
Komparasi Model Analisis Sentimen Pada Twitter Terhadap Kemahalan Minyak Goreng dengan Metode *Naive Bayes* dan *Support Vector Machine*

Moh.Aminullah Al Fachri¹, Ummi Athiyah^{2*}

¹Program Studi Informatika, Institut Teknologi Telkom Purwokerto

²Program Studi Sains Data, Institut Teknologi Telkom Purwokerto

^{1,2}Jl. DI Pandjaitan 128 Purwokerto Selatan, Banyumas 53147, Indonesia

E-mail: 18102166@ittelkom-pwt.ac.id¹, ummi@ittelkom-pwt.ac.id²

Info Naskah:

Naskah masuk: 30 Jan 2023

Direvisi: 25 Juni 2023

Diterima: 27 Juni 2023

Abstrak

Akhir tahun 2021 masyarakat dihebohkan dengan persediaan minyak goreng yang berkurang secara drastis dan harga yang mahal. Hal tersebut membuat masyarakat banyak membicarakannya melalui media sosial seperti twitter. Kebebasan pada twitter menimbulkan banyak respon dari masyarakat. Banyaknya respon negatif dan positif pada twitter membuat perbandingan kedua respon tersebut sulit diamati. Penelitian ini bertujuan untuk mengetahui perbandingan respon positif dan respon negatif. *Machine learning* dengan metode *naive bayes* dan *support vector machine* mampu mengatasi masalah tersebut. Penelitian yang dilakukan meneliti bagaimana perbandingan antara respon positif dan respon negatif serta metode mana yang mempunyai akurasi lebih tinggi. Data yang digunakan adalah 10.000 data cuitan berbahasa Indonesia. Pengujian model dilakukan dengan 1839 data uji. metode *Naive Bayes* mendapatkan akurasi 74,06% dengan hasil prediksi dua tweet positif dan 1837 tweet negatif. Metode SVM dilakukan pengujian pada kernel linear, polynomial, RBF dan sigmoid. Kernel dengan nilai akurasi tertinggi adalah kernel sigmoid dengan akurasi 81.8% dengan hasil prediksi 266 tweet positif dan 1573 tweet negatif.

Keywords:

machine learning;

naive bayes;

text mining;

SVM;

twitter;

Abstract

At the end of 2021, people are shocked by the drastically reduced supply of cooking oil and high prices. This makes people talk about it a lot through social media like Twitter. Freedom on Twitter raises many responses from the public. The number of negative and positive responses on Twitter makes comparisons between the two responses difficult to observe. This study aims to determine the comparison of positive responses and negative responses. Machine learning with the naive bayes method and support vector machine is able to overcome this problem. The research conducted examines how the comparison between positive responses and negative responses and which method has higher accuracy. The data used is 10,000 Indonesian language tweets. Model testing was carried out with 1839 test data. The Naive Bayes method gets an accuracy of 74.06% with the results of predicting two positive tweets and 1837 negative tweets. The SVM method was tested on linear, polynomial, RBF, and sigmoid kernels. The kernel with the highest accuracy value is the sigmoid kernel with an accuracy of 81.8% with the predicted results of 266 positive tweets and 1573 negative tweets.

*Penulis korespondensi:

Ummi Athiyah

E-mail: ummi@ittelkom-pwt.ac.id

1. Pendahuluan

Sembilan bahan pokok atau sering disebut sebagai sembako merupakan hal yang dibutuhkan oleh manusia guna memenuhi kebutuhan pangannya. Beberapa faktor dapat mempengaruhi dalam pemenuhan kebutuhan tersebut, seperti ketersediaan stok, harga dan daya beli masyarakat. Faktor-faktor tersebut saling berkaitan, saat harga produk naik maka daya beli masyarakat akan berkurang. Minyak goreng merupakan bahan yang tidak dapat digunakan secara berulang kali karena dapat menyebabkan penyakit pada manusia sehingga kebutuhan minyak goreng terus bertambah [1]. Berdasarkan data dari Survei Ekonomi Nasional (SUSENAS) dari Badan Pusat Statistik pada tahun 2015 – 2020 konsumsi minyak goreng sawit pada tingkat rumah tangga mengalami kenaikan rata-rata 2,32 setiap tahunnya [2].

