

PERBANDINGAN METODE NAÏVE BAYES DAN SUPPORT VECTOR MACHINE UNTUK ANALISIS SENTIMEN PADA ULASAN PENGGUNA APLIKASI LINKEDIN

¹⁾Gishella Septania Al-Husna, ²⁾Dian Asmarajati, ³⁾Iman Ahmad Ihsanuddin, ⁴⁾Rina Mahmudati

^{1,2,3,4)}Universitas Sains Al-Qur'an

¹⁾gishella.septania@gmail.com

INFO ARTIKEL	ABSTRAK
Riwayat Artikel : Diterima : 22 Mei 2024 Disetujui : 31 Mei 2024	<p>Dengan semakin sedikitnya informasi lowongan pekerjaan dalam bentuk cetak, teknologi informasi yang berkembang pesat membuat lowongan pekerjaan lebih mudah ditemukan secara digital melalui aplikasi, media sosial, atau website. Namun, lowongan pekerjaan dari sumber yang tidak jelas dapat menimbulkan penipuan. LinkedIn adalah salah satu aplikasi terpercaya untuk mencari lowongan pekerjaan. Penelitian ini membandingkan dua metode dalam analisis sentimen ulasan pengguna LinkedIn dari Google Play Store, yaitu Naïve Bayes dan Support Vector Machine (SVM). Data dikumpulkan melalui web scraping dan diproses dengan text pre-processing yang mencakup data cleaning, case folding, stopword removal, tokenizing, dan stemming. Hasil menunjukkan bahwa Naïve Bayes menghasilkan akurasi 88%, precision 88%, recall 85%, dan f1-score 86%, sementara SVM menghasilkan akurasi 90%, precision 89%, recall 88%, dan f1-score 88%. SVM terbukti lebih efektif dalam analisis sentimen dibandingkan Naïve Bayes.</p>
Kata Kunci : Lowongan Pekerjaan, LinkedIn, Analisis Sentimen, Naïve Bayes, Support Vector Machine.	

ARTICLE INFO	ABSTRACT
Article History : Received : May 22, 2024 Accepted : May 31, 2024	<p><i>With the decreasing availability of job vacancy information in print media, the rapid advancement of information technology has made job vacancies easier to find digitally through applications, social media, or websites. However, job vacancies from unclear sources can lead to scams. LinkedIn is one of the trusted applications for job hunting. This study compares two methods for sentiment analysis of LinkedIn user reviews from the Google Play Store, namely Naïve Bayes and Support Vector Machine (SVM). Data were collected via web scraping and processed with text pre-processing, including data cleaning, case folding, stopword removal, tokenizing, and stemming. Results show that Naïve Bayes achieved an accuracy of 88%, precision of 88%, recall of 85%, and f1-score of 86%, while SVM achieved an accuracy of 90%, precision of 89%, recall of 88%, and f1-score of 88%. SVM proved to be more effective in sentiment analysis compared to Naïve Bayes.</i></p>
Keywords: Job Vacancies, LinkedIn, Sentiment Analysis, Naïve Bayes, Support Vector Machine	

1. PENDAHULUAN

Jumlah pengangguran di Indonesia terbilang masih cukup besar, berdasarkan data Badan Pusat Statistik (BPS) mencapai 8,42 juta orang pada Agustus 2022. Dengan demikian, tingkat pengangguran terbuka (TPT) di Indonesia sebesar 5,86% (Sadya, 2022). Salah satu penyebabnya dikarenakan masih banyak pencari kerja yang kesulitan mendapatkan pekerjaan. Jumlah pencari kerja di Indonesia pada tahun 2022, BPS mencatat ada 937.176 orang. Jumlah tersebut mengalami penurunan 65,76% dibandingkan pada tahun sebelumnya yang sebanyak 2,74 juta orang (Simbolon, 2023). Hal itu terjadi karena ketidakseimbangan antara lapangan pekerjaan dengan jumlah tenaga kerja, serta kemampuan dan keterampilan para pencari kerja yang tidak memadai atau tidak sesuai dengan kriteria suatu perusahaan. Terutama para fresh graduate yang belum memiliki cukup pengalaman dan keterampilan yang optimal untuk mendapatkan pekerjaan. Terkadang para pencari kerja juga tidak cukup aktif untuk mencari informasi mengenai lowongan pekerjaan yang tersedia.

