



@is The Best :

Accounting Information Systems and
Information Technology Business Enterprise
Volume 6, Nomor 2 (2021) Hal. 91-105
ISSN: 2252-9853 (Print) | ISSN: 2656-808X (Online)
<https://ojs.unikom.ac.id/index.php/aisthebest/index>

Terakreditasi Peringkat 4, SK No.: 28/E/KPT/2019

DOI: <https://doi.org/10.34010/aisthebest.v6i2.4919>

Analisis Sentimen Kepuasan Pengguna Aplikasi WhatsApp Menggunakan Algoritma *Naïve Bayes* Dan *Support Vector Machine*

Acep Saepulrohman¹, Sudin Saepudin², Dudih Gustian³

^{1,2,3} Program Studi Sistem Informasi, Universitas Nusa Putra, Sukabumi, Jawa Barat, Indonesia

Email: acep.saepulrohman_si17@nusaputra.ac.id

ABSTRACT

Information and communication technology was currently growing rapidly, one of which was chat or instant messaging applications such as WhatsApp, line and telegram. In October 2020, the majority of instant messaging app users was WhatsApp app users, with a total of 2 billion users. Even though the WhatsApp application was in the top ranking and got the highest score, but this could not be used as a measured of satisfaction because there were still negative views on the WhatsApp application, some users assumed that WhatsApp often had errors when used, then other problems that arise such as the network that the user used was unstable. To conduct an analysis of this, a sentiment analysis approached was needed to categorize user comments into positive or negative. This studied used the naïve bayes algorithm with support vector machine in analyzing positive and negative comments on the satisfaction of users of the WhatsApp application on the google played store. From the results of tests carried out on 1500 user commented data, the evaluation of the model used 10 folded cross validation shows that the leveled of accuracy for WhatsApp application user satisfaction based on the naïve bayes algorithm was 70.40% and support vector machine was 77.00%, while the auc valued naïve bayes was 0.585 and support vector machine was 0.876. From these results, the svm algorithm could be used for researched with the same data characteristics.

Keywords: sentiment analysis, naïve bayes, support vector machine, WhatsApp, google played store reviews

ABSTRAK

Teknologi informasi dan komunikasi saat ini sangat berkembang pesat, salah satunya Aplikasi Chat atau pesan instan seperti WhatsApp, Line dan Telegram. Pada bulan Oktober 2020, mayoritas pengguna aplikasi pesan instan adalah pengguna aplikasi WhatsApp, dengan total 2 miliar pengguna. Sekalipun aplikasi WhatsApp tersebut masuk dalam peringkat teratas dan mendapat skor tertinggi, akan tetapi hal tersebut tidak dapat dijadikan tolak ukur kepuasan karena masih terdapat pandangan yang negatif terhadap aplikasi WhatsApp, sebagian pengguna menganggap bahwa WhatsApp seringkali eror pada saat digunakan, kemudian masalah lain yang muncul seperti jaringan yang digunakan pengguna tidak stabil. Untuk melakukan analisis mengenai hal tersebut diperlukan pendekatan analisis sentimen guna mengkategorikan komentar pengguna menjadi positif atau negatif. Penelitian ini menggunakan algoritma *Naïve Bayes* dengan *Support Vector Machine* dalam menganalisa komentar positif dan negatif terhadap kepuasan pengguna aplikasi WhatsApp di *Google Play Store*. Dari hasil pengujian yang dilakukan terhadap 1500 data komentar pengguna, evaluasi model menggunakan *10 Fold Cross Validation* menunjukkan bahwa tingkat keakuratan untuk kepuasan pengguna aplikasi WhatsApp berdasarkan algoritma *Naïve Bayes* adalah sebesar 70,40% dan *Support Vector Machine* sebesar 77,00%, sedangkan nilai AUC *Naïve Bayes* sebesar 0,585 dan *Support Vector Machine* adalah 0,876. Dari hasil tersebut algoritma *Support Vector Machine* dapat digunakan untuk penelitian dengan karakteristik data yang sama.

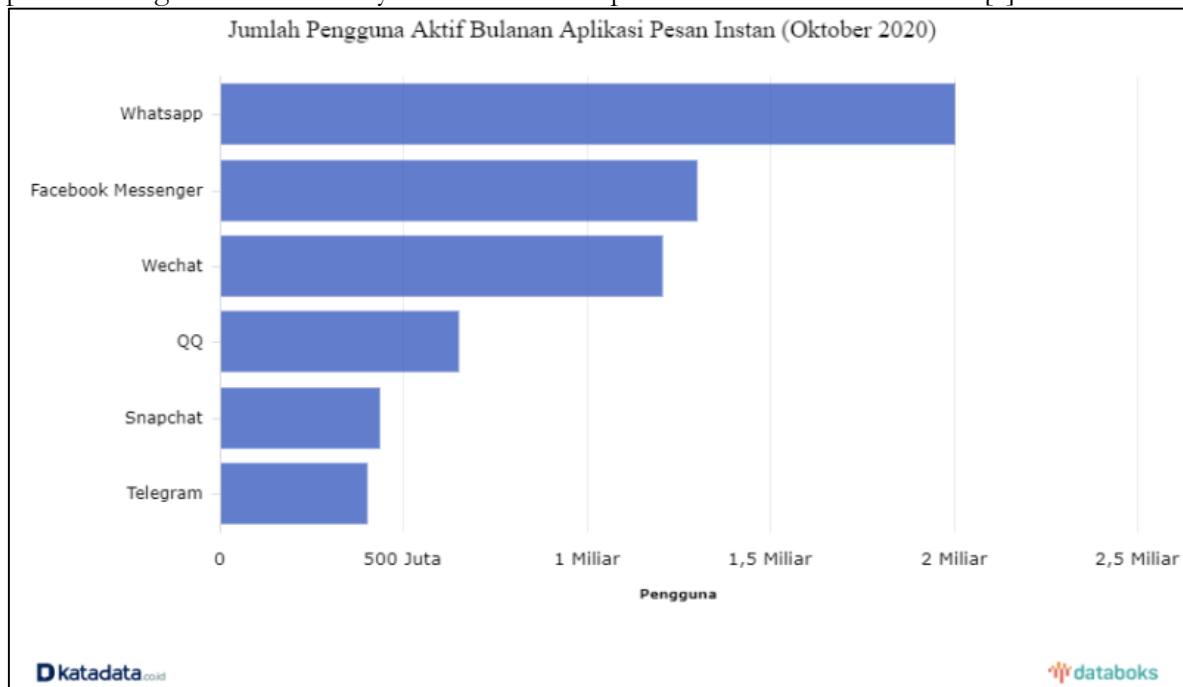
Kata Kunci: Analisis Sentimen, Naïve Bayes, Support Vector Machine, WhatsApp, Ulasan Google Play Store.

Pendahuluan

Perkembangan teknologi saat ini sangat pesat, termasuk perkembangan teknologi informasi dan komunikasi. Teknologi telekomunikasi adalah contohnya. Perkembangan teknologi telekomunikasi telepon pintar (*smartphone*) juga sangat rumit dan berdampak sangat penting bagi masyarakat [1]. Indonesia merupakan Negara peringkat ke-6 dengan jumlah pengguna *smartphone* terbanyak di seluruh dunia, dengan jumlah pengguna sebanyak 73.155.000 atau 27,4% dari total

penduduk Indonesia yakni sebanyak 266.795.000 jiwa [2]. Pada era Revolusi Industri 4.0 juga percepatan teknologi yang semakin massif, peran teknologi informasi dalam berbagai bidang dalam suatu unit juga berkembang pesat. Teknologi semakin maju dan kemajuan tersebut harus dimanfaatkan sebaik mungkin. Tantangan yang dihadapi dalam revolusi industri 4.0 antara lain: 1) Masalah keamanan teknologi informasi, 2) Keandalan dan stabilitas mesin produksi, 3) Kurangnya keterampilan yang memadai, 4) Keengganan untuk berubah oleh pemangku kepentingan, 5) Kehilangan banyak pekerjaan karena perubahan menjadi otomatisasi [3].

