

Analisis Sentimen Kepuasan Pengguna Aplikasi BCA Mobile menggunakan Metode *Naïve Bayes* dan *Support Vector Machine* (SVM)

July Aldren Marpaung¹, Mariza Devega², Yuhelmi³

^{1,2,3}Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Lancang Kuning

^{1,2,3}Jl. Yos Sudarso KM. 8 Rumbai, Pekanbaru, Riau, telp. 0811 753 2015

e-mail: aldrenmarpaung70@gmail.com, marizadevega@unilak.ac.id,
yuhelmi@unilak.ac.id

Abstrak

Aplikasi *BCA Mobile* dirilis pada tahun 2011 yang dikelola oleh PT Bank Central Asia Tbk. Pada awalnya aplikasi *BCA Mobile* hanya dapat digunakan oleh pengguna *BlackBerry*, namun pada tahun 2012 aplikasi *BCA Mobile* dapat digunakan juga oleh pengguna *Android* dan *iOS*. Namun seiring waktu pengguna melaporkan bahwa ketika mereka mencoba untuk melakukan transaksi melalui aplikasi *BCA Mobile*, mereka mengalami kendala yang menunjukkan bahwa transaksi tidak dapat diproses dan pesan *error*. Dari permasalahan yang ada, maka dilakukan penelitian yang bertujuan untuk mengetahui tanggapan pengguna berdasarkan ulasan di *google play store* terhadap *BCA mobile*. Pada penelitian ini menggunakan algoritma *Naïve Bayes* dan *Support Vector Machine* (SVM), dimana data yang di scraping berjumlah 7.000 data ulasan. Setelah melalui tahapan *Pre – Processing* dan pelabelan data maka di dapatkan hasil data *valid* berjumlah 788 data ulasan. Data yang di dapat akan dibagi menjadi data *Training* 80% dan data *Testing* 20%. Pada proses klasifikasi algoritma *Naïve Bayes* menghasilkan akurasi 71.0% dan algoritma *Support Vector Machine* (SVM) 83.2%. Jadi dari kedua metode yang di gunakan dalam proses analisis sentimen yang memiliki akurasi lebih akurat adalah algoritma *Support Vector Machine* (SVM) dengan akurasi 83.2%.

Kata kunci: Analisis Sentimen, Klasifikasi, *Scraping* Data Ulasan, SVM

Abstract

The *BCA Mobile* app was released in 2011 and is managed by PT Bank Central Asia Tbk. Initially, the *BCA mobile* app was only available for *BlackBerry* users, but in 2012 it was also available for *Android* and *iOS* users. However, as time passed, users that when they attempted to make transactions through the *bca mobile* app, they experienced an impediment that indicated that the transaction could not be processed and an error message. In this study, *Naïve Bayes* and *Support Vector Machine* (SVM) algorithms are used, where data is scraped out of 7,000 review data. The data can be divided into 80% *Training* data and 20% *Testing* data. In the algorithm classification process *Naïve Bayes* produced an accuracy of 71.0% and the *Support Vector Machine* (SVM) algorytm of 83.2%. So of the two methods used in the sentiment analysis process that possessed the more accurate is the support vector machine (SVM) with an accurate 83.2%.

Keywords: Sentiment Analysis, Classification, *Scraping* Review Data, SVM

1. PENDAHULUAN

Perkembangan teknologi saat ini mengalami pertumbuhan yang cepat, termasuk penggunaan aplikasi di sektor perbankan yang semakin meluas untuk melakukan transaksi secara mobile tanpa harus pergi ke bank secara fisik.

M-banking merupakan suatu teknologi platform yang disediakan oleh lembaga keuangan atau bank, yang memungkinkan penggunaanya untuk melakukan berbagai

transaksi keuangan, seperti memeriksa saldo, mentransfer uang, melakukan transaksi, dan membayar tagihan, dari jarak jauh kapan pun dan di mana pun dengan menggunakan perangkat seluler seperti ponsel atau asisten digital pribadi. M-banking merupakan perkembangan dari perbankan daring yang memungkinkan pelanggan untuk melaksanakan seluruh transaksi perbankan melalui platform mobile.

Mobile banking adalah suatu layanan terbaru yang dipersembahkan oleh bank, yang memungkinkan pengguna untuk melakukan transaksi perbankan melalui *smartphone*. Ini adalah fasilitas perbankan yang menggunakan perangkat komunikasi bergerak seperti ponsel, dengan menyediakan aplikasi unggulan untuk bertransaksi perbankan secara praktis.

