SENTIMENT ANALYSIS THE DAMAGE ESAF FRAME WITH SUPPORT VECTOR MACHINE AND IMPACT ON HONDA MOTORCYCLE SALES

Kinanthi Putri Ariyani[1], Aviolla Terza Damaliana[2], Trimono Trimono[3]
Data Science, East Java Veteran National Development University
Jl. Rungkut Madya, Gn. Anyar, Kec. Gn. Anyar, Surabaya, East Java, Indonesia
E-mail: *[1]**anthiputri@gmail.com**, [2]aviolla.terza.sada@upnjatim.ac.id,
[3]trimono.stat@upnjatim.ac.id

Abstract

Damage to the Enhanced Smart Architecture Frame (eSAF) on Honda motorcycles has triggered consumer concerns and has become a public spotlight. This study analyzes public sentiment towards the problem using the Support Vector Machine (SVM) and its impact on sales at one of the dealerships in Surabaya. The data used was in the form of comments from Twitter social media which were classified into two classes, namely positive and negative. Based on the results of the analysis, the majority of 589 public sentiments (59.7%) tended to be negative towards the problem of damage to the eSAF frame, while 397 public sentiments (40.3%) showed positive sentiment. Sales results showed significant fluctuations after this issue emerged, along with increasing negative sentiment. SVM models with a Linear kernel provide the best results with 85% accuracy, 84% precision, 85% recall, and 85% f1-score. SVM was chosen because it excels in text classification compared to algorithms such as K-Nearest Neighbors (KNN), C4.5, and Naïve Bayes, and has been applied in areas such as face detection, bioinformatics, and text processing. This research provides insights for manufacturers to improve product quality, improve customer service, and restore public trust. In addition, the use of the Support Vector Machine algorithm in sentiment analysis can be a reference for similar research in other fields.

Keyword: *eSAF Frame, Sales, Sentiment Analysis, Support Vector Machine*

INTRODUCTION

Technological developments influence various sectors, including transportation, which is essential in Indonesia (Hasbi & Sugiyono, 2023). Motorcycle ownership rose from 115 million units in 2020 to 132 million in 2023 (Badan Pusat Statistik, 2024). PT Astra Honda Motor (AHM) dominated the market with a 78% share in 2023 (CNN Indonesia, 2024) and has long been the top-selling brand (Alexandro et al., 2022). However, the eSAF frame issue in August 2023, involving rust and fractures, raised concerns about product quality and trust, despite AHM's clarification and five-year warranty (Hasbi & Sugiyono, 2023; Utama & Sakti, 2024; Syafi'i & Wiranata, 2024).

Social media plays a key role in shaping public opinion. Sentiment analysis is widely used to measure customer perception. Fitriyah et al. (2020) achieved 79.19% accuracy using SVM for analyzing Gojek user sentiment, while Herwinsyah & Witanti (2022) reported 73.6% for COVID-19 sentiment. Studies on the eSAF issue are still limited, with Syafi'i & Wiranata (2024) using Naïve Bayes and reaching only 70.27% accuracy.

This study applies Support Vector Machine (SVM) to analyze public sentiment regarding the eSAF issue. SVM is known for its high performance in text classification, outperforming methods like Naïve Bayes, KNN, and C4.5 (Husada & Paramita, 2021).

Unlike previous studies that focused only on sentiment classification, this research combines sentiment data from social media with real motorcycle sales data. This integrated approach provides a more comprehensive view of how public sentiment can influence actual purchasing behavior and offers valuable insights for crisis management and consumer behavior analysis.

