

Optimasi *Support Vector Machine* (SVM) Menggunakan *Naive Bayes* dan *Decision Tree* untuk Klasifikasi Tema Tugas Akhir Mahasiswa Sistem Informasi

Arfika ^{1*}

^{1*} Program Studi Teknik Komputer, Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) Abulyatama Aceh, Kabupaten Aceh Besar, Provinsi Aceh, Indonesia.

*Correspondence email:
Arfika_bna@gmail.com

Received: 19 May 2024
Accepted: 18 July 2024
Published: 1 August 2024

Full list of author information is available at the end of the article.

Abstract

The School of Informatics and Computer Management (STMIK) Abulyatama faces challenges in classifying student thesis topics. The current manual classification process is inefficient and ineffective. Therefore, the implementation of data mining techniques is needed to manage this data, specifically to automate the classification process. This study aims to optimize the Support Vector Machine (SVM) model by integrating Naive Bayes and Decision Tree algorithms to improve the accuracy of thesis topic classification. Based on the analysis, it can be concluded that both SVM and K-Means can be utilized by decision-makers to categorize thesis topics as a decision support system. Naive Bayes and Decision Tree were shown to optimize SVM and enhance its accuracy. Naive Bayes achieved the highest accuracy at 82.50%, while Decision Tree recorded an accuracy of 62.50%, making Naive Bayes the most suitable model for thesis classification.

Keywords: Support Vector Machine; Naive Bayes; Decision Tree.

Abstrak

Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) Abulyatama mengalami kesulitan dalam mengelompokkan tema tugas akhir mahasiswa. Proses pengklasifikasian yang dilakukan secara manual selama ini tidak efektif dan efisien. Oleh karena itu, diperlukan penerapan teknik data mining untuk mengelola data tersebut, khususnya untuk mengotomatiskan proses klasifikasi. Penelitian ini bertujuan untuk mengoptimalkan model Support Vector Machine (SVM) dengan mengintegrasikan algoritma Naive Bayes dan Decision Tree guna meningkatkan akurasi klasifikasi tema tugas akhir. Berdasarkan hasil analisis, dapat disimpulkan bahwa model SVM dan K-Means dapat digunakan oleh para pengambil kebijakan untuk mengategorikan tema tugas akhir sebagai sistem pendukung keputusan. Naive Bayes dan Decision Tree terbukti dapat mengoptimalkan SVM dan meningkatkan akurasi. Model Naive Bayes memiliki akurasi tertinggi sebesar 82,50%, sementara Decision Tree mencatatkan akurasi sebesar 62,50%, sehingga Naive Bayes merupakan model yang paling sesuai untuk klasifikasi tugas akhir.

Kata Kunci: Support Vector Machine; Naive Bayes; Decision Tree.



1. Pendahuluan

Pengambilan mata kuliah tugas akhir merupakan syarat wajib bagi mahasiswa untuk memperoleh gelar sarjana di perguruan tinggi. Jumlah mahasiswa yang mengikuti mata kuliah tugas akhir semakin meningkat, menyebabkan kesulitan dalam pengelolaan dan pengelompokan data oleh pihak akademik. STMIK Abulyatama adalah salah satu institusi yang mengalami kendala dalam pengelompokan tema tugas akhir mahasiswa. Kesulitan ini muncul karena jumlah data yang besar dan variasi tema yang beragam sesuai dengan penelitian yang dilakukan oleh masing-masing mahasiswa. Proses klasifikasi yang selama ini dilakukan secara manual sangat tidak efisien dan tidak efektif, sehingga diperlukan penerapan teknik data mining untuk mengelola data tersebut, khususnya untuk keperluan pengklasifikasian. Data mining sendiri adalah proses ekstraksi pola yang menarik dari data berjumlah besar yang sebelumnya tidak diketahui dan memiliki nilai yang berguna. Proses ini terbukti dapat membantu pengelompokan berbagai topik tugas akhir, sehingga dapat diidentifikasi tren penelitian yang terjadi dari tahun ke tahun di universitas. Beberapa penelitian sebelumnya menunjukkan bahwa model klasifikasi yang diterapkan melalui data mining mampu mencapai tingkat akurasi yang cukup tinggi, bahkan mendekati 100%. Selain untuk klasifikasi, data mining juga dapat diterapkan untuk prediksi yang bermanfaat dalam pengelolaan data dalam jumlah besar.