Pengguna melaporkan bahwa ketika mereka mencoba untuk melakukan transaksi melalui aplikasi *BCA Mobile*, mereka menerima pesan *error* dengan kode *error* 200 dan kode *error* 205. Pesan tersebut menunjukkan bahwa transaksi tidak dapat diproses pada saat itu, dan disarankan untuk mencoba beberapa saat lagi. Selain itu, beberapa nasabah juga diminta untuk melakukan registrasi ulang pada aplikasi *BCA Mobile*. Dengan adanya permasalahan *error* tersebut apakah masyarakat tetap merasa nyaman dan tetap bertahan menggunakan *BCA mobile* atau tidak. Dari permasalahan yang diangkat, maka peneliti tertarik melakukan penelitian untuk mengetahui tanggapan masyarakat berdasarkan sentimen di *play store* terhadap *BCA mobile*.



Gambar 1. Error pada Aplikasi BCA Mobile

Hal ini bertujuan untuk menilai ulasan aplikasi *BCA Mobile* di *Play Store*, mengingat *Bank BCA* adalah salah satu bank terbesar di Asia dengan banyak pengguna di Indonesia. Penelitian ini dilakukan untuk menilai kualitas pelayanan *BCA Mobile* dalam hal cek saldo, transfer dana, dan pembayaran tagihan.

2. METODE PENELITIAN

Metode Penelitian yang digunakan adalah metode Naïve Bayes dan Support Vector Machine (SVM).

2.1. Naïve Bayes

Algoritma Naïve Bayes adalah salah satu teknik klasifikasi yang didasarkan pada metode probabilitas dan statistik yang diperkenalkan oleh ilmuwan Inggris Thomas Bayes.

Metode ini melibatkan perkiraan peluang di masa depan berdasarkan pengalaman di masa lalu, yang dikenal sebagai Teorema Bayes. Teorema ini digabungkan dengan pendekatan "naïve" yang mengasumsikan bahwa setiap atribut dalam data set adalah

independen satu sama lain. Dalam penerapannya, algoritma *Naïve Bayes Classifier* mengasumsikan bahwa tidak ada hubungan langsung antara satu atribut dengan atribut lainnya, meskipun dalam kenyataannya atribut tersebut mungkin saling terkait. Salah satu keunggulan dari algoritma *Naïve Bayes Classifier* adalah kemudahannya dalam penggunaan karena memiliki proses perhitungan yang singkat, digunakan untuk mengklasifikasikan atribut yang tidak relevan. (Sigid Widodo et al., 2023).

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

Keterangan :

X : Data dengan *Class* yang belum diketahui

H : Hipotesis data merupakan suatu *class* spesifik

$P(H|X)$: Probabilitas hipotesis H berdasar kondisi X (*posteriori probabilitas*)

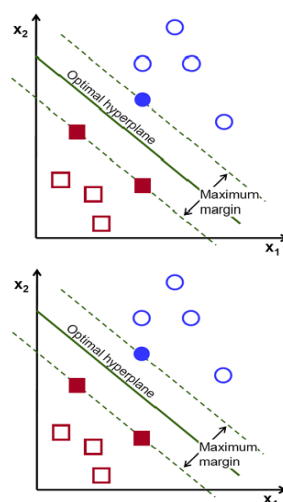
$P(H)$: Probabilitas hipotesis H (*prior probabilitas*)

$P(X|H)$: Probabilitas X berdasarkan kondisi pada hipotesis H

$P(X)$: Probabilitas X.

2.2. Support Vector Machine

Support Vector Machine (SVM) adalah sebuah metode yang umum digunakan untuk melakukan klasifikasi pada data teks dengan tingkat akurasi yang lebih tinggi. SVM merupakan salah satu teknik dalam pembelajaran berbimbing yang digunakan untuk proses klasifikasi. Dalam model klasifikasi, SVM memiliki konsep yang lebih kuat dan terdefinisi secara matematis. SVM bertujuan untuk menemukan hyperplane optimal dengan cara memaksimalkan jarak antara kelas-kelas data. Hyperplane merupakan fungsi yang digunakan untuk memisahkan antara kelas-kelas tersebut. Dalam dimensi dua, hyperplane disebut sebagai garis, sedangkan dalam dimensi tiga disebut sebagai bidang. Lebih lanjut, dalam klasifikasi data dengan dimensi yang lebih tinggi disebut sebagai hyperplane. (Pratiwi et al., 2021).



Gambar 2. Konsep SVM Mencari Hyperlane Terbaik

3. HASIL DAN PEMBAHASAN

Berikut akan di jelaskan hasil dan pembahasan dalam jurnal penelitian ini.

3.1. Pengambilan Data

Pengambilan data pada penelitian ini dilakukan dengan cara *Scraping* data ulasan aplikasi BCA Mobile yang berjumlah 7.000 ulasan pengguna, yang berada pada *Google Play Store* menggunakan *Goole Colab* dengan bahasa pemrograman *Python*. Selanjutnya mengimport paket yang diperlukan seperti *import pandas as pd* dan *import numpy as np* dimana fungsi nya adalah untuk menghubungkan *library* yang akan di gunakan dalam melakukan *Scraping* data.

