

Implementasi Algoritma SVM Dan Naive Bayes Untuk Analisis Sentimen Pada Ulasan Pengguna Aplikasi Onedrive

Andrian Eko Widodo^{*1}, Fanny Fatma Wati², Nadiyah Hidayati³

^{1,2,3}Universitas Bina Sarana Informatika

Email: ¹andrian.aeo@bsi.ac.id, ²fanny.ffw@bsi.ac.id, ³nadiyah.nyy@bsi.ac.id

Abstrak

Teknologi modern kini menyediakan solusi untuk menyimpan data secara online, yang dikenal sebagai cloud storage, seperti OneDrive yang memungkinkan pengguna untuk menyimpan, berbagi, dan mengakses berkas secara fleksibel dari berbagai platform. Dengan populernya OneDrive, peningkatan penggunaannya membawa implikasi besar terhadap pengalaman pengguna dan keberhasilan bisnis Microsoft. Persepsi dan pengalaman pengguna sangat penting dalam menilai kualitas aplikasi, karena ulasan di platform seperti Google Play Store memberikan gambaran nyata tentang kepuasan dan kebutuhan pengguna. Untuk memastikan sebuah aplikasi dapat digunakan dan diterima oleh pengguna, diperlukan pengalaman pengguna yang optimal. Seiring meningkatnya jumlah pengguna di Google Play Store, ulasan yang diberikan juga terus bertambah. Ulasan ini sering menjadi sumber informasi penting tentang produk atau aplikasi perangkat lunak tertentu. Akan tetapi, banyaknya ulasan yang beragam dapat membawa pengaruh baik maupun buruk bagi pengembang. Oleh karena itu, diperlukan sistem otomatis untuk memproses data ulasan tersebut dengan menggunakan analisis sentimen. Penelitian ini bertujuan untuk menganalisis sentimen pengguna terhadap aplikasi OneDrive yang terdapat di Google Play Store. Untuk melihat kecenderungan opini pengguna, apakah bersifat negatif atau positif. Dalam penelitian ini, digunakan metode Naive Bayes dan Support Vector Machine (SVM) untuk mengelompokkan ulasan pengguna ke dalam sentimen positif atau negatif. Proses penelitian meliputi perumusan masalah, perancangan penelitian, pengumpulan data ulasan, pengolahan data, penyajian data dalam bentuk tabel, dan analisis data. Peneliti berharap hasil dari penelitian ini dapat membantu pengembang aplikasi dalam mengidentifikasi kelebihan dan kekurangan aplikasi mereka, sekaligus menjadi bahan evaluasi untuk perbaikan di masa depan.

Kata Kunci: Analisa Sentimen, Aplikasi OneDrive, Algoritma SVM, Naive Bayes

Abstract

Modern technology now provides solutions for storing data online, known as cloud storage, such as OneDrive which allows users to store, share and access files flexibly from various platforms. With the popularity of OneDrive, its increasing use has major implications for Microsoft's user experience and business success. User perception and experience are very important in assessing the quality of an app, as reviews on platforms like Google Play Store provide a real picture of user satisfaction and needs. To ensure an application can be used and accepted by users, an optimal user experience is needed. As the number of users on the Google Play Store increases, the reviews provided also continue to increase. These reviews are often an important source of information about a particular software product or application. However, the large number of mixed reviews can have both good and bad influences on developers. Therefore, an automatic system is needed to process the review data using sentiment analysis. This research aims to analyze user sentiment towards the OneDrive application on the Google Play Store. To see the trend of user opinions, whether they are negative or positive. In this research, the Naive Bayes and Support Vector Machine (SVM) methods were used to group user reviews into positive or negative sentiment. The research process includes problem formulation, research design, review data collection, data processing, data presentation in tabular form, and data analysis. Researchers hope that the results of this research can help application developers identify the strengths and weaknesses of their applications, as well as provide evaluation material for future improvements.

Keywords: Sentiment Analysis, OneDrive Application, SVM Algorithm, Naive Bayes

1. PENDAHULUAN

Di era masyarakat yang berpusat pada data dan informasi, arsip sangat penting karena setiap kegiatan, baik individu maupun institusi, selalu terkait dengan arsip. Hal tersebut menyebabkan pentingnya mengelola dan menjaga dokumen secara efisien untuk mencegah kehilangan atau kerusakan

(1). Kemajuan teknologi saat ini menawarkan solusi penyimpanan data secara daring yang disebut cloud storage. Layanan ini memungkinkan data disimpan melalui internet dan diakses dari berbagai perangkat (2). Contohnya adalah aplikasi OneDrive, yang memungkinkan pengguna menyimpan, berbagi, dan mengakses berkas secara fleksibel. Popularitas OneDrive di berbagai platform meningkatkan pengalaman pengguna dan keberhasilan bisnis Microsoft. Persepsi dan pengalaman pengguna sangat penting untuk menilai kualitas aplikasi, karena ulasan di platform seperti Google Play Store mencerminkan kepuasan dan kebutuhan pengguna.