Aplikasi chat yang saat ini sedang gencar berkembang di dalamnya termasuk WhatsApp. Aplikasi pesan instan tersebut memiliki kelebihan dan fungsi masing-masing, karena pada dasarnya aplikasi ini digunakan oleh masyarakat luas dan dapat diakses dari seluruh dunia [4].



Gambar 1. Jumlah Pengguna Aktif Bulanan Aplikasi Pesan Instan

Pada Gambar 1 diketahui bahwa jumlah pengguna *Global* Aplikasi pesan instan pada bulan Oktober 2020 terbanyak yaitu pada pengguna aplikasi WhatsApp dengan jumlah pengguna sejumlah 2 milyar [5]. Banyak jenis aplikasi *chatting* dengan berbagai fungsi. Namun, di Indonesia salah satu aplikasi yang paling populer adalah WhatsApp. Hal ini dikarenakan fungsi-fungsi dalam aplikasi WhatsApp memberikan banyak kemudahan bagi pengguna untuk melakukan percakapan privat maupun grup [6]. Peranan teknologi informasi tidak terlepas dari usaha manajemen perusahaan terkait untuk membangun usahanya agar dapat bersaing di kancah nasional [7]. Namun terdapat permasalahan yang cukup signifikan, sekalipun aplikasi WhatsApp tersebut masuk dalam peringkat teratas dan mendapat skor tertinggi. Akan tetapi hal tersebut tidak dapat dijadikan tolak ukur kepuasan karena masih terdapat pandangan yang negatif terhadap aplikasi WhatsApp, sebagian pengguna menganggap bahwa WhatsApp seringkali error pada saat digunakan. Selain hal tersebut, Kemudian masalah lainnya yakni muncul koneksi jaringan yang digunakan pengguna tidak stabil. Berdasarkan beberapa alasan di atas, maka diperlukan penelitian sentimen analisis yang mengkaji ulasan pengguna WhatsApp, baik berupa kepuasan maupun kekecewaan penggunanya [8]. Untuk meneliti dan menganalisis objektivitas hal tersebut maka diperlukan suatu metode dan analisis untuk mengklasifikasikan komentar pengguna kedalam beberapa kategori, dimana dalam

penelitian ini metode yang digunakan yaitu metode *Naïve Bayes* dan *Support Vector Machine*.

Analisis Sentimen adalah proses mengekstraksi, mengolah data maupun memahami suatu data tekstual secara otomatis atau biasa disebut analisis sentimen merupakan analisis yang biasanya mengolah data opini atau suatu pandangan publik yang nantinya akan menghasilkan kesimpulan dari data yang dianalisis tersebut [9]. Kesimpulan tersebut dapat berupa persentase sentimen positif, negatif ataupun netral.

Naive Bayes adalah teknik untuk memprediksi kemungkinan [10]. Metode ini merupakan metode yang cukup sederhana [11], cara kerja *Naive Bayes* adalah memprediksi kemungkinan kejadian di masa mendatang berdasarkan data yang tersedia sebelumnya yang dapat mengolah data dalam jumlah yang besar dengan tingkat akurasi nilai yang tinggi. Ada beberapa teknik yang diselesaikan dalam mengklasifikasikan dokumen teks, yaitu: Menentukan nilai kemungkinan kejadian setiap kata dalam klasifikasi dokumen, Menentukan nilai probabilitas kemunculan setiap kata pada setiap kategori dokumen, Menentukan kelas dokumen pengujian perhitungan berdasarkan proses pertama dan kedua [12]. *Support Vector Machine* adalah metode yang efisien dengan cara kerja mengkategorikan serta meregresi beberapa masalah linier maupun bukan linear. Kelebihan menggunakan aplikasi ini ialah memiliki dimensi yang relatif tinggi menggunakan fungsi kernel yang digunakan [13][14].

Penelitian yang telah dilakukan oleh Khoirul Abbi Rokhman, Berlilana, dan Primandani Arsi, dalam penelitiannya yang berjudul Perbandingan Metode *Support Vector Machine* Dan *Decision Tree* untuk Analisis Sentimen Review Komentar Pada Aplikasi *Transportasi Online*. Memperoleh hasil akurasi sebesar 90.20% untuk metode *Support Vector Machine* sedangkan 89.80% untuk metode *Decision Tree*. Dapat disimpulkan untuk metode *Support Vector Machine* nilai akurasinya lebih tinggi dibandingkan metode *Decision Tree* [15]. Selain itu berdasarkan penelitian Aputra, S. A., Didi Rosiyadi, Windu Gata, & Syepri Maulana Husain yang berjudul Analisis Sentimen *E Wallet* Pada Google Play store Menggunakan *Naive Bayes* Berbasis *Particle Swarm Optimization*, meneliti tentang ulasan sentimen positif dan negatif pengguna terhadap aplikasi OVO. Perolehan hasil penelitian yang dilakukan ialah Validasi silang pada NB dengan FS adalah 83.60 % untuk akurasi dan 0.801 untuk AUC Sedangkan NB tanpa FS adalah 82.30 % tingkat akurasi 0.780 untuk AUC [16]. Penelitian terkait dilakukan oleh Rian Tineges dkk, pada tahun 2020 dengan judul Analisis Sentimen Terhadap Layanan Indihome Berdasarkan Twitter Dengan Metode Klasifikasi (SVM), Masalah dari penelitian ini adalah sulitnya mengklasifikasikan tanggapan atau komentar positif dan negatif masyarakat terhadap layanan indihome, karena hal tersebut dibutuhkan sebuah penelitian mengenai metode atau cara untuk dapat merubah sampel menjadi berbagai informasi pengetahuan, salah satunya menggunakan metode analisis sentimen menggunakan SVM. Hasil yang didapat dari analisis tersebut ialah presisi 86%, akurasi 87%, pengingat 95% , nilai eror 13 % dengan nilai f1 90% [17].

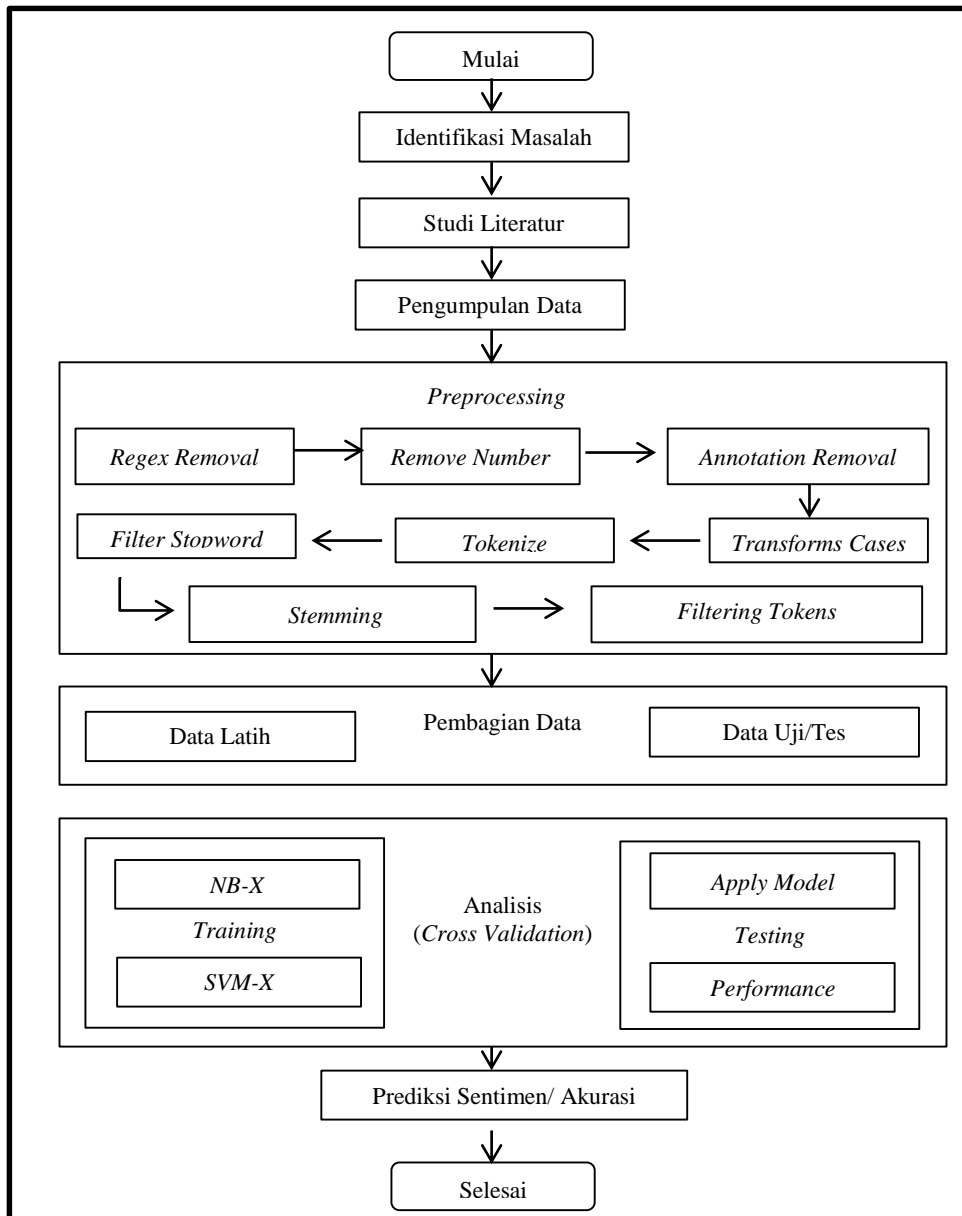
Penelitian selanjutnya oleh Umi Rofiqoh, dkk pada tahun 2017 dengan judul Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada *Twitter* Dengan Metode *Support Vector Machine* dan *Lexicon Based Features*, masalah dari penelitian ini adalah penurunan layanan dan berpindahnya konsumen ke penyedia jasa yang lain, sehingga dicari solusi untuk mengetahui kepuasan pelanggan melalui twitter. Untuk melakukan penelitian tersebut dianalisis menggunakan SVM berbasis *lexicon based features* yang diperbarui melalui fitur TF-IDF. Penelitian ini menghasilkan tingkat akurasi 79% bernilai degree 2, nilai konstanta belajar sebesar 0,0001 dan jumlah literasi yang dilakukan maksimum sejumlah 50 kali. Metode analisis tanpa bantuan lexicon based memiliki tingkat akurasi 84% [18].

