
Perbandingan Algoritma Klasifikasi Analisis Sentimen Pengguna Aplikasi Getcontact Dalam Pencegahan Penipuan Online

Hermanto^{1*}, Riza Fahlapi², Antonius Yadi Kuntoro³, Taufik Asra⁴

¹Teknologi Informasi, Politeknik Aisyiyah Pontianak, Jl. Ampera, Sungai Jawi, Kec. Pontianak Kota, Kota Pontianak, Kalimantan Barat, 78114, Indonesia

²Teknologi Informasi, Universitas Bina Sarana Informatika, Jl. Kramat Raya No.98, RT.2/RW.9, Kwitang, Kec. Senen, Kota Jakarta Pusat, Daerah Khusus Ibukota Jakarta, 10450, Indonesia

³Sistem Informasi, Universitas Nusa Mandiri, Jl. Jatiwaringin No. 2, Cipinang Melayu, Makasar Jakarta Timur, 13620, Indonesia

⁴Rekayasa Perangkat Lunak, Universitas Bina Sarana Informatika, Jl. Kramat Raya No.98, RT.2/RW.9, Kwitang, Kec. Senen, Kota Jakarta Pusat, Daerah Khusus Ibukota Jakarta, 10450, Indonesia

***Email Korespondensi:**
hermanto.hm@polita.ac.id

Abstrak

Penipuan online merujuk pada berbagai tindakan penipuan yang dilakukan melalui internet dengan tujuan untuk mendapatkan keuntungan finansial atau informasi pribadi secara curang. Kewaspadaan akan pentingnya keamanan untuk diri sendiri dan orang-orang yang kita kenal, perlu terus kita sebar luaskan dimana saat ini banyak terjadi modus penipuan online yang sangat beragam. Salah satu aplikasi yang cukup dikenal oleh masyarakat diantaranya seperti aplikasi GetContact adalah sebuah aplikasi yang dirancang untuk memberikan informasi tentang panggilan yang masuk, identifikasi panggilan spam atau penipuan, dan menyediakan layanan berkaitan dengan daftar kontak telepon yang telah didaftarkan juga oleh sesama pengguna aplikasi tersebut. Dalam penelitian ini, peneliti akan menganalisis sentimen komentar pengguna aplikasi Getcontact dengan membandingkan hasil pengujian terhadap algoritma klasifikasi yaitu Naïve Bayes Classifier dan SVM. Proses penelitian ini akan dimulai dengan pengambilan sampel data menggunakan teknik scrapping pada google playstore dan data komentar pengguna aplikasi Getcontact proses menggunakan RapidMiner. Setelah proses preprocessing dan pengujian model dengan dua metode textmining dengan menggunakan algoritma yaitu SVM dan Naive Bayes, hasil evaluasi dan validasi menunjukkan bahwa Naive Bayes memiliki tingkat akurasi yang lebih tinggi dari pada SVM. Untuk Naive Bayes, nilai akurasi mencapai 82.97% dengan nilai AUC sebesar 0.500, sedangkan untuk SVM, nilai akurasi adalah 78,00% dengan nilai AUC sebesar 0.926. Hasil ini menunjukkan bahwa Naive Bayes lebih unggul dalam mengklasifikasikan komentar pengguna pada aplikasi Getcontact di google play sebagai komentar positif dan negative.

Kata Kunci: *GetContact; Naive Bayes Classifier; Penipuan Online; Textmining; Support Vektor Machine*

Abstract

Online fraud refers to various fraudulent acts carried out over the internet with the aim of fraudulently obtaining financial gain or personal information. We need to continue to spread awareness about the importance of security for ourselves and the people we know, where currently there are many different modes of online fraud. One application that is well known to the public is the GetContact application, which is an application designed to provide information about incoming calls, identify spam or fraudulent calls, and provide services related to a list of telephone contacts that have been registered by fellow users of the application. In this research, researchers will analyze the sentiment of comments from users of the Getcontact application by comparing the test results of classification algorithms, namely Naïve Bayes Classifier and SVM. This research process will begin with data sampling using the scrapping technique on Google Playstore and processing data from users of the Getcontact application using RapidMiner. After the preprocessing process and model testing

with two textmining methods using algorithms, namely SVM and Naive Bayes, the evaluation and validation results show that Naive Bayes has a higher level of accuracy than SVM. For Naive Bayes, the accuracy value reached 82.97% with an AUC value of 0.500, while for SVM, the accuracy value was 78.00% with an AUC value of 0.926. These results show that Naive Bayes is superior in classifying user comments on the Getcontact application on Google Play as positive and negative comments.

