

Analisis Sentimen Terhadap Kinerja Menteri Keuangan Purbaya Yudhi Sadewa Menggunakan Algoritma K-Nearest Neighbors

Renoka Tresna Pramuda^{*1}, Fida Maisa Hana², Agung Prihandono³

^{1,2,3}Ilmu Komputer, Fakultas Sains dan Teknologi, Universitas Muhammadiyah Kudus, Jl. Ganesha Raya No.1, Purwosari, Kota Kudus, Kudus, Jawa Tengah, (0291) 437218
e-mail: ^{*1}renokatresna29@gmail.com, ²fidamaisa@umkudus.ac.id, ³agungdono@umkudus.ac.id

Abstrak

Perkembangan media sosial meningkatkan partisipasi publik dalam menyampaikan opini terhadap kebijakan dan kinerja pejabat pemerintah. Penelitian ini bertujuan untuk menganalisis sentimen terhadap kinerja Purbaya Yudhi Sadewa sebagai Menteri Keuangan berdasarkan komentar TikTok. Data diperoleh melalui teknik scraping dari video TikTok dalam periode satu bulan setelah pelantikan dan menghasilkan 1317 komentar setelah proses pembersihan data. Tahapan pengolahan meliputi preprocessing, pelabelan data menggunakan metode manual dan lexicon-based, ekstraksi fitur menggunakan TF-IDF serta pemodelan menggunakan algoritma K-Nearest Neighbors. Penentuan parameter optimal dilakukan menggunakan GridSearchCV dengan 10-Fold Cross Validation, yang menghasilkan nilai K terbaik 13 untuk data dengan labeling manual dan 15 untuk data dengan labeling lexicon-based. Hasil pengujian menggunakan Holdout Validation (80:20) menunjukkan bahwa model KNN dengan labeling manual memperoleh akurasi tertinggi sebesar 94,32%, sedangkan model dengan lexicon-based memperoleh akurasi sebesar 75,38%. Hasil penelitian menunjukkan bahwa pelabelan manual menghasilkan performa klasifikasi yang lebih baik dan lebih konsisten dibandingkan metode lexicon-based. Temuan ini mengindikasikan bahwa kualitas pelabelan data berpengaruh signifikan terhadap performa klasifikasi sentimen, terutama pada komentar media sosial yang bersifat informal dan ambigu.

Kata kunci— Analisis Sentimen, K-Nearest Neighbors, TikTok, Menteri Keuangan

1. PENDAHULUAN

Seiring berkembangnya teknologi informasi dan komunikasi, Media sosial kini telah menjadi menjadi sarana utama bagi masyarakat untuk mengungkapkan pendapat, berbagi pengalaman, serta menyampaikan opini terhadap berbagai permasalahan, termasuk kebijakan dan kinerja para pejabat pemerintah [1]. Salah satu media sosial yang populer saat ini adalah TikTok. Menurut data dari The Global Statistics, pada tahun 2025 TikTok mencatat sebanyak 108 juta pengguna berusia 18 tahun ke atas di Indonesia, menjadikan Indonesia sebagai salah satu pasar terbesar bagi TikTok di tingkat global. TikTok adalah platform media sosial yang memungkinkan penggunaannya dapat membuat dan membagikan video pendek, dilengkapi dengan berbagai fitur seperti musik, filter, serta berbagai fitur lainnya [2]. Selain berfungsi sebagai media hiburan dan konten kreatif, TikTok juga dimanfaatkan sebagai sarana untuk menyampaikan pendapat, termasuk kritik terhadap berbagai isu sosial dan politik [3].

Menteri Keuangan Republik Indonesia (Menkeu) merupakan pimpinan Kementerian Keuangan Republik Indonesia. Saat ini, jabatan Menteri Keuangan Republik Indonesia dipegang oleh Purbaya Yudhi Sadewa sejak 8 September 2025, menggantikan Sri Mulyani [4]. Sebagai Menteri Keuangan, Purbaya telah terlibat dalam sejumlah kebijakan dan respons terhadap kondisi ekonomi nasional. Munculnya figur baru dalam pemerintahan Indonesia memunculkan beragam respon publik sekaligus

harapan terhadap arah kebijakan yang akan dijalankan, terutama di media sosial seperti TikTok. Oleh karena itu, diperlukan analisis sentimen untuk memahami kecenderungan opini publik terhadap kinerja dan kebijakan yang diambil Purbaya sebagai Menteri Keuangan.

