



Analisis Sentimen Masyarakat terhadap Pembangunan Ibu Kota Nusantara (IKN) menggunakan Algoritma *Multinomial Naive Bayes* dan *SMOTE*

Kevin Kristian Bram Barus¹, Kevinro Parmonangan Manalu¹, Mona Ghassan¹, Okta Br Hutasoit¹, Riza Saskia Rani¹, Vagle Siburian¹, Zaqiatunnisa^{1,*}

¹Universitas Mandiri Bina Prestasi, Medan, Indonesia

Informasi Artikel

Sejarah Artikel:

Submit: 02 Maret 2026

Revisi: 10 Maret 2026

Diterima: 17 Maret 2026

Diterbitkan: 30 Maret 2026

Kata Kunci

Analisis Sentimen, Machine Learning, Naive Bayes, TF-IDF, SMOTE

Correspondence

E-mail: zaqiatunnisa15@gmail.com*

A B S T R A K

Penelitian ini membahas analisis sentimen masyarakat terhadap pembangunan Ibu Kota Nusantara (IKN) di Indonesia yang menjadi salah satu isu besar dan banyak diperbincangkan di media sosial. Berbagai opini yang muncul menunjukkan respons yang beragam, mulai dari dukungan, harapan, hingga kritik dan penolakan. Kondisi ini menjadikan analisis sentimen sebagai pendekatan yang relevan untuk memahami persepsi publik secara lebih sistematis dan berbasis data. Penelitian ini bertujuan untuk mengidentifikasi dan mengklasifikasikan sentimen masyarakat terhadap IKN menggunakan pendekatan *Machine Learning* dengan algoritma *Multinomial Naive Bayes* yang berbasis pada *Teorema Bayes*. Data yang digunakan berupa komentar pengguna media sosial yang kemudian diproses melalui beberapa tahapan, yaitu *preprocessing* (pembersihan teks), pelabelan data berbasis kata kunci, ekstraksi fitur menggunakan metode TF-IDF untuk mengubah data teks menjadi bentuk numerik, serta penyeimbangan data menggunakan metode *SMOTE* untuk mengatasi ketidakseimbangan jumlah data antar kelas. Hasil penelitian menunjukkan bahwa model *Machine Learning* yang dibangun mampu mengklasifikasikan sentimen dengan cukup baik, terutama dalam mendeteksi sentimen positif, meskipun masih terdapat beberapa kesalahan dalam mengidentifikasi sentimen negatif. Secara keseluruhan, penelitian ini menunjukkan bahwa penerapan *Machine Learning* menggunakan algoritma *Naive Bayes* efektif dalam analisis sentimen dan berpotensi menjadi alat bantu dalam memahami opini publik terhadap kebijakan pemerintah, khususnya terkait pembangunan Ibu Kota Nusantara (IKN).

Abstract

This study discusses public sentiment analysis regarding the development of Indonesia's Nusantara Capital City (IKN), which has become one of the major issues widely discussed on social media. The various opinions expressed reflect diverse responses, ranging from support and expectations to criticism and opposition. This situation makes sentiment analysis a relevant approach for understanding public perceptions in a more systematic and data-driven manner. The objective of this study is to identify and classify public sentiment toward IKN using a Machine Learning approach with the Multinomial Naive Bayes algorithm, which is based on Bayes' Theorem. The data used consist of social media user comments, which were processed through several stages, including text preprocessing, keyword-based data labeling, feature extraction using the TF-IDF method to transform textual data into numerical representations, and data balancing using the Synthetic Minority Over-sampling Technique (SMOTE) to address class imbalance. The results indicate that the developed Machine Learning model is capable of classifying sentiments with satisfactory performance, particularly in detecting positive sentiments, although some misclassifications still occurred in identifying negative sentiments. Overall, this study demonstrates that the application of Machine Learning using the Naive Bayes algorithm can be effectively utilized for sentiment analysis and has the potential to serve as a useful tool for understanding public opinion regarding government policies, particularly those related to the development of the Nusantara Capital City (IKN).