Keywords: *GetContact; Naive Bayes Classifier; Online Fraud; Textmining; Support Vector Machine*

1. Pendahuluan

Penipuan *online* merujuk pada berbagai tindakan penipuan yang dilakukan melalui internet dengan tujuan untuk mendapatkan keuntungan finansial atau informasi pribadi secara curang. Kewaspadaan akan pentingnya keamanan untuk diri sendiri dan orang-orang yang kita kenal, perlu terus kita sebar luaskan dimana saat ini banyak terkadi modus Penipuan *online* yang sangat beragam, contoh yang terjadi belakangan ini adalah *phishing*, para pelaku penipu dengan segala Upaya mencoba untuk dapat memperoleh informasi rahasia personal dan sensitif seperti: tanggal lahir, data keuangan dan kata sandi yang dilakukan dengan menyamar sebagai entitas yang tepercaya baik secara lansung maupun melalui media sosial, ragam keahlian pada pelaku mulai dari penjualan barang (palsu) atau penawaran investasi (bodong). Penipuan online yang hampir jarang disadari saat ini juga bisa mencakup penipuan bermodus cinta, pelakunya berpura-pura menjadi pasangan interaktif, baik dan perhatian dengan tujuan untuk memeras korban yang telah sebelumnya memberikan informasi pribadinya (Alfauzi et al., 2019; Noor Rahmad, 2019). Penipuan semacam ini sangat merugikan individu serta berdampak materi yang juga psikologis korban, ancaman bisnis, dan organisasi, untuk menghindarinya, sangat penting bagi setiap individu untuk dapat meningkatkan kesadaran risiko ini, berperilaku secara sadar untuk dapat memahami taktik penipuan, dan selalu menjaga kehati-hatian saat berinteraksi atau bertransaksi secara *online*.

Pada umumnya para pelaku penipuan dapat menjadi lembaga resmi atau pribadi yang terlihat baik dan familiar atau pihak yang sah, seperti lembaga pemerintah, keuangan atau perusahaan terkemuka lainnya dan akan meminta informasi pribadi seperti nomor telepon, rekening atau kartu kredit, dan yang paling dikhawatirkan adalah tanpa disadari korban memberikan kode OTP (*One-Time Password*) yang didukung informasi bank. Teknik intimidasi atau pemaksaan yang sering pelaku gunakan untuk tak sedikit membuat korban takut dan terpaksa kooperatif untuk mengikti apa yang diperitahka oleh pelaku. Pada awalnya pelaku penipua dapat berpura-pura menjadi penjual yang kompeten dan memiliki legalitas, menawarkan produk terbaru atau layanan yang sangat menarik dengan iming-iming harga murah dan memiliki prosepek yang menguntungkan, namun setelah terjadi proses pembayaran yang telah dilakukan oleh korban, pelaku menghilang tanpa memberikan barang dan/atau layanan yang diperjanjikan. Ada juga pelaku yang memiliki modus dengan menghubungi korban untuk memberi tahu korban dengan modus bahwasanya korban telah memenangkan hadiah atau kontes, tetapi perlu membayar sejumlah uang atau memberikan informasi pribadi terlebih dahulu untuk dapat mengklaim hadiah atau kontes yang dimenangkan.

Salah satu aplikasi yang cukup dikenal oleh masyarakat diantaranya seperti aplikasi GetContact adalah sebuah aplikasi yang dirancang untuk memberikan informasi tentang panggilan yang masuk, identifikasi panggilan spam atau penipuan, dan menyediakan layanan berkaitan dengan daftar kontak telepon yang telah didaftarkan juga oleh sesama pengguna aplikasi tersebut. Selain itu pengguna dapat mengidentifikasi panggilan *spam* atau penipuan secara otomatis setelah meyetujui aktifitar panggilan untuk dapat dikontrol oleh aplikasi tersebut. Ketika seseorang menerima panggilan dari nomor yang telah diidentifikasi sebagai spam oleh komunitas yang juga pengguna GetContact, aplikasi ini dapat memberi tahu pengguna dan memperingatkan pengguna berupa informasi nomor telepon yang dapat digunakan untuk memberikan peringatan kepada pengguna lain ketika nomor-nomor ini muncul dalam panggilan masuk pada perangkat telepon. Guna melindungi diri dari pelaku penipuan *online* lewat telepon, sangat penting untuk tidak memberikan informasi data pribadi maupun data yang dapat berhubugan dengan transaksi keuangan kepada pihak yang tidak dikenal, baru dikenal atau yang meragukan (Rusmana, 2015). Pastikan selalu melakukan verifikasi terkait identitas penelepon, hindari mengklik tautan atau mengunduh aplikasi yang dikirimka

melalui sumber yang tidak dikenal dan di percaya serta pertimbangkan untuk megaktifkan dan menggunakan aplikasi keamanan yang dapat membantu identifikasi dan memblokir panggilan spam atau beridikasi penipuan. Kesadaran, kewaspadaa dan kehati-hatian merupakan kunci dalam mencegah terjadinya penipuan. Textmining adalah proses data yang diperlukan dari kumpulan data teks. ekstraksi mencakup pemrosesan teks secara otomatis untuk mengidentifikasi, analisis pola dan tren dari informasi dokumen atau teks yang besar dan beragam. Banyak algoritma yang dapat digunakan untuk klasifikasi teks, analisis sentimen, pengelompokan teks pada proses ekstraksi data teks dalam konteks *textmining*(Herianto Herianto, 2018).

