

IMPLEMENTASI MACHINE LEARNING SEBAGAI ANALISIS KEPUASAN PELANGGAN TERHADAP PENGGUNAAN APLIKASI KAI ACCESS

Febrina Tesalonika Nugraha¹⁾, Hendry²⁾,

1. Universitas Kristen Satya Wacana, Indonesia
2. Universitas Kristen Satya Wacana, Indonesia

Article Info

Kata Kunci: *KAI Access; Machine Learning, Orange*

Keywords: *KAI Access; Machine Learning, Orange*

Article history:

Received 15 May 2023

Revised 29 May 2023

Accepted 12 June 2023

Available online 1 December 2023

DOI :

<https://doi.org/10.29100/jipi.v8i4.4185>

* Corresponding author.

Febrina Tesalonika Nugraha

E-mail address:

672019309@student.uksw.edu

ABSTRAK

Revolusi Industri 4.0 adalah sebuah *disruption* era sebagaimana hal ini ditandai dari cara kerja yang berpindah atau berubah baik salah satunya adalah pada bidang transportasi. Kereta api adalah transportasi darat yang berjalan dengan rangkaian kendaraan lainnya dan bergerak diatas rel. Dengan mengikuti perkembangan zaman seperti yang sudah di kemukakan maka KAI meningkatkan pelayanan dengan berinovasi untuk memmbuat aplikasi *KAI Access*. Dari jumlah data awal sebanyak 1561 lalu dilakukan cleansing data hingga didapatkan 1117 data. Hasil dari penelitian ini adalah pengguna *KAI Access* dominan memiliki sentiment positif yang dinilai mayoritas memberikan rating puas terhadap penggunaan aplikasi *KAI Access*. Dalam hal performa metode Support Vector Machine memiliki tingkat akurasi tertinggi dibandingkan dua pembanding lainnya yaitu K-NN dan *Decision Tree*. Penelitian ini memberikan wawasan tentang pengaplikasian *machine learning* sebagai model prediksi kepuasan pelanggan terhadap penggunaan *KAI Access*.

ABSTRACT

The Industrial Revolution 4.0 is a disruption era as this is marked by the way of work that moves or changes, one of which is in the field of transportation. Train is land transportation that runs with a series of other vehicles and moves on rails. By following the development of the times as stated above, KAI is improving its services by innovating to create the *KAI Access* application. From the initial data of 1561, data cleansing was carried out to obtain 1117 data. The results of this study are that dominant *KAI Access* users have positive sentiments, which are considered by the majority to give satisfaction ratings to the use of the *KAI Access* application. In terms of performance, the Support Vector Machine method has the highest level of accuracy compared to the other two comparisons, namely K-NN and *Decision Tree*. This research provides insight into the application of machine learning as a predictive model of customer satisfaction with the use of *KAI Access*.

I. PENDAHULUAN

REVOLUSI Industri 4.0 adalah sebuah disruption era sebagaimana hal ini ditandai dari cara kerja yang berpindah atau berubah. Perubahan ini adalah perubahan cara kerja dari konvensional menjadi modern yang dilakukan melalui pendekatan-pendekatan digital. Era Revolusi Industri 4.0 ini memberikan dampak yang luas bagi dunia, tidak terkecuali Indonesia. Indonesia yang merupakan negara berkembang mendapatkan dampak dari Revolusi Industri 4.0 dengan dilihat dari perkembangan baik dalam ekonomi, sosial, hingga transportasi. Pertumbuhan atau perkembangan dalam era digital semacam ini dibarengi dengan kemunculan perusahaan-perusahaan teknologi yang menunjang peningkatan perubahan secara massif. Peningkatan ini berorientasi pada inovasi serta efisiensi yang semakin mempermudah masyarakat sebagai konsumen. Berkaitan dengan perkembangan perusahaan-perusahaan yang berbasis digital ini memberikan dampak maraknya penggunaan berbagai macam aplikasi daring. Aplikasi daring ini secara massif berkembang dalam bidang sektor jasa transportasi jika dilihat dari perpektif penyedia layanan. Salah satu transportasi yang juga berkembang pesat saat ini adalah kereta api. Kereta api adalah transportasi darat yang berjalan dengan rangkaian kendaraan lainnya dan bergerak diatas rel. Kereta api bergerak dengan lokomotif yang dikendalikan oleh masinis, PT. KAI adalah Badan Usaha Milik Negara yang bergerak pada bidang transportasi rangkaian kereta api. Karena sifatnya sebagai angkutan massal yang efektif dan memiliki jadwal keberangkatan dan kedatangan yang beragam maka beberapa negara menjadikan kereta api

