

## ANALISIS SENTIMEN BERBASIS ASPEK PADA ULASAN APLIKASI KAI ACCESS MENGGUNAKAN METODE SUPPORT VECTOR MACHINE

Gracia Radiena<sup>1</sup>, Adi Nugroho<sup>2</sup>

<sup>1,2</sup>Fakultas Teknologi Informasi, Universitas Kristen Satya Wacana, Salatiga

[1672019260@student.uksw.edu](mailto:1672019260@student.uksw.edu), [2adi.nugroho@uksw.edu](mailto:2adi.nugroho@uksw.edu)

### ABSTRAK

PT Kereta Api Indonesia melakukan inovasi dengan meluncurkan aplikasi yang diberi nama KAI Access. Aplikasi KAI Access memiliki fitur pemesanan tiket, *ticket rescheduling*, pembatalan tiket hingga *e-boarding pass*. Tujuan dari penelitian ini adalah untuk mengetahui sentimen dari sebuah produk *mobile*. Opini terkait Aplikasi KAI Access dapat digunakan PT Kereta Api Indonesia sebagai parameter kunci untuk mengetahui tingkat kepuasan publik sekaligus bahan evaluasi bagi PT Kereta Api Indonesia. Berdasarkan hasil pengujian yang telah dilakukan pada ulasan pengguna aplikasi KAI Access dengan total 8.078 ulasan, lebih banyak pengguna memberikan opini positif dalam aspek *satisfaction* dan opini negatif pada aspek *learnability*, *efficiency*, dan *errors*. Digunakan model CRISP-DM (*Cross Industry Standard Process for Data Mining*) dan algoritma *Support Vector Machine* untuk melakukan klasifikasi. Hasil klasifikasi terbaik diperoleh nilai *accuracy*, *precision*, *recall*, dan *F-measure* yang dihasilkan dari tiap aspek yaitu untuk *Learnability* 94.73%, 100.00%, 89.50%, dan 94.64%, *Efficiency* 94.38%, 72.00%, 100.00%, dan 94.46%, *Errors* 85.13%, 97.11%, 72.41%, dan 82.96%, *Satisfaction* 87.26%, 98.46%, 73.78%, dan 84.20%.

**Kata kunci :** *KAI Access, Analisis Sentimen, Support Vector Machine*

### ABSTRACT

*PT Kereta Api Indonesia innovates by launching an application called KAI Access. The KAI Access application has features for ticket ordering, ticket rescheduling, ticket cancellation and e-boarding pass. The purpose of this study is to determine the sentiment of a mobile. Opinion regarding the KAI Access Application can be used by PT Kereta Api Indonesia as a key parameter to determine the level of public satisfaction as well as evaluation material for PT Kereta Api Indonesia. Based on the results of tests conducted on user reviews of the KAI Access application with a total of 8,078 reviews, more users give positive opinions on the satisfaction and negative opinions on the learnability, efficiency and errors. Model CRISP-DM (Cross Industry Standard Process for Data Mining) and Support Vector Machine algorithm are used to perform classification. The best classification results obtained accuracy, precision, recall, and F-measure resulting from each aspect, namely for Learnability 94.73%, 100.00%, 89.50%, and 94.64%, Efficiency 94.38%, 72.00%, 100.00%, and 94.46%, Errors 85.13%, 97.11%, 72.41%, and 82.96%, Satisfaction 87.26%, 98.46%, 73.78%, and 84.20%.*

**Keywords:** *KAI Access, Sentiment Analysis, Support Vector Machine*



## 1. PENDAHULUAN

Perkembangan dan kemajuan Teknologi Informasi dan Komunikasi (TIK) sangat diperlukan guna untuk memudahkan dan menyelesaikan berbagai masalah yang dihadapi oleh manusia dengan cepat. Selain itu, masyarakat saat ini menginginkan semuanya serba praktis dan tidak menyita banyak waktu. Pada zaman sekarang ini sarana transportasi merupakan suatu bagian yang tidak dapat dipisahkan dan sangat dibutuhkan masyarakat dalam kehidupan sehari-hari.