RESEARCH METHODS

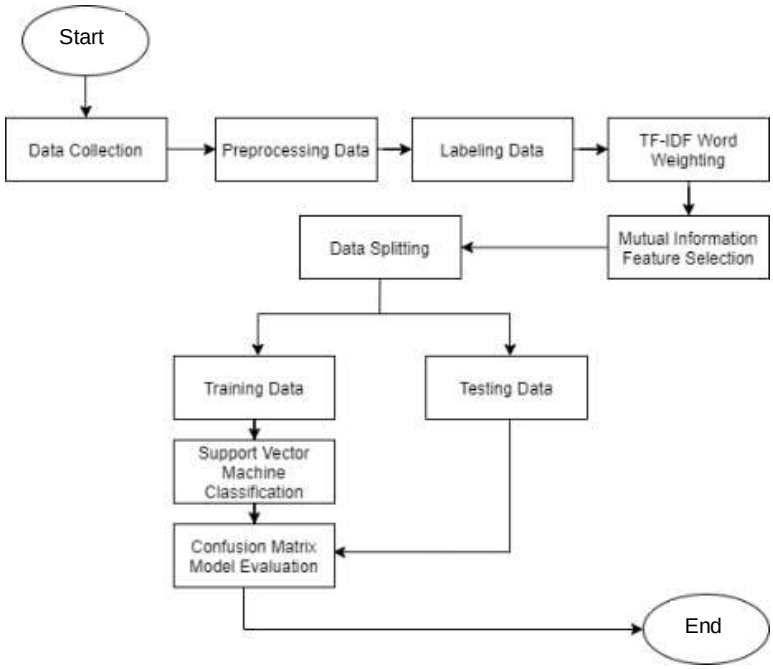

Figure 1. Research Flowchart

In this study, sentiment analysis will be carried out related to the problem of eSAF frames on Honda motorcycles. This analysis uses the Support Vector Machine (SVM) model to classify sentiment into two categories, namely positive and negative.

1. **Data Collection**
   The research data was collected from Twitter between August 15, 2023, and July 20, 2024, using the Tweet Harvest tool and a Twitter API authentication token. The data includes public comments and sentiment labels (positive or negative) from threads discussing the eSAF frame. A total of 986 entries were gathered and saved in CSV format under the filename "rangka_esaf.csv".

2. **Preprocessing Data**
   - Data Cleansing
     Removes irrelevant elements such as numbers, special characters, punctuation, links, hashtags, emojis, etc., using the Regular Expression (RE) library.
   - Case Folding
     Converts all text to lowercase to standardize formatting and avoid case-sensitive differences.
   - Tokenizing
     Splits text into smaller units (tokens), typically words, using the NLTK library for easier analysis.
   - Normalization
     Replaces informal/slang words with formal ones using a reference file (*normalisasi.xlsx*) and the RE library, based on KBBI.
   - Filtering (Stopword Removal)
     Removes meaningless common words like conjunctions and prepositions using the NLTK library and manual additions.
   - Stemming
     Reduces words to their root form by removing prefixes and suffixes to capture the core meaning.

3. Labelling Data

The data labelling stage is an important step in the analysis of public opinion because it aims to categorize the text in the categories of positive and negative opinions. To perform this stage, a positive.txt and negative.txt file is used that contains positive and negative words. Data labelling is done by taking a txt file and reading each text carefully to adjust to the data to be analysed.

4. TF-IDF Word Weighting

The TF-IDF weighting stage or Term Frequency - Inverse Document Frequency is a process to measure the level of importance of a word in each tweet (Addiga & Bagui, 2022). The calculation formula of TF-IDF can be seen in equation 1:

$$W_{t,d} = tf(w,d) \times idf(w,D) \quad (1)$$

Where, $W_{t,d}$ is the TF-IDF weight for the word w in the document d, d is the number of documents containing the word w, D is the total set of documents used.

5. Mutual Information Feature Selection

The feature selection stage using Mutual Information (MI) aims to select the words that are most relevant to sentiment labels in text analysis. MI measures how much information from a word can help predict sentiment. The higher the MI value, the stronger the relationship between the word and the targeted sentiment (Hanafi et al., 2020). By combining TF-IDF and MI, the most influential features can be selected, while the less informative features are ignored. The calculation formula of MI can be seen in equation 2:

$$I(U,C) = \sum_{et\in\{1,0\}} \sum_{ec\in\{1,0\}} P(U=et, C=ec) \log_2 \frac{P(U=et, C=ec)}{P(U=et)P(C=ec)} \quad (2)$$

Where, U are random variables with $et = 1$ (documents containing terms t) and $et = 0$ (documents do not contain terms t), and C are random variables with $ec = 1$ (documents included in the class c) and $ec = 0$ (documents not included in the class c).

6. Support Vector Machine Classification

Support Vector Machine (SVM) is a machine learning method for classification. In the context of sentiment analysis, SVM functions to separate data into different categories based on relevant characteristics. SVM works by obtaining an optimal hyperplane that is able to separate data in the feature space with the greatest margin between two different classes. The hyperplane acts as a separator of tweets with positive sentiment and negative sentiment (Husada & Paramita, 2021). The calculation formula of the SVM can be seen in equation 3:

$$f(x) = \sum_{i=1}^{n} \alpha_i \, y_i K(x_i, x) + b \quad (3)$$

Where α_i is the coefficient specified during the training process to govern the contribution of each training data to the model's decisions, y_i is the class label of the ith training data, $K(x_i, x)$ is the kernel function that measures the similarity between the training data x_i and the input data x, and b is the bias used to set the position of the decision boundary (decision limit). In the classification process using SVM, problems often arise that produce non-optimal results, so that the classification of data is not good. To solve this, you can use kernel functions (Husada & Paramita, 2021). The kernel function functions to transform data into higher dimensional spaces, so that data that is non-linear can be separated linearly (Praghakusma & Charibaldi, 2021). Here's the formula for calculating Linear, Polynomial, and RBF kernels:

Table 1. Kernel Formula

Kernel Name	Kernel Formula
Linear	$K(x_i, x) = x_i^T x$
Polynomial	$K(x_i, x) = \left(\gamma\left(x_i^T x\right) + C\right)^d$
RBF	$K(x_i, x) = exp(-\gamma\|x_i - x\|^2)$

7. Confusion Matrix Model Evaluation

The confusion matrix is a key tool for evaluating classification model performance by comparing predicted and actual labels (Indransyah et al., 2022). It presents four values: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN), which provide insights into the model's accuracy and help identify its strengths and weaknesses.

Table 2. Confusion Matrix

Actual Data	Prediction Data	
	True	False
True	TP	FN
False	FP	MR

In the confusion matrix, the predicted classification results will be compared with the actual data class. The calculation formula of the accuracy, recall, precision, and f1-score values can be seen in equations 4 to 7:

$$Accuracy = \frac{TN + TP}{TP + FP + FN + TN} \quad (4)$$

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

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

RESULTS AND DISCUSSION

1. Data Collection

In this study, the data to be used comes from social media Twitter. Comment data was taken using the keywords "honda rangka esaf", "rangka esaf", and public comments on a thread or a collection of several tweets with the topic of eSAF frame. In Table 3 is an example of the data obtained:

Table 3. Comment Data

No	Raw Data
1	@welovehonda Hebat bet si ini ngasih garansi 5 tahun buat rangka juga tanpa limit kilometer pula ahhhh terbaik memang honda
…	…
986	@detikcom CANGGIH TAPI KEROPOS BUAT APA?!! SEMAKIN KE SINI MOTOR HONDA TAMBAH JELEK BENER KUALITASNYA!

2. Preprocessing Data

In the data preprocessing stage, 2 data were taken as examples from a total of 986 available data. The following is the data preprocessing process that has been carried out:

- Data Cleansing

Table 4. Data Cleansing

No	Raw Data	Data Cleansing
1	@welovehonda Hebat bet si ini ngasih garansi 5 tahun buat rangka juga tanpa limit kilometer pula ahhhh terbaik memang honda	Hebat bet si ini ngasih garansi tahun buat rangka juga tanpa limit kilometer pula ahhhh terbaik memang honda
2	@detikcom CANGGIH TAPI KEROPOS BUAT APA?!! SEMAKIN KE SINI MOTOR HONDA TAMBAH JELEK BENER KUALITASNYA!	CANGGIH TAPI KEROPOS BUAT APA SEMAKIN KE SINI MOTOR HONDA TAMBAH JELEK BENER KUALITASNYA

- Case Folding

Table 5. Case Folding

No	Data Cleansing	Case Folding
1	Hebat bet si ini ngasih garansi tahun buat rangka juga tanpa limit kilometer pula ahhhh terbaik memang honda	hebat bet si ini ngasih garansi tahun buat rangka juga tanpa limit kilometer pula ahhhh terbaik memang honda
2	CANGGIH TAPI KEROPOS BUAT APA SEMAKIN KE SINI MOTOR HONDA TAMBAH JELEK BENER KUALITASNYA	canggih tapi keropos buat apa semakin ke sini motor honda tambah jelek bener kualitasnya