sebagai transportasi utama lokal, antarkota maupun antarnegara. Pelayanan yang berkualitas menjadi bagian yang perlu diperhatikan oleh perusahaan penyedia transportasi dari solusi masalah transportasi perkotaan yang semakin kompleks, waktu menjadi salah satu pelayanan prioritas yang harus diperhatikan untuk memberikan kepercayaan kepada konsumen. Dengan mengikuti perkembangan zaman seperti yang sudah di kemukakan maka KAI meningkatkan pelayanan dengan berinovasi untuk membuat aplikasi KAI *Access* yang tersedia pada *IOS*, *Android* dan *Windows*. KAI *Access* memudahkan pengguna untuk memesan tiket serta mengubah jadwal maupun membatalkan jadwal keberangkatan. KAI terus berusaha untuk meningkatkan pelayanan agar para pengguna merasa puas dengan pelayanan yang ada. Salah satu cara yang dilakukan oleh PT KAI dalam mengetahui kepuasan pelanggan adalah dengan memberikan media penilaian kepada para pelanggan. Penelitian ini fokus terhadap penilaian pengguna KAI *Access* yang dianggap penting untuk strategi bisnis PT KAI sehingga dapat terus memenuhi kepercayaan para pelanggan terhadap pelayanan PT KAI. Adapun yang menjadi tujuan dilakukannya penelitian ini adalah untuk mengetahui tingkat kepuasan pengguna KAI *Access* dengan menggunakan metode *data mining* digunakan untuk menemukan pengetahuan tersembunyi pada *database*. [1] *Data mining* menggambarkan pengumpulan teknik dengan tujuan untuk menemukan pola yang tidak diketahui pada data untuk mencapai sebuah hasil, sehingga *data mining* dapat diartikan sebuah proses pengumpulan data yang menghasilkan keluaran berupa pengetahuan sebuah hasil dari kumpulan data dengan tujuan tertentu. Penelitian ini akan membahas pengaplikasian *machine learning* sebagai model prediksi kepuasan pelanggan terhadap penggunaan KAI *Access* yang lebih akurat karena dapat menghitung dengan mempertimbangkan tingkat akurasi dengan menggunakan tiga algoritma. *Machine learning* didefinisi sebagai aplikasi sebuah komputer dan algoritma matematika yang diadopsi dengan cara pembelajaran dari data dan menghasilkan prediksi. [2] Algoritma pada *machine learning* menggunakan teknik statistik untuk menemukan pola, seringkali data yang digunakan berukuran besar. *Machine learning* memiliki cara kerja mekanisme seperti manusia dalam belajar dan mengambil keputusan yang tepat, dalam perkembangan algoritma *machine learning* telah berkembang berdasarkan output yaitu *supervised learning*. *Supervised learning* merupakan algoritma yang membangkitkan sebuah fungsi yang memetakan *input* ke *output* yang diinginkan. [3]

Penelitian ini dilakukan berdasarkan dari penelitian sebelumnya dengan tujuan untuk menganalisis jumlah pembaca yang maksimal. Hasil yang didapatkan dari penelitian ini adalah Berdasarkan tingkat akurasi dengan menggunakan algoritma support vector machine memiliki kinerja yang lebih baik dibandingkan dengan algoritma *naïve bayes* sehingga dinilai mampu membantu peneliti dalam mendapatkan tingkat akurasi yang lebih tinggi dibandingkan dengan *naïve bayes*. Sampel yang digunakan terlalu kecil sehingga dapat menyebabkan penelitian tidak dapat menggambarkan kondisi populasi sesungguhnya. Berdasarkan penjelasan dari hasil penelitian tersebut menjadi pertimbangan bagi peneliti untuk menggunakan metode Support Vector Machine menjadi perbandingan metode pada penelitian ini.[4] Penelitian sebelumnya kedua dilakukan dengan tujuan untuk membantu penyelesaian terkait permasalahan *Drop Out* siswa, dengan memberikan wawasan baru dengan memanfaatkan teknologi berupa *machine learning*. Teknik yang digunakan dalam penelitian ini yaitu pemodelan *support vector machine*, dari hasil yang didapatkan berdasarkan pemodelan yang telah digunakan yaitu Support Vector Machine memiliki tingkat akurasi prediksi yang memuaskan. Penelitian ini masih menggunakan perhitungan secara manual dengan data yang cukup banyak. Oleh karena itu, peneliti menggunakan tools Orange Data Mining untuk melakukan preprocessing sebagai pendekatan pada penelitian dengan data yang cukup banyak. [5] Penelitian sebelumnya ketiga dilakukan dengan tujuan untuk memprediksi kemungkinan pengunduran diri calon mahasiswa baru. Penelitian ini dilakukan dengan menggunakan metode *data mining* dengan algoritma Decision Tree. Penelitian ini dilakukan sebanyak tiga kali penelitian pertama dilakukan dengan perhitungan manual menghasilkan akurasi sebanyak dan penelitian terakhir dilakukan pemisahan *data training* dan *testing* dengan menghasilkan akurasi tertinggi sebesar 84.50% namun atribut yang digunakan pada penelitian ini terbatas. Dari penjabaran penelitian tersebut menjadi pertimbangan bagi penulis dalam melakukan penelitian dengan menggunakan perbandingan metode Decision Tree dengan metode lainnya. [6] Penelitian sebelumnya yang ketujuh dilakukan oleh Yahya dan Winda tujuan dari penelitian ini adalah untuk mengklasifikasi hasil penjualan dari Lombok Vape on untuk mengetahui efektivitas penjualan perbulannya agar toko tersebut dapat mengetahui pasang surut pemasukan yang didapat. Pengklasifikasi tidak menggunakan model apapun untuk dicocokkan dan hanya berdasarkan pada memori. Hasil dari penelitian ini adalah pengolahan dilakukan dengan menggunakan 9 *validation* dari pengolahan terdapat *K-Fold Validation* 6 dengan hasil akurasi sebesar 86.48% dan AUC sebesar 0.874. Dari penelitian tersebut penulis memutuskan untuk menggunakan metode K-NN sebagai salah satu metode perbandingan pada penelitian ini. [7]