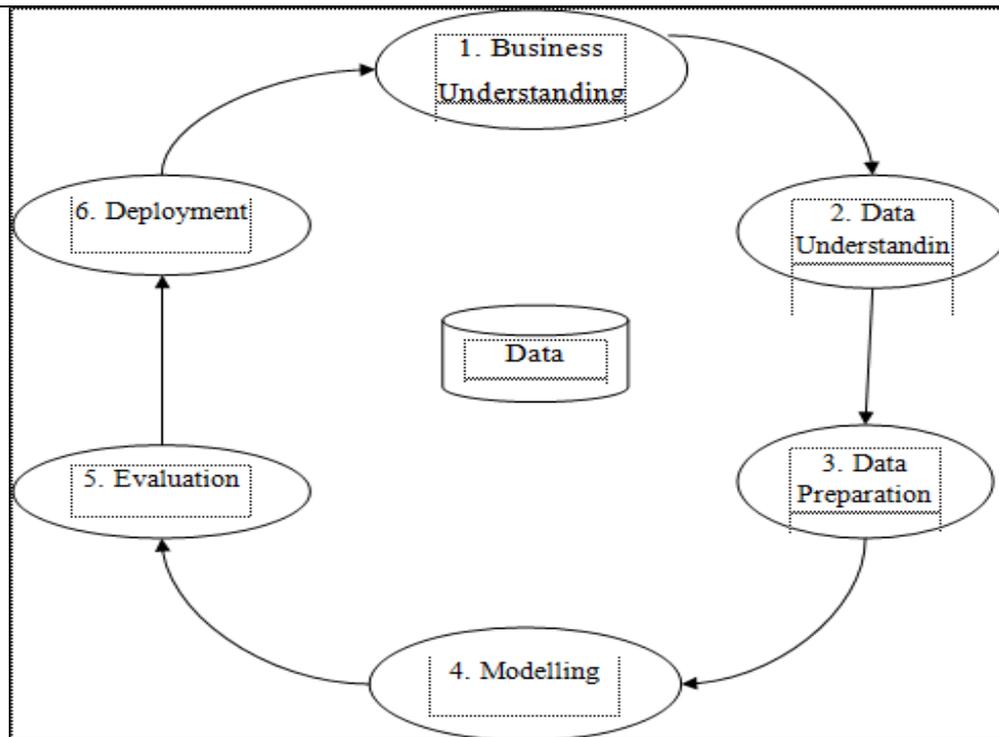
Agar layanan aplikasi diterima pengguna, aplikasi harus menawarkan Pengalaman pengguna yang positif mencakup pandangan dan reaksi individu terhadap produk, sistem, atau layanan (3). Dengan semakin banyaknya pengguna aplikasi di Google Play Store setiap tahun, maka jumlah ulasan yang diberikan juga terus bertambah (4). Ulasan ini, sering digunakan sebagai sumber informasi ketika mencari informasi tentang produk atau aplikasi perangkat lunak tertentu (5). Banyaknya ulasan yang beragam dapat memberikan dampak baik maupun buruk bagi pengembang. Oleh karena itu, diperlukan sistem otomatis untuk mengelola data ulasan tersebut. Salah satu teknik yang digunakan adalah analisis sentimen, yaitu proses otomatis untuk memproses data teks dan memperoleh informasi mengenai sentimen dalam suatu opini (6). Analisis ini bertujuan untuk melihat kecenderungan opini atau pendapat mengenai suatu masalah atau objek, apakah kecenderungannya negatif atau positif (7).

Pada penelitian ini analisa sentimen dilakukan untuk melihat ulasan dari pengguna aplikasi OneDrive di Google Play Store, ulasan tersebut dikategorikan menjadi positif dan negatif. Salah satu teknik untuk menganalisis sentimen adalah dengan menggunakan metode Naive Bayes Classifier, yang merupakan metode klasifikasi berdasarkan teorema Bayes. Metode ini dipilih karena cocok untuk jumlah masukan besar serta unggul dalam kecepatan dan kesederhanaan (8). Support Vector Machine (SVM) adalah metode yang digunakan untuk klasifikasi dan regresi, baik untuk masalah linear maupun non-linear. Metode ini unggul dalam menerapkan pemisahan linear pada data input non-linear berdimensi tinggi, SVM didasarkan pada konsep menemukan fungsi pemisah (*hyperplane*) yang dapat memisahkan dua kelas secara optimal (9). Banyak penelitian telah melaporkan bahwa SVM adalah salah satu metode paling akurat untuk klasifikasi teks (10).