PT Kereta Api Indonesia (Persero) merupakan satu-satunya perusahaan BUMN milik pemerintah dibawah dinas perhubungan yang bergerak dibidang perkereta apian, setiap perusahaannya terletak dipulau Jawa maupun di pulau Sumatera. PT Kereta Api Indonesia juga melakukan inovasi dengan meluncurkan aplikasi yang diberi nama KAI Access. Aplikasi KAI Access memiliki fitur pemesanan tiket, ticket rescheduling, pembatalan tiket hingga e-boarding pass. Kereta api seakan menjadi salah satu alat transportasi favorit masyarakat Indonesia terbukti dengan semakin banyaknya layanan khusus Kereta Api Indonesia di berbagai perangkat Android, IOS, dan Windows Phone. Saat ini telah muncul bermacam-macam jenis dari aplikasi mobile yang ada, salah satunya adalah mobile ticketing. Mobile ticketing merupakan aplikasi mobile yang juga dapat digunakan untuk pemesanan serta pembelian tiket. Beberapa aplikasi tersebut seperti Treveloka, Tiket.com, Pegipegi, Tokopedia dan juga aplikasi favorit yang akan kita bahas yaitu KAI Access. Hal inilah yang menjadi salah satu faktor yang peneliti ingin teliti untuk mempelajari lebih lanjut tentang KAI Access dan tingkat kepuasan pengguna

terhadap aplikasi transportasi darat tersebut. Aspek penting yang perlu diperhatikan dalam mengembangkan aplikasi ini adalah persepsi pengguna terhadap keluhan pemesanan tiket kereta api. Pada saat pengguna merasa puas dengan suatu aplikasi, maka pengguna terus menggunakan aplikasi tersebut.

Penelitian ini berfokus pada analisis sentimen dengan ulasan pengguna Aplikasi KAI Access terhadap pemesanan tiket, menunjukkan bahwa penting untuk mempertimbangkan kualitas layanan, serta dianggap sebagai strategi bisnis, yang tujuannya adalah untuk memenuhi kepercayaan pengguna terhadap pelayanan yang diterima, keinginan dan kebutuhan pengguna aplikasi KAI Access. Tujuan penelitian ini adalah untuk mengetahui sentimen sebuah produk *mobile*. Selain itu perhitungan klasifikasi analisis sentimen dapat digunakan sebagai bahan evaluasi PT Kereta Api Indonesia yang bertujuan untuk meningkatkan produk serta kepuasan pelanggan.

Sentimen Analisis adalah proses untuk mengekstraksi data tekstual secara otomatis sehingga diperoleh informasi untuk mengetahui kecenderungan penilaian pada suatu objek, terdiri dari ulasan positif atau negatif [1]. Sentimen masyarakat dapat dijadikan acuan untuk pengambilan keputusan suatu produk [2].

Sentimen analisis dapat dilakukan dengan menggunakan metode klasifikasi. Beberapa metode yang paling umum digunakan adalah *Support Vector Machine* (SVM), *k-Nearest Neighbor* (k-NN), dan *Naive Bayes* (NB). Beberapa penelitian terkait analisis sentimen telah dilakukan, salah satunya adalah Pravina, dkk. menganalisis sentimen tentang opini maskapai penerbangan pada dokumen

Twitter menggunakan algoritma Support Vector Machine (SVM). Data diambil untuk dilakukannya penelitian ini merupakan data dari Twitter. Klasifikasi sentimen dilakukan dengan menggunakan fitur Lexicon Based, yang dapat menerima pendapat dalam bahasa selain bahasa Indonesia (bahasa Inggris digunakan dalam penelitian ini) untuk menganalisis sentimen. Algoritma support vector machine digunakan untuk melakukan klasifikasi. Hasil penelitian mengungkapkan parameter optimal dan dampak penggunaan Lexicon Based Features. Penelitian ini menggunakan parameter C dengan nilai 10 dan learning rate 0,03, serta menggunakan Lexicon Based Features dengan 50 iterasi. Hasil accuracy sebesar 40%, precision 40%, 100% recall, dan f-measure sebesar 57,14% [3].

Penelitian yang dilakukan oleh Wahyudi dan Kusumawardhana menganalisis sentimen pada review aplikasi Grab di Google Play Store menggunakan Support Vector Machine. Digunakan lebih dari 1.000 review pengguna yang diperoleh dari aplikasi Grab Indonesia di Google Play Store. Hasil Ulasan positif yang paling sering ditinjau adalah “ovo”, sedangkan Hasil Ulasan negatif yang paling sering ditinjau adalah “driver”. Akurasi analisis menggunakan Support Vector Machine sebesar 85,54% [4].

Penelitian Parasati, dkk. Membahas mengenai analisis sentimen menggunakan metode *Naïve Bayes Classifier* pada ulasan pelanggan restoran bakso president Malang berbasis aspek. Sebanyak 2.152 ulasan pelanggan tahun 2012 - 2019 diperoleh dari situs TripAdvisor dan Google Review menggunakan teknik web scraping. Hasil akurasi berdasarkan aspek yaitu Makanan 88%, Layanan 76% Layanan, dan Atmosfir 84% [5].