Pada akhir tahun 2021 hingga April 2022 masyarakat Indonesia mengalami kesulitan dalam mendapatkan minyak goreng. Kelangkaan stok minyak goreng di berbagai tempat ini tentunya berdampak terhadap harga dan daya beli masyarakat. Keadaan dilapangan saat itu permintaan minyak goreng dari masyarakat tetap sedangkan stok di lapangan semakin berkurang. Menurut ilmu ekonomi kelangkaan terjadi karena sumber daya ekonomi tidak dapat memenuhi kebutuhan pasar yang tak terbatas. Kelangkaan stok yang tersedia akan membuat harga produk semakin tinggi [3].

Dikutip dari datareportal.com jumlah pengguna twitter di Indonesia berada di peringkat 5 teratas setelah Amerika Serikat, Jepang, India dan Brazil. Pada bulan Januari jumlah akun twitter yang aktif di Indonesia sebesar 19,9 juta akun. Di Indonesia twitter berada di peringkat 6 sosial media terpopuler dan berada di peringkat 5 sosial media favorit di Indonesia. Twitter merupakan salah satu pilihan dari beberapa media sosial yang legal di Indonesia. Jack Dorsey pada Maret tahun 2006 mendirikan twitter dan pertama kali diluncurkan pada bulan Juli 2006. Saat itu twitter berhasil berada di sepuluh besar sosial media yang paling sering digunakan sejak diluncurkan [4].

Kebebasan dalam mengungkapkan gagasan pada twitter dimanfaatkan oleh pengguna salah satunya adalah mengungkapkan keresahan pengguna atas apa yang sedang ramai dibicarakan. Topik yang ramai salah satunya adalah mengenai mahalannya harga minyak goreng di Indonesia. Topik tersebut mendapatkan berbagai tanggapan dari pengguna baik itu tanggapan negatif atau tanggapan positif. Dari banyaknya *tweet* yang ada tentunya akan timbul suatu masalah yaitu sulit untuk mengamati berapa persen tanggapan positif dan negatif.

Pemrosesan Bahasa alami adalah penerapan ilmu komputer yang mendalami sistem bagaimana komputer dapat berinteraksi dengan manusia. Interaksi antara manusia dan komputer dapat dilakukan dengan media Bahasa. Komputer tidak akan memahami Bahasa manusia secara langsung sehingga dibutuhkan perantara untuk mengubah bahasa tersebut dengan aturan gramatikal dan semantiknya [4]. Analisis sentiment berguna dalam pengolahan Bahasa alami, teks mining serta komputasi linguistic dimana hal tersebut digunakan untuk menentukan opini pada sebuah topik. Analisis sentimen menghasilkan dua atau lebih polaritas, umumnya adalah polaritas negatif dan polaritas

positif. Polaritas tersebut dapat digunakan sebagai bahan pertimbangan untuk dijadikan sebuah keputusan [5].

Penelitian yang dilakukan oleh Elly Indrayuni dengan topik *sentiment* analisis terhadap review produk kosmetik telah dilakukan pada tahun 2019. Peneliti mengambil 200 komentar yang terdiri dari 100 komentar positif dan 100 komentar negatif data dari website yang berbahasa Indonesia yaitu <https://femaledaily.com/>. Pemilihan komentar positif dan negatif tersebut dilakukan peneliti secara manual. Pengolahan data digunakan metode *Naive Bayes* yang dioptimalkan dengan menambah proses N-Gram pada tahap *preprocessing*. Dari hasil pengujian yang dilakukan didapatkan tingkat akurasi sebesar 90.50% dan nilai AUC sebesar 0.715 [6].

