

Penerapan *Data mining* Klasifikasi Tingkat Kepuasan Mahasiswa Terhadap Pelayanan Akademik Menggunakan Metode *Naïve Bayes* Dan *Support Vector Machine* (Studi Kasus Program Studi Sistem Informasi Universitas Labuhanbatu)

¹Dewi Antika, ²Syaiful Zuhri Harahap, ³Rahma Muti Ah, ⁴Angga Putra Juledi

^{1,2,3,4}Sistem Informasi, Fakultas Sains Dan Teknologi, Universitas Labuhanbatu

Email : 1dewiantikauuy@gmail.com, 2syaiifulzuhriharahap@gmail.com,
3rmuthea5@gmail.com, 4anggaj19@gmail.com

Corresponding Author : dewiantikauuy@gmail.com

Abstract

This study was conducted to classify public satisfaction levels using the Support Vector Machine (SVM) algorithm as the primary data analysis method. The objective of this study was to obtain an accurate and reliable prediction model for determining the Satisfaction and Dissatisfaction categories based on the available data. The theoretical basis used refers to the concept of machine learning, specifically SVM, which works by forming an optimal hyperplane to separate data classes. In addition, model evaluation theories such as the Confusion Matrix were used to objectively measure prediction performance. The research methodology included data collection, pre-processing, dividing the dataset into training and test data, and training the SVM model. Evaluation was conducted using accuracy, sensitivity, and specificity metrics to assess the model's ability to predict data accurately. The results and discussion indicate that the SVM successfully classified the majority of data correctly, with the Satisfaction class having a perfect prediction rate while the Dissatisfaction class still had a small error. Further analysis indicated the need for SVM parameter optimization to improve accuracy in the minority class. The conclusion of this study states that the SVM has good performance in classifying public satisfaction data, although it still requires refinement in recognizing certain class patterns. This finding opens up opportunities for developing more adaptive methods to improve predictive performance.

Keywords : *Support Vector Machine, classification, public satisfaction, model evaluation, Confusion Matrix.*

1. Pendahuluan

Peningkatan kualitas pelayanan akademik merupakan salah satu aspek penting dalam menunjang mutu pendidikan tinggi (Manaransyah, Rahman, & Rachmawaty, 2023). Kepuasan mahasiswa terhadap layanan akademik menjadi indikator keberhasilan institusi dalam memberikan pelayanan yang sesuai dengan harapan. Seiring berkembangnya teknologi informasi, pemanfaatan data mining telah menjadi pendekatan yang efektif untuk menganalisis data dalam jumlah besar guna mendapatkan pola dan informasi yang berguna (Fahmi Kamal, Widi Winarso, & Lia Mardiani, 2020). Dalam konteks pendidikan,

data mining dapat dimanfaatkan untuk mengevaluasi tingkat kepuasan mahasiswa terhadap pelayanan yang diberikan, sehingga pihak institusi dapat mengambil kebijakan perbaikan yang lebih tepat sasaran. Penelitian ini difokuskan pada pengklasifikasian tingkat kepuasan mahasiswa terhadap pelayanan akademik dengan memanfaatkan teknik data mining. Dua metode klasifikasi yang digunakan dalam penelitian ini adalah Naïve Bayes dan Support Vector Machine (SVM). Kedua metode ini dipilih karena telah terbukti memiliki performa yang baik dalam pengolahan data tekstual dan numerik, serta mampu menangani masalah klasifikasi dengan tingkat akurasi yang tinggi. Dengan pendekatan ini, diharapkan sistem dapat memberikan gambaran prediktif mengenai kepuasan mahasiswa berdasarkan data survei atau kuesioner yang telah dikumpulkan. Berdasarkan tinjauan literatur, beberapa penelitian sebelumnya telah menerapkan metode klasifikasi untuk analisis kepuasan, namun belum banyak yang secara spesifik menerapkannya dalam konteks pelayanan akademik pada Program Studi Sistem Informasi, khususnya di Universitas Labuhanbatu. Selain itu, perbandingan antara metode Naïve Bayes dan SVM dalam konteks ini masih jarang ditemukan, padahal kedua metode memiliki karakteristik yang berbeda dan potensi akurasi yang kompetitif. Dengan demikian, terdapat celah penelitian (research gap) yang penting untuk dijelajahi guna mengetahui metode mana yang paling efektif untuk klasifikasi tingkat kepuasan mahasiswa. Penelitian ini menawarkan solusi dengan membangun model klasifikasi tingkat kepuasan mahasiswa menggunakan metode Naïve Bayes dan Support Vector Machine, serta membandingkan performa kedua metode tersebut (Pratama, Yanris, Nirmala, & Hasibuan, 2023) (Anggriandi, Utami, & Ariatmanto, 2023). Model ini akan dibangun berdasarkan data hasil kuesioner yang mencerminkan persepsi mahasiswa terhadap pelayanan akademik. Dengan adanya model ini, pihak program studi dapat mengidentifikasi faktor-faktor utama yang mempengaruhi kepuasan mahasiswa dan merumuskan strategi perbaikan yang lebih tepat dan berbasis data. Selain itu, hasil penelitian ini diharapkan dapat memberikan kontribusi terhadap pengembangan sistem pendukung keputusan di lingkungan pendidikan tinggi.