Penelitian Fitriana, dkk. Menganalisis sentimen opini vaksin *Covid-19* pada media sosial Twitter menggunakan SVM dan *Naive Bayes*. Hasil penelitian diperoleh bahwa algoritma Support Vector Machine memiliki performa lebih baik dengan akurasi 90,47%, presisi 90,23%, recall 90,78% dan algoritma *Naïve Bayes* dengan akurasi 88,64%, presisi 87,32%, recall 88,13%, dengan selisih akurasi 1,83%, presisi 2.91% dan recall 2.65%. Dari segi waktu, algoritma *Naive Bayes* mempunyai performa lebih baik dari SVM dengan nilai 8,1 detik dibanding dengan 11 detik untuk SVM. Hasil analisis untuk sentimen *Naïve Bayes* adalah 8,76% netral, 42,92% negatif, 48,32% positif, untuk SVM 10,56% netral, 41,28% negatif, positif 48,16% [6].

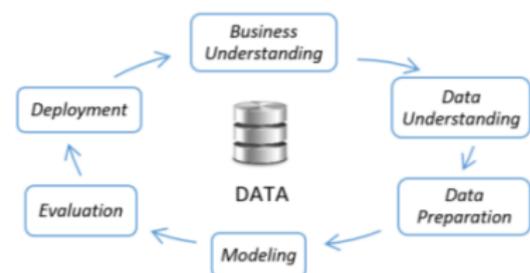
Metode SVM memiliki keunggulan tersendiri, seperti yang ditunjukkan oleh sejumlah penelitian sebelumnya. Algoritma *Support Vector Machine* (SVM) merupakan teknik yang populer untuk mengklasifikasikan teks. Support Vector Machine adalah suatu teknik untuk melakukan prediksi, baik dalam kasus klasifikasi maupun regresi. Menemukan *hyperplane* optimal yang memberi jarak atau pemisah antar dua kelas merupakan cara kerja *Support Vector Machine* [7]. Pada penelitian ini dilakukan analisis terkait Aplikasi KAI Access dilihat dari beberapa macam aspek. Namun, jumlah dataset ulasan dari masing-masing aspek yang digunakan mengalami *imbalance class*. Model yang menggunakan data tidak seimbang menghasilkan akurasi prediksi yang rendah [8]. *Synthetic Minority Over-sampling Technique* (SMOTE) merupakan salah satu cara untuk menangani *imbalance class*. SMOTE merupakan metode untuk membangkitkan data minoritas sehingga setara dengan data mayoritas [9].

Penelitian lain yang dilakukan Iskandar dan Nataliani membahas analisis sentimen Gadget berbasis aspek dengan perbandingan Naïve Bayes, SVM, dan k-NN. Menerapkan teknik SMOTE untuk mengatasi imbalance class. Dataset yang didapatkan berjumlah 9.597 komentar terkait Gadget Samsung Galaxy Z Flip 3 di Youtube. Hasil penelitian menunjukkan bahwa lebih banyak pengguna memberikan opini desain secara positif dan harga, spesifikasi, dan citra merk secara negatif. Hal ini terbukti dengan menggunakan model CRISP-DM bahwa model klasifikasi Support Vector Machine (SVM) menunjukkan hasil terbaik jika dibandingkan dengan Naive Bayes (NB), dan k-Nearest Neighbor (k-NN). Dilihat dari empat aspek, rata-rata accuracy SVM 96.43%. Aspek desain 94.40%, aspek harga 97.44%, aspek spesifikasi 96.22%, dan aspek citra merk 97.63% [9].

Berdasarkan sejumlah penelitian terdahulu tentang metode *Support Vector Machine* (SVM) dan penggunaan *Synthetic Minority Over-sampling Technique* (SMOTE) untuk mengatasi *imbalance class*. Dalam penelitian ini akan diuraikan metode klasifikasi *Support Vector Machine* (SVM) menggunakan SMOTE. Disamping itu, analisis sentimen pada penelitian ini berdasarkan aspek usability dari Nielsen (1994) yaitu *learnability*, *memorability*, *efficiency*, and *errors* [10]. Akan tetapi pada penelitian ini aspek *memorability* diperlakukan sebagai aspek yang sama dengan *learnability* karena memiliki arti yang hampir sama. *Output* dari penelitian ini adalah untuk mengetahui tingkat *accuracy*, *recall*, *precision*, dan *F-measure* terbaik dalam mengklasifikasi ulasan aplikasi KAI Access.