1. Pendahuluan

Perkembangan teknologi digital yang semakin pesat, khususnya dalam bidang komunikasi dan informasi, telah membuat media sosial menjadi salah satu sarana utama bagi masyarakat untuk menyampaikan opini dan pandangannya terhadap berbagai isu. Media sosial kini tidak hanya digunakan sebagai alat komunikasi, tetapi juga sebagai ruang diskusi terbuka yang memungkinkan masyarakat berbagi informasi dan mengekspresikan pendapat secara cepat dan luas [4].

Salah satu isu yang banyak menarik perhatian publik adalah pembangunan Ibu Kota Nusantara (IKN) di Indonesia. Kebijakan ini menimbulkan berbagai tanggapan dari masyarakat, mulai dari dukungan karena dianggap dapat mendorong pemerataan pembangunan, hingga kritik yang berkaitan dengan dampak ekonomi, lingkungan, dan sosial. Beragam opini tersebut banyak ditemukan di *platform* media sosial seperti TikTok yang memiliki tingkat interaksi tinggi melalui fitur komentar, sehingga dapat mencerminkan persepsi masyarakat secara langsung [8].

Namun, banyaknya data komentar yang terus bertambah setiap waktu membuat proses analisis secara manual menjadi kurang efektif, memakan waktu, dan berpotensi menimbulkan subjektivitas. Oleh karena itu, diperlukan metode yang mampu mengolah data secara otomatis, cepat, dan objektif. Salah satu pendekatan yang dapat digunakan adalah analisis sentimen, yaitu: teknik dalam bidang *Natural Language Processing* (NLP) yang bertujuan untuk mengidentifikasi dan mengelompokkan opini masyarakat ke dalam kategori tertentu, seperti positif dan negatif [9][10].

Seiring berkembangnya teknologi, analisis sentimen banyak dikombinasikan dengan metode *Machine Learning* untuk meningkatkan kemampuan klasifikasi data teks [3][10]. Dalam penelitian ini, digunakan algoritma *Multinomial Naive Bayes* yang merupakan implementasi dari *Teorema Bayes* dalam proses klasifikasi. Algoritma ini dipilih karena memiliki kemampuan yang baik dalam mengolah data teks, efisien dalam proses komputasi, serta mampu menghasilkan performa yang cukup baik pada tugas klasifikasi sentimen [2][4]. Dengan memanfaatkan *Machine Learning* berbasis *Naive Bayes*, data komentar yang awalnya tidak terstruktur dapat diolah menjadi informasi yang lebih sistematis sehingga dapat memberikan gambaran mengenai kecenderungan opini masyarakat terhadap pembangunan IKN [2][10].

2. Metode Penelitian

2.1. Tahapan Penelitian

Penelitian ini dilakukan melalui beberapa tahapan, yaitu pengumpulan data komentar TikTok terkait IKN, *preprocessing* data, pelabelan sentimen, ekstraksi fitur menggunakan TF-IDF, penyeimbangan data dengan *SMOTE*, pembagian data menjadi training dan testing, pelatihan model menggunakan algoritma *Multinomial Naive Bayes*, serta evaluasi model menggunakan *accuracy*, *classification report*, dan *confusion matrix*.

1. Pengumpulan Data

Dataset berasal dari komentar TikTok terkait IKN dalam format Excel. Data komentar TikTok yang digunakan dikumpulkan dalam rentang waktu tertentu dengan jumlah data awal yang besar. Setelah dilakukan *preprocessing* dan pembersihan data, ukuran dataset disesuaikan untuk mendapatkan data yang valid dan representatif.