Naive Bayes adalah salah satu algoritma klasifikasi *textmining*. Naïve Bayes menghitung probabilitas sebuah dokumen atau teks termasuk dalam suatu kelas berdasarkan kemunculan kata-kata atau fitur-fitur tertentu dalam teks. Algoritma Naïve Bayes memiliki asumsi "naif" bahwa semua kata atau fitur dalam teks adalah independen satu sama lain, meskipun dalam suatu kalimat, kata-kata sering kali berkaitan satu sama lain. Meskipun demikian, pada proses *textmining* Naive Bayes setelah dilakukan klasifikasi akan hasil yang tinggi dalam klasifikasi teks, terutama pada proses klasifikasi kategori spam email atau analisis sentimen ulasan pada suatu produk(Hermanto et al., 2020) .

Support Vector Machine (SVM) dalam *textmining* juga digunakan secara luas, dalam proses klasifikasi teks. SVM berfokus pada pencarian *hyperplane* terbaik yang memisahkan kalimat-kalimat dari berbagai kelas. Dalam konteks *textmining*, perbedaharaan kalimat direpresentasikan sebagai vektor fitur, dengan setiap kata atau fitur yang mungkin dalam teks menjadi dimensi dalam vektor. SVM berusaha untuk menemukan *hyperplane* yang memisahkan kalimat-kalimat dari berbagai kelas dengan jarak maksimal antara *hyperplane* dan kalimat-kalimat pendukungnya. Ini memungkinkan SVM untuk mengatasi data teks yang rumit dan tidak linier dengan baik(Hermanto et al., 2022, 2020; Kuntoro et al., 2022) .

Penelitian yang dilakukan oleh Astrid Noviriandini dalam jurnal bertajuk “Klasifikasi Support Vector Machine Berdasarkan *Particle Swarm Optimization* (PSO) untuk Analisis Sentimen Pengguna Aplikasi PeduliLindungi” mengklasifikasikan sentimen terhadap review pengguna aplikasi PeduliLindungi saat ini Bernama Satu Sehat. Tingkat akurasi yang ditunjukkan oleh algoritma ini cederug tinggi menunjukkan, sehingga menunjukan bahwa metode ini dapat menjadi pendekatan yang baik dan dapat dipercaya untuk menganalisis dan mengklasifikasikan sentimen berdasarkan masukan pengguna dari aplikasi tersebut, yang pada akhirnya dapat membantu pengguna mengambil keputusan terkait pemilihan aplikasi yang dipergunakan. Hasil pengujian diperoleh dengan nilai akurasi = 93,0% dan AUC = 0,977. Oleh karena itu, implementasi *Support Vector Machine* (SVM) berbasis PSO(Noviriandini et al., 2022).

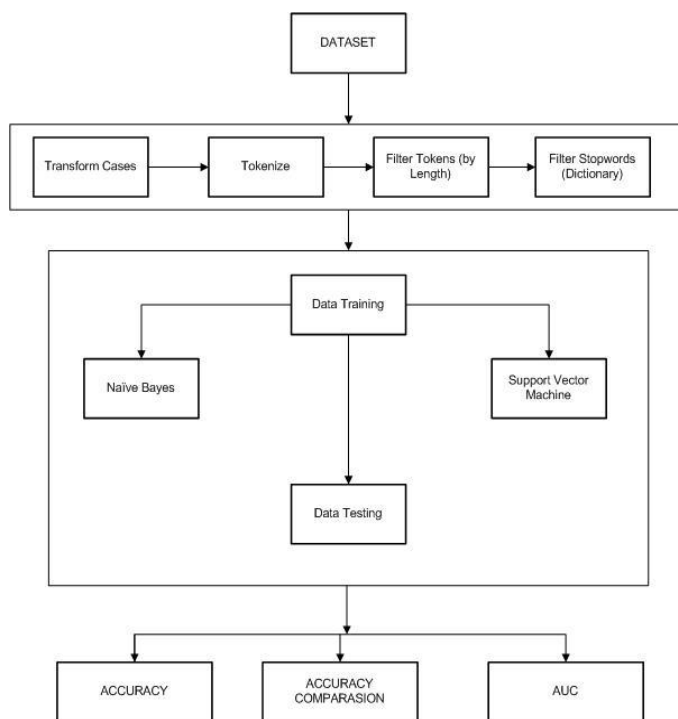
Penelitian lain yang juga dilakukan oleh Hermanto dalam jurnal berjudul “Analisis tingkat kepuasan pengguna aplikasi JMO menggunakan algoritma Naïve Bayes Classifier. Melalui penelitian ini, hasil analisis kepuasan pengguna aplikasi JMO memberikan kontribusi positif terhadap penggunaanya(Noviriandini et al., 2023). Tingkat akurasi yang tinggi menunjukkan potensi algoritma ini sebagai alat yang efektif dalam memberikan mengevaluasi sentimen dan tingkat kepuasan pengguna aplikasi, tentunya dapat menjadi informasi berharga bagi pengembang dan penyelenggara aplikasi untuk meningkatkan kualitas dan pengalaman pengguna hasil akurasi 86,27 %. Selain itu ada juga penelitian menggunakan algoritma naïve bayes dengan mengembangkan sebuah sistem rekomendasi jurusan berbasis web menggunakan algoritma Naive Bayes Gaussian guna mengatasi masalah kebingungan siswa dalam memilih jurusan di STIKI Malang(Perkasa & Eka Purwiantono, 2023). Terdapat keterbatasan informasi dan bimbingan karir yang membuat siswa lulusan SMA/K sulit menentukan pilihan jurusan yang sesuai dengan minat dan potensinya. Dalam penelitian ini, data latih dari 107 mahasiswa aktif dan lulusan STIKI Malang digunakan untuk memberikan rekomendasi berdasarkan berbagai atribut seperti jenis kelamin, jurusan, skill, hobi, alasan kuliah, alasan memilih program studi, ketertarikan terhadap matematika, dan ketertarikan terhadap bahasa Inggris. Metode Naive Bayes Gaussian berhasil mengklasifikasikan data dengan akurasi mencapai 87,85%, yang dapat mengatasi masalah ketidakpastian dalam pemilihan jurusan(Hermanto et al., 2019, 2020).