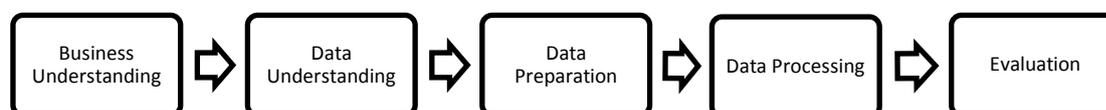
Berdasarkan pada latar belakang yang telah dijabarkan, penelitian ini dilakukan untuk mencari tahu tingkat kepuasan pelanggan terkait dari sentiment yang telah diberikan oleh pengguna aplikasi KAI *Access* dengan menggunakan perbandingan algoritma Support Vector Machine, Decision Tree dan K-Nearest Neighbour. Penelitian ini diharapkan mampu menilai bagaimana sentiment pengguna aplikasi KAI *Access*, serta memberikan

wawasan pada PT KAI dalam pengembangan aplikasi KAI Access. Pada penelitian ini data diambil dari data yang dimiliki oleh PT KAI terkait keseluruhan penilaian pengguna KAI Access yang berisi komentar dan rating pengguna sebanyak 1156 data dilakukan pembersihan data hingga terdapat 1117 data. Dalam penelitian ini menggunakan orange data mining untuk preprocess pada penggunaan data yang banyak

II. METODE PENELITIAN

Pada penelitian ini akan membandingkan tingkat akurasi paling tinggi dari metode yang digunakan yaitu Support Vector Machine, Decision Tree dan K-NN dengan tujuan dari penelitian ini untuk menilai kepuasan penggunaan aplikasi KAI Access dari metode dengan tingkat akurasi yang paling tinggi, sentiment yang muncul pada pengguna KAI Access positif, netral atau negatif. Pada penelitian ini penulis mengambil data dari PT KAI terkait komentar dan rating pengguna aplikasi KAI Access dari App Store, Play Store dan Windows Store.

Support vector machine merupakan suatu teknik untuk melakukan prediksi baik dalam regresi maupun klasifikasi digunakan untuk mendapatkan fungsi pemisah yang optimal untuk memisahkan observasi yang memiliki nilai variable target yang berbeda. [8] Karakteristik *support vector machine* secara umum secara prinsip merupakan linear *clasifier*. K-NN adalah Teknik yang sederhana dan efisien dalam bidang pengenalan pola. K-NN mengklasifikasi hasil *query instance* berdasarkan mayoritas jarak dari kategori yang ada pada data. Metode K-NN salah satu metode klasifikasi terhadap sekumpulan data yang berdasarkan mayoritas dari kategori dan tujuan untuk mengklasifikasi obyek baru berdasar atribut dan sample sample dari training data. [9] Decision Tree adalah sebuah struktur data yang terdiri dari node dan rusuk. Setiap node menggambarkan node pembagi dimana tiap node merupakan satu masukan dan memiliki paling sedikit dua keluaran node akar menyatakan node paling atas dari pohon. Pohon keputusan tiap leaf node menyatakan keadaan harus diisi pada tiap puncak pohon dimana menggambarkan nilai kelas data. Pembuatan keputusan berarti memilih alternatif keputusan yang tersedia. Decision Tree atau pohon keputusan memperlihatkan probabilitas yang akan mempengaruhi alternatif pada keputusan tersebut disertai dengan estimasi hasil akhir yang di dapat bila mengambil alternatif keputusan. Decision Tree adalah model visual untuk menyederhanakan proses pembuatan keputusan secara rasional. [10] Decision Tree memiliki kemampuan dalam memperhitungkan berbagai alternatif pemecahan masalah termasuk faktor-faktor yang mempengaruhi serta nilai resiko dan nilai informasi dalam alternatif keputusan. [11] Metode penelitian diperlukan pada setiap penelitian agar memudahkan peneliti dalam menyelesaikan penelitian, berikut tahap metodologi penelitian dan penjelasan terperinci secara sistematis sebagai berikut:



Gambar. 1. Metode penelitian

A. Bussiness Understanding

Peneliti telah melakukan diskusi dengan PT. KAI terkait pemahaman permasalahan yang sering terjadi yaitu mengenai data penilaian pada playstore maupun *Appstore* terhadap penggunaan aplikasi KAI *Access*, dengan adanya penelitian ini maka akan menambah wawasan pada PT. KAI terkait analisis kepuasan pengguna terhadap penggunaan aplikasi KAI *Access* dengan menggunakan *machine learning*.

B. Data Understanding

Pada tahap ini dilakukan pemahaman data yang didapatkan terkait dengan penilaian pengguna KAI *Access* yang didapatkan dari data yang dimiliki oleh PT. KAI terkait penilaian melalui rating dan komentar yang diberikan oleh pengguna KAI *Access* dan dilakukan analisis terkait *variable* pada data yaitu *sentiment* sebagai target dibagi menjadi *sentiment negative*, *sentiment positif* dan *sentiment netral*. *Review text* sebagai meta yang berisikan komentar dari pengguna KAI *Access* yang berisi penilaian dalam bentuk kata-kata saran maupun kritik. Pemahaman data juga dilakukan guna memahami akan pengambilan data tertentu pada dataset yang hanya dapat digunakan, karena tidak semua data dapat digunakan atau sesuai pada analisis.

C. Data Preparation

Data preparation (persiapan data) pada proses ini data yang akan dilanjutkan pada proses selanjutnya disiapkan pada tahap ini. data mentah pada tahap ini telah dikumpulkan untuk diubah menjadi data yang siap digunakan. pembersihan data adalah proses menghilangkan data yang tidak konsisten atau tidak relevan, data yang diperoleh dari *database* suatu perusahaan maupun hasil eksperimen, memiliki data yang tidak sempurna seperti data yang

hilang, data yang tidak *valid* atau juga hanya sekedar salah dalam penginputan. Pada tahap ini dilakukan pembersihan pada data penilaian pengguna KAI *Access* khususnya pada bagian *review text* sebagai meta karena terdapat beberapa rating tanpa disertai dengan saran atau kritik sehingga dilakukan *cleansing* agar hasil yang didapatkan nantinya dapat maksimal. [12]

D. Data Processing

Data processing atau pengolahan data adalah proses yang dilakukan setelah dilakukan pemahaman data hingga pembersihan data sehingga data siap untuk diolah sesuai kebutuhan dengan output yang mudah dipahami pembaca. Pada tahap ini dilakukan pengolahan sebanyak 1117 data menggunakan tools orange machine learning dan menggunakan metode *Support Vector Machine*, *K-Nearest Neighbor* dan *Decision Tree*. Hasil dari pengolahan data yang telah dilakukan pada tools orange dapat dilihat pada *widget Test and score* yang akan menampilkan nilai tertinggi dari ketiga metode yang digunakan.

E. Evaluation

Evaluasi adalah pertimbangan dan pengamatan ulang yang dilakukan oleh peneliti sebelum mencapai hasil akhir. Pada tahap ini dilakukan evaluasi terlebih dahulu guna meminimalisir kesalahan yang ada evaluasi dilakukan berdasarkan hasil tingkat akurasi yang dihasilkan berdasarkan dari data kepuasan pelanggan terhadap penggunaan KAI *Access*. Peneliti melakukan pengamatan ulang dari awal proses *cleansing data* dengan meneliti satu persatu data agar tidak ada kesalahan dalam penginputan, lalu melakukan trial and error kembali hingga didapatkan hasil akhir yang sama sehingga akan dijadikan *output* yang *valid*.