2. Landasan Teori

Data Mining

Data mining adalah proses untuk menemukan pola, hubungan, dan informasi penting dari kumpulan data yang besar dengan menggunakan teknik statistik, matematika, dan kecerdasan buatan (Fatma & Harahap, 2024) (Alam, Alana, & Juliane, 2023). Dalam konteks pendidikan, data mining sering disebut sebagai Educational Data Mining (EDM), yang bertujuan untuk menganalisis data akademik mahasiswa guna meningkatkan kualitas pembelajaran dan layanan institusi pendidikan (Sriwinarti & Juniarti, 2021). Melalui teknik data mining, institusi dapat menggali wawasan tersembunyi dari data yang sebelumnya dianggap tidak bermakna, seperti data kepuasan, kehadiran, nilai, dan perilaku mahasiswa (Diansyah, 2022).

Model Klasifikasi

Model klasifikasi merupakan salah satu metode utama dalam data mining yang digunakan untuk memprediksi kelas atau kategori dari suatu data berdasarkan pola yang

ditemukan dari data latih (Wahyudi et al., 2022) (Mulyanto, Susanti, Rossi, Wajiran, & Borman, 2021). Dalam konteks kepuasan mahasiswa, model klasifikasi berguna untuk memetakan data survei ke dalam kategori tertentu, misalnya puas atau tidak puas (Gatto, Maulana Awangga, & Andarsyah, 2023) (Susetyoko, Wiratmoko Yuwono, & Elly Purwantini, 2022). Model ini bekerja dengan mempelajari karakteristik dari masing-masing kelas berdasarkan atribut-atribut tertentu, kemudian menerapkan pengetahuan tersebut untuk mengklasifikasikan data baru yang belum dikenali.

Metode Naïve Bayes

Naïve Bayes adalah metode klasifikasi yang didasarkan pada Teorema Bayes dengan asumsi bahwa setiap fitur atau atribut data bersifat independen satu sama lain (Nugraha & Romadhony, 2023).v. Meskipun asumsi ini jarang benar dalam kenyataan, metode ini telah terbukti cukup efektif dan efisien (Poerwandono & Perwitosari, 2023)visien dalam berbagai aplikasi klasifikasi, termasuk pengolahan data teks, email filtering, dan prediksi opini pengguna. Keunggulan utama dari metode ini terletak pada kesederhanaannya serta kecepatan dalam memproses data yang besar (Ratih, Retnaningsih, & Dewi, 2022).

Metode Support Vector Machine (SVM)

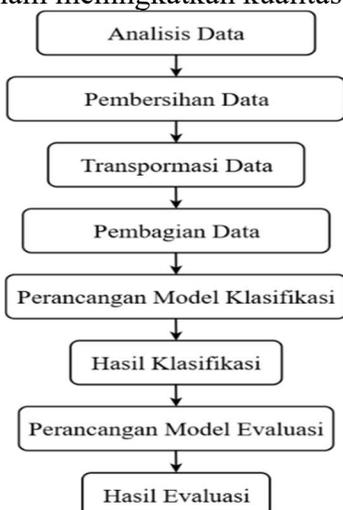
Support Vector Machine (SVM) adalah algoritma klasifikasi yang bekerja dengan mencari hyperplane terbaik yang mampu memisahkan data ke dalam dua atau lebih kelas dengan margin maksimum (Pattnaik & Parvathi, 2022) (Husein, Sipahutar, Dashuah, & Hutauruk, 2023). SVM sangat efektif dalam menangani data berdimensi tinggi dan mampu mengatasi masalah klasifikasi yang kompleks dengan menggunakan teknik kernel. Dengan pendekatan ini, SVM tidak hanya memisahkan data secara linear, tetapi juga dapat mengklasifikasikan data yang tidak linear melalui transformasi ke ruang fitur berdimensi lebih tinggi.

3. Metodologi Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksploratif dan komparatif untuk menganalisis serta membandingkan performa dua algoritma klasifikasi, yaitu Naïve Bayes dan Support Vector Machine (SVM), dalam mengklasifikasikan tingkat kepuasan mahasiswa terhadap pelayanan akademik (Diana Dewi, Qisthi, Lestari, & Putri, 2023) (Arsi, Hidayati, & Nurhakim, 2022). Data yang digunakan diperoleh melalui penyebaran kuesioner kepada mahasiswa Program Studi Sistem Informasi Universitas Labuhanbatu. Kuesioner ini terdiri dari sejumlah pertanyaan yang mencakup aspek-aspek pelayanan akademik seperti pelayanan administrasi, ketersediaan dosen pembimbing, akses informasi akademik, dan kemudahan sistem layanan digital. Data hasil kuesioner kemudian diolah, dibersihkan, dan dipersiapkan dalam format yang sesuai untuk proses klasifikasi.

Langkah-langkah dalam metodologi ini meliputi pengumpulan data, preprocessing data (seperti normalisasi dan encoding), pembagian data menjadi data latih dan data uji, serta pembangunan model klasifikasi menggunakan algoritma Naïve Bayes dan SVM. Selanjutnya, dilakukan proses evaluasi performa model menggunakan metrik akurasi, presisi, recall, dan F1-score untuk menilai tingkat keberhasilan klasifikasi masing-masing metode.