Penelitian Dudih Gustian, dkk pada tahun 2019 dengan judul Sistem Pakar Dengan Algoritma *Naive Bayes* Untuk Prediksi Hasil Produksi Ayam Broiler Plasma (Studi Kasus: PT. Sekawan Sinar Surya), Penelitian ini menganalisis mengenai parameter yang berpengaruh terhadap hasil produksi para peternak dan perusahaan menggunakan algoritma NB. Penelitian ini menggunakan perangkat Weka. Yang menghasilkan tingkat akurasi sebesar 96,36 % dengan nilai ROC mendekati angka 1 yaitu sebesar 0,9995 [19].

Penelitian ini memberikan solusi yaitu berupa informasi pada pihak manajemen WhatsApp agar tetap konsisten dalam pelayanan pada pengguna sehingga tingkat kepercayaan pengguna tetap terjaga dan meningkatkan kualitas aplikasi demi tercapainya kepuasan pengguna WhatsApp.

Metode

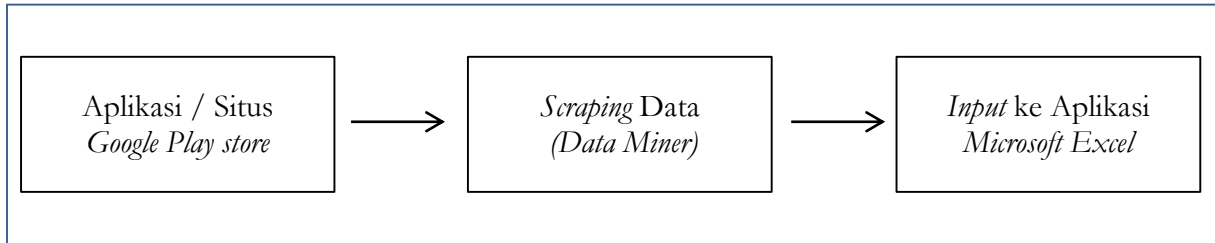
Untuk menganalisa akurasi terbaik pada ulasan aplikasi WhatsApp digunakan metode algoritma *Naïve Bayes* dan *Support Vector Machine*. Tahapan untuk menghasilkan akurasi yang unggul ialah terdapat pada gambar 2 berikut:



Gambar 2. Diagram Alir Metode Penelitian

Pengumpulan Data

Pengumpulan data dilakukan dengan dua tahap, tahap pertama adalah mengumpulkan data yang akan dijadikan data set dan tahap kedua membagi dataset tersebut menjadi dua data yaitu data latih/*training* dan data uji/*testing*. Data opini bersumber dari *extension browser chrome Dataminer* yang didapatkan dari ulasan pengguna WhatsApp di google play store.



Gambar 3. Proses Pengumpulan Ulasan Komentari

Pemrosesan Data

Persiapan data yang telah didapat atau *preprocessing* memiliki beberapa kegiatan didalamnya, diantaranya mengolah, membersihkan data agar siap diolah pada proses selanjutnya yaitu proses pemodelan, langkah-langkah *preprocessing* antara lain sebagai berikut:

1. Penghapusan Ekspresi Reguler (*Regex Removal*)
Proses *Regex Removal* yaitu proses menghapus ekspresi *regular* yang ada dalam ulasan komentar.
2. Penghapusan Angka (*Remove Number*)
Proses *Remove Number* merupakan proses menghapus angka yang ada dalam ulasan komentar.
3. Penghapusan tanda @ (*Annotation Removal*)
Proses *Annotation Removal* merupakan proses menghilangkan tanda *att* atau @ (*annotation*) yang ada dalam ulasan komentar
4. *Transform Cases*
Proses *Transform Cases* yang dilakukan yaitu mengubah ulasan komentar menjadi huruf kecil agar terdapat keseragaman huruf.
5. *Tokenizing*
Proses *Tokenizing* adalah cara untuk pemisahan teks menjadi potongan-potongan yang disebut token. Tokenisasi berfungsi untuk memisahkan baris kata dalam kalimat, paragraf atau halaman menjadi potongan kata tunggal.
6. *Filter Stopword*
Kata-kata biasa yang umumnya muncul dalam jumlah besar dan dianggap tidak memiliki makna. Tahap ini merupakan interaksi untuk memisahkan kata-kata biasa, misalnya “yang”, “di”, “ke”, *stopwords* sangat penting, khususnya *stopwords* digunakan untuk mengurangi jumlah kata yang harus disiapkan.
7. *Stemming*
Pada tahapan *stemming*, yaitu mengubah kata-kata yang memiliki imbuhan menjadi kata dasar, diharapkan untuk menghapus kata-kata dengan ejaan yang kurang tepat.
8. *Filtering Tokens*
Proses penghapusan kata-kata dengan panjang huruf tertentu.
9. Pelabelan
Pada tahap ini penulis melakukan pelabelan secara manual menggunakan aplikasi Excel

terhadap ulasan komentar pengguna aplikasi WhatsApp. Kategori label pada data training dimulai dari memberi label pada data opini di aplikasi microsoft excel. Data tersebut dikategorikan kedalam dua label sentimen yaitu positif dan negatif. Label Positif diberikan kepada ulasan dengan sentimen positif terhadap aspek, label negatif diberikan kepada ulasan dengan sentimen negatif terhadap aspek.