Penelitian ini bertujuan untuk memprediksi label sentimen pada ulasan pengguna aplikasi OneDrive di Google Play Store menggunakan metode Naive Bayes dan Support Vector Machine sebagai model klasifikasi. Ulasan tersebut akan dikategorikan ke dalam dua kelompok, yaitu positif dan negatif. Peneliti berharap hasil penelitian ini dapat memberikan bantuan kepada pengembang aplikasi dalam mengevaluasi kekurangan dan kelebihan aplikasi mereka, serta menjadi referensi untuk evaluasi di masa depan.

2. METODE PENELITIAN

Penelitian ini dilaksanakan melalui beberapa tahapan yang dikembangkan dari model *Cross-Standard Industry for Data Mining* (CRISP-DM), yang terdiri dari enam tahap, yaitu *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modelling*, *Evaluation*, dan *Deployment*. Gambar 1 di bawah ini menunjukkan model yang diusulkan dalam penelitian ini dengan menerapkan metode CRISP-DM.



Gambar 1 Metode Penelitian

1) Pemahaman Bisnis (*Business Understanding*)

Tahapan ini merupakan langkah pertama dari CRISP-DM, di mana penelitian difokuskan pada pencapaian tujuan penelitian. Pada tahap ini, peneliti mengumpulkan data ulasan yang ada di Google Play. Selain itu, pada tahap *Business Understanding*, dilakukan pemahaman terhadap model yang digunakan dalam proses pengolahan data, dengan pendekatan pengkategorian terbaik yang akan diperoleh melalui perbandingan hasil dari algoritma. Klasifikasi dilakukan dengan menggunakan algoritma SVM dan NB.

2) Pemahaman Data (*Data Understanding*)

Pada tahap ini, dilakukan pengambilan data mentah sesuai dengan atribut yang diperlukan. Data diperoleh dari ulasan aplikasi OneDrive yang diambil dari halaman *review* aplikasi OneDrive. Berikut adalah beberapa contoh ulasan aplikasi OneDrive yang tersedia di Google Play Store:

No	Komentar	Label
1	Saya sudah membeli langganan premium 100 GB uang saya melalui shopee pay sudah terpotong 32rb tapi setelah saya cek one drive penyimpanan saya tidak bertambah dan langganannya belum aktif, bagaimana solusinya ya ?? Kalau semisalkan ada masalah di microsoftnya harusnya uang saya dikembalikan!	Negatif
2	Parah akun di kunci sudah isi formulir sudah ke pusat bantuan sdh minta email tp tetap di kunci	Negatif
3	Ngasih free kok pelit, kalau mau maju dan bisa bersaing dengan cloud lainnya, kasih lah space penyimpanan besar, sama google aja punya banyak, ada google foto, Drive, belum lain aplikasi pihak 3 terabox	Negatif
4	lengkap untuk penyimpanan foto dan dokumen saya	Positif

5	sangat bermanfaat bagi pengguna yg kurang ruang penyimpanan eksternal.	Positif
6	Sangat bagus dan bermanfaat untuk semua dalam penyimpanan	Positif

3) Persiapan Data (*Data Preparation*)

Pada tahap ini, dilakukan proses persiapan data yang bertujuan untuk memperoleh data yang bersih dan siap digunakan dalam penelitian. Data awal yang diperoleh kemudian melalui proses pembersihan untuk menghapus kata-kata duplikat. Hasil dari pelabelan dan pengkategorian yang dilakukan oleh tim ahli bertujuan untuk menormalkan status data dalam penelitian text mining. Berikut adalah hasil dari tahap ini:

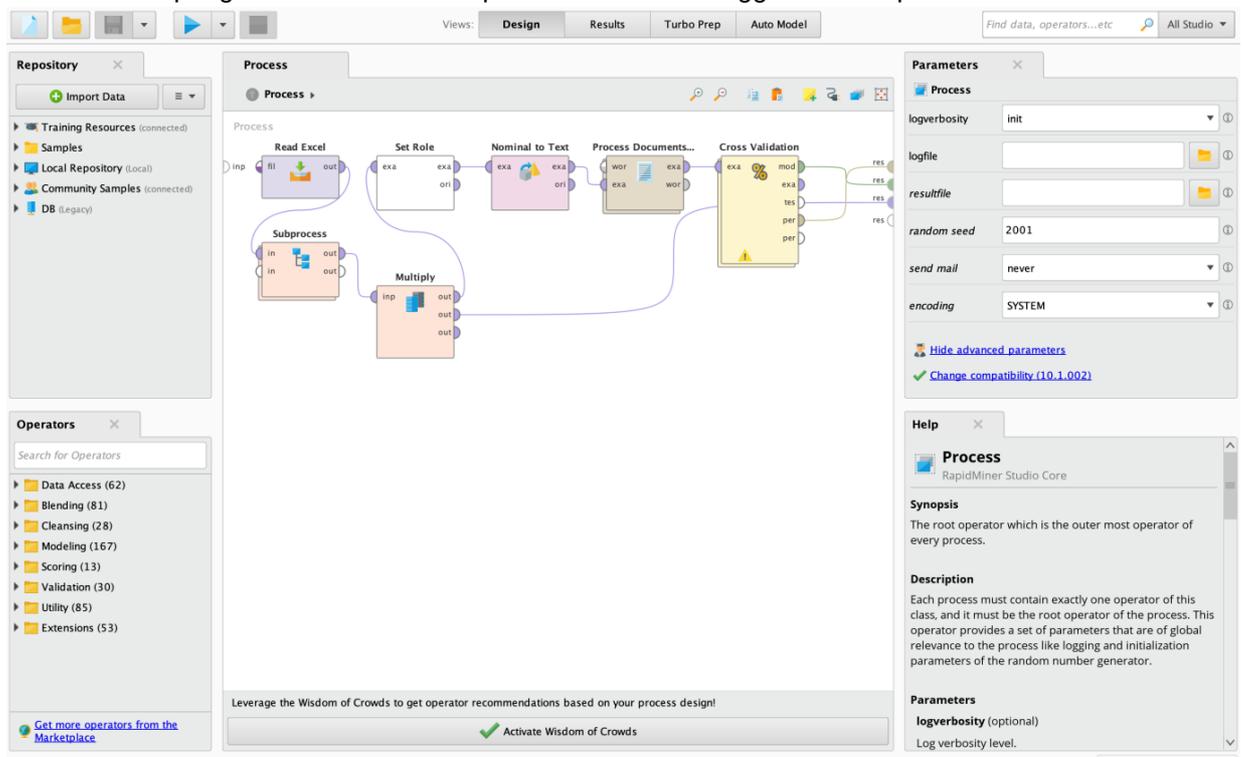
Tabel 1. Tabel Hasil Proses Pelabelan

Label Kategori	Jumlah Kategori
Negatif	500
Positif	500

Sumber: Hasil Penelitian

4) Pemodelan (*Modeling*)

Modelling adalah tahap pemilihan teknik mining dengan menentukan algoritma yang akan digunakan untuk mengolah data. Tools yang digunakan dalam tahap ini adalah Rapid Miner. Hasil pengujian model yang dilakukan bertujuan untuk mengklasifikasikan data ulasan menggunakan algoritma SVM dan Naive Bayes, dengan menerapkan model *Partitioning* dan *Cross Validation*. Berikut adalah desain model pengolahan data ulasan aplikasi OneDrive menggunakan Rapid Miner.