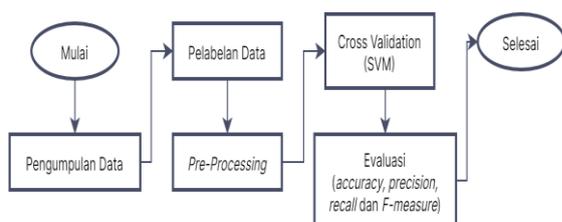
## 2. METODE PENELITIAN

Penelitian ini menggunakan metode CRISP-DM (*Cross-Industry Standard Process for Data Mining*). CRISP-DM adalah metode yang dapat diterapkan pada strategi pemecahan masalah umum dan metodologi yang menyediakan standar baku untuk data mining [11]. CRISP-DM adalah standar proses data mining sebagai strategi pemecahan masalah secara umum dari bisnis atau unit penelitian. Metode CRISP-DM (*Cross-Industry Standard Process for Data Mining*) terdiri dari beberapa tahapan, yaitu *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation* dan *deployment* [12]. Langkah-langkah dalam CRISP-DM dapat dilihat pada Gambar 1. Penelitian ini dilakukan hanya sampai pada tahap *evaluation* dan tidak dilanjutkan pada tahap *deployment* yaitu tahap implementasi pada *tools*.



Gambar 1. Model CRISP-DM

Penelitian ini menerapkan model CRISP-DM (*Cross-Industry Standard Process for Data Mining*). Tahapan penelitian ini seperti yang ditunjukkan pada Gambar 2. Tahapan penelitian terdiri dari pengumpulan data ulasan KAI Access, pelabelan data, *pre-processing*, *cross validation*, dan tahap evaluasi berdasarkan ulasan pengguna aplikasi KAI Access.



**Gambar 2.** Tahapan Penelitian

### Business Understanding

Pada tahapan *business understanding* ini, dilakukan pemahaman terkait objek atau masalah yang diangkat selama penelitian. Objek atau masalah yang diangkat adalah seberapa besar *accuracy*, *presisi*, *recall*, *f-measure* yang dihasilkan dari penerapan algoritma *Support Vector Machine*, dimana dataset yang digunakan diperoleh dari ulasan pengguna yang berkaitan dengan aplikasi KAI Access.

### Data Understanding

Pada tahap *data understanding*, dilakukan proses pengambilan data mentah sesuai dengan atribut yang dibutuhkan. Data diperoleh langsung dari pihak PT Kereta Api Indonesia (Persero). Data dikumpulkan dari Oktober 2021, hingga April 2022 sebanyak 8.078 data ulasan.

Ulasan pengguna yang tidak mencakup penelitian seperti ulasan mengenai nama versi aplikasi, *device* yang digunakan, dan ulasan lainnya tidak digunakan sebagai dataset. Setelah proses seleksi selesai, ada sekitar 1.261 ulasan yang didapatkan. Dataset yang digunakan diberi label secara manual, kemudian dilakukan proses identifikasi aspek. Empat kategori aspek yang diteliti yaitu *Learnability* (Mudah dipelajari), *Efficiency* (Efisien), *Errors* (Pencegahan kesalahan), dan *Satisfaction* (Kepuasan pengguna). Aspek ditentukan berdasarkan hasil analisis data ulasan pada aplikasi KAI Access. Aspek *Learnability*,

menjelaskan tingkat kemudahan pengguna untuk memenuhi task-task dasar ketika pertama kali mereka melihat atau menggunakan aplikasi KAI Access. Aspek *Efficiency*, menjelaskan tingkat kecepatan pengguna dalam menggunakan aplikasi. Aspek *Errors*, menjelaskan jumlah error yang ditemui pengguna selama penggunaan aplikasi. Aspek *Satisfaction*, adalah tingkat kepuasan pengguna dalam menggunakan aplikasi KAI Access.

Aspek yang dihasilkan diklasifikasikan menjadi dua kategori, yaitu positif dan negatif. Aspek positif ditandai dengan 1, negatif dengan -1 dan ulasan yang tidak mencakup ruang lingkup penelitian dengan 0. Berikut adalah contoh pelabelan ulasan terhadap masing-masing aspek ditunjukkan pada Tabel 1.

**Tabel 1.** Labeling Dataset

No	Ulasan	Learn ability	Efficiency	Errors	Satisfaction
1	Simpel, praktis, mudah dipakainya	1	0	0	0
2	Memudahkan untuk mengakses pemesanan pembatalan dan perubahan jadwal hanya dari Mobile	0	1	0	0
3	habis update ga bisa booking tiket, error ketika klik kereta yang dipilih,perbaiki dong, update lagi tetep gabisa mohon tanggapan	0	0	-1	0
4	sangat membantu untuk membeli tiket kereta api, dan sangat praktis	0	0	0	1

### Data Preparation

Tahap ini merupakan proses persiapan data yang tujuannya adalah untuk mendapatkan data bersih yang siap digunakan untuk penelitian. Beberapa tahap *pre-processing* yang dilakukan adalah *Transform Case*, merupakan proses mengubah semua huruf pada teks menjadi huruf kecil semua atau sebaliknya. *Tokenizing*, dilakukan pemecahan kalimat menjadi kata. *Stopwords*, pada tahap ini dilakukan penghapusan kata-kata yang dianggap tidak sesuai atau sering muncul seperti: ‘dari’, ‘dan’, ‘ini’, dan lainnya. *Stemming* adalah proses mengubah kata yang berimbuhan menjadi kata dasar.

### Modeling

Pada tahap *modeling* ini, model dibuat dengan menerapkan klasifikasi pada dataset ulasan yang sudah diproses pada tahap *pre-processing*. Algoritma klasifikasi yang digunakan adalah *Support Vector Machine*, dan untuk mengatasi *imbalance class* digunakan SMOTE. Pada penelitian ini digunakan *tools* RapidMiner versi 9.10.

### Evaluation

Pada tahap evaluasi, pengujian dilakukan dengan menggunakan *Cross Validation* (*number of folds* = 10). Dalam tahap validasi untuk pelatihan data terdapat dua subproses yaitu subproses *training* dan subproses *testing*. Setelah pelatihan data, maka dihasilkan evaluasi performa model *Support Vector Machine* oleh *tools* RapidMiner. Rumus yang digunakan yaitu *accuracy* (1), *precision* (2), *recall* (3) dan *F-measure* (4) sebagai berikut:

$$accuracy = \frac{TP+TN}{(TP+FP+FN+TN)} \quad (1)$$

$$precision = \frac{TP}{TP+FP} \quad (2)$$

$$recall = \frac{TP}{TP+FN} \quad (3)$$

$$F - measure = \frac{2x\ precision \times recall}{precision+recall} \quad (4)$$

Ukuran evaluasi model klasifikasi, yaitu *True Positif* (TP) adalah jumlah data dengan nilai positif yang diprediksi dengan benar sebagai positif, *True Negatif* (TN) adalah jumlah data yang memiliki nilai negatif dan diprediksi dengan benar menjadi negatif, *False Positif* (FP) merupakan jumlah data negatif tetapi diprediksi positif, *False Negatif* (FN) merupakan kumpulan data yang memiliki nilai positif tetapi diprediksi negatif [13].

### 3. HASIL DAN PEMBAHASAN

Bab ini menjelaskan hasil dan pembahasan performa algoritma *Support Vector Machine* dengan metode CRISP-DM yang digunakan dalam penelitian ini. Pada penelitian ini data yang digunakan diperoleh langsung dari pihak PT Kereta Api Indonesia (Persero). Setelah proses seleksi, 1.261 ulasan diperoleh dan dikelompokkan berdasarkan aspek *Learnability*, *Efficiency*, *Errors*, dan *Satisfaction*. Selain itu, pelabelan dilakukan secara manual menggunakan label positif dan negatif, seperti yang ditunjukkan pada Tabel 2.

Tabel 2. Dataset Berdasarkan Aspek

Sentimen	Learnability	Efficiency	Errors	Satisfaction
Positif	22	25	7	310
Negatif	182	159	326	230
Total	204	184	333	540

#### Pre-processing

Sebelum digunakan untuk data klasifikasi, terlebih dahulu diproses pada tahap *pre-processing*. Penelitian ini menggunakan beberapa tahapan *pre-processing*, yaitu

*Transform Case, Tokenizing, Stemming, Stopwords.* Tahap *transform case*, dilakukan perubahan yaitu mengubah huruf kapital menjadi bentuk *lower case* atau huruf kecil. Hasil proses *transform case* ditunjukkan pada Tabel 3.

**Tabel 3.** Tahap Transform Case

Sebelum	Sesudah
Tolong diperbaiki	tolong diperbaiki
kualitasnya saya sudah	kualitasnya saya sudah
update tapi lemot banget	update tapi lemot banget
banget gabisa pesan tiket	banget gabisa pesan tiket

Tahap selanjutnya adalah *Tokenizing*. Pada tahap ini dilakukan proses memecah data ulasan yang masih dalam bentuk kalimat, menjadi potongan karakter kata-kata yang membentuknya. Hasil dari proses *tokenizing* dapat dilihat pada Tabel 4.

**Tabel 4.** Tahap Tokenizing

Sebelum	Sesudah
tolong diperbaiki	'tolong' 'diperbaiki'
kualitasnya saya sudah	'kualitasnya' 'saya' 'sudah'
update tapi lemot banget	'update' 'tapi' 'lemot' 'banget'
banget gabisa pesan tiket	'banget' 'gabisa' 'pesan' 'tiket'

Tahap *Stemming* berfungsi untuk mengembalikan kata menjadi bentuk kata dasar dari sebuah kata [14]. Hasil dari proses *stemming* dapat dilihat pada Tabel 5.

**Tabel 5.** Tahap Stemming

Sebelum	Sesudah
'tolong' 'diperbaiki'	'tolong' 'perbaiki'
'kualitasnya' 'saya' 'sudah'	'kualitas' 'saya' 'sudah'
'update' 'tapi' 'lemot'	'update' 'tapi' 'lemot'
'lemot' 'banget' 'banget'	'banget' 'banget'
'gabisa' 'pesan' 'tiket'	'gabisa' 'pesan' 'tiket'

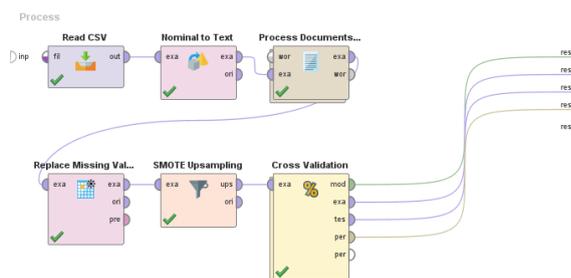
Pada proses *Stopwords*, kata-kata yang tidak relevan akan dihapus sehingga dihasilkan sekelompok teks yang memiliki arti dan berkaitan dengan klasifikasi sentimen [15]. Hasil dari proses *stopwords* dapat dilihat pada Tabel 6.

**Tabel 6.** Tahap Stopwords

Sebelum	Sesudah
'tolong' 'perbaiki'	'tolong' 'perbaiki'
'kualitas' 'saya' 'sudah'	'kualitas' 'update'
'update' 'tapi' 'lemot'	'lemot' 'gabisa' 'pesan'
'banget' 'banget'	'tiket'
'gabisa' 'pesan' 'tiket'	

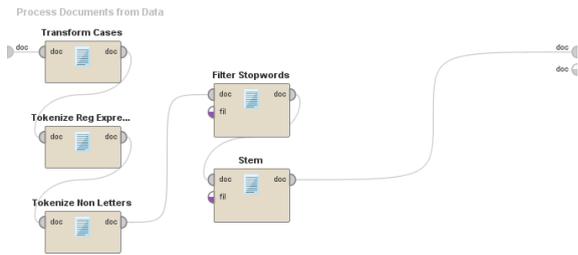
### Pembuatan Model Klasifikasi

Tahap pembuatan model klasifikasi adalah tahap dimana klasifikasi untuk dataset ulasan yang telah melalui tahap *pre-processing* digunakan untuk membuat model. Algoritma klasifikasi yang digunakan dalam penelitian ini adalah *Support Vector Machine*. *Tools* yang digunakan dalam penelitian ini adalah RapidMiner versi 9.10. Tahapan proses ditunjukkan pada Gambar 3.



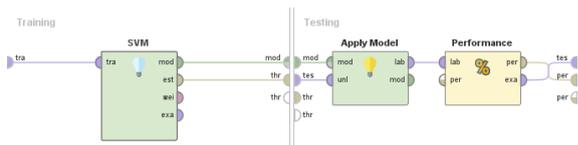
**Gambar 3.** Proses Pembuatan Model Klasifikasi

Operator *Read CSV* digunakan untuk membaca data ulasan yang sudah dilabeli secara manual dalam bentuk *Microsoft Excel* (.CSV). Proses dilanjutkan dengan operator *Nominal to Text*, yang berguna untuk mengubah atribut dari nilai nominal menjadi string. Kemudian data diproses dengan operator *Process Documents* atau yang dikenal dengan tahapan data *preparation* (*pre-processing*). Beberapa langkahnya ditunjukkan pada Gambar 4.



**Gambar 4.** Proses dalam Operator Process Document

Data diproses lebih lanjut setelah melewati operator *Process Document* kemudian operator *Replace Missing Values*, yang berguna untuk mengganti *missing value* pada atribut. Untuk menyeimbangkan dataset agar seimbang, digunakan operator SMOTE up sampling. Selanjutnya untuk evaluasi digunakan *10 fold-cross validation* yang akan mengulang pengujian sebanyak 10 kali dan hasil pengukuran adalah nilai rata-rata dari 10 kali pengujian. *10 folds Cross Validation* adalah pilihan terbaik untuk mendapatkan hasil validasi yang akurat melalui percobaan dan pembuktian teoritis [16].