3.2. Pre – Processing

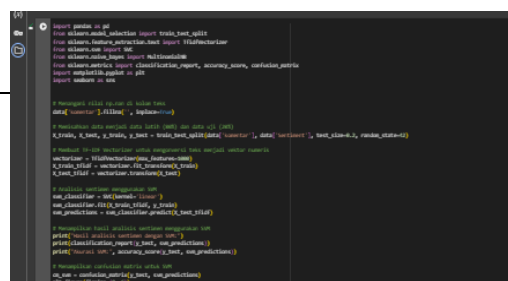
Pada *Pre-Processing* terdapat lima tahapan proses yang akan di lakukan sebbagai berikut :

- a. *Data Cleaning*
Pada gambar dibawah ini menampilkan peroses dalam melakukan *Data Cleaning* yang bertujuan untuk membersihkan data dan nanti nya akan di lakukan tahap *Case Folding*.
- b. *Case Folding*
Setelah di lakukan *Data Cleaning* kemudian akan dilakukan *Case Folding*, tahapan ini bertujuan untuk menormalisasi kata kedalam bentuk huruf kecil (*Lower Text*) tetapi masih dalam bentuk kalimat yang sama.
- c. *Normalisasi Teks*
Setelah di lakukan *Case Folding* kemudian akan dilakukan *Normalisasi Teks*, tahapan ini bertujuan untuk menormalisasi kata singkat kedalam bentuk kata normal, normalisasi teks bertujuan untuk mempermudah dalam melakukan analisis.
- d. *Tokenization*
Setelah di lakukan *Normalisasi Teks*, data yang di dapat sebelumnya akan di lakukan tahapan *Tokenization*. Hal ini bertujuan untuk memotong kata sehingga menjadi bentuk tunggal, guna memudahkan dalam mengidentifikasi suatu kalimat.
- e. *Stopword Removal*
Setelah di lakukan *Tokenization*, data yang di dapat sebelumnya akan di lakukan tahapan *Stopword Removal* yaitu untuk menghapus kata sambung yang tidak memiliki makna, contoh kata yang akan di hilangkan adalah “yang”, “dan”, “di”, “dari”, dll.
- f. *Stemming*
Setelah semua tahapan dilakukan kemudian akan di lakukan proses *Stemming*, hal ini bertujuan untuk menghilangkan kata imbuhan baik yang berada di awal maupun di akhir guna mendapatkan kata dasar nya.

3.3. Pelabelan Data

Setelah semua proses *Pre – Processing* di lakukan, maka tahap selanjutnya adalah pelabelan data, pelabelan ini di lakukan pada setiap data ulasan yang telah di peroleh sebelumnya. Hal ini di lakukan untuk mengetahui apakah ulasan tersebut dapat di katakan sentimen *Positif* atau *Negatif*.

3.4. Klasifikasi Algoritma Naïve Bayes Dan Support Vector Machine (SVM)

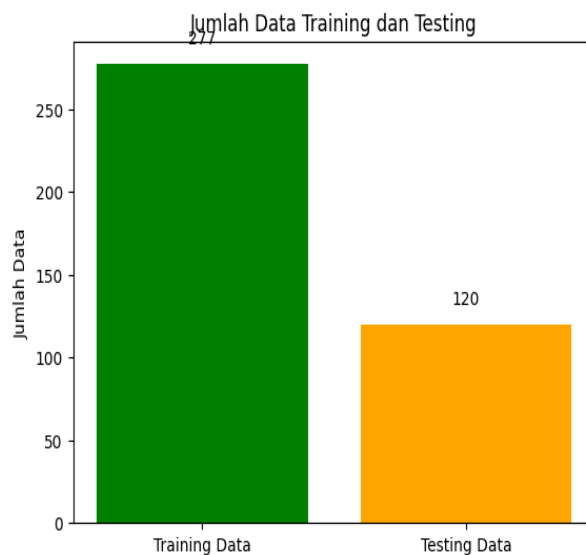


Gambar 3. Proses Klasifikasi *Naïve Bayes* dan *Support Vector Machine* (SVM)

Berikut akan di buat pembagian data (70% dan 30%, 80% dan 20%, 90% dan 10%).

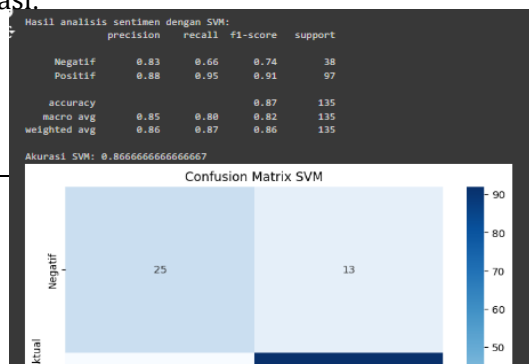
a. Pembagian 70% data Testing dan 30% data Training

Data yang akan di klasifikasikan di bagi menjadi data Training dan data Testing. Data sebanyak 448 akan di kelompokkan menjadi 70% data Training dan 30% data Testing. Maka di peroleh 277 data Training dan 120 data Testing.