Gambar 2. Model Pengolahan Data

5) Evaluation

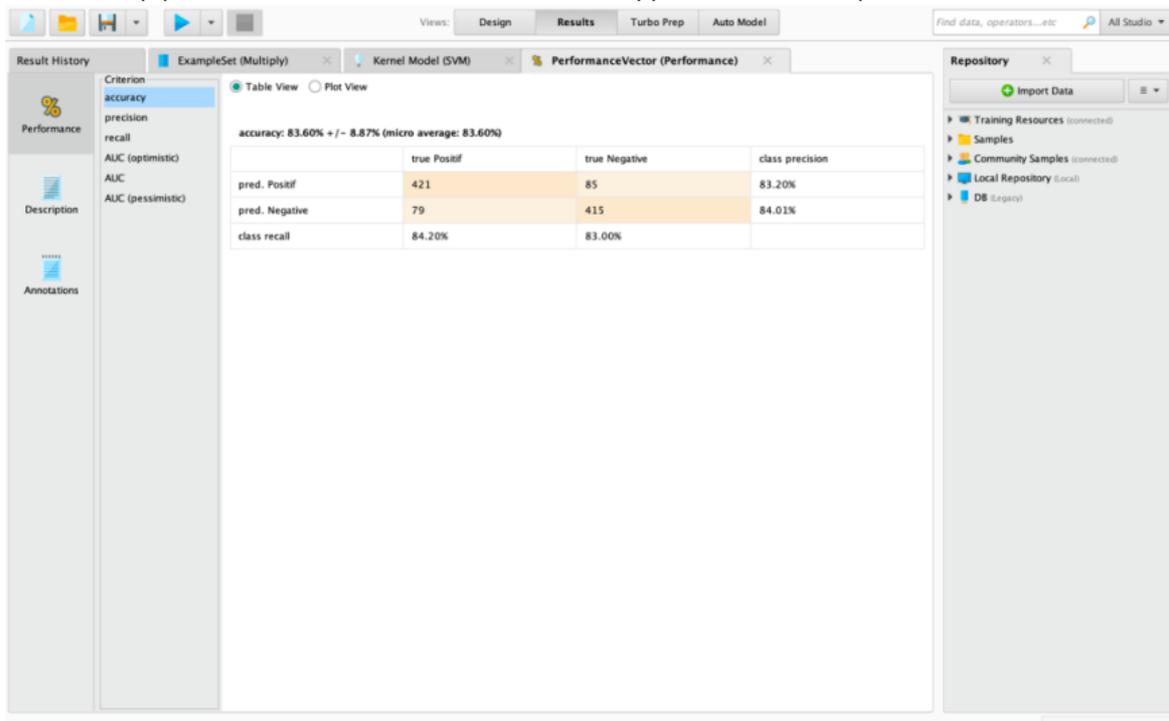
Evaluasi dilakukan berdasarkan nilai akurasi, AUC, dan *confussion matrix*. *Confusion Matrix* akan menggambarkan hasil akurasi mulai dari prediksi positif yang benar, prediksi positif yang salah, prediksi negatif yang benar dan prediksi negatif yang salah. Akurasi akan dihitung dari seluruh hasil prediksi yang benar (baik prediksi positif dan negatif) dibandingkan dengan seluruh data testing. Semakin tinggi nilai akurasi, semakin baik pula model yang dihasilkan.

3. HASIL DAN PEMBAHASAN

Hasil pengujian model dengan klasifikasi SVM dan Naive Bayes digunakan untuk mengukur nilai *accuracy*, *precision*, *recall*, dan AUC. *Cross validation* juga diterapkan sebagai model validasi. Data yang dianalisis adalah ulasan aplikasi OneDrive. Berikut ini adalah hasil dari confusion matrix dan kurva ROC:

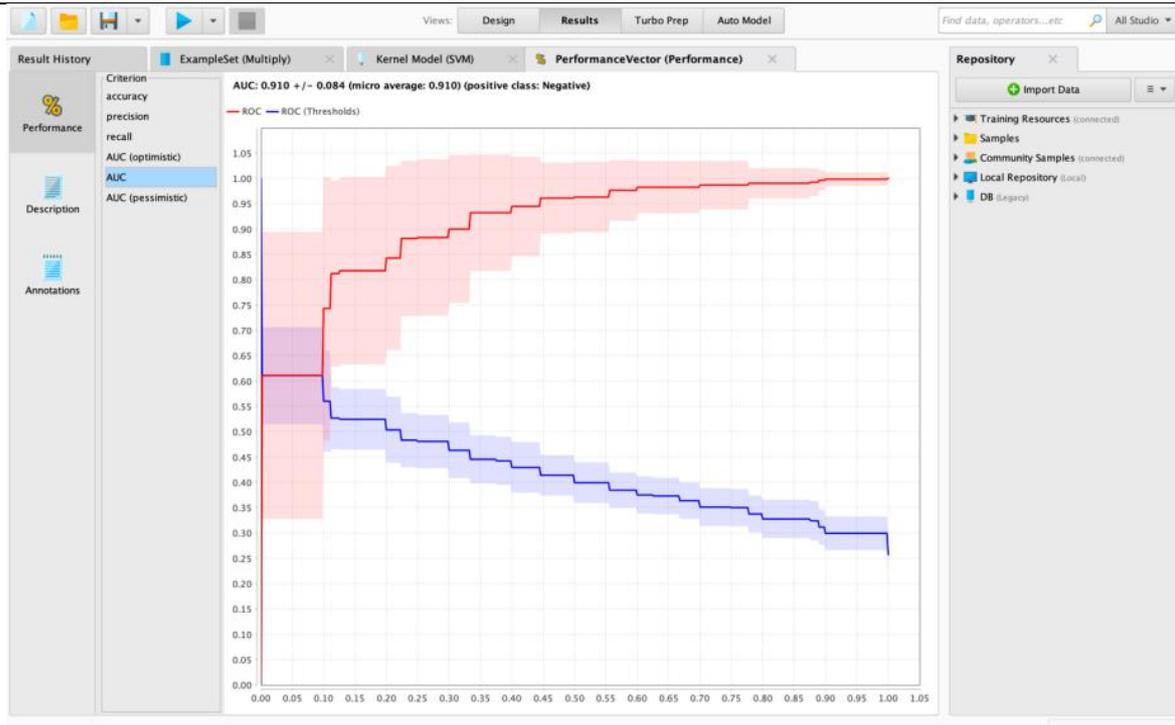
1) Hasil Evaluasi dengan metode SVM

Dalam tahap pemodelan ini, teknik klasifikasi data *Support Vector* diterapkan..