Tabel 1. Contoh Set Pelabelan

| Komentar | Sentimen |
|--|----------|
| Dengan adanya aplikasi WhatsApp sangat membantu saya sekali, dalam hal panggilan video, Mengirim pesan, atau yg lainnya. makasih WhatsApp 😊😊 | POSITIF |
| Aplikasi ini bagus sekali apalagi ada pandemi corona ini wah hebat banget bisa berkomunikasi pada yang diluar sana saya suka dengan aplikasi WhatsApp wah keren sekali ya mari kita coba bersama sama ya supaya yang diluar sana tidak kesepian yak yuk dimulai dari sekarang gaess Terima kasih | POSITIF |
| Kok habis di update malah jadi keluar tulisan WhatsApp tidak menanggapi kenapa jadi gitu sih | NEGATIF |
| Saya sangat suka wa ini bisa chat teman saya dan bisa membuat grup wow bangett Terima kasih WhatsApp messenger karena ada nya aplikasi ini buat saya nggak bosan dan sekolah online terima kasih | POSITIF |
| Aku udah update tapi kenapa ya fitur barunya engga ada sama sekali | NEGATIF |
| Saya pribadi rasa puas menggunakan akun WhatsApp ini. | POSITIF |
| Gak berguna system silent sound WhatsApp di mode video call walau di silent suara tetap aja kita yang silent masih mendengar lawan bicara alias gak bisa arah harusnya penelpon ataupun penerima tidak bisa saling dengar. | NEGATIF |

Pemodelan Data

Pada tahap analisis ini metode yang digunakan adalah algoritma NB dan SVM yang kemudian dilakukan tahapan pembagian dataset. Pembagian tersebut dilakukan menggunakan 10 *fold cross validation* yang terdiri dari proses *training* dan *testing*. Setelah itu, dilakukan analisis menggunakan kedua algoritma NB dan SVM dan AUC berdasarkan *performance*. *Tools* yang digunakan adalah aplikasi *rapidminer*.

Klasifikasi Naive Bayes adalah konsep probabilitas yang bisa digunakan untuk menentukan kelas dokumen teks dan dapat mengolah data dalam jumlah besar dengan hasil akurasi yang tinggi [12]. Sedangkan klasifikasi SVM memiliki keunggulan mampu mengidentifikasi hyperplane terpisah yang memaksimalkan margin antara kedua kelas yang berbeda [20]. Namun Support Vector Machine memiliki kekurangan pada masalah pemilihan parameter atau fitur yang sesuai. Pemilihan fitur sekaligus parameter setup pada SVM berpengaruh signifikan terhadap hasil akurasi klasifikasi [21].

Evaluasi

Evaluasi bertujuan untuk memutuskan tingkat akurasi menggunakan metode NB dan SVM berupa akurasi dan AUC (Area under curve). Hasil dari pengujian tersebut berupa kelas prediksi yang akan dibandingkan dengan kelas aktual pada data uji dengan menggunakan *confusion matrix* yang tertera pada gambar 2. yaitu *false positive rate* (FP rate), *false negative rate* (FN rate), *true positive rate* (TP rate) dan *true negative rate* (TN rate) sebagai indikator. Evaluasi ini bertujuan untuk memahami fungsi dari model yang sebelumnya telah dibuat.

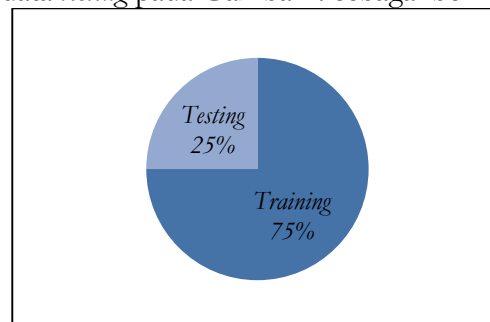
Tabel 2. *Confusion Matrix*

| <i>Fact</i> | <i>Prediction</i> | |
|------------------|--------------------|--------------------|
| | <i>Positif (P)</i> | <i>Negatif (N)</i> |
| <i>False (F)</i> | FP | FN |
| <i>True (T)</i> | TP | TN |

Hasil dan Pembahasan

Pengumpulan Data

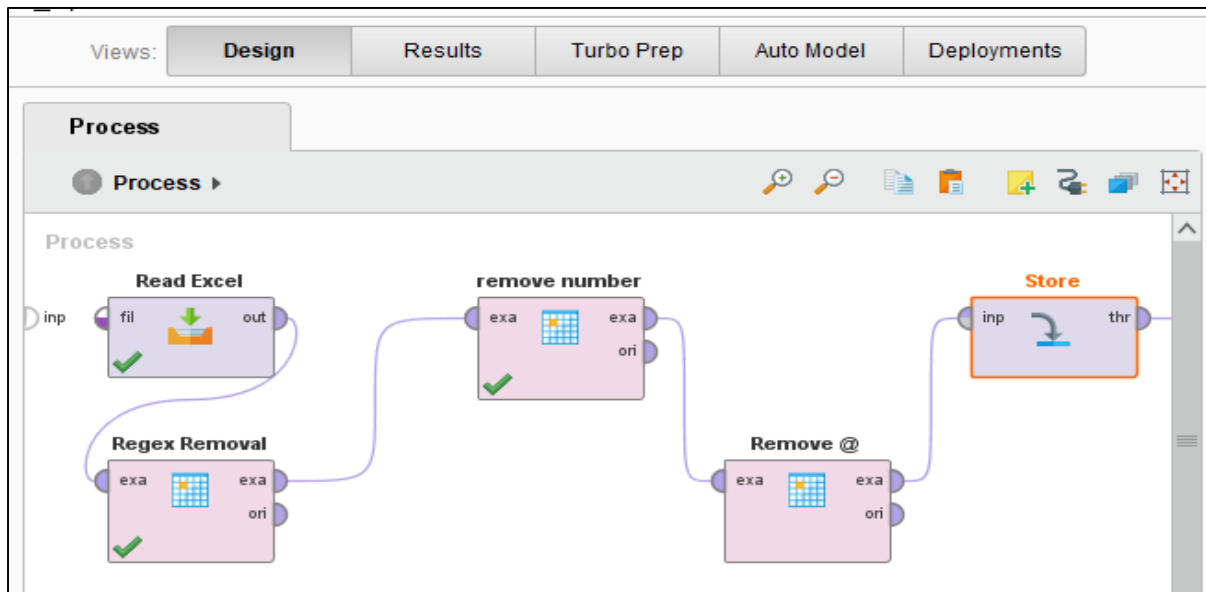
Dataset diambil dari ulasan pengguna WhatsApp pada *play store* dengan cara *web scrapping* atau *extension browser chrome Dataminer*. Berdasarkan komposisi yang tepat sesuai dengan kaidah data mining perbandingan antara data *training* dan data *testing* yaitu 3 : 1, dengan komposisi data *training* dipilih sebanyak 1.500 ulasan, dan data *testing* dipilih sebanyak 500 ulasan. Masing-masing data yang telah di *scrap* dari *playstore* tersebut diberi nama *training_WhatsApp.xlsx* dan *testing_WhatsApp.xlsx*. Data tersebut merupakan data yang diambil paling terbaru, disajikan perbandingan data *training* dan data *testing* pada Gambar 4 sebagai berikut:



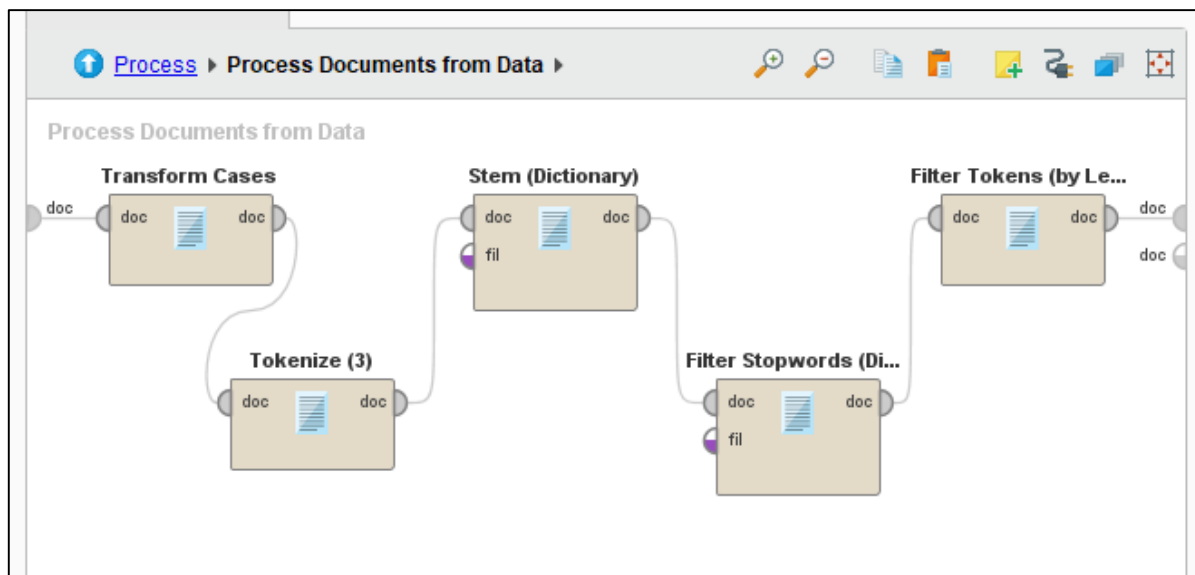
Gambar 4. Perbandingan data *training* dan data *testing*