Dalam konteks penelitian ini, analisis sentimen komentar pengguna aplikasi Getcontact dengan membandingkan hasil pengujian algoritma klasifikasi yaitu Naïve Bayes Classifier (NB) dan Support Vector

Machine (SVM). Proses analisis akan dimulai dengan pengambilan data sampel menggunakan teknik *scrapping* di Google Playstore serta pengolahan data *feedback* pengguna pada aplikasi Getcontact menggunakan RapidMiner yang melibatkan pembuatan model dari data awal untuk pemisahan data menjadi data sampel dan data uji. Hasil analisis akan diukur secara presisi menggunakan matriks konfusi (*confusion matrix*), serta metrik lain seperti Presisi dan AUC yang akan direpresentasikan melalui kurva.

2. Metode Penelitian

Penelitian ini menggunakan data ulasan pengguna dari aplikasi Getcontact di Google Play untuk mengevaluasi status ulasan-ulasan pengguna di platform Google. Penelitian ini menggunakan data komentar dengan total 360 komentar pengguna aplikasi dalam 3 (tiga) bulan terakhir. Penelitian ini bertujuan untuk membandingkan 2 (dua) algoritma yang digunakan, yaitu menggunakan algoritma NB dan SVM. Tujuan utamanya adalah untuk membandingkan kinerja kedua algoritma ini dalam hal akurasi dan nilai Area Under the Curve (AUC). Hasil pengujian ini akan diintegrasikan ke dalam sistem informasi sehingga dapat mengkategorikan komentar pengguna aplikasi Getcontact, menentukan apakah komentar tersebut positif atau negatif. Adapun metode yang diusulkan dapat digambarkan dalam kerangka refleksi sebagai berikut:

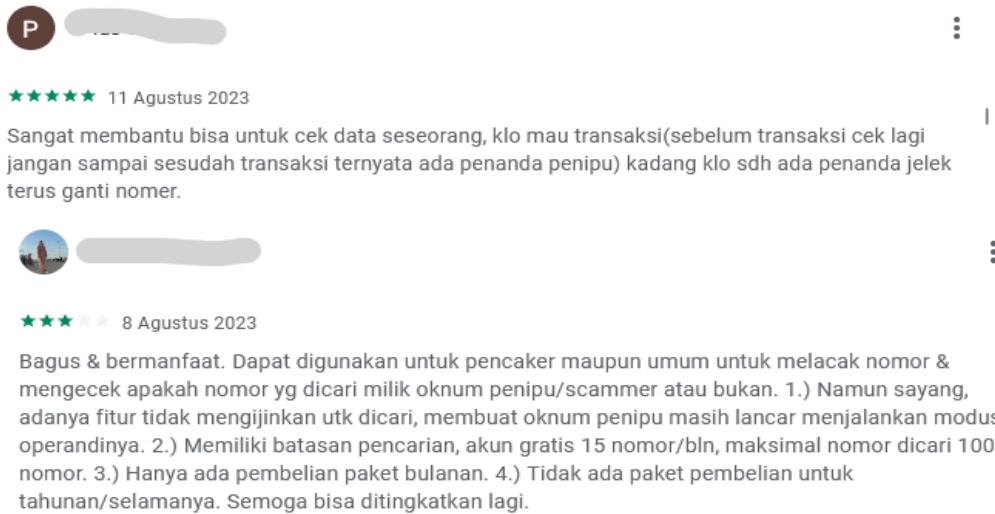


Gambar 1. Kerangka Pikir

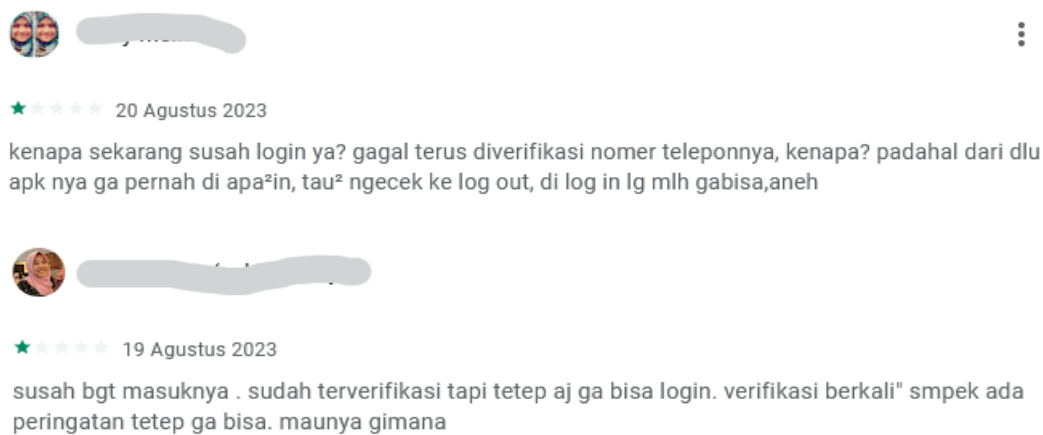
3. Hasil

Penelitian ini menggunakan data yang diperoleh dari ulasan pengguna aplikasi Getcontact di Google Play yang berisi komentar mengenai penggunaan aplikasi ini. Data ini digunakan oleh penulis untuk membuat kumpulan data yang membedakan komentar positif dan negatif. Tujuan dari penelitian ini adalah melakukan analisis sentimen untuk memahami pandangan pengguna mengenai penggunaan aplikasi Getcontact, khususnya dalam hal pencegahan penipuan *online* melalui metode klasifikasi textmining akan diterapkan pada penelitian ini.

Berikut contoh komentar pengguna aplikasi Getcontact pada google play:



Gambar 2 : Komentar Positif



Gambar 3. Komentar Negatif

Tahap *Understanding* pemahaman data merupakan langkah awal dalam proses penelitian ini, dimana informasi yang akan digunakan untuk penelitian lebih dipahami. Data dikumpulkan dari *review* pengguna aplikasi Getcontact di Google Play menggunakan teknik data *scraping* dan disimpan dalam format file Excel. Sumber data utama penelitian ini adalah komentar dari pengguna aplikasi GetConctact di Google Play, dengan total 360 data. Dari komentar-komentar tersebut, 202 komentar negatif dan 158 komentar positif, sehingga menciptakan keseimbangan yang diperlukan untuk penelitian *textmining*. Seluruh data komentar baik positif maupun negatif telah dikelompokkan dan disimpan dalam format .xls.

Tahap *data preparation* adalah langkah awal dalam proses pengolahan data awal (*preprocessing*). *Preprocessing* yaitu tahapan memproses data atau dokumen. Dalam tahapan *preprocessing* terdapat beberapa bagian yaitu *cleaning*, *case folding*, *parsing/tekonizing*, *filtering* kemudian tahapan *stemming* untuk mendapatkan kata dasar yang akan diklasifikasikan.

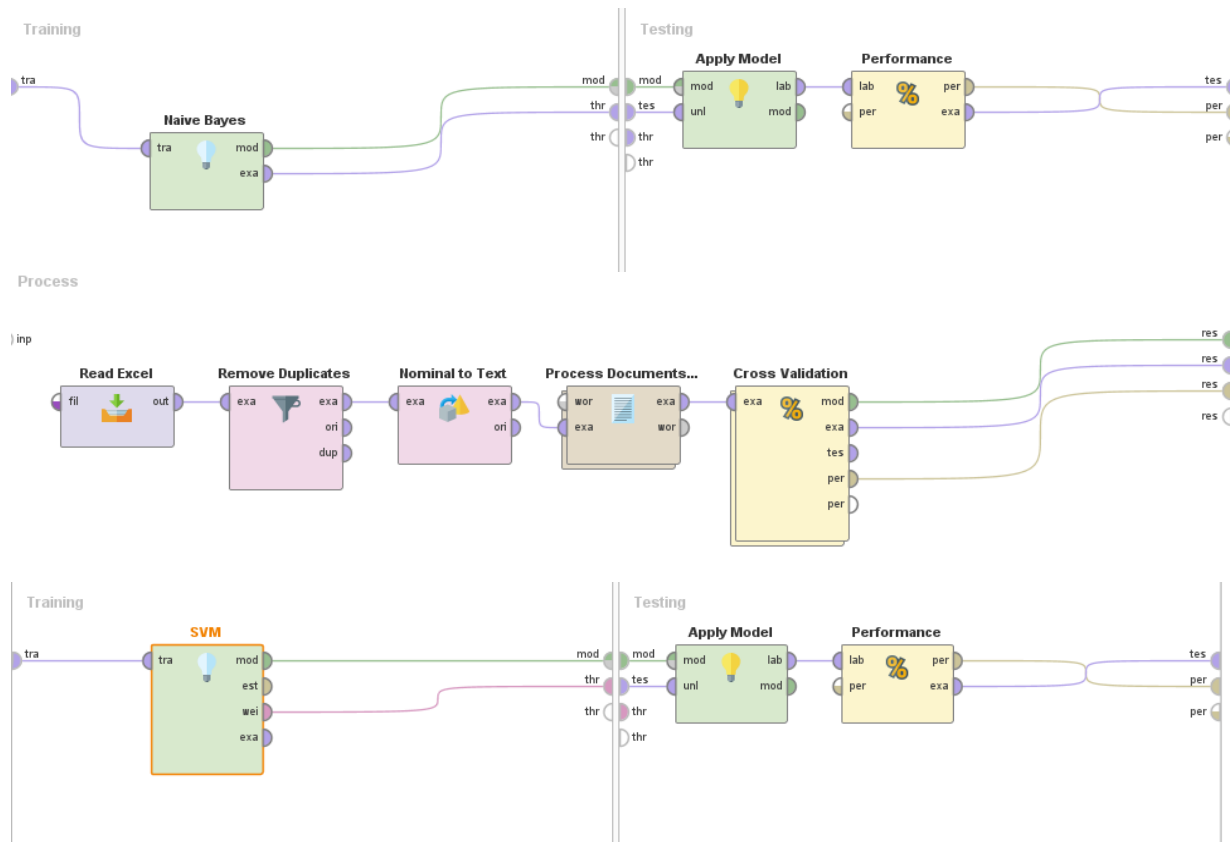
Proses himpunan data ini penting untuk menjaga integritas dan kebersihan data. Diperlukan ketika terdapat entri atau data dengan nilai atau karakteristik yang sama atau mirip, yang jika dibiarkan dapat mengganggu