Penelitian yang dilakukan oleh Taofik Krisdiyanto dan Erry Maricha Oki Nurharyanto dengan topik sentimen analisis masyarakat terhadap penerapan PPKM. Data yang digunakan berupa teks dari tagar “ppkm” di twitter sejumlah 1000 baris data. Data tersebut dilakukan tahap *preprocessing* untuk mendapatkan data yang matang, kemudian diklasifikasikan kedalam 2 klasifikasi yaitu positif dan negatif. Untuk mengklasifikasikan data tersebut peneliti menggunakan metode *naive bayes*. Hasil penelitian didapatkan 99% data *tweet* dari masyarakat termasuk kedalam polaritas positif dan sisanya polaritas negatif. [7].

Penelitian yang dilakukan oleh Rachmad Mahendrajaya, Ghulam Asrofi Buntoro dan Moh. Bhanu Setyawan pada 2019 menganalisis sentimen masyarakat terhadap pengguna Go-Pay. Penelitian ini menggunakan data dari twitter sebanyak 1210 baris data. *Dataset* tersebut dianalisis dengan metode pelabelan *lexicon based* kemudian diklasifikasikan dengan metode *Support Vector Machine*. Dari hasil pengujian dengan *confusion matrix* didapatkan hasil akurasi sebesar 89,17% dengan jumlah ulasan positif sebanyak 923 dan 287 ulasan negatif [8].

Rian Tinages, Agung Triayudi dan Ira Diana Sholihati pada tahun 2020 melakukan penelitian terhadap sentimen analisis pelanggan indihome. Penelitian ini dilakukan menggunakan data dari twitter, diambil dari *tweet* yang menyebutkan username @IndiHome sejumlah 1000 baris data kemudian dilakukan tahap *preprocessing* dan didapatkan *dataset* matang sebanyak 1400 baris. Tahap analisis dilakukan menggunakan metode *Support Vector Machine*. Hasil *confusion matrix* didapatkan tingkat akurasi sebesar 87% dengan 367 *tweet* memiliki polaritas positif dan 83 *tweet* memiliki polaritas negatif [9].

Berdasarkan penelitian - penelitian yang sudah dilakukan di atas metode *Naive bayes* dan *Support vector machine* digunakan sebagai metode klasifikasi dengan sebagian besar *dataset* yang digunakan adalah data teks. Dari hasil pengujian yang dilakukan kedua metode tersebut mampu menghasilkan tingkat akurasi yang tinggi. Sehingga metode tersebut cocok untuk dijadikan sebagai metode penelitian untuk menganalisis sentimen masyarakat terhadap mahalannya harga minyak goreng. Penelitian yang dilakukan oleh Elly Indrayuni dan penelitian yang dilakukan oleh Rachmad Mahendrajaya akan dijadikan sebagai referensi utama dalam penelitian yang akan dilakukan.

Berdasarkan masalah tersebut pada penelitian ini akan dilakukan analisis sentimen data *tweet* pengguna dari twitter

dengan tema mahal nya harga minyak goreng di Indonesia. Dalam penelitian ini akan menggunakan dua buah metode sebagai bahan perbandingan yaitu *Naive bayes* dan *Support vector machine* (SVM). *Dataset* yang digunakan akan melewati proses *preprocessing* untuk mengolah *dataset* menjadi lebih matang sehingga mampu memberikan hasil yang lebih baik. *Dataset* akan diklasifikasikan menjadi 2 kelas yaitu kelas positif dan kelas negatif. Berdasarkan penelitian yang dilakukan oleh Muhammad Akbar S dengan metode SVM membagi *dataset* dengan perbandingan data latih dan data tes 75%:25%, 80%:20% dan 90%:10% menghasilkan tingkat akurasi tertinggi pada perbandingan 80%:20% dibandingkan dua perbandingan lainnya [10]. Penelitian yang dilakukan oleh Winda Christina Widyaningtyas dengan metode *Naive Bayes* membagi data latih dan data uji dengan perbandingan 20%:80%, 50%:50% dan 80%:20% menghasilkan tingkat akurasi tertinggi pada perbandingan 80%:20% dibandingkan dua perbandingan lainnya [11]. Berdasarkan hal tersebut dalam penelitian ini dilakukan pembagian data dengan persentase 80% data latih dan 20% data uji .