Pemrosesan Data

Proses berikutnya adalah tahap preparing maupun persiapan, di dalam tahap ini terdapat beberapa langkah kegiatan yaitu membangun dan membersihkan data yang nantinya akan digunakan pada tahap pemodelan data. Langkah-langkah tersebut dimulai dari tahapan penghapusan ekspresi, menghilangkan tanda at (@), penghapusan angka, proses tokenisasi, penghilangan imbuhan, pengubahan huruf besar menjadi kecil, penghilangan kata sambung yang tidak memiliki arti, menghilangkan kata dengan panjang huruf tertentu. Dalam *preprocessing* ini terbagi dalam beberapa langkah yang terdapat pada Gambar 5 dan Gambar 6 sebagai berikut:



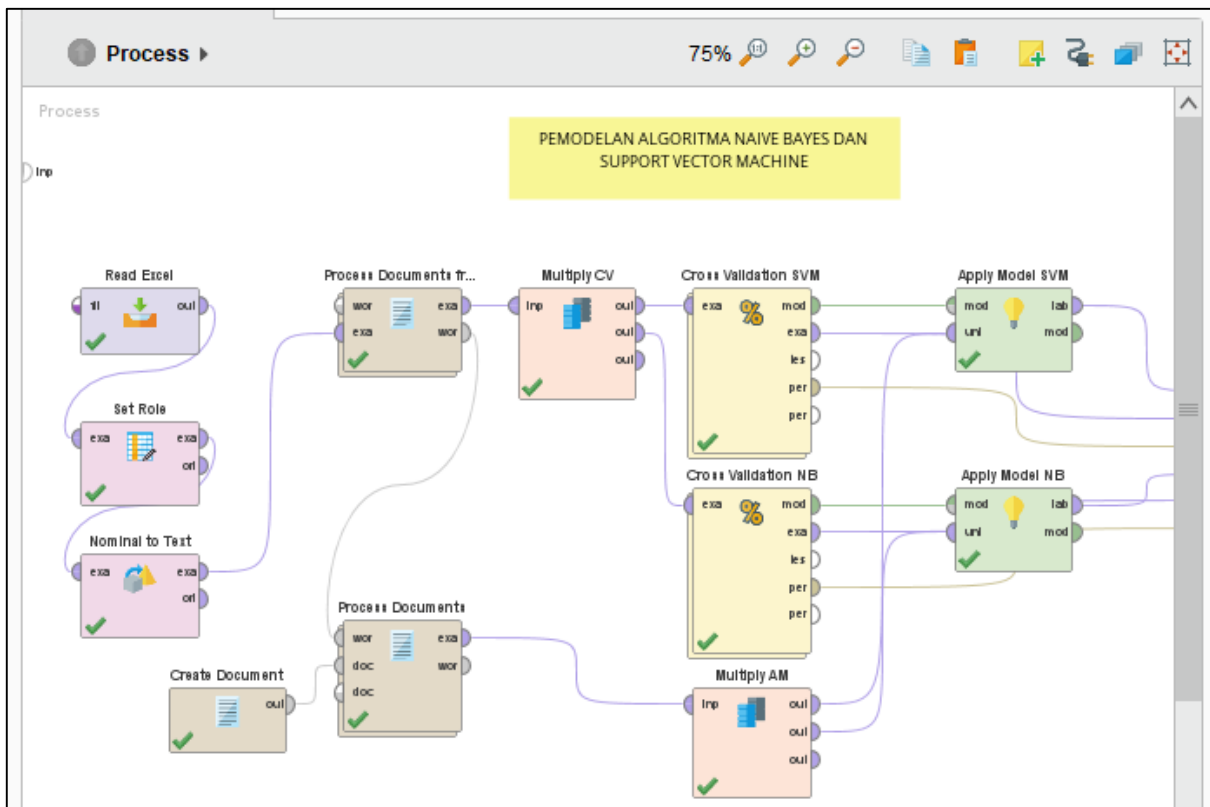
Gambar 5. Pengolahan data awal menggunakan *Tool RapidMiner*.



Gambar 6. Pengolahan data lanjutan menggunakan *Tool RapidMiner*.

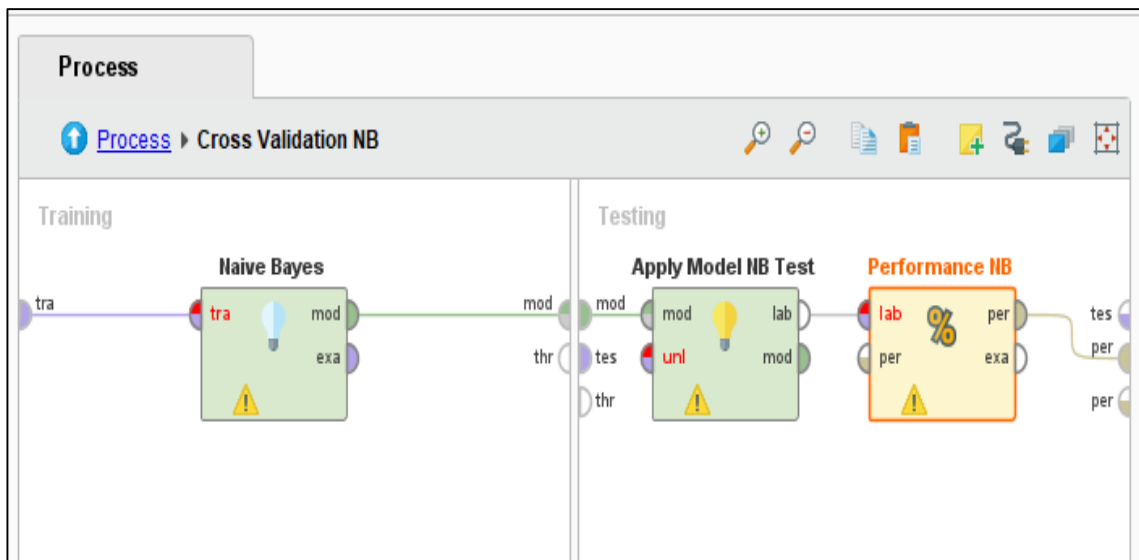
Pemodelan Data

Berikut ini disajikan hasil dari pemodelan data yang diberi label positif dan negatif menggunakan algoritma NB dan SVM yang terdapat pada Gambar 7.

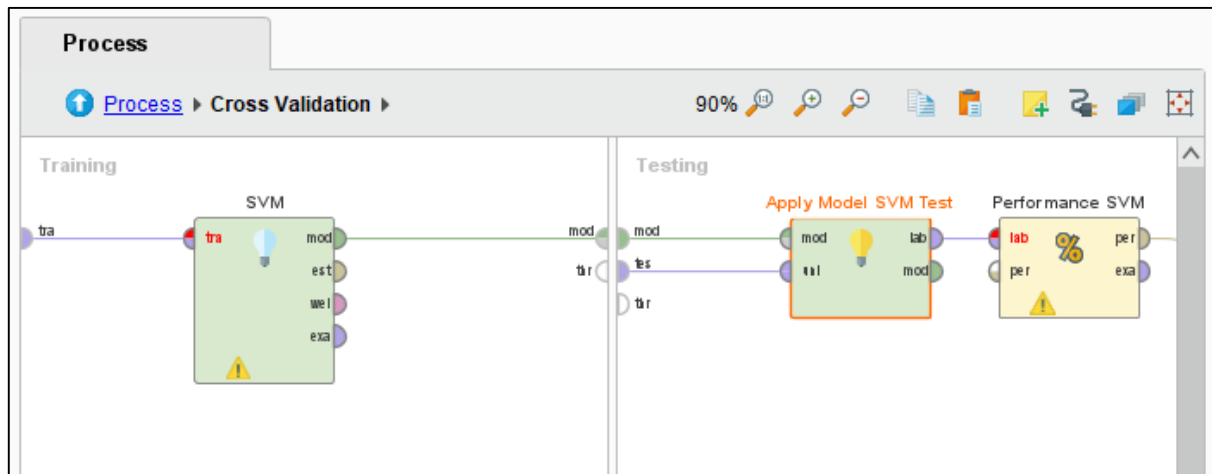


Gambar 7. Modelling Naive Bayes dan Support Vector Machine.

