

Sentimen Analisis *Review* Pada Aplikasi Edlink Sevima Di *GooglePlaystore* Menggunakan Metode SVM

Rivan Kurnia¹, Teguh Nurhadi Suharsono²

^{1,2}Fakultas Teknik, S1-Teknik Informatika, Universitas Sangga Buana Ypkp, Kota Bandung, Indonesia

Email: rivank076@gmail.com, teguh.nurhadi@usbypkp.ac.id

Submitted: 05/10/2024; Accepted: 20/10/2024; Published: 24/10/2024

Abstrak — Kemajuan teknologi dan informasi di era industri 4.0 memiliki dampak signifikan pada proses pembelajaran di berbagai lembaga pendidikan. Aplikasi Edlink Sevima adalah salah satu platform e-learning yang banyak digunakan dan dievaluasi oleh pengguna melalui ulasan di Google Play Store. Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna terhadap aplikasi Edlink Sevima menggunakan metode Support Vector Machine (SVM). Metode SVM dipilih karena kemampuannya mengklasifikasikan data teks dengan akurasi tinggi. Data ulasan dikumpulkan dari Google Play Store dan diproses menggunakan teknik preprocessing untuk mengoptimalkan analisis sentimen. Hasil penelitian menunjukkan bahwa metode SVM mampu mengklasifikasikan ulasan menjadi sentimen positif, negatif, dan netral dengan akurasi terbaik sebesar 75%. Sentimen negatif memiliki precision sebesar 0.71 dan recall sebesar 0.79, sedangkan sentimen positif memiliki precision sebesar 0.78 dan recall sebesar 0.87. Diharapkan hasil dari penelitian ini dapat memberikan sumbangan dalam pemahaman tentang ulasan pengguna aplikasi di *Google Play Store*. Walaupun demikian, penelitian ini memiliki keterbatasan, seperti metode yang digunakan dan data yang dikumpulkan. Oleh karena itu, penelitian ini dapat dijadikan dasar untuk penelitian lanjutan dalam bidang analisis sentimen dari ulasan aplikasi di *Google Play Store*.

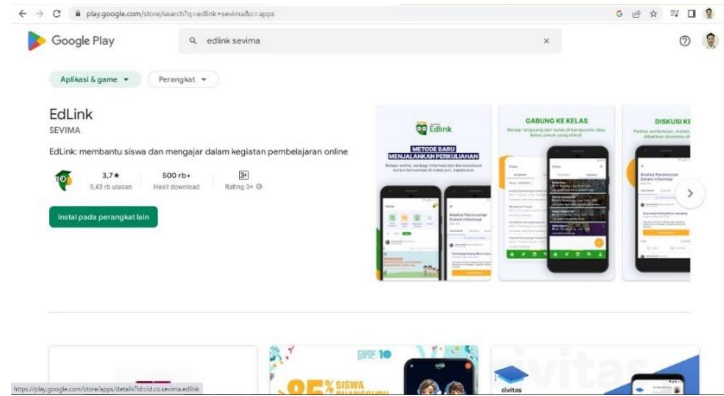
Kata kunci— Analisis Sentimen, Edlink Sevima, Support Vector Machine, Google Play Store, E-learning.

I. PENDAHULUAN

Pesatnya perkembangan teknologi dan informasi telah memberikan dampak yang signifikan terhadap proses pembelajaran di sejumlah lembaga pendidikan pada Era Industri 4.0. Dimana salah satu jenis pemanfaatan inovasi data yang telah diterapkan dalam bidang pelatihan adalah *e-learning*. *E-Learning* merupakan media pembelajaran yang dirancang dengan memanfaatkan inovasi komputer dan web sehingga materi pembelajaran dapat berjalan tanpa gangguan atau keterbatasan waktu.[1] Analisis sentimen merupakan proses memahami, mengekstraksi, dan mengolah data tekstual secara otomatis untuk memperoleh informasi yang terkandung dalam sebuah kalimat opini.[2]

Selain itu, terdapat pula beberapa aplikasi yang digunakan sebagai media pembelajaran elektronik seperti *Moodle*, *Google Classroom*, *Sevima Edlink*, *Schoolgy*, dan pembelajaran elektronik lainnya. *Sevima Edlink* merupakan platform seluler berbasis Android dan *website* yang secara khusus ditujukan untuk lebih meningkatkan peluang pertumbuhan dalam bidang pendidikan. Tahap ini menyediakan berbagai elemen termasuk konten pembelajaran cerdas, tugas daring, percakapan, dan masih banyak lagi. Aplikasi ini dapat digunakan oleh para pendidik di tingkat sekolah maupun di tingkat perguruan tinggi dalam mengerjakan latihan pembelajaran dan memiliki akses yang mudah karena dapat diakses di berbagai tahap termasuk *Google Play Store*. [3] Di era komputerisasi ini, Ulasan pengguna di berbagai tahapan, terutama *Google Play Store*, sangat memengaruhi pandangan *individu* terhadap suatu aplikasi.

Meskipun aplikasi ini telah mendapat banyak perhatian dan digunakan secara umum, kualitas dan kinerja aplikasi ini sering dinilai oleh pengguna melalui survei di *Google Play Store*. Ulasan-ulasan ini memberikan wawasan tentang kepuasan pengguna, masalah teknis, dan fitur-fitur yang diinginkan. Pengembang dapat memperoleh pemahaman yang lebih baik tentang sentimen pengguna dan membuat penyesuaian yang diperlukan dengan menganalisis ulasan-ulasan ini. Selain itu, ada beberapa aplikasi yang bekerja di bidang yang sama yang juga terus meningkatkan dan merencanakan agar aplikasi tersebut memiliki reputasi yang baik sehingga lebih banyak orang mengakses aplikasi tersebut. Aplikasi *Edlink Sevima* memiliki ulasan yang dapat dikategorikan masih rendah dengan peringkat 3,7 dengan 5 ribu ulasan dan dengan hasil unduhan lebih dari 500 ribu unduhan sebagaimana dibuktikan oleh hasil gambar pencarian pada Gambar 1.1 diplatform *Google Play Store*. [4]



Gambar 1. 1 Penilaian Aplikasi Edlink Sevima

Gambar dibawah menggambarkan bagaimana pengalaman pengguna dalam menggunakan aplikasi berdasarkan evaluasi mereka. Pada aplikasi Edlink Sevima masih belum *ideal* baik dari segi *login* maupun pengumpulan tugas, hal ini terlihat dari berbagai kelompok pengguna yang merasakan kekecewaan yang dirasakan oleh pengguna. Seperti pada Gambar 2 di bawah ini merupakan contoh informasi ulasan pengguna untuk aplikasi Edlink Sevima melalui *Google Play Store*. [5]

Metode *Support Vector Machine* (SVM) adalah salah satu metode yang populer dalam analisis sentimen karena kemampuannya dalam mengklasifikasikan data teks dengan tingkat akurasi yang tinggi. Metode ini bekerja dengan mencari hyperplane optimal yang memisahkan data ke dalam kategori yang berbeda. Dengan memanfaatkan SVM, analisis sentimen dapat dilakukan untuk mengklasifikasikan setiap ulasan menjadi kategori sentimen positif, negatif dan netral. Akan tetapi, pada saat yang sama analisis sentimen pada ulasan balik dari pengguna Edlink Sevima juga memiliki tantangan tersendiri. [6] Ulasan balik pengguna di aplikasi Edlink Sevima sering kali ulasan ditulis dalam berbagai bahasa yang tidak baku, menggunakan bahasa sehari-hari, serta mengandung singkatan dan frase unik. Oleh karena itu, rating yang menjadi label untuk menganalisis apakah ulasan tersebut positif, negatif dan negatif. Dan metode *Support Vector Machine* dengan *Teknik preprocessing* data yang menjadi kunci untuk mengoptimalkan hasil analisis sentimen. Dalam konteks tersebut, penelitian ini bertujuan untuk menganalisis sentimen *review* pada aplikasi Edlink Sevima di *Google PlayStore* menggunakan metode *Support Vector Machine*. Dengan melakukan analisis ini, diharapkan akan ditemukan pola-pola sentimen yang muncul dalam respon pengguna terhadap aplikasi Edlink Sevima, serta kata-kata kunci yang berpengaruh dalam membentuk sentimen. [7]

II. METODE PENELITIAN

A. Metodologi Pengumpulan Data

Metode pengumpulan data yang dilakukan sebagai bahan penelitian ini yaitu dengan menggunakan beberapa metode, antara lain :

1. Observasi

Data *review* akan diambil dari Edlink Sevima di platform *Google Play Store*. Kriteria pemilihan ulasan yang digunakan adalah ulasan yang memiliki rating dan komentar yang lengkap. Metode pengambilan data dengan mengunduh data ulasan dari *Google Play Store* menggunakan *Google Collabs* dan menghapus ulasan yang tidak memenuhi kriteria pemilihan ulasan.

2. Studi Pustaka

Metode pengumpulan data study pustaka dilakukan dengan mengumpulkan beberapa informasi dari berbagai referensi seperti buku, artikel, makalah, jurnal ilmiah, dan penelitian terdahulu terkait dengan analisis sentimen.

B. Metode Pengembangan Sistem

Metode pengembangan sistem merupakan pendekatan atau teknik dalam ilmu komputer dan pemrosesan bahasa alami yang digunakan untuk mengevaluasi, mengidentifikasi, dan mengklasifikasikan sentimen atau opini yang terkandung dalam teks, ulasan, atau konten lainnya. Tujuan utama dari metode ini adalah untuk memahami apakah suatu teks mengandung sentimen positif dan negatif. Dalam analisis sentimen, metode ini biasanya melibatkan pengguna algoritma *machine learning* atau teknik pemrosesan bahasa alami untuk memproses teks dan mengambil informasi yang berkaitan dengan sentimen yang diungkapkan. Metode analisis sentimen dapat digunakan dalam berbagai konteks, seperti menganalisis produk, berita, tanggapan media sosial, dan lain sebagainya.

Metode yang digunakan untuk mengembangkan sistem adalah pendekatan *Support Vector Machine* (SVM). SVM adalah salah satu pendekatan yang umum digunakan untuk mengklasifikasikan teks berdasarkan sentimen yang terkandung di dalamnya, seperti ulasan positif dan negatif. SVM juga adalah sebuah algoritma *machine learning* yang dapat digunakan untuk pemisahan dua kelas atau lebih dalam ruang fitur yang lebih tinggi.

III. HASIL DAN PEMBAHASAN

A. Deskripsi Sistem

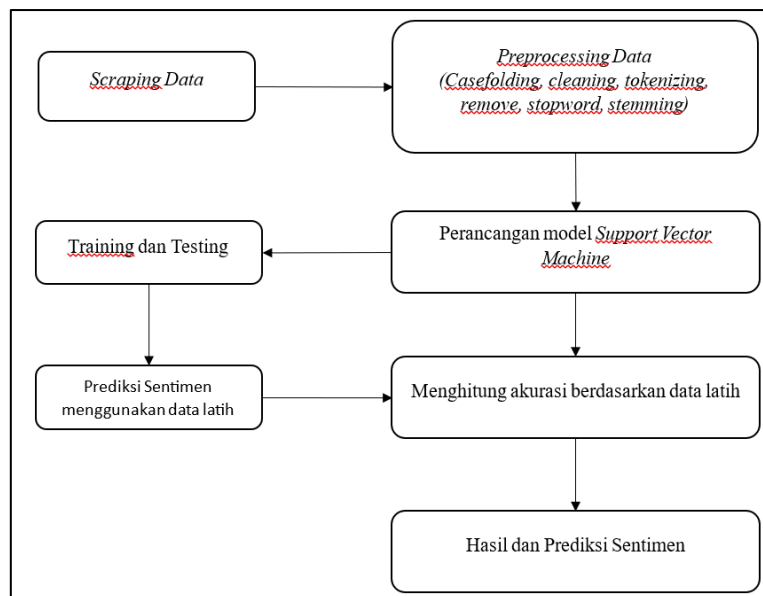
Sistem dalam penelitian ini berfokus pada analisis sentiment *review* pada aplikasi Edlink Sevima. Dengan menggunakan metode *Support Vector Machine* (SVM), analisis sentimen adalah studi ilmiah tentang analisis teks, pemrosesan bahasa alami, dan linguistik komputasional untuk mengidentifikasi, mengekstrak, dan mempelajari informasi subjektif dari data tekstual secara ilmiah. Ulasan pengguna akan dibagi menjadi analisis sentimen positif, negatif dan netral, oleh sistem ini. Sasaran utama sistem ini diharapkan dapat memberikan pemahaman yang lebih mendalam tentang bagaimana ulasan pengguna aplikasi Edlink Sevima membahas kekuatan dan kelemahan aplikasi. Tim pengembangan aplikasi Edlink Sevima dapat menggunakan informasi ini untuk merumuskan kebijakan dengan memahami persepsi dan respons terhadap umpan balik pengguna.

Fungsi Fungsi utama dalam analisis sistem pada penelitian yang dilakukan, yaitu :

1. Sistem memungkinkan pengguna untuk mengumpulkan data *review* pada aplikasi Edlink Sevima di *Google Play Store*. Pengumpulan data menggunakan teknik *scraping* dengan *library python google play scraper*.
2. Data yang telah diambil kemudian diproses untuk mempersiapkan data sebelum diolah oleh algoritma *Support Vector Machine*. Proses meliputi penghapusan karakter yang tidak diinginkan, penggantian kata yang tidak relevan dengan sentimen, dan normalisasi data.
3. Menggunakan metode TF-IDF (Term Frequency-Inverse Document Frequency) untuk mengubah tes *review* menjadi representasi numerik yang dapat dimengerti oleh metode *Support Vector Machine*. Ekstraksi fitur ini memungkinkan analisis teks menggunakan algoritma *Machine Learning*.
4. Melalui penggunaan metode *Support Vector Machine*, sistem mampu mengklasifikasikan sentiment pada setiap *review* menjadi kategori positif, negatif dan netral.
5. Hasil analisis sentiment ditampilkan melalui antarmuka pengguna yang intuitif. Pengguna dapat melihat visualisasi, statistik, dan pola sentimen yang muncul dalam *review* pada aplikasi Edlink Sevima.

B. Prosedur Analisis

Prosedural penelitian mengusulkan proses analisis sentiment terhadap *review* pada aplikasi Edlink Sevima dengan memperhatikan beberapa variable seperti total ulasan atau variable lainnya yang mempengaruhi dalam sistem untuk mengklasifikasi suatu teks sentiment menjadi positif, negatif dan netral. Tahapan prosedur untuk melakukan penelitian dapat dilihat pada gambar berikut.



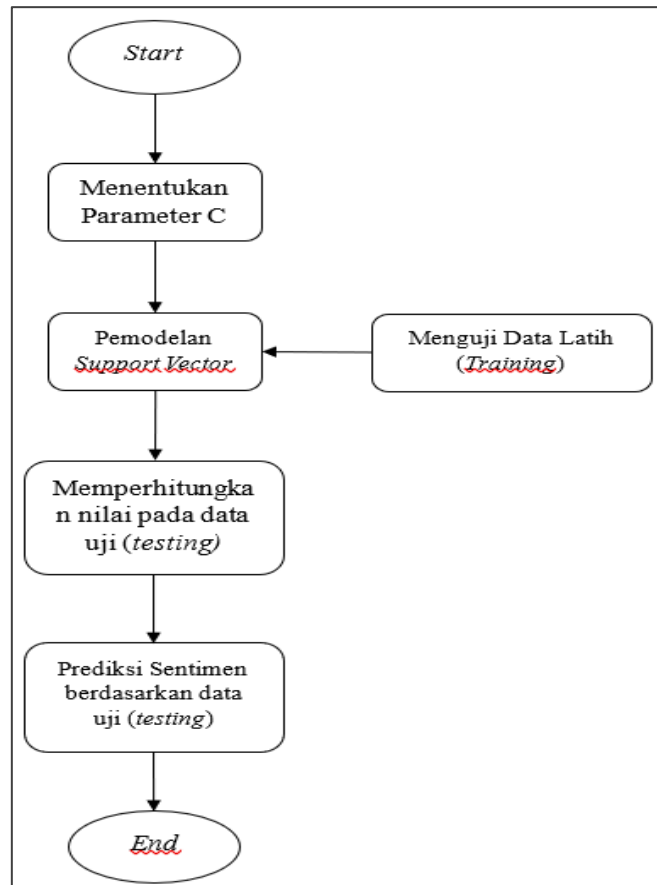
Gambar 3. 1 Tahapan Prosedur

Setelah data berhasil diperoleh, agar ulasan sebagai kalimat yang memiliki makna dalam analisis sentimen, akan dilakukan *Preprocessing data*. Proses ini dimulai dengan *casefolding* untuk melakukan penyeragaman kata menjadi huruf kecil, diikuti dengan *cleaning* untuk menghapus elemen yang tidak relevan, seperti karakter khusus dan tautan. Selanjutnya, dilakukan *tokenizing* untuk memecah kalimat menjadi kata-kata individu, diikuti dengan penghapusan *stopword* untuk menghilangkan istilah yang tidak penting. Tahap berikutnya adalah *stemming*, yaitu mengambil kata dasar dari kata kerja atau kata sifat. Pada tahap akhir *preprocessing*, fitur akan diekstraksi untuk digunakan dalam model. Setelah melakukan *preprocessing*, model

Support Vector Machine akan dibangun untuk melatih dan menguji model tersebut. Kemudian, sentimen akan diprediksi berdasarkan data yang telah diuji, dan nilai akurasi dari model akan dihitung untuk mengevaluasi kinerjanya.

C. Support Vector Machine (SVM)

SVM adalah salah satu metode dalam supervised machine learning untuk klasifikasi dengan tugas utama untuk mengatasi permasalahan klasifikasi dengan linear maupun non-linear. Metode ini memiliki fungsi untuk mencari hyperplane terbaik dengan memaksimalkan sebuah jarak antar kelas. SVM adalah salah satu metode dalam supervised machine learning untuk klasifikasi dengan tugas utama untuk mengatasi permasalahan klasifikasi dengan linear maupun non-linear. Metode ini memiliki fungsi untuk mencari hyperplane terbaik dengan memaksimalkan sebuah jarak antar kelas. Untuk perhitungan hyperplane pada SVM akan menggunakan rumus berikut.



Gambar 3.2 Alur Metode Support Vector Machine

Berdasarkan pada gambar tersebut dapat disimpulkan beberapa proses dalam metode tersebut dimulai dari menentukan parameter “c” lalu akan dilakukan pemodelan support vector dan hyperplane pada data training sehingga akan dilakukan perhitungan nilai pada setiap data testing yang dibutuhkan, kemudian akan dilakukan prediksi sentimen positif, negatif dan netral.

D. Pengumpulan Data

Tahap pengumpulan data dilakukan dengan menggunakan teknik *Scraping* dengan library python *Google Play Scraper*. Data yang diambil sejumlah 2.153 review dengan sorting *Most Relevant*. Pada python, perlu melakukan instalasi library *Google Play Scraper*, dilanjutkan dengan *scraping* data. Dan gambar dibawah menjelaskan tentang untuk mengolah data ulasan aplikasi yang berhasil diambil dalam bentuk dataframe, selanjutnya data disimpan menjadi file csv.

| reviewId | userName | userImage | content | score | thumbsUpCount | reviewCreatedVersion | at | replyContent |
|----------|---|--------------------------------------|---|-------|---------------|----------------------|---------------------|---|
| 0 | 6289d1d8-4055-4582-a44c-e6bbe0bf50e4 Ihham sharif | lh.googleusercontent.com/a-/ALV-U... | Aplikasinya ok, ui juga simpel. Tapi gaada tem... | 3 | 3 | 4.8.1 | 2024-05-06 11:16:45 | Halo Kak Ihham. Terima kasih atas ulasannya. B... |
| 1 | 976bdef-bd85-4541-bc08-8d57144ef7d7 gilang abitio | lh.googleusercontent.com/a-/ALV-U... | Tolong min ini!!! kenapa aplikasi saya pas say... | 3 | 4 | 4.8.1 | 2024-05-25 10:56:08 | Halo Kak Gilang, mohon maaf atas ketidaknyaman... |
| 2 | 0a05ce45-6f90-4d36-99f6-ecbead972e07 22.Noviana Andriani | lh.googleusercontent.com/a-/ALV-U... | Saya kasih ulasan bintang 3 karena ketika meng... | 3 | 0 | 4.8.1 | 2024-07-22 02:21:34 | Halo Kak Noviana, mohon maaf untuk ketidaknyam... |
| 3 | e1f96c10-ca28-46e0-960f-47e47b48e9d5 Tara Alamsyah | lh.googleusercontent.com/a-/ALV-U... | Berubah pikiran, ketika saya memanage pemerks... | 1 | 16 | 4.8.1 | 2024-08-09 04:35:23 | Hai Kak Tara, mohon maaf untuk ketidaknyamanan... |

Gambar 3.3 Data Berhasil Diambil Dalam Bentuk DataFrame

E. Preprocessing Data

Preprocessing data adalah proses penting dalam analisis data yang dilakukan sebelum data dapat diolah lebih lanjut atau dimodelkan. Tujuannya adalah untuk membersihkan, mengatur dan mempersiapkan data agar siap digunakan dalam analisis, eksplorasi, dan pemodelan. Preprocessing data memiliki peran krusial dalam menjaga kualitas dan keandalan hasil analisis, serta memastikan bahwa data yang digunakan sesuai dengan tujuan penelitian. review yang didapatkan kemudian dilakukan processing.[8] Berikut ini adalah tahapan pada preprocessing data, yaitu :

1. Tahap Labeling/Sentiment

Pada Pada proses labeling teks telah ditentukan bahwa terdapat tiga label sentimen, yaitu positif, netral dan negatif. Dengan menggunakan library tersebut, tidak lagi melakukan pelabelan secara manual. Yang dimana, penetapan tiga sentiment ini berdasarkan score yang diberikan oleh pengguna Edlink Sevima. Untuk score 5-4 adalah sentiment positif dengan kode 2, score 3 adalah netral dengan kode 1, dan rating 1-2 adalah negatif dengan kode 0. Proses ini penting dalam analisis sentimen karena memungkinkan pengelompokan ulasan ke dalam katagori positif, netral dan negatif, yang selanjutnya dapat digunakan untuk analisis lebih lanjut, dapat dilihat pada gambar berikut.[9]

| content | score | Year | Month | Day | sentiment |
|--|-------|------|-------|-----|-----------|
| 2039 Sangat bermanfaat untuk menunjang proses perku... | 5 | 2017 | 2 | 23 | 2 |
| 2033 sangat bermanfaat sekali. masukan dari saya, p... | 5 | 2017 | 3 | 14 | 2 |
| 2058 Coba dulu | 3 | 2017 | 3 | 30 | 1 |
| 2032 Apa ini hanya untuk dosen saja?bagaimana untuk... | 4 | 2017 | 5 | 3 | 2 |
| 2036 Bagus tpi kok forceclose yah stelah seminggu d... | 5 | 2017 | 5 | 20 | 2 |

Gambar 3.4 Tahap Labeling

2. Tahap Casefolding

Di tahapan membersihkan seperti clean tanda baca, nomor dan lain-lain. Casefolding adalah salah satu tahapan penting dalam preprocessing data teks yang dilakukan dalam rangka membersihkan dan mengatur data sebelum dilakukan analisis lebih lanjut. Pada tahap ini, seluruh huruf dalam teks diubah menjadi huruf kecil. Tujuannya adalah untuk mencapai konsistensi dalam teks dan menghindari perbedaan yang timbul akibat variasi kapitalisasi. Tujuan utama dari langkah ini adalah untuk mencapai konsistensi dalam teks, sehingga kata yang sama dengan variasi kapitalisasi dapat dianggap sebagai satu entitas Langkah casefolding teks memiliki manfaat penting dalam proses preprocessing data.[10]

3. Tahapan StopWord

Tahapan ini adalah proses dalam pemrosesan bahasa alami (Natural Language Processing, NLP) yang bertujuan untuk menghapus kata-kata umum yang tidak memberikan makna penting dalam anallisis teks. Kata-kata ini, yang dikenal sebagai "StopWords" sering kali mencakup kata-kata seperti "dan," "atau," "adalah," "dari," dan "yang." Meskipun kata-kata ini memiliki fungsi gramatikal yang penting dalam kalimat, mereka biasanya tidak memberikan informasi yang signifikan untuk analisis sentimen atau pengklasifikasian teks.[11]

4. Tahap Tokenizing

Pada tahapan tokenize merupakan proses dalam pemrosesan bahasa alami yang bertujuan untuk membagi teks menjadi unit-unit yang lebih kecil yang disebut token. Token ini bisa berupa kata, frasa, atau bahkan karakter, tergantung pada tujuan analisis. Proses tokenizing merupakan langkah awal yang penting dalam analisis teks karena memungkinkan pengolahan data yang lebih terstruktur dan terorganisir. Tujuan nya agar mempermudah analisis dan mendukung proses selanjutnya yang mana dapat mengurangi kompleksitas. Berikut bisa dilihat perbedaannya pada gambar berikut setelah di lakukan script StopWord.[12]

| | content | score | Year | Month | Day | sentiment | content_token |
|------|--|-------|------|-------|-----|-----------|--|
| 2039 | sangat bermanfaat untuk menunjang proses perku... | 5 | 2017 | 2 | 23 | 2 | [bermanfaat, menunjang, proses, perkuliahan] |
| 2033 | sangat bermanfaat sekali. masukan dari saya, p... | 5 | 2017 | 3 | 14 | 2 | [bermanfaat, masukan, mengetik, kolom, komenta... |
| 2058 | coba dulu | 3 | 2017 | 3 | 30 | 1 | [coba] |
| 2032 | apa ini hanya untuk dosen saja?bagaimana untuk... | 4 | 2017 | 5 | 3 | 2 | [dosen, guru, smk, sma, pas, daftar, ko, yg, s... |
| 2036 | bagus tpi kok forceclose yah setelah seminggu d... | 5 | 2017 | 5 | 20 | 2 | [bagus, tpi, forceclose, yah, setelah, seminggu... |

Gambar 3.5 Hasil Tokenizing setelah di StopWord

5. Tahapan *Stemming*

Pada tahapan *Stemming* bertujuan untuk mengurangi kata-kata ke bentuk dasarnya atau akar kata (*Stem*). Proses ini penting dalam analisis teks, termasuk analisis sentimen, karena membantu dalam menyederhanakan variasi kata yang memiliki makna yang sama. Dengan mengubah kata-kata ke bentuk dasarnya, model analisis dapat lebih efektif dalam memahami konteks dan makna dari teks yang dianalisis. Tujuan dari *stemming* untuk mengurangi variasi kata dalam banyak bahasa, kata dapat memiliki berbagai bentuk berdasarkan konjugasi, deklinasi, atau afiksasi.[13]

Stemming juga membantu mengurangi variasi ini sehingga kata-kata yang memiliki makna serupa dapat dikelompokkan bersama. Dengan menggunakan bentuk dasar kata, model analisis dapat lebih fokus pada makna yang relevan daripada membingungkan variasi kata yang tidak perlu.[14] Dan juga mengurangi jumlah bentuk kata yang berbeda, *stemming* juga membantu dalam mengurangi dimensi data, yang dapat mempercepat proses pelatihan model dan mengurangi kebutuhan akan sumber daya komputasi. Contohnya seperti kata-kata "belajar," "belajarannya," dan "belajarnya" akan diubah menjadi bentuk dasar "belajar" setelah proses stemming. Dengan cara ini, semua variasi kata yang berkaitan dengan konsep belajar dapat dikelompokkan bersama, yang membantu dalam analisis sentimen yang lebih efektif. Perbandingan teks setelah dilakukan *stemming* dapat dilihat pada gambar berikut ini.[15]

| | content | score | Year | Month | Day | sentiment | content_token | stemmed |
|------|--|-------|------|-------|-----|-----------|--|--|
| 2039 | sangat bermanfaat untuk menunjang proses perku... | 5 | 2017 | 2 | 23 | 2 | [bermanfaat, menunjang, proses, perkuliahan] | [manfaat, tunjang, proses, kuliah] |
| 2033 | sangat bermanfaat sekali. masukan dari saya, p... | 5 | 2017 | 3 | 14 | 2 | [bermanfaat, masukan, mengetik, kolom, komenta... | [manfaat, masuk, etik, kolom, komentar, text, ... |
| 2058 | coba dulu | 3 | 2017 | 3 | 30 | 1 | [coba] | [coba] |
| 2032 | apa ini hanya untuk dosen saja? bagaimana untuk... | 4 | 2017 | 5 | 3 | 2 | [dosen, guru, smk, sma, pas, daftar, ko, yg, s... | [dosen, guru, smk, sma, pas, daftar, ko, yg, s... |
| 2036 | bagus tpi kok forceclose yah setelah seminggu d... | 5 | 2017 | 5 | 20 | 2 | [bagus, tpi, forceclose, yah, setelah, seminggu... | [bagus, tpi, forceclose, yah, setelah, minggu, ... |

Gambar 3.6 Perbandingan Setelah Tahap *Stemming*

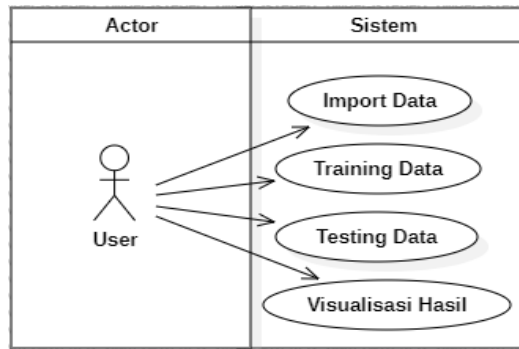
F. Perancangan Sistem Dan Prosedural

Perancangan sistem ini untuk mengalisa sentimen terhadap ulasan di Edlink Sevima . Lalu diperoleh hasil katagori terdapat dua tipe sentimen yaitu sentimen positif, sentimen negatif dan sentiment netral. Setelah dilakukan pengkatagorian tersebut, selanjutnya akan dilakukan pemodelan svm dengan tujuan melatih sistem untuk melakukan prediksi sentimen lalu menerapkan data uji (*testing*). Pengujian akan dilakukan dengan metode *Support Vector Machine* yang telah dilatih untuk melakukan prediksi sentimen. Pada nilai akurasi dalam memprediksi sentimen acuan yang dapat digunakan adalah nilai dari metode evaluasi *confusion matrix* seperti nilai *precision*, *recall*, dan *f1 score*.

Setelah melakukan pelatihan model menggunakan *Support Vector Machine* jenis linear berikutnya adalah mengembangkan sebuah aplikasi sederhana yang bertujuan untuk memprediksi sentimen pada data uji (*testing*) ataupun data baru yang akan dimasukkan oleh pengguna. Untuk memudahkan dalam memahami perancangan prosedural, akan dijelaskan pada beberapa diagram seperti *use case diagram*, *activity diagram*, *sequence diagram* dan *class diagram*.

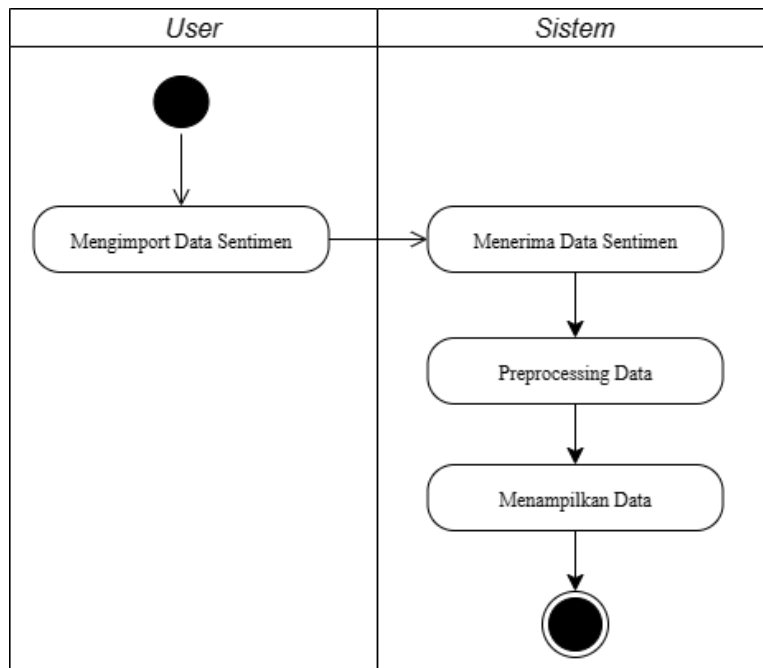
G. Use Case Diagram, Activity Diagram, Dan Sequence Diagram

Pada *use case*, peran yang dimiliki oleh *user* yaitu mengimport data ulasan yang telah dilabeli secara otomatis pada tahap sebelumnya yaitu labeling pada data latih, baik skenario 1 dengan data *Training* sebesar 70% dan data *Testing* sebesar 30%, skenario 2 dengan data *Training* sebesar 80% dan data *Testing* sebesar 20%, dan skenario 3 dengan data *Training* sebesar 90% dan data *Testing* sebesar 10%. *Use case diagram user* dapat dilihat pada gambar berikut.



Gambar 3.7 Use Case Diagram User

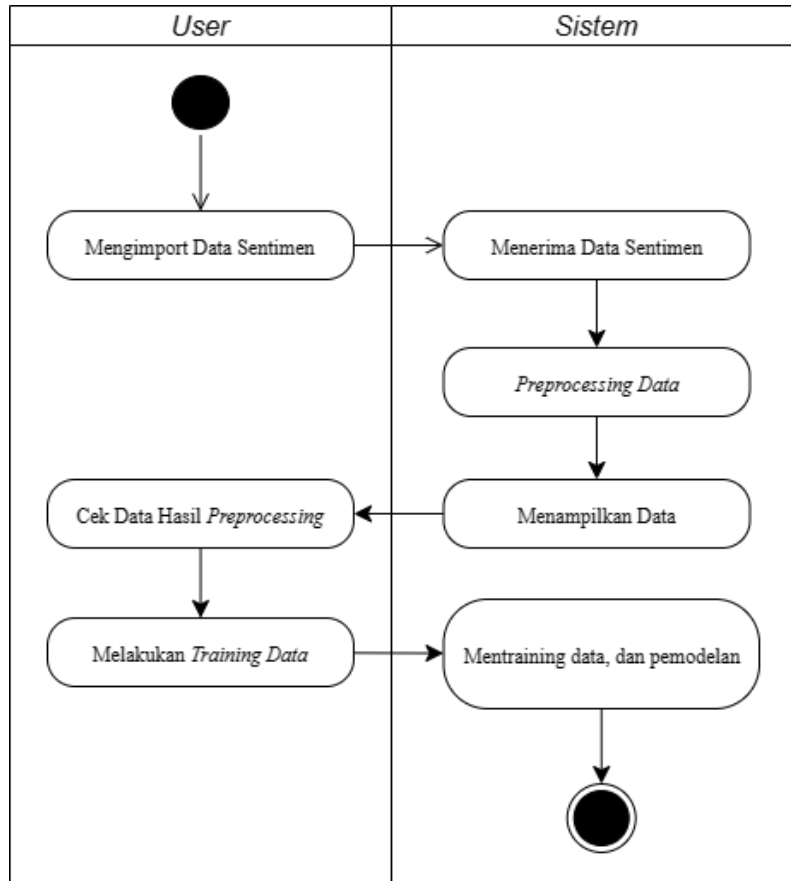
Pada *activity diagram* melakukan import data, aktifitas *user* dalam melakukan import data pada sistem dapat dilihat pada gambar berikut.



Gambar 3.8 Activity Diagram Melakukan Import Data

Pada gambar di atas dapat dilihat bahwa *user* telah berada pada halaman awal aplikasi, *user* memasukkan data untuk diprediksi hasil sentimen oleh sistem, lalu sistem akan mempreprocessing data tersebut sebelum menampilkan hasil prediksi sentimen.

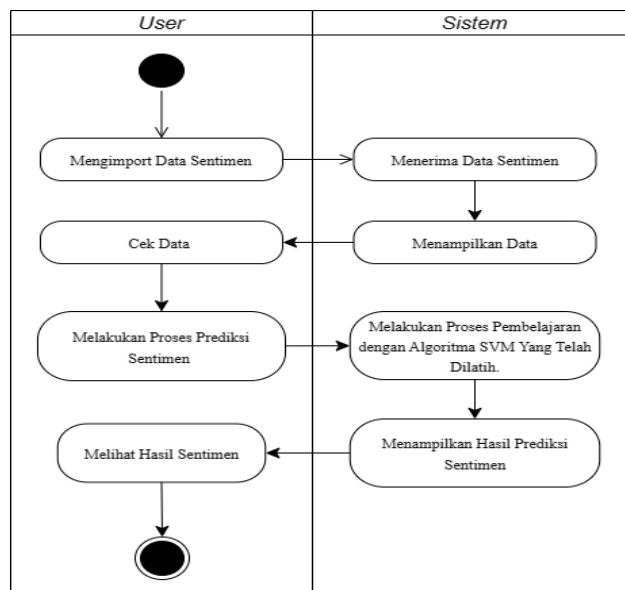
Activity Diagram melakukan *training data*, Aktifitas *user* untuk melakukan *training data* pada sistem dapat dilihat pada gambar berikut.



Gambar 3.9 Activity Diagram Melakukan Training Data

Pada gambar diatas dapat dilihat bahwa pengguna telah melakukan *training data* lalu sistem menerima hasil *training data* dan pemodelan, lalu *user* dapat melakukan pengujian untuk melakukan prediksi sentimen.

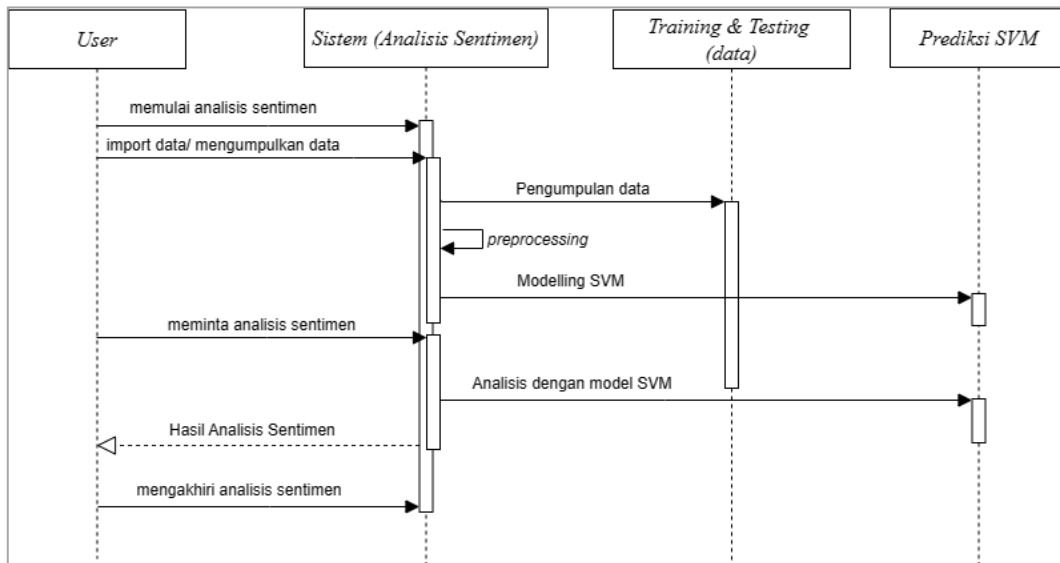
Activity Diagram melakukan *testing data*, Aktifitas *user* dalam melakukan *testing data* pada sistem dapat dilihat pada gambar berikut.



Gambar 3.10 Activity Diagram Memprediksi Sentimen

Pada gambar diatas merupakan proses yang berhubungan antara *user* dan sistem, melalui dari sistem menampilkan hasil *preprocessing*, kemudian pengguna melakukan *training data* untuk prediksi sentimen dengan menggunakan pembelajaran *Support Vector Machine* yang telah dilatih hingga sistem menampilkan hasil prediksi sentimen terhadap data baru ke *user*.

Sequence Diagram merupakan gambar yang menunjukkan bagaimana komponen dalam sistem berinteraksi antara satu dan lainnya melalui pesan dalam urutan waktu. *Sequence diagram* dapat dilihat pada gambar berikut.



Gambar 3.11 Sequence Diagram

Penjelasan terhadap komponen dalam *Sequence Diagram* pada gambar 3.11 diatas, sebagai berikut :

1. *User*/pengguna adalah aktor yang memulai proses analisis sentimen terhadap *review* di aplikasi Edlink Sevima. *User* memulai analisis, mengumpulkan data dan meminta hasil analisis.
2. Sistem (analisis sentimen) adalah bagian pada sistem yang bertanggung jawab atas analisis sentimen yang melibatkan tiga tahap utama seperti pengumpulan data sentimen, *preprocessing data*, dan pembuatan model prediksi *Support Vector Machine*. Sistem kemudian melakukan analisis sentimen terhadap data yang diinginkan oleh *user*/pengguna.
3. *Training* dan *Testing* merupakan representasi dari data sentimen yang telah dikumpulkan dengan sistem yang mengambil data dari sumber kemudian melakukan proses pelatihan dan pengujian data.
4. Prediksi *Support Vector Machine* adalah pemodelan SVM yang telah dilatih oleh sistem untuk melakukan analisis sentimen. Pemodelan ini digunakan oleh sistem untuk memprediksi sentimen berdasarkan data yang diberikan.

Interaksi dalam *sequence diagram* pada gambar 3.11, sebagai berikut :

1. *User* memulai proses analisis sentimen dengan mengirim pesan ke sistem.
2. Sistem mulai bekerja untuk mengumpulkan data sentimen dari data *training* dan *testing*.
3. Setelah mengumpulkan data, sistem melakukan *preprocessing* data untuk membersihkan dan mempersiapkannya.
4. Kemudian sistem melatih model *Support Vector Machine* dengan data yang telah diproses.
5. *User* meminta analisis sentimen dengan mengirim pesan ke sistem.
6. Sistem menggunakan model *Support Vector Machine* untuk menganalisis sentimen.
7. Hasil analisis sentimen dikirim kembali ke *user*/pengguna.

H. Pengujian Sistem

Dalam pengujian sistem, akan dijelaskan mengenai pelaksanaan dari model *Support Vector Machine* untuk mengetahui sentiment pada ulasan di Edlink Sevima. Pada bab ini juga menguraikan langkah-langkah dalam upaya memperoleh model analisa sentimen dengan melakukan pemodelan *Support Vector Machine*.

I. Pembuatan DataSet untuk Model SVM

Setelah *review* berhasil diterapkan tahap *preprocessing*, selanjutnya adalah mempersiapkan data. Untuk membaca data yang sebelumnya telah dilalui tahap *preprocessing* dapat dilihat pada gambar dibawah ini.

```

1 import pandas as pd
2 df=pd.read_csv('HasilPreprocessing_data_Rivan_2agustus.csv')
3 df = df[['text_string', 'sentiment']]
4 df

```

| | text_string | sentiment |
|------|---|-----------|
| 0 | manfaat tunjang proses kuliah | 2 |
| 1 | manfaat masuk etik kolom komentar text area ny... | 2 |
| 2 | coba | 1 |
| 3 | dosen guru daftar mahasiswa dosen | 2 |
| 4 | bagus forceclose setelah minggu pakai | 2 |
| ... | ... | ... |
| 2148 | kasih ulas bintang quiz edlink klik tombol nex... | 1 |
| 2149 | mantap | 2 |
| 2150 | aplikasi ribet | 0 |
| 2151 | bagus | 2 |
| 2152 | mantap | 2 |

2153 rows x 2 columns

Gambar 3.12 Dataset Sentimen

Lalu memeriksa jumlah pada masing-masing sentimen. Contoh data sentimen yang digunakan peneliti menggunakan jumlah data positif 1049, data negatif 888 dan data netral 216, yang dapat dilihat pada gambar dibawah ini.

```

Jumlah masing-masing sentiment

1 #Untuk melihat masing-masing data
2 jml = df['sentiment'].value_counts()
3 print(jml)

```

```

sentiment
2    1049
0     888
1     216
Name: count, dtype: int64

```

Gambar 3.13 Jumlah masing-masing sentimen

J. Evaluasi Skenario

Proses evaluasi pada penelitian akan menggunakan *Confussion matrix* untuk menyatakan jumlah data uji yang benar diklasifikasikan dan jumlah data uji yang salah diklasifikasikan. Hasil evaluasi *confussion matrix* pada skenario 1, skenario 2, dan skenario 3 dapat dilihat pada gambar dibawah ini berikut gambarnya.

```

Confusion Matrix Skenario 1:
[[212  0  54]
 [ 44  0  21]
 [ 52  0 263]]

```

Gambar 3.14 Hasil Confusion Matrix Skenario 1

| Classification Report Skenario 1: | | | | |
|-----------------------------------|-----------|--------|----------|---------|
| | precision | recall | f1-score | support |
| 0 | 0.69 | 0.80 | 0.74 | 266 |
| 1 | 1.00 | 0.00 | 0.00 | 65 |
| 2 | 0.78 | 0.83 | 0.81 | 315 |
| accuracy | | | 0.74 | 646 |
| macro avg | 0.82 | 0.54 | 0.51 | 646 |
| weighted avg | 0.76 | 0.74 | 0.70 | 646 |

Gambar 3.15 Hasil Evaluasi Pada Skenario 1

Dari hasil evaluasi pengujian pada skenario 1, dapat diketahui nilai nilai *accuracy* metode *Support Vector Machine* sebesar 74%. Pada sentimen negatif *precision* memiliki nilai sebesar 0.69, pada sentimen netral *precision* memiliki nilai sebesar 1.00 dan untuk sentimen positif *precision* memiliki nilai 0.78. kemudian untuk sentimen negatif *recall* memiliki nilai sebesar 0.80, untuk sentimen netral memiliki nilai *recall* sebesar 0.00 dan untuk sentimen positif memiliki nilai *recall* sebesar 0.83. *f1-score* untuk sentimen negatif memiliki nilai sebesar 0.74, *f1-score* pada sentiment netral memiliki nilai sebesar 0.00 dan *f1-score* pada sentiment positif memiliki nilai sebesar 0.81. untuk kelas negatif jumlah data aktual *support* terdapat 266 data, untuk kelas netral jumlah data aktual *support* terdapat 65 data dan untuk kelas positif jumlah data aktual terdapat *support* 315 data.

```
Confusion Matrix Skenario 2:
[[138  1  39]
 [ 28  0  15]
 [ 32  0 178]]
```

Gambar 3.16 Hasil Confusion Matrix Skenario 2

| Classification Report Skenario 2: | | | | |
|-----------------------------------|-----------|--------|----------|---------|
| | precision | recall | f1-score | support |
| 0 | 0.70 | 0.78 | 0.73 | 178 |
| 1 | 0.00 | 0.00 | 1.00 | 43 |
| 2 | 0.77 | 0.85 | 0.81 | 210 |
| accuracy | | | 0.73 | 431 |
| macro avg | 0.49 | 0.54 | 0.85 | 431 |
| weighted avg | 0.66 | 0.73 | 0.80 | 431 |

Gambar 3.17 Hasil Evaluasi Pada Skenario 2

Dari hasil evaluasi pengujian pada skenario 2, dapat diketahui nilai nilai *accuracy* metode *Support Vector Machine* sebesar 73%. Pada sentimen negatif *precision* memiliki nilai sebesar 0.70, pada sentimen netral *precision* memiliki nilai sebesar 0.00 dan untuk sentimen positif *precision* memiliki nilai 0.77. kemudian untuk sentimen negatif *recall* memiliki nilai sebesar 0.78, untuk sentimen netral memiliki nilai *recall* sebesar 0.00 dan untuk sentimen positif memiliki nilai *recall* sebesar 0.85. *f1-score* untuk sentimen negatif memiliki nilai sebesar 0.73, *f1-score* pada sentiment netral memiliki nilai sebesar 1.00 dan *f1-score* pada sentiment positif memiliki nilai sebesar 0.81. untuk kelas negatif jumlah data aktual *support* terdapat 178 data, untuk kelas netral jumlah data aktual *support* terdapat 43 data dan untuk kelas positif jumlah data aktual terdapat *support* 210 data.

```
Confusion Matrix Skenario 3:
[[70  0 19]
 [15  0  7]
 [14  0 91]]
```

Gambar 3.18 Hasil Confusion Matrix Skenario 3

| Classification Report Skenario 3: | | | | |
|-----------------------------------|-----------|--------|----------|---------|
| | precision | recall | f1-score | support |
| 0 | 0.71 | 0.79 | 0.74 | 89 |
| 1 | 1.00 | 0.00 | 0.00 | 22 |
| 2 | 0.78 | 0.87 | 0.82 | 105 |
| accuracy | | | 0.75 | 216 |
| macro avg | 0.83 | 0.55 | 0.52 | 216 |
| weighted avg | 0.77 | 0.75 | 0.71 | 216 |

Gambar 3.19 Hasil Evaluasi Pada Skenario 3

Dari hasil evaluasi pengujian pada skenario 3, dapat diketahui nilai nilai *accuracy* metode *Support Vector Machine* sebesar 75%. Pada sentimen negatif *precision* memiliki nilai sebesar 0.71, pada sentimen netral *precision* memiliki nilai sebesar 1.00 dan untuk sentimen positif *precision* memiliki nilai 0.78. kemudian untuk sentimen negatif *recall* memiliki nilai sebesar 0.79, untuk sentimen netral memiliki nilai *recall* sebesar 0.00 dan untuk sentimen positif memiliki nilai *recall* sebesar 0.87. *f1-score* untuk sentimen negatif memiliki nilai sebesar 0.74, *f1-score* pada sentiment netral memiliki nilai sebesar 0.00 dan *f1-score* pada sentiment positif memiliki nilai sebesar 0.82. untuk kelas negatif jumlah data aktual *support* terdapat 89 data, untuk kelas netral jumlah data aktual *support* terdapat 22 data dan untuk kelas positif jumlah data aktual terdapat *support* 105 data. Dari hasil evaluasi pengujian dari ketiga skenario dapat diambil kesimpulan bahwa skenario terbaik berada pada skenario 3 yaitu dengan data latih sebesar 90% dan data uji sebesar 10%, dengan akurasi 75%.

K. Pengujian Prediksi Sentimen

```

23 # Menampilkan hasil prediksi untuk teks pertama
24 print(f"Hasil klasifikasi dari teks \"{text1[0]}\" adalah {prediksi1[0]}")
25
26 # Menampilkan hasil prediksi untuk teks kedua
27 print(f"Hasil klasifikasi dari teks \"{text2[0]}\" adalah {prediksi2[0]}")
28
29 # Menampilkan hasil prediksi untuk teks ketiga
30 print(f"Hasil klasifikasi dari teks \"{text3[0]}\" adalah {prediksi3[0]}")
31
Hasil klasifikasi dari teks "manfaat tunjang proses kuliah" adalah 2
Hasil klasifikasi dari teks "tolong join kelas aplikasi tutup mohon" adalah 0
Hasil klasifikasi dari teks "bingun aplikasi langsung menu halama tolong tanggap trims" adalah 1

```

Gambar 3.20 Hasil Pengujian Prediksi Sentimen

Pada gambar 3.20 diatas merupakan hasil klasifikasi analisis sentimen ulasan pada Edlink Sevima berhasil dilakukan. Contoh jika *user* memasukkan kalimat yang mengandung kata pada label positif maka hasil klasifikasi akan positif yaitu dengan kode angka '2', sebaliknya jika user memasukkan kata kalimat yang mengandung kata pada label negatif maka hasil klasifikasi akan negatif dengan kode angka '0' dan jika memasukkan kata kalimat yang mengandung kalimat netral maka hasil klasifikasi akan netral dengan kode angka '1'. Yang dimana data testing diambil dari *sentiment* dan sedangkan data training yang diambil dari *text_string* yang telah dilatih sebelumnya menggunakan metode SVM setelah dilakukan preprocessing data.