2. *Preprocessing* Data

Langkah-langkah *preprocessing* data meliputi menghapus URL dan simbol, mengubah teks menjadi huruf kecil, menghapus stopwords, dan membersihkan karakter tidak penting. Teknik *SMOTE* diterapkan untuk mengatasi ketidakseimbangan jumlah data antar kelas sentimen. *SMOTE* bekerja dengan menambah data sintetis pada kelas minoritas, sehingga model tidak bias terhadap kelas mayoritas dan performa klasifikasi menjadi lebih baik.

3. Labeling

Proses *labeling* dilakukan secara otomatis berdasarkan kata kunci yang terdapat pada komentar. Komentar yang mengandung kata-kata positif seperti "bagus", "mantap", dan "setuju" diberi label positif (1), sedangkan komentar lainnya diberi label negatif/netral (0). Hasil pelabelan digunakan sebagai target klasifikasi pada model *Naive Bayes*.

4. Ekstraksi Fitur

Setelah proses *preprocessing*, data teks diubah menjadi bentuk numerik melalui proses ekstraksi fitur. Pada penelitian ini digunakan metode TF-IDF (*Term Frequency-Inverse Document Frequency*) untuk mengukur tingkat kepentingan suatu kata dalam dokumen. Hasil ekstraksi fitur digunakan sebagai input dalam proses pelatihan model klasifikasi sentimen.

5. Penyeimbangan Data (*SMOTE*)

Chawla dkk. mengemukakan bahwa metode *SMOTE* (*Synthetic Minority Over-sampling Technique*) dapat digunakan untuk mengatasi ketidakseimbangan data dengan menghasilkan sampel sintetis pada kelas minoritas sehingga meningkatkan performa model klasifikasi [5].

6. Pembagian Data

Setelah proses *preprocessing*, *labeling*, ekstraksi fitur TF-IDF, dan penyeimbangan data menggunakan *SMOTE* selesai dilakukan, dataset dibagi menjadi data latih (*training data*) dan data uji (*testing data*). Pembagian data dilakukan dengan rasio 80% untuk data latih dan 20% untuk data uji menggunakan metode *train-test split*. Data latih digunakan untuk melatih model *Multinomial Naive Bayes*, sedangkan data uji digunakan untuk mengevaluasi kinerja model dalam mengklasifikasikan sentimen.

7. Training Model

Pada tahap ini, data latih (*training data*) digunakan untuk melatih model klasifikasi menggunakan algoritma *Multinomial Naive Bayes*. Algoritma ini dipilih karena memiliki performa yang baik dalam pengolahan data teks serta mampu melakukan klasifikasi dengan cepat dan efisien. Proses pelatihan dilakukan dengan mempelajari pola hubungan antara fitur TF-IDF dan label sentimen yang telah diberikan sebelumnya, sehingga model dapat digunakan untuk memprediksi sentimen pada data baru.

8. Evaluasi

Evaluasi model dilakukan untuk mengukur kemampuan algoritma *Multinomial Naive Bayes* dalam mengklasifikasikan sentimen komentar terhadap IKN. Pengujian dilakukan menggunakan data uji (*testing data*) dengan metrik evaluasi berupa *accuracy*, *classification report*, dan *confusion matrix*. Hasil evaluasi digunakan untuk mengetahui tingkat ketepatan model serta mengidentifikasi kesalahan klasifikasi antara sentimen positif dan negatif/netral.

2.2. Naive Bayes

Naive Bayes merupakan algoritma klasifikasi berbasis probabilitas yang mengasumsikan bahwa setiap fitur bersifat independen. Menurut Wiguna dan Ginting, Algoritma *Naive Bayes* dikenal memiliki kecepatan komputasi yang tinggi dan sangat efisien dalam menangani data teks yang

besar... meskipun didasarkan pada asumsi independensi yang sederhana, algoritma ini terbukti memiliki akurasi sangat baik [1].