Setelah itu, melakukan tahapan validasi silang pada setiap metode *Naive Bayes* dan *Support Vector Machine* disajikan pada gambar 8 dan gambar 9.



Gambar 8. Cross Validation NB



Gambar 9. *Cross Validation SVM.*

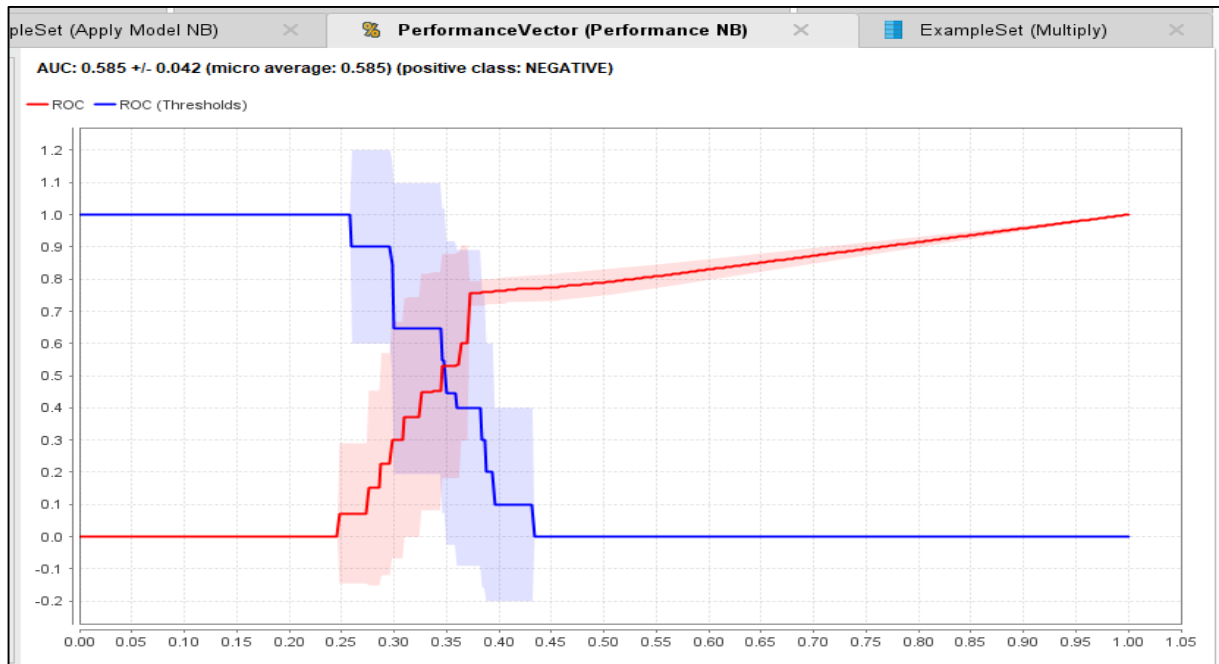
Evaluasi

Proses selanjutnya adalah melakukan proses evaluasi dengan cara membandingkan dua metode yang dianalisis yaitu algoritma NB dan SVM berupa akurasi dan AUC (*Area under curve*). Dilakukannya evaluasi ini adalah untuk mengetahui manfaat dari model yang dibuat pada tahap sebelumnya. Evaluasi yang dilakukan adalah dengan cara *confusion matrix* yaitu *false positive rate* (FP rate) , *false negative rate* (FN rate), *true positive rate* (TP rate) dan *true negative rate* (TN rate) sebagai indikator.

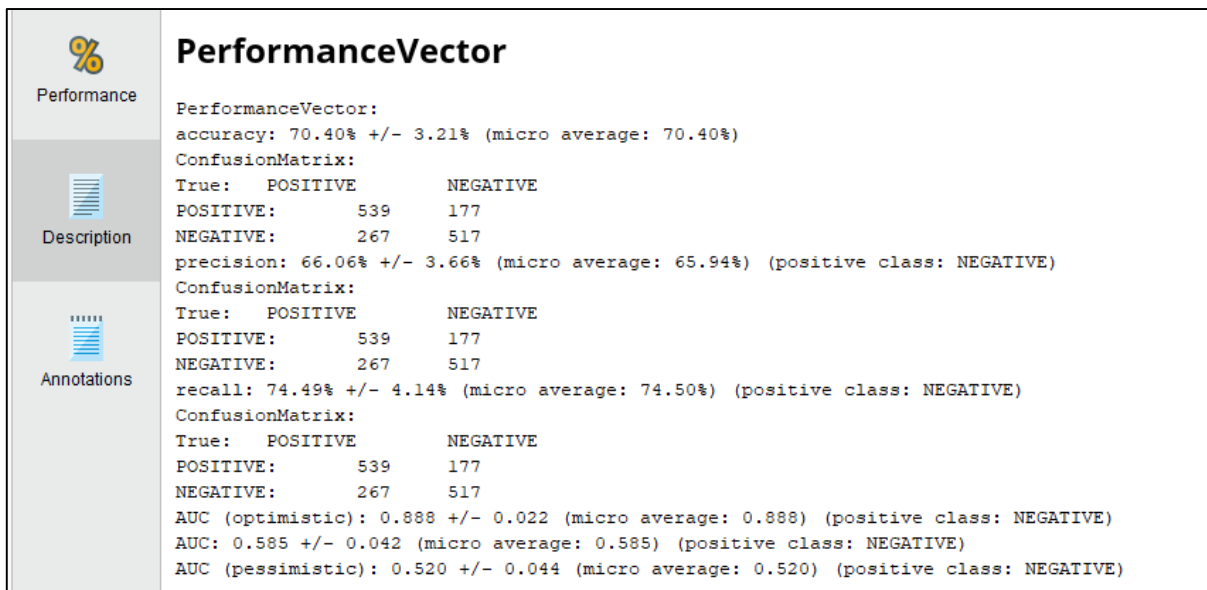
Berdasarkan hasil pengujian model NB didapatkan nilai akurasi yang terdapat pada Gambar 10 menunjukkan Jumlah *true positive rate* adalah 539 record dikategorikan sebagai label positif dan *false positive rate* adalah 267 record dikategorikan sebagai label negatif. Kemudian jumlah *true negative rate* adalah 517 record dikategorikan sebagai label negatif dan *false negative rate* adalah 177 record dikategorikan sebagai label positif. Gambar 10 menunjukkan tingkat akurasi algoritma NB adalah sebesar 70,40% dan pada Gambar 11 nilai AUC sebesar 0,585.

| | true POSITIVE | true NEGATIVE | class precision |
|----------------|---------------|---------------|-----------------|
| pred. POSITIVE | 539 | 177 | 75.28% |
| pred. NEGATIVE | 267 | 517 | 65.94% |
| class recall | 66.87% | 74.50% | |