III. HASIL DAN PEMBAHASAN

Pada bagian ini adalah hasil dari analisis yang telah dilakukan oleh penulis untuk perbandingan dari metode *Support Vector Machine*, *Decision Tree*, dan *K-NN* dalam pengklasifikasian sentiment pada komentar pengguna KAI *Access*. Hasil dan pembahasan pada penelitian ini akan dilakukan sebagai berikut:

1. Business Understanding

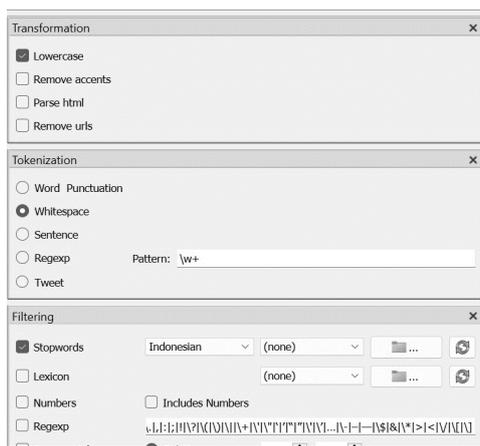
Pada tahap *Business Understanding* proses awal yaitu pengambolan data yang digunakan pada penelitian ini. data yang digunakan pada penelitian ini adalah data yang diperoleh dari PT Kereta Api Indonesia terkait penilaian pengguna aplikasi KAI *Access* pada *App Store*, *Play Store* dan juga *Windows Store* sebanyak 1117 data. Diharapkan dapat menambah wawasan pada PT. KAI terkait analisis kepuasan pengguna terhadap penggunaan aplikasi KAI *Access* dengan menggunakan *machine learning*.

2. Data Understanding

Pada tahanan Pemahaman Data (Pemahaman Data) dilakukan pengolahan dataset yang digunakan pada tahap selanjutnya. Pemahaman data yang dilakukan terkait dengan penilaian pengguna KAI *Access* yang didapatkan dari data yang dimiliki oleh PT. KAI terkait penilaian melalui rating dan komentar yang diberikan oleh pengguna KAI *Access* dan dilakukan analisis terkait *variable* pada data yaitu sentiment sebagai target dibagi menjadi *sentiment negative*, *sentiment positif* dan *sentiment netral*. *Review text* sebagai meta yang berisikan komentar dari pengguna KAI *Access* yang berisi penilaian dalam bentuk kata-kata saran maupun kritik. Pemahaman data juga dilakukan guna memahami akan pengambilan data tertentu pada dataset yang hanya dapat digunakan, karena tidak semua data dapat digunakan atau sesuai pada analisis.

3. Data Preparation

Langkah selanjutnya adalah *data preparation* (persiapan data) pada proses ini data yang akan dilanjutkan pada proses selanjutnya disiapkan pada tahap ini. data mentah pada tahap ini telah dikumpulkan untuk diubah menjadi data yang siap digunakan. Adapun tahap yang dilakukan pada tahap ini dengan menggunakan *Orange Data Mining* sebagai berikut :



Gambar. 2. Preprocess pada Orange

a. Transformation

Transformasi adalah proses mengubah data menjadi bentuk yang dibutuhkan sesuai process data. Proses transformasi, setiap huruf pada data input akan diubah sesuai dengan tahapannya. Pada tahap transformation dalam penelitian ini lowercase berfungsi untuk mengubah semua teks menjadi huruf kecil.

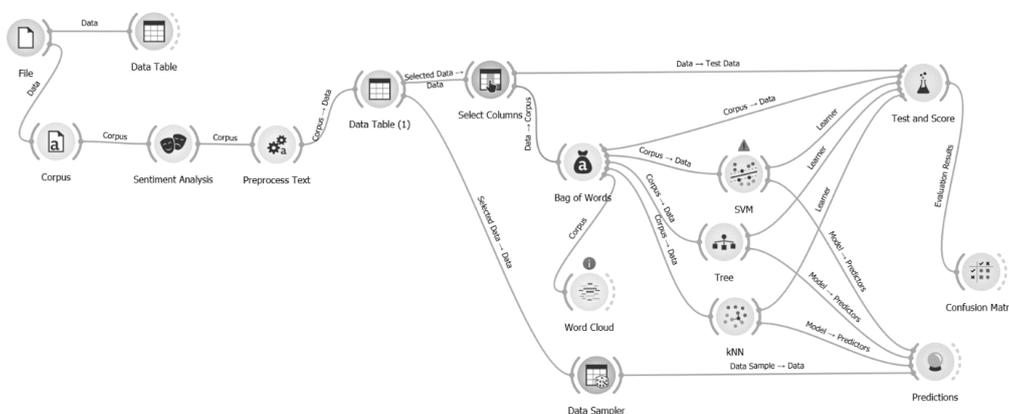
b. Tokenization

Tokenization adalah proses untuk memisahkan deretan kata dalam kalimat, paragraf atau halaman menjadi potongan kata tunggal yang berdiri sendiri. Pemecahan kata dan kalimat berdasarkan pada spasi di dalam kalimat. Tahap ini juga menghilangkan karakter dan dan symbol selain a-z dihilangkan. [13]

c. Filtering

Tahap terakhir adalah filtering atau pembersihan data adalah sebuah proses yang digunakan untuk memperbaiki, mendeteksi ataupun menghapus dataset yang korup atau tidak akurat. Proses filtering penting untuk mencegah data duplikat dan membuat data lebih terstruktur. [14]