Pemilihan algoritma *Naive Bayes* didasarkan pada kemampuannya yang efisien dan cepat dalam mengolah data teks besar. Algoritma ini cocok untuk kasus analisis sentimen karena menggunakan pendekatan probabilistik yang sederhana namun efektif, terutama dalam klasifikasi teks dengan fitur independen. Dalam implementasinya, algoritma *Naive Bayes* dapat di lihat pada script yang sudah di analisis pada *google colab* berbasis python di bawah ini.

```
import pandas as pd
import re
import nltk
from nltk.corpus import stopwords
from sklearn.model_selection import train_test_split
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.naive_bayes import MultinomialNB
from sklearn.metrics import classification_report, accuracy_score, confusion_matrix

# --- TAMBAHAN LIBRARY SMOTE ---
from imblearn.over_sampling import SMOTE
# --- TAMBAHAN LIBRARY PLOTTING ---
import matplotlib.pyplot as plt
import seaborn as sns

#1. Menghubungkan Google Colab
from google.colab import drive
drive.mount('/content/drive')

#2. Load Dataset
excel_file_path = '/content/drive/MyDrive/data_IKN/komentar Tiktok IKN Data.xlsx'
df = pd.read_excel(excel_file_path)

#3. Simple Labeling (Tetap sama)
def manual_label(text):
    positive_words = ['bagus', 'mantap', 'setuju', 'terima kasih', 'dukung', 'siap']
    text = str(text).lower()
    if any(positive_word in text for positive_word in positive_words):
        return 1 # Positif
    else:
        return 0 # Negatif/Netral

#11. Grafik Batang Distribusi Prediksi
pred_counts = pd.Series(pred.value_counts())
# Ensure labels are in the correct order for class 0 and 1
pred_labels = ['Negatif', 'Positif']
pred_values = [pred_counts.get(0, 0), pred_counts.get(1, 0)] # Get counts for 0 and 1, default to 0 if not present

plt.figure(figsize=(7, 5))
sns.barplot(x=pred_labels, y=pred_values, palette='viridis')
plt.title('Distribusi Kelas Prediksi')
plt.xlabel('Sentimen')
plt.ylabel('Jumlah')
plt.show()

#11. Pie Chart Distribusi Prediksi
plt.figure(figsize=(7, 7))
plt.pie(pred_values, labels=pred_labels, autopct='%1.1f%%', startangle=90, colors=['lightcoral', 'lightyellow'])
plt.title('Proporsi Kelas Prediksi')
plt.axis('equal') # Equal aspect ratio ensures that pie is drawn as a circle.
plt.show()

# Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount('/content/drive', force_remount=True).
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data] Package stopwords is already up-to-date!
Jumlah data setelah SMOTE: 2718
Distribusi Kelas:
0    1359
1    1359

return 0 # Negatif/Netral

df['label'] = df['text'].apply(manual_label)

#4. Preprocessing Function (Tetap sama)
nltk.download('stopwords')
stop_words = set(stopwords.words('indonesian'))

def clean_text(text):
    text = re.sub(r'http[s]?://(?:[a-zA-Z]+\.)+[a-zA-Z]+', '', str(text), flags=re.MULTILINE)
    text = re.sub(r'@user', '', text)
    text = re.sub(r'#[a-zA-Z]+', '', text)
    text = text.lower()
    text = ' '.join(word for word in text.split() if word not in stop_words)
    return text

df['cleaned_text'] = df['text'].apply(clean_text)

# --- PERUBAHAN MULAI DI SINI ---

#5. Feature Extraction (TF-IDF) dilakukan SEBELUM Split agar bisa di-SMOTE
tfidf_vectorizer = TfidfVectorizer(max_features=5000)
X_tfidf = tfidf_vectorizer.fit_transform(df['cleaned_text'])
y = df['label']

#6. PENERAPAN SMOTE
# Di sini kita bikin data kelas 1 jadi seimbang dengan kelas 0
smote = SMOTE(random_state=42)
X_resampled, y_resampled = smote.fit_resample(X_tfidf, y)

print("Jumlah data setelah SMOTE: (len(y_resampled))")
print("Distribusi kelas: (y_resampled.value_counts())")

#7. Split Data (Pakai data yang sudah di-resample)
X_train, X_test, y_train, y_test = train_test_split(X_resampled, y_resampled, test_size=0.2, random_state=42)

#8. Model Training (Naive Bayes)
model = MultinomialNB()
model.fit(X_train, y_train)

#9. Model Evaluation
y_pred = model.predict(X_test)

print("\n--- Evaluasi Model (Setelah SMOTE) ---")
print(f"Accuracy: (accuracy_score(y_test, y_pred) * 100.2f)%"
print(f"Classification Report:\n", classification_report(y_test, y_pred))
print(f"Confusion Matrix:\n", confusion_matrix(y_test, y_pred))

# --- VISUALISASI TAMBAHAN ---

#10. Visualisasi Confusion Matrix (kotak)
cm = confusion_matrix(y_test, y_pred)
plt.figure(figsize=(8, 6))
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues', cbar=False,
            xticklabels=['Predicted Negatif', 'Predicted Positif'],
            yticklabels=['Actual Negatif', 'Actual Positif'])
plt.title('Confusion Matrix')
plt.xlabel('Predicted label')
plt.ylabel('True label')

print("\n--- Evaluasi Model (Setelah SMOTE) ---")
print(f"Accuracy: (accuracy_score(y_test, y_pred) * 100.2f)%"
print(f"Classification Report:\n", classification_report(y_test, y_pred))
print(f"Confusion Matrix:\n", confusion_matrix(y_test, y_pred))

# --- VISUALISASI TAMBAHAN ---

#10. Visualisasi Confusion Matrix (kotak)
cm = confusion_matrix(y_test, y_pred)
plt.figure(figsize=(8, 6))
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues', cbar=False,
            xticklabels=['Predicted Negatif', 'Predicted Positif'],
            yticklabels=['Actual Negatif', 'Actual Positif'])
plt.title('Confusion Matrix')
plt.xlabel('Predicted label')
plt.ylabel('True label')
```

