

ANALISIS SENTIMEN PENDAPAT MASYARAKAT TERKAIT PEMBLOKIRAN REKENING NGANGGUR OLEH PPATK PADA MEDIA SOSIAL TWITTER MENGGUNAKAN ALGORITMA NAIVE BAYES

Silvana Putriani Wijaya¹ Irfan Ali²

Program Studi Manajemen Informatika¹
Program Studi Rekayasa Perangkat Lunak²

STMIK IKMI Cirebon
<https://ikmi.ac.id/page/18/?lang=de>
Putrisilvana408@gmail.com
irfanali0.0@gmail.com

(*) Corresponding Author : Putrisilvana408@gmail.com

Published: 30 Desember 2025

Abstract— The Financial Transaction Reports and Analysis Center (PPATK) policy regarding blocking idle (inactive) accounts has generated pros and cons among the public, especially on Twitter. This study aims to analyze public sentiment regarding this policy using text mining methods. The data used were Twitter tweets relevant to the topic of blocking idle accounts. The research process was carried out through several stages, namely data collection, text cleaning (preprocessing), feature extraction using Term Frequency–Inverse Document Frequency (TF-IDF), and sentiment classification using the Naive Bayes algorithm. Model performance evaluation was carried out using accuracy, precision, recall, and F1-score metrics. The results showed that the Naive Bayes model was able to provide good classification performance with an netral 96,8%, positif 2,3%, negatif 0,9% on the test data. The distribution of public sentiment was dominated by negative sentiment, indicating public concern about the potential for incorrect blocking and the impact of this policy on account holders. A small portion of opinions were neutral or positive, reflecting the existence of parties who support PPATK's steps in preventing financial crimes. This research is expected to provide input for policy makers to develop clearer and more transparent public communication strategies, as well as provide a space for objections for affected communities.

Keywords: Sentiment Analysis, Naïve Bayes Algorithm, Twitter

Abstrak- Kebijakan Pusat Pelaporan dan Analisis Transaksi Keuangan (PPATK) terkait pemblokiran rekening idle (tidak aktif) menuai pro dan kontra di kalangan masyarakat, terutama di Twitter. Penelitian ini bertujuan menganalisis sentimen publik terkait kebijakan ini menggunakan metode text mining. Data yang digunakan berupa cuitan Twitter yang relevan dengan topik pemblokiran rekening idle. Proses penelitian dilakukan melalui beberapa tahapan, yaitu pengumpulan data, pembersihan teks (preprocessing), ekstraksi fitur menggunakan Term Frequency–Inverse Document Frequency (TF-IDF), dan klasifikasi sentimen menggunakan algoritma Naive Bayes. Evaluasi kinerja model dilakukan menggunakan metrik akurasi, presisi, recall, dan F1-score. Hasil penelitian menunjukkan bahwa model Naive Bayes mampu memberikan kinerja klasifikasi yang baik dengan netral 20,33%, positif 24,17%, dan negatif 55,50% pada data uji. Sebaran sentimen publik didominasi oleh sentimen negatif, yang menunjukkan kekhawatiran publik terhadap potensi pemblokiran yang tidak tepat dan dampak kebijakan ini terhadap pemegang rekening. Sebagian kecil opini bersifat netral atau positif, mencerminkan adanya pihak-pihak yang mendukung langkah PPATK dalam mencegah kejahatan keuangan. Penelitian ini diharapkan dapat memberikan masukan bagi para pembuat kebijakan untuk mengembangkan strategi komunikasi publik yang lebih jelas dan transparan, serta menyediakan ruang bagi masyarakat yang terdampak untuk menyampaikan keberatan.