Informasi lowongan pekerjaan dalam bentuk media cetak sekarang ini pun sudah mulai sedikit. Kecanggihan teknologi informasi dunia yang berkembang semakin pesat, menjadikan media informasi banyak ditemukan dalam bentuk digital. Begitu pula informasi lowongan pekerjaan dengan mudah didapatkan pada aplikasi pekerjaan, media sosial, ataupun website yang dapat diakses dengan adanya internet. Namun lowongan pekerjaan yang sumbernya tidak jelas, belum tentu terjamin keasliannya, karena kejahatan dapat menyertainya seperti adanya penipuan dari lowongan pekerjaan palsu. Terkadang masih ada orang yang sengaja membuat lowongan pekerjaan palsu dengan tujuan penipuan, untuk mengambil data pribadi dan meminta dana dari orang yang sudah melamar lowongan pekerjaan palsu tersebut. Jadi para pencari kerja harus berhati-hati dan memperhatikan dari mana sumber informasi lowongan pekerjaan yang akan dilamarnya.

Cara supaya terhindar dari kejahatan, lebih baik mencari lowongan pekerjaan dari sumber yang sudah terpercaya dan resmi, salah satunya yaitu pada platform LinkedIn. LinkedIn dapat

diakses pada website maupun aplikasi berbasis desktop dan mobile, yang dapat diunduh di Google Play Store dan App Store. Aplikasi LinkedIn sendiri sudah mencapai 1 miliar lebih pengunduh, dengan rating 4,3 dan terdapat 2,62 juta lebih ulasan berdasarkan data pada Google Play Store. LinkedIn adalah jaringan profesional online terbesar di dunia. LinkedIn bisa digunakan untuk menemukan lowongan dan kesempatan magang, terhubung dan memperkuat hubungan profesional, serta mempelajari keahlian yang dibutuhkan demi kesuksesan karier (LinkedIn). Menggunakan LinkedIn melalui aplikasi, lebih mempermudah para pencari kerja terutama fresh graduate dalam mengakses berbagai informasi lowongan pekerjaan yang terbaru dan bisa mendapatkan notifikasi secara realtime.

Meskipun begitu, tidak menutup kemungkinan tidak adanya permasalahan dalam penggunaan aplikasi LinkedIn. Pasti ada pro kontra yang muncul mengenai kelebihan dan kekurangan dari aplikasi, sehingga para pengguna memberikan berbagai macam ulasan mereka. Suatu ulasan atau online customer reviews adalah ulasan yang diberikan oleh konsumen terkait dengan informasi dari evaluasi suatu produk tentang berbagai macam aspek (Mo, Li, & Fan, 2015). Ulasan yang diberikan oleh pengguna LinkedIn yaitu seperti “Sulit banget loginnya kenapa ya? udah ganti gmailpun juga masih nggak bisa login. Harus verifikasi data padahal sudah kirim KTP masih tetap muncul verifikasi data terus”, “Kok update nya makin nggak bagus, sering error. Upload foto lama dan error saat di akhir loading”, “Kuis keahlian adakan bahasa Indonesia” dan lain sebagainya. Beberapa ulasan tersebut yang diambil dari Google Play Store kebanyakan terkendala di masalah teknis aplikasi, namun ada juga ulasan pengguna yang awam. Mereka masih belum paham dengan fitur-fitur dan bagaimana penggunaan aplikasi LinkedIn dengan efektif, sehingga pengguna menjadi tidak aktif menggunakannya.

Para pengguna tidak hanya memberikan ulasan mengenai permasalahan pada aplikasi, tetapi juga memberikan pujian dan kepuasan dalam menggunakan aplikasi LinkedIn. Terdapat banyak sekali ulasan positif dan negatif yang diberikan oleh para pengguna aplikasi

sebelumnya, tetapi bagi calon pengguna yang melihat ulasan tersebut tidak mengetahui lebih banyak ulasan positif atau negatifnya, sehingga mereka belum mengetahui gambaran kualitas dari aplikasi LinkedIn seperti apa. Oleh karena itu, dengan menganalisis sentimen ulasan pengguna aplikasi, yang mengklasifikasikan ulasan ke positif atau negatif, dapat membantu calon pengguna mengetahui kecenderungan ulasan dan perspektif dari pengguna aplikasi LinkedIn sebelumnya. Begitu juga bagi pengembang aplikasi, dapat melihat ulasan negatif saja dengan mudah, sehingga dapat dijadikan masukan untuk mengembangkan aplikasi LinkedIn yang lebih baik lagi.

