

Terakreditasi SINTA Peringkat 3

Surat Keputusan Direktur Jenderal Pendidikan Tinggi, Riset, dan Teknologi Nomor 225/E/KPT/2022 masa berlaku mulai Vol.7 No. 1 tahun 2022 s.d Vol. 11 No. 2 tahun 2026

Terbit online pada laman web jurnal:
<http://publishing-widyagama.ac.id/ejournal-v2/index.php/jointecs>



Vol. 8 No. 2 (2023) 77 - 84

JOINTECS

(Journal of Information Technology and Computer Science)

e-ISSN:2541-6448

p-ISSN:2541-3619

Sentimen Analisis Aplikasi Belajar Online Menggunakan Klasifikasi SVM

Adi Ariyo Munandar¹, Farikhin², Catur Edi Widodo³

¹Program Studi Sistem Informasi, Sekolah Pascasarjana, Universitas Diponegoro Semarang

²Program Studi Matematika, Fakultas Sains Dan Matematika, Universitas Diponegoro Semarang

³Program Studi Fisika, Fakultas Sains Dan Matematika, Universitas Diponegoro Semarang

¹adiariyomunandar@gmail.com, ²farikhin@lecturer.undip.ac.id, ³caturediwido@lecturer.undip.ac.id

Abstract

Google Play Store is where a wide variety of applications are available, whether paid or not. Google Play Store page is a place for application users to express opinions, reviews and ratings. Ruang Guru, Zenius and Quipper are available on the platform. Analysis was carried out using sentiment analysis and SVM algorithm. Data was obtained using data scraping techniques, using help of google-play-scraper library. Web scraping process is divided into 3 stages namely Fetching, Extraction, and Transformation. Data collected is 30,000 data which is divided into 10,000 data for each application. Research begins with data preprocessing stage which includes normalization, case folding, cleaning, tokenizing, and stopwords. then data is divided into 90% training data and 10% test data. Training data is labeled with values 1, 0, and -1. Value 1 means positive, value 0 means neutral and -1 means negative. Results of classification sentiment using SVM show that Ruang Guru has highest positive value compared to Zenius and Quipper. However, user response equally gives a positive value for application. Accuracy value of research shows that sentiment classification data with SVM has an average accuracy for Ruang Guru of 99%, Zenius of 96%, and Quipper of 82%.

Keywords: Ruang Guru; Zenius; Quipper; Sentiment Analysis; SVM.

Abstrak

Google Play Store adalah tempat berbagai macam aplikasi tersedia, baik berbayar ataupun tidak. Halaman Google Play Store menjadi tempat pengguna aplikasi untuk menyampaikan pendapat, ulasan dan penilaian. Ruang Guru, Zenius dan Quipper tersedia di platform tersebut. Data pada ulasan, menjadi sangat bermanfaat untuk dianalisa. Analisa dilakukan dengan menggunakan sentimen analisis dan algoritma SVM. Data diperoleh dengan menggunakan teknik scraping data, dengan menggunakan bantuan library google-play-scraper. Proses web Scraping, dibagi menjadi 3 tahap yaitu Fetching, Extraction, dan Transformation. Data dikumpulkan sebanyak 30.000 data, yang dibagi menjadi 10.000 data Ruang Guru, Zenius dan Quipper. Penelitian diawali dengan Tahap preproccesing data meliputi normalisasi, case folding, cleaning, tokenizing, dan Stopword. kemudian data dibagi menjadi 90% data latih dan 10% data uji. Data latih diberi label dengan nilai 1, 0, dan -1. Nilai 1 berarti positif, nilai 0 berarti netral dan -1 berarti negatif. Hasil sentimen klasifikasi menggunakan SVM, menunjukkan bahwa Ruang Guru memiliki nilai positif tertinggi dibandingkan Zenius dan Quipper. Akan tetapi, respon pengguna sama-sama memberikan nilai positif untuk aplikasi tersebut. Nilai akurasi dari penelitian menunjukkan bahwa, data Klasifikasi sentimen dengan SVM, mempunyai akurasi rata-rata untuk Ruang Guru sebesar 99%, Zenius sebesar 96%, dan Quipper sebesar 82%.

Kata kunci: Ruang Guru; Zenius; Quipper; Sentimen Analisis; SVM.



1. Pendahuluan

Pendidikan merupakan proses perubahan tingkah laku dan cara berfikir seseorang untuk menjadi lebih baik dari sebelumnya. Pendidikan memiliki peranan yang sangat penting dalam proses perubahan hidup manusia yang didasari pada akal dan pengetahuan yang diberikan. Sehingga melahirkan berbagai macam konsep dan pengetahuan yang mendalam, salah satunya mengenai teknologi. Perkembangan teknologi dan pendidikan saat ini berdampak dan sangat maju dengan pesat, berbagai bentuk inovasi teknologi untuk pendidikan sangat bermanfaat bagi kehidupan manusia. Teknologi dan pendidikan tidak dapat dipisahkan sesuai dengan perkembangan zaman karena teknologi dapat membantu proses perkembangan pada bidang pendidikan [1].

Kemajuan teknologi saat ini, menjadi salah satu alat penunjang kehidupan sehari-hari umat manusia, baik dibidang ekonomi, sosial maupun pendidikan. Perkembangan teknologi dibidang pendidikan ditandai adanya aplikasi belajar *online*, sebagai salah satu sarana untuk memenuhi tuntutan belajar anak dan masyarakat yang masih menempuh pendidikan. Aplikasi belajar *online* dapat memberikan alternatif pilihan bagi siswa yang memiliki akses jaringan internet untuk memperoleh layanan bantuan belajar yang efektif, efisien, dan interaktif secara optimal[2]. Dampak dari perkembangan teknologi pendidikan yaitu, munculnya perusahaan penyedia berbagai macam Platform aplikasi, seperti Quipper, Zenius dan Ruangguru[3].

Google Play Store merupakan pasar terbuka untuk berbagai macam aplikasi seperti toko online, permainan, buku elektronik, kelas belajar, dan lain-lain, yang tersedia di *platform* tersebut. Negara Indonesia memiliki banyak perusahaan pembuat aplikasi belajar, seperti, Ruangguru, Zenius, Quipper, dan lain-lain. Setiap aplikasi memiliki kelemahan dan kelebihan masing-masing. Sebagian besar pengguna menyampaikan hal tersebut melalui halaman keluhan dan saran yang tersedia di halaman *Google Play Store*[4]. Setiap halaman aplikasi di *Google Play Store* memiliki bagian komentar di mana pengguna menyampaikan kritik membangun mereka terhadap aplikasi yang telah diunduh dan digunakan[5].

Halaman *Google Play Store*, terdapat sebuah ulasan mengenai pendapat pengguna yang telah menggunakan aplikasi tersebut[6]. Ulasan mengenai aplikasi merupakan ruang yang digunakan pengembang aplikasi untuk memperbaiki kinerja aplikasi. sehingga pengembang aplikasi dapat mengambil sebuah keputusan berdasarkan rekomendasi yang diberikan oleh pelanggan[7]. Kelebihan dan kekurangan dari masing-masing *brand* aplikasi belajar *online* masih perlu dianalisis lebih dalam, agar dapat menguntungkan dan mempermudah penggunaannya. *Google PlayStore* menjadi tempat para penggunanya memberikan *review* jujur terhadap penilaian dari kualitas masing-masing *brand* bimbingan belajar *online* (*bimbel online*) [8].

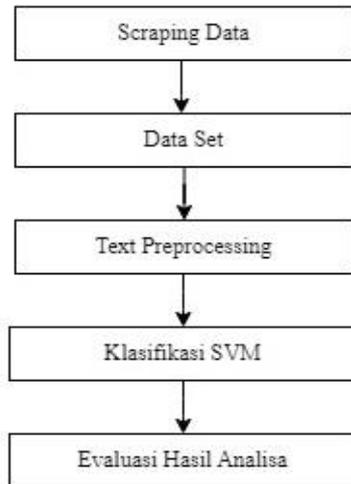
Sentimen analisis merupakan proses pengolahan data yang diperoleh dari berbagai macam data untuk mendapatkan sebuah informasi. Analisis sentimen digunakan untuk menemukan atau menganalisa umpan balik pengguna mengenai sebuah produk dan layanan[9]. Analisis sentimen adalah instrumen terbaik untuk menentukan apakah evaluasi itu positif atau negatif. Hasil dari sentimen analisis didapatkan dari keterkaitan hubungan saling mempengaruhi aspek kognitif, psikologis dan social [10]. Para pengguna tidak hanya menggambarkan isi dari konten tentang acara tetapi juga terkait perasaan penggunaannya [11]. Hampir semua perusahaan profesional dan non-teknis menggunakan teknik sentimen analisis untuk mengartikan maksud dari masukan dan keinginan pelanggan mengenai bagaimana kualitas dari produk atau layanannya [12].

Penelitian sebelumnya mengenai sentimen analisis, yang dilakukan Omar Alqaryouti dkk, melakukan analisa terhadap aplikasi *smart government*, yang berguna untuk meningkatkan kinerja aplikasi dan layanan pemerintah. Evaluasi dilakukan dengan menganalisa ribuan data, kemudian diklasifikasi menggunakan SVM dan leksikon SVM. Hasilnya, bahwa leksikon SVM lebih baik [13]. Mohammad Al-Smadi dkk, melakukan analisis sentimen dari ulasan Hotel Arab. Penelitian dilakukan dengan mengimplementasikan *deep recurrent neural network(RNN)* dan *Support Vector Machine (SVM)*. Hasil evaluasi, menunjukkan bahwa metode SVM lebih baik daripada metode RNN[14]. Ziedhan Alifio Diekson dkk, membahas kepuasan pelanggan terhadap layanan Traveloka. Dataset dikumpulkan dari platform media sosial yang terdiri dari 1200 *tweet* terkait Traveloka. Penelitian ini menggunakan tiga metode klasifikasi: *Support Model Vektor (SVM)*, *Regresi Logistik*, dan *Närve Bayes*. Hasil menunjukkan, dari 1200 *tweets* kebanyakan memberikan nilai positif terhadap traveloka. Hasilnya menunjukkan bahwa SVM lebih baik akurasi dalam menentukan sentimen *tweet* tentang Traveloka[15].

Penelitian bertujuan mengimplementasikan dan meningkatkan performa akurasi dari kinerja algoritma SVM dalam konteks analisis sentimen. SVM merupakan algoritma klasifikasi yang bagus dan efektif, terutama dalam kasus pemisahan linear dan non-linear. Sentimen analisis adalah teknik dalam digunakan untuk mengidentifikasi, memahami, dan mengklasifikasikan sentimen atau opini dalam teks. Penelitian ini, akan dilakukan dengan mengembangkan sentiment analisis menggunakan algoritma *Support Vector Machine* dalam menentukan aplikasi belajar *online* terbaik dengan mengukur kepuasan dari konsumen dan berdasarkan pengalaman pengguna. Penelitian memanfaatkan ulasan yang tersedia di *Google Play Store*. Hasil dari analisis pada penelitian ini dapat membandingkan hasil sentimen antara dua atau lebih dari *brand* aplikasi belajar *online* berdasarkan tingkat akurasi dan klasifikasi sentimen analisis menggunakan SVM.

2. Metode Penelitian

Pada tahap ini, akan dijelaskan mengenai proses dan tahapan mengenai metode penelitian yang digunakan. Tahapan penelitian dijelaskan secara runtut pada Gambar 1. kerangka Umum Penelitian. Berikut alur penelitian.



Gambar 1. Kerangka Umum Penelitian

Metode penelitian, dimulai dari tahapan pengumpulan data yang dilakukan dengan teknik scraping data di *Google PlayStore*. Kemudian, masuk ke tahapan pengolahan data. Pengolahan data dilakukan dengan teknik preprocessing data, yang meliputi *cleansing*, *case folding*, *tokenize* dan *filter stopwords*. Selanjutnya, data akan diklasifikasi menggunakan algoritma SVM, sehingga menghasilkan keluaran berupa sentimen negatif dan positif.

2.1. Scraping Data

.Teknik *Scraping* merupakan teknik untuk mengubah data web yang tidak terstruktur menjadi data terstruktur yang dapat disimpan dan dianalisis dalam database atau spreadsheet pusat. Proses *web Scraping*, dibagi menjadi 3 tahap yaitu *Fetching*, *Extraction*, dan *Transformation*. Tahap *Fetching* dilakukan dengan akses melalui protokol HTTP, untuk mengirim dan menerima permintaan dari *server web*. Kemudian dilakukan *Extraction*, melalui halaman HTML dengan melakukan parsing data. Setelah data terkumpul, dilakukan *Transformation* data, untuk mendapatkan data yang terstruktur [16].

2.2. Data Set

Pengumpulan data dilakukan dengan menggunakan data yang diperoleh dari Platform Google Play Store sebagai penyedia aplikasi. Data dikumpulkan dari ulasan beberapa aplikasi belajar *online* seperti Ruangguru, Zenius dan Quipper. Proses pengumpulan data dengan menggunakan bantuan library dari pemrograman Python. *google-play-scraper* merupakan *library* yang digunakan penambangan data dari *Google PlayStore*, dilakukan dengan menggunakan Token, untuk meminta

hak akses penggunaan data. Pengambilan data dilakukan dengan membangun sistem *Scraping* data. Data hasil *Scraping* ini dimasukkan ke *database* untuk diproses lebih lanjut.

2.3. Text Preprocessing

Tahap teks preprocessing digunakan untuk memperoleh data yang sesuai dengan proses dari sistem yang akan dibangun, dimana sudah ada pengurangan kosa kata, menghilangkan *noise*, membuat data lebih terstruktur sehingga dihapuskan dapat mempermudah dan mempercepat proses. Adapun tahapan yang dilakukan pada teks preprocessing diantaranya yaitu, *normalisasi*, *case folding*, *cleaning*, *tokenizing* dan *Stopword*.

Normalisasi, sebuah data teks diperlukan pengecekan apakah sebuah kata termasuk dalam kamus bahasa Indonesia. *Case Folding*, mengubah huruf besar menjadi kecil untuk menyamakan atau menyeragamkan karakter pada data teks[17]. *Cleaning*, *cleaning* merupakan proses membersihkan sebuah *review* atau penilaian dari kata-kata yang tidak diperlukan untuk mengurangi proses noise pada proses klasifikasi. *Tokenizing*, proses *tokenizing* dan *Stopword* adalah tahap pemotongan string masukan berdasarkan kata-kata yang menyusun atau dengan kata lain pemecahan kalimat menjadi kata-kata kemudian diubah menjadi bagian token-token yang akan digunakan dalam klasifikasi SVM dan penghilang kata sambung atau kata umum yang besar dan tak mempunyai makna [18].

2.4. Klasifikasi SVM

Support Vector Machine (SVM) adalah salah satu algoritma yang sangat berguna dalam klasifikasi data yang besar. Ide Algoritma SVM diciptakan oleh Vladimir Vapnik dan sebagai algoritma yang efisien dan bisa diterapkan di berbagai bidang[19]. SVM merupakan pembelajaran dengan analisa data dan mengenali sebuah pola. Standar dari SVM yaitu mengambil himpunan data masukan dan memprediksi untuk setiap masukkan yang diberikan, hasil kemungkinan masukkan adalah anggota dari salah satu kelas dari dua kelas yang ada, yang mana membuat algoritma SVM sebagai penggolong *non probabilistic linier biner*, dikarenakan algoritma SVM sebuah pengklasifikasi, dengan kemudian diberi himpunan pelatihan yang mana masing-masing ditandai sebagai milik salah satu dari dua kategori[20]. klasifikasi linier SVM dinotasikan sebagai formula perumusan berikut [21]:

$$f(x) = W^T X + b \quad (1)$$

Dalam klasifikasi SVM, fungsi klasifikasi dilambangkan dengan $f(x)$. Fungsi $f(x)$ memprediksi kelas target y berdasarkan fitur input x . W adalah vektor bobot, X adalah vektor fitur masukan, dan b adalah bias. Sehingga menurut Vapnik dan Cortes (1995) diperoleh persamaan formula perumusan sebagai berikut:

$$[(W^T \cdot x_i) + b] \geq 1 \text{ untuk } y_i = +1 \quad (2)$$

$$[(W^T \cdot x_i) + b] \leq -1 \text{ untuk } y_i = -1 \quad (3)$$

Tabel 1. Confusion Matrix

Kelas Prediksi			
Kelas Aktual	Kelas Positif	Kelas Positif	Kelas Negatif
		TP (True Positive)	FP (False Positive)
	Kelas Negatif	FN (False Negative)	TN (True Negative)

Tabel 2. Data Set Penelitian

Belajar Online	Token	Jumlah Data
Ruang Guru	com.ruangguru.livestudents	10.000
Zenius	net.zenius.mobile	10.000
Quipper	com.quipper.school.assignment	10.000

Tabel 3. Contoh Hasil Scraping

Id	Nama	Ulasan
bc7e11c2-47be-48a6-be9e-537bf0862134	Anindya	aplikasinya bagus bangettt, ngebantu banget buat aku yang cara belajarnya memang audiovisual. tapi abis di update kok malah lag & ga bisa nonton vidio. aku coba install ulang, pas mau log in lagi malah ga bisa, selalu ke force stop sendiri. mohon perbaikannya ya 🙏
17447195-f12c-47e3-8a2e-fee8fe42a19d	Muhammad Iqbal	Mengapa setelah update aplikasi jadi sering force close dan stuck layar hitam, padahal saya sudah berulang kali instal dan uninstall. Saya sudah hapus cache aplikasi ruanggurunya dan restart hp saya, tapi hasilnya tetap tidak bisa di akses. Mohon kepada ruangguru untuk memperbaiki aplikasinya, kemungkinan ini adalah bug.

Klasifikasi dilakukan dengan dataset latih yang dinotasikan dengan $f(x)$ = himpunan data *training*, $k_i = 1, 2, \dots, n$ dan, y_i = label kelas dari x^i . Pengelompokan menggunakan hyperplane linier seperti terlihat pada rumus. Parameter w menotasikan *vector* terhadap *hyperplane*, b adalah *offset*[22].

2.5. Evaluasi Hasil Analisa

Tahap evaluasi menjadi salah satu tahap penting dalam membangun sebuah program sistem, dengan adanya tahap evaluasi maka pengguna akan mengetahui setiap point akurasi, presisi bahkan *error* dari sebuah program sistem yang dibuatnya. Salah satu teknik untuk mengetahui hasil evaluasi dari program sentimen analisis yaitu bisa menggunakan *Confusion Matrix*. *Confusion Matrix* merupakan sebuah metode yang dapat mempermudah program untuk dapat menentukan nilai penting dalam menimbang nilai akurasi, presisi, recall dan error dari sebuah program sistem, yang diilustrasikan pada Tabel 1. *Confusion Matrix*[23].

Confusion Matrix terdiri dari kelas prediksi dan kelas aktual. Tabel 1 menggambarkan kinerja klasifikasi pada hasil prediksi dan nilai aktual dalam penelitian. Matrix menampilkan jumlah data yang diklasifikasikan dengan

benar dan salah dalam setiap kelasnya. Berdasarkan bentuk dari *confusion matrix* diatas, maka dengan mudah untuk dapat mengetahui nilai-nilai penting yang memudahkan peneliti untuk menghitung akurasi, recall dan presisi dari sebuah program sistem, maka berikut formula perumusannya:

$$\text{Clasification Akurasi} = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \times 100\% \quad (4)$$

$$\text{Recall} = \frac{T_p}{T_p + F_p} \times 100\% \quad (5)$$

$$\text{Akurasi} = \frac{T_p}{T_p + F_n} \times 100\% \quad (6)$$

Metode *Confusion Matrix* merupakan metode yang populer saat ini untuk melaporkan kinerja pengklasifikasian dengan menghitung akurasi dan *recall*[24]. Hasil analisa disimpulkan dengan *confusion matrix* yang dinotasikan T_p , T_n , F_p , dan F_n . T_p merupakan Jumlah kelas positif yang diklasifikasi positif. T_n penunjukkan Jumlah kelas positif yang diklasifikasi negatif. F_p menunjukkan Jumlah kelas negatif yang diklasifikasi positif. F_n adalah Jumlah kelas negatif yang diklasifikasi negatif.

3. Hasil dan Pembahasan

Hasil dan pembahasan merupakan penjelasan hasil dari penelitian yang telah dilakukan mengenai Sentimen dan SVM terhadap data Ruang Guru, Zenius dan Quipper. Pada tahapan ini, hasil akan dijelaskan dari Pengumpulan Data, *Preprocessing* Data, Hasil dari Klasifikasi Sentimen Analisis menggunakan SVM dan Evaluasi terhadap hasil yang telah diperoleh dalam penelitian. Berikut penjelasan mengenai hasil dan pembahasan dari setiap langkah penelitian.

3.1. Pengumpulan Data

Data pada penelitian ini, diperoleh dengan cara melakukan *scraping* data. Teknik *scraping* data digunakan dengan bantuan library *google play scraper* yang tersedia di bahasa pemrograman *python*. Library dihubungkan dengan menggunakan API dan token yang terdapat pada Tabel 2. Data yang dikumpulkan sebanyak 30.000 data dengan kategori *most relevant*. Hal ini dilakukan, agar data yang diperoleh adalah data yang terbaik dari setiap *review* produk.

Penelitian dilakukan dengan mengumpulkan 10.000 data setiap produknya. Setiap data dibagi menjadi 90% data latih dan 10% data uji untuk diproses kedalam penelitian. Seperti yang terlihat pada Tabel 2 menunjukkan token dan jumlah data yang diperoleh. Data yang terkumpul menjadi sebuah dataset yang dihimpun dalam lembar kerja CSV. Kemudian data diproses kedalam pemrosesan teks. Tabel 3 menunjukkan hasil dari *scraping* data yang telah diperoleh dan akan dijadikan sebuah dataset penelitian.

3.2. Hasil Preprocessing Data

Tahap selanjutnya, yaitu *preprocessing* data. Data yang telah diperoleh dijadikan sebuah dataset yang siap untuk

Tabel 4. Contoh Hasil Normalisasi

Text	Hasil
Saya sangat membantu dgn aplikasi ini.. krena kata2nya yg sopan dan mudah di mengerti.. pokoknya baguslah. Cumana yg saya sarankan koinnya dimurahkan dikitlah.. soalnya kemahalan hehhe☺☺	Saya sangat membantu dgn aplikasi ini.. krena kata2nya yg sopan dan mudah di mengerti.. pokoknya baguslah. Cumana yg saya sarankan koinnya dimurahkan dikitlah.. soalnya kemahalan hehhe☺☺
zenius keren asiik dan bikin semngat belajar !	zenius keren asiik dan bikin semngat belajar !
Selalu berbuat yg terbaik utk kemajuan.....	Selalu berbuat yg terbaik utk kemajuan.....

Tabel 5. Contoh Hasil Case Folding

Text	Hasil
Saya sangat membantu dgn aplikasi ini.. krena kata2nya yg sopan dan mudah di mengerti.. pokoknya baguslah. Cumana yg saya sarankan koinnya dimurahkan dikitlah.. soalnya kemahalan hehhe☺☺	saya sangat membantu dgn aplikasi ini.. krena kata2nya yg sopan dan mudah di mengerti.. pokoknya baguslah. cumana yg saya sarankan koinnya dimurahkan dikitlah.. soalnya kemahalan hehhe☺☺
zenius keren asiik dan bikin semngat belajar !	zenius keren asiik dan bikin semngat belajar !
Selalu berbuat yg terbaik utk kemajuan.....	selalu berbuat yg terbaik utk kemajuan.....

diolah kedalam sistem. Data akan melalui beberapa tahapan untuk membersihkan, memformat, dan menyiapkan data teks sebelum menggunakannya sebagai input untuk model analisis sentimen. Data akan diperbaiki dan menjadi data yang terstruktur. Tahap preprocessing data meliputi *normalisasi*, *case folding*, *cleaning*, *tokenizing*, dan *Stopword*.

Tahap awal yaitu normalisasi, data yang diproses dipastikan bahwa data mengandung bahasa indonesia. Selanjutnya, *casefolding* dilakukan untuk menyeragamkan data. Kata atau kalimat yang ada diubah kedalam bentuk huruf kecil semua (*lower case*). Contoh hasil Normalisasi dan *casefolding* terlihat pada Tabel 4 dan Tabel 5.

Cleansing merupakan proses pembersihan sebuah kalimat atau kata dari beberapa karakter yang tidak bermakna. Proses, pembersihan dilakukan dengan menghilangkan *URL* dari dalam konten, menghapus *hashtag* (#) dan mention (@) dan tanda baca seperti titik(.),koma(,), dan lain-lain. Karakter emoticon juga dihilangkan. Cleansing dilakukan untuk membersihkan teks dari karakter khusus, menghilangkan *noise* yang tidak relevan dan mempermudah proses pembuatan *tokenize stopwords*. Berikut adalah contoh dari hasil proses cleansing yang dilakukan, seperti yang terlihat pada Tabel 6.

3.3. Klasifikasi Sentimen Analisis Menggunakan SVM

Sentimen analisis merupakan pengolahan sebuah kata atau kalimat untuk mendapatkan informasi. Data Ruang guru, Zenius dan Quipper diolah untuk mendapatkan

Tabel 6. Contoh Hasil Cleansing

Text	Hasil
saya sangat membantu dgn aplikasi ini.. krena kata2nya yg sopan dan mudah di mengerti.. pokoknya baguslah. cumana yg saya sarankan koinnya dimurahkan dikitlah.. soalnya kemahalan hehhe☺☺	saya sangat membantu dgn aplikasi ini krena katanya yg sopan dan mudah di mengerti pokoknya baguslah cumana yg saya sarankan koinnya dimurahkan dikitlah soalnya kemahalan hehhe
zenius keren asiik dan bikin semngat belajar !	zenius keren asiik dan bikin semngat belajar
selalu berbuat yg terbaik utk kemajuan.....	selalu berbuat yg terbaik utk kemajuan

Tabel 7. Contoh Hasil Tokenize

Text	Hasil
saya sangat membantu dgn aplikasi ini krena katanya yg sopan dan mudah di mengerti pokoknya baguslah cumana yg saya sarankan koinnya dimurahkan dikitlah soalnya kemahalan hehhe	['saya', 'sangan', 'membantu', 'dgn', 'aplikasi', 'ini', 'krena', 'katanya', 'yg', 'sopan', 'dan', 'mudah', 'di', 'mengerti', 'pokoknya', 'baguslah', 'cumana', 'yg', 'saya', 'sarankan', 'koinnya', 'dimurahkan', 'dikitlah', 'soalnya', 'kemahalan', 'hehhe']
zenius keren asiik dan bikin semngat belajar	['zenius', 'keren', 'asiik', 'dan', 'bikin', 'semngat', 'belajar']
selalu berbuat yg terbaik utk kemajuan	['selalu', 'berbuat', 'yg', 'terbaik', 'utk', 'kemajuan']

Tokenize merupakan tahapan untuk mengubah data ke dalam struktur text menjadi kata. Kemudian *Stopword* merupakan proses untuk menghilangkan kata-kata yang jumlahnya besar namun tidak memiliki makna. Struktur kata yang telah diubah, akan menjadi sebuah token-token. Token ini, berfungsi untuk membantu proses perhitungan dan menjadi perhitungan ke dalam sebuah array data. Proses *tokenize* dan *Stopword*, dilihatkan dalam Tabel 7 dan Tabel 8 dalam mendapatkan sebuah informasi mengenai produk tersebut. Setelah melalui tahap *preprocessing*, data siap diolah dan dianalisa untuk menghasilkan sebuah informasi, berupa sentimen negatif, netral dan positif. Klasifikasi dilakukan dengan membagi data menjadi data latih dan data uji. Penelitian melakukan percobaan perbandingan 90% data latih dan 10% data uji. Data latih diberi label dengan nilai 1, 0, dan -1. Nilai 1 berarti positif, nilai 0 berarti netral dan -1 berarti negatif. Hasil klasifikasi sentimen dilihatkan pada Gambar 2.

Analisa sentimen pada Gambar 2, menunjukkan bahwa data yang ada berhasil diolah dan diproses menjadi sebuah informasi. Data sebanyak 30.000, yang dibagi menjadi 10.000/produk, dapat begitu mudah dianalisa dengan bantuan sentimen analisis. Nilai positif banyak diberikan pengguna kepada lembaga belajar *online* Ruang Guru. Nilai netral dan negatif diberikan pengguna kepada lembaga Quipper. Akan tetapi tanggapan pengguna mengenai Quipper, Zenius dan Ruang Guru, semuanya memberikan respon positif. Hal ini, membuktikan bahwa masyarakat pada umumnya dan khususnya peserta didik, sangat menyambut dengan baik kemajuan teknologi dan pendidikan.

Tabel 8. Contoh Hasil Stopword

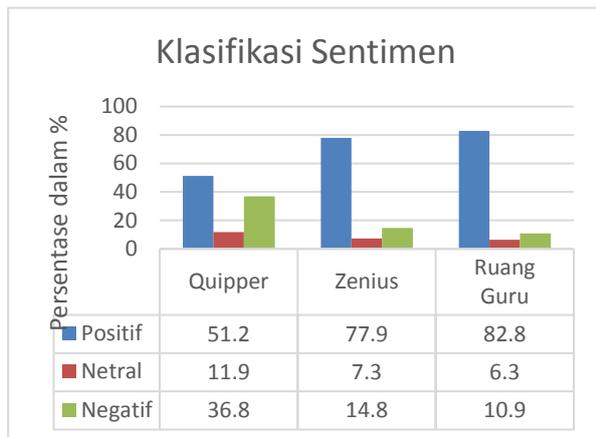
Text	Hasil
saya sangat membantu dgn aplikasi ini krena katanya yg sopan dan mudah di mengerti pokoknya baguslah cumana yg saya sarankan koinnya dimurahkan dikitlah soalnya kemahalan hehhe	['sangan', 'membantu', 'dgn', 'aplikasi', 'krena', 'katanya', 'yg', 'sopan', 'mudah', 'mengerti', 'pokoknya', 'baguslah', 'cumana', 'yg', 'sarankan', 'koinnya', 'dimurahkan', 'dikitlah', 'soalnya', 'kemahalan', 'hehhe']
zenius keren asiik dan bikin semngat belajar	['zenius', 'keren', 'asiik', 'bikin', 'semngat', 'ajar']
selalu berbuat yg terbaik utk kemajuan	['berbuat', 'yg', 'terbaik', 'utk', 'kemajuan']

	precision	recall	f1-score	support
-1	0.95	0.96	0.95	810
0	0.99	0.97	0.98	754
1	0.94	0.95	0.95	773
accuracy			0.96	2337
macro avg	0.96	0.96	0.96	2337
weighted avg	0.96	0.96	0.96	2337

Gambar 4. Hasil Evaluasi Zenius

	precision	recall	f1-score	support
-1	0.80	0.75	0.77	520
0	0.82	0.88	0.85	518
1	0.84	0.83	0.84	499
accuracy			0.82	1537
macro avg	0.82	0.82	0.82	1537
weighted avg	0.82	0.82	0.82	1537

Gambar 5. Hasil Evaluasi Quipper



Gambar 2. Hasil Klasifikasi Sentimen

	precision	recall	f1-score	support
-1	0.99	0.99	0.99	837
0	0.99	0.99	0.99	785
1	0.98	0.98	0.98	862
accuracy			0.99	2484
macro avg	0.99	0.99	0.99	2484
weighted avg	0.99	0.99	0.99	2484

Gambar 3. Hasil Evaluasi Ruang Guru

3.4 Hasil dan Evaluasi

Tahap akhir dari penelitian yaitu hasil dan evaluasi terhadap klasifikasi sentimen analisis menggunakan SVM. Hasil dan evaluasi didapatkan dengan merumuskan nilai akurasi, recall dan precision. Nilai tersebut didasarkan pada perhitungan sentimen dan SVM yang telah dilakukan. Gambar 2, Gambar 3, dan Gambar 4, merupakan evaluasi hasil dari proses klasifikasi sentimen yang telah dijalankan.

Nilai akurasi digunakan untuk mengukur seberapa tepat, data sentimen melakukan prediksi klasifikasinya dengan benar. Penelitian sentimen analisis menggunakan SVM, menunjukkan bahwa nilai akurasi yang dilakukan pada data RuangGuru menunjukkan sebesar 99% dan menjadi lebih tinggi dibandingkan data ulasan zenius dan Quipper. Nilai *precision*, *recall* dan *f1-score*, juga menunjukkan bahwa penelitian dengan menggunakan data Ruangguru mendapatkan nilai yang baik, yaitu dengan rata-rata 99%.

4. Kesimpulan

Berdasarkan penelitian yang sudah dilakukan, Sentimen analisis menggunakan algoritma SVM dapat diimplementasikan dengan baik dalam mengklasifikasikan Ruang Guru, Zenius, dan Quipper. Hasil dari penelitian menunjukkan bahwa, data Klasifikasi sentimen dengan SVM, mempunyai akurasi rata-rata untuk Ruang Guru sebesar 99%, Zenius sebesar 96%, dan Quipper sebesar 82%. Penelitian ini berhasil mendapatkan dan meningkatkan performance measure pada akurasi SVM dengan nilai akurasi 99%. Hasil ini, didapatkan dengan membandingkan data pada Ruang Guru, Zenius dan Quipper.

Penelitian menggunakan perbandingan data 90%:10%. Kedepannya, diharapkan bisa mencoba performa SVM dan sentimen analisis dengan membandingkan perbandingan data latih dan data uji yang lain.

Daftar Pustaka

- [1] A. Z. Shoumi, "Peran Multimedia Dalam Pendidikan Pada Aplikasi Ruang Guru," in *Prosiding Seminar Nasional Cendekiawan*, 2019, hal. 2. doi: 10.25105/semnas.v0i0.5809.
- [2] A. Ramadhayanti, "Analisis Strategi Belajar Dengan Metode Bimbel Online Terhadap Kemampuan Pemahaman Kosakata Bahasa Inggris dan Pronunciation (Pengucapan/pelafalan) Berbahasa Remaja Saat Ini," *KREDO J. Ilm. Bhs. dan Sastra*, vol. 2, no. 1, hal. 39–52, 2018, doi: 10.24176/kredo.v2i1.2580.
- [3] A. F. Hayati, "Perbedaan Hasil Belajar Siswa Menggunakan Bimbingan Belajar Online," *J. Inov. Pendidik. Ekon.*, vol. 10, no. 1, hal. 79–85, 2020, doi: 10.24036/011085130.
- [4] M. I. Ahmadi, F. Apriani, M. Kurniasari, S. Handayani, dan D. Gustian, "Sentiment Analysis Online Shop on the Play Store Using Method Support Vector Machine (Svm)," in *Seminar Nasional ...*, 2020, hal. 196–203. [Daring]. Tersedia pada: <http://jurnal.upnyk.ac.id/index.php/semnasif/articl>

- e/view/4101
- [5] M. M. Jawad Soumik, S. Salvi Md Farhavi, F. Eva, T. Sinha, dan M. S. Alam, "Employing machine learning techniques on sentiment analysis of google play store bangla reviews," in *2019 22nd International Conference on Computer and Information Technology, ICCIT 2019*, 2019, hal. 1–5. doi: 10.1109/ICCIT48885.2019.9038348.
- [6] R. Wahyudi dan G. Kusumawardana, "Analisis Sentimen pada Aplikasi Grab di Google Play Store Menggunakan Support Vector Machine," *J. Inform.*, vol. 8, no. 2, hal. 200–207, 2021, doi: 10.31294/ji.v8i2.9681.
- [7] S. Venkatakrishnan, A. Kaushik, dan J. K. Verma, "Sentiment Analysis on Google Play Store Data Using Deep Learning," in *Applications of Machine Learning*, 2020, hal. 15–30. doi: 10.1007/978-981-15-3357-0_2.
- [8] L. Kirtibas Singh dan R. Renuga Devi, "Student feedback sentiment analysis: A review," in *Materials Today: Proceedings*, 2021, no. xxxx. doi: 10.1016/j.matpr.2020.10.782.
- [9] Q. Li, S. Shah, R. Fang, A. Nourbakhsh, dan X. Liu, "Tweet Sentiment Analysis by Incorporating Sentiment-Specific Word Embedding and Weighted Text Features," *Proc. - 2016 IEEE/WIC/ACM Int. Conf. Web Intell. WI 2016*, hal. 568–571, 2017, doi: 10.1109/WI.2016.0097.
- [10] K. Jindal dan R. Aron, "Materials Today: Proceedings A systematic study of sentiment analysis for social media data," in *Materials Today: Proceedings*, 2021, no. xxxx. doi: 10.1016/j.matpr.2021.01.048.
- [11] S. Vashishtha dan S. Susan, "Fuzzy rule based unsupervised sentiment analysis from social media posts," *Expert Syst. Appl.*, vol. 138, hal. 112834–112849, 2019, doi: 10.1016/j.eswa.2019.112834.
- [12] P. Chitra dkk., "Materials Today: Proceedings Sentiment analysis of product feedback using natural language processing," 2021. doi: 10.1016/j.matpr.2020.12.1061.
- [13] O. Alqaryouti, N. Siyam, A. A. Monem, dan K. Shaalan, "Aspect-based sentiment analysis using smart government review data," *Appl. Comput. Informatics*, vol. 3, hal. 11–23, 2019, doi: 10.1016/j.aci.2019.11.003.
- [14] M. Al-Smadi, O. Qawasmeh, M. Al-Ayyoub, Y. Jararweh, dan B. Gupta, "Deep Recurrent neural network vs. support vector machine for aspect-based sentiment analysis of Arabic hotels' reviews," *J. Comput. Sci.*, vol. 27, hal. 386–393, 2018, doi: 10.1016/j.jocs.2017.11.006.
- [15] Z. A. Diekson, M. R. B. Prakoso, M. S. Q. Putra, M. S. A. F. Syaputra, S. Achmad, dan R. Sutoyo, "Sentiment analysis for customer review: Case study of Traveloka," *Procedia Comput. Sci.*, vol. 216, no. 2022, hal. 682–690, 2023, doi: 10.1016/j.procs.2022.12.184.
- [16] M. A. Khder, "Web scraping or web crawling: State of art, techniques, approaches and application," *Int. J. Adv. Soft Comput. its Appl.*, vol. 13, no. 3, hal. 144–168, 2021, doi: 10.15849/ijasca.211128.11.
- [17] E. Fitri, "Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine," *J. Transform.*, vol. 18, no. 1, hal. 71, 2020, doi: 10.26623/transformatika.v18i1.2317.
- [18] A. P. Nardilasari, A. L. Hananto, S. S. Hilabi, dan B. Priyatna, "Analisis Sentimen Calon Presiden 2024 Menggunakan Algoritma SVM," vol. 7, no. 1, hal. 11–18, 2024.
- [19] D. Wang dan Y. Zhao, "Using News to Predict Investor Sentiment: Based on SVM Model," *Procedia Comput. Sci.*, vol. 174, no. 2019, hal. 191–199, 2020, doi: 10.1016/j.procs.2020.06.074.
- [20] A. P. Giovani, A. Ardiansyah, T. Haryanti, L. Kurniawati, dan W. Gata, "Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi," *J. Teknoinfo*, vol. 14, no. 2, hal. 115, 2020, doi: 10.33365/jti.v14i2.679.
- [21] S. Bo, S. J. Song, dan W. Cheng, "A new algorithm of support vector machine based on weighted feature," *Proc. 2009 Int. Conf. Mach. Learn. Cybern.*, vol. 3, no. July, hal. 1616–1620, 2009, doi: 10.1109/ICMLC.2009.5212256.
- [22] C. O. Sianturi, S. C. Sianturi, dan A. H. Mondolang, "Klasifikasi Citra Daun Anggur Menggunakan SVM Kernel Linear," vol. 7, no. 1, hal. 19–26, 2023.
- [23] I. Markoulidakis, G. Kopsiaftis, I. Rallis, dan I. Georgoulas, "Multi-Class Confusion Matrix Reduction method and its application on Net Promoter Score classification problem," *ACM Int. Conf. Proceeding Ser.*, no. Cx, hal. 412–419, 2021, doi: 10.1145/3453892.3461323.
- [24] M. Heydarian dan T. E. Doyle, "MLCM: Multi-Label Confusion Matrix," vol. 2, hal. 19083–19095, 2022.

Halaman ini sengaja dikosongkan