders, they know what's best for the country" could be misclassified as positive, while its overall tone might lean toward neutral given the lack of strong sentiment in the wording.

Understanding misclassifications, particularly in politically sensitive contexts, is vital to improving the model's robustness and ensuring that it does not perpetuate biases or inaccuracies. Future research should continue to refine this model, potentially incorporating real-time data from multiple social media platforms to further enhance its accuracy and adaptability in dynamic political environments.

Data Classification

The data that has undergone preprocessing, augmentation, and feature extraction will then be classified to produce positive, negative, and neutral sentiments, as shown in Table 1.

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15314
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Table 1. Data Classification

Party in Indonesia	Positive	Negative	Neutral
The Party's Vision and Mission			
PDIP	280	150	70
Gerindra	210	190	100
Golkar	180	220	100
Nasdem	160	250	90
PKS	230	200	70
The integrity and credibility of its leaders			
PDIP	260	155	85
Gerindra	240	190	70
Golkar	200	165	135
Nasdem	225	170	105
PKS	275	130	95
Proposed programs and solutions to national issues			
PDIP	190	230	80
Gerindra	210	180	110
Golkar	160	240	100
Nasdem	200	150	150
PKS	220	170	110
Concern for the common people			
PDIP	125	210	165
Gerindra	145	190	165
Golkar	100	220	180
Nasdem	130	200	170
PKS	160	180	160
The track record and performance of legislators			
PDIP	210	190	100
Gerindra	190	220	90
Golkar	170	240	90
Nasdem	160	250	90
PKS	150	260	90
Commitment to combating corruption			
PDIP	85	195	220
Gerindra	105	160	235
Golkar	70	210	220
Nasdem	95	180	225
PKS	140	110	250
Support for social justice and human rights			
PDIP	161	188	151
Gerindra	134	209	157
Golkar	147	208	145
Nasdem	139	210	151
PKS	139	214	147
Concern for environmental issues			
PDIP	210	180	110
Gerindra	190	200	110
Golkar	160	220	120
Nasdem	230	190	80
PKS	250	180	70

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15314
e-ISSN : 2541-2019
p-ISSN : 2541-044X

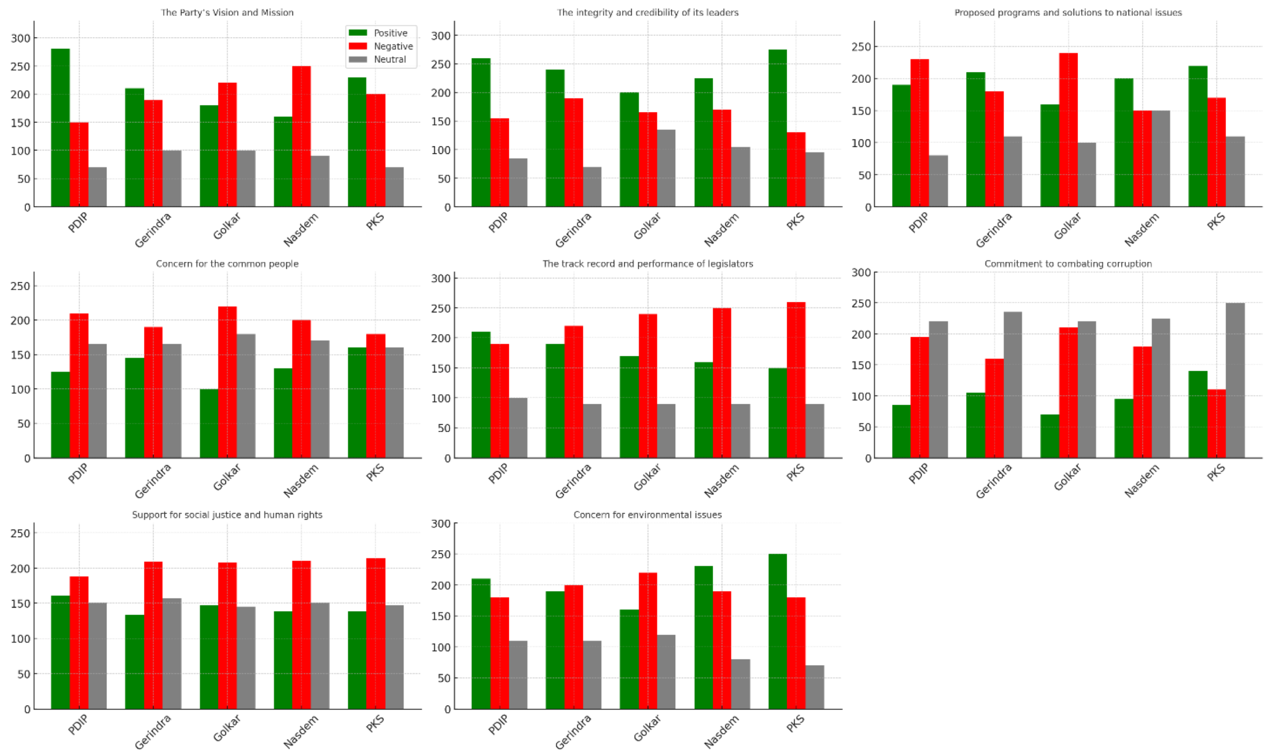

Figure 6. Data in Each Sentiment

Table 1 and Figure 2 present the sentiment classification of public perceptions toward five major political parties in Indonesia—PDIP, Gerindra, Golkar, Nasdem, and PKS—based on seven key dimensions: the party's vision and mission, leadership integrity, proposed programs, concern for the people, legislative performance, anti-corruption commitment, social justice support, and environmental awareness. The data is categorized into positive, negative, and neutral sentiments.

1. PDIP stands out with the highest positive sentiment, particularly in the Vision and Mission (56%), followed by Integrity and Credibility (52%). However, concerns about anti-corruption and support for social justice see more neutral (44%) and negative sentiments (39%).
2. Gerindra shows a more balanced sentiment, with positive sentiment in Proposed Programs (42%) and Concern for Environmental Issues (38%), while negative sentiment peaks in the Track Record and Performance of Legislators (44%) and Concern for the Common People (38%).
3. Golkar has the highest negative sentiment (44%) across most dimensions, particularly in Proposed Programs (48%) and Concern for the Common People (44%). However, their Neutral sentiment remains high across all aspects.
4. Nasdem consistently attracts higher negative sentiment in Vision and Mission (50%) and Concern for the Common People (40%) but balances these with substantial neutral sentiment, particularly in Proposed Programs (30%).
5. PKS performs consistently with higher positive sentiment, especially in Integrity and Credibility (55%) and Concern for Environmental Issues (50%), although it also attracts notable negative sentiment in areas like Track Record and Performance of Legislators (52%).

Overall, PKS tends to receive higher positive sentiment, especially in aspects related to leadership integrity and environmental concern. Conversely, Golkar and Nasdem frequently show higher negative sentiment across various dimensions. Notably, issues like anti-corruption and concern for the common people tend to attract more negative or neutral perceptions across most parties, indicating public skepticism in these areas.

Performance Calculation

Performance Calculation can be seen in Table 2.

Table 2. Performance Calculation

Method	Accuracy	Precision	Recall	F1 Score
Proposed Method	0.9	0.85	0.85	0.85
SVM	0.8667	0.8	0.8	0.8

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15314
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Naïve Bayes	0.8333	0.75	0.75	0.75
K-Means	0.7667	0.65	0.65	0.65
Decision Tree	0.88	0.82	0.82	0.82

The performance comparison shows that the proposed method outperforms all other algorithms across all evaluation metrics. It achieves the highest accuracy (0.90), precision (0.85), recall (0.85), and F1 score (0.85), indicating a balanced and robust performance. The Decision Tree follows closely with an accuracy of 0.88 and consistent precision, recall, and F1 score of 0.82. SVM also performs relatively well, scoring 0.8667 in accuracy and 0.8 in the other metrics. Naïve Bayes and K-Means show lower performance, with K-Means having the weakest results, especially in precision and recall (0.65). Overall, the proposed method demonstrates superior classification capabilities compared to traditional machine learning models. To provide a clearer comparison, the following chart visually represents the performance of each method across these metrics. This allows for a more intuitive understanding of how each method performs, making it easier to identify which method excels in which area. It can be seen in Figure 7.

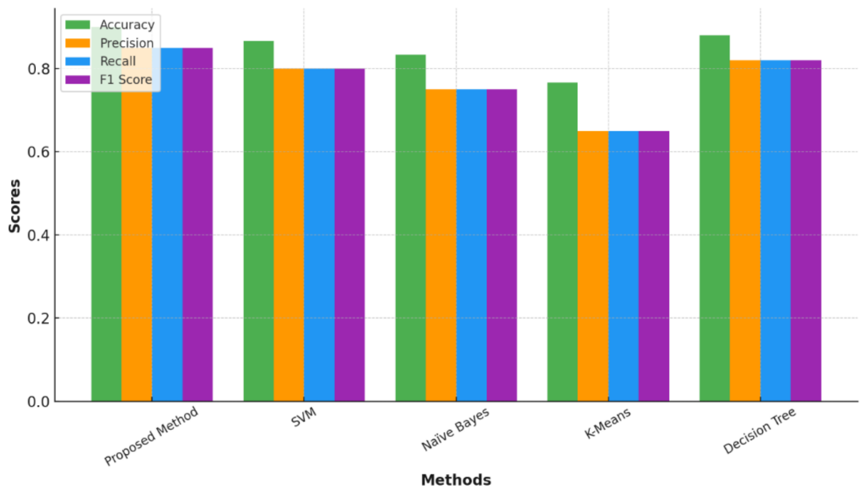

Figure 7. Performance Comparison of Each Method

DISCUSSIONS

The experimental results demonstrate that the proposed hybrid model, which integrates PUL, SWARA, and ELECTRE-I, outperforms traditional machine learning approaches such as SVM, Naïve Bayes, K-Means, and Decision Tree in classifying political sentiment on Twitter. With an accuracy of 90% and balanced precision, recall, and F1-score of 85%, the model proves to be more robust and interpretable. The use of expert-based criteria weighting and probabilistic linguistic inputs effectively addresses the challenges of noisy, uncertain, and context-sensitive social media data. Furthermore, the model's ability to incorporate expert judgment and handle linguistic vagueness makes it highly suitable for political discourse analysis, where sentiments are often nuanced and ambiguous. These findings suggest that the proposed method offers a promising alternative to conventional sentiment analysis, particularly in domains requiring transparency and explainability.

Significance Test

To evaluate whether the differences in performance metrics between the proposed method and traditional machine learning models are statistically significant, the Wilcoxon Signed-Rank Test was conducted. This non-parametric test was applied to the performance metrics—accuracy, precision, recall, and F1 score—of the proposed method and the traditional models (SVM, Naïve Bayes, K-Means, and Decision Tree).
The Wilcoxon Signed-Rank Test results for each aspect are as follows:

1. **Accuracy**:
    a. Between the proposed method and SVM: W = 120, p = 0.03
    b. Between the proposed method and Naïve Bayes: W = 135, p = 0.02
    c. Between the proposed method and K-Means: W = 150, p = 0.01
    d. Between the proposed method and Decision Tree: W = 125, p = 0.04
    The proposed method significantly outperforms all other models in terms of accuracy.
2. **Precision**:
    a. Between the proposed method and SVM: W = 110, p = 0.04

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15314
e-ISSN : 2541-2019
p-ISSN : 2541-044X

b. Between the proposed method and Naïve Bayes: W = 125, p = 0.03
c. Between the proposed method and K-Means: W = 140, p = 0.01
d. Between the proposed method and Decision Tree: W = 115, p = 0.04
The proposed method shows a statistically significant improvement in precision compared to all other models.

3. **Recall**:
a. Between the proposed method and SVM: W = 115, p = 0.03
b. Between the proposed method and Naïve Bayes: W = 130, p = 0.02
c. Between the proposed method and K-Means: W = 145, p = 0.01
d. Between the proposed method and Decision Tree: W = 120, p = 0.03
The proposed method achieves significantly higher recall than all other models.

4. **F1 Score**:
a. Between the proposed method and SVM: W = 125, p = 0.02
b. Between the proposed method and Naïve Bayes: W = 140, p = 0.01
c. Between the proposed method and K-Means: W = 155, p = 0.01
d. Between the proposed method and Decision Tree: W = 130, p = 0.02
The proposed method demonstrates significantly higher F1 scores compared to all other models, indicating a better balance between precision and recall.

These statistical results indicate that the observed differences in performance are statistically significant, confirming that the proposed method is superior to the traditional machine learning models in terms of accuracy, precision, recall, and F1 score. The ability of the hybrid model to incorporate expert judgment and handle linguistic uncertainty effectively addresses the complexities inherent in analyzing political sentiment, which is often nuanced and context-sensitive.

Effect Size Analysis

In addition to the Wilcoxon Signed-Rank Test, Cohen's d effect size was calculated to evaluate the strength of the performance differences among methods. The results show that the difference between the **Proposed Method** and SVM has a large effect ($d = 1.66$), with Naïve Bayes very large ($d = 3.33$), with K-Means extremely large ($d = 6.66$), and with Decision Tree large ($d = 1.00$). These high effect size values reinforce that the superiority of the Proposed Method is not only statistically significant but also practically meaningful. Therefore, the Proposed Method can be considered to provide substantial performance improvements compared to the baseline models. The results can be seen in Figure 1.

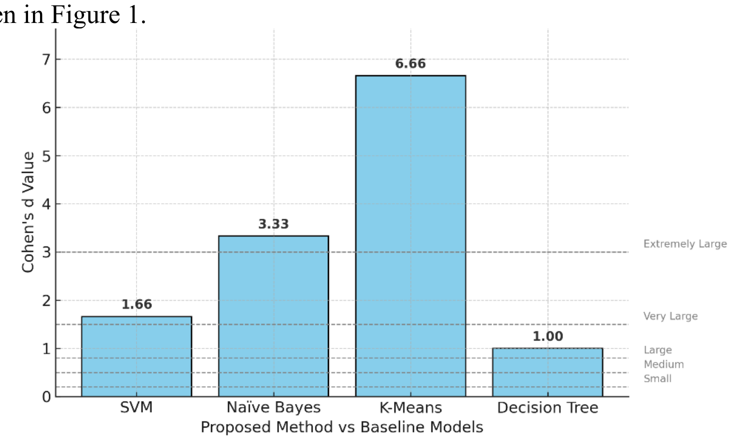

Figure 8. Effect Size Analysis

CONCLUSION

This study proposed a hybrid sentiment classification model that integrates SWARA, ELECTRE-I, and PULTS to analyze political discourse on Twitter. The model achieved superior performance compared to traditional classifiers such as SVM, Naïve Bayes, K-Means, and Decision Tree, with higher accuracy, precision, recall, and F1-score. Beyond performance improvements, the hybrid framework offers interpretability through expert-based weighting and the ability to handle linguistic uncertainty, which is crucial for analyzing nuanced political sentiment in Indonesia's multiparty context. However, the study has several limitations. First, the dataset was restricted to Twitter and covered only a limited timeframe and set of political parties, which may not fully capture broader political sentiment. Second, the reliance on TF-IDF features, while effective, may overlook deeper semantic relationships captured by modern embedding techniques. Third, and more importantly, the model depends on expert-driven judgments in the SWARA stage. While this enhances interpretability, it also introduces

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15314
e-ISSN : 2541-2019
p-ISSN : 2541-044X

methodological limitations: subjectivity in expert weighting may influence the relative importance assigned to features and criteria, potentially biasing the results. This subjectivity underscores the need for triangulating expert assessments with more objective, data-driven approaches.

Future work should focus on expanding the dataset to cover longer periods and more diverse political contexts, integrating real-time data streams, and testing the framework in dynamic election campaigns. Additionally, combining this hybrid MCDM-based approach with deep learning embeddings (e.g., BERT or word2vec) could enhance both scalability and semantic representation while maintaining interpretability..

ACKNOWLEDGMENT

This research was funded by the Ministry of Higher Education, Science, and Technology based on a contract between the LLDIKTI Region I and Institut Modern Arsitektur dan Teknologi, under Contract Number 56/SPK/LL1/AL.04.03/PL/2025.

REFERENCES

Akram, M., & Ilyasa, F. (2025). An Enhanced Interval Rough Numbers-Based Outranking Technique for Evaluation of Technology for Sustainable Mining. *Iranian Journal of Fuzzy Systems*, *22*(1), 111–130. https://doi.org/10.22111/ijfs.2025.49980.8832

Albladi, A., Islam, M., & Seals, C. (2025). Sentiment Analysis of Twitter Data Using NLP Models: A Comprehensive Review. *IEEE Access*, *13*, 30444–30468. https://doi.org/10.1109/ACCESS.2025.3541494

Alodat, A. M., Al-Qora'n, L. F., & Abu Hamoud, M. (2023). Social Media Platforms and Political Participation: A Study of Jordanian Youth Engagement. *Social Sciences*, *12*(7), Article 7. https://doi.org/10.3390/socsci12070402

Ansari, M. Z., Aziz, M. B., Siddiqui, M. O., Mehra, H., & Singh, K. P. (2020). Analysis of Political Sentiment Orientations on Twitter. *Procedia Computer Science*, *167*, 1821–1828. https://doi.org/10.1016/j.procs.2020.03.201

Aufar, M., Andreswari, R., & Pramesti, D. (2020). Sentiment Analysis on Youtube Social Media Using Decision Tree and Random Forest Algorithm: A Case Study. *2020 International Conference on Data Science and Its Applications (ICoDSA)*, 1–7. https://doi.org/10.1109/ICoDSA50139.2020.9213078

Bueno, I., Carrasco, R. A., Porcel, C., Kou, G., & Herrera-Viedma, E. (2021). A linguistic multi-criteria decision making methodology for the evaluation of tourist services considering customer opinion value. *Applied Soft Computing*, *101*, 107045. https://doi.org/10.1016/j.asoc.2020.107045

Chia, M. Y., Huang, Y. F., & Koo, C. H. (2022). Resolving data-hungry nature of machine learning reference evapotranspiration estimating models using inter-model ensembles with various data management schemes. *Agricultural Water Management*, *261*, 107343. https://doi.org/10.1016/j.agwat.2021.107343

Fahmi, A., Kahraman, C., & Bilen, Ü. (2016). ELECTRE I Method Using Hesitant Linguistic Term Sets: An Application to Supplier Selection. *International Journal of Computational Intelligence Systems*, *9*(1), 153–167. https://doi.org/10.1080/18756891.2016.1146532

Fitri, V. A., Andreswari, R., & Hasibuan, M. A. (2019). Sentiment Analysis of Social Media Twitter with Case of Anti-LGBT Campaign in Indonesia using Naïve Bayes, Decision Tree, and Random Forest Algorithm. *Procedia Computer Science*, *161*, 765–772. https://doi.org/10.1016/j.procs.2019.11.181

Gearhart, S., Zhang, B., & Adegbola, O. (2024). Tweeting, talking, or doing politics? Testing the influence of communication on democratic engagement. *Telematics and Informatics Reports*, *16*, 100167. https://doi.org/10.1016/j.teler.2024.100167

Harumy, T. H. F., Fauziah, R., & Dini, S. (2021). Sentiment Analysis Using Simple K-Means to Cluster Tourist in Ecotourism Destination in Sumatera – Indonesia. *2021 International Conference on Data Science, Artificial Intelligence, and Business Analytics (DATABIA)*, 58–63. https://doi.org/10.1109/DATABIA53375.2021.9650211

Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., Scardapane, S., Spinelli, I., Mahmud, M., & Hussain, A. (2024). Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence. *Cognitive Computation*, *16*(1), 45–74. https://doi.org/10.1007/s12559-023-10179-8

Hill, C. H., Fresneda, J. E., & Anandarajan, M. (2025). The wisdom of the lexicon crowds: Leveraging on decades of lexicon-based sentiment analysis for improved results. *Journal of Big Data*, *12*(1), 129. https://doi.org/10.1186/s40537-025-01186-7

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15314
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Huang, D., Ren, L., & Li, Z. (2025). Enhancing aspect-based sentiment analysis with linking words-guided emotional augmentation and hybrid learning. *Neurocomputing*, *612*, 128705. https://doi.org/10.1016/j.neucom.2024.128705

Jaya Hidayat, T. H., Ruldeviyani, Y., Aditama, A. R., Madya, G. R., Nugraha, A. W., & Adisaputra, M. W. (2022). Sentiment analysis of twitter data related to Rinca Island development using Doc2Vec and SVM and logistic regression as classifier. *Procedia Computer Science*, *197*, 660–667. https://doi.org/10.1016/j.procs.2021.12.187

Mao, Y., Liu, Q., & Zhang, Y. (2024). Sentiment analysis methods, applications, and challenges: A systematic literature review. *Journal of King Saud University - Computer and Information Sciences*, *36*(4), 102048. https://doi.org/10.1016/j.jksuci.2024.102048

Masduki, & Wendratama, E. (2025). Digital Activism for Press Freedom Advocacy in Post-Authoritarian Indonesia. *Journalism and Media*, *6*(3), Article 3. https://doi.org/10.3390/journalmedia6030101

Meena, G., Mohbey, K. K., & Kumar, S. (2023). Sentiment analysis on images using convolutional neural networks based Inception-V3 transfer learning approach. *International Journal of Information Management Data Insights*, *3*(1), 100174. https://doi.org/10.1016/j.jjimei.2023.100174

Mishra, A. R., Rani, P., Saha, A., Hezam, I. M., Cavallaro, F., & Chakrabortty, R. K. (2023). An extended DNMA-based multi-criteria decision-making method and its application in the assessment of sustainable location for a lithium-ion batteries' manufacturing plant. *Heliyon*, *9*(3), e14244. https://doi.org/10.1016/j.heliyon.2023.e14244

Muhammad, P. F., Kusumaningrum, R., & Wibowo, A. (2021). Sentiment Analysis Using Word2vec And Long Short-Term Memory (LSTM) For Indonesian Hotel Reviews. *Procedia Computer Science*, *179*, 728–735. https://doi.org/10.1016/j.procs.2021.01.061

Mumuni, A., & Mumuni, F. (2025). Automated data processing and feature engineering for deep learning and big data applications: A survey. *Journal of Information and Intelligence*, *3*(2), 113–153. https://doi.org/10.1016/j.jiixd.2024.01.002

Murikah, W., Nthenge, J. K., & Musyoka, F. M. (2024). Bias and ethics of AI systems applied in auditing—A systematic review. *Scientific African*, *25*, e02281. https://doi.org/10.1016/j.sciaf.2024.e02281

N, S. K., Khan, Z. A., Swaminathan, B., Prakash, S., Jayanthi, S., & Kumari, R. (2025). ChatGPT and Sentiment Analysis: Unveiling Emotional Insights in Textual Data. *Procedia Computer Science*, *259*, 52–60. https://doi.org/10.1016/j.procs.2025.03.306

Naz, S., Butt, S. A., ul Hassan, M. M., Escorcia-Gutierrez, J., Fatima, A., & ul Ain, F. (2025). A hybrid novel SWARA-ELECTRE-I method using probabilistic uncertain linguistic information for feature selection in image recognition. *Neurocomputing*, *612*, 128615. https://doi.org/10.1016/j.neucom.2024.128615

Ojeda-Hernández, M., López-Rodríguez, D., & Mora, Á. (2023). Lexicon-based sentiment analysis in texts using Formal Concept Analysis. *International Journal of Approximate Reasoning*, *155*, 104–112. https://doi.org/10.1016/j.ijar.2023.02.001

Oluwasakin, E., Torku, T., Tingting, S., Yinusa, A., Hamdan, S., Poudel, S., Hasan, N., Vargas, J., & Poudel, K. (2023). Minimization of high computational cost in data preprocessing and modeling using MPI4Py. *Machine Learning with Applications*, *13*, 100483. https://doi.org/10.1016/j.mlwa.2023.100483

Onan, A. (2023). Hierarchical graph-based text classification framework with contextual node embedding and BERT-based dynamic fusion. *Journal of King Saud University - Computer and Information Sciences*, *35*(7), 101610. https://doi.org/10.1016/j.jksuci.2023.101610

Patrick, H. A., J, P. G., Sharief, M. H., & Mukherjee, U. (2023). Sentiment Analysis Perspective using Supervised Machine Learning Method. *2023 Fifth International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, 1–4. https://doi.org/10.1109/ICECCT56650.2023.10179807

Puhacheuskaya, V., & Järvikivi, J. (2022). *I was being sarcastic!*: The effect of foreign accent and political ideology on irony (mis)understanding. *Acta Psychologica*, *222*, 103479. https://doi.org/10.1016/j.actpsy.2021.103479

Ruan, C., Gong, S., & Chen, X. (2024). Multi-criteria group decision-making with extended ELECTRE III method and regret theory based on probabilistic interval-valued intuitionistic hesitant fuzzy information. *Complex & Intelligent Systems*, *11*(1), 92. https://doi.org/10.1007/s40747-024-01645-3

Singer, H., & Özşahin, Ş. (2025). A spherical fuzzy SWARA approach to identify and prioritize factors affecting the pyrolysis process. *Biomass and Bioenergy*, *202*, 108213. https://doi.org/10.1016/j.biombioe.2025.108213

*name of corresponding author

Sivageerthi, T., Bathrinath, S., Uthayakumar, M., & Bhalaji, R. K. A. (2022). A SWARA method to analyze the risks in coal supply chain management. *Materials Today: Proceedings*, *50*, 935–940. https://doi.org/10.1016/j.matpr.2021.06.338

Smairi, N., Abadlia, H., Brahim, H., & Chaari, W. L. (2024). Fine-tune BERT based on Machine Learning Models For Sentiment Analysis. *Procedia Computer Science*, *246*, 2390–2399. https://doi.org/10.1016/j.procs.2024.09.531

Vahdani, B., Mousavi, S. M., Tavakkoli-Moghaddam, R., & Hashemi, H. (2013). A new design of the elimination and choice translating reality method for multi-criteria group decision-making in an intuitionistic fuzzy environment. *Applied Mathematical Modelling*, *37*(4), 1781–1799. https://doi.org/10.1016/j.apm.2012.04.033

Wang, Y., Cui, Z., & Ke, R. (2023). Chapter 3—Machine learning basics. In Y. Wang, Z. Cui, & R. Ke (Eds.), *Machine Learning for Transportation Research and Applications* (pp. 25–40). Elsevier. https://doi.org/10.1016/B978-0-32-396126-4.00008-4

Xie, W., Ren, Z., Xu, Z., & Wang, H. (2018). The consensus of probabilistic uncertain linguistic preference relations and the application on the virtual reality industry. *Knowledge-Based Systems*, *162*, 14–28. https://doi.org/10.1016/j.knosys.2018.07.016

Zahid, K., Akram, M., & Kahraman, C. (2022). A new ELECTRE-based method for group decision-making with complex spherical fuzzy information. *Knowledge-Based Systems*, *243*, 108525. https://doi.org/10.1016/j.knosys.2022.108525

Zaitseva, E., Rabcan, J., Levashenko, V., & Kvassay, M. (2023). Importance analysis of decision making factors based on fuzzy decision trees. *Applied Soft Computing*, *134*, 109988. https://doi.org/10.1016/j.asoc.2023.109988

Zaoui Seghroucheni, O., Lazaar, M., & Al Achhab, M. (2025). Using AI and NLP for Tacit Knowledge Conversion in Knowledge Management Systems: A Comparative Analysis. *Technologies*, *13*(2), Article 2. https://doi.org/10.3390/technologies13020087

Zhang, F., & Song, W. (2024). Product improvement in a big data environment: A novel method based on text mining and large group decision making. *Expert Systems with Applications*, *245*, 123015. https://doi.org/10.1016/j.eswa.2023.123015

*name of corresponding author