2. Metode

Penelitian dilakukan dengan beberapa tahapan secara berurutan seperti tertera pada gambar 2.1. Dimulai dengan tahap identifikasi masalah dan studi literatur untuk menemukan pokok masalah dan penelitian terdahulu sebagai dasar penunjang penelitian. Langkah selanjutnya adalah tahap pengambilan data dengan cara *crawling* pada twitter. Tahap *crawling* dilakukan satu kali dengan mengambil 10.000 data berbahasa Indonesia dengan kata pencarian “minyak goreng mahal”. Hasil pengambilan data berupa data teks yang terdiri dari *tweet*, *username*, dan *date*. Data yang akan diproses dalam tahap analisis adalah data *tweet*.

Tahap selanjutnya adalah tahap mengkonversi data mentah menjadi data matang dengan tahap yang dikenal dengan *preprocessing*, tahapan *preprocessing* yang dilakukan diantaranya:

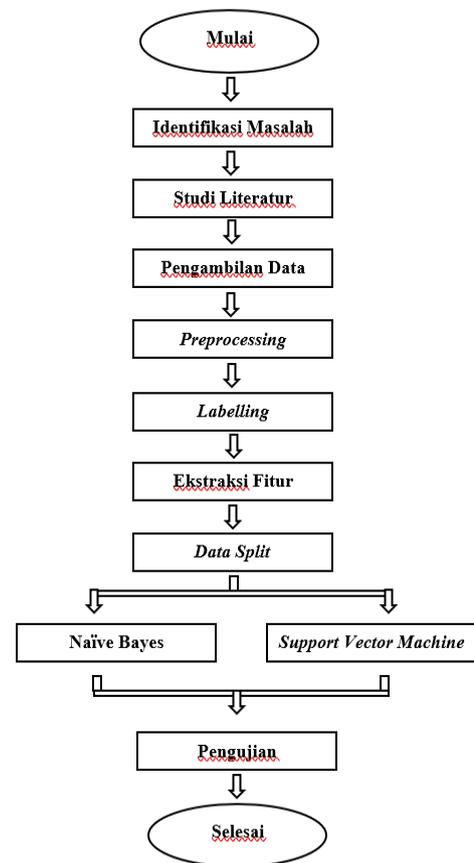
- 1) *Case folding* adalah tahap mengubah huruf ke bentuk huruf kecil semua.
- 2) *Filtering* adalah tahap membersihkan data yang dilakukan dengan menghilangkan beberapa karakter seperti username, tagar, emoji, simbol dan lain-lain.
- 3) Konversi *Three or More* adalah tahap mengkonversi karakter sama lebih dari tiga dan letaknya berurutan menjadi 1 karakter.
- 4) Konversi Kata Slang adalah tahap mengubah kata slang atau kata gaul menjadi kata tertentu sesuai dengan kamus yang dibuat.
- 5) *Stemming* adalah tahap mengubah kata berimbuhan menjadi kata dasarnya.

Tahap selanjutnya adalah tahap *labelling* atau pelabelan, metode *Naive Bayes* serta SVM merupakan bagian dari *supervised learning* sehingga diperlukan label sebagai bahan untuk membangun model. Tahap pelabelan dilakukan menggunakan metode *lexicon-based* dengan bantuan *library* Textblob. TextBlob memiliki 2919 kata berbahasa Inggris dimana masing-masing kata sudah memiliki polaritas masing-masing. Dalam menentukan

polaritas TextBlob akan melihat apakah kalimat tersebut memiliki skor positif atau negatif [12]

Tahap selanjutnya adalah tahap ekstraksi fitur yaitu tahap pembobotan dengan metode *Term Frequency Inverse Document Frequency* (TF-IDF). Pembobotan ini akan mengubah data teks menjadi numerik untuk dapat dianalisis dengan *Naive Bayes* dan SVM. *Supervised learning* selain membutuhkan label sebagai pembangunan model juga membutuhkan data latih sehingga dilakukan tahap *data split* dengan membagi 9194 data hasil *preprocessing* menjadi data latih dan data tes.