Gambar 3. Hasil Accuracy Dengan Metode SVM

Hasil dari analisis kami menunjukkan bahwa akurasi dengan penggunaan metode *Support Vector Machine* (SVM) hasil analisis sentimen terhadap aplikasi OneDrive yang kami terapkan adalah sebesar 83,60%. Dari total 500 data ulasan yang kami gunakan, terdapat 423 data ulasan positif yang diprediksi dengan benar oleh SVM, terdapat 85 data ulasan negatif dari pengguna yang terprediksi sebagai positif, dan 79 data ulasan positif yang terprediksi sebagai negatif, sementara 415 data ulasan negatif yang sesuai dengan prediksi negatif. *Confusion matrix* yang terbentuk dari proses pengolahan data ulasan aplikasi ini memberikan penilaian terhadap kinerja klasifikasi berdasarkan objek yang benar dan salah. Selain *confusion matrix*, tingkat akurasi juga didukung oleh kurva ROC. Berikut ini adalah tampilan kurva ROC dari algoritma SVM:



Gambar 4. ROC curve dari algoritma SVM

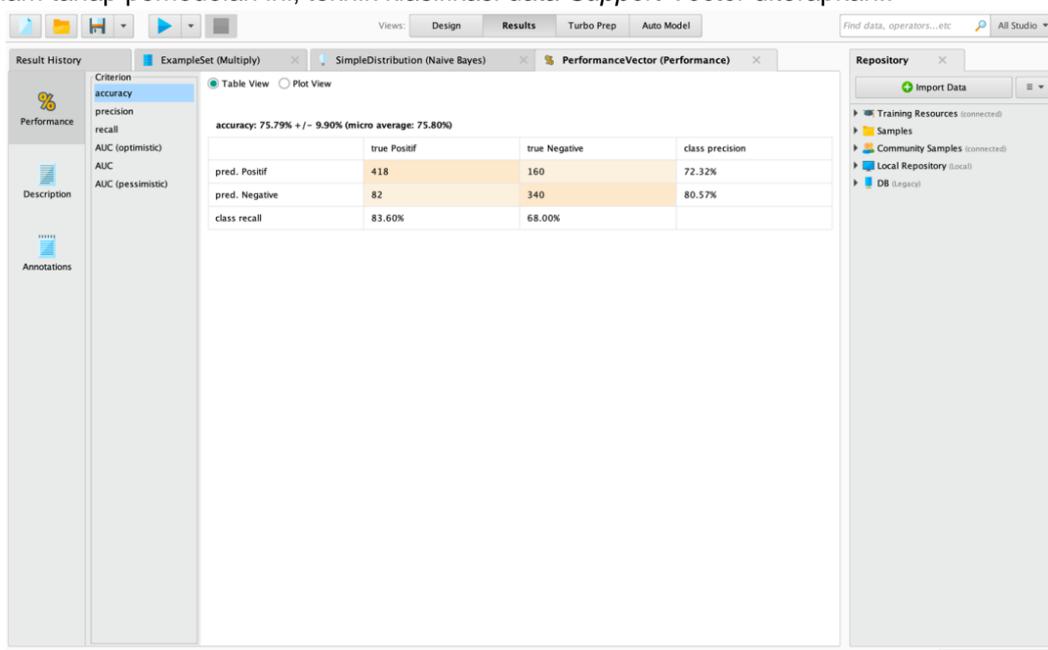
Kurva ROC SVM yang dihasilkan dari pengolahan dataset ulasan aplikasi menunjukkan nilai AUC yang dihasilkan dari gambar 4 adalah sebesar sebesar 0,910. +/- 0,084 (micro average: 0,910) (Positive class: Negatif). Berdasarkan penjelasan di atas, dapat disimpulkan bahwa pengolahan data menggunakan metode SVM pada ulasan aplikasi OneDrive menghasilkan hasil sebagai berikut:

Tabel 2. Nilai hasil Accuracy, Recall, AUC

	Accuracy	Recall	Precision	AUC
SVM	83,60%	83,06%	84,90%	0,910

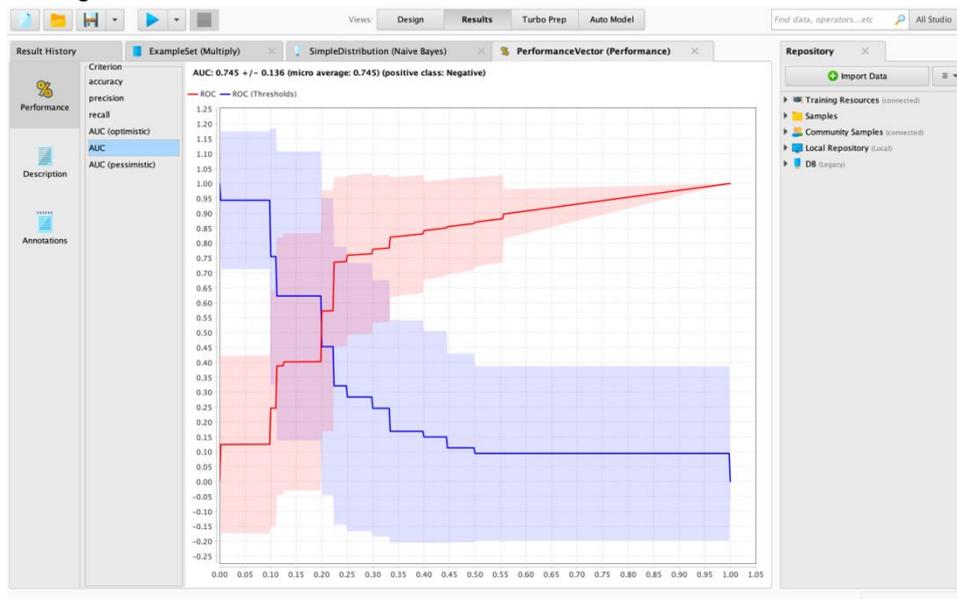
2) Hasil Evaluasi dengan metode NB

Dalam tahap pemodelan ini, teknik klasifikasi data *Support Vector* diterapkan..



Gambar 5. Hasil Accuracy Dengan Metode NB

Hasil dari analisis kami menunjukkan bahwa akurasi dengan penggunaan metode *Naïve Bayes* (NB) hasil analisis sentimen terhadap aplikasi OneDrive yang kami terapkan adalah sebesar 75,79%. Dari total 500 data ulasan yang kami gunakan, terdapat 418 data ulasan positif yang diprediksi dengan benar oleh NB, terdapat 180 data ulasan negatif dari pengguna yang terprediksi sebagai positif, dan 82 data ulasan positif yang terprediksi sebagai negatif, sementara 340 data ulasan negatif yang sesuai dengan prediksi negatif. *Confusion matrix* yang terbentuk dari proses pengolahan data ulasan aplikasi memberikan penilaian terhadap kinerja klasifikasi berdasarkan objek yang benar dan salah. Selain *confusion matrix*, tingkat akurasi juga diperkuat dengan kurva ROC. Berikut ini adalah tampilan kurva ROC dari algoritma NB:



Gambar 6. ROC curve dari algoritma NB

Kurva ROC NB yang dihasilkan dari pengolahan dataset ulasan aplikasi menunjukkan nilai AUC sebesar 0,745 dengan margin +/- 0,136 (*micro average*: 0,745) (*Positive class*: *Negatif*). Berdasarkan penjelasan di atas, dapat disimpulkan bahwa pengolahan data menggunakan metode NB pada ulasan aplikasi OneDrive menghasilkan hasil sebagai berikut:

Tabel 3. Nilai hasil Accuracy, Recall, Precision dan AUC

	Accuracy	Recall	Precision	AUC
NB	75,79%	67,88%	80,92%	0,745

4. KESIMPULAN

Dapat disimpulkan bahwa pengolahan data menggunakan metode SVM dan NB menunjukkan bahwa metode SVM menghasilkan nilai akurasi terbaik dibandingkan dengan metode NB, dengan akurasi sebesar 83,60%. Hal ini karena metode SVM mampu mengenali pola distribusi kata dalam kalimat berdasarkan data komentar yang diolah menggunakan *RapidMiner Analytics Platform*, yang digunakan untuk menentukan apakah komentar tersebut positif atau negatif. Sementara itu, akurasi terendah diperoleh dengan menggunakan algoritma *Naive Bayes*, karena metode ini didasarkan pada probabilitas kemunculan kata.