Analisis sentimen atau opinion mining merupakan proses memahami, mengekstrak dan mengolah data tekstual secara otomatis untuk mendapatkan informasi sentimen yang terkandung dalam suatu kalimat opini. Analisis sentimen dilakukan untuk melihat pendapat atau kecenderungan opini terhadap sebuah masalah atau objek oleh seseorang, apakah cenderung berpandangan atau beropini negatif atau positif (Rozi, Pramono, & Dahlan, 2012). Pada penelitian analisis sentimen ini, penulis melakukan perbandingan menggunakan 2 metode machine learning yaitu Naïve Bayes dan Support Vector Machine, untuk mengetahui metode mana yang tingkat akurasi lebih tinggi. Persamaan dari metode Naïve Bayes dan Support Vector Machine yaitu sama-sama algoritma supervised learning (pembelajaran yang diawasi) dan dapat digunakan untuk masalah klasifikasi linier, sehingga keduanya dapat dibandingkan. Naïve Bayes merupakan pengklasifikasian dengan bentuk model probabilistik dan statistik yang disederhanakan dengan berdasar pada teorema Bayes dengan asumsi bahwa setiap atribut bersifat bebas (Fikri, Sabrila, & Azhar, 2020). Kelebihannya cepat, efisien, dan mudah untuk dibuat, sedangkan kelemahannya yaitu kebebasan antar atribut membuat akurasi menjadi berkurang. Support Vector Machine (SVM) adalah metode klasifikasi dan regresi yang biasa digunakan untuk masalah linear dan nonlinear (Herlinawati, Yuliani, Faizah, Gata, & Samudi, 2020). Kelebihannya cocok untuk ruang dimensi tinggi, sedangkan kelemahannya yaitu tidak cocok untuk dataset dalam jumlah yang besar.

Pada penelitian sebelumnya dengan judul “Penerapan Algoritma Naïve Bayes untuk Analisis Sentimen Penggunaan Aplikasi Jobstreet” dilakukan pada tahun 2022. Penelitian tersebut berisi tentang klasifikasi analisis sentimen pada ulasan pengguna aplikasi Jobstreet dengan menggunakan metode Naïve Bayes dengan oversampling SMOTE. Hasil pengujian yang didapatkan memiliki nilai accuracy sebesar 90%, precision sebesar 95% dan recall sebesar 93%. Berdasarkan data ulasan, sentimen terhadap aplikasi Jobstreet mendapatkan 83,2% respon positif dan 16,8% sisanya respon negatif, yang berarti opini pengguna lebih cenderung positif (Widodo, Matondang, & Prasvita, 2022). Pada penelitian tersebut, juga hanya menggunakan 1 metode klasifikasi saja yaitu metode Naïve Bayes, sehingga tingkat akurasi tidak dapat dibandingkan dengan metode yang lain untuk mengetahui apakah akurasi metode tersebut ada yang lebih baik atau tidak.

Berdasarkan latar belakang tersebut, maka penulis melakukan penelitian dengan membandingkan 2 metode klasifikasi untuk analisis sentimen dengan objek aplikasi yang berbeda, yaitu dengan judul “Perbandingan Metode Naïve Bayes dan Support Vector Machine untuk Analisis Sentimen pada Ulasan Pengguna Aplikasi LinkedIn”.

2. METODE

Penelitian ini menggunakan metode eksperimen dengan pendekatan kuantitatif untuk membandingkan akurasi metode Naïve Bayes dan Support Vector Machine (SVM) dalam analisis sentimen ulasan pengguna aplikasi LinkedIn di Google Play Store. Data dikumpulkan melalui web scraping menggunakan Google Colab, mengumpulkan 2000 ulasan pengguna. Data kemudian dilabeli menjadi sentimen positif dan negatif, diikuti dengan proses text pre-processing yang meliputi data cleaning, case folding, stopword removal, tokenizing, dan stemming. Selanjutnya, data dibagi menjadi data latih dan data uji, dengan pembobotan TF-IDF dilakukan untuk setiap kata dalam ulasan.

Pada tahap analisis data, kedua metode diuji menggunakan confusion matrix untuk menentukan akurasi. Penelitian ini

menemukan bahwa metode Naïve Bayes dan SVM memiliki perbedaan dalam tingkat akurasi, dengan tujuan akhir untuk menentukan metode mana yang lebih baik dalam mengklasifikasikan sentimen ulasan pengguna aplikasi LinkedIn. Hasil analisis menunjukkan perbandingan akurasi yang lebih baik di antara kedua metode ini, memberikan wawasan tentang efektivitas masing-masing dalam analisis sentimen aplikasi berbasis pengguna.

3. HASIL DAN PEMBAHASAN

Data ulasan pengguna dikumpulkan melalui teknik web scraping dari laman penginstalan aplikasi LinkedIn di Google Play Store. Penulis menggunakan library Google Play Scraper untuk mengumpulkan data ulasan yang paling relevan sebanyak 2000 data ulasan dalam bahasa Indonesia dan berasal dari negara Indonesia. Data yang terkumpul kemudian dilabeli menjadi sentimen positif dan negatif. Proses pelabelan ini bertujuan untuk memisahkan ulasan berdasarkan sentimen yang diekspresikan oleh pengguna, sehingga memudahkan dalam tahap analisis selanjutnya.

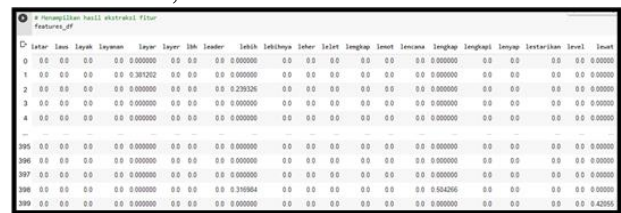


id	username	ulasan	rating	at
1	Adhira S Rini	Ya LinkedIn memang sih mau login aja susah banget ya... udah verifikasi berkali-kali malah dia ga bisa login, cuma diulang-ulang	1	2023-06-18 2:55:52
2	Septia hadi	Kemana pembayaran saldo premium otomatis? Harusnya kan konfirmasi dulu Ulang segebu gitu dipotong otomatis kan	1	2023-06-01 7:31:11
3	Ahlin Dwi Pratama Naphidhu	Tha bta idat bisa login, dan pada saat Verifikasi ID untuk login tidak. Terjadi kesalahan pada saat upload. Masih juga li	1	2023-06-24 12:53:12
4	Dipen Surya Suciarta	Verifikasi sangat lama juga. Padahal udah muncul keterangan "verifikasi selesai anda akan mendapatkan bahwa anda i	1	2023-05-08 12:28:21
5	Faiul Rahman Alpa Maulana	Ini gimana sih, udah berhasil upload postingan gagal mulu. Tolong diing diperbaiki. Padahal jangin lu lancar dan ubi pa	2	2023-05-27 13:20:14
6	Pengembar Yusubda	Suati banget logiknya kenapa ya? udah gnti grupdan juga masih ngga bisa login. Harus verifikasi data padahal sudah	1	2023-04-30 7:17:25
7	Adinda Salfitri	Berapa waktu lagi tolong diperbaiki ini kaget menggorek sekali, karena lagi butuh setting	3	2023-06-26 12:42:42
8	Septi yena	Bener2 aplikasi nya kacau. Masa mau login aja verifikasi ribet bgt, udah bta biar ga bisa akun dhack. Tp ini sih keterlalu	1	2023-06-08 13:42:40
9	Hartono Adu	Tidak bisa masuk... sudah verifikasi menggunakan ktp dan sim dengan foto foto yg benar... sudah muncul terdapat masalah	1	2023-06-16 11:03:11
10	Muhammad Ridy	Sangat buruk, akun saya terkunci kaku. Bagaimana bisa? Sabarannya aplikasi sebesar ini bisa merenggang hal hal seperti	1	2023-06-16 11:03:11
11	Adi Nugroho	Parah, mau dah di upload pake identitas asli tetap ga bisa, malah ada notifikasi "terjadi masalah, harap coba lagi", kan a	2	2023-06-14 3:39:43
12	Asyifa Asyifa	Ini kenapa sih, Aplikasi ny close sendiri mulok. Mau up foto susah klu. Berkali-kali close sendiri	3	2023-06-19 9:51:39
13	Fauzi Khusnulhidayati Alhyar	Logiknya sulu, udah setengah jam lebih ngalotin waktu buat login keamanannya cuman... Mau di verifikasi atau di apps sama sma	3	2023-05-13 13:56:52
14	Rindo Agustin	Wti ini kenapa akun gua ga bisa login? Bisa ga kalau login itu dibuat simple aja. Ini dibuat verifikasi udah benar gua i	3	2023-05-12 0:40:08
15	Rizkaul Fathmah Azharra	King mau login ga bisa bisa sih? Baka apaknya baka kaku. Mau post juga baka kaku.	1	2023-07-02 0:40:48
16	Dijaa Nuyana	Dia memarahi verikasi, membuka iklan benci dan muncul pelajaran... membangun koneksi dengan seluruh pengguna	5	2023-01-03 7:07:02
17	Faiyur Rahman	Indonesia udah banyak pengangguran, ditambah lagi di tpu ama aplikasi beginian, bukan ngasih solusi, malah nambah	1	2023-06-27 1:27:59
18	Sato Nurhidayah	Login aja di persulit, pas kat ulasan lain juga pada sama... pada susah login li	1	2023-07-04 9:14:00
19	0059 Katarina Prati	Ini gimana dah... waktu udah di masuki ny dip nya malah di susah masuki ny tp nya lagi... yaa kan udah tentatif aja... ud	1	2023-06-26 13:16:10
20	Diah Anggrani Pujiandungum	Mohon dapat diperbaiki ketika login, sudah memverifikasi tetap gagal	1	2023-07-06 11:14:31

Gambar 1. Hasil scraping data ulasan

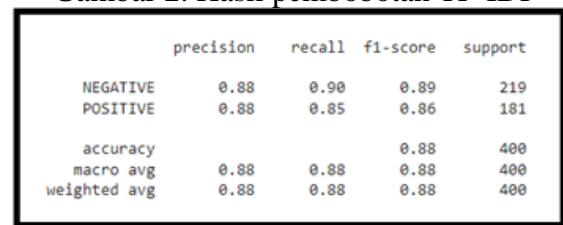
Proses text pre-processing meliputi beberapa tahap yaitu data cleaning, case folding, stopword removal, tokenizing, dan stemming. Pada tahap data cleaning dan case folding, teks diubah menjadi huruf kecil dan tanda baca yang tidak diperlukan dihilangkan. Proses stopword removal dilakukan untuk menghapus kata-kata yang tidak memiliki makna penting. Tokenizing memisahkan teks menjadi kata-kata individu (token), dan stemming mengubah kata-kata menjadi bentuk dasar mereka menggunakan library Sastrawi. Data ulasan yang telah bersih dibagi menjadi data latih (80%) dan data uji (20%). Pembagian ini dilakukan untuk memastikan bahwa model dapat diuji dengan data yang belum pernah dilihat sebelumnya, sehingga evaluasi performa model lebih akurat.

Setelah proses pre-processing, dilakukan pembobotan TF-IDF (Term Frequency-Inverse Document Frequency) untuk menghitung bobot dari setiap kata dalam ulasan. TF-IDF membantu dalam mengidentifikasi kata-kata penting yang sering muncul di satu dokumen tetapi jarang muncul di dokumen lain. Pada tahap pemodelan, metode Naïve Bayes digunakan untuk membangun model klasifikasi sentimen dari data latih. Model ini kemudian diuji menggunakan data uji untuk mengevaluasi performanya. Hasil dari confusion matrix menunjukkan bahwa metode Naïve Bayes memiliki tingkat akurasi sebesar 88%, precision sebesar 88%, recall sebesar 85%, dan f1-score sebesar 86%.



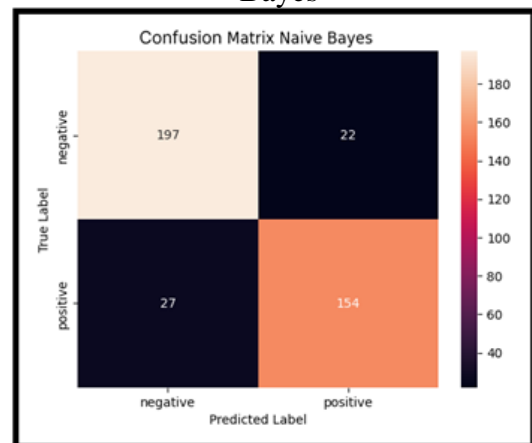
term	tf	idf	tfidf
1	0.0	0.0	0.0
2	0.0	0.0	0.0
3	0.0	0.0	0.0
4	0.0	0.0	0.0
5	0.0	0.0	0.0
6	0.0	0.0	0.0
7	0.0	0.0	0.0
8	0.0	0.0	0.0
9	0.0	0.0	0.0
10	0.0	0.0	0.0
11	0.0	0.0	0.0
12	0.0	0.0	0.0
13	0.0	0.0	0.0
14	0.0	0.0	0.0
15	0.0	0.0	0.0
16	0.0	0.0	0.0
17	0.0	0.0	0.0
18	0.0	0.0	0.0
19	0.0	0.0	0.0
20	0.0	0.0	0.0
21	0.0	0.0	0.0
22	0.0	0.0	0.0
23	0.0	0.0	0.0
24	0.0	0.0	0.0
25	0.0	0.0	0.0
26	0.0	0.0	0.0
27	0.0	0.0	0.0
28	0.0	0.0	0.0
29	0.0	0.0	0.0
30	0.0	0.0	0.0
31	0.0	0.0	0.0
32	0.0	0.0	0.0
33	0.0	0.0	0.0
34	0.0	0.0	0.0
35	0.0	0.0	0.0
36	0.0	0.0	0.0
37	0.0	0.0	0.0
38	0.0	0.0	0.0
39	0.0	0.0	0.0
40	0.0	0.0	0.0
41	0.0	0.0	0.0
42	0.0	0.0	0.0
43	0.0	0.0	0.0
44	0.0	0.0	0.0
45	0.0	0.0	0.0
46	0.0	0.0	0.0
47	0.0	0.0	0.0
48	0.0	0.0	0.0
49	0.0	0.0	0.0
50	0.0	0.0	0.0
51	0.0	0.0	0.0
52	0.0	0.0	0.0
53	0.0	0.0	0.0
54	0.0	0.0	0.0
55	0.0	0.0	0.0
56	0.0	0.0	0.0
57	0.0	0.0	0.0
58	0.0	0.0	0.0
59	0.0	0.0	0.0
60	0.0	0.0	0.0
61	0.0	0.0	0.0
62	0.0	0.0	0.0
63	0.0	0.0	0.0
64	0.0	0.0	0.0
65	0.0	0.0	0.0
66	0.0	0.0	0.0
67	0.0	0.0	0.0
68	0.0	0.0	0.0
69	0.0	0.0	0.0
70	0.0	0.0	0.0
71	0.0	0.0	0.0
72	0.0	0.0	0.0
73	0.0	0.0	0.0
74	0.0	0.0	0.0
75	0.0	0.0	0.0
76	0.0	0.0	0.0
77	0.0	0.0	0.0
78	0.0	0.0	0.0
79	0.0	0.0	0.0
80	0.0	0.0	0.0
81	0.0	0.0	0.0
82	0.0	0.0	0.0
83	0.0	0.0	0.0
84	0.0	0.0	0.0
85	0.0	0.0	0.0
86	0.0	0.0	0.0
87	0.0	0.0	0.0
88	0.0	0.0	0.0
89	0.0	0.0	0.0
90	0.0	0.0	0.0
91	0.0	0.0	0.0
92	0.0	0.0	0.0
93	0.0	0.0	0.0
94	0.0	0.0	0.0
95	0.0	0.0	0.0
96	0.0	0.0	0.0
97	0.0	0.0	0.0
98	0.0	0.0	0.0
99	0.0	0.0	0.0
100	0.0	0.0	0.0