4. Data Processing



Gambar. 3. Design pada tools orange untuk menghitung tingkat accuracy model prediksi

Pada Gambar 3 merupakan modelling pada tools orange yang membandingkan model algoritma yang diuji. Proses keberhasilan pada tools orange menggunakan widget Test and Score. Data yang digunakan sebanyak 1117 data yang telah melalui proses cleansing data kemudian dilakukan processing. Pada Langkah ini sesuai dengan algoritma yang dibahas yaitu metode Support Vector Machine, Decision Tree dan K-NN. Data yang telah diimport pada file csv telah melalui tahap cleansing data ditampilkan pada data tabel. Pada data tersebut yang akan diproses sebagai input untuk menguji pada tahap test and score.

Name	Type	Role	Values
8 Sentimen	C categorical	target	negatif, netral, positif
9 Developer Rep...	N numeric	skip	
10 Feature 1	S text	skip	
11 Review Submi...	S text	skip	
12 Review Last ...	S text	skip	
13 Review Text	S text	meta	
14 Developer Rep...	S text	skip	

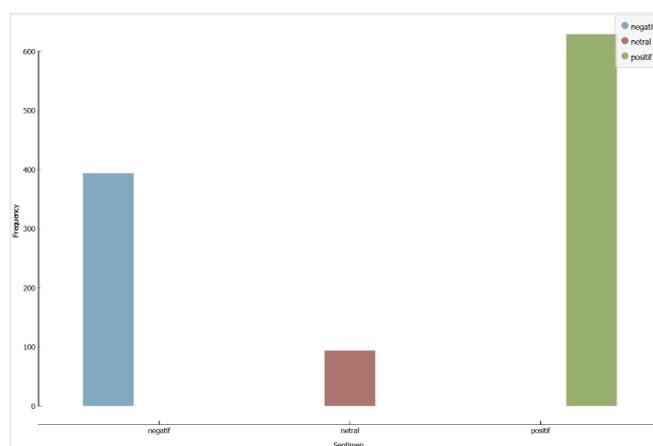
Gambar. 4. Proses label data

Pada gambar 4 tahap awal data yang ada pada *widget* file akan menampilkan seluruh atribut yang ada pada data dan focus pada atribut yang digunakan pada *processing* data. Pada sentiment akan dibagi menjadi 3 values yaitu sentiment negatif, sentiment netral, dan sentiment positif lalu diubah menjadi *type categorical* dan menjadi target pada process penelitian ini. Pada atribut *review text* yang berisikan komentar pengguna aplikasi KAI Access telah dilakukan *cleansing* data dari 1561 data menjadi 1117 data yang sudah siap digunakan.

Model	AUC	CA	F1	Precision	Recall
kNN	0.690	0.606	0.494	0.671	0.606
Tree	0.848	0.782	0.768	0.759	0.782
SVM	0.888	0.809	0.774	0.742	0.809

Gambar. 5. Hasil Test and Score

Gambar 5 merupakan tampilan dari *widget test and score pada tools orange*, pada penelitian ini proses pengujian menerapkan *K-Fold Cross Validation (K=10)* yang dapat di atur pada *widget test and score* pada gambar di atas. Pada *widget* tersebut juga diperlihatkan hasil evaluasi dari ketiga algoritma, dimana hasil perbandingan ketiga algoritma memprediksi tingkat kepuasan pelanggan dengan menunjukkan nilai AUC, CAM F1, *Precision* dan *Recall*. Penelitian kali ini hanya menampilkan nilai AUC dalam melihat kinerja dari model prediksi yang menganalisa. Hasil evaluasi menunjukkan bahwa ketiga model algoritma memiliki nilai akurasi yang berbeda. *K-Nearest Neighbor* memiliki nilai CA 0.690 yang dinilai paling rendah dibandingkan dengan *Decision Tree* dengan nilai 0.848 dan *Support Vector Machine (SVM)* yang dinilai memiliki nilai paling tinggi dibandingkan kedua algoritma lainnya yaitu 0.888. hal tersebut menunjukkan bahwa metode terbaik yang yang dapat digunakan sebagai analisis kepuasan pelanggan terhadap penggunaan KAI Access yaitu *Support Vector Machine (SVM)*.