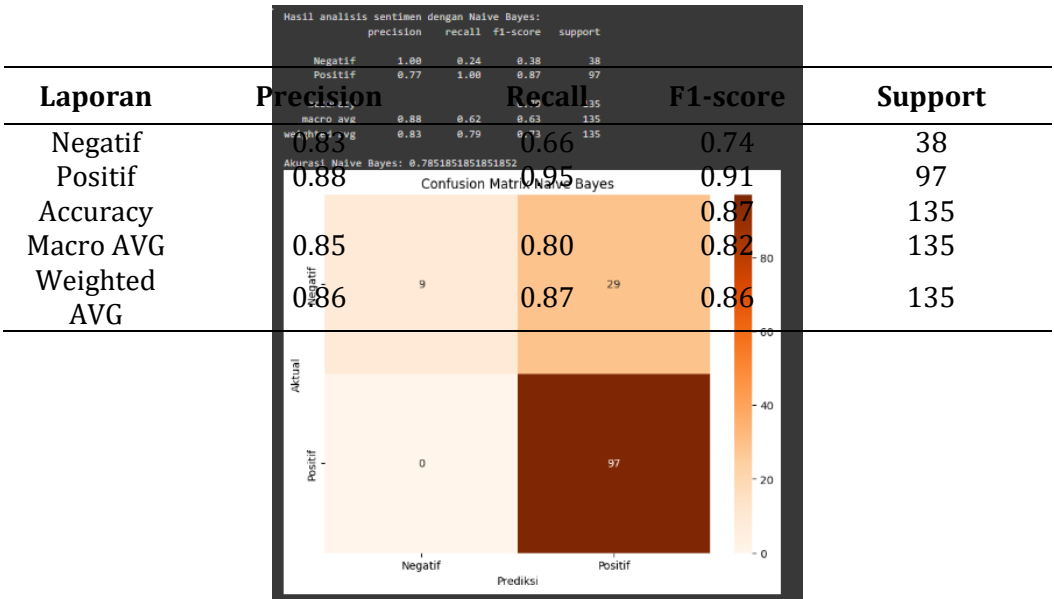
Gambar 4. Pembagian Data 70% dan 30%

Gambar diatas merupakan pembagian data Training dan Testing, yang nantinya akan di lakukan klasifikasi.



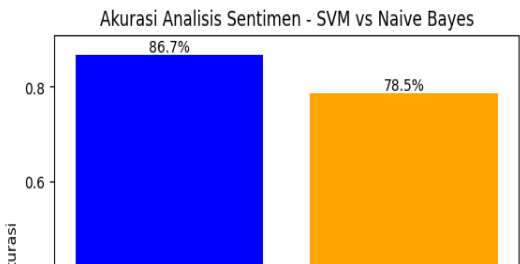
Gambar 5. Confusion Matrix SVM

TABEL 1. Tabel Hasil Confusion Matrix SVM



Gambar 6. Confusion Matrix Naïve Bayes

TABEL 2. Tabel Hasil Confusion Matrix Naïve Bayes



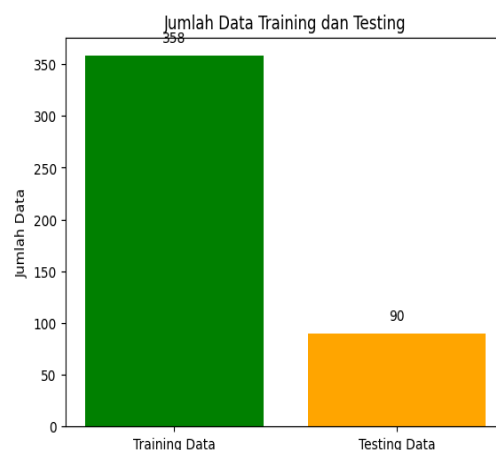
Laporan	Precision	Recall	F1-score	Support
Negatif	1.00	0.24	0.38	38
Positif	0.77	1.00	0.87	97
Accuracy			0.79	135
Macro AVG	0.88	0.62	0.63	135
Weighted AVG	0.83	0.79	0.73	135

Gambar 7. Hasil Akurasi 70% dan 30%

Gambar diatas merupakan hasil akurasi yang di peroleh dari pembagian data 70% dan 30%, dimana *Support Vector Machine* (SVM) merupakan tingkat akurasi tertinggi dengan *persentase* 86.7%, dan *Naïve Bayes* 78.5%.