Kata kunci: Sentiment Analysis, Naïve Bayes Algorithm, Twitter

INTRODUCTION

Perkembangan teknologi informasi yang sangat pesat telah mengubah cara masyarakat

menyampaikan dan membentuk opini publik. Media sosial tidak hanya berfungsi sebagai sarana komunikasi, tetapi juga sebagai ruang diskusi terbuka yang merefleksikan persepsi, sikap, dan emosi masyarakat terhadap berbagai isu sosial, ekonomi, dan kebijakan publik. Salah satu platform media sosial yang banyak digunakan untuk menyampaikan pendapat secara singkat dan real time adalah Twitter. Karakteristik Twitter yang bersifat terbuka, cepat, dan berbasis teks menjadikannya sumber data yang potensial untuk dianalisis menggunakan pendekatan text mining dan analisis sentimen [1][2]

Opini publik yang berkembang di media sosial memiliki pengaruh yang signifikan terhadap citra institusi dan keberhasilan implementasi suatu kebijakan. Oleh karena itu, analisis opini publik menjadi penting untuk memahami respons masyarakat secara objektif dan berbasis data. Pendekatan analisis sentimen memungkinkan pengelompokan opini masyarakat ke dalam kategori positif, negatif, dan netral, sehingga dapat memberikan gambaran umum mengenai kecenderungan sikap publik terhadap suatu isu tertentu [3][4]

Dalam konteks kebijakan publik di Indonesia, Pusat Pelaporan dan Analisis Transaksi Keuangan (PPATK) memiliki peran strategis dalam menjaga stabilitas dan keamanan sistem keuangan nasional. Salah satu kebijakan yang menimbulkan perhatian publik adalah pemblokiran rekening nganggur atau rekening tidak aktif (dormant). Kebijakan ini bertujuan untuk mencegah tindak kejahatan keuangan, seperti pencucian uang dan penyalahgunaan rekening. Namun, implementasi kebijakan tersebut juga menimbulkan pro dan kontra di kalangan masyarakat, khususnya terkait transparansi, keadilan, dan dampaknya terhadap nasabah [5]

Media sosial Twitter menjadi sarana utama masyarakat dalam mengekspresikan pendapat mereka terhadap kebijakan pemblokiran rekening nganggur oleh PPATK. Banyak pengguna Twitter menyampaikan keluhan, kritik, maupun dukungan melalui cuitan singkat yang tersebar luas dalam waktu singkat. Kondisi ini menghasilkan volume data teks yang besar dan tidak terstruktur, sehingga diperlukan teknik pengolahan data yang efektif untuk Mengekstraksi informasi yang bernilai dari opini-opini tersebut [6]

Text mining merupakan salah satu pendekatan yang digunakan untuk mengolah dan menganalisis data teks tidak terstruktur. Tahapan dalam text mining meliputi pengumpulan data, preprocessing teks, ekstraksi fitur, dan proses klasifikasi. Pada tahap ekstraksi fitur, metode Term Frequency Inverse Document Frequency (TF IDF)

sering digunakan untuk merepresentasikan teks dalam bentuk numerik agar dapat diproses oleh algoritma klasifikasi [7]. Pendekatan ini terbukti efektif dalam berbagai penelitian analisis sentimen berbasis media sosial.

Salah satu algoritma klasifikasi yang banyak digunakan dalam analisis sentimen adalah Naive Bayes. Algoritma ini didasarkan pada teori probabilitas Bayes dengan asumsi independensi antar fitur. Meskipun sederhana, Naive Bayes memiliki keunggulan dalam hal efisiensi komputasi dan kinerja yang cukup baik pada data teks pendek seperti tweet. Beberapa penelitian menunjukkan bahwa Naive Bayes mampu memberikan hasil klasifikasi yang kompetitif dibandingkan algoritma lain pada analisis sentimen opini publik [8][9]

Penggunaan Naive Bayes sebagai metode klasifikasi dalam penelitian ini juga sejalan dengan karakteristik program studi D3 Manajemen Informatika yang menekankan pada aspek implementatif dan pemanfaatan teknologi informasi dalam pemecahan masalah. Dengan menggunakan algoritma yang relatif sederhana dan mudah dipahami, penelitian ini diharapkan dapat memberikan gambaran yang jelas mengenai proses analisis sentimen, mulai dari pengolahan data hingga interpretasi hasil.

Berdasarkan latar belakang tersebut, penelitian ini berfokus pada analisis sentimen pendapat masyarakat terkait kebijakan pemblokiran rekening nganggur oleh PPATK pada media sosial Twitter menggunakan algoritma Naive Bayes. Penelitian ini diharapkan dapat memberikan kontribusi akademik dalam penerapan text mining pada analisis opini publik, serta kontribusi praktis sebagai bahan pertimbangan bagi pemangku kebijakan dalam memahami respons masyarakat terhadap kebijakan yang diterapkan.

MATERIALS AND METHODS

Penelitian ini menggunakan pendekatan kuantitatif deskriptif dengan metode text mining untuk menganalisis sentimen opini publik terhadap kebijakan pemblokiran rekening nganggur oleh Pusat Pelaporan dan Analisis Transaksi Keuangan (PPATK) pada media sosial Twitter. Pendekatan ini dipilih karena mampu mengolah data teks tidak terstruktur dalam jumlah besar secara sistematis dan objektif, sehingga dapat memberikan gambaran kecenderungan sentimen masyarakat berdasarkan data aktual yang diperoleh dari media sosial.

Data yang digunakan dalam penelitian ini berupa cuitan (tweet) berbahasa Indonesia yang berkaitan dengan kebijakan pemblokiran rekening nganggur oleh PPATK. Pengumpulan data dilakukan melalui proses scraping Twitter dengan

menggunakan kata kunci yang relevan, seperti "PPATK", "rekening nganggur", "rekening dormant", dan "pemblokiran rekening". Data yang telah terkumpul kemudian diseleksi untuk menghilangkan tweet duplikat, spam, serta konten yang tidak relevan dengan topik penelitian, sehingga diperoleh dataset yang layak untuk dianalisis.

Tahap selanjutnya adalah preprocessing data, yang bertujuan untuk membersihkan teks dan mengurangi noise pada data tweet. Preprocessing meliputi case folding untuk mengubah seluruh teks menjadi huruf kecil, tokenisasi untuk memecah kalimat menjadi kata-kata, serta stopword removal untuk menghapus kata umum yang tidak memiliki makna signifikan dalam analisis sentimen. Proses ini penting untuk meningkatkan kualitas data dan kinerja model klasifikasi.

Setelah preprocessing, data teks direpresentasikan dalam bentuk numerik melalui proses ekstraksi fitur menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF). Metode ini digunakan untuk memberikan bobot pada setiap kata berdasarkan frekuensi kemunculannya dalam dokumen dan keseluruhan dataset, sehingga kata-kata yang lebih informatif memiliki pengaruh lebih besar dalam proses klasifikasi sentimen.

Klasifikasi sentimen dilakukan menggunakan algoritma Naïve Bayes, yang dipilih karena kesederhanaan, efisiensi komputasi dan kemampuannya dalam menangani data teks pendek seperti tweet. Dataset dibagi menjadi data latih dan data uji untuk membangun dan menguji model klasifikasi. Tweet diklasifikasikan ke dalam tiga kelas sentimen, yaitu positif, negatif, dan netral, berdasarkan probabilitas kemunculan kata pada masing-masing kelas.

Evaluasi kinerja model dilakukan untuk mengukur tingkat keberhasilan algoritma Naïve Bayes dalam mengklasifikasikan sentimen opini publik. Metode evaluasi yang digunakan meliputi confusion matrix, accuracy, precision, recall, dan F1-score. Hasil evaluasi ini digunakan sebagai dasar analisis pada tahap pembahasan untuk menilai efektivitas metode yang digunakan serta mengidentifikasi keterbatasan penelitian.

RESULTS AND DISCUSSION

Tahap awal penelitian dimulai dengan pengumpulan data berupa tweet yang berkaitan dengan kebijakan pemblokiran rekening nganggur oleh PPATK. Pengambilan data dilakukan melalui proses scraping Twitter menggunakan kata kunci yang relevan, seperti PPATK, rekening nganggur, rekening dormant, dan pemblokiran rekening. Proses ini menghasilkan sejumlah tweet mentah

yang masih mengandung data duplikat, spam, serta konten yang tidak sepenuhnya relevan. Pada tahap ini, penulis dapat menyisipkan gambar berupa hasil scraping awal atau daftar tweet mentah sebagai ilustrasi proses pengumpulan data.

Setelah data terkumpul, dilakukan proses seleksi untuk memastikan bahwa hanya tweet yang relevan dengan topik penelitian yang digunakan. Tweet duplikat dihapus, begitu pula dengan tweet yang tidak mengandung opini atau tidak berkaitan langsung dengan kebijakan PPATK. Tahap ini bertujuan untuk meningkatkan kualitas dataset agar hasil analisis lebih akurat. Gambar yang dapat disisipkan pada bagian ini berupa perbandingan jumlah data sebelum dan sesudah seleksi. Hasil proses scraping data tersebut bisa di lihat di gambar 1 di bawah ini



```

SCRAPING DATA
Add blockquote

import pandas as pd

data = pd.read_csv("/content/dataset.csv")
data.info()

<<Class 'pandas.core.frame.DataFrame'>
RangeIndex: 692 entries, 0 to 691
Data columns (total 1 columns):
 #   Column  Non-Null Count  Dtype
---  ---
 0   Tweet   692 non-null    object
dtypes: object(1)
memory usage: 5.5+ KB

```

Gambar 1 hasil sraping data

Data tweet yang telah diseleksi kemudian melalui tahap preprocessing. Proses ini meliputi case folding untuk mengubah seluruh teks menjadi huruf kecil, tokenisasi untuk memecah kalimat menjadi kata kata, serta stopword removal untuk menghilangkan kata kata umum yang tidak memiliki makna signifikan. Preprocessing bertujuan untuk mengurangi noise pada data teks dan menyiapkan data agar siap diproses pada tahap berikutnya. Pada bagian ini, penulis dapat menyisipkan gambar yang menunjukkan contoh tweet sebelum dan sesudah preprocessing.

Tahap preprocessing data bertujuan untuk membersihkan dan menyiapkan teks agar dapat diproses oleh algoritma Naïve Bayes secara optimal.

Setelah preprocessing, data teks direpresentasikan dalam bentuk numerik menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF). Metode ini memberikan bobot pada setiap kata berdasarkan frekuensi kemunculannya dalam tweet dan keseluruhan dataset. Kata-kata yang sering muncul tetapi bersifat umum akan memiliki bobot lebih kecil, sedangkan kata-kata yang lebih spesifik akan memiliki bobot lebih besar. Tahap ini sangat penting karena hasil pembobotan TF-IDF menjadi dasar bagi proses klasifikasi sentimen. Gambar yang dapat disisipkan pada bagian ini berupa tampilan matriks TF-IDF atau hasil proses ekstraksi fitur.

Hasil proses preprocessing data tersebut bisa di lihat di gambar 2 di bawah ini

```

PREPROCESSING DATA

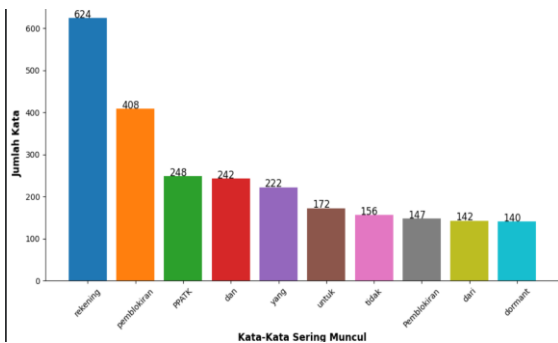
import pandas as pd

data = pd.read_csv("dataset.csv")
data.info()

<<class 'pandas.core.frame.DataFrame'>
RangeIndex: 692 entries, 0 to 691
Data columns (total 1 columns):
 #   Column  Non-Null count  Dtype
---  ---
 0   Tweet    692 non-null     object
dtypes: object(1)
memory usage: 5.5+ KB
    
```

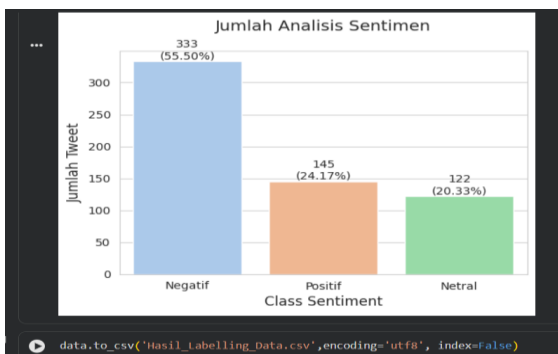
Gambar 2 hasil preprocessing data

Tahap selanjutnya adalah wordcloud sebelum preprocessing Pada tahap ini dilakukan visualisasi awal data dalam bentuk wordcloud sebelum proses preprocessing diterapkan. Wordcloud digunakan untuk melihat kata kata yang paling sering muncul dalam tweet mentah terkait isu pemblokiran rekening oleh PPKATK. Hasil proses wordcloud sebelum preprocessing tersebut bisa di lihat di gambar 3 di bawah ini



Gambar 3 hasil wordcloud sebelum preprocessing

Pelabelan data dengan metode lexicon based dilakukan untuk memberikan label sentimen secara otomatis pada setiap tweet sebelum proses pelatihan model Naïve Bayes. Metode ini menggunakan daftar kata positif dan negatif yang telah disusun dalam kamus sentimen (lexicon) Bahasa Indonesia. Hasil proses pelabelan data dengan metode lexicon based tersebut bisa di lihat di gambar 4 di bawah ini

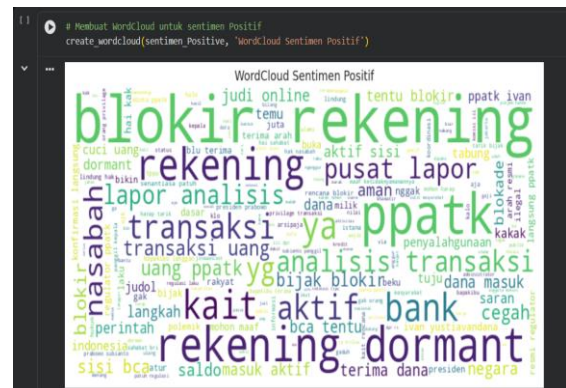


Gambar 4 hasil pelabelan data dengan metode lexicon based

Wordcloud sentimen dibuat untuk memvisualisasikan kata-kata yang paling sering muncul dalam setiap kategori sentimen positif, negatif, dan netral setelah proses pelabelan lexicon-based. Visualisasi ini membantu memahami pola bahasa yang digunakan masyarakat dalam merespons isu pemblokiran rekening oleh PPKATK. Hasil wordcloud tersebut bisa di lihat di gambar 5,6,dan 7 di bawah ini



Gambar 5 hasil wordcloud negatif



Gambar 6 hasil wordcloud positif



Gambar 7 wordcloud netral

Tahap selanjutnya adalah klasifikasi sentimen menggunakan algoritma Naive Bayes.

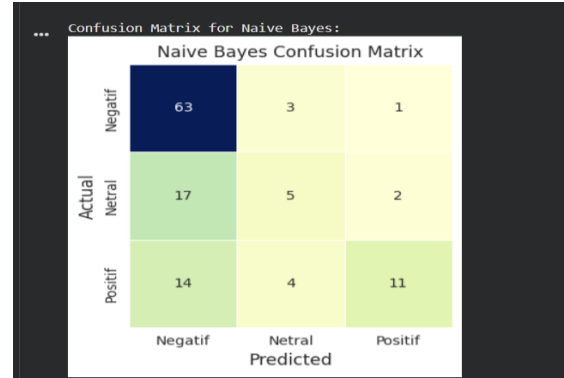
Dataset dibagi menjadi data latih dan data uji untuk membangun dan menguji model klasifikasi. Algoritma Naive Bayes menghitung probabilitas setiap tweet untuk termasuk ke dalam kelas sentimen positif, negatif, atau netral berdasarkan kata-kata yang terkandung di dalamnya. Hasil klasifikasi menunjukkan bahwa sebagian besar tweet cenderung dikategorikan sebagai sentimen negatif, sementara sentimen positif dan netral muncul dalam proporsi yang lebih kecil. Pada bagian ini, penulis dapat menyisipkan gambar hasil klasifikasi atau tampilan proses model pada perangkat lunak yang digunakan.

Evaluasi kinerja model dilakukan untuk mengetahui seberapa baik algoritma Naive Bayes dalam mengklasifikasikan sentimen. Metode evaluasi yang digunakan meliputi confusion matrix, accuracy, precision, recall, dan F1 score. Hasil evaluasi menunjukkan bahwa model mampu mengenali sebagian besar data pada kelas tertentu, namun masih memiliki keterbatasan dalam membedakan sentimen yang jumlah datanya relatif sedikit. Gambar confusion matrix atau tabel hasil evaluasi dapat disisipkan pada bagian ini untuk memperjelas hasil pengujian.

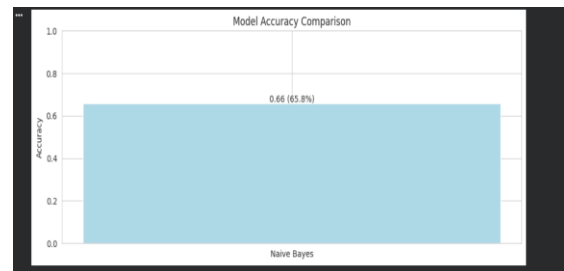
Berdasarkan hasil klasifikasi, sentimen negatif mendominasi opini publik di Twitter terkait kebijakan pemblokiran rekening nganggur oleh PPATK. Hal ini menunjukkan bahwa sebagian besar pengguna Twitter lebih banyak membagikan informasi atau berita tanpa menyertakan opini emosional yang kuat. Sentimen positif dan netral muncul dalam jumlah yang lebih kecil, yang mengindikasikan adanya dukungan dan kritik, meskipun tidak mendominasi percakapan publik.

Dominasi sentimen negatif juga dapat dipengaruhi oleh karakteristik data Twitter, di mana banyak tweet berupa retweet atau kutipan berita. Selain itu, keterbatasan jumlah data dan ketidakseimbangan kelas sentimen turut memengaruhi hasil klasifikasi. Temuan ini sejalan dengan penelitian sebelumnya yang menyatakan bahwa analisis sentimen terhadap isu kebijakan publik cenderung menghasilkan proporsi sentimen negatif yang tinggi.

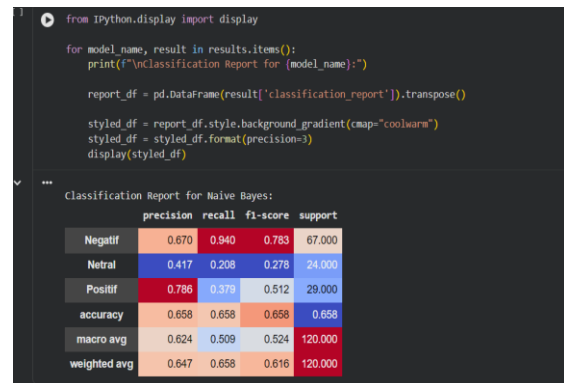
Secara keseluruhan, hasil penelitian ini menunjukkan bahwa metode text mining dengan algoritma Naive Bayes dapat digunakan untuk menganalisis opini publik di media sosial, meskipun masih memiliki keterbatasan yang dapat dikembangkan lebih lanjut pada penelitian selanjutnya. Hasil klasifikasi sentimen menggunakan algoritma Naive Bayes tersebut bisa dilihat di gambar 8, 9, dan 10 di bawah ini



Gambar 8 naïve bayes confusion matrix



Gambar 9 naïve bayes model accuracy comparison



Gambar 10 naïve bayes classification report

CONCLUSION

Berdasarkan hasil penelitian dan pembahasan yang telah diuraikan pada Bab IV, dapat disimpulkan bahwa analisis sentimen opini publik terhadap kebijakan pemblokiran rekening nganggur oleh PPATK pada media sosial Twitter berhasil dilakukan menggunakan pendekatan text mining dengan algoritma Naive Bayes. Penelitian ini mampu mengolah data tweet berbahasa Indonesia secara sistematis melalui tahapan pengumpulan data, preprocessing teks, ekstraksi fitur

menggunakan TF IDF, proses klasifikasi, serta evaluasi kinerja model.

Hasil klasifikasi sentimen menunjukkan bahwa opini publik di Twitter terkait kebijakan pemblokiran rekening nganggur oleh PPATK didominasi oleh sentimen negatif dengan persentase 55,50%, sementara sentimen positif 24,17% dan negatif 20,33%. Dominasi sentimen negatif mengindikasikan bahwa sebagian besar pengguna Twitter lebih banyak menyampaikan informasi atau membagikan berita tanpa mengekspresikan sikap emosional yang kuat terhadap kebijakan tersebut. Temuan ini menunjukkan bahwa diskursus publik di Twitter cenderung bersifat informatif dibandingkan persuasif atau konfrontatif.

Dari sisi kinerja model, algoritma Naïve Bayes menghasilkan nilai akurasi sebesar 67.000%, dengan precision rata rata tertimbang sebesar 0,670%, recall sebesar 0,940%, dan F1 score sebesar 0,783%. Nilai recall yang relatif tinggi menunjukkan bahwa model cukup mampu mengenali data pada kelas sentimen yang ada, khususnya kelas sentimen negatif yang mendominasi dataset. Namun, rendahnya nilai akurasi dan precision menunjukkan bahwa model masih memiliki keterbatasan dalam membedakan kelas sentimen minoritas, yaitu sentimen positif dan netral.

Keterbatasan kinerja tersebut dipengaruhi oleh ketidakseimbangan distribusi kelas pada dataset, dimana jumlah tweet dengan sentimen negatif jauh lebih besar dibandingkan dengan sentimen positif dan netral. Kondisi ini menyebabkan algoritma Naive Bayes cenderung memprediksi data ke dalam kelas mayoritas untuk memaksimalkan probabilitas, sehingga berdampak pada rendahnya performa klasifikasi secara keseluruhan. Meskipun demikian, hasil penelitian ini menunjukkan bahwa Naïve Bayes tetap dapat digunakan sebagai metode dasar (baseline) dalam analisis sentimen opini publik pada media sosial.

Secara keseluruhan, penelitian ini membuktikan bahwa penerapan text mining dengan algoritma Naïve Bayes dapat memberikan gambaran umum mengenai kecenderungan sentimen masyarakat terhadap kebijakan publik, khususnya kebijakan pemblokiran rekening

nganggur oleh PPATK. Hasil penelitian ini diharapkan dapat menjadi referensi awal bagi penelitian selanjutnya serta menjadi masukan bagi pemangku kebijakan dalam memahami respons masyarakat terhadap kebijakan yang diterapkan di ranah publik digital.

REFERENCE

- [1] C. C. Aggarwal and C. Zhai, *Mining Text Data*. Springer, 2018.
- [2] B. Liu, *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge University Press, 2020.
- [3] R. Feldman and J. Sanger, *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge University Press, 2020.
- [4] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. Cambridge University Press, 2022.
- [5] F. Rahman, Y. Prasetyo, and A. Wibowo, "Sentiment analysis of public policy using Twitter data: A case study in Indonesia," *Procedia Comput. Sci.*, vol. 216, pp. 321–330, 2023.
- [6] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, 2020.
- [7] L. Zahrotun, "Text preprocessing in Indonesian sentiment analysis," *J. Inf. Syst. Eng. Bus. Intell.*, vol. 6, no. 1, pp. 1–10, 2020.
- [8] Z. Rustam and D. Talita, "Comparison of Naive Bayes and Support Vector Machine for sentiment analysis on social media," *J. Phys. Conf. Ser.*, vol. 1722, no. 1, 2021.
- [9] R. Nugraha and A. Setiawan, "Analisis sentimen opini publik di Twitter menggunakan algoritma Naive Bayes," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 10, no. 2, pp. 215–224, 2023.