Analisis sentimen merupakan salah satu teknik dalam bidang Natural Language Processing (NLP) yang digunakan untuk mengidentifikasi komentar pengguna terhadap suatu produk atau layanan, termasuk di antaranya aplikasi [5]. Tujuan dari analisis sentimen adalah untuk mengenali opini seseorang yang bersifat positif, negatif atau netral yang kemudian dapat digunakan untuk pengambilan Keputusan [6]. Analisis sentimen tidak hanya memberikan gambaran umum mengenai sikap publik terhadap suatu isu, tetapi juga dapat dimanfaatkan untuk mendeteksi kata yang sering digunakan serta mengungkap pola persepsi masyarakat secara menyeluruh [7].

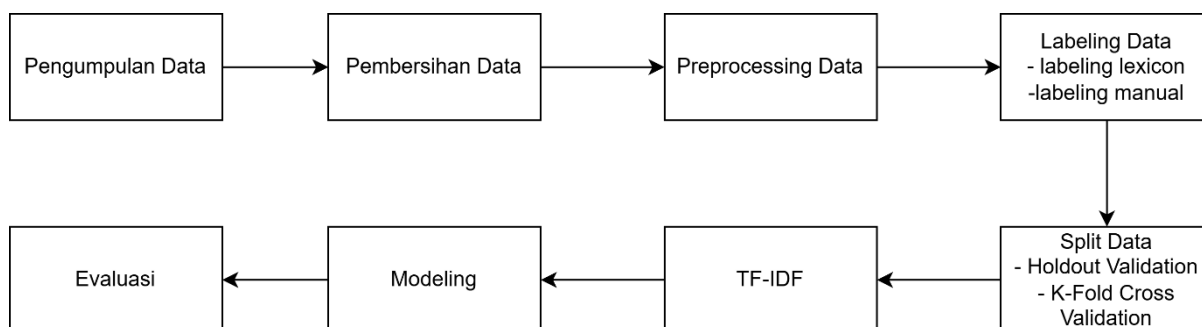
Metode yang digunakan pada penelitian ini adalah K-Nearest Neighbors (KNN). Algoritma ini digunakan untuk melakukan proses klasifikasi terhadap suatu objek berdasarkan data pelatihan dengan mempertimbangkan jarak terdekat antar data [8]. Dalam penerapannya, algoritma KNN memiliki sejumlah keunggulan, antara lain kesederhanaan dan mudah untuk dipahami, sifatnya yang non-parametrik, fleksibilitas dalam penyesuaian model, serta ketahanannya terhadap noise pada data [9]. Algoritma ini sederhana tetapi efektif dalam mengatasi masalah klasifikasi berupa teks termasuk pada kasus analisis komentar tiktok [10].

Penelitian sebelumnya menunjukkan bahwa algoritma KNN telah berhasil dan mampu menghasilkan akurasi yang lebih tinggi daripada algoritma *Naive Bayes*. Dengan jumlah 800 data algoritma KNN menghasilkan akurasi 88,12%, Precision sebesar 93,98% dan nilai Recall 81,53%, sehingga metode KNN mengungguli metode *Naive Bayes* dengan akurasi 82,27%, precision 86,36% dan recall 76,93% [11]. Dalam penelitian lain, Komparasi Metode SVM, KNN dan NBC Pada Analisis Sentimen memberikan hasil bahwa metode KNN lebih unggul dibandingkan SVM dan NBC dalam hal akurasi pengujian dengan akurasi 93,4%, sedangkan untuk SVM 93,1% dan NBC 87,9% [12]. Sementara itu, penelitian lain membandingkan antara algoritma *Naive Bayes*, KNN dan *Decision Tree* pada analisis sentimen ulasan aplikasi KitaLulus dengan 600 data ulasan dari Twitter. Setelah dilakukan SMOTE pada tahap preprocessing, menunjukkan bahwa algoritma KNN merupakan algoritma terbaik dengan akurasi 83,33%, precision 80,36% dan recall 71,09% [13].

Diharapkan penelitian ini dapat menjadi acuan bagi Menteri Keuangan dalam menanggapi berbagai respon masyarakat terhadap kebijakan yang diambil. Melalui temuan ini, evaluasi terhadap kebijakan yang memicu perhatian maupun kritik publik dapat diperbaiki. Selain itu, penelitian ini diharapkan mampu menambah referensi dalam kajian analisis sentimen melalui penerapan algoritma K-Nearest Neighbor (KNN).

2. METODE PENELITIAN

Penelitian ini melakukan analisis sentimen terhadap kinerja Menteri Keuangan Purbaya Yudhi Sadewa menggunakan algoritma KNN. Tahapan penelitian meliputi pengumpulan data, pembersihan data, preprocessing data, pelabelan data, split data, ekstraksi fitur menggunakan TF-IDF, modeling menggunakan KNN serta evaluasi dengan confusion matrix dan classification report.



Gambar 1. Diagram Tahapan Penelitian

Berdasarkan Gambar 1, tahapan penelitian dimulai dari pengumpulan data secara langsung menggunakan API TikTok dengan teknik scraping menggunakan tools *TikTok Comments Scraper*. Data yang diperoleh berupa komentar dari video TikTok dengan hasil pencarian menggunakan kata kunci “purbaya yudhi sadewa”, “kebijakan purbaya”, “kinerja purbaya”. Pengumpulan data dilakukan dalam periode sebulan setelah pelantikan Purbaya sebagai Menteri Keuangan. Sehingga diperoleh sebanyak 1319 data dan disimpan dalam format csv yang kemudian akan diproses lebih lanjut.

Tahapan berikutnya pembersihan data untuk mengatasi data kosong atau tidak terisi. Selain itu juga dilakukan pemeriksaan terhadap data duplikat untuk menghindari komentar yang muncul lebih dari sekali. Setelah data dibersihkan, kemudian masuk ke tahap preprocessing data. Menurut [14] tahapan preprocessing meliputi:

- a. Case Folding
Mengubah seluruh huruf menjadi huruf kecil (*lowercase*) untuk menjaga konsistensi dan mencegah perbedaan makna akibat variasi kapitalisasi.
- b. Tokenizing
Memecah kalimat menjadi per kata atau token agar proses analisis menjadi lebih mudah dan terstruktur.
- c. Normalisasi
Menyatukan berbagai bentuk kata yang memiliki makna sama meskipun ejaannya berbeda.
- d. Stopword Removal
Menghapus kata-kata umum seperti “yang”, “dan”, atau “di” yang dianggap tidak memberikan kontribusi signifikan terhadap makna kalimat.
- e. Stemming
Mengembalikan setiap kata ke bentuk dasarnya menggunakan algoritma Sastrawi, sehingga analisis menjadi lebih efisien dan konsisten secara morfologis.

Proses pelabelan data menggunakan 2 metode yaitu labeling manual dan *labeling lexicon-based*. Labeling manual dilakukan secara langsung oleh ahli yang mengklasifikasikan setiap komentar ke dalam kategori positif, negatif dan netral. Pelabelan manual menghasilkan distribusi sentimen yang lebih seimbang dan lebih mencerminkan kondisi data [15]. *Lexicon-based* yaitu metode yang memberi label berdasarkan kamus kata yang sudah diberi nilai sentimen positif, negatif dan netral [16]. Setiap kata dicocokkan dengan kamus sentimen, lalu skornya dijumlahkan atau dirata-ratakan untuk setiap komentar [17]. Langkah selanjutnya membagi data menjadi dua bagian yaitu training dan testing menggunakan metode holdout validation dan k-fold cross validation. Setelah itu ekstraksi fitur menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). TF-IDF merupakan teknik pembobotan kata yang dilakukan dengan menghitung frekuensi kemunculan suatu kata dalam sebuah dokumen serta memperhatikan seberapa sering istilah tersebut muncul di seluruh kumpulan dokumen [18]. Hasil pembobotan TF-IDF tersebut selanjutnya membentuk vektor numerik yang merepresentasikan masing-masing komentar.

Dalam penelitian ini dilakukan pemodelan menggunakan algoritma KNN. Pada proses pemodelan, ditentukan nilai K sebagai jumlah tetangga terdekat yang digunakan dalam pengambilan keputusan klasifikasi. Penentuan nilai K terbaik pada K-Nearest Neighbors dilakukan menggunakan GridSearchCV dengan 10-Fold Cross Validation. Setiap kandidat nilai K diuji dengan membagi data training menjadi 10 fold, kemudian dilatih di 9 fold dan diuji pada 1 fold. Akurasi tiap fold dihitung, lalu dirata-ratakan untuk mendapatkan performa setiap K . Nilai K dengan rata-rata akurasi tertinggi dipilih sebagai parameter terbaik untuk model KNN.

Tahapan terakhir evaluasi, evaluasi merupakan proses yang bertujuan untuk menguji performa dari model [19]. *Confusion Matrix* digunakan untuk mengevaluasi kinerja dari algoritma KNN yang sudah diterapkan dalam proses analisis sentimen. Berdasarkan *Confusion Matrix* tersebut, dihitung beberapa metrik evaluasi, yaitu *akurasi*, *precision*, *recall*, dan *F1-score*.

3. HASIL DAN PEMBAHASAN

Data penelitian diperoleh sejumlah 1319 baris dan 17 kolom. Setelah pembersihan, termasuk penghapusan data kosong, tersisa 1.317 data. Dalam penelitian ini hanya kolom 'text' yang digunakan karena kolom tersebut berisi komentar dari pengguna. Sampel data penelitian dapat dilihat di Tabel 1.

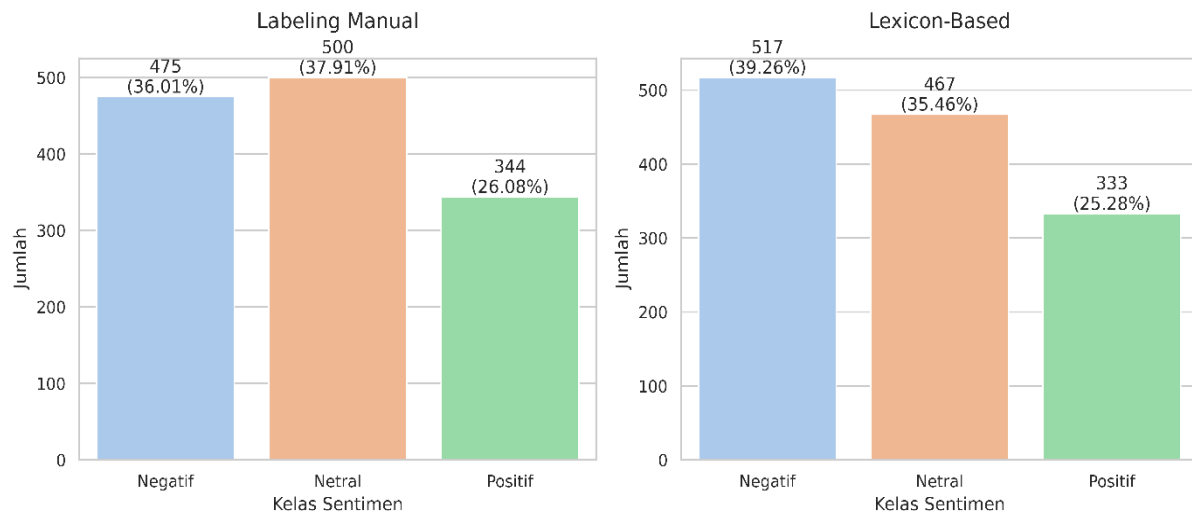
Tabel 1. Sampel Data

No	Text
1	????? dompet rakyat jelata masih kosong bos.. ihsg itu buat yg berduit saya gak ngurus.. pikirlah gimana penjualan UMKM bisa naik kencang
2	sabar atuh ini bapaknya cuma satu, baru kerja sebulan, rakyat Indonesia ada ratusan juta
3	keren menkeu kita ini..
...	...
1317	bri minta lagi tuhh pakk 😊😊

Tabel 2 menyajikan hasil dari preprocessing data. Hasil akhir menunjukkan bahwa data menjadi lebih terorganisir dan relevan untuk proses analisis sentimen lebih lanjut. Gambar 2 menampilkan perbandingan distribusi kelas sentimen yang dihasilkan dari labeling manual dan metode lexicon-based. Visualisasi ini bertujuan untuk menunjukkan perbedaan karakteristik hasil pelabelan dari kedua metode yang digunakan dalam penelitian. Hasil labeling manual menunjukkan sebanyak 475 komentar (36,01%) dikategorikan negatif, 500 komentar (37,91%) netral dan 344 komentar (26,08%) sebagai positif. Lexicon-based menunjukkan hasil yang berbeda, sebanyak 517 komentar (39,26%) dikategorikan sebagai negatif, 467 komentar (35,46%) dikategorikan sebagai netral dan 333 komentar (25,28%) sebagai positif. Labeling manual cenderung lebih mengidentifikasi komentar yang bersifat netral, terutama pada teks yang maknanya ambigu. Metode lexicon-based cenderung lebih menandai komentar sebagai negatif daripada positif, dengan distribusi yang merata antara negatif dan netral. Perbedaan distribusi ini menunjukkan bahwa kedua metode memiliki karakteristik pelabelan yang berbeda dalam menginterpretasikan sentimen.

Tabel 2. Hasil Preprocessing

Tahapan	Hasil Preprocessing
text	????? dompet rakyat jelata masih kosong bos.. ihsg itu buat yg berduit saya gak ngurus.. pikirlah gimana penjualan UMKM bisa naik kencang
Case Folding/Cleaning	dompet rakyat jelata masih kosong bos ihsg itu buat yg berduit saya gak ngurus pikirlah gimana penjualan umkm bisa naik kencang
Tokenizing	[dompet, rakyat, jelata, masih, kosong, bos, ihsg, itu, buat, yg, berduit, saya, gak, ngurus, pikirlah, gimana, penjualan, umkm, bisa, naik, kencang]
Normalisasi	dompet, rakyat, jelata, masih, kosong, bis, ihsg, itu, buat, yang, berduit, saya, tidak, mengurus, pikirlah, bagaimana, penjualan, umkm, bisa, naik, kencang
Stopword	dompet, rakyat, jelata, kosong, ihsg, berduit, tidak, mengurus, pikirlah, penjualan, umkm, naik, kencang
Stemming	dompet, rakyat, jelata, kosong, ihsg, duit, tidak, urus, pikir, jual, umkm, naik, kencang



Gambar 2. Hasil Perbandingan Labeling Data

Tabel 3 dan Tabel 4 menunjukkan hasil Split data yang dilakukan pada tiap hasil perbandingan labeling data. Pada labeling manual metode HoldOut Validation membagi data dengan beberapa perbandingan, mulai 50:50 sampai 90:10. Hasil pengujian menunjukkan meskipun jumlah data latih yang lebih besar dapat membantu model mengenali pola dengan lebih baik, porsi data uji yang terlalu kecil dapat membuat proses evaluasi kurang stabil dan meningkatkan risiko overfitting. Sedangkan hasil pengujian K-Fold Cross Validation menunjukkan performa yang lebih konsisten karena seluruh data digunakan secara bergantian sebagai data latih dan data uji.

Tabel 3. Hasil Split Data Berdasarkan Labeling Manual

Perbandingan (%)	Holdout Validation		K-Fold Cross Validation	
	KNN (%)	K-Fold	K-Fold	KNN(%)
50:50	67,98%	1		95,45%
55:45	70,49%	2		87,87%
60:40	74,57%	3		65,15 %
65:35	78,31%	4		63,63 %
70:30	81,06%	5		62,12%
75:25	86,06%	6		68,18%
80:20	94,32%	7		65,90%
83:17	92,86%	8		61,83%
85:15	93,94%	9		64,12%
90:10	93,94%	10		59,54%

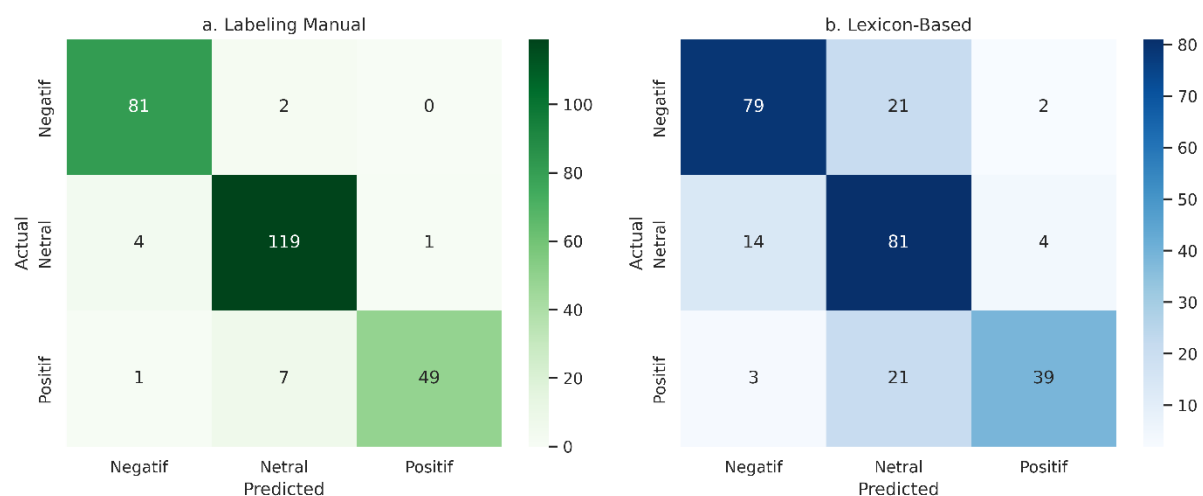
Sementara itu, metode Holdout Validation pada labeling lexicon menunjukkan bahwa meskipun penambahan data latih dapat meningkatkan performa model, hasil klasifikasi berbasis lexicon belum mampu representasi pola sentimen secara optimal dibandingkan labeling manual. Sedangkan hasil K-Fold Cross Validation pada labeling lexicon juga menunjukkan kestabilan yang cukup baik, namun tingkat akurasinya masih sedikit lebih rendah dibandingkan labeling manual.

Tabel 4. Hasil Split Data Berdasarkan Labeling Lexicon

Perbandingan (%)	Holdout Validation		K-Fold Cross Validation	
	KNN (%)	K-Fold	K-Fold	KNN(%)
50:50	63,28%	1	1	74,24%
55:45	64,42%	2	2	77,27%
60:40	64,71%	3	3	63,63%
65:35	66,16%	4	4	63,63%
70:30	69,19%	5	5	64,39%
75:25	72,12%	6	6	68,93%
80:20	75,38%	7	7	61,36%
83:17	72,77%	8	8	59,54%
85:15	74,24%	9	9	63,35%
90:10	74,24%	10	10	63,35%

Gambar 3 menunjukkan confusion matrix model KNN berdasarkan labeling manual dan metode lexicon-based. Pada data labeling manual kelas negatif, sebanyak 81 data berhasil diklasifikasikan dengan benar sebagai negatif, 2 data diklasifikasikan netral dan tidak ada yang salah ke positif. Pada kelas netral, 119 data berhasil diprediksi benar, 4 data diprediksi netral dan 1 data yang salah ke positif. Sementara pada kelas positif, sebanyak 49 data berhasil diprediksi benar, namun 7 data diprediksi netral dan 1 sebagai negatif. Prediksi model pada data labeling manual lebih terkonsentrasi pada diagonal utama dengan jumlah kesalahan klasifikasi yang relatif rendah pada seluruh kelas sentimen. Hal ini menunjukkan bahwa label yang diberikan secara manual memiliki konsistensi yang lebih baik sehingga pola sentimen yang dipelajari model menjadi lebih jelas.

Sedangkan pada metode lexicon-based, pada kelas negatif, terdapat 79 data yang berhasil diprediksi dengan benar. Sementara itu, terdapat 21 data diprediksi netral dan 2 data diprediksi positif. Pada kelas netral, model berhasil memprediksi 81 data dengan benar. Namun, terdapat 14 data yang diprediksi negatif dan 4 data yang diprediksi positif. Pada kelas positif model berhasil memprediksi dengan benar sebanyak 39 data namun terdapat 21 data yang salah diprediksi sebagai netral dan 3 data yang diprediksi sebagai negatif. Dari hasil tersebut masih terlihat kesalahan klasifikasi yang lebih sering terjadi, terutama pada perpindahan prediksi antara kelas negatif, netral dan positif. Hal ini menunjukkan bahwa metode lexicon-based cenderung mengalami kesulitan dalam menangkap konteks kalimat atau penggunaan kata yang memiliki makna berbeda pada konteks komentar.



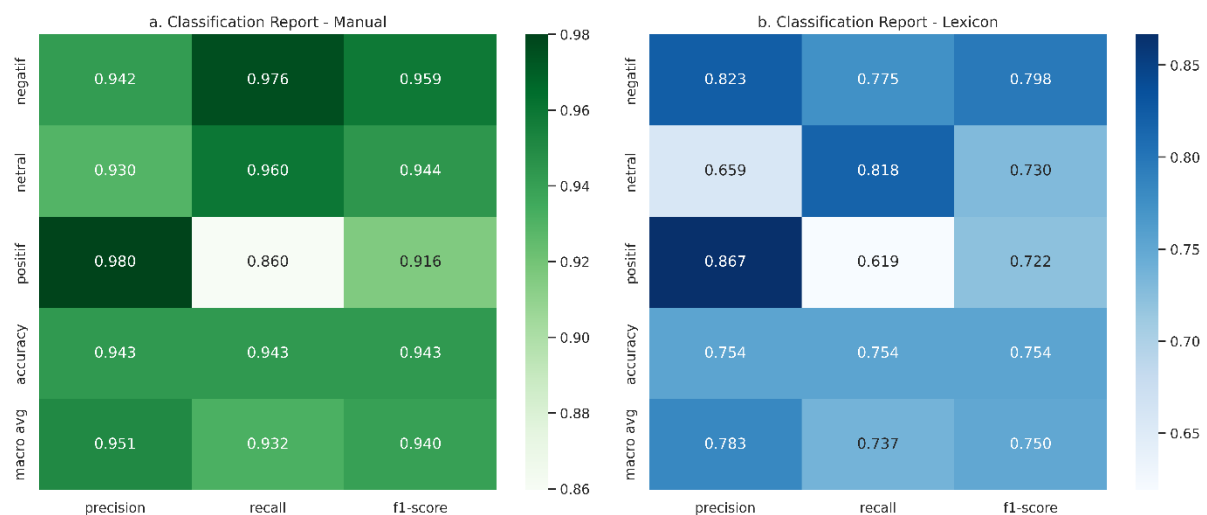
Gambar 3. Confusion Matrix: a. Labeling Manual dan b. Labeling Lexicon-Based

Gambar 4 menunjukkan perbandingan classification report model KNN pada data yang dilabeli secara manual dan menggunakan metode lexicon-based. Model yang dilatih menggunakan hasil

labeling manual menunjukkan performa yang lebih baik pada seluruh metrik evaluasi dibandingkan dengan model yang menggunakan metode lexicon-based. Hal ini terlihat dari nilai precision, recall dan F1-score yang relatif tinggi dan konsisten pada seluruh kelas sentimen, serta akurasi keseluruhan mencapai 94%.

Pada labeling manual, nilai F1-score yang tinggi pada ketiga kelas menunjukkan bahwa model mampu menjaga keseimbangan antara precision dan recall. Hasil ini mengindikasikan bahwa label yang diberikan secara manual mampu merepresentasikan sentimen secara lebih akurat sehingga pola yang dipelajari model menjadi lebih jelas dan konsisten. Selain itu, perbedaan nilai antar kelas relatif kecil, yang menunjukkan bahwa model tidak mengalami kesulitan dalam membedakan sentimen.

Sedangkan pada metode lexicon-based diperoleh akurasi 75% dengan nilai F1-score yang lebih rendah pada seluruh kelas. Penurunan performa paling terlihat pada kelas netral dan positif, yang menunjukkan bahwa model masih mengalami kesulitan dalam membedakan kedua kategori tersebut. Hal ini mengindikasikan bahwa pelabelan lexicon-based belum sepenuhnya mampu menangkap konteks dan makna komentar secara tepat, sehingga menghasilkan label yang kurang konsisten dibandingkan labeling manual.



Gambar 4. Classification Report: a. Labeling Manual dan b. Labeling Lexicon-Based

4. KESIMPULAN

Berdasarkan hasil penelitian, dapat disimpulkan bahwa analisis sentimen terhadap kinerja Menteri Keuangan Purbaya Yudhi Sadewa menggunakan Algoritma K-Nearest Neighbors pada platform TikTok menghasilkan performa model yang cukup baik. Hasil penelitian menunjukkan bahwa labeling manual menghasilkan performa yang lebih baik dibandingkan metode lexicon-based. Pada tahap penentuan parameter menggunakan K-Fold Cross Validation, diperoleh nilai K terbaik sebesar 13 untuk data dengan labeling manual dan 15 untuk data dengan labeling lexicon-based. Nilai K tersebut kemudian digunakan pada proses Holdout Validation untuk membangun model klasifikasi. Hasil pengujian menunjukkan bahwa model KNN dengan labeling manual dan K=13 mampu mencapai akurasi tertinggi sebesar 96% pada skenario Holdout Validation dengan pembagian data 80:20. Sementara model yang menggunakan labeling lexicon-based dengan k=15 pada skenario yang sama hanya memperoleh akurasi sebesar 75%. Selain itu, hasil K-Fold Cross Validation menunjukkan bahwa labeling manual memiliki tingkat konsistensi yang lebih stabil meskipun nilai akurasi cenderung lebih rendah dibandingkan metode Holdout Validation.

Berdasarkan evaluasi confusion matrix dan classification report, model dengan labeling manual menunjukkan nilai precision, recall dan F1-score yang tinggi serta relatif seimbang pada seluruh kelas

senetimen. Sebaliknya, model dengan labeling lexicon-based menghasilkan nilai evaluasi yang lebih rendah, terutama pada kelas netral dan positif. Hal ini menunjukkan bahwa kualitas pelabelan data memiliki pengaruh yang signifikan terhadap performa model klasifikasi sentimen. Labeling manual mampu menghasilkan label yang lebih representatif terhadap konteks komentar, sedangkan pendekatan lexicon-based masih memiliki keterbatasan dalam memahami bahasa pada media sosial TikTok yang cenderung informal, kontekstual dan ambigu.

SARAN

Untuk penelitian selanjutnya disarankan untuk pengolahan stopwords dan normalisasi perlu dilakukan lebih ketat untuk menghasilkan data yang lebih akurat. Penambahan data juga direkomendasikan agar hasil analisis sentimen lebih luas dan mencerminkan opini masyarakat lebih menyeluruh.

DAFTAR PUSTAKA

- [1] A. Khaidar, "Analisis Sentimen Di Instagram Terhadap Menteri Keuangan Purbaya Yudhi Sadewa," *JITET (jurnal Inform. dan Tek. Elektro Terap.*, vol. 13, no. 3, 2025.
- [2] N. Apriliani, N. Suarna, and W. Prihartono, "Analisis Sentimen Review Penggunaan Tiktok Melalui Pendekatan Algoritma Naïve Bayes," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 6, pp. 3725–3731, 2024, doi: 10.36040/jati.v7i6.8299.
- [3] S. Firdaus, N. Naila, A. D. Pramesti, C. K. Sari, and D. F. Azzahra, "Tiktok Sebagai Media Sosial dalam Melakukan Kritik terhadap Pembangunan di Lampung," *Indones. J. Soc. Dev.*, vol. 6, pp. 1–13, 2024.
- [4] Kompas.com, "Profil Purbaya Yudhi Sadewa, Menteri Keuangan yang Baru Dilantik." Accessed: Jan. 17, 2026. [Online]. Available: <https://money.kompas.com/read/2025/09/08/160543726/profil-purbaya-yudhi-sadewa-menteri-keuangan-yang-baru-dilantik>
- [5] S. Rihastuti, A. Rosyidi, A. Surakarta, and R. Forest, "Progres Pembangunan Ikn Dengan Metode Random Forest," *J. Comput. Sci. Technol.*, vol. 5, no. 1, pp. 19–23, 2025.
- [6] A. D. A. Putra and S. Juanita, "Sentiment Analysis on User Reviews of the Bibit and Bareksa Application with the KNN Algorithm," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 8, no. 2, pp. 636–646, 2021.
- [7] V. Agustina and A. Herliana, "Analisis Sentimen Publik atas Kebijakan Efisiensi Anggaran 2025 dengan Text Mining dan Natural Language Processing," *J. MEDIA Inform. [JUMIN]*, vol. 6, no. 3, pp. 2182–2194, 2025.
- [8] A. Budianita, N. Iman, F. M. Hana, and C. B. Hakim, "Komparasi Algoritma K-Nearest Neighbor dan Naive Bayes pada Klasifikasi Tingkat Kualitas Udara Kota Tangerang Selatan," *J. Inform. dan Rekayasa Perangkat Lunak*, vol. 6, no. 1, pp. 320–327, 2024.
- [9] M. Saifurridho, M. Martanto, and U. Hayati, "Analisis Algoritma K-Nearest Neighbor terhadap Sentimen Pengguna Aplikasi Shopee," *J. Inform. Terpadu*, vol. 10, no. 1, pp. 21–26, 2024, doi: 10.54914/jit.v10i1.1054.
- [10] V. S. Saputra, A. Ridwan, T. G. Pratama, K. Kudus, and J. Tengah, "Rancang Bangun Sistem

- Rekomendasi Buku Berbasis Item-Based Collaborative Filtering Menggunakan Algoritma K-Nearest,” *JITET (Jurnal Inform. dan Tek. Elektro Ter.)*, vol. 13, no. 1, pp. 1540–1545, 2025.
- [11] S. D. Prasetyo, S. S. Hilabi, and F. Nurapriani, “Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN,” *J. KomtekInfo*, vol. 10, pp. 1–7, 2023, doi: 10.35134/komtekinfo.v10i1.330.
- [12] I. G. Hendrayana, D. G. H. Divayana, and M. W. A. Kesiman, “Komparasi Metode Svm, K-Nn Dan Nbc Pada Analisis Sentimen,” *J. Indones. Manaj. Inform. dan Komun.*, vol. 4, no. 1, pp. 191–198, 2023, doi: 10.35870/jimik.v4i1.157.
- [13] R. M. Harun and F. Fahmi, “Perbandingan Algoritma Naïve Bayes , KNN , dan Decision Tree Pada Analisis Sentimen Terhadap Ulasan Aplikasi KitaLulus,” *Build. Informatics, Technol. Sci.*, vol. 6, no. 3, 2024, doi: 10.47065/bits.v6i3.6367.
- [14] K. A. Tompunu and A. Wedhasmara, “Analisis Sentimen Masyarakat Pada Komentar Instagram Terhadap Program Pemerintah Kota Palembang Dalam Pencapaian SDG 6 Menggunakan Algoritma Naive Bayes,” *J. Teknol. dan Sist. Inf. Univrab*, vol. 11, no. 1, pp. 476–493, 2026.
- [15] A. S. Rizkia, Wufron, and F. F. Roji, “Sentiment Analysis of Coretax : A Comparison of Manual , Transformers- Based , and Lexicon-Based Data Labeling on IndoBERT Performance Analisis Sentimen Coretax : Perbandingan Pelabelan Data Manual ,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 5, no. July, pp. 1037–1048, 2025.
- [16] K. Miranda and R. R. Suryono, “Analisis Sentimen Pinjaman Online : Studi Komparatif Algoritma Naïve Bayes , Decision Tree , dan KNN,” *Edumatic J. Pendidik. Inform.*, vol. 9, no. 2, pp. 372–381, 2025, doi: 10.29408/edumatic.v9i2.30142.
- [17] S. A. Nugraha, “Penerapan Lexicon Based Untuk Analisis Sentimen Masyarakat Indonesia Terhadap Danantara,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 9, no. 3, pp. 4949–4957, 2025.
- [18] M. H. Mahendra, D. T. Murdiansyah, and K. M. Lhaksmana, “Analisis Sentimen Tweet COVID-19 Menggunakan Metode K-Nearest Neighbors dengan Ekstraksi Fitur TF-IDF dan CountVectorizer Dike : Jurnal Ilmu Multidisiplin,” *Dike J. Ilmu Multidisiplin*, vol. 1, pp. 37–43, 2023.
- [19] T. Wahyudi, Rudiman, and A. N. Verdhkha, “Klasifikasi Sentimen X-Twitter Perihal Pindahan Ibu Kota Indonesia Menggunakan Ekstraksi Fitur Tf-Idf Dan Metode Support Vector Machine (SVM),” *J. Keilmuan dan Apl. Bid. Tek. Inform.*, vol. 18, no. 2, pp. 185–199, 2024.