Analisis data dilakukan dengan metode *Naive Bayes* serta SVM menggunakan data latih sebanyak 7355 baris data. *Naive Bayes* adalah metode klasifikasi yang diambil dari teorema Bayes. Metode ini bekerja dengan cara melakukan analisis terhadap data latih yang sudah terdapat label di masing-masing baris data. Algoritma hasil pelatihan atau model dapat digunakan untuk mengklasifikasikan data baru berdasarkan polaritas yang ada.



Gambar 1. Diagram Alir Penelitian

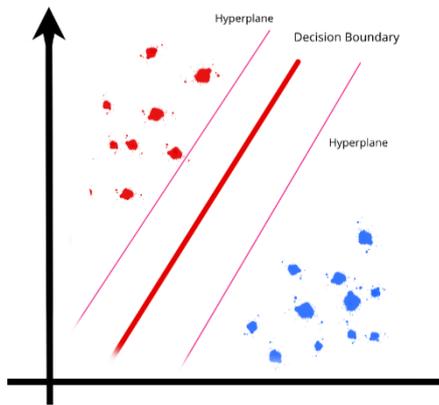
Bentuk dasar teorema Bayes dapat diamati dalam persamaan 1. *posterior probability* adalah peluang terjadinya H jika terdapat sebuah *evidence* A. *Likelihood* adalah probabilitas A terhadap hipotesis H. *Prior* adalah probabilitas hipotesis tanpa dipengaruhi oleh hal apapun. Nilai *Evidence* dalam praktiknya selalu bernilai konstan sehingga dapat diabaikan. Perhitungan hanya terfokus terhadap perkalian antara nilai *prior* dan nilai *likelihood* ketika *evidence* diabaikan[13].

$$P(A) = \frac{P(A|H)P(H)}{P(A)} \quad (1)$$

Keterangan:

- P(H|A) : posterior probability
- P(H) : prior probability
- P(A|H) : likelihood
- P(A) : evidence

Support vector machine pada dasarnya bekerja dengan linear classifier atau bisa disebut dengan klasifikasi biner. Sebagai bagian dari supervised learning, Support vector machine akan bekerja dengan data latih yang sudah memiliki label kelas yang jelas. Hasil dari pelatihan Support vector machine akan mampu memprediksi data baru akan masuk ke dalam kelas tertentu. Untuk menentukan hasil prediksi metode Support vector machine menggunakan decision boundary atau batas keputusan, ilustrasi mengenai decision boundary dapat diamati dalam gambar 2.



Gambar 2. Decision Boundary

Metode Support vector machine juga mampu melakukan klasifikasi non-linear yang dilakukan dengan menggunakan kernel. Klasifikasi dengan menggunakan fungsi kernel digunakan ketika mengoperasikan ruang fitur berdimensi tinggi yang terpisah secara non-linear. Fungsi kernel yang dapat digunakan dalam klasifikasi menggunakan Support vector machine sebagai berikut [14].

Kernel Polynomial

$$k(\hat{x}_i, \hat{x}_j) = (\hat{x}_i \hat{x}_j + 1)^p \quad (2)$$

Kernel RBF

$$k(\hat{x}_i, \hat{x}_j) = \exp \quad (3)$$

$$\left(-\frac{\|\hat{x}_i - \hat{x}_j\|^2}{2\sigma^2} \right) \quad (3)$$

Kernel Sigmoid

$$k(\hat{x}_i, \hat{x}_j) = \tanh(\alpha \hat{x}_i \hat{x}_j + \beta) \quad (4)$$

Model yang dihasilkan dari tahap analisis kemudian dilakukan pengujian dengan confusion matrix untuk menemukan nilai accuracy dengan persamaan 5 [15]:

$$\frac{TruePs + TrueNg}{\sum \text{ data}} \quad (5)$$

untuk menemukan nilai precision dengan persamaan 6.

$$\frac{TruePs}{TruePs + FalsePs} \quad (6)$$

untuk menemukan nilai recall dengan persamaan 7.

$$\frac{TruePs}{TruePs + FalseNg} \quad (7)$$

untuk menemukan nilai F1-Score dengan persamaan 8.