Gambar 2. Hasil pembobotan TF-IDF

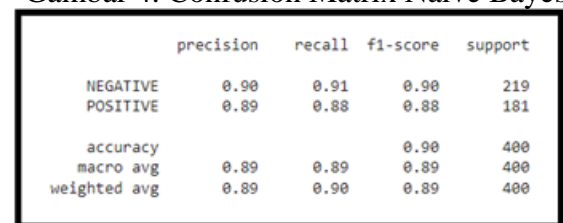


	precision	recall	f1-score	support
NEGATIVE	0.88	0.90	0.89	219
POSITIVE	0.88	0.85	0.86	181
accuracy			0.88	400
macro avg	0.88	0.88	0.88	400
weighted avg	0.88	0.88	0.88	400

Gambar 3. Nilai performa Pemodelan Naïve Bayes

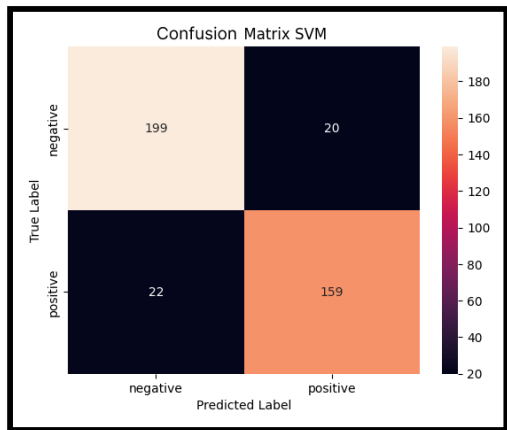


Gambar 4. Confusion Matrix Naïve Bayes



	precision	recall	f1-score	support
NEGATIVE	0.90	0.91	0.90	219
POSITIVE	0.89	0.88	0.88	181
accuracy			0.90	400
macro avg	0.89	0.89	0.89	400
weighted avg	0.89	0.90	0.89	400

Gambar 5. Nilai performa pemodelan SVM



Gambar 6. Confusion Matrix SVM

Metode SVM juga digunakan untuk membangun model klasifikasi sentimen dari data latih. Hasil evaluasi menggunakan confusion matrix menunjukkan bahwa SVM memiliki tingkat akurasi sebesar 89%, precision sebesar 88%, recall sebesar 88%, dan f1-score sebesar 88%. Hasil ini menunjukkan bahwa SVM memiliki performa yang sedikit lebih baik dibandingkan dengan Naïve Bayes. Berdasarkan hasil evaluasi, metode SVM menunjukkan performa yang lebih baik dibandingkan dengan Naïve Bayes dalam mengklasifikasikan sentimen ulasan pengguna aplikasi LinkedIn. Dengan tingkat akurasi, precision, recall, dan f1-score yang lebih tinggi, SVM terbukti lebih efektif dalam analisis sentimen untuk data ulasan aplikasi ini.

4. PENUTUP

4.1. Kesimpulan

Berdasarkan pada hasil penelitian yang telah penulis lakukan dapat diambil kesimpulannya, yaitu hasil analisis sentimen pada ulasan pengguna aplikasi LinkedIn menggunakan metode Naïve Bayes, menghasilkan nilai accuracy sebesar 88%, precision sebesar 88%, recall sebesar 85%, dan f1-score sebesar 86%. Sedangkan metode Support Vector Machine menghasilkan nilai yang lebih tinggi, yaitu accuracy sebesar 90%, precision sebesar 89%, recall sebesar 88%, dan f1-score sebesar 88%. Terlihat dari perbandingan performa dua metode tersebut, maka metode Support Vector Machine memiliki tingkat akurasi yang lebih baik dibandingkan dengan metode Naïve Bayes. Hal itu ditunjukkan dengan metode Support Vector Machine yang lebih unggul dalam memprediksi

label sentimen, karena terdapat lebih banyak jumlah label prediksi yang benar dan jumlah label prediksi yang salah lebih sedikit daripada metode Naïve Bayes.

4.2. Saran

Untuk penelitian selanjutnya, disarankan untuk memperluas jumlah data ulasan yang digunakan serta melibatkan ulasan dalam berbagai bahasa untuk menguji generalisasi model pada data multibahasa. Selain itu, dapat diintegrasikan metode ensemble learning untuk melihat apakah kombinasi dari beberapa algoritma dapat meningkatkan akurasi dan performa analisis sentimen. Penggunaan teknik deep learning seperti LSTM atau BERT juga dapat dieksplorasi untuk membandingkan hasil dengan metode machine learning tradisional, serta menyertakan analisis terhadap aspek-aspek tertentu dari ulasan seperti fitur aplikasi yang paling banyak dikritik atau dipuji oleh pengguna.

5. DAFTAR PUSTAKA

- Creswell, J. W. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Thousand Oaks, CA: Sage.
- Fikri, A., Sabrila, M., & Azhar, M. (2020). Penerapan algoritma Naïve Bayes pada analisis sentimen. *Jurnal Ilmiah Teknologi Informasi*, 14(2), 89-97.
- Hanafi, A., & Ferdiansyah, A. (2020). Preprocessing data teks. *Jurnal Informatika*, 8(1), 45-53.
- Herlinawati, E., Yuliani, N., Faizah, G., Gata, T., & Samudi, R. (2020). Penerapan metode Support Vector Machine pada klasifikasi data. *Jurnal Sistem Informasi*, 12(3), 123-130.
- Karsito, & Susanti, A. (2019). Confusion matrix untuk pengukuran performa metode klasifikasi. *Jurnal Teknologi Informasi*, 11(2), 78-85.
- Maturidi, A. (2014). *Metodologi penelitian: Pendekatan teori dan aplikasi*. Jakarta: Pustaka Al-Bayan.
- Mo, Z., Li, Y., & Fan, P. (2015). Customer reviews on e-commerce platforms:

- Sentiment analysis and data mining. *Journal of Business Research*, 68(5), 1034-1041.
- Nurjanah, S., Perdana, T., & Fauzi, A. (2017). Pembobotan kata dengan TF-IDF pada analisis sentimen. *Jurnal Informatika*, 9(2), 145-153.
- Oliver, T. (2022). Google Colaboratory: Panduan lengkap. Retrieved from <https://colab.research.google.com/>
- Putra, M. P. (2020). Analisis data dan pengambilan keputusan. Jakarta: Media Ilmu.
- Ramadinah, N. (2021). Pengukuran performa model klasifikasi menggunakan confusion matrix. *Jurnal Informatika*, 13(1), 45-53.
- Rozi, S., Pramono, E., & Dahlan, M. (2012). Opinion mining: Penerapan analisis sentimen dalam data teks. *Jurnal Informatika*, 7(1), 67-75.
- Sadya, T. (2022). Tingkat pengangguran terbuka di Indonesia 2022. Badan Pusat Statistik.
- Simbolon, F. (2023). Jumlah pencari kerja di Indonesia tahun 2022. Badan Pusat Statistik.
- Suyanto. (2017). Machine learning dan aplikasinya. Yogyakarta: Penerbit Andi.
- Tineges, T., Triayudi, & Sholihati, N. (2020). Metode SVM pada klasifikasi data teks. *Jurnal Informatika*, 10(2), 652-653.
- Trivusi, E. (2022). Jenis-jenis Support Vector Machine. *Jurnal Informatika*, 11(3), 87-95.
- Turland, T. (2010). Web scraping dan aplikasi dalam pengumpulan data. *Journal of Web Technology*, 6(1), 33-45.
- Vijayarani, S., Ilamathi, M. J., & Nithya, M. (2015). Preprocessing techniques for text mining: An overview. *International Journal of Computer Science & Communication Networks*, 5(1), 7-16.
- Wandani, D., Fauziah, N., & Andrianingsih, T. (2021). Text mining dalam analisis sentimen. *Jurnal Informatika*, 15(1), 12-23.
- Widodo, B. K., Matondang, N. H., & Prasvita, D. S. (2022). Penerapan algoritma Naïve Bayes untuk analisis sentimen penggunaan aplikasi Jobstreet. *Techno.COM*, 14(2), 89-97.
- Yunefri, Fadrial, & Sutejo, D. (2021). Pembobotan kata pada analisis sentimen dengan TF-IDF. *Jurnal Informatika*, 9(1), 45-53.