Support Vector Machine (SVM) merupakan salah satu algoritma pembelajaran mesin yang digunakan secara luas dalam berbagai aplikasi klasifikasi dan regresi. SVM berfungsi untuk menemukan hyperplane terbaik yang dapat memisahkan data dalam ruang fitur ke dalam dua kategori yang berbeda. Algoritma ini sangat efektif dalam menangani data dengan dimensi yang tinggi, seperti yang dibuktikan dalam berbagai penelitian di bidang analisis sentimen, klasifikasi teks, dan pengenalan pola. Samantri dan Afiyati (2024) membandingkan algoritma SVM dengan *Random Forest* dalam analisis sentimen kebijakan kenaikan harga BBM di Indonesia, menunjukkan bahwa SVM memiliki keunggulan dalam mengklasifikasikan data dengan akurasi yang cukup tinggi dibandingkan dengan metode lain. Penelitian ini memperlihatkan kekuatan SVM dalam menganalisis data yang kompleks dengan dimensi yang tinggi, terutama dalam kasus analisis sentimen. SVM juga digunakan dalam bidang medis, seperti yang ditunjukkan oleh Widyawati *et al.* (2023) dalam studi prediksi kanker paru-paru. Penelitian ini membandingkan performa SVM dengan *Naive Bayes* dan *Decision Tree*, dan hasilnya menunjukkan bahwa SVM dapat memberikan akurasi yang tinggi dalam klasifikasi data medis, terutama ketika dibandingkan dengan model lain seperti *Decision Tree* dan *Naive Bayes*. Hal ini menegaskan keunggulan SVM dalam menangani data yang tidak hanya besar tetapi juga sangat bervariasi. Apriansyah *et al.* (2024) menyoroti optimalisasi nilai K dalam algoritma KNN untuk klasifikasi spam dan HAM dalam teks SMS, tetapi tetap mengakui keunggulan SVM dalam masalah klasifikasi yang memerlukan pengenalan pola yang presisi. Sementara itu, Amini dan Setiawan (2024) menggunakan SVM dalam analisis sentimen di Twitter terkait kandidat Wakil Presiden Indonesia pada tahun 2024, dan menunjukkan bahwa algoritma SVM sangat efektif dalam menangani teks dan sentimen yang kompleks di media sosial. Penggunaan SVM dalam pengenalan pola juga diaplikasikan oleh Legito *et al.* (2023) dalam pengenalan flora menggunakan *Convolutional Neural Networks* (CNN). Sementara penelitian ini berfokus pada CNN, SVM tetap merupakan metode yang efektif dalam tugas klasifikasi gambar sederhana, terutama ketika fitur-fitur yang jelas dapat diekstraksi. Suspendi *et al.* (2024) menerapkan pembelajaran mendalam (deep learning) untuk memprediksi pasar saham, tetapi menunjukkan bahwa dalam beberapa kasus, SVM lebih cepat dalam mengklasifikasikan data dan memberikan hasil yang dapat diandalkan tanpa memerlukan model yang kompleks seperti deep learning. Hal ini memperkuat argumen bahwa SVM masih relevan untuk aplikasi yang membutuhkan efisiensi dan akurasi tinggi dengan sumber daya komputasi yang terbatas. Qubra dan Saputra (2024) menunjukkan keberhasilan SVM dalam klasifikasi berita hoax menggunakan metode *Naive Bayes* sebagai perbandingan. Dalam penelitian ini, meskipun *Naive Bayes* memberikan hasil yang baik, SVM tetap unggul dalam klasifikasi yang memerlukan keputusan yang lebih akurat pada data yang lebih rumit. Penelitian lain oleh Syah *et al.* (2023) mengaplikasikan SVM bersama dengan algoritma lain seperti AdaBoost dan Particle Swarm Optimization untuk analisis sentimen di Twitter terkait IndihomeCare, menunjukkan bahwa SVM dapat memberikan akurasi yang tinggi ketika dikombinasikan dengan teknik optimasi. Pernama dan Purnomo (2023) juga menggunakan SVM untuk menganalisis risiko pinjaman, membandingkannya dengan Artificial Neural Network dan *Naive Bayes*. Hasilnya menunjukkan bahwa SVM mampu menghasilkan prediksi risiko yang lebih akurat. Hafid *et al.* (2023) serta Fatmasari *et al.* (2023) juga menggunakan SVM dalam analisis sentimen dan klasifikasi komentar di media sosial dan layanan akademik. Penelitian mereka mendukung keandalan SVM dalam menangani data teks dan memberikan hasil klasifikasi yang lebih akurat dibandingkan dengan algoritma lain. SVM terus menjadi salah satu algoritma utama dalam berbagai

aplikasi klasifikasi data besar dan kompleks. Keunggulan utamanya terletak pada kemampuannya untuk bekerja dengan baik pada data berdimensi tinggi, serta kemampuannya untuk memisahkan kelas data dengan margin yang maksimal. Ini menjadikan SVM sebagai pilihan yang sangat baik untuk berbagai tugas klasifikasi, baik dalam bidang akademik maupun industri.

Naive Bayes adalah algoritma pembelajaran mesin berbasis probabilistik yang sering digunakan dalam klasifikasi teks, analisis sentimen, dan deteksi pola. Algoritma ini bekerja dengan asumsi bahwa setiap fitur bersifat independen satu sama lain, sebuah asumsi yang dikenal sebagai *naive independence*. Meskipun sederhana, *Naive Bayes* telah terbukti efektif dalam berbagai aplikasi data mining, khususnya dalam menangani dataset yang besar dan kompleks. Firmansyah dan Lestariningsih (2024) dalam studi mereka mengenai analisis sentimen ulasan aplikasi di Google Playstore menggunakan *Naive Bayes*, menemukan bahwa algoritma ini mampu memberikan hasil yang cukup akurat dalam mengklasifikasikan ulasan aplikasi dengan metode yang sederhana. Algoritma ini telah digunakan pula dalam sistem pakar, seperti yang diterapkan oleh Tyar dan Wahyuddin (2022) untuk mendeteksi hama pada tanaman alpukat dan oleh Efendi *et al.* (2022) dalam deteