Available online at www.coins-indocompt.org

COmputing and INformation Systems Journal

E-ISSN: 3109-3248

Vol. 1, No. 3, December, 2025

PUBLIC SENTIMENT ANALYSIS OF THE ANTAM GOLD CASE ON SOCIAL MEDIA X USING THE SUPPORT VECTOR MACHINE (SVM) METHOD

Achmad Raudhotul Hikmah[1], Nur Qodariyah Fitriyah[2], Dudi Irawan.[3*]

Muhammadiyah University of jember, Jember 68121, Indonesia

achmadraudhotulhikmah@gmail.com

Abstract

The development of social media, especially X (Twitter), has made it an important source for observing public opinion. One of the issues widely discussed is the Antam gold case related to the alleged misuse of official stamps and the circulation of illegal gold. This study analyzes public sentiment toward this issue by classifying tweets into positive, negative, and neutral categories. Data were collected through crawling using the keyword "kasus emas antam" during the period March 1–May 7, 2025, resulting in 716 tweets. The research process includes preprocessing, feature extraction using N-gram and TF-IDF, and classification using Support Vector Machine (SVM). Evaluation was carried out using a confusion matrix to calculate accuracy, precision, recall, and F1-score. The results show that the SVM model achieved an accuracy of 85%, precision of 91%, recall of 78%, and an F1-score of 80%, indicating good performance in classifying sentiment. This research is expected to provide an overview of public perception and serve as a reference in the application of sentiment analysis on other social issues.

Keywords : Sentiment Analysis, Support Vector Machine (SVM), Antam Gold, TF-IDF, N-gram.

1. Introduction

The development of technology and social media provides a space for the public to express opinions quickly and openly on various public issues [1]. Social media platform X has become one of the most widely used data sources in sentiment analysis research due to its active, real-time nature and open conversations, allowing public opinion to be collected in large quantities, supported by the availability of APIs that facilitate data retrieval [2]. In contrast, data from the Google Play Store are used more selectively because the reviews available are specific to certain applications and less representative of public opinion compared to X [3].

The Antam gold case, which occurred due to alleged misuse of the official Logam Mulia (LM) stamp on illegal gold during the period 2010–2021, attracted widespread public attention. Although the gold was not counterfeit, the practice of producing illegal gold with the Antam stamp without authorization affected the company's reputation as well as public trust in gold investment. Therefore, understanding public sentiment is crucial to assessing public perception and the social and economic impacts of the case.

Sentiment analysis allows the classification of public opinions into positive, negative, and neutral categories through computational approaches. One effective method used in sentiment analysis is the Support Vector Machine (SVM), first introduced by Vapnik and known for producing high-accuracy classification results [4]. To address

* Corresponding author: author@institute.xxx .

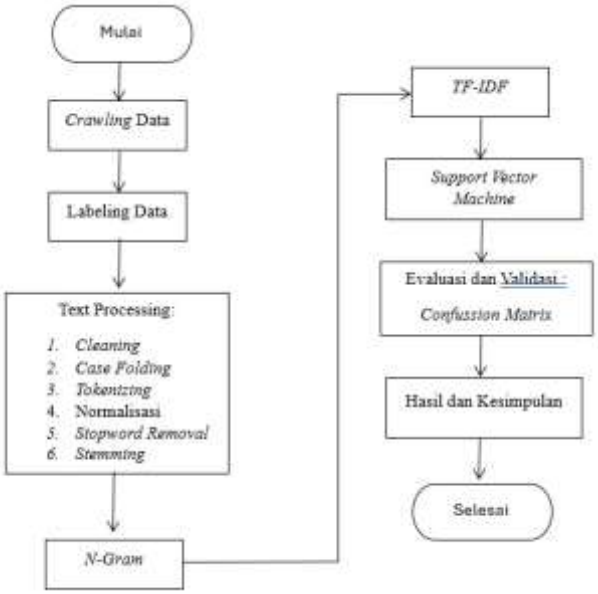
http://dx.doi.org/10. xxxxx/HIJ-xxxxxx

nonlinear data problems, SVM has been developed using kernel functions, making it more adaptive to textual data [5]. Various studies have demonstrated that SVM has a strong mathematical foundation and provides better classification performance compared to other methods in sentiment analysis [6,7].

Based on the above, this study aims to analyze public sentiment regarding the Antam gold case using data from platform X and to evaluate the performance of the Support Vector Machine method in the sentiment classification process.

2. Method

This study employs a quantitative approach using sentiment analysis to classify public opinions regarding the Antam gold case. Research data were obtained from social media platform X by utilizing data crawling techniques to collect posts relevant to the research topic. Platform X was selected due to its active, real-time nature and the ability to retrieve large amounts of data through the Application Programming Interface (API).



Figur 1. Research Stages

2.1. Data Collection

Data penelitian bersumber dari platform media sosial X karena bersifat aktif, real-time, dan percakapannya terbuka sehingga mampu merepresentasikan opini publik secara luas.

2.2. Tet Preprocessing

Preprocessing cleans comments from noise to prepare them for analysis, including Cleaning, Case Folding, Tokenizing, Normalization, Stopword Removal, and Stemming.

2.2.1 Cleaning

At this stage, the text data is cleaned by removing irrelevant or disruptive elements, such as punctuation, symbols, numbers, URLs, mentions, hashtags, emoticons, extra spaces, or inconsistent formatting. This process aims to produce clean text and reduce noise.

2.2.2 Case Folding

In the Case Folding stage, all letters in the text are converted to lowercase. This step is carried out to standardize the text format, ensuring that differences between uppercase and lowercase letters do not affect the analysis process.

2.2.3 Tokenizing

In the Tokenization stage, the text is split into words or the smallest units called tokens. This process is essential for preparing the text data for further analysis, as it facilitates the identification of words, patterns, and frequencies within the text.

2.2.4 Normalization

The normalization process aims to standardize non-standard words in accordance with the Indonesian Dictionary (KBBI), including abbreviations. This step involves converting abbreviations, slang, or informal word variations into a more formal and standardized form, ensuring consistency of the text for further analysis.

2.2.5 Stopword Removal

In the Stopword Removal stage, common words that carry little significance in the analysis, such as “di” (in), “dan” (and), “yang” (which/that), and similar terms, are removed. This process helps reduce noise in the text and focuses the analysis on more meaningful and relevant words. An example of the Stopword Removal process is presented in the following table.

2.2.6 N-Gram

After preprocessing, the comment data were converted into features using N-Grams (Unigrams, Bigrams, Trigrams) to capture word patterns and context. These features were then used to train an SVM model to classify sentiment into positive, negative, or neutral categories.

2.2.7 TF-IDF

After the text was converted into N-Gram representations, the next step was to transform the data into numerical vectors using the TF-IDF weighting method. This technique aims to evaluate the importance of a word within a document relative to the entire corpus in the dataset. Words that frequently appear in one document but rarely in others are assigned higher weights. In this study, TF-IDF was used to assign weights to features extracted from unigrams and bigrams in the crime report description texts. These weights were then used as input for the SVM algorithm to enhance the model’s classification accuracy.

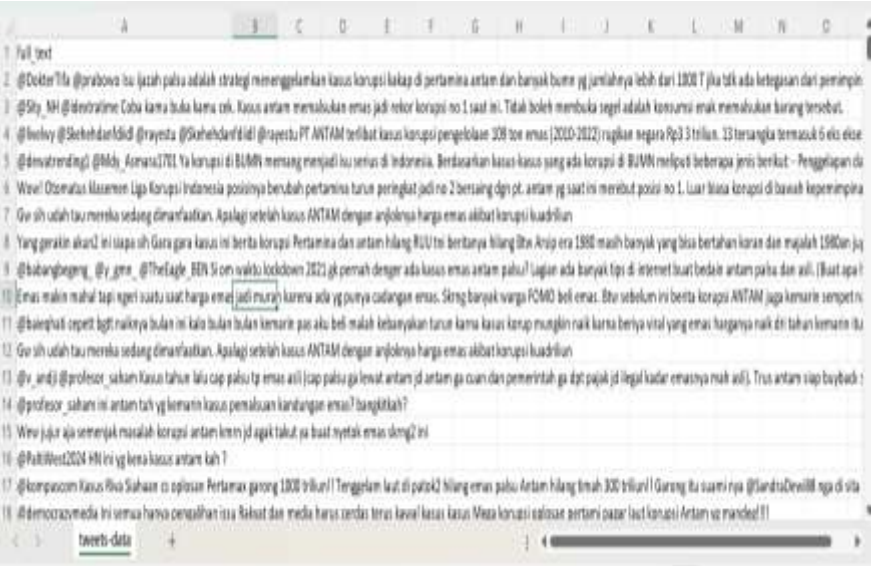
2.2.8 Support Vector Machine Modeling

After obtaining the TF-IDF values for each document, the next step is vectorization, which involves assigning a numerical representation to each word in the document.

3. Results

3.1. Data Collection

At this stage, data crawling was performed on comments from platform X using Google Colab, resulting in a dataset of 716 diverse comments.



3.2. Data Labeling

After the data were obtained through the crawling process, the next stage involved manually labeling the data by an NLP (Natural Language Processing) expert, Mr. Hardian Oktavianto, S.Si., M.Kom, and a language expert, Dr. Eka Nova Ali Vardani, M.Pd.

	A	B
1	full_text	lebeling
2	@DokterTifa @prabowo Isu ijazah palsu adalah strategi menenggelamkan kas	negatif
3	@Sity_NH @idextratime Coba kamu buka kamu cek. Kasus antam memalsuka	netral
4	@livelwy @Skehehdanfdiidl @rayestu @Skehehdanfdiidl @rayestu PT ANTAM	netral
5	@dewatrending1 @Mdy_Asmara1701 Ya korupsi di BUMN memang menjadi i	netral
6	Wow! Otomatus klasemen Liga Korupsi Indonesia posisinya berubah pertamin	netral
7	Gw sih udah tau mereka sedang dimanfaatkan. Apalagi setelah kasus ANTAM	negatif
8	Yang gerakin akun2 ini siapa sih Gara gara kasus ini berita korupsi Pertamina d	netral
9	@babangbegeng_ @y_gmn_ @TheEagle_BEN Si om waktu lockdown 2021 gk p	netral
10	Emas makin mahal tapi ngeri suatu saat harga emas jadi murah karena ada yg	negatif
11	@baieqhati cepett bgtt naiknya bulan ini kalo bulan bulan kemarin pas aku bel	netral
12	Gw sih udah tau mereka sedang dimanfaatkan. Apalagi setelah kasus ANTAM	negatif
13	@v_andji @profesor_saham Kasus tahun lalu cap palsu tp emas asli (cap palsu	netral
14	@profesor_saham ini antam tuh yg kemarin kasus pemalsuan kandungan ema	netral
15	Wew jujur aja semenjak masalah korupsi antam kmrn jd agak takut ya buat ny	negatif
16	@PaltiWest2024 HN ini yg kena kasus antam kah ?	netral
17	@kompascom Kasus Riva Siahaan cs oplosan Pertamax garong 1000 triliun!! Te	netral
18	@democrazymedia Ini semua hanya pengalihan issu Rakyat dan media harus	negatif

3.2.1 Cleaning

Cleaning is the stage of removing non-alphabetic characters, such as numbers, #, @, emojis, and website links.

```
Hasil Cleaning (10 data pertama):
                                                full_text  \
0  @DokterTifa @prabowo Isu ijazah palsu adalah s...
1  @Sity_NH @idextratime Coba kamu buka kamu cek....
2  @livelwy @Skehehdanfdiidl @rayestu @Skehehdanf...
3  @dewatrending1 @Mdy_Asmara1701 Ya korupsi di B...
4  Wow! Otomatus klasemen Liga Korupsi Indonesia ...
5  Gw sih udah tau mereka sedang dimanfaatkan. Ap...
6  Yang gerakin akun2 ini siapa sih Gara gara kas...
7  @babangbegeng_ @y_gmn_ @TheEagle_BEN Si om wak...
8  Emas makin mahal tapi ngeri suatu saat harga e...
9  @baieqhati cepett bgtt naiknya bulan ini kalo ...

                                               clean_text
0  Isu ijazah palsu adalah strategi menenggelamka...
1  Coba kamu buka kamu cek Kasus antam memalsukan...
2  PT ANTAM terlibat kasus korupsi pengelolaan to...
3  Ya korupsi di BUMN memang menjadi isu serius d...
4  Wow Otomatus klasemen Liga Korupsi Indonesia p...
5  Gw sih udah tau mereka sedang dimanfaatkan Apa...
6  Yang gerakin akun ini siapa sih Gara gara kasu...
7  Si om waktu lockdown gk pernah denger ada kasu...
8  Emas makin mahal tapi ngeri suatu saat harga e...
9  cepett bgtt naiknya bulan ini kalo bulan bulan...
```

3.2.2 Case Folding

At this stage, case folding converts all letters in the text to lowercase.

```
                                            case_folding  \
0  isu ijazah palsu adalah strategi menenggelamka...
1  coba kamu buka kamu cek kasus antam memalsukan...
2  pt antam terlibat kasus korupsi pengelolaan to...
3  ya korupsi di bumn memang menjadi isu serius d...
4  wow otomatus klasemen liga korupsi indonesia p...
5  gw sih udah tau mereka sedang dimanfaatkan apa...
6  yang gerakin akun ini siapa sih gara gara kasu...
7  si om waktu lockdown gk pernah denger ada kasu...
8  emas makin mahal tapi ngeri suatu saat harga e...
9  cepett bgtt naiknya bulan ini kalo bulan bulan...

                                           tokenizing
0  [isu, ijazah, palsu, adalah, strategi, menengg...
1  [coba, kamu, buka, kamu, cek, kasus, antam, me...
2  [pt, antam, terlibat, kasus, korupsi, pengelol...
3  [ya, korupsi, di, bumn, memang, menjadi, isu, ...
4  [wow, otomatus, klasemen, liga, korupsi, indon...
5  [gw, sih, udah, tau, mereka, sedang, dimanfaat...
6  [yang, gerakin, akun, ini, siapa, sih, gara, g...
7  [si, om, waktu, lockdown, gk, pernah, denger, ...
8  [emas, makin, mahal, tapi, ngeri, suatu, saat,...
9  [cepett, bgtt, naiknya, bulan, ini, kalo, bula...
```

3.2.3 Tokenizing

Tokenization is the process of breaking sentences into individual words.

	tokenizing
0	[isu, ijazah, palsu, adalah, strategi, menengg...
1	[coba, kamu, buka, kamu, cek, kasus, antam, me...
2	[pt, antam, terlibat, kasus, korupsi, pengelol...
3	[ya, korupsi, di, bumn, memang, menjadi, isu, ...
4	[wow, otomatus, klasemen, liga, korupsi, indon...
5	[gw, sih, udah, tau, mereka, sedang, dimanfaat...
6	[yang, gerakin, akun, ini, siapa, sih, gara, g...
7	[si, om, waktu, lockdown, gk, pernah, denger, ...
8	[emas, makin, mahal, tapi, ngeri, suatu, saat,...
9	[cepett, bgtt, naiknya, bulan, ini, kalo, bula...

	normalisasi
0	[isu, ijazah, palsu, adalah, strategi, menengg...
1	[coba, kamu, buka, kamu, cek, kasus, antam, me...
2	[pt, antam, terlibat, kasus, korupsi, pengelol...
3	[ya, korupsi, di, badan usaha milik negara, me...
4	[wow, otomatus, klasemen, liga, korupsi, indon...
5	[aku, sih, udah, tau, mereka, sedang, dimanfaa...
6	[yang, gerakin, akun, ini, siapa, sih, gara, g...
7	[si, orang, waktu, lockdown, tidak, pernah, de...
8	[emas, makin, mahal, tapi, ngeri, suatu, saat,...
9	[cepett, banget, naiknya, bulan, ini, kalo, bu...

3.2.4 Normalization

The normalization process is intended to standardize non-standard words according to the Indonesian Dictionary (KBBI), including abbreviations.

	normalisasi
0	[isu, ijazah, palsu, adalah, strategi, menengg...
1	[coba, kamu, buka, kamu, cek, kasus, antam, me...
2	[pt, antam, terlibat, kasus, korupsi, pengelol...
3	[ya, korupsi, di, badan usaha milik negara, me...
4	[wow, otomatus, klasemen, liga, korupsi, indon...
5	[aku, sih, udah, tau, mereka, sedang, dimanfaa...
6	[yang, gerakin, akun, ini, siapa, sih, gara, g...
7	[si, orang, waktu, lockdown, tidak, pernah, de...
8	[emas, makin, mahal, tapi, ngeri, suatu, saat,...
9	[cepett, banget, naiknya, bulan, ini, kalo, bu...

	stopword_removal
0	[isu, ijazah, palsu, strategi, menenggelamkan,...
1	[coba, buka, cek, antam, memalsukan, emas, rek...
2	[pt, antam, terlibat, korupsi, pengelolaan, to...
3	[ya, korupsi, badan usaha milik negara, isu, s...
4	[wow, otomatus, klasemen, liga, korupsi, indon...
5	[sih, udah, tau, dimanfaatkan, antam, anjlokny...
6	[gerakin, akun, sih, gara, gara, berita, korup...
7	[si, orang, lockdown, denger, emas, antam, pal...
8	[emas, mahal, ngeri, harga, emas, murah, cadan...
9	[cepett, banget, naiknya, kalo, kemarin, pas, ...

3.2.5 Stopword Removal

This stage consists of eliminating words deemed insignificant within a sentence.

	stopword_removal
0	[isu, ijazah, palsu, strategi, menenggelamkan,...
1	[coba, buka, cek, antam, memalsukan, emas, rek...
2	[pt, antam, terlibat, korupsi, pengelolaan, to...
3	[ya, korupsi, badan usaha milik negara, isu, s...
4	[wow, otomatus, klasemen, liga, korupsi, indon...
5	[sih, udah, tau, dimanfaatkan, antam, anjlokny...
6	[gerakin, akun, sih, gara, gara, berita, korup...
7	[si, orang, lockdown, denger, emas, antam, pal...
8	[emas, mahal, ngeri, harga, emas, murah, cadan...
9	[cepett, banget, naiknya, kalo, kemarin, pas, ...

	stemming
0	[isu, ijazah, palsu, strategi, tenggelam, koru...
1	[coba, buka, cek, antam, palsu, emas, rekor, k...
2	[pt, antam, libat, korupsi, kelola, ton, emas,...
3	[ya, korupsi, badan usaha milik negara, isu, s...
4	[wow, otomatus, klasemen, liga, korupsi, indon...
5	[sih, udah, tau, manfaat, antam, anjlok, harga...
6	[gerakin, akun, sih, gara, gara, berita, korup...
7	[si, orang, lockdown, denger, emas, antam, pal...
8	[emas, mahal, ngeri, harga, emas, murah, cadan...
9	[cepett, banget, naik, kalo, kemarin, pas, bel...

3.2.6 Stemming

```
                                        stopword_removal  \
0  [isu, ijazah, palsu, strategi, menenggelamkan,...
1  [coba, buka, cek, antam, memalsukan, emas, rek...
2  [pt, antam, terlibat, korupsi, pengelolaan, to...
3  [ya, korupsi, badan usaha milik negara, isu, s...
4  [wow, otomatus, klasemen, liga, korupsi, indon...
5  [sih, udah, tau, dimanfaatkan, antam, anjlokny...
6  [gerakin, akun, sih, gara, gara, berita, korup...
7  [si, orang, lockdown, denger, emas, antam, pal...
8  [emas, mahal, ngeri, harga, emas, murah, cadan...
9  [cepett, banget, naiknya, kalo, kemarin, pas, ...

                                             stemming
0  [isu, ijazah, palsu, strategi, tenggelam, koru...
1  [coba, buka, cek, antam, palsu, emas, rekor, k...
2  [pt, antam, libat, korupsi, kelola, ton, emas,...
3  [ya, korupsi, badan usaha milik negara, isu, s...
4  [wow, otomatus, klasemen, liga, korupsi, indon...
5  [sih, udah, tau, manfaat, antam, anjlok, harga...
6  [gerakin, akun, sih, gara, gara, berita, korup...
7  [si, orang, lockdown, denger, emas, antam, pal...
8  [emas, mahal, ngeri, harga, emas, murah, cadan...
9  [cepett, banget, naik, kalo, kemarin, pas, bel...
```

Stemming is a process that restores words to their base form by removing prefixes, infixes, and suffixes.

3.2.7 N-Gram

The N-Gram method is used to extract text features in order to identify word patterns and sequences. In this study,

```
                                             unigram  \
0  [isu, ijazah, palsu, strategi, tenggelam, koru...
1  [coba, buka, cek, antam, palsu, emas, rekor, k...
2  [pt, antam, libat, korupsi, kelola, ton, emas,...
3  [ya, korupsi, badan usaha milik negara, isu, s...
4  [wow, otomatus, klasemen, liga, korupsi, indon...

                                              bigram  \
0  [isu ijazah, ijazah palsu, palsu strategi, str...
1  [coba buka, buka cek, cek antam, antam palsu, ...
2  [pt antam, antam libat, libat korupsi, korupsi...
3  [ya korupsi, korupsi badan usaha milik negara,...
4  [wow otomatus, otomatus klasemen, klasemen lig...

                                             trigram
0  [isu ijazah palsu, ijazah palsu strategi, pals...
1  [coba buka cek, buka cek antam, cek antam pals...
2  [pt antam libat, antam libat korupsi, libat ko...
3  [ya korupsi badan usaha milik negara, korupsi ...
4  [wow otomatus klasemen, otomatus klasemen liga...
```

feature representation is enhanced by combining unigrams, bigrams, and trigrams in the text.

3.2.8 TF-IDF

TF-IDF is used to assign weights to words based on their importance within a document, allowing the most relevant words to be distinguished in text analysis.

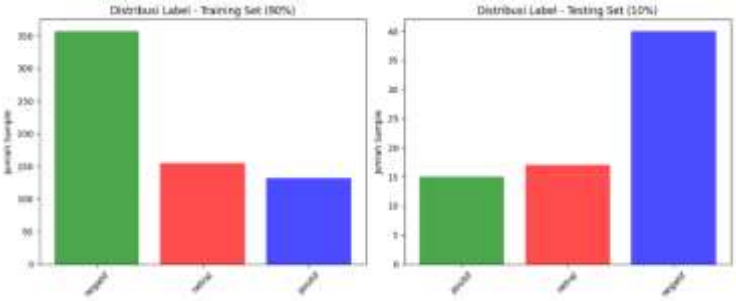
```
Top 10 unigrams:
antam                          0.050600
emas                           0.026510
korupsi                        0.023997
pt                             0.020451
pertamina                      0.019457
palsu                          0.017084
bang                           0.013522
rugi                           0.012123
ya                             0.010173
aja                            0.009346

Top 10 bigrams:
pt antam                       0.019721
emas antam                     0.013997
antam bang                     0.013459
korupsi antam                  0.010908
pertamina antam                0.009952
rugi negara                    0.007917
emas palsu                     0.007906
antam rugi                     0.006887
rp kuadriliun                  0.006811
negara rp                      0.006657

Top 10 trigrams:
rugi negara rp                 0.006657
antam rugi negara              0.006108
negara rp kuadriliun           0.004700
alih isu korupsi               0.004585
palsu antam kak                0.004321
korupsi pertamina antam        0.003550
ton emas antam                 0.003143
pt antam korupsi               0.003041
antam bener gaksi              0.003014
menantu mulyono amp            0.003014
```

3.2.9 Support Vector Machine

Sentiment classification was performed using a Support Vector Machine (SVM) with a linear kernel on features derived from N-grams and TF-IDF, with a 90:10 split for training and testing data based on expert-labeled



annotations.