akurasi dan efektivitas analisis data. Tindakan ini untuk memastikan konsistensi dataset, mengoptimalkan hasil analisis dan meningkatkan kehandalan informasi yang dihasilkan.

Tahap *nominal to text* melibatkan konversi atau transformasi variabel nominal ke dalam bentuk teks atau kata-kata yang lebih dapat dipahami. Ini berguna dalam analisis data atau pemodelan karena variabel nominal perlu diubah menjadi bentuk teks agar hasil analisis dapat lebih mudah diinterpretasikan. Tahap *Transform Case* adalah proses mengubah karakter dalam teks menjadi huruf besar, huruf kecil, huruf kapitalisasi awal, atau kombinasi huruf besar dan kecil.

Proses penghapusan *stopword* dari teks melibatkan penggunaan kamus atau daftar kata-kata *stopword* yang telah ditentukan sebelumnya. Tindakan ini membantu membersihkan teks dari kata-kata umum dan sering muncul, meningkatkan kualitas analisis dengan memfokuskan pada kata-kata yang lebih berarti dan penting. Dengan melakukan langkah-langkah ini, data menjadi lebih siap untuk digunakan dalam analisis selanjutnya, dengan menghilangkan duplikasi dan mengubah variabel nominal menjadi bentuk yang lebih mudah dimengerti.

Dalam penelitian ini, komentar pengguna terkait kepuasan penggunaan aplikasi Getcontact di Google Play Store dianalisis menggunakan algoritma NB dan SVM framework Rapid Miner. Desain crawling dataset dan analisis sentimen pada Rapid Miner dijelaskan dalam gambar yang tidak ditampilkan dalam kalimat ini:



Gambar 4. Desain Model Algoritma NB dan SVM

4. Pembahasan

Pada pengujian ini, digunakan data komentar pengguna aplikasi Getcontact yang telah melalui proses *preprocessing*. Data komentar diambil dari operator menggunakan fitur *Read Excel* karena format data harus berupa file Excel (.xlsx). Proses selanjutnya melibatkan konversi file menjadi dokumen melalui *Process Documents from Files*. Tahap validasi terdiri dari data *training* dan data *testing*.

Selanjutnya, model algoritma NB dan SVM digunakan untuk melakukan perhitungan pada data yang telah dilabeli dalam dua kategori, yaitu komentar positif dan negatif. Setelah modelnya dijalankan, proses akan menghasilkan nilai akurasi dan AUC sebagai hasil dari analisis data tersebut.

accuracy: 82.97% +/- 6.39% (micro average: 82.92%)

	true Negatif	true positif	class precision
pred. Negatif	170	28	85.86%
pred. positif	34	131	79.39%
class recall	83.33%	82.39%	

Gambar 5. Confusion Matrix Naïve Bayes

$$Acc (Accuracy) = \frac{TP + TN}{TP + TN + FP + FN} + \frac{131 + 170}{131 + 170 + 28 + 34} = \frac{301}{363} = 0,8297$$

Berdasarkan hasil analisis data, terdapat 131 *record* yang benar-benar diklasifikasikan sebagai Positif (TP), terdapat 170 *record* yang benar-benar diklasifikasikan sebagai Negatif (TN), 28 *record* yang salah diklasifikasikan sebagai *False* Positif (FP), 34 *record* yang benar-benar diklasifikasikan sebagai *False* Negatif (FN). Dengan menggunakan algoritma *Naïve Bayes*, tingkat akurasi mencapai 82.97% berdasarkan tabel tersebut.

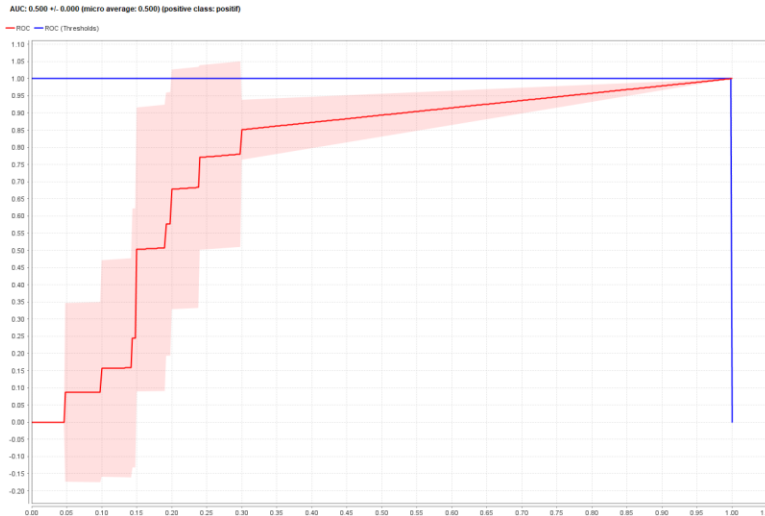
accuracy: 78.00% +/- 6.88% (micro average: 77.96%)

	true Negatif	true positif	class precision
pred. Negatif	126	2	98.44%
pred. positif	78	157	66.81%
class recall	61.76%	98.74%	