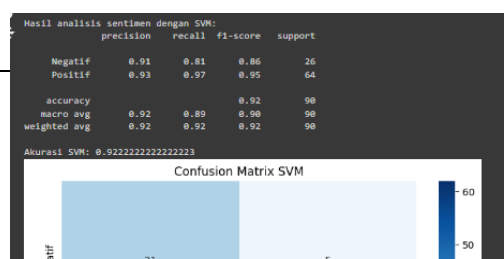
b. Pembagian 80% data Training dan 20% data Testing

Data yang akan di klasifikasikan di bagi menjadi data Training dan data Testing. Data sebanyak 448 akan di kelompokkan menjadi 80% data Training dan 20% data Testing. Maka di peroleh 358 data Training dan 90 data Testing.



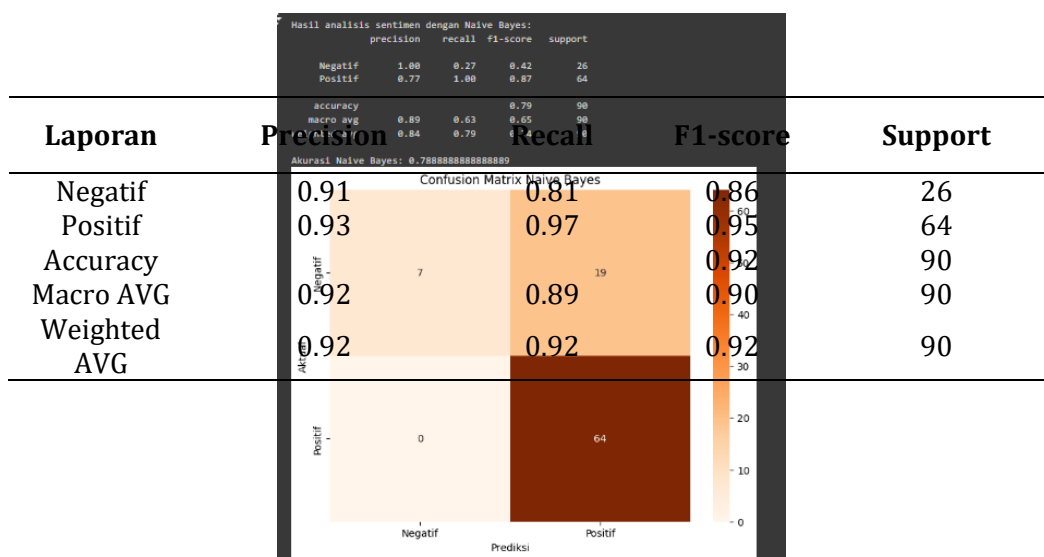
Gambar 8. Pembagian Data 80% dan 20%

Gambar diatas merupakan pembagian data Training dan Testing, yang nantinya akan di lakukan klasifikasi.



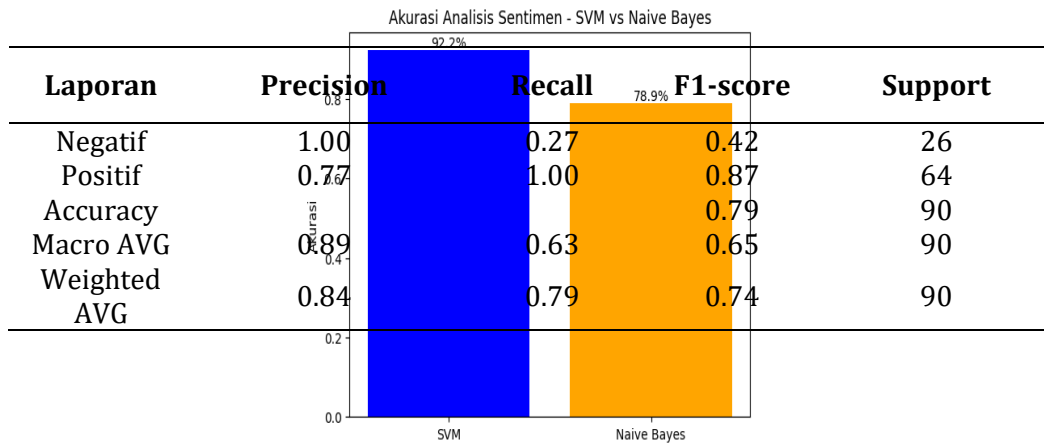
Gambar 9. Confusion Matrix SVM

TABEL 3. Tabel Hasil Confusion Matrix SVM



Gambar 10. Confusion Matrix Naïve Bayes

TABEL 4. Tabel Hasil Confusion Matrix Naïve Bayes

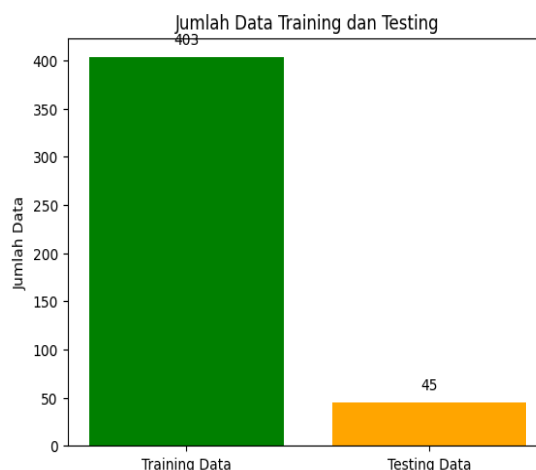


Gambar 11. Hasil Akurasi 80% dan 20%

Gambar diatas merupakan hasil akurasi yang di peroleh dari pembagian data 80% dan 20%, dimana *Support Vector Machine* (SVM) merupakan tingkat akurasi tertinggi dengan *persentase* 92.2%, dan *Naive Bayes* 78.9%.