- Tokenizing

Table 6. Tokenizing

No	Case Folding	Tokenizing
1	hebat bet si ini ngasih garansi tahun buat rangka juga tanpa limit kilometer pula ahhhh terbaik memang honda	['hebat', 'bet', 'si', 'ini', 'ngasih', 'garansi', 'tahun', 'buat', 'rangka', 'juga', 'tanpa', 'limit', 'kilometer', 'pula', 'ahhhh', 'terbaik', 'memang', 'honda']
2	canggih tapi keropos buat apa semakin ke sini motor honda tambah jelek bener kualitasnya	['canggih', 'tapi', 'keropos', 'buat', 'apa', 'semakin', 'ke', 'sini', 'motor', 'honda', 'tambah', 'jelek', 'bener', 'kualitasnya']

- Normalization

Table 7. Normalization

No	Tokenizing	Normalization
1	['hebat', 'bet', 'si', 'ini', 'ngasih', 'garansi', 'tahun', 'buat', 'rangka', 'juga', 'tanpa', 'limit', 'kilometer', 'pula', 'ahhhh', 'terbaik', 'memang', 'honda']	['hebat', 'banget', 'si', 'ini', 'memberi', 'garansi', 'tahun', 'untuk', 'rangka', 'juga', 'tanpa', 'limit', 'kilometer', 'pula', 'ahhhh', 'terbaik', 'memang', 'honda']
2	['canggih', 'tapi', 'keropos', 'buat', 'apa', 'semakin', 'ke', 'sini', 'motor', 'honda', 'tambah', 'jelek', 'bener', 'kualitasnya']	['canggih', 'tapi', 'keropos', 'untuk', 'apa', 'semakin', 'ke', 'sini', 'motor', 'honda', 'tambah', 'jelek', 'benar', 'kualitasnya']

- Filtering

Table 8. Filtering

No	Normalization	Filtering
1	['hebat', 'banget', 'si', 'ini', 'memberi', 'garansi', 'tahun', 'untuk', 'rangka', 'juga', 'tanpa', 'limit', 'kilometer', 'pula', 'ahhhh', 'terbaik', 'memang', 'honda']	['hebat', 'banget', 'garansi', 'rangka', 'limit', 'kilometer', 'terbaik', 'honda']
2	['canggih', 'tapi', 'keropos', 'untuk', 'apa', 'semakin', 'ke', 'sini', 'motor', 'honda', 'tambah', 'jelek', 'benar', 'kualitasnya']	['canggih', 'keropos', 'motor', 'honda', 'jelek', 'kualitasnya']

- Stemming

Table 9. Stemming

No	Filtering	Stemming
1	['hebat', 'banget', 'garansi', 'rangka', 'limit', 'kilometer', 'terbaik', 'honda']	['hebat', 'banget', 'garansi', 'rangka', 'limit', 'kilometer', 'baik', 'honda']
2	['canggih', 'keropos', 'motor', 'honda', 'jelek', 'kualitasnya']	['canggih', 'keropos', 'motor', 'honda', 'jelek', 'kualitas']

3. Labeling Data

In the data labeling stage, 2 data were taken as examples from a total of 986 available data. The following are the results of the data labeling that has been carried out:

Table 10. Labeling Data

No	Text	Sentiment
1	hebat banget garansi rangka limit kilometer baik honda	Positive
2	canggih keropos motor honda jelek kualitas	Negative

This data labeling results in 589 negative sentiments and 397 positive sentiments.

4. TF-IDF Word Weighting

	abad	abai	acara	action	ada	adik	adil	adnoh	adv	after	...	would	yamaha	yamahaa	yaris	yasudah	ylki	you	your	youtuber	zaman
0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
981	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
982	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
983	0.0	0.453291	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
984	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
985	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

986 rows × 1000 columns

Figure 2. TF-IDF Weighting Results

The Term Frequency-Inverse Document Frequency (TF-IDF) method transforms text into numerical values by calculating the importance of each word in a document relative to all documents. TF measures how often a word appears in a document, while IDF reduces the weight of words that commonly appear across documents. After preprocessing (tokenization, lowercasing, stopword removal, and stemming), the TF-IDF vectorizer generated a matrix with 986 rows (documents) and 1000 columns (most frequent words). A high TF-IDF value, such as 0.453291 in row 983, indicates that the word is highly relevant in that document. Meanwhile, words with a value of 0 are considered irrelevant or absent.