Gambar 1. Implementasi Program Analisis Sentimen IKN Menggunakan Python

3. Hasil dan Pembahasan

Penelitian ini menerapkan konsep *Machine Learning*. Raschka dan Mirjalili menyatakan bahwa *Machine Learning* merupakan pendekatan yang memungkinkan komputer mempelajari pola dari data sehingga dapat digunakan untuk melakukan prediksi maupun klasifikasi secara otomatis [3].

Model yang digunakan adalah *Multinomial Naive Bayes*, yang dilatih menggunakan data komentar yang telah melalui tahap *preprocessing*, *labeling*, ekstraksi fitur TF-IDF, dan penyeimbangan data menggunakan *SMOTE*.

Hasil pengujian menunjukkan nilai *True Positive* (277), *True Negative* (207), *False Positive* (60), dan *False Negative* (0). Hasil tersebut menunjukkan bahwa model *Machine Learning* yang dibangun mampu mengklasifikasikan sentimen dengan cukup baik, terutama dalam mendeteksi komentar positif.

```
Distribusi kelas:
label
0    1359
1    1359
Name: count, dtype: int64

--- Evaluasi Model (Setelah SMOTE) ---
Accuracy: 88.97%
Classification Report:
              precision    recall  f1-score   support

     0       1.00        0.78    0.87       267
     1       0.82        1.00    0.90       277

 accuracy          0.91    0.89    0.89       544
 macro avg          0.91    0.89    0.89       544
 weighted avg          0.91    0.89    0.89       544

Confusion Matrix:
[[207  60]
 [  0 277]]
```

