

Analisis Sentimen Pengguna Aplikasi Shopee Menggunakan Metode Naive Bayes Classifier dan K-NN

Yumarlin MZ^{*1}, Jemmy Edwin Bororing², Sri Rahayu³, Jeffry Andhika Putra⁴

^{1,2,3,4} Program Studi Informatika Fakultas Teknik Universitas Janabadra

E-mail: ^{*1}yumarlin@janabadra.ac.id, ²jemmy@janabadra.ac.id, ³ayu.dj@janabadra.ac.id

⁴jeffry@janabadra.ac.id

Abstrak

Perkembangan aplikasi e-commerce mengalami kemajuan pesat dalam beberapa tahun terakhir. Aplikasi e-commerce memberikan pengalaman belanja yang lebih mudah, nyaman, dan personal bagi pengguna. Fitur-fitur seperti pencarian produk yang efisien, ulasan pelanggan, rekomendasi produk dan keamanan pembayaran. Shopee adalah salah satu platform e-commerce yang populer di Indonesia dan memberikan pengguna akses yang mudah untuk berbelanja secara online dengan berbagai pilihan produk dan penawaran menarik. Tujuan penelitian ini untuk mengetahui analisis sentimen pengguna aplikasi Shopee berdasarkan data ulasan yang di dapat dari situs website google play menggunakan metode Naive Bayes Classifier dan K-Nearest Neighbour (K-NN) untuk mengklasifikasikan ulasan berdasarkan komentar sentimen positif, sentimen negatif dan sentiment netral. Hasil penelitian dengan menerapkan metode Naive Bayes Classifier di dapat nilai akurasi sebesar 75.97%, dengan prediksi komentar positif sebesar 742, komentar negative 519 dan komentar netral sebesar 86. Dan metode K-Nearest Neighbor nilai akurasi sebesar 16.69%, dengan prediksi komentar positif sebesar 154, komentar negative 80 dan komentar netral sebesar 62. Analisis Sentimen aplikasi shopee berdasarkan komentar pengguna google play store menunjukkan tingkat kepuasan konsumen baik di lihat dari besarnya nilai respon komentar positif berdasarkan hasil perhitungan machine learning yang sudah dilakukan.

Kata Kunci : Shopee, Analisis Sentimen, Naive Bayes Classifier, K-NN

1. PENDAHULUAN

Pertumbuhan yang signifikan dalam industri e-commerce telah memicu peningkatan penggunaan aplikasi berbasis e-commerce, termasuk aplikasi Shopee. Shopee merupakan salah satu perusahaan yang bergerak dibidang aplikasi e-commerce mulai beroperasi pada akhir Juni 2015 di Indonesia guna memenuhi permintaan dan gaya hidup masyarakat di Indonesia [1]. Hasil survei Jakpat pada semester pertama di tahun 2022 untuk tren e-commerce Indonesia dengan melibatkan 1.420 responden di seluruh Indonesia, Shopee berhasil mengungguli tujuh platform e-commerce lain [2]. Aplikasi Shopee sebagai platform belanja online telah memberikan manfaat besar bagi konsumen, seiring dengan jumlah pengguna yang semakin meningkat, penting untuk memahami sentimen dan pendapat pelanggan terhadap layanan Shopee. Analisis sentimen telah menjadi pendekatan yang efektif untuk menganalisis data yang berhubungan dengan opini, sikap, dan emosi pelanggan terhadap produk atau layanan.