5. Mutual Information Feature Selection

Table 11. Mutual Information Selection Results

It	Featured	Mutual Information Score
0	rangka	0.166173
1	honda	0.148213
2	motor	0.113336
…	…	…
97	mending	0.006852
98	temu	0.006826
99	lucu	0.006826

Some of the features with the highest Mutual Information scores include the words "rangka" (score 0.166173), "honda" (score 0.148213), and "motor" (score 0.113336). These words show a high relevance to the topic being analyzed. After the feature selection is carried out, weighting is carried out on the features that have been selected and ranked in value. The results of the weighting of the selected features can be seen in table 12:

Table 12. Weighting Results in Feature Selection

It	Featured	Average TF-IDF
0	rangka	0.056683
1	honda	0.053075
2	motor	0.049465
…	…	…
97	ringan	0.003750
98	pasar	0.003731
99	model	0.003602

The feature with the highest score in the TF-IDF weighting calculation was obtained on the word "rangka" (score 0.056683). This calculation is used to facilitate the modeling process by counting words that have been selected and converted into numerical representations. After performing the TF-IDF calculations on the selected features, the next step is to proceed to the modeling stage.

6. Support Vector Machine Classification

Prior to modelling, exploration of various combinations of hyperparameters was carried out for each type of kernel used in SVM, namely Linear, Polynomial, and Radial Basis Function (RBF). The parameters explored in Grid Search include C with a value range of [0.01, 0.1, 0.5, 1, 10, 20, 30, 40, 50, 100] for all kernels; degree [1, 2, 3, 4, 5] for polynomial kernels; and gamma [0.001, 0.01, 0.1, 1, 10] for the RBF kernel. After a Grid Search, the best combination of parameters obtained can be seen in Figure 3:

```
Hasil Grid Search untuk setiap kernel:
Kernel: linear
  Parameter Terbaik: {'C': 40, 'kernel': 'linear'}
Kernel: poly
  Parameter Terbaik: {'C': 100, 'degree': 1, 'kernel': 'poly'}
Kernel: rbf
  Parameter Terbaik: {'C': 20, 'gamma': 1, 'kernel': 'rbf'}
```

Figure 3. Grid Search

In this case, the SVM model leverages three common kernel types, namely Linear, Polynomial, and RBF to help classify data. The value of x is taken from the results of the weighting with the pre-selected features and the value of y is taken from the label column containing the positive and negative sentiment classes. The data was then divided into training data and test data with a ratio of 80:20 with the parameter stratify=y to ensure that the distribution of sentiment classes remained proportional.

- **Linear Kernel**
  In linear kernels, the C parameter controls the trade-off between margin and misclassification. Although C was set to 40 through Grid Search, it yielded lower accuracy (84%) than C = 10, which achieved 85% on testing and 83% on training data. The SVM model with a linear kernel generates a hyperplane based on feature coefficients $w_1, w_2, \ldots, w_n$ and a bias b, which define the decision boundary separating positive and negative classes. The performance evaluation can be seen in Figure 4:

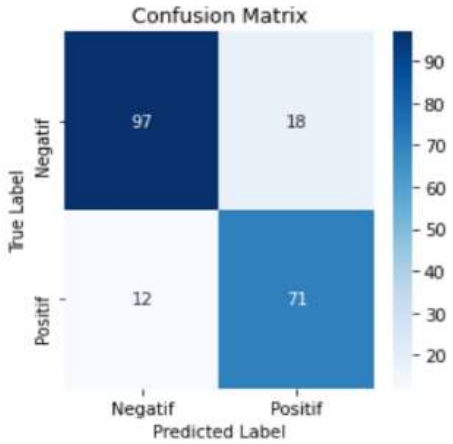

Figure 4. Confusion Matrix Linear Kernel

- True Positive (TP): 71 (positive prediction and true positive)
- False Positive (FP): 18 (positive prediction, but should be negative)
- False Negative (FN): 12 (negative prediction, but should be positive)
- True Negative (TN): 97 (negative prediction and negative true prediction)

- **Polynomial Kernel**
  In a polynomial kernel, the parameter degree=1 determines the polynomial degree of the kernel which in this case is a linear polynomial (degree 1). The value of C=100 is used to set penalties for classification errors. With a large C, the model is expected to pay close attention to classification errors in the training data, providing a narrower margin but striving to reduce errors. The performance evaluation can be seen in Figure 5:

P-ISSN: 2622-6901
E-ISSN: 2622-6375

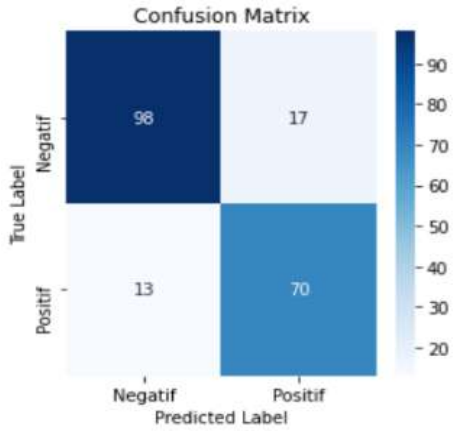

Figure 5. Confusion Matrix Polynomial Kernel

- True Positive (TP): 70 (positive prediction and true positive)
- False Positive (FP): 17 (positive prediction, but should be negative)
- False Negative (FN): 13 (negative prediction, but should be positive)
- True Negative (TN): 98 (negative prediction and negative true prediction)

- **RBF Kernel**
  In the RBF kernel, the gamma=1 parameter controls the form of the RBF kernel function which affects how far the influence of each point data points affects the classification decision. The parameter C=20 is used to set the penalty for classification errors. The performance evaluation can be seen in Figure 6:

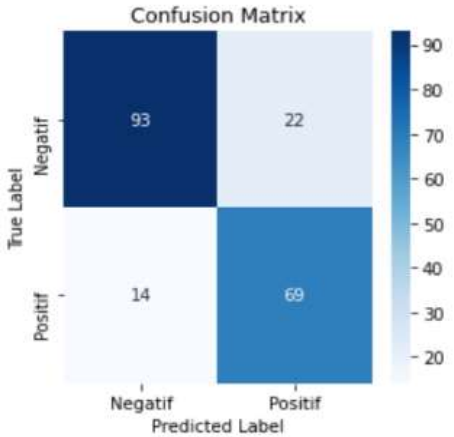

Figure 6. Confusion Matrix RBF Kernel

- True Positive (TP): 69 (positive prediction and true positive)
- False Positive (FP): 22 (positive prediction, but should be negative)
- False Negative (FN): 14 (negative prediction, but should be positive)
- True Negative (TN): 93 (negative prediction and negative true prediction)

P-ISSN: 2622-6901
E-ISSN: 2622-6375

7. Support Vector Machine Test Summary

Table 13. Test Summary

Kernel	Parameters	Training Accuracy	Testing Accuracy	Precision Avg.	Recall Avg.	F1-Score Avg.
Linear	cost = 10	83%	85%	84%	85%	85%
Polynomial	degree = 1 cost = 100	84%	84%	84%	85%	85%
RBF	gamma = 1 cost = 20	91%	82%	81%	82%	82%

Grid Search results show that the Linear kernel (cost = 10) delivers the best performance, with 85% testing accuracy and evaluation metrics (Precision, Recall, F1-Score) averaging 84–85%, indicating strong generalization. The Polynomial kernel (degree = 1, cost = 100) achieved similar results with 84% accuracy, but the Linear kernel slightly outperformed it in testing. The RBF kernel (gamma = 1, cost = 20) had the highest training accuracy (91%) but lower testing accuracy (82%) and average metrics of 81–82%, suggesting overfitting. Thus, the Linear kernel is the most optimal for balanced and reliable performance.