IV. KESIMPULAN DAN SARAN

Kesimpulan yang dapat diambil dari hasil penelitian yang dilakukan dapat disimpulkan yaitu, sebagai berikut :

1. Metode *Support Vector Machine* mampu digunakan sebagai pendekatan untuk memprediksi analisis sentimen *review* pada Edlink Sevima.
2. Hasil nilai akurasi terbaik berada pada skenario 3 yaitu dengan nilai akurasi metode *Support Vector Machine* sebesar 75%. Pada sentimen negatif *precision* memiliki nilai sebesar 0.71, pada sentimen netral *precision* memiliki nilai sebesar 1.00 dan untuk sentimen positif *precision* memiliki nilai 0.78. kemudian untuk sentimen negatif *recall* memiliki nilai sebesar 0.79, untuk sentimen netral memiliki nilai *recall* sebesar 0.00 dan untuk sentimen positif memiliki nilai *recall* sebesar 0.87. *f1-score* untuk sentimen negatif memiliki nilai sebesar 0.74, *f1-score* pada sentiment netral memiliki nilai sebesar 0.00 dan *f1-score* pada sentiment positif memiliki nilai sebesar 0.82. untuk kelas negatif jumlah data aktual *support* terdapat 89 data, untuk kelas netral jumlah data aktual *support* terdapat 22 data dan untuk kelas positif jumlah data aktual terdapat *support* 105 data.
3. Dari hasil analisis sentimen menunjukkan hasil visualisasi bahwa benar melakukan prediksi pada teks baru yang telah dimasukkan.

Pada ini terdapat beberapa kekurangan, maka dari itu peneliti ingin memberikan saran agar pemahaman tentang analisis sentimen berkembang, berikut beberapa saran yang dapat peneliti simpulkan :

1. Agar praproses data jauh lebih baik tingkatkan teknik praproses data dengan menggunakan *stemming*, *lemmatization*, dan penghapusan kata-kata umum yang tidak bermakna (*stopwords*) dan gunakan teknik fitur ekstraksi yang lebih kompleks seperti TF-IDF atau *word embedding* misalnya *Word2Vec* atau *Bert*. Karena metode ini dapat menangkap konteks kata jauh lebih baik, yang berpotensi meningkatkan performa model *Support Vector Machine*.
2. Melakukan perbandingan kinerja dengan metode analisis sentimen lainnya, seperti *Random Forst*, *Naive Bayes*, dan *Decision tree*, yang dapat memberikan wawasan tentang efektivitas metode yang digunakan.
3. Disarankan untuk menyertakan visualisasi seperti confusion matrix, precision-recall curve, dan ROC curve dalam laporan evaluasi model. Visualisasi ini akan memberikan wawasan yang lebih jelas tentang performa model dan area yang perlu diperbaiki.

UCAPAN TERIMA KASIH

Saya ingin mengucapkan banyak terima kasih kepada saudari perempuan saya Ermala Yudarini, S.Sos., dan saudara laki-laki saya Briptu Sugi Hartono, yang selalu memberikan dukungan moral dan moril, serta semangat dan doa tanpa henti dalam setiap langkah selama saya menempuh perkuliahan saya dari awal sampai akhir. Tak lupa juga, saya mengucapkan banyak terima kasih juga kepada semua pihak yang turut serta berkontribusi, sehingga skripsi ini dapat terselesaikan dengan baik dan sukses.

REFERENSI

- [1] S. G. Rezeki, M. I. P. Nasution, and S. S. A. Sundari, "Application of E-Learning As a Learning Medium for Information Technology," *Romeo Review of Multidisciplinary Education Culture and Pedagogy*, vol. 2, no. 1, pp. 1–6, 2023, doi: 10.55047/romeo.v2i1.518.

- [2] T. N. Suharsono, A. Fauzan, and R. Mardiaty, "Sentiment Analysis of Covid-19 on Indonesian Twitter by Implementing the Naïve Bayes Method," in *2022 8th International Conference on Wireless and Telematics (ICWT)*, 2022, pp. 1–4. doi: 10.1109/ICWT55831.2022.9935134.
- [3] R. Maula, "Cara Kerja Aplikasi Sevima edLink Sebagai Solusi Media Pembelajaran Di Masa Pandemi," *AL-Hikmah: Jurnal Studi Islam*, vol. 1, no. 4, pp. 108–115, 2020.
- [4] S. S. Meilindha, "Efektivitas Penggunaan Aplikasi Sevima Edlink Dalam Proses Pembelajaran Pada Mahasiswa Prodi PAI FTIK IAIN Langsa," *Al-Ikhtibar: Jurnal Ilmu Pendidikan*, vol. 4, no. 2, pp. 528–537, 2017, [Online]. Available: <https://journal.iainlangsa.ac.id/index.php/ikhtibar/article/view/3208>
- [5] A. I. Tanggraeni and M. N. N. Sitokdana, "Analisis Sentimen Aplikasi E-Government pada Google Play Menggunakan Algoritma Naïve Bayes," *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, vol. 9, no. 2, pp. 785–795, 2022, doi: 10.35957/jatisi.v9i2.1835.
- [6] T. M. P. Aulia, N. Arifin, and R. Mayasari, "Perbandingan Kernel Support Vector Machine (Svm) Dalam Penerapan Analisis Sentimen Vaksinisasi Covid-19," *SINTECH (Science and Information Technology) Journal*, vol. 4, no. 2, pp. 139–145, 2021.
- [7] A. N. Ulfah and M. K. Anam, "Analisis Sentimen Hate Speech Pada Portal Berita Online Menggunakan Support Vector Machine (SVM)," *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, vol. 7, no. 1, pp. 1–10, 2020, doi: 10.35957/jatisi.v7i1.196.
- [8] S. Styawati, N. Hendrastuty, and A. R. Isnain, "Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine," *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 6, no. 3, pp. 150–155, 2021, doi: 10.30591/jpit.v6i3.2870.
- [9] R. H. Fauzi, "Pembangunan Kamus Bahasa Indonesia sebagai Sumber Daya Natural Language Processing (NLP) Bahasa Indonesia," vol. 2017, pp. 193–200, 2017.
- [10] I. Darajat, "Efektivitas TPS (Think Pair Share) dan STAD (Student Team Achievement Division) Menggunakan Sevima Edlink terhadap Prestasi Belajar PAI Siswa Kelas VIII di ...," 2024, [Online]. Available: http://etheses.iainponorogo.ac.id/id/eprint/27800%0Ahttp://etheses.iainponorogo.ac.id/27800/1/201200316_ILMI DAROJAT_ETHESIS.pdf
- [11] Y. X. Chu, X. G. Liu, and C. H. Gao, "Multiscale models on time series of silicon content in blast furnace hot metal based on Hilbert-Huang transform," *Proceedings of the 2011 Chinese Control and Decision Conference, CCDC 2011*, pp. 842–847, 2011, doi: 10.1109/CCDC.2011.5968300.
- [12] R. Feldman and J. Sanger, "Advanced Approaches in Analyzing Unstructured Data," *The Text Mining Handbook, Cambridge University*, 2007.
- [13] N. Herlinawati, Y. Yuliani, S. Faizah, W. Gata, and S. Samudi, "Analisis Sentimen Zoom Cloud Meetings di Play Store Menggunakan Naïve Bayes dan Support Vector Machine," *CESS (Journal Comput. Eng. Syst. Sci., vol. 5, no. 2, p. 293, 2020, doi: 10.24114/cess.v5i2.18186, 2020.*
- [14] R. Munawarah, O. Soesanto, and M. R. Faisal, "Penerapan Metode Support Vector Machine dalam Mengklasifikasikan Tweet Ujaran Kebencian Terhadap Tokoh Publik pada Twitter," *Kumpulan jurnaL Ilmu Komputer (KLIK)*, vol. 04, no. 01, pp. 103–113, 2016.
- [15] R. A. L. A. Renal and S. Alam, "Komparasi Payment DIgital untuk Analisis Sentimen Berdasarkan Ulasan di Google Playstore Menggunakan Metode Support Vector Machine," *STORAGE: Jurnal Ilmiah Teknik dan Ilmu Komputer*, vol. 2, no. 3, pp. 118–128, 2023.