Gambar 10. Nilai *Accuracy* Algoritma NB

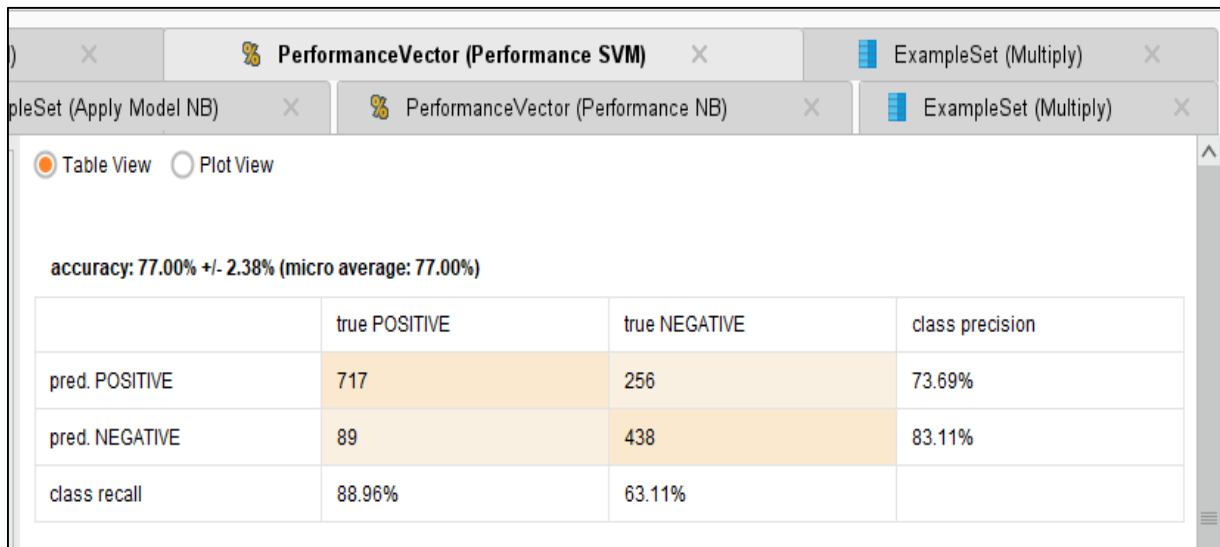


Gambar 11. Nilai AUC (*Area Under Curve*) NB



Gambar 12. Performance Vector Algoritma NB

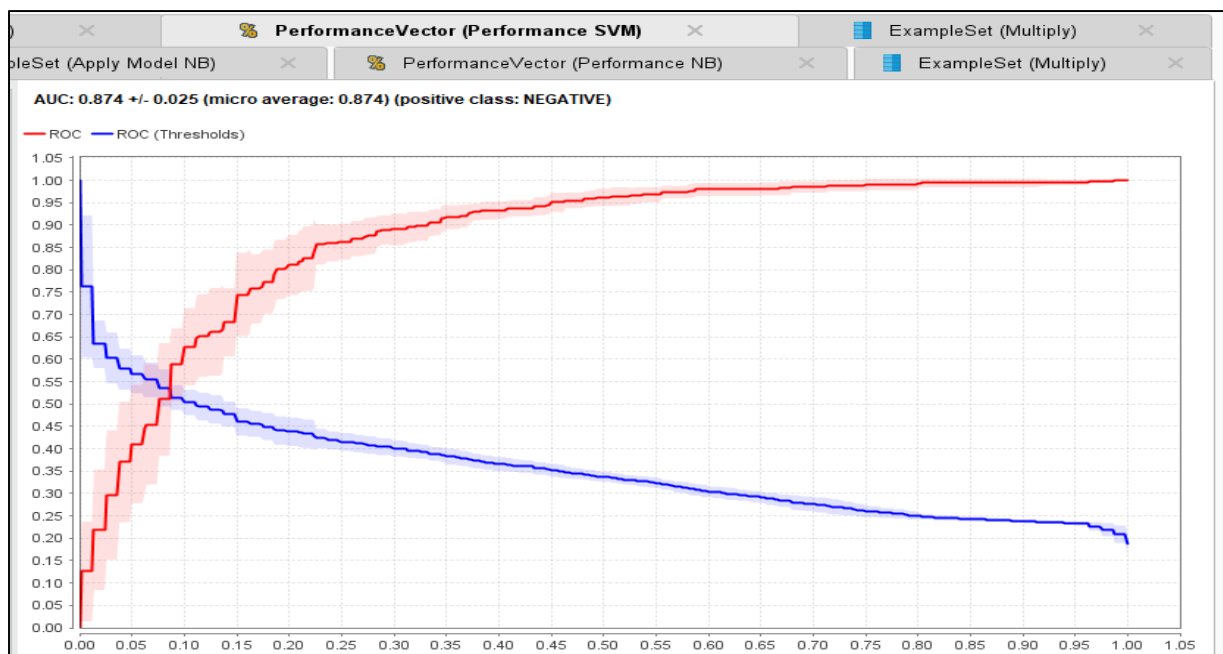
Berdasarkan hasil pengujian model SVM didapatkan nilai akurasi yang terdapat pada Gambar 13 menunjukkan Jumlah *true positive rate* adalah 717 *record* dikategorikan sebagai label positif dan *false positive rate* adalah 89 *record* dikategorikan sebagai label negatif. Kemudian jumlah *true negative rate* adalah 438 *record* dikategorikan sebagai label negatif dan *false negative rate* adalah 256 *record* dikategorikan sebagai label positif. Gambar 13 menunjukkan tingkat akurasi algoritma SVM adalah sebesar 77,00% dan pada Gambar 14 nilai AUC sebesar 0,874.



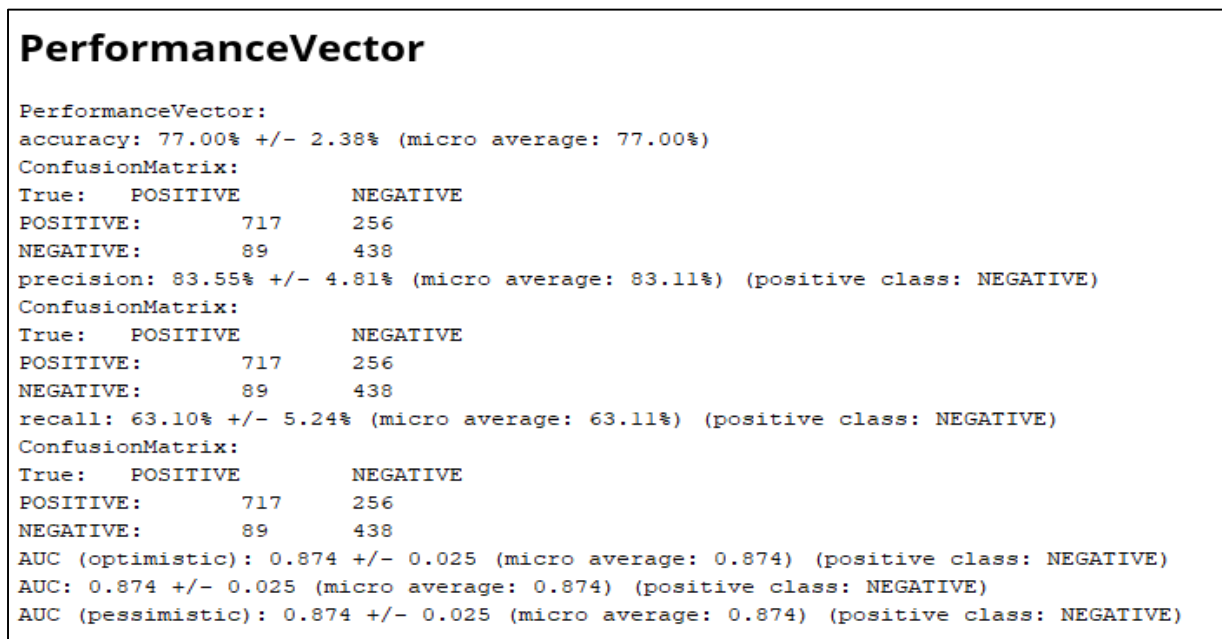
The screenshot shows a software window titled "PerformanceVector (Performance SVM)". It displays performance metrics for an SVM model. The accuracy is 77.00% with a standard deviation of +/- 2.38% (micro average: 77.00%). Below this, a table provides a detailed breakdown of the model's performance.

| | true POSITIVE | true NEGATIVE | class precision |
|----------------|---------------|---------------|-----------------|
| pred. POSITIVE | 717 | 256 | 73.69% |
| pred. NEGATIVE | 89 | 438 | 83.11% |
| class recall | 88.96% | 63.11% | |