Dengan membandingkan hasil dari kedua metode, penelitian ini bertujuan untuk mengetahui algoritma mana yang lebih optimal dan akurat dalam memetakan tingkat kepuasan mahasiswa berdasarkan data yang tersedia, serta memberikan dasar pengambilan keputusan bagi pengelola program studi dalam meningkatkan kualitas layanan akademik.



Gambar 1. Alur Penelitian

4. Hasil Dan Pembahasan

Analisis Data

Analisis data adalah tahap awal dalam penelitian yang bertujuan untuk memahami isi, struktur, dan karakteristik data yang telah dikumpulkan, sehingga dapat memberikan gambaran awal mengenai kondisi dan pola yang ada. Pada penelitian ini, analisis data dilakukan terhadap hasil kuesioner mahasiswa Program Studi Sistem Informasi Universitas Labuhanbatu yang berisi penilaian terhadap berbagai aspek pelayanan akademik. Proses ini melibatkan identifikasi variabel-variabel penting, pemeriksaan distribusi data, serta pengamatan tren atau anomali yang mungkin memengaruhi hasil klasifikasi. Dengan analisis yang tepat, peneliti dapat menentukan atribut yang relevan untuk digunakan pada tahap berikutnya, sehingga proses klasifikasi menjadi lebih efektif dan akurat.

Tabel 1. Data Kuesioner mahasiswa Program Studi Sistem Informasi

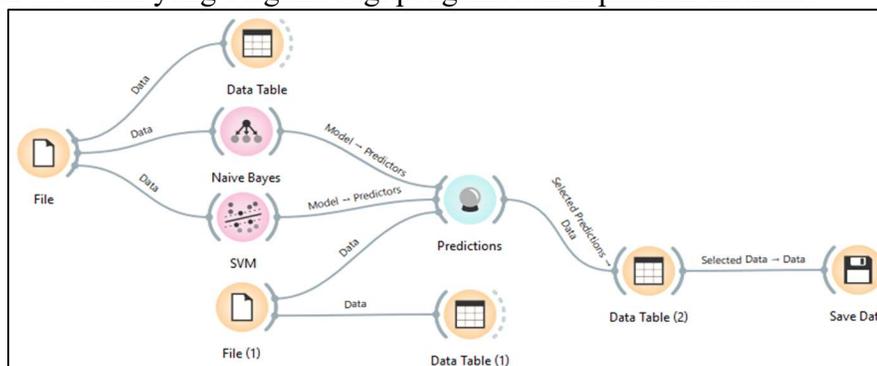
Sarana dan Prasarana					Kemudahan Administrasi					Layanan Akademik				
P1	P2P	P3	P4	P5	P1	P2P	P3	P4	P5	P1	P2P	P3	P4	P5
3	3	3	2	3	4	3	3	3	4	3	3	3	2	3
4	4	4	4		4	4	4	4	4	4	4	4	4	4
3	3	2	1	2	3	3	3	2	2	2	3	2	2	4
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
3	4	3	3	3	3	3	3	3	3	2	3	3	2	3
3	3	2	1	4	4	4	4	4	4	3	4	4	4	4
4	4	4	4	4	4	3	3	4	4	4	4	4	4	4
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
3	3	4	2	3	4	3	3	3	4	4	4	3	3	4
3	3	3	3	3	4	4	4	4	4	4	3	3	4	4
2	3	3	3	3	4	3	3	3	4	3	4	2	4	3
4	4	3	3	3	3	3	3	3	3	3	3	3	3	3
2	2	3	1	3	4	3	3	3	3	3	3	3	3	3
.....
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
3	3	3	3	4	4	3	3	3	4	3	3	3	3	3

Tabel di atas merupakan hasil pengumpulan data kuesioner tingkat kepuasan mahasiswa Program Studi Sistem Informasi Universitas Labuhanbatu terhadap pelayanan

akademik yang mencakup tiga aspek utama, yaitu sarana dan prasarana, kemudahan administrasi, dan layanan akademik. Setiap aspek diukur melalui beberapa indikator (P1, P2P, P3, dst.) dengan skala penilaian kualitatif seperti Sangat Baik, Baik, Cukup, dan Kurang. Data ini memuat identitas responden berupa nama lengkap dan jenis kelamin, yang kemudian dihubungkan dengan jawaban mereka pada setiap indikator. Informasi ini nantinya akan digunakan dalam proses analisis data mining dengan metode klasifikasi Naïve Bayes dan Support Vector Machine untuk mengelompokkan mahasiswa berdasarkan tingkat kepuasan mereka, sekaligus mengidentifikasi faktor-faktor yang paling memengaruhi persepsi mereka terhadap kualitas pelayanan akademik.

Perancangan Model Klasifikasi

Perancangan model klasifikasi merupakan tahap penting dalam penelitian ini yang bertujuan untuk membangun sistem prediksi tingkat kepuasan mahasiswa terhadap pelayanan akademik. Pada tahap ini, data yang telah melalui proses pembersihan dan transformasi dikembangkan menggunakan dua metode klasifikasi, yaitu *Naïve Bayes* dan *Support Vector Machine* (SVM). Model dirancang dengan mempertimbangkan pemilihan atribut relevan dari variabel Sarana dan Prasarana, Kemudahan Administrasi, dan Layanan Akademik yang menjadi faktor utama dalam menentukan kategori kepuasan mahasiswa. Selanjutnya, data dibagi menjadi set training untuk melatih model dan set testing untuk menguji performa prediksi. Proses perancangan ini melibatkan pemilihan parameter, pelatihan algoritma, serta evaluasi menggunakan metrik akurasi untuk memastikan bahwa model mampu mengklasifikasikan data baru dengan tingkat ketepatan yang tinggi dan memberikan wawasan yang berguna bagi pengambilan keputusan institusi.



Gambar 2. Model Support Vector Machine

Pada gambar di atas merupakan model klasifikasi yang dirancang menggunakan dua algoritma, yaitu *Naive Bayes* dan *Support Vector Machine* (SVM). Proses dimulai dari node File yang memuat dataset awal, kemudian data ditampilkan melalui Data Table dan secara paralel digunakan untuk melatih kedua model klasifikasi tersebut. Output model dari Naive Bayes dan SVM dikirim ke node Predictions untuk melakukan prediksi pada data uji. Hasil prediksi ditampilkan kembali melalui Data Table (2), dan jika diperlukan, data tersebut dapat disimpan menggunakan node Save Data. Alur ini menunjukkan proses lengkap mulai dari input data, pelatihan model, prediksi, hingga penyimpanan hasil akhir.

Hasil Klasifikasi

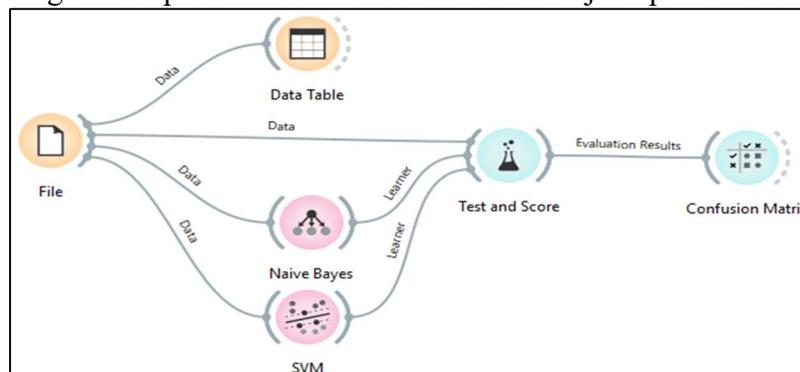
Hasil klasifikasi akan disajikan dalam bentuk tabel yang memperlihatkan prediksi tingkat kepuasan mahasiswa berdasarkan data testing. Tabel tersebut memudahkan analisis perbandingan antara hasil prediksi model dengan data aktual untuk menilai akurasi dan performa metode klasifikasi yang digunakan.

Dari total 30 responden, terdapat 25 mahasiswa (83,33%) yang masuk ke dalam kategori “Puas” dan 5 mahasiswa (16,67%) yang masuk ke dalam kategori “Tidak Puas”. Hal ini menunjukkan bahwa mayoritas responden merasa pelayanan akademik sudah baik, terutama pada aspek Kemudahan Administrasi dan Layanan Akademik yang sebagian besar mendapat penilaian “Baik”. Namun, terdapat beberapa mahasiswa yang merasa tidak puas, umumnya disebabkan oleh penilaian “Kurang Baik” pada satu atau lebih indikator, khususnya di Sarana dan Prasarana atau Layanan Akademik.

Distribusi data ini juga memberikan gambaran awal tentang pola yang mungkin memengaruhi tingkat kepuasan mahasiswa. Misalnya, mahasiswa yang menilai semua indikator sebagai “Baik” cenderung masuk ke kategori “Puas”, sedangkan yang memberikan penilaian “Kurang Baik” pada dua atau lebih indikator cenderung masuk ke kategori “Tidak Puas”. Informasi seperti ini sangat penting untuk proses pelatihan model klasifikasi, karena membantu algoritma mengenali hubungan antara variabel input (penilaian indikator) dan output (kategori kepuasan). Dengan data ini, model dapat mempelajari pola yang ada dan memprediksi tingkat kepuasan mahasiswa baru secara lebih akurat.

Perancangan Model Evaluasi

Perancangan model evaluasi dilakukan untuk mengukur kinerja sistem klasifikasi dalam memprediksi tingkat kepuasan mahasiswa secara akurat dan objektif. Proses ini dimulai dengan membagi data menjadi dua bagian, yaitu data latih untuk membangun model dan data uji untuk mengukur performanya. Selanjutnya, digunakan metrik evaluasi seperti akurasi, precision, recall, dan F1-score untuk mengetahui seberapa baik model mengenali kelas “Puas” dan “Tidak Puas”. Confusion matrix juga digunakan untuk memvisualisasikan jumlah prediksi benar dan salah pada masing-masing kelas, sehingga dapat diketahui pola kesalahan model. Dengan perancangan evaluasi yang terstruktur, pengujian dapat memberikan gambaran menyeluruh tentang kekuatan dan kelemahan model, serta menjadi dasar perbaikan agar hasil prediksi lebih andal dan sesuai tujuan penelitian.



Gambar. 3. Model Evaluasi

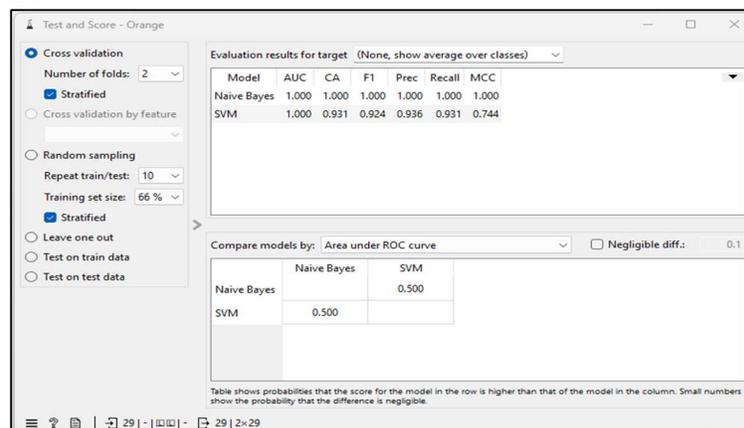
Pada gambar di atas merupakan model evaluasi yang telah dirancang untuk membandingkan kinerja dua algoritma klasifikasi, yaitu Naive Bayes dan Support Vector Machine (SVM), menggunakan data yang diimpor dari sebuah file. Data yang masuk terlebih dahulu dapat ditampilkan melalui Data Table untuk memastikan kebenaran dan kelengkapannya. Selanjutnya, data tersebut dikirim ke kedua algoritma pembelajaran (Naive Bayes dan SVM) yang berperan sebagai learner. Hasil pembelajaran dari kedua algoritma kemudian dievaluasi melalui Test and Score, yang menghitung metrik evaluasi seperti akurasi, presisi, recall, dan F1-score. Akhirnya, Confusion Matrix digunakan untuk menampilkan perbandingan prediksi dan data sebenarnya secara visual, sehingga memudahkan analisis performa masing-masing algoritma dalam memproses data.

Hasil Evaluasi

Hasil evaluasi akan menghasilkan Test and Score yang menampilkan metrik kinerja seperti akurasi, presisi, recall, dan F1-score dari setiap algoritma yang diuji. Selain itu, Confusion Matrix akan memberikan gambaran visual perbandingan antara prediksi model dan data sebenarnya, sehingga memudahkan identifikasi kesalahan klasifikasi.

Hasil Test and Score

Test and Score adalah proses evaluasi model yang menghitung metrik kinerja seperti akurasi, presisi, recall, dan F1-score berdasarkan data uji. Fitur ini membantu membandingkan efektivitas beberapa algoritma secara kuantitatif untuk menentukan model yang paling optimal.



Gambar 4. Hasil Evaluasi

Berdasarkan hasil evaluasi Test and Score, metode Naive Bayes menunjukkan performa yang sangat baik dengan nilai AUC, CA (Classification Accuracy), F1-score, Precision, Recall, dan MCC (Matthews Correlation Coefficient) masing-masing sebesar 1.000. Nilai-nilai sempurna ini menunjukkan bahwa model Naive Bayes mampu mengklasifikasikan data dengan akurasi 100% tanpa kesalahan pada data uji. AUC sebesar 1.000 mengindikasikan kemampuan diskriminasi yang sempurna antara kelas “Puas” dan “Tidak Puas”, sedangkan MCC yang juga bernilai 1.000 menandakan model ini memiliki keseimbangan yang sempurna dalam memprediksi kedua kelas.

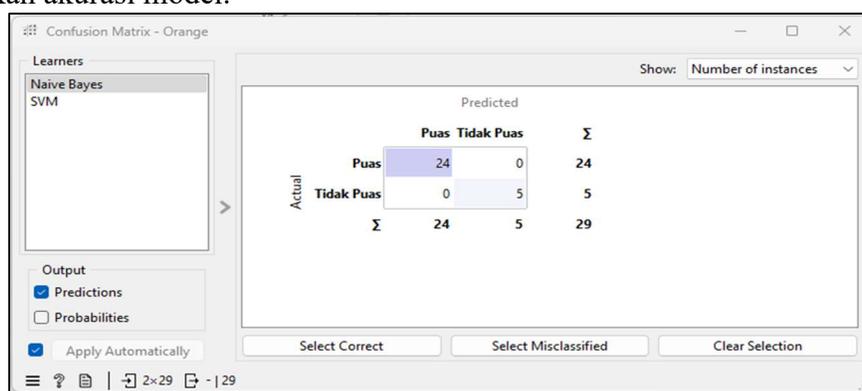
Metode SVM (*Support Vector Machine*) juga memberikan performa yang baik, tetapi sedikit lebih rendah dibandingkan Naive Bayes. Nilai AUC sebesar 1.000 menunjukkan kemampuan pemisahan kelas yang sama baiknya seperti Naive Bayes. Namun, nilai CA sebesar 0.931 menunjukkan adanya sedikit kesalahan prediksi dibandingkan akurasi sempurna pada Naive Bayes. Nilai F1-score sebesar 0.924, Precision sebesar 0.936, dan Recall sebesar 0.931 mengindikasikan bahwa SVM masih cukup efektif, tetapi ada ketidaksempurnaan dalam mengklasifikasikan data, terutama pada keseimbangan antara presisi dan sensitivitas. MCC sebesar 0.744 juga menunjukkan bahwa kekuatan prediksi SVM terhadap kedua kelas tidak seimbang secara sempurna.

Jika dibandingkan secara langsung, terlihat bahwa Naive Bayes unggul dalam semua metrik dibandingkan SVM. Naive Bayes mendapatkan skor sempurna di semua metrik, sementara SVM memiliki penurunan performa khususnya pada CA, F1, Precision, Recall, dan MCC. Hal ini menunjukkan bahwa untuk dataset ini, Naive Bayes lebih mampu menangkap pola data secara konsisten dibandingkan SVM, mungkin karena data memiliki distribusi yang sesuai dengan asumsi probabilistik Naive Bayes.

Berdasarkan perbandingan nilai probabilitas pada tabel "Compare models by Area under ROC curve", nilai 0.500 di kedua arah antara Naive Bayes dan SVM menunjukkan bahwa tidak ada perbedaan signifikan dalam AUC mereka, karena keduanya memiliki kemampuan diskriminasi yang sama sempurnanya. Namun, perbedaan mencolok pada metrik lain seperti MCC dan akurasi membuktikan bahwa Naive Bayes lebih stabil dan andal pada dataset ini, sehingga dapat dipilih sebagai model yang lebih optimal untuk prediksi kepuasan responden.

Hasil Confusion Matrix Metode Naïve Bayes

Confusion Matrix adalah tabel evaluasi yang menunjukkan jumlah prediksi benar dan salah dari suatu model berdasarkan kelas sebenarnya. Matriks ini memudahkan analisis jenis kesalahan yang terjadi, seperti *false positive* dan *false negative*, sehingga membantu meningkatkan akurasi model.



Gambar 5. Hasil Confusion Matrix Metode Naïve Bayes

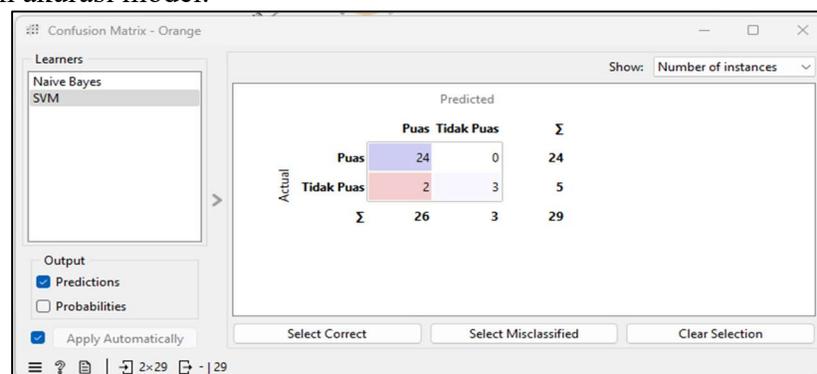
Pada gambar di atas merupakan hasil evaluasi model Naive Bayes menggunakan Confusion Matrix pada data klasifikasi kepuasan, dengan dua kelas yaitu Puas dan Tidak Puas. Terlihat bahwa dari total 24 data aktual dengan label Puas, seluruhnya (24) berhasil diprediksi dengan benar sebagai Puas (True Positive) dan tidak ada yang salah

diklasifikasikan menjadi Tidak Puas (False Negative). Begitu juga pada kelas Tidak Puas, dari 5 data aktual seluruhnya diprediksi dengan benar sebagai Tidak Puas (True Negative) dan tidak ada yang salah diprediksi menjadi Puas (False Positive).

Hasil ini menunjukkan bahwa model Naive Bayes mampu mengklasifikasikan data dengan tingkat akurasi sempurna, yaitu 100%, karena tidak terdapat kesalahan prediksi sama sekali. Kondisi ini juga mengindikasikan bahwa data latih dan data uji memiliki pola yang sangat jelas atau model sudah mampu mempelajari karakteristik masing-masing kelas dengan baik. Namun, perlu diperhatikan bahwa akurasi sempurna seperti ini bisa jadi disebabkan oleh dataset yang relatif kecil atau kurang bervariasi, sehingga pengujian dengan data yang lebih banyak dan beragam penting untuk memastikan performa model tetap konsisten.

Hasil Confusion Matrix Metode Support Vector Machine (SVM)

Confusion Matrix adalah tabel evaluasi yang menunjukkan jumlah prediksi benar dan salah dari suatu model berdasarkan kelas sebenarnya. Matriks ini memudahkan analisis jenis kesalahan yang terjadi, seperti *false positive* dan *false negative*, sehingga membantu meningkatkan akurasi model.



Gambar 6. Hasil Confusion Matrix Metode Support Vector Machine (SVM)

Pada gambar di atas merupakan hasil evaluasi model Support Vector Machine (SVM) menggunakan Confusion Matrix. Dari total 29 data uji, model memprediksi 24 data dengan label Puas secara benar (True Positive) dan 3 data dengan label Tidak Puas secara benar (True Negative). Namun, terdapat 2 data yang seharusnya Tidak Puas tetapi diprediksi sebagai Puas (False Positive), serta tidak ada kasus Puas yang salah diprediksi menjadi Tidak Puas (False Negative = 0). Secara umum, hal ini menunjukkan bahwa model memiliki tingkat akurasi yang cukup tinggi, karena mayoritas prediksi sesuai dengan label aktualnya.