Gambar 6. Confusion Matrix SVM

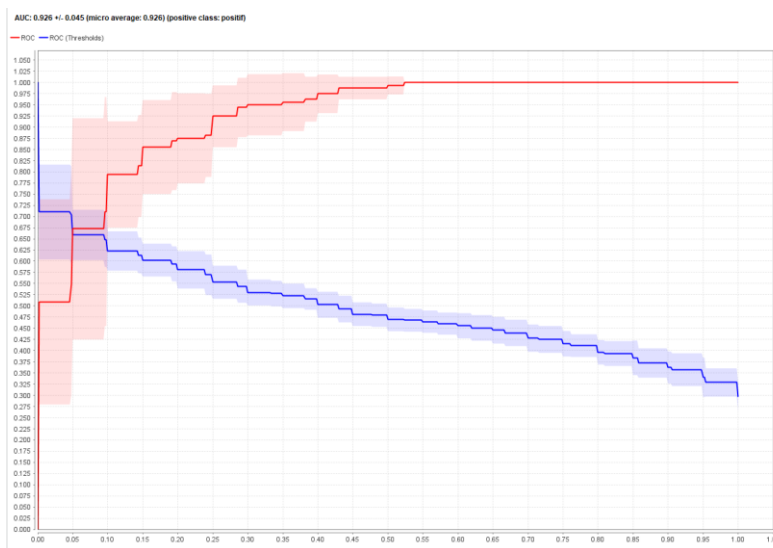
$$Acc (Accuracy) = \frac{TP + TN}{TP + TN + FP + FN} + \frac{157 + 126}{157 + 126 + 2 + 78} = \frac{283}{363} = 0,7797$$

Berdasarkan hasil analisis data, terdapat 157 *record* yang diklasifikasikan sebagai *True Positif* (TP), 126 *record* yang diklasifikasikan sebagai *True Negatif* (TN), 2 *record* yang salah diklasifikasikan sebagai *False Positif* (FP), 78 *record* yang diklasifikasikan sebagai *False Negatif* (FN). Berdasarkan tabel di atas penggunaan algoritma SVM mencapai tingkat akurasi 78.00%.



Gambar 7. Nilai AUC dalam Algoritma NB

Kurva ROC (Receiver Operating Characteristic) dari model Naïve Bayes menunjukkan nilai AUC (Area Under Curve) sebesar 0.500. Nilai AUC ini menunjukkan bahwa hasil diagnosanya dapat diklasifikasikan sebagai "Excellent Classification." Nilai ilai AUC yang mendekati 1.0 dianggap sangat baik, dan nilai 0.500 menunjukkan tingkat performa yang baik didalam klasifikasi data.



Gambar 8. Nilai AUC dalam Algoritma SVM

Kurva ROC (Receiver Operating Characteristic) dari model Naïve Bayes menunjukkan nilai AUC 0.926. Nilai AUC ini menunjukkan bahwa hasil diagnosanya dapat diklasifikasikan sebagai "Excellent Classification". Dalam konteks evaluasi model, nilai AUC yang mendekati 1.0 biasanya dianggap sangat baik, dan nilai 0.926 menunjukkan tingkat performa yang baik dalam mengklasifikasikan data.

Berdasarkan hasil analisis dari masing masing algoritma diatas, maka dapat dirangkum hasilnya seperti tabel berikut:

Tabel 1. Perbandingan Performance Algoritma

	Support Vector Machine	Naive Bayes
Akurasi	78.00%	82.97%
AUC	0.926	0.500

Dari hasil perbandingan performance kedua algoritma diatas NB (82,79% akurasi) dan Algoritma SVM (78,00% akurasi), maka hasil pengujian NB memiliki akurasi yang lebih tinggi daripada Algoritma SVM. Dengan demikian, dapat disimpulkan bahwa Algoritma NB memiliki kinerja yang baik dalam tugas klasifikasi Pada penelitian ini sehingga dapat digunakan untuk memberikan solusi terhadap permasalahan analisis sentimen komentar pengguna aplikasi Getcontact pada google playstore dalam mengklasifikasikan komentar komentar positif dan negatif(Tinaliah & Elizabeth, 2022).

5. Kesimpulan

Pada penelitian ini, setelah melalui proses preprocessing dan pengujian model dengan dua metode text mining menggunakan algoritma yaitu SVM dan NB, hasil evaluasi dan validasi menunjukkan bahwa NB memiliki tingkat akurasi yang lebih tinggi dibandingkan SVM. Untuk NB nilai akurasinya mencapai 82,97% dengan nilai AUC sebesar 0,500, sedangkan untuk SVM nilai akurasinya mencapai 78,00% dengan nilai AUC sebesar 0,926.