3.2.10 Evaluation

The Evaluasi The performance of the SVM model was evaluated using accuracy, precision, recall, and F1-score to measure the effectiveness of sentiment classification.

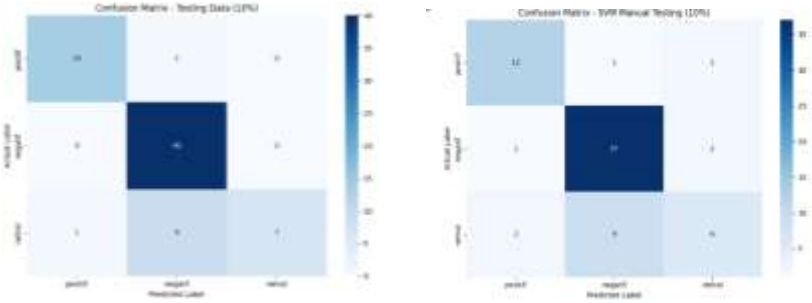


Figure 4.12 shows that the SVM model with N-Gram features performs better compared to the model without N-Gram. The model with N-Gram achieved an accuracy of 85%, performing very well on negative sentiment, reasonably well on positive sentiment, but still facing challenges in classifying neutral sentiment. In contrast, the SVM model without N-Gram yielded a lower accuracy of 76%, with a notable decline in performance, particularly for the neutral class. These results indicate that incorporating N-Gram features enhances the model's overall sentiment classification capability, although neutral sentiment classification remains a challenge.

4. Conclusion

This study demonstrates that the Support Vector Machine (SVM) method is capable of effectively classifying public sentiment regarding the Antam gold case on social media platform X into positive, negative, and neutral categories. The data processing workflow, which included text preprocessing, feature extraction using N-grams, TF-IDF weighting, and SVM modeling, achieved strong classification performance with an accuracy of 85%. These results confirm that SVM is a reliable method for analyzing text-based public opinion on social media, although challenges remain in optimally classifying neutral sentiment.

References

[1] El Farabi, Q. N. S. (2023). Kehadiran Media Sosial Dan Partisipasi Politik Bagi Pemilih Pemula. CommLine, 7(2), 112. https://doi.org/10.36722/cl.v7i2.1381.

[2] Dini, P. (2025). Pengaruh Komunikasi Digital Melalui Twitter Terhadap Persepsi. 5(1), 16–31. https://doi.org/10.53363/bureau.v5i1.501

[3] Ernianti Hasibuan, Heriyanto EA. Analisis Sentimen Pada Ulasan Aplikasi Amazon Shopping di Google Play Store Menggunakan Naive Bayes Classifier. Jurnal Teknik dan Science. 2022;1(3):13–24.

[4] Irawan, D., Perkasa, E. B., Yurindra, Y., Wahyuningsih, D., & Helmud, E. (2021). Perbandingan Klassifikasi SMS Berbasis Support Vector Machine, Naive Bayes Classifier, Random Forest dan Bagging Classifier. Jurnal Sisfokom (Sistem Informasi Dan Komputer), 10(3), 432–437. https://doi.org/10.32736/sisfokom.v10i3.1302.

[5] Fide, S., Suparti, S., & Sudarno, S. (2021). Analisis Sentimen Ulasan Aplikasi Tiktok Di Google Play Menggunakan Metode

Support Vector Machine (Svm) Dan Asosiasi. Jurnal Gaussian, 10(3), 346–358. https://doi.org/10.14710/j.gauss.v10i3.32786.

[6] Tinaliah, T., & Elizabeth, T. (2022). Analisis Sentimen Ulasan Aplikasi PrimaKu Menggunakan Metode Support Vector Machine. JATISI (Jurnal Teknik Informatika Dan Sistem Informasi), 9(4), 3436–3442. https://doi.org/10.35957/jatisi.v9i4.3586.

[7] Arsi, P., & Waluyo, R. (2021). Sentiment Analysis of Discourse on Moving the Indonesian Capital City Using the Support Vector Machine (SVM) Algorithm. Jurnal Teknologi Informasi Dan Ilmu Komputer, 8(1), 147. https://doi.org/10.25126/jtiik.202183944.