Hasil ini juga mengindikasikan bahwa SVM cukup baik dalam mengenali kelas Puas, terbukti dengan nilai prediksi benar mencapai 100% untuk kelas tersebut (24 dari 24). Namun, untuk kelas Tidak Puas, kinerjanya sedikit menurun karena masih terdapat kesalahan klasifikasi sebesar 40% pada kelas ini (2 dari 5 salah diprediksi). Meski begitu, secara keseluruhan SVM masih mampu memisahkan data dengan baik, yang terlihat dari distribusi prediksi benar yang dominan pada diagonal utama matriks.

5. Kesimpulan

Berdasarkan hasil klasifikasi menggunakan algoritma Support Vector Machine (SVM), model mampu memberikan prediksi yang cukup akurat terhadap data uji dengan dua kategori kelas yaitu Puas dan Tidak Puas. Dari total 29 data uji, model berhasil mengklasifikasikan 24 data Puas secara tepat dan 3 data Tidak Puas secara benar, sehingga mayoritas prediksi berada pada posisi diagonal utama Confusion Matrix yang menandakan ketepatan klasifikasi. Meskipun demikian, masih terdapat kesalahan pada 2 data yang seharusnya berlabel Tidak Puas namun diprediksi sebagai Puas, yang menunjukkan adanya sedikit kelemahan model dalam mengenali karakteristik kelas Tidak Puas. Tidak adanya kesalahan pada prediksi kelas Puas (False Negative = 0) menjadi indikasi bahwa SVM sangat efektif dalam mengidentifikasi data dengan kategori tersebut, dengan tingkat sensitivitas yang sangat tinggi terhadap kelas positif. Secara keseluruhan, hasil evaluasi ini mengonfirmasi bahwa model SVM yang digunakan memiliki performa yang baik dan dapat diandalkan untuk proses klasifikasi serupa, meskipun peningkatan akurasi pada kelas minoritas (Tidak Puas) masih diperlukan agar hasil prediksi lebih seimbang.

6. Daftar Pustaka

- Alam, A., Alana, D. A. F., & Juliane, C. (2023). Comparison Of The C.45 And Naive Bayes Algorithms To Predict Diabetes. *Sinkron*, 8(4), 2641–2650. <https://doi.org/10.33395/sinkron.v8i4.12998>
- Andrianto, R., & Irawan, F. (2023). Implementasi Metode Regresi Linear Berganda Pada Sistem Prediksi Jumlah Tonase Kelapa Sawit di PT . Paluta Inti Sawit. *Jurnal Pendidikan Tambusai*, 7(1), 2926–2934.
- Anggriandi, D., Utami, E., & Ariatmanto, D. (2023). Comparative Analysis of CNN and CNN-SVM Methods For Classification Types of Human Skin Disease. *Sinkron*, 8(4), 2168–2178. <https://doi.org/10.33395/sinkron.v8i4.12831>
- Arifuddin, N. A., Pinastawa, I. W. R., Anugraha, N., & Pradana, M. G. (2023). Classification of Stroke Opportunities with Neural Network and K-Nearest Neighbor Approaches. *Sinkron*, 8(2), 688–693. <https://doi.org/10.33395/sinkron.v8i2.12228>
- Arsi, P., Hidayati, L. N., & Nurhakim, A. (2022). *Komparasi Model Klasifikasi Sentimen Issue Vaksin Covid-19 Berbasis Platform Instagram*. 6, 459–466. <https://doi.org/10.30865/mib.v6i1.3509>
- Diana Dewi, D., Qisthi, N., Lestari, S. S. S., & Putri, Z. H. S. (2023). Perbandingan Metode Neural Network Dan Support Vector Machine Dalam Klasifikasi Diagnosa Penyakit Diabetes. *Cerdika: Jurnal Ilmiah Indonesia*, 3(09), 828–839. <https://doi.org/10.59141/cerdika.v3i09.662>
- Diansyah, S. (2022). *Jurnal Sistim Informasi dan Teknologi Klasifikasi Tingkat Kepuasan Pengguna dengan Menggunakan Metode K-Nearest Neighbour (KNN)*. 4, 1–3. <https://doi.org/10.37034/jsisfotek.v4i1.114>
- Fahmi Kamal, Widi Winarso, & Lia Mardiani. (2020). Peningkatan Kepuasan Mahasiswa Melalui Kualitas Pelayanan Akademik (Studi Kasus Pada Fakultas Keguruan Dan Ilmu Pendidikan Universitas Islam As-Syafi'iyah Jakarta). *Jurnal Ilmiah Akuntansi Dan Manajemen*, 16(1), 33–45. <https://doi.org/10.31599/jiam.v16i1.111>