**Gambar 5.** Desain Evaluasi Algoritma SVM

**Evaluasi Model Klasifikasi**

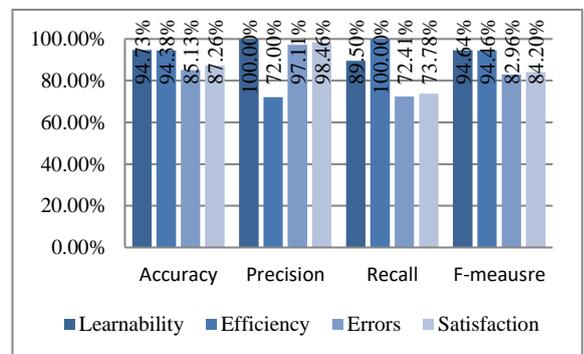
Setelah melalui proses validasi menggunakan algoritma *Support Vector Machine*, proses evaluasi menghasilkan *accuracy*, *recall*, *precision*, dan *F-measure* dari model yang digunakan. Proses pengujian hasil klasifikasi ditunjukkan pada Tabel 7 yang menunjukkan hasil pengujian klasifikasi sentimen untuk masing-masing aspek. Nilai *accuracy* paling tinggi terdapat pada aspek *Learnability* yaitu 94.73%, sedangkan nilai *accuracy* paling rendah

terdapat pada aspek *Errors* yaitu 85.13%. Nilai *precision* tertinggi pada pada aspek *Learnability* yaitu 100.00% dan nilai *precision* terendah terdapat dalam aspek *Efficiency* 94.38% yaitu 72.00%. Nilai *recall* tertinggi terdapat pada aspek *Efficiency* yaitu 100.00% dan nilai *recall* terendah pada aspek *Errors* yaitu 72.41%. Nilai *F-measure* tertinggi pada aspek *Efficiency* yaitu 94.64% dan nilai *F-measure* terendah terdapat dalam aspek *Errors* yaitu 82.96%. Hasil pengujian penggunaan *Support vector machine* (SVM) pada klasifikasi ini menunjukkan seberapa baik model dan performa yang dihasilkan.

**Tabel 7.** Hasil Evaluasi Sentimen Berbasis Aspek

Aspek	Accuracy	Precision	Recall	F-measure
Learnability	94.73%	100.00%	89.50%	94.64%
Efficiency	94.38%	72.00%	100.00%	94.46%
Errors	85.13%	97.11%	72.41%	82.96%
Satisfaction	87.26%	98.46%	73.78%	84.20%

Berdasarkan hasil pengujian perbandingan rata-rata tiap aspek, *Learnability* memiliki rata-rata tertinggi diantara aspek lainnya, kecuali untuk hasil *recall*, *Efficiency* memiliki rata-rata tertinggi. Grafik perbandingan rata-rata model klasifikasi dapat dilihat pada Gambar 6.



**Gambar 6.** Perbandingan rata-rata

#### 4. KESIMPULAN

Berdasarkan hasil yang diperoleh setelah melakukan penelitian ini, maka dapat diambil kesimpulan penerapan model dengan menggunakan algoritma *Support Vector Machine* (SVM) terhadap dataset yang sudah didapatkan sebanyak 1.261 ulasan dari Oktober 2021 hingga April 2022. Penggunaan *Support Vector Machine* (SVM) dapat dijadikan opsi dalam melakukan analisis sentimen level aspek untuk penilaian kepuasan pengguna aplikasi KAI Access. Terbukti berdasarkan pengujian dan nilai dari akurasi yang dihasilkan dari aspek *Learnability* sebesar 94,73%, aspek *Efficiency* sebesar 94,38%, aspek *Errors* sebesar 85,13%, dan aspek *Satisfaction* sebesar 87,26%. Secara umum sentimen positif pada aplikasi KAI Access berkaitan dengan bisa membeli tiket tanpa perlu pergi ke stasiun. Sentimen negatif terhadap aplikasi KAI Access terkait adanya error pada aplikasi setelah update dan kesulitan mendapatkan tiket kereta saat Lebaran.