REFERENSI

1. Jalinur J, Yudisman SN. Penyimpanan Arsip Digital Berbasis Cloud Sebagai Mitigasi Bencana di Indonesia. JIPIS : Jurnal Ilmu Perpustakaan dan Informasi Islam. 2023;2(2).
2. Arswanda MAP, Caesar C, Sihombing J, Andri A, Prima Laia A. EVALUASI PENGALAMAN MAHASISWA MIKROSKIL PADA APLIKASI ONEDRIVE MENGGUNAKAN UEQ. SIFO Mikroskil. 2022;23(2622–8130):1–5.
3. Viski Izabal S, Aknuranda I, Az-Zahra HM. Evaluasi dan Perbaikan User Experience Menggunakan User Experience Questionnaire (UEQ) dan Focus Group Discussion (FGD) pada Situs Web FILKOM Apps Mahasiswa Fakultas Ilmu Komputer Universitas Brawijaya. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer [Internet]. 2018;2(9):3224–32. Available from: <http://j-ptiik.ub.ac.id>
4. Oktafani M, Prasetyaningrum TP. IMPLEMENTASI SUPPORT VECTOR MACHINE UNTUK ANALISIS SENTIMEN KOMENTAR APLIKASI TANDA TANGAN DIGITAL. Jurnal Sistem Informasi Dan Bisnis Cerdas (SIBC). 2022;15(1):10–9.
5. Herlinawati N, Yuliani Y, Faizah S, Gata W, Samudi S. ANALISIS SENTIMEN ZOOM CLOUD MEETINGS DI PLAY STORE MENGGUNAKAN NAÏVE BAYES DAN SUPPORT VECTOR MACHINE. CESS (Journal of Computer Engineering System and Science). 2020;5(2):2502–714.
6. Ginabila G, Fauzi A. Analisis Sentimen Terhadap Pemutar Musik Online Spotify Dengan Algoritma Naive Bayes dan Support Vector Machine. Jurnal Ilmiah ILKOMINFO - Jurnal Ilmu Komputer dan Informatika. 2023;6(2):2621–4962.
7. Khofifah W, Rahayu DN, Yusuf AM. Analisis Sentimen Menggunakan Naive Bayes Untuk Melihat Review Masyarakat Terhadap Tempat Wisata Pantai Di Kabupaten Karawang Pada Ulasan Google Maps. Jurnal Interkom: Jurnal Publikasi Ilmiah Bidang Teknologi Informasi dan Komunikasi. 2022 Jan 29;16(4):28–38.
8. Gunawan F, Fauzi MA, Adikara PP. ANALISIS SENTIMEN PADA ULASAN APLIKASI MOBILE MENGGUNAKAN NAIVE BAYES DAN NORMALISASI KATA BERBASIS LEVENSHTAIN DISTANCE (STUDI KASUS APLIKASI BCA MOBILE). Systemic. 2017;03(02):1–6.
9. Sari EDN, Irhamah I. Analisis Sentimen Nasabah Pada Layanan Perbankan Menggunakan Metode Regresi Logistik Biner, Naive Bayes Classifier (NBC), dan Support Vector Machine (SVM). JURNAL SAINS DAN SENI ITS. 2019;8(2):2337–3520.
10. Santoso I, Gata W, Paryanti AB. Penggunaan Feature Selection di Algoritma Support Vector Machine untuk Sentimen Analisis Komisi Pemilihan Umum. Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi). 2017;1(3):364–70.
11. Sitorus MR. PENGGUNAAN TEKNOLOGI KOMPUTASI KANTOR BERBASIS CLOUD UNTUK KOLABORASI TIM YANG EFISIEN. 2024.
12. Hakim ZR, Sugiyono S. Analisa Sentimen Terhadap Kereta Cepat Jakarta – Bandung Menggunakan Algoritma Naive Bayes Dan K-Nearest Neighbor. Jurnal Sains dan Teknologi. 2024 Apr 2;5(3):939–45.
13. Pratama AD, Hendry H. ANALISA SENTIMEN MASYARAKAT TERHADAP PENGGUNAAN CHATGPT MENGGUNAKAN METODE SUPPORT VECTOR MACHINE (SVM). JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika). 2024 Feb 23;9(1):327–38.
14. Aji S, Warjiyono W, Pratmanto D, Ardiansyah A, Widodo AE, Faqih H, et al. Review Sentiment Analysis of World Class Hotel Using Naive Bayes Classifier And Particle Swarm Optimization Method. In: ICCSET. European Alliance for Innovation n.o.; 2019.
15. Pratmanto D, Rousyati R, Wati FF, Widodo AE, Suleman S, Wijianto R. App Review Sentiment Analysis Shopee Application in Google Play Store Using Naive Bayes Algorithm. In: Journal of Physics: Conference Series. IOP Publishing Ltd; 2020.