- Fatma, N., & Harahap, S. Z. (2024). *Analysis of Public Interest in Automatic Motorcycles Using KNN and Neural Network Methods*. 8(2), 1178–1187.
- Gatto, P. A., Maulana Awangga, R., & Andarsyah, R. (2023). Diagnosis Penyakit Demam Berdarah Menggunakan Naïve Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(3), 1676–1681. <https://doi.org/10.36040/jati.v7i3.6891>
- Husein, A. M., Sipahutar, B., Dashuah, R., & Hutauruk, E. (2023). Sentiment Analysis Od Face To Face School Policy On Twitter Social Media With Support Vector Machine(SVM). *Sinkron*, 8(1), 480–486. <https://doi.org/10.33395/sinkron.v8i1.11950>
- Manaransyah, G., Rahman, A., & Rachmawaty, I. K. (2023). Pengaruh Kualitas Pengajaran, Kualitas Pelayanan akademik dan Lingkungan Belajar Virtual pada Kepuasan Mahasiswa Pascasarjana dalam Perkuliahan Daring. *Missio Ecclesiae*, 12(2), 121–132. <https://doi.org/10.52157/me.v12i2.205>
- Mawaddah, A., Dar, M. H., & Yanris, G. J. (2023). Analysis of the SVM Method to Determine the Level of Online Shopping Satisfaction in the Community. *Sinkron*, 8(2), 838–855. <https://doi.org/10.33395/sinkron.v8i2.12261>
- Mulyanto, A., Susanti, E., Rossi, F., Wajiran, W., & Borman, R. I. (2021). Penerapan Convolutional Neural Network (CNN) pada Pengenalan Aksara Lampung Berbasis Optical Character Recognition (OCR). *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 7(1), 52. <https://doi.org/10.26418/jp.v7i1.44133>
- Naufal, M. F., Arifin, T., & Wirjawan, H. (2023). Analisis Perbandingan Tingkat Performa Algoritma SVM, Random Forest, dan Naïve Bayes untuk Klasifikasi Cyberbullying pada Media Sosial. *Jurnal Riset Sistem Informasi Dan Teknik Informatika (JURASIK)*, 8, 82. Retrieved from <https://tunasbangsa.ac.id/ejurnal/index.php/jurasik>
- Nugraha, A. B., & Romadhony, A. (2023). Identification of 10 Regional Indonesian Languages Using Machine Learning. *Sinkron*, 8(4), 2203–2214. <https://doi.org/10.33395/sinkron.v8i4.12989>
- Pattnaik, G., & Parvathi, K. (2022). Machine learning-based approaches for tomato pest classification. *Telkomnika (Telecommunication Computing Electronics and Control)*, 20(2), 321–328. <https://doi.org/10.12928/TELKOMNIKA.v20i2.19740>
- Poerwandono, E., & Perwitosari, J. (2023). Penerapan Data Mining Untuk Penilaian Kinerja Karyawan Di PT. Riksa Dinar Djaya Menggunakan Metode Naïve Bayes Classification (Edhy Poerwandono 1 , Faizal Joko Perwitosari 2) Penerapan Data Mining Untuk Penilaian Kinerja Karya Di PT Riksa Dinar Djaya Men. *Jurnal Sains Dan Teknologi*, 5(1), |pp. Retrieved from <https://doi.org/10.55338/saintek.v5i1.1416>
- Pratama, H. A., Yanris, G. J., Nirmala, M., & Hasibuan, S. (2023). *Implementation of Data Mining for Data Classification of Visitor Satisfaction Levels*. 8(3), 1832–1851.
- Ramadhan, A., Susetyo, B., & Indahwati. (2019). Penerapan Metode Klasifikasi Random Forest Dalam Mengidentifikasi Faktor Penting Penilaian Mutu Pendidikan. *Jurnal Pendidikan Dan Kebudayaan*, 4(2), 169–182. <https://doi.org/10.24832/jpnk.v4i2.1327>
- Ratih, I. D., Retnaningsih, S. M., & Dewi, V. M. (2022). Klasifikasi Kualitas Tanah Menggunakan Metode Naive Bayes Classifier. *Jurnal Aplikasi Matematika Dan*

- Statistik*), 1(1), 11–20.
- Septiarini, A., Rizqi Saputra, Andi Tejawati, & Masna Wati. (2021). Deteksi Sarung Samarinda Menggunakan Metode Naive Bayes Berbasis Pengolahan Citra. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(5), 927–935. <https://doi.org/10.29207/resti.v5i5.3435>
- Sikumbang, E. D., Ariani, F., Handayani, T., & Ramanda, K. (2022). Penerapan Algoritma C4.5 Untuk Menentukan Tingkat Kepuasan Pelanggan Kartu Telkomsel Prabayar. *Jurnal Sains Komputer & Informatika*, 6(September), 811–820.
- Sriwinarti, N. K., & Juniarti, P. (2021). *Analisis Metode K-Nearest Neighbors (K-NN) Dan Naive Bayes Dalam Memprediksi Kelulusan Mahasiswa (Analysis of K-Nearest Neighbors (K-NN) and Naive Bayes Methods in Predicting Student Graduation)*. 3(2), 106–112.
- Susetyoko, R., Wiratmoko Yuwono, & Elly Purwantini. (2022). Model Klasifikasi Pada Seleksi Mahasiswa Baru Penerima KIP Kuliah Menggunakan Regresi Logistik Biner. *Jurnal Informatika Polinema*, 8(4), 31–40. <https://doi.org/10.33795/jip.v8i4.914>
- Wahyudi, A., Ovelia Tampubolon, S., Afrilia Putri, N., Ghassa, A., Rasywir, E., & Kisbianty, D. (2022). Penerapan Data Mining Algoritma Naive Bayes Clasifier Untuk Mengetahui Minat Beli Pelanggan Terhadap INDIHOME. *Jurnal Informatika Dan Rekayasa Komputer(JAKAKOM)*, 2(2), 240–247. <https://doi.org/10.33998/jakakom.2022.2.2.111>