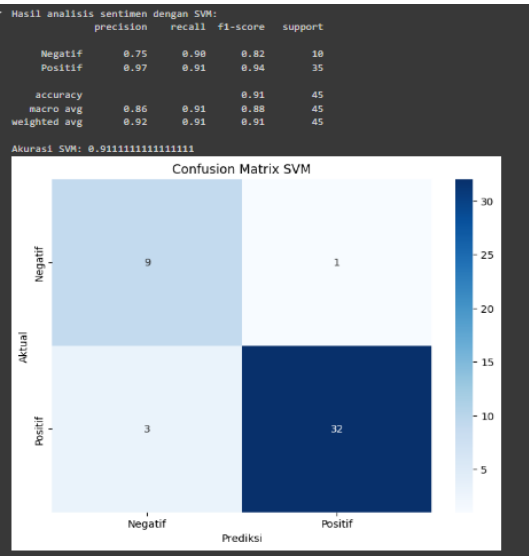
c. Pembagian 90% data Training dan 10% data Testing

Data yang akan di klasifikasikan di bagi menjadi data Training dan data Testing. Data sebanyak 448 akan di kelompokkan menjadi 90% data Training dan 10% data Testing. Maka di proleh 403 data Training dan 45 data Testing.



Gambar 12. Pembagian Data 90% dan 10%

Gambar diatas merupakan pembagian data Training dan Testing, yang nantinya akan di lakukan klasifikasi.



Gambar 13. Confusion Matrix SVM

TABEL 5. Tabel Hasil Confusion Matrix SVM

Laporan

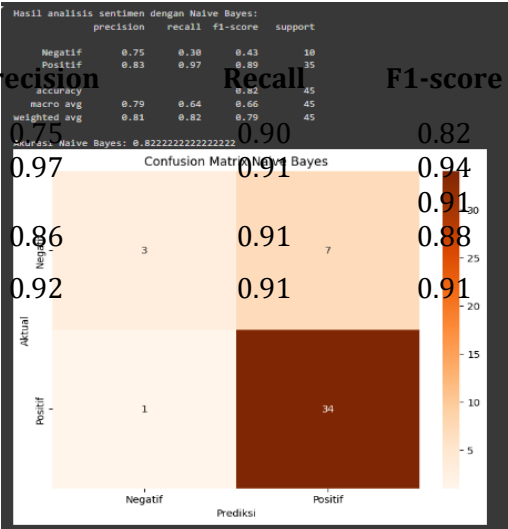
Negatif
Positif
Accuracy
Macro AVG
Weighted
AVG

Precision

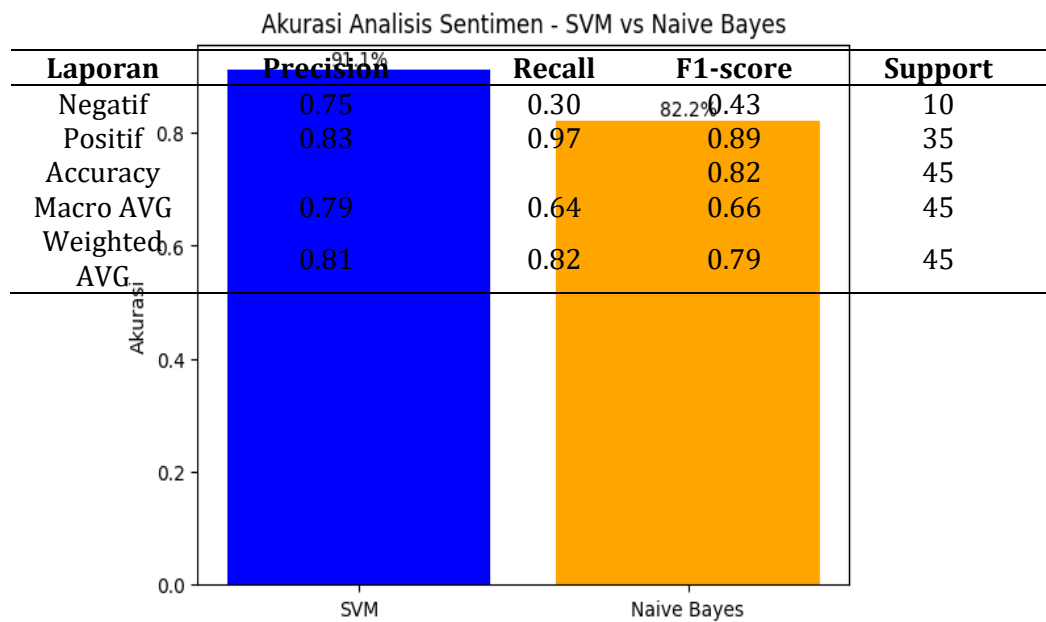
Recall

F1-score

Support



Gambar 14. Confusion Matrix Naïve Bayes
Tabel 6. Tabel Hasil Confusion Matrix SVM