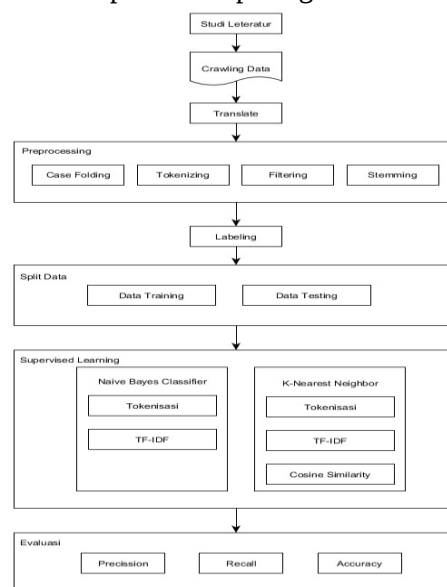
Analisis sentimen merupakan studi komputasi mengenai pendapat, perilaku dan emosi seseorang terhadap entitas, menggunakan pendekatan yang efektif untuk mengidentifikasi dan menganalisis sentimen yang terkandung dalam data teks, termasuk ulasan dan komentar pengguna aplikasi [3]. Kepuasan pengguna dapat didefinisikan sebagai suatu tingkat perasaan seorang pengguna sebagai hasil perbandingan antara harapan pengguna tersebut akan sebuah produk dengan hasil nyata yang diperoleh si pengguna dari produk tersebut [4].

Penelitian untuk Analisis Sentimen yang menggunakan data dari pengguna google play store sudah banyak dilakukan, seperti [5] menganalisis review pengguna aplikasi Grab pada Google Play Store, analisis review pengguna ini menggunakan metode Support Vector Machine (SVM). Hasil dari analisis sentimen menghasilkan akurasi 85,54% dan Hasil Review positif yang paling sering diulas adalah ovo, sedangkan review negatif yang paling sering diulas adalah driver. Penelitian selanjutnya [6] aplikasi Indodax sebagai platform jual beli aset kripto, melakukan analisis sentimen terhadap ulasan pada aplikasi Indodax yang diberikan oleh pengguna di google play store. Menggunakan metode Support Vector Machine yang menghasilkan akurasi sebesar 85% pada rasio pembagian data sebesar 80:20. Sentimen Analisis pada Ulasan Google Play Store menggunakan metode Naive Bayes [7] akan membandingkan teks ulasan yang diberikan oleh pengguna, hasil penelitian ini didapat akurasi setinggi 78,9%, untuk ulasan yang memiliki rating bintang 5 dan bintang 1. Tetapi akurasi menurun mencapai 73,7% untuk semua rating ulasan terkecuali bintang 3.

Naive Bayes Classifier dan K-NN (K-Nearest Neighbor), dua metode pendekatan yang populer dalam analisis sentimen. Naive Bayes Classifier merupakan salah satu metode machine learning, konsep dasar yang digunakan yaitu melakukan klasifikasi dengan perhitungan nilai probabilitas untuk menentukan kategori dari suatu dokumen[8]. K-Nearest Neighbor menggunakan teknik lazy learning dalam melakukan proses klasifikasi terhadap objek berdasarkan data jurnal pembelajaran yang jaraknya paling dekat dengan objek tersebut, termasuk kedalam metode supervised learning [9]. Tujuan dari penelitian ini melakukan analisis sentimen pengguna aplikasi Shopee berdasarkan data ulasan yang di dapat dari situs website google play untuk mengklasifikasikan sentimen pelanggan pada ulasan atau komentar positif, negatif, atau netral dengan menerapkan metode Naive Bayes Classifier dan K-NN.

2. METODE PENELITIAN

Metode penelitian yang digunakan untuk analisis sentimen aplikasi shopee bagi pengguna google play store terdapat 8 tahapan yaitu (1) *Studi Literatur* (2) *Crawling*, (3) *Translate*, (4) *Preprocessing*, (5) *Labeling*, (6) *Split Data*, (7) *Supervised Learning*, (8) *Evaluasi*. Berikut tahapan metode penelitian dapat dilihat pada gambar 1.



Gambar 1. Metode Penelitian

2.1. Studi Literatur

Sebelum melakukan proses penelitian langsung terhadap kondisi yang terjadi, dalam tahapan ini melihat beberapa referensi baik buku, jurnal penelitian, media sosial, portal berita dan lain sebagainya yang akan digunakan untuk penyusunan penelitian ini.

2.2. Crawling

Di dalam tahapan *Crawling*, data mentah yang digunakan diambil dari Google Play Store pada laman aplikasi Shopee dengan teknik *scraping* [10]. Teknik *scraping* pada proses ini dilakukan secara otomatis menggunakan *extention* dari Google Chrome Browser yaitu Data Miner.

2.3. Translate

Setelah proses *crawling* data selesai dilakukan, selanjutnya data komentar yang masih berbahasa Indonesia ditranslasikan ke Bahasa Inggris dengan tujuan untuk memudahkan proses pada tahapan *preprocessing*.

2.4. Preprocessing

Preprocessing adalah proses perubahan bentuk data yang tidak terstruktur menjadi data yang terstruktur sesuai kebutuhan untuk proses dalam *text mining* [11]. Tahapan *Preprocessing* yang digunakan yakni: *Case Folding* yaitu tahap mengubah semua huruf dalam dokumen menjadi huruf kecil dan karakter non-huruf akan dihilangkan, *Tokenizing* yang merupakan tahap pemotongan string input berdasarkan tiap kata yang menyusunnya dan secara garis besar memecah sekumpulan karakter dalam suatu teks ke dalam satuan kata, *Filtering* dengan mengambil kata-kata penting dari hasil tahap *tokenizing* yang dapat dilakukan dengan menghilangkan *stoplist* atau *stopword*, dan *Stemming* yang mentransformasi suatu kata menjadi kata dasarnya (*root word*) dengan menggunakan aturan-aturan tertentu [12].

2.5. Labelling

Di dalam tahapan *Labeling* penulis menggunakan kamus lexicon SentiWordnet untuk menentukan setiap kelas pada setiap komentar. Tahapan ini berjalan dengan cara membuat kamus kata opini (lexicon) terlebih dahulu. Kata-kata yang terdapat pada kamus tersebut akan digunakan untuk mengidentifikasi kata positif dan negatif pada suatu kalimat. Pada tahap ini dilakukan pembagian ke dalam tiga kelas sentimen. Data masuk pada kelas positif jika skor yang didapat > 0 , masuk pada kelas negatif jika skor yang di dapat < 0 dan jika skor $= 0$ masuk pada kelas netral.

2.6. Split Data

Data hasil *Preprocessing* yang dikategorisasi dan diklasifikasikan ke dalam sentimen positif, negatif dan netral, kemudian dibagi menjadi data latih dan data uji pada data yang telah memiliki label kelas dengan perbandingan 80%:20%. Data latih digunakan untuk pembentukan model klasifikasi [13]. Data yang digunakan sebagai data latih sebanyak 80% dari keseluruhan data yang didapatkan setelah proses *Labeling*. Selanjutnya, model klasifikasi yang sudah dilatih akan diuji menggunakan algoritma K-Nearest Neighbor (KNN) untuk mengetahui performa klasifikasi. Data uji yang digunakan sebanyak 20% dari dari keseluruhan data yang didapatkan setelah proses *Labeling*.

2.7. Supervised Learning

Algoritma K-Nearest Neighbor (KNN) digunakan sebagai metode klasifikasi terhadap data latih dan data uji dengan pembobotan menggunakan metode Term Frequency – Inverse Document Frequency (TF-IDF) [14]. Metode TF-IDF merupakan metode untuk menentukan frekuensi relatif dari setiap kata-kata atau token-token yang setiap kata tersebut akan diberikan pembobotan berupa nilai berdasarkan penting atau tidaknya suatu kata dalam dokumen berdasarkan jumlah kemunculan kata dalam suatu dokumen dan mengukur kata-kata tersebut terhadap keseluruhan dokumen yang ada.

2.8. Evaluasi

Proses evaluasi bertujuan untuk mengetahui *accuracy* dari algoritma yang digunakan terhadap data untuk menganalisa sentimen pengguna terhadap aplikasi *Shopee*. Formulasi dari perhitungan nilai *accuracy* yakni :

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \times 100 \quad (1)$$

Hasil dari nilai *accuracy* yang didapat untuk menunjukkan rasio dari dataset yang dapat terklasifikasi dengan benar oleh sistem pada penelitian ini untuk aplikasi *shopee* dari keseluruhan dokumen yang ada.

3. HASIL DAN PEMBAHASAN

3.1. Crawling Data

Data yang digunakan dalam penelitian ini merupakan data ulasan atau review yang diberikan oleh pengguna *Shopee* melalui situs Google Play Store. Data yang disajikan dalam tabel 1 diperoleh dengan menggunakan teknik *scraping* berjumlah 11.212 yang merupakan ulasan dari tanggal 16-18 September 2022.

Tabel 1. Data Ulasan dari Akun *Shopee*

Nama Pengguna	Tanggal	Komentar /Ulasan
Fenie Fenie	16 September 2022	Masih tetep suka sama shopee.. banyak promonya. TAPI PLIIIIIISSSS, JANGAN PELIT KOIN DONG UTK IMBALAN REVIEW NYA.. Masa' dari yg awalnya klo kasih review bagus dpt 100 koin, turun 50 koin, sekarang jd cuma 25 koin. Pliiiiiiisss balikin lagi deh koin gratisnya. Biar lebih banyak lg yg download.
Fariz Pramanthy	17 September 2022	Tolong kalau misal ada list ekspedisi diberi tanda ekspedisi mana, supaya customer bisa pilih, kadang ada beberapa ekspedisi yang pengiriman nya lebih memakan waktu untuk daerah tertentu, dibanding brand sebelah lebih lengkap sebelah keterangan ekspedisi nya, Thanks
Diky saputra	18 September 2022	kenapa ya kak, sekrng gak bisa ubah alamat pengiriman, dulu bisa sekarang gak bisa.. padahal udh di update ke versi baru, kasian entr yg ngantar karna no hp nya salah dan g bisa d ubah.

3.2. Data Translation

Komentar yang sebelumnya berbahasa Indonesia kemudian diubah menggunakan google translate ke dalam bahasa Inggris. Data ulasan yang sudah diubah ke dalam bahasa Inggris dapat dilihat pada tabel 2.

Tabel 2. Data Ulasan Translation

Ulasan Sebelum di translate	Ulasan setelah di translate
Masih tetep suka sama shopee.. banyak promonya. TAPI PLIIIIIISSSS, JANGAN PELIT KOIN DONG UTK IMBALAN REVIEW NYA.. Masa' dari yg awalnya klo kasih review bagus dpt 100 koin, turun 50 koin, sekarang jd cuma 25 koin. Pliiiiiisss balikin lagi deh koin gratisnya. Biar lebih banyak lg yg download.	still like shopee.. lots of promos. but pliiiiiiisss, don't be stingy with coins for review rewards.. from the beginning, if you give a good review, you get 100 coins, down 50 coins, now it's only 25 coins. pliiiiiiisss, return the free coins again. let more people download.
Tolong kalau misal ada list ekspedisi diberi tanda ekspedisi mana, supaya customer bisa pilih, kadang ada beberapa ekspedisi yang pengiriman nya lebih memakan waktu untuk daerah tertentu, dibanding brand sebelah lebih lengkap sebelah keterangan ekspedisi nya, Thanks	Please, if there is a list of expeditions marked with which expedition, so that customers can choose, sometimes there are some expeditions whose delivery takes more time for certain areas, compared to the next brand which is more complete next to the description of the expedition, Thanks
kenapa ya kak, sekrng gak bisa ubah alamat pengiriman, dulu bisa sekarang gak bisa.. padahal udh di update ke versi baru, kasian entr yg ngantar karna no hp nya salah dan g bisa d ubah.	why, sis, now I can't change the shipping address, before now I can't... even though it's been updated to a new version, sorry for the entry that delivered because the cellphone number was wrong and couldn't be changed.

3.3. Preprocessing

Sebanyak 11.212 data yang telah didapatkan dari tahapan *crawling* dan telah ditranslasi, selanjutnya melalui beberapa langkah dalam tahapan *preprocessing* dan menghasilkan 11.083 data.

3.4. Labelling

Setelah proses *preprocessing* data kemudian di ambil data *training* yang telah diklasifikasikan kedalam sentimen (positif, netral dan negatif). Dari proses labeling mendapatkan hasil analisis sentiment berupa sentimen positif sebanyak 5746 data, sentimen negatif sebanyak 2689 data dan hasil sentimen netral sebanyak 432 data.

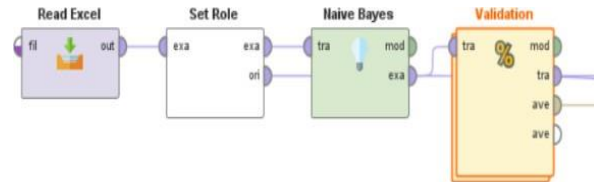
3.5. Split Data

Split data membagi dataset menjadi menggunakan tools Rapid Miner : (1). *Data Training* : *Training* data diperoleh dari hasil *Preprocessing* yang kemudian dikategorikan dan diklasifikasikan kedalam sentimen positif dan negatif. Dimana dari langkah tersebut didapatkan data sebanyak 80% dari 11.212 data yaitu 8.970 data. (2). *Data Testing* : dalam proses ini, data yang sudah di *training* akan diuji menggunakan metode algoritma *Naïve Bayes Classifier* dan *K-Nearest Neighbor* (KNN) kemudian data akan diolah dan dibandingkan sehingga menghasilkan data yang mendapatkan nilai akurasi tinggi yang akan di rekomendasikan. Dimana data yang akan di uji sebanyak 20% dari 11.212 data yaitu 2.242 data.

3.6. Analisis Hasil

3.6.1. Metode Naïve Bayes Classifier

Dengan menggunakan tools Rapid Miner, data yang sudah di Labelling diimplementasikan ke dalam metode naive Bayes Classifier, prosesnya dapat dilihat pada gambar 2, berikut ini.



Gambar 2. Proses Naïve Bayes Classifier

Dari Proses Naive Bayes Classifier di dapat hasil keluaran *true positif*, *true negatif*, dan *true netral* dapat dilihat pada tabel 3, berikut ini.

Tabel 3. Hasil Analisis Sentimen Naive Bayes Classifier

	True Positif	True Negatif	True Netral	Class Precision
Perd.Positif	742	10	0	98.76%
Pred.Negatif	401	519	0	56.41%
Pred.Netral	6	9	86	85.15%
Class recall	64.58%	96.74%	100%	

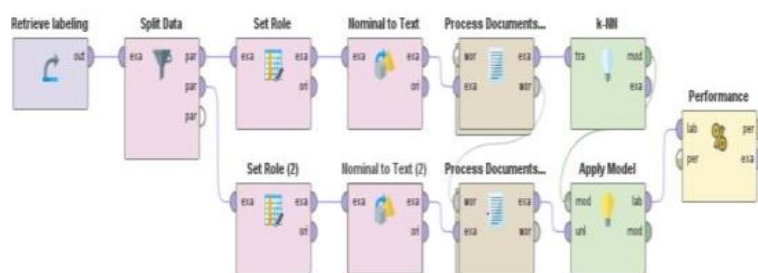
Hasil dari penerapan metode Naive Bayes Classifier pada tabel 3 menunjukkan untuk data ulasan positif machine learning memprediksi komentar positif sebesar 742, komentar negatifnya 401 dan netral sebesar 6, dengan nilai recall sebesar 64.58% dan precision sebesar 98.76%. Untuk prediksi ulasan negative machine learning memprediksi komentar positif sebesar 10, komentar negative 519 dan netralnya sebesar 9, dengan nilai recall sebesar 96.74% dan precision sebesar 56.41%. Sedangkan untuk ulasan netral machine learning memprediksi komentar positif sebesar 0, komentar negatifnya 0 dan komentar netralnya sebesar 86 dengan nilai recall sebesar 100% dan precisionnya sebesar 85.15%. Selanjutnya dilakukan perhitungan nilai *accuracy* yakni :

$$Accuracy = \frac{742 + 519 + 86}{742 + 10 + 0 + 401 + 519 + 0 + 6 + 9 + 86} = \frac{1347}{1773} \times 100 = 75,97$$

Nilai *accuracy* yang di dapat sebesar 75.97% hasil klasifikasi dataset menggunakan metode Naive Bayes Classifier.

3.6.2. K-NN (K- Nearest Neighbor)

Dengan menggunakan tools Rapid Miner, dataset yang sudah siap digunakan di split menjadi data training dan data testing, prosesnya dapat di lihat pada gambar 3, berikut ini.



Gambar 3. Proses Implementasi Metode K-Nearest Neighbor

Dari Proses K-Nearest Neighbor di dapat hasil keluaran *true positif*, *true negatif*, dan *true netral* dapat dilihat pada tabel 4, berikut ini.

Tabel 4. Nilai Analisis Sentimen K-Nearset Neighbor

	True Positif	True Negatif	True Netral	Class Precision
Perd.Positif	154	56	12	69.37%
Pred.Negatif	39	80	12	61.07%
Pred.Netral	956	420	62	4.37%
Class recall	13.40%	14.87%	72.09%	

Hasil dari penerapan metode K- Nearest Neighbor (K-NN) dari tabel 4 menunjukkan untuk data ulasan positif dari machine learning memprediksi komentar positif sebesar 154, komentar negatifnya 39 dan netral sebesar 956, dengan nilai recall sebesar 13.40% dan precisionnya sebesar 69.3%. Untuk ulasan negative machine learning prediksi komentar positif sebesar 56, komentar negative 80 dan netralnya sebesar 420, dengan nilai recall sebesar 14.87% dan precisionnya sebesar 61.07%. Sedangkan untuk ulasan netral machine learning memprediksi komentar positif sebesar 12, komentar negatifnya 12 dan komentar netralnya sebesar 62 dengan nilai recall sebesar 72.09% dan precisionnya sebesar 4.37%. Selanjutnya dilakukan perhitungan nilai *accuracy* dari hasil tabel 4 diatas :

$$Accuracy = \frac{154 + 80 + 62}{154 + 56 + 12 + 39 + 80 + 12 + 956 + 402 + 62} = \frac{296}{1773} \times 100 = 16.69 \%$$

Nilai *accuracy* yang di didapat sebesar 16,69% dari penerapan metode K-NN.

Berdasarkan hasil dari tabel 3 dan tabel 4 untuk penerapan metode dalam mengklasifikasikan data ulasan pengguna Shopee dengan perbandingan data latih dan data uji sebesar 80% : 20% dari data set yang di peroleh dengan hasil klasifikasi sentiment positif sebanyak 5746, negative sebanyak 2689, dan netral sebanyak 432. Tingkat akurasi Naive Bayes Classifier sebesar 75.97%, dan tingkat akurasi pada K-Nearest Neighbor sebesar 16.69%, menunjukkan metode Naive Bayes Classifier memiliki performa yang lebih baik di bandingkan metode K-NN untuk analisis sentimen aplikasi shopee dalam penelitian ini.

4. KESIMPULAN

Berdasarkan hasil dan pembahasan yang dilakukan terhadap pengguna aplikasi Shopee dari Google Play Store, maka dapat ditarik kesimpulan sebagai berikut :

1. Penerapan metode Naive Bayes Classifier dari hasil perhitungan nilai akurasi sebesar 75.97% dengan prediksi komentar positif sebesar 742, komentar negative 519 dan komentar netral sebesar 86.
2. Penerapan metode K-NN (K-Nearest Neighbor) dari hasil perhitungan nilai akurasi mendapatkan nilai sebesar 16.69%, dengan prediksi komentar positif sebesar 154, komentar negative 80 dan komentar netral sebesar 62.
3. Dalam penelitian ini menunjukkan bahwa metode Naive Bayes Classifier berdasarkan nilai akurasi yang di dapat performanya lebih baik dari metode K-Nearset Neighbor.
4. Analisis Sentimen untuk aplikasi shopee dari pengguna google play store menunjukkan respon yang baik untuk komentar positif yang memiliki nilai tertinggi berdasarkan prediksi perhitungan machine learning yang sudah dilakukan.