8. Visualization

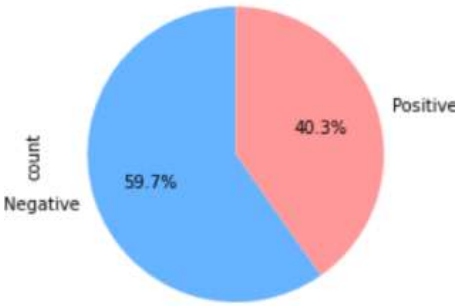

Figure 7. Sentiment Distribution

Figure 8. Word Cloud Positive Sentiment

Figure 9. Word Cloud Negative Sentiment

9. Sales Data Graph

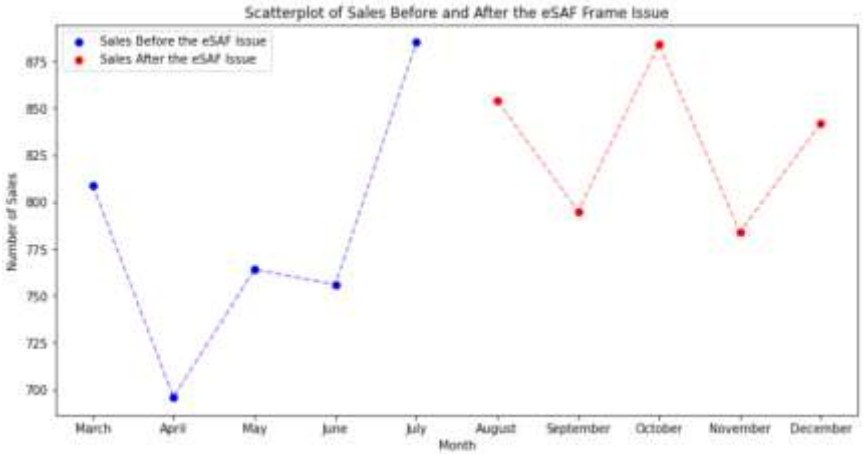

Figure 10. Sales Charts

Based on Figure 10 compares motorcycle sales with eSAF frames at a Surabaya dealer before and after the frame damage issue. Sales from March to July (blue dots) show a relatively stable trend with a peak in July. In contrast, sales from August to December (red dots) display a more volatile pattern, with fluctuations such as a drop in September, a spike in October, and another drop in November. This shift suggests a possible influence of rising negative sentiment on consumer purchasing behaviour.

CONCLUSION

Sentiment analysis on the eSAF frame issue shows 59.7% negative and 40.3% positive opinions, indicating a dominant negative perception that may affect purchasing decisions. This aligns with fluctuating sales data from a dealer in Surabaya, suggesting that negative sentiment impacts consumer confidence, even if sales declines are not always drastic.

Model evaluation using three SVM kernels—Linear, Polynomial, and RBF—revealed that the Linear kernel performed best, with 85% testing accuracy and 83% training accuracy. The C parameter also influenced performance, where C = 10 outperformed C = 40. The Polynomial kernel showed balanced results (84%), while the RBF kernel overfit the data, achieving 91% training but only 82% testing accuracy. These results highlight the importance of proper parameter selection, and suggest that regularization or hyperparameter tuning could further enhance model performance.

SUGGESTION

To improve model performance, it is recommended to optimize parameter selection, particularly for the RBF kernel, to prevent overfitting and enhance generalization. Regularization and data imbalance handling should also be considered. Accurate labelling by individuals with strong language skills and contextual understanding is essential for producing reliable training data and analysis results.

BIBLIOGRAPHY

Addiga, A., & Bagui, S. (2022). Sentiment Analysis on Twitter Data Using Term Frequency-Inverse Document Frequency. *Journal of Computer and Communications*, 117-128.

Agatha, A., Paramita, S., & Sudarto. (2022). Opini Publik Netizen terhadap Pencemaran Nama Baik di Media Online. *Jurnal Koneksi*, 278-286.

Alexandro, R., Uda, T., Hariatama, F., & Sinaga, B. D. (2022). Analysis of Percentage Frequency Distribution Towards Satisfaction from Users of Honda Motorcycles. *International Journal of Social Science and Business*, 191-198.

Fitriyah, N., Warsito, B., & Maruddani, D. A. (2020). Analisis Sentimen Gojek Pada Media Sosial Twitter Dengan Klasifikasi Support Vector Machine (SVM). *Jurnal Gaussian*, 376-390.