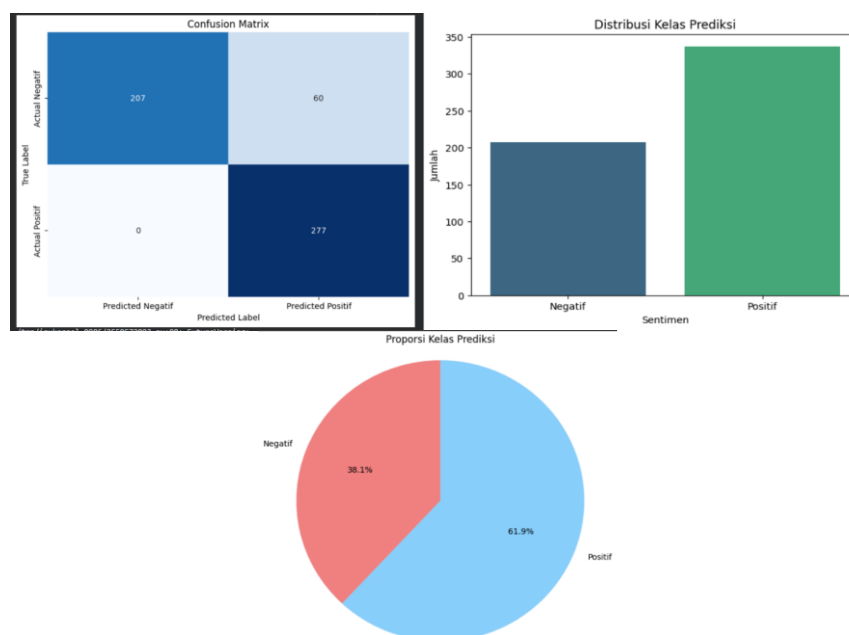
Gambar 2. Arsitektur Badan Manusia

Secara keseluruhan, penerapan *Machine Learning* dengan algoritma *Naive Bayes* terbukti efektif untuk analisis sentimen terhadap IKN. Namun, masih diperlukan pengembangan lebih lanjut untuk mengurangi kesalahan klasifikasi pada komentar negatif atau netral sehingga performa model dapat menjadi lebih baik. Kalimat ini lebih menonjolkan bahwa yang kalian bangun adalah model *Machine Learning* untuk klasifikasi sentimen, bukan sekadar analisis statistik biasa.

3.1. Hasil Confusion Matrix

Hasil *confusion matrix* menunjukkan bahwa model memiliki performa yang sangat baik dalam mendeteksi sentimen positif dengan nilai *False Negative* sebesar 0, namun masih terdapat kesalahan dalam klasifikasi sentimen negatif yang ditunjukkan oleh nilai *False Positive* sebesar 60. Selain *confusion matrix*, evaluasi model juga menggunakan metrik *precision*, *recall*, dan *f1-score* agar penilaian performa model lebih komprehensif dan detail.

Kesalahan klasifikasi sentimen negatif dapat mempengaruhi interpretasi hasil analisis, terutama dalam konteks pengambilan keputusan kebijakan. Oleh karena itu, peningkatan akurasi pada sentimen negatif sangat penting untuk memperoleh gambaran opini masyarakat yang lebih akurat dan seimbang.



Gambar 3. Visualisasi *Confusion Matrix* Model *Multinomial Naive Bayes*

3.2. Kelebihan dan Kekurangan

Penelitian ini memiliki beberapa kelebihan yang mendukung efektivitas analisis sentimen yang dilakukan. Pertama, penelitian menggunakan pendekatan *Machine Learning* yang mampu mengklasifikasikan sentimen secara otomatis sehingga proses analisis menjadi lebih cepat dan efisien dibandingkan metode manual. Kedua, algoritma *Multinomial Naive Bayes* yang digunakan memiliki proses komputasi yang sederhana, cepat, dan efektif dalam menangani data teks berukuran besar. Ketiga, penerapan metode *Term Frequency-Inverse Document Frequency* (TF-IDF) berhasil mengubah data teks menjadi representasi numerik yang dapat diproses secara optimal oleh model klasifikasi. Selain itu, penggunaan teknik *Synthetic Minority Over-sampling Technique* (SMOTE) mampu mengatasi permasalahan ketidakseimbangan kelas pada data sehingga meningkatkan performa model dalam melakukan klasifikasi sentimen.