$$\frac{2 * precision * recall}{precision + recall} \quad (8)$$

3. Hasil dan Pembahasan

3.1 Hasil Preprocessing

Berdasarkan 10.000 baris data yang digunakan setelah melewati tahap preprocessing dan tahap pelabelan didapatkan 9194 baris data matang. Tahap pelabelan menghasilkan 6829 memiliki polaritas negatif dan 2365 memiliki label positif. Data tersebut dibagi menjadi dua bagian yaitu data latih dan data tes. Hasil pembagian data didapatkan 7355 baris data training dan 1839 baris data testing. Data training memiliki 1888 baris data dengan label positif dan 5467 baris data dengan label negatif. Data testing memiliki 477 baris data dengan label positif dan 1362 baris data dengan label negatif.

3.2 Hasil Analisis Naïve bayes

Tahap analisis menggunakan metode Naive Bayes menghasilkan nilai True Positive sebanyak 1, False Positive sebanyak 1, False Negative sebanyak 476, True Negative sebanyak 1361 atau tertera pada tabel 1.

Tabel 1. Confusion Matrix Naive Bayes

| | | NILAI SEBENARNYA | |
|---------|---------|------------------|---------|
| | | Positif | Negatif |
| PREDIKS | Positif | 1 | 1 |
| | Negatif | 476 | 1361 |

Berdasarkan tabel 2 dan menggunakan persamaan 5,6,7,8 didapatkan nilai accuracy sebesar 74,06% , nilai precision sebesar 50%, nilai recall sebesar 0,2% dan fi-score 0,4%.

3.3 Hasil Analisis Support Vector Machine

Tahap analisis menggunakan metode SVM pada masing masing kernel didapatkan hasil yang berbeda sesuai dengan tabel 2.

Tabel 2. *Confusion Matrix SVM*

| | | SEBENARNYA | |
|----------|-------------------|------------|---------|
| | | Positif | Negatif |
| P | Linear | Positif | 205 |
| | | Negatif | 66 |
| R | Polynomial | Positif | 6 |
| | | Negatif | 3 |
| E | RBF | Positif | 471 |
| | | Negatif | 1359 |
| D | Sigmoid | Positif | 135 |
| | | Negatif | 38 |
| I | | Positif | 324 |
| | | Negatif | 1324 |
| K | | Positif | 205 |
| | | Negatif | 61 |
| S | | Positif | 272 |
| | | Negatif | 1301 |
| I | | Positif | 205 |
| | | Negatif | 66 |

1) *Kernel Linear*

Kernel linear mendapatkan hasil True Positif sebanyak True Positive sebanyak 205, False positive sebanyak 66, False Negative sebanyak 272 dan True Negative 1296. Hasil evaluasi yang didapatkan adalah accuracy sebesar 81,6%, precision sebesar 75,6%, recall sebesar 42,9% dan F1-Score sebesar 54,8%.

2) *Kernel Polynomial*

Kernel Polynomial menghasilkan True Positive sebanyak 6, False positive sebanyak 3, False Negative sebanyak 471 dan True Negative 1359. Hasil evaluasi kernel polynomial menghasilkan accuracy 74,2%, precision 66,6%, recall 1,25% dan F1-Score 2,4%

3) *Kernel RBF*

Kernel RBF menghasilkan True Positive sebanyak 135, False positive sebanyak 38, False Negative sebanyak 324 dan True Negative 1324. Hasil evaluasi kernel RBF menghasilkan accuracy 79,3%, precision 78%, recall 28,3% dan F1-Score 41,5%