Gambar 13. Nilai Akurasi Algoritma SVM



Gambar 14. Nilai AUC (*Area Under Curve*) Algoritma SVM



Gambar 15. *Performance Vector Algoritma SVM*

Perolehan hasil dari masing-masing model algoritma dapat dilihat pada tabel berikut ini :

Tabel 3. *Confusion Matrix* NB dan SVM

| Metode | <i>True Positive</i> | <i>False Positive</i> | <i>True Negative</i> | <i>False Negative</i> |
|------------|----------------------|-----------------------|----------------------|-----------------------|
| <i>NB</i> | 539 | 177 | 517 | 267 |
| <i>SVM</i> | 717 | 258 | 438 | 89 |

Tabel 4. Perbandingan Performa algoritma NB dan SVM

| Metode | Akurasi | AUC (<i>Area under curve</i>) |
|------------|---------|---------------------------------|
| <i>NB</i> | 70,40% | 0,585 |
| <i>SVM</i> | 77,00% | 0,874 |

Kesimpulan

Berdasarkan hasil analisis kepuasan pengguna aplikasi WhatsApp pada *play store* dengan algoritma NB dan SVM melalui proses *preprocessing* dan *modeling* serta proses evaluasi pengkategorian ulasan berlabel positif dan ulasan berlabel negatif dapat diketahui bahwa hasil akhir nilai akurasi NB adalah sebesar 70,40% dan SVM sebesar 77,00% dan nilai AUC pada algoritma NB sebesar 0,585 dan SVM adalah 0,876. SVM memiliki tingkat akurasi yang unggul 6,6%, daripada NB. Maka metode SVM dapat digunakan pada karakteristik data yang sama. Penelitian ini memberikan solusi yaitu berupa informasi pada pihak manajemen WhatsApp agar tetap konsisten dalam pelayanan pada pengguna sehingga tingkat kepercayaan pengguna tetap terjaga dan meningkatkan kualitas aplikasi demi tercapainya kepuasan pengguna WhatsApp. Hasil penelitian akan semakin akurat apabila penulis menghindari kata-kata yang bermakna sama atau bermakna

ganda dan menggunakan dataset yang lebih banyak agar menambah keakuratan data yang akan diteliti.

Ucapan Terima Kasih

Penelitian ini dapat dilaksanakan dengan baik berkat bantuan dari berbagai pihak, untuk itu peneliti mengucapkan terimakasih kepada Bapak Sudin Saepudin, M.Kom dan Bapak Dudih Gustian, S.T., M.Kom yang telah memberikan bimbingan dan arahnya selama pengerjaan penelitian ini.

Daftar Pustaka

- [1] A. Habibie *et al.*, “Analisis Preferensi Konsumen Dalam Memilih Smartphone Di Indonesia Consumer Analysis of Preferences in Choosing Smartphone in,” vol. 7, no. 1, pp. 114–124, 2020.
- [2] Newzoo, “Top Countries by Smartphone Users.” .
- [3] Supriyati, R. S. Bahri, and E. Komarudin, “Computerized of International Financial Report Standard for Good Governance in Small Medium Enterprises,” in *IOP Conference Series: Materials Science and Engineering*, 2019, vol. 662, no. 5, doi: 10.1088/1757-899X/662/5/052009.
- [4] V. Cipher, A. E. S. Dan, R. C. Untuk, and A. Pesan, “Implementasi Kriptografi Dengan Algoritma,” vol. 1, no. 3, pp. 931–937, 2018.
- [5] Y. Pusparisa, “Whatsapp Kuasai Pasar Aplikasi Pesan Instan Global 2020,” *Databoks*, 2020.
- [6] F. Marleny and Mambang, “Analisis Pengguna Whatsapp Terhadap Kesalahan Mengirim Pesan Teks Menggunakan Metode Klasifikasi,” *J. Teknol. Inf. Univ. Lambung Mangkurat*, vol. 5, no. 1, pp. 19–24, 2020, doi: 10.20527/jtiulm.v5i1.43.
- [7] T. S. Gunardi, Dedi Sulistiyo S, “Usulan Peta Strategis Teknologi Informasi Menggunakan Pendekatan Analisis Critical Success Factor (CSF) dan IT Balanced Scorecard,” *Http://Jurnal.Unikom.Ac.Id/Jurnal/Usulan-Peta-Strategi.2V*, vol. 10, no. 1, pp. 143–153, 2012.
- [8] A. P. Giovani, A. Ardiansyah, T. Haryanti, L. Kurniawati, and W. Gata, “Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi,” *J. Teknoinfo*, vol. 14, no. 2, p. 115, 2020, doi: 10.33365/jti.v14i2.679.
- [9] A. Erfina and Y. H. Putra, “Irony Sentence Detection Techniques Using Fuzzy Historical Classifier,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 662, no. 6, 2019, doi: 10.1088/1757-899X/662/6/062004.
- [10] D. C. Rini, Y. Farida, and D. Puspitasari, “Klasifikasi Menggunakan Metode Hybrid Bayesian-Neural Network (Studi Kasus: Identifikasi Virus Komputer),” *J. Mat. MANTIK*, vol. 1, no. 2, p. 38, 2016, doi: 10.15642/mantik.2016.1.2.38-43.
- [11] A. P. Fadillah and B. Hardiyana, “Penerapan Naïve Bayes Classifier Untuk Pemilihan Konsentrasi Mata Kuliah,” *J. Teknol. dan Inf. UNIKOM*, vol. 8, no. 2, 2018, doi: 10.34010/jati.v8i2.1039.
- [12] A. R. T. Lestari, R. S. Perdana, and M. A. Fauzi, “Analisis Sentimen Tentang Opini Pilkada DKI 2017 Pada Dokumen Twitter Berbahasa Indonesia Menggunakan Naïve Bayes dan Pembobotan Emoji,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 1, no. 12, pp. 1718–1724, 2017.
- [13] D. J. Haryanto, L. Muflikhah, and M. A. Fauzi, “Analisis Sentimen Review Barang Berbahasa Indonesia Dengan Metode Support Vector Machine Dan Query Expansion,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 2, no. 9, pp. 2909–2916, 2018.
- [14] I. Santoso, W. Gata, and A. B. Paryanti, “Penggunaan Feature Selection di Algoritma Support Vector Machine untuk,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 1, no.

- 10, pp. 5–11, 2019.
- [15] K. A. Rokhman, B. Berlilana, and P. Arsi, “Perbandingan Metode Support Vector Machine dan Decision Tree untuk Analisis Sentimen Review Komentar Pada Aplikasi Transportasi Online,” *J. Inf. Syst. Manag.*, vol. 2, no. 2, pp. 1–7, 2021, doi: <https://doi.org/10.24076/joism.2021v3i1.341>.
- [16] S. A. Aaputra, “Sentiment Analysis Analisis Sentimen E-Wallet Pada Google Play Menggunakan Algoritma Naive Bayes Berbasis Particle Swarm Optimization,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 3, no. 3, pp. 377–382, 2019.
- [17] 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, 2020, doi: 10.30865/mib.v4i3.2181.
- [18] U. Rofiqoh, R. S. Perdana, and M. A. Fauzi, “Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada Twitter Dengan Metode Support Vector Machine dan Lexion Based Feature,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 1, no. 12, pp. 1725–1732, 2017.
- [19] D. Gustian, I. Suciati, and S. Saepudin, “Sistem Pakar Dengan Algoritma Naive Bayes Untuk Prediksi Hasil Produksi Ayam Broiler Plasma (Studi Kasus : Pt . Sekawan Sinar Surya),” no. 12. .
- [20] D. A. Kristiyanti, A. H. Umam, M. Wahyudi, R. Amin, and L. Marlinda, “Comparison of SVM Naive Bayes Algorithm for Sentiment Analysis Toward West Java Governor Candidate Period 2018-2023 Based on Public Opinion on Twitter,” *2018 6th Int. Conf. Cyber IT Serv. Manag. CITSM 2018*, no. Citism, pp. 1–6, 2019, doi: 10.1109/CITSM.2018.8674352.
- [21] S. Sarafrazi and H. Nezamabadi-Pour, “Facing the classification of binary problems with a GSA-SVM hybrid system,” *Math. Comput. Model.*, vol. 57, no. 1–2, pp. 270–278, 2013, doi: 10.1016/j.mcm.2011.06.048.