Gambar 15. Hasil Akurasi 90% dan 10%

Gambar diatas merupakan hasil akurasi yang di peroleh dari pembagian data 90% dan 10%, dimana *Support Vector Machine* (SVM) merupakan tingkat akurasi tertinggi dengan *persentase* 91.1%, dan *Naive Bayes* 82.2%.

d. Pembahasan

Berdasarkan confusion matriks yang dihasilkan maka dapat dilihat hasil dari nilai akurasi yang didapatkan.

Tabel 7. Tabel Hasil Pembahasan

Hasil	Naïve Bayes			Support Vector Machine (SVM)		
Data Training	70%	80%	90%	70%	80%	90%
Data Testing	30%	20%	10%	30%	20%	10%
Akurasi	78.5%	78.9%	82.2 %	86.7%	92.2%	91.1%

Pada Tabel diatas menunjukan beberapa hasil akurasi yang di proleh dari pembagian (70% dan 30%, 80% dan 20%, 90% dan 10%) dengan menggunakan 7.000 data ulasan

yang di *Scrapping* dari *Google Playstore*. Setelah melalui tahapan tahap *Preprocessing*, data awal yang berjumlah 7000 kini telah berkurang menjadi 4484, kemudian di lakukan pelabelan data, data yang bernilai netral akan di *drop* atau di hapus, dan menghasilkan data *valid* yang berjumlah 448. Dari beberapa tahapan diatas maka dapat di simpulkan dalam penelitian ini metode *Support Vector Machine (SVM)* lebih unggul dengan tingkat akurasi lebih tinggi dari pada metode *Naïve Bayes*.

4. KESIMPULAN

Algoritma Naïve Bayes dan algoritma *Support Vector Machine (SVM)* dapat digunakan dalam menganalisis ulasan aplikasi BCA Mobile dengan menghasilkan 2 klasifikasi diantaranya kelas positif dan kelas negatif. Berdasarkan hasil penelitian yang telah dilakukan dapat di simpulkan dalam penelitian ini menggunakan metode *Support Vector Machine (SVM)* lebih akurat dengan akurasi yang di hasilkan lebih tinggi, dibandingkan dengan menggunakan metode *Naïve Bayes*.

DAFTAR PUSTAKA

- [1] <https://sharingvision.com/insight/survey-2022-sharing-vision-mobile-banking-menjadi-primadona/>
 - [2] [https://jabar.tribunnews.com/2023/05/14/layanan-mobile banking-bca-error-trending-di twitter-begini-kata-admin-halobca](https://jabar.tribunnews.com/2023/05/14/layanan-mobile-banking-bca-error-trending-di-twitter-begini-kata-admin-halobca)
 - [3] https://www.detik.com/jatim/bisnis/d-6719790/mobile-banking-bca-eror-nasabah-tak-bisa-cek-saldo-hingga-transfer?utm_campaign=detikcomsocmed&utm_medium=oa&utm_source=twitter&utm_content=detikjatim
 - [4] Ahmad Dzulhijjah, D., Sanjaya, H., Said Wahyudi Hidayat, A., Yulistia Alwanda, A., & Utami, E. (2023). Perbandingan Metode Random Forest dan KNN pada Analisis Sentimen Twitter. *Smart Comp: Jurnalnya Orang Pintar Komputer*, 12(3), 767–772. <https://doi.org/10.30591/smartcomp.v12i3.5106>
 - [5] Bañez-Coronel, M., Ayhan, F., Tarabochia, A. D., Zu, T., Perez, B. A., Tusi, S. K., Pletnikova, O., Borchelt, D. R., Ross, C. A., Margolis, R. L., Yachnis, A. T., Troncoso, J. C., Ranum, L. P. W., Roos, R. A. C., Perez, M., Jin, W., Le, D., Carlozzi, N., Dayalu, P., ... Frank, S. (2018).
 - [6] Febriana, S. (2023). *Skripsi Analisis Sentimen Pengguna Mobile Banking Bca Pada Media Sosial Twitter Menggunakan Naïve Bayes Dan Support Vector Machine (Svm)*. 11190930000034, 1–100. <https://repository.uinjkt.ac.id/dspace/handle/123456789/75077>
 - [7] Fitriyanti, F., Hum, M., Kn, M., Wulan Rachmanti, V., & Mulya Amalia, N. (2022). *Pengaruh Inovasi Digital (Mobile Banking) Bank Bca Dalam Industri Perbankan Pada Tahun 2020-2022 Terhadap Kepuasan Nasabah*. June.
 - [8] Gunawan, F., Fauzi, M. A., & Adikara, P. P. (2017). Analisis Sentimen Pada Ulasan Aplikasi Mobile Menggunakan Naive Bayes dan Normalisasi Kata Berbasis Levenshtein Distance (Studi Kasus Aplikasi BCA Mobile). *Systemic: Information System and Informatics Journal*, 3(2), 1–6. <https://doi.org/10.29080/systemic.v3i2.234>
 - [9] Jogianto, M., Romney, M., & Jogiyanto, M. (2015). *BAB 2 TINJAUAN PUSTAKA DAN DASAR TEORI 2.1. Pengertian sistem*. 2010, 5–19.
 - [10] Pratiwi, R. W., H, S. F., Dairoh, D., Afidah, D. I., A, Q. R., & F, A. G. (2021). Analisis Sentimen Pada Review Skincare Female Daily Menggunakan Metode Support Vector
-