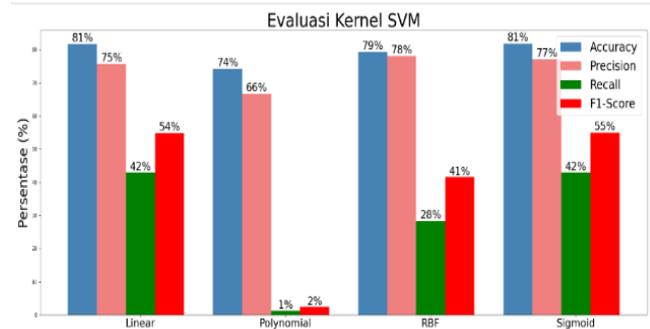
4) *Kernel Sigmoid*

Kernel RBF menghasilkan True Positive sebanyak 205, False positive sebanyak 61, False Negative sebanyak 272 dan True Negative 1301. Hasil evaluasi kernel RBF menghasilkan accuracy 81,8%, precision 77%, recall 42,9% dan F1-Score 55%

Nilai accuracy menunjukkan perbandingan antara hasil prediksi benar dengan keseluruhan data tes, kernel yang memiliki accuracy terbaik adalah kernel sigmoid yang mampu memprediksi 1506 baris data set dengan benar. Nilai precision menunjukkan bahwa perbandingan antara True Positif dan data yang diprediksi positif. Kernel RBF adalah kernel dengan nilai precision tertinggi dengan nilai True Positif 135 dan data diprediksi positif 173. Nilai recall menunjukkan perbandingan antara True Positif dengan baris dataset yang berlabel positif. Kernel linear dan sigmoid memiliki nilai True Positif yang sama sehingga juga memiliki nilai recall yang. Kernel linear dan sigmoid merupakan kernel dengan nilai recall tertinggi dengan nilai True Positif 205. Nilai F1-Score menunjukkan tinggi atau

rendahnya nilai precision dan recall karena F1-Score bergantung terhadap nilai precision dan recall.

Hasil evaluasi SVM dengan accuracy tertinggi dibandingkan dengan hasil evaluasi metode Naive Bayes memiliki perbedaan yang cukup signifikan. Selisih accuracy antara kernel sigmoid pada SVM dan Naive Bayes mencapai 7,8% dan selisih precision mencapai 27% atau dapat diamati pada grafik gambar 3. selisih tersebut menunjukkan bahwa metode SVM kernel sigmoid dengan dataset yang digunakan lebih baik dalam memprediksi label dataset.



Gambar 3. Perbandingan Naive Bayes dan SVM

4. Kesimpulan

Berdasarkan hasil pengujian yang dilakukan dapat ditarik kesimpulan bahwa kedua metode memiliki hasil analisis yang berbeda. Perbedaan hasil evaluasi terjadi karena perbedaan prediksi masing-masing metode. Metode Naive bayes menghasilkan akurasi sebesar 74,06% sedangkan metode SVM dengan kernel sigmoid menghasilkan akurasi sebesar 81,6%. Hasil evaluasi dapat ditingkatkan kembali dengan tahap preprocessing yang lebih detail untuk didapatkan data yang lebih bersih. Selain memperbaiki preprocessing juga diperlukan pelabelan secara manual dengan menggunakan tenaga ahli pada bidang Bahasa.

Ucapan Terimakasih

Terima kasih disampaikan kepada seluruh civitas Institut Teknologi Telkom Purwokerto atas segala dukungan materil maupun non-materil dalam menyelesaikan penelitian ini.

Daftar Pustaka

- [1] H. N. Muhammad, F. Nikmah, N. U. Hidayah, and A. K. Haqiqi, "Arang Aktif Kayu Leucaena Leucocephala sebagai Adsorben Minyak Goreng Bekas Pakai (Minyak Jelantah)," *Phys. Educ. Res. J.*, vol. 2, no. 2, p. 123, Aug. 2020, doi: 10.21580/perj.2020.2.2.6176.
- [2] T. Kurniasih, G. A. Utama, and S. R. Rahmad, *Distribusi Perdagangan Komoditas Minyak Goreng Indonesia 2021*. Jakarta, 2021.
- [3] J. Sipayung, "Pengaruh Persepsi Kelangkaan dan Antisipasi Penyesalan Terhadap Urgensi untuk Membeli, Perilaku Penimbunan dan Penyembunyian dalam Toko dengan Daya Saing dan Motivasi Belanja Hedonis Sebagai Variabel Moderasi (Studi pada Store UNIQLO & H&M)," ATMA JAYA YOGYAKARTA, 2020.
- [4] R. Fajar, S. Program, P. Rekeyasa, N. Lunak, and R.