Naïve Bayes terbukti unggul dalam mengklasifikasikan komentar pengguna pada aplikasi Getcontact di Google Play menjadi positif dan negatif. Pada penelitian lain yang menggunakan algoritma SVM dan Naïve Bayes dalam mereview data feedback pengguna aplikasi KAI Access, diperoleh hasil bahwa SVM memiliki akurasi sebesar 73,36% dan AUC sebesar 0,794, sedangkan Naïve Bayes memiliki akurasi sebesar 67,10% dan AUC. sebesar 0,573. Oleh karena itu, penerapan Naïve Bayes dalam penelitian ini menunjukkan tingkat akurasi yang lebih tinggi, sehingga berpotensi menjadi solusi untuk klasifikasi sentimen komentar pengguna aplikasi di Google Play.

Untuk penelitian lanjutan, disarankan untuk menggunakan jumlah data yang lebih besar dalam periode lebih dari 3 bulan di Google Play untuk meningkatkan akurasi. Selain itu, penambahan metode klasifikasi lain seperti Algoritma C-45 dan KNN juga dianjurkan untuk membandingkan dengan metode Supervised Learning lainnya

Referensi

- Alfauzi, R. Z., Poerbaningtyas, E., & Oktavia, C. A. (2019). *Parental Control System Aktivitas Santri Berbasis Android (Studi Kasus : Pondok Pesantren Mahasiswa Baitul Jannah).* 07, 97–103.
- Herianto Herianto. (2018). Penerapan Text-Mining Untuk Mengidentifikasi Pengguna Twitter Terhadap Fenomena Peran DPR RI. *JURNAL SAINS & TEKNOLOGI /*, 8(2).
- Hermanto, Antonius Yadi Kuntoro, & Taufik Asra. (2022). Klasifikasi Keluhan Pengguna Kai Access Untuk Pemesanan Tiket Dengan Algoritma Svm Dan Naïve Bayes. *Sinkron*, 6(2).
- Hermanto, H., Mustopa, A., & Kuntoro, A. Y. (2020). Algoritma Klasifikasi Naive Bayes Dan Support Vector Machine Dalam Layanan Komplain Mahasiswa. *JITK (Jurnal Ilmu Pengetahuan Dan Teknologi Komputer)*, 5(2), 211–220. <https://doi.org/10.33480/jitk.v5i2.1181>
- Hermanto, Sandra Jamu Kuryanti, & Siti Nur Khasanah. (2019). Comparison of Naïve Bayes Algorithm, C4.5 and Random Forest for Service Classification OjekOnline. *Journal Publications & Informatics Engineering Research*, 3(2).
- Kuntoro, A. Y., Asra, T., Sistem, S., Fakultas, I., Informatika, T., Mandiri, U. N., Timur, J., Studi, P., Komputer, T., Teknik, F., Bina, U., Informatika, S., Studi, P., Perangkat, R., Fakultas, L., Informatika, T., Bina, U., & Informatika, S. (2022). Klasifikasi Keluhan Pengguna KAI Access Untuk Pemesanan. *JIKA (Jurnal Informatika) Universitas Muhammadiyah Tangerang*, 161–169.

- Noor Rahmad. (2019). Kajian Hukum terhadap Tindak Pidana Penipuan Secara Online. *JURNAL HUKUM EKONOMI SYARIAH*, 3(2).
- Noviriandini, A., Hermanto, H., & Ayu Ambarsari, D. (2023). Analisis Tingkat Kepuasan Pengguna Aplikasi JMO (Jamsostek Mobile) Menggunakan Algoritma Naive Bayes Classifier. *Reputasi: Jurnal Rekayasa Perangkat Lunak*, 4(1), 33–37. <https://doi.org/10.31294/reputasi.v4i1.1986>
- Noviriandini, A., Hermanto, H., & Yudhistira, Y. (2022). Klasifikasi Support Vector Machine Berbasis Particle Swarm Optimization Untuk Analisa Sentimen Pengguna Aplikasi Pedulilindungi. *JIKA (Jurnal Informatika)*, 6(1), 50. <https://doi.org/10.31000/jika.v6i1.5681>
- Perkasa, K. B. P. Y., & Eka Purwiantono, F. (2023). Sistem Rekomendasi Jurusan Menggunakan Algoritma Naïve Bayes Gaussian Berbasis Web. *J-INTECH*, 11(2), 361–370. <https://doi.org/10.32664/j-intech.v11i2.1090>
- Rusmana, A. (2015). Penipuan Dalam Interaksi Melalui Media Sosial (Kasus Peristiwa Penipuan melalui Media Sosial dalam Masyarakat Berjejaring). *Jurnal Kajian Informasi Dan Perpustakaan*, 3(2), 187. <https://doi.org/10.24198/jkip.v3i2.9994>
- Tinaliah, T., & Elizabeth, T. (2022). Analisis Sentimen Ulasan Aplikasi PrimaKu Menggunakan Metode Support Vector Machine. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 9(4), 3436–3442. <https://doi.org/10.35957/jatisi.v9i4.3586>