- Machine (SVM). *Journal of Informatics, Information System, Software Engineering and Applications (INISTA)*, 4(1), 40–46. <https://doi.org/10.20895/inista.v4i1.387>
- [11] Ritonga, A. S., & Purwaningsih, E. S. (2018). Penerapan Metode Support Vector Machine (SVM) Dalam Klasifikasi Kualitas Pengelasan Smaw (Shield Metal Arc Welding). *Ilmiah Edutic*, 5(1), 17–25.
- [12] Rizky Pratama, M., Ramadhan, Y. R., & Komara, M. A. (2023). Analisis Sentimen BRLmo dan BCA Mobile Menggunakan Support Vector Machine dan Lexicon Based. *Jutisi: Jurnal Ilmiah Teknik Informatika Dan Sistem Informasi*, 12(3), 1439–1450.
- [13] Rozi, F., Sukmana, F., & Adani, M. N. (2021). Pengelompokkan Judul Buku dengan Menggunakan Algoritma K-Nearest Neighbor (K-NN) dan Term Frequency – Inverse Document Frequency (TF-IDF). *JIMP: Jurnal Informatika Merdeka Pasuruan*, 6(3), 1–5.
- [14] Ruslim, K. I., Adikara, P. P., & Indriati. (2019). Analisis Sentimen Pada Ulasan Aplikasi Mobile Banking Menggunakan Metode Support Vector Machine dan Lexicon Based Features. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 3(7), 6694–6702.
- [15] Sari, W. F., Rahim, R., & Adrianto, F. (2023). Analisis Sentiment Review Pengguna Bca Mobile Menggunakan Teks Mining. *Cakrawala*, 6, 981–987.
- [16] Sigid Widodo, A. Z. M., Pandu Kusuma, A., & Dwi Puspitasari, W. (2023). Analisis Algoritma Naive Bayes Classifier (Nbc) Pada Klasifikasi Tingkat Minat Barang Di Toko Violet Cell. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(1), 87–94. <https://doi.org/10.36040/jati.v7i1.5692>
- [17] Sipayung, E. M., Maharani, H., & Zefanya, I. (2016). Perancangan Sistem Analisis Sentimen Komentar Pelanggan Menggunakan Metode Naive Bayes Classifier. *Jurnal Sistem Informasi (JSI)*, 8(1), 2355–4614. <http://ejournal.unsri.ac.id/index.php/jsi/index>
- [18] Sunarko, G. S., Wasino, & Sutrisno, T. (2023). Klasterisasi Sentimen Ulasan Pengguna Aplikasi Bca Mobile Pada Platform Google Play Store Dengan Algoritma K-Means Clustering. *Jurnal Ilmu Komputer Dan Sistem Informasi*, 11(1). <https://doi.org/10.24912/jiksi.v11i1.24145>
- [19] Timur, S., Aryati, R., & Sibaroni, Y. (2023). Analisis Perbandingan Model Kernel Support Vector Machine dalam Analisis Sentimen Opini Pengguna Bank BCA di Twitter. 2023.
- [20] Widayat, W. (2021). Analisis Sentimen Movie Review menggunakan Word2Vec dan metode LSTM Deep Learning. *Jurnal Media Informatika Budidarma*, 5(3), 1018. <https://doi.org/10.30865/mib.v5i3.3111>
- [21] Yunanto, I., & Yulianto, S. (2022). Twitter Sentiment Analysis Pedulilindungi Application Using Naïve Bayes and Support Vector Machine. *Jurnal Teknik Informatika (Jutif)*, 3(4), 807–814. <https://doi.org/10.20884/1.jutif.2022.3.4.292>