DAFTAR PUSTAKA

- [1] Fitri, Margaretha. (2017) “Analisis Hubungan Antara Motif Dengan Tingkat Kepuasan Pengguna Aplikasi Shopee Sebagai Media Berbelanja Online Pada Shopeeholics Di Kota Samarinda,” e-Journal Ilmu Komunikasi, 5 (4) 2017: 26–40ISSN (Cetak) 2502-5961, ISSN (Online) 2502-597X,ejournal.ilkom.fisip-unmu l.ac.id.
- [2] Galih A. Palupi, 28 Agustus 2022, Shope Masih Menjadi E-Commerce Pilihan Utama Masyarakat Indonesia, <https://goodstats.id/article/jakpat-shopee-masih-rajai-e-commerce-pilihan-masyarakat-indonesia-tahun-2022-scYdn>
- [3] Dinda. P, S., dan W, Wibowo. (2022) “Analisis Sentimen Ulasan Aplikasi Buzzbreak Menggunakan Metode Naïve Bayes Classifier pada Situs Google Play Store,” Jurnal Sains dan Seni ITS Vol. 11, No. 2, pp. 2337-3520 (2301-928X Print)
- [4] Kotler. P., A. (2018). Marketing Managemet. Global Edition. Pearson.
- [5] Rizki. Wahyudi, dan Gilang, Kusumawardhana. (2021) “Analisis Sentimen pada review Aplikasi Grab di Google Play Store Menggunakan Support Vector Machine,JURNAL INFORMATIKA, Vol. 8 No. 2. ISSN: 2355-6579 | E-ISSN: 2528-2247.
- [6] Alfio. Kusuma, Ermatita, dan H. Nurramdhani, Irmanda. (2022) “Analisis Sentimen Pada Ulasan Aplikasi Indodax di Google Play Store Menggunakan Metode Support Vector Machine,” Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA) Jakarta-Indonesia,. e-ISSN 2962-6129.
- [7] Edyt. Daryfayi. P. D., dan Ibnu, Asror. (2020) “Sentimen Analisis pada Ulasan Google Play Store Menggunakan Metode Naïve Bayes,” e-Proceeding of Engineering : Vol.7, No.2 .pp 8400.ISSN : 2355-9365.
- [8] N. Tri. Romadloni, Imam. Santoso, dan Sularso, Budilaksono. (2019) “Perbandingan Metode Naive Bayes, Knn Dan Decision Tree Terhadap Analisis Sentimen Transportasi Krl Commuter Line,” Jurnal IKRA-ITH Informatika Vol 3 No 2. ISSN 2580-4316.
- [9] Wachid, D. et al. (2023) “Analisis Sentimen Penerapan Kurikulum Merdeka Pada Pengguna Twitter Menggunakan Metode K-Nearest Neighbor Dengan Forward Selection,” Smart Comp, Vol. 12, No. 1. P-ISSN: 2089-676X. E-ISSN: 2549-0796.
- [10] Krotov. V., Johnson. L., dan Silva. L., “Legality and Ethics of Web Scraping,” *Communications of the Association for Information Systems*, vol. 47, pp. 539–563, 2020, doi:org/10.17705/1CAIS.04724.

- [11] Takdirillah, R. (2020) “Penerapan Data Mining Menggunakan Algoritma Apriori Terhadap Data Transaksi Sebagai Pendukung Informasi Strategi Penjualan,” *Edumatic : Jurnal Pendidikan Informatika*, vol. 4(1), pp. 37–46. <https://doi.org/10.29408/edumatic.v4i1.2081>
- [12] Senthikumar, Maheswari. (2019) “Rule Based Morphological Variation Removable Stemming Algorithm,” *International Journal of Recent Technology and Engineering*, 8(4), 1809–1814, <https://doi.org/10.35940/ijrte.C6200.118419>
- [13] Arifiyanti. A. A., dan Wahyuni. E, D. (2020) “SMOTE: Metode Penyeimbang Kelas pada Klasifikasi Data Mining,” *SCAN - Jurnal Teknologi Informasi dan Komunikasi*, 15(1), 34–39. <https://doi.org/10.33005/scan.v15i1.1850>
- [14] Fahrur. Rozi, Farid. Sukmana, M. Nabil, Adani. (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* Vol.6, No.3, P-ISSN : 2502-5716, E-ISSN : 2503-1945