Gambar. 6. Distribution

Processing data telah selesai dilakukan Langkah selanjutnya adalah menghubungkan hasil dari *widget test and score* ke *widget distribution* pada *Orange Data Mining*. Hasil sentiment digunakan untuk menampilkan sentiment yang telah diolah lalu data akan divisualisasikan berdasarkan dari masing-masing sentiment *positif*, *negative*, dan netral yang ditampilkan melalui *widget distribution* guna memudahkan pembaca dalam memahami hasil

visualisasi. Berdasarkan dari hasil visualisasi yang ada pada Gambar 6 dapat dilihat bahwa respon dari pengguna aplikasi KAI Access dominan pada kategori sentiment positif. Distribusi pada sentimen yang ditunjukkan terlihat perbandingan yang jauh antara sentiment positif dan netral. Dari hasil akhir pada visualisasi tersebut terlihat bahwa pengguna KAI Access dominan memberikan komentar positif dalam penilaian terhadap aplikasi KAI Access. Hal ini cukup menunjukkan bahwa pengguna aplikasi KAI Access cenderung puas terhadap aplikasi KAI Access.

5. Evaluation

Confusion Matrix adalah sebuah metode yang digunakan untuk perhitungan akurasi recall, precision dan error rate. *Confusion matrix* pada ketiga metode pada penelitian ini terdapat negative, netral dan positif dimana menunjukkan presentase dari setiap sentiment. [15]

	negatif	netral	positif	Σ
negatif	73.2 %	NA	10.5 %	394
netral	15.9 %	NA	3.6 %	94
positif	10.9 %	NA	85.9 %	629
Σ	441	0	676	1117

Gambar. 7. Hasil widget confusion matrix SVM

Pada Gambar 7 merupakan tampilan dari *widget confusion matrix* dengan menampilkan *proportion of actual* yaitu prediksi komentar *matrix* ini berfokus pada *proportion of actual positif* presentase data pada algoritma *Support Vector Machine* adalah sebesar 85,9%

	negatif	netral	positif	Σ
negatif	73.6 %	56.2 %	10.6 %	394
netral	12.9 %	25.0 %	4.5 %	94
positif	13.4 %	18.8 %	84.9 %	629
Σ	402	48	667	1117

Gambar. 8. Hasil widget confusion matrix Decision Tree

Pada Gambar 8 merupakan tampilan dari *widget confusion matrix* dengan menampilkan *proportion of actual* yaitu prediksi komentar presentase data pada algoritma *Decision Tree* lebih rendah dari algoritma *Support Vector Machine* adalah sebesar 85,9%

	negatif	netral	positif	Σ
negatif	96.0 %	100.0 %	32.4 %	394
netral	4.0 %	0.0 %	8.6 %	94
positif	0.0 %	0.0 %	59.0 %	629
Σ	50	1	1066	1117

Gambar. 9. Hasil widget confusion matrix K-NN

Pada Gambar 9 merupakan tampilan dari *widget confusion matrix* dengan menampilkan *proportion of actual* yaitu prediksi komentar presentase data pada algoritma *KNN* lebih rendah dari algoritma *Support Vector Machine* dan *Decision Tree* adalah sebesar 59,0%

TABEL 1
PERBANDINGAN HASIL METODE

Metode	Accuracy	Precision	Recall
Support Vector Machine	89%	74%	80%
Decision Tree	85%	76%	78%
K-NN	70%	67%	60%

Pada tabel 1 ditunjukkan bahwa metode Support Vector Machine memiliki nilai akurasi yang lebih tinggi dibandingkan dengan kedua metode lainnya. Hal ini menunjukkan bahwa Support Vector Machine dapat menjadi alternatif yang baik dalam proses klasifikasi namun performa pada setiap metode yang digunakan juga tergantung dari atribut yang dimiliki dari setiap dataset yang digunakan. Penelitian sebelumnya juga dilakukan oleh Indrayanti dan tim (2022) bahwa penerapan SVM dilakukan dengan menggunakan *tools Orange Data Mining* dengan tingkat *error*

terendah [16]. Oleh karena itu, metode Support Vector Machine diidentifikasi memiliki potensi untuk memberikan hasil yang lebih baik pada penelitian yang berbeda.

IV. KESIMPULAN DAN SARAN

Hasil dari pengujian yang telah dilakukan adalah nilai *AUC K-Nearest Neighbor* sebesar 0.690 sedangkan *Decision Tree* sebesar 0.848 dan *Support Vector Machine* sebesar 0.888. Berdasarkan nilai CA tersebut dapat disimpulkan bahwa *Support Vector Machine* merupakan metode terbaik untuk menganalisis kepuasan pelanggan terhadap penggunaan aplikasi KAI Access dibandingkan dengan *K-Nearest Neighbor* dan *Decision Tree*. Penelitian selanjutnya disarankan dapat membandingkan dengan menggunakan metode lain dan dengan menggunakan *tools orange* atau dapat menerapkan *tools* lain. Pada pengujian ini juga didapatkan hasil bahwa pengguna KAI Access dominan memberikan sentiment positif terhadap penggunaan aplikasi KAI Access yang dinilai puas dalam menggunakan aplikasi KAI Access. Adapun saran untuk penelitian selanjutnya, dapat memperbanyak atribut sebagai penunjang dan menggunakan Teknik pengolahan lebih canggih guna menambah wawasan baru untuk meningkatkan kinerja.