Di sisi lain, penelitian ini juga memiliki beberapa keterbatasan yang perlu menjadi perhatian. Proses pelabelan data masih menggunakan metode berbasis kata kunci sehingga belum mampu memahami konteks kalimat secara menyeluruh, yang berpotensi menyebabkan kesalahan dalam pemberian label sentimen. Selain itu, model masih mengalami kekeliruan dalam mengklasifikasikan sebagian komentar, terutama yang memiliki karakteristik sentimen negatif atau netral dengan makna yang ambigu. Keterbatasan lainnya adalah data penelitian hanya berasal dari satu *platform* media sosial sehingga hasil analisis belum sepenuhnya merepresentasikan opini masyarakat secara luas. Di samping itu, model yang dikembangkan belum mampu mendeteksi makna bahasa yang kompleks, seperti sarkasme, ironi, maupun ungkapan ambigu, sehingga akurasi klasifikasi pada kondisi tersebut masih perlu ditingkatkan.

4. Kesimpulan

Penelitian ini berhasil menerapkan *Machine Learning* menggunakan algoritma *Multinomial Naive Bayes* untuk melakukan analisis sentimen terhadap komentar masyarakat mengenai pembangunan Ibu Kota Nusantara (IKN). Proses analisis dilakukan melalui tahapan *preprocessing*, *labeling*, ekstraksi fitur menggunakan TF-IDF, penyeimbangan data menggunakan SMOTE, pelatihan model, dan evaluasi. Hasil penelitian menunjukkan bahwa model mampu mengklasifikasikan sentimen dengan cukup baik, terutama dalam mendeteksi sentimen positif. Hal ini menunjukkan bahwa algoritma *Naive Bayes* efektif digunakan untuk analisis sentimen pada data teks. Meskipun demikian, masih terdapat beberapa kesalahan klasifikasi pada sentimen negatif atau netral sehingga diperlukan pengembangan lebih lanjut untuk meningkatkan performa model. Secara keseluruhan, penelitian ini membuktikan bahwa penerapan *Machine Learning* dapat membantu memahami opini masyarakat terhadap pembangunan IKN secara lebih cepat, sistematis, dan objektif.

Daftar Pustaka

- [1] Wiguna, D.P., & Ginting, S. E. (2026). Analisis Sentimen E-Commerce di Indonesia dengan Algoritma *Naive Bayes* (Studi Kasus Pada Platform Shopee, Tokopedia, Bukalapak, dan Lazada). *AICOM: Artificial Intelligence and Computing*, vol. 01, no. 04), 109-114.
- [2] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. New York, NY, USA: Cambridge University Press, 2008.
- [3] S. Raschka and V. Mirjalili, *Python Machine Learning: Machine Learning and Deep Learning with Python, Scikit-Learn, and TensorFlow*. Birmingham, UK: Packt Publishing, 2022.
- [4] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. Waltham, MA, USA: Morgan Kaufmann, 2012.
- [5] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321-357, 2002.
- [6] G. Salton and C. Buckley, "Term-Weighting Approaches in Automatic Text Retrieval," *Information Processing & Management*, vol. 24, no. 5, pp. 513-523, 1988.

- [7] T. Joachims, "Text Categorization with Support Vector Machines: Learning with Many Relevant Features," in *Proceedings of the European Conference on Machine Learning*, Berlin, Germany, 1998, pp. 137–142.
- [8] Kementerian Sekretariat Negara Republik Indonesia, "*Ibu Kota Nusantara (IKN)*," 2024.
- [9] S. Bird, E. Klein, and E. Loper, *Natural Language Processing with Python*. Sebastopol, CA, USA: O'Reilly Media, 2009.
- [10] A. Rajput and M. Ahmad, "Sentiment Analysis Using *Machine Learning* Techniques: A Review," *International Journal of Computer Applications*, vol. 177, no. 29, pp. 1–7, 2020.