#### DAFTAR PUSTAKA

- [1] I. F. Rozi, S. Hadi, and E. Achmad, "Implementasi Opinion Mining ( Analisis Sentimen ) untuk Ekstraksi Data Opini Publik pada Perguruan Tinggi," vol. 6, no. 1, pp. 37–43, 2012.
- [2] D. S. Utami and A. Erfina, "Analisis Sentimen Pinjaman Online di Twitter Menggunakan Algoritma Support Vector Machine (SVM)," *SISMATIK (Seminar Nas. Sist. Inf. dan Manaj. Inform.,* vol. 1, no. 1, pp. 299–305, 2021.
- [3] A. M. Pravina, I. Cholissodin, and P. P. Adikara, "Analisis Sentimen Tentang Opini Maskapai Penerbangan pada Dokumen Twitter Menggunakan Algoritme Support Vector Machine (SVM)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.,* vol. 3, no. 3, pp. 2789–2797, 2019, [Online]. Available: <http://j-ptiik.ub.ac.id>.
- [4] R. Wahyudi and G. Kusumawardana, "Analisis Sentimen pada Aplikasi Grab di Google Play Store Menggunakan Support Vector Machine," *J. Inform.,* vol. 8, no. 2, pp. 200–207, 2021, doi: 10.31294/ji.v8i2.9681.
- [5] W. Parasati, F. Abdurrachman Bachtar, and N. Y. Setiawan, "Analisis Sentimen Berbasis Aspek pada Ulasan Pelanggan Restoran Bakso President Malang dengan Metode Naïve Bayes Classifier," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.,* vol. 4, no. 4, pp. 1090–1099, 2020, [Online]. Available: <http://j-ptiik.ub.ac.id>.
- [6] F. Fitriana, E. Utami, and H. Al Fatta, "Analisis Sentimen Opini Terhadap Vaksin Covid - 19 pada Media Sosial Twitter Menggunakan Support Vector Machine dan Naive Bayes," *J. Komtika (Komputasi dan Inform.,* vol. 5, no. 1, pp. 19–25, 2021, doi: 10.31603/komtika.v5i1.5185.
- [7] R. Tineges, A. Triayudi, and I. D. Sholihati, "Analisis Sentimen Terhadap Layanan Indihome Berdasarkan Twitter Dengan Metode Klasifikasi Support Vector Machine (SVM)," *J. Media Inform. Budidarma,* vol. 4, no. 3, p. 650, 2020, doi: 10.30865/mib.v4i3.2181.
- [8] O. Heranova, "Synthetic Minority Oversampling Technique pada Averaged One Dependence Estimators untuk Klasifikasi Credit Scoring," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi),* vol. 3, no. 3, pp. 443–450, 2019, doi: 10.29207/resti.v3i3.1275.
- [9] J. W. Iskandar and Y. Nataliani, "Perbandingan Naïve Bayes, SVM, dan k-NN untuk Analisis Sentimen Gadget Berbasis Aspek," *J. RESTI (Rekayasa Sist. dan Teknol.*

- Informasi*), vol. 5, no. 6, pp. 1120–1126, 2021, doi: 10.29207/resti.v5i6.3588.
- [10] W. Handiwidjojo and L. Ernawati, “Pengukuran Tingkat Ketergunaan ( Usability ) Sistem Informasi Keuangan,” *Juisi*, vol. 02, no. 01, pp. 49–55, 2016.
- [11] A. D. Adhi Putra, “Analisis Sentimen pada Ulasan pengguna Aplikasi Bibit Dan Bareksa dengan Algoritma KNN,” *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 8, no. 2, pp. 636–646, 2021, doi: 10.35957/jatisi.v8i2.962.
- [12] J. M. D. Syahid Achyar, “Perbandingan algoritma untuk klasifikasi analisis sentimen terhadap Genose pada media sosial Twitter,” *semanTIK*, vol. 7, no. 1, pp. 9–16, 2021, doi: 10.5281/zenodo.5034916.
- [13] R. Umar, I. Riadi, and Purwono, “Perbandingan Metode SVM, RF dan SGD untuk Penentuan Model Klasifikasi Kinerja Programmer pada Aktivitas Media Sosial,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 4, no. 2, pp. 329–335, 2020.
- [14] L. Asri *et al.*, “Analisis Sentimen Opini Publik Berita Kebakaran Hutan Melalui Komparasi Algoritma Support Vector Machine Dan K-Nearest Neighbor Berbasis Particle Swarm Optimization,” *J. Pilar Nusa Mandiri*, vol. 13, no. 1, pp. 103–112, 2017, [Online]. Available: [www.tribunnews.com](http://www.tribunnews.com).
- [15] M. A. K. Neighbors, A. E. Pratama, A. Ariesta, and G. Gata, “Analisis Sentimen Masyarakat terhadap Tim Nasional Indonesia pada Piala AFF 2020,” vol. 10, pp. 187–196, 2022.
- [16] H. S. Utama, D. Rosiyadi, B. S. Prakoso, and D. Ariadarma, “Analisis Sentimen Sistem Ganjil Genap di Tol Bekasi Menggunakan Algoritma Support Vector Machine,” *J. RESTI (Rekayasa Sist. dan Teknol.*
- Informasi*), vol. 3, no. 2, pp. 243–250, 2019, doi: 10.29207/resti.v3i2.1050.