DAFTAR PUSTAKA

- [1] A. Roihan, P. A. Sunarya, and A. S. Rafika, "Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper," *IJCIT (Indonesian J. Comput. Inf. Technol.*, vol. 5, no. 1, pp. 75–82, 2020, doi: 10.31294/ijcit.v5i1.7951.
- [2] F. Sodik, B. Dwi, and I. Kharisudin, "Perbandingan Metode Klasifikasi Supervised Learning pada Data Bank Customers Menggunakan Python," vol. 3, pp. 689–694, 2020.
- [3] R. Oethario, S. Kasus, K. Kemangi, R. Oethario, S. Karnila, and R. Oethario, "Data mining," vol. 13, no. 1, pp. 80–89, 2013.
- [4] U. Riyanto, "ANALISIS PERBANDINGAN ALGORITMA NAIVE BAYES DAN SUPPORT VECTOR MACHINE DALAM MENGLASIFIKASIKAN JUMLAH PEMBACA," pp. 62–72, 2018.
- [5] S. Sains, M. Siswa, D. Out, and S. Kasus, "Penerapan Teknik Data Mining dengan Metode Support Vector Machine," vol. 3, no. 1, 2017.
- [6] S. Kom and M. Kom, "PENERAPAN DECISION TREE UNTUK MENGANALISIS MAHASISWA BARU," vol. 7, no. 1, pp. 8–14, 2016.
- [7] W. P. Hidayanti, "Penerapan Algoritma K-Nearest Neighbor Untuk Klasifikasi Efektivitas Penjualan Vape (Rokok Elektrik) pada 'Lombok Vape On' Pendahuluan dihasilkan tidak stabil dan tidak mampu diprediksi Dari penelitian yang dilakukan, berusaha untuk mengklasifikasikan," vol. 3, no. 2, 2020.
- [8] S. D. Di and K. Magelang, "1, 2, 3 1," vol. 3, pp. 811–820, 2014.
- [9] M. Mahasiswa and B. Di, "Implementasi algoritma k-nearest neighbor untuk menentukan mahasiswa berprestasi di stmk kristen neumann indonesia," pp. 56–61.
- [10] I. Alwiah Musdar, H. Angriani, and S. KHARISMA Makassar, "Implementasi Metode Decision Tree Dalam Menentukan Pemberian Kredit Mobil Menggunakan Visual Basic (Studi Kasus UD PUTRA MAS Makassar)," *Jtriste*, vol. 4, no. 1, pp. 55–67, 2017.
- [11] J. Ilmiah, I. Asia, and F. Pakaja, "STUDY DECISION TREE / POHON KEPUTUSAN SEBAGAI SEBUAH ALAT BANTU PENDUKUNG SISTEM DALAM PROSES PENGAMBILAN KEPUTUSAN PENJUALAN PADA CV. KHAN SETIA UTAMA, PONDOK CABE-DEPOK Sekolah Tinggi Manajemen Informatika dan Komputer ASIA Malang Jurnal Ilmiah Teknol," vol. 3, no. 2, pp. 51–69, 2009.
- [12] L. Robinson, "Implementasi Metode Generalized Vector Space Model Pada Aplikasi Information Retrieval untuk Pencarian Informasi Pada Kumpulan Dokumen Teknik Elektro Di UPT BPI LIPI," *J. Ilm. Komput. dan Inform.*, 2019.
- [13] P. Purnamasari, W. Suyitno, R. E. Indrajit, and M. Fauzi, "Penerapan Data Mining Dalam Menangani," vol. 1, no. 2, pp. 53–60, 2017.
- [14] N. Putu, A. Widiari, I. M. Agus, D. Suarjaya, and D. P. Githa, "Teknik Data Cleaning Menggunakan Snowflake untuk Studi Kasus Objek Pariwisata di Bali," vol. 8, no. 2, pp. 137–145, 2020.
- [15] A.- Arini, L. K. Wardhani, and D.- Octaviano, "Perbandingan Seleksi Fitur Term Frequency & Tri-Gram Character Menggunakan Algoritma Naive Bayes Classifier (Nbc) Pada Tweet Hashtag #2019gantipresiden," *Kilat*, vol. 9, no. 1, pp. 103–114, 2020, doi: 10.33322/kilat.v9i1.878.
- [16] I. Indriyanti, N. Ichsan, H. Fatah, T. Wahyuni, and E. Ermawati, "Implementasi Orange Data Mining Untuk Prediksi Harga Bitcoin," *J. Responsif Ris. Sains dan Inform.*, vol. 4, no. 2, pp. 118–125, 2022, doi: 10.51977/jti.v4i2.762.