- Bengkalis, "Implementasi Algoritma *Naive Bayes* Terhadap Analisis Sentimen Opini Film Pada Twitter," vol. 3, no. 1, pp. 50–59, 2018.
- [5] S. Azza Amira, S. Utama, and D. Muhammad Hanif Fahmi, "Penerapan Metode *Support vector machine* untuk Analisis Sentimen pada Review Pelanggan Hotel," *Edu Komputika*, vol. 7, no. 2, pp. 40–48, 2020.
- [6] E. Indrayuni, "Klasifikasi Text Mining Review Produk Kosmetik Untuk Teks Bahasa Indonesia Menggunakan Algoritma *Naive Bayes*," *J. KHATULISTIWA Inform.*, vol. VII, no. 1, pp. 29–36, 2019.
- [7] T. Krisdiyanto, "Analisis Sentimen Opini Masyarakat Indonesia Terhadap Kebijakan PPKM pada Media Sosial Twitter Menggunakan *Naive bayes* Clasifiers," *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.*, vol. 7, no. 1, p. 32, 2021, doi: 10.24014/coreit.v7i1.12945.
- [8] R. Mahendrajaya, G. A. Buntoro, and M. B. Setyawan, "Analisis Sentimen Pengguna Gopay Menggunakan Metode Lexicon Based dan *Support Vector Machine*," *Komputek*, vol. 3, no. 2, p. 52, 2019, doi: 10.24269/jkt.v3i2.270.
- [9] R. Tineges, A. Triayudi, and I. D. Sholihati, "Analisis Sentimen Terhadap Layanan Indihome Berdasarkan Twitter Dengan Metode Klasifikasi *Support vector machine* (SVM)," *J. MEDIA Inform. BUDIDARMA*, vol. 4, no. 3, p. 650, Jul. 2020, doi: 10.30865/mib.v4i3.2181.
- [10] M. Akbar S, "Perbandingan Algoritme *Naive bayes* Classifier Dan K-Nearest Neighbors Pada Prediksi Pergerakan Mata Uang Dollar Amerika (Usd) Terhadap Harga Emas," Universitas Islam Indonesia Yogyakarta, 2020.
- [11] W. C. Widyaningtyas, S. Al Faraby, and Adiwijaya, "Klasifikasi Sentiment Analysis pada Review Film Berbahasa Inggris dengan Menggunakan Metode Doc2Vec dan *Support vector machine* (SVM) Sentiment Analysis Classification of Movie Review in English Language using Doc2Vec and *Support vector machine* (SVM)," *e-Proceeding Eng.*, vol. 5, no. 1, pp. 1570–1578, 2018.
- [12] A. S. Ritonga and E. S. Purwaningsih, "Penerapan Metode *Support vector machine* (SVM) dalam Klasifikasi Kualitas Pengelasan SMAW (Shield Metal Arc Welding)," *J. Ilm. Educ.*, vol. 5, no. 1, pp. 17–25, 2018.
- [13] R. G. Rafsanjani, N. Hidayat, and R. K. Dewi, "Diagnosis Penyakit Hati Menggunakan Metode *Naive Bayes* Dan Certainty Factor," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 11, pp. 4478–4482, 2018.
- [14] N. T. Putri, I. D. Wijaya, and A. T. H. Retno, "Analisis Sentimen Opini Masyarakat Terhadap Pembangunan Infrastruktur Kota Malang Melalui Twiter Dengan Menggunakan Metode *Support Vector Machine*," *Semin. Inform. Apl. Polinema*, pp. 118–123, 2020.
- [15] H. Hozairi, A. Anwari, and S. Alim, "Implementasi Orange Data Mining Untuk Klasifikasi Kelulusan Mahasiswa Dengan Model K-Nearest Neighbor, Decision Tree Serta *Naive Bayes*," *Netw. Eng. Res. Oper.*, vol. 6, no. 2, p. 133, 2021, doi: 10.21107/nero.v6i2.237.