Hanafi, A., Adiwijaya, & Astuti, W. (2020). Klasifikasi Multi Label Pada Hadis Bukhari Terjemahan Bahasa Indonesia Menggunakan Mutual Information dan K-Nearest Neighbor. *Jurnal SISFOKOM (Sistem Informasi dan Komputer)*, 357-364.

Hasbi, M. R., & Sugiyono, H. (2023). Problematika Penggunaan Rangka Enhanced Smart Architecture Frame Pada Sepeda Motor yang Cacat Produksi (Studi Kasus Kerusakan Rangka Motor Matic Honda). *Jurnal Interpretasi Hukum*, 712-720.

Herwinsyah, & Witanti, A. (2022). Analisis Sentimen Masyarakat Terhadap Vaksinasi COVID-19 Pada Media Sosial Twitter Menggunakan Algoritma Support Vector Machine (SVM). *Jurnal Sistem Informasi dan Informatika (SIMIKA)*, 59-67.

Husada, H. C., & Paramita, A. S. (2021). Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM). *Jurnal Teknologi Informasi dan Komunikasi*, 18-26.

Indonesia, C. (2024, March 28). *78 Persen Motor Baru Terjual di Indonesia 2023 adalah Merek Honda*. Retrieved September 20, 2024, from CNN Indonesia: https://www.cnnindonesia.com/otomotif/20240327184707-595-1079768/78-persen-motor-baru-terjual-di-indonesia-2023-adalah-merek-honda

Indransyah, R., Chrisnanto, Y. H., & Sabrina, P. N. (2022). Sentimen Pergelaran Motogp di Indonesia Menggunakan Algoritma Correlated Naïve Bayes Clasifier. *Infotech Journal*, 60-66.

Lubis, N. R., & Zahara, F. (2024). Perlindungan Konsumen Terhadap Pembelian Sepeda Motor Baru Mengenai Kerusakan Rangka Esaf Ditinjau Dari Perspektif Ibnu Taimiyah dan Undang-Undang Nomor 8 Tahun 1999 Tentang Perlindungan Konsumen. *UNES LAW REVIEW*, 6970-6980.

Maharani, C. A., Warsito, B., & Santoso, R. (2023). Analisis Sentimen Vaksin COVID-19 Pada Twitter Menggunakan Recurrent Neural Network (RNN) dengan Algoritma Long Short-Term Memory (LSTM). *Jurnal Gaussian*, 403-413.

Praghakusma, A. Z., & Charibaldi, N. (2021). Komparasi Fungsi Kernel Metode Support Vector Machine untuk Analisis Sentimen Instagram dan Twitter (Studi Kasus Komisi Pemberantasan Korupsi). *Jurnal Sarjana Teknik Informatika, 9*, 33-42.

Statistik, B. P. (2024, Februari 20). *Jumlah Kendaraan Bermotor Menurut Provinsi dan Jenis Kendaraan (unit), 2023*. Retrieved September 19, 2024, from Badan Pusat Statistik (BPS – Statistics Indonesia): https://www.bps.go.id/id/statistics-table/3/VjJ3NGRGa3dkRk5MTlU1bVNFOTVVbmQyVURSTVFUMDkjMw==/jumlah-kendaraan-bermotor-menurut-provinsi-dan-jenis-kendaraan--unit---2023.html?year=2023

Syafi'i, A. C., & Wiranata, A. D. (2024). Analisis Sentimen Terhadap Rangka E-SAF Honda Pada Media Sosial X Dengan Algoritma Naïve Bayes. *KLIK: Kajian Ilmiah Informatika dan Komputer*, 57-66.

Utama, M. A., & Sakti, M. (2024). Consumer Protection for Honda Vehicle Users with Frame eSAF Damage Based on the Principles of Consumer Safety and Security. *Journal of Law, Politic, and Humanities (JLPH)*, 1325-1331.

Yoga, Z. (2023, September 11). *Pedagang Motor Bekas: Efek Viral Rangka eSAF Honda Keropos, Konsumen Lebih Pilih Opsi Lain*. Retrieved Maret 2, 2025, from OTO by CarDekho SEA: https://www.oto.com/berita-motor/pedagang-motor-bekas-efek-viral-rangka-esaf-honda-keropos-konsumen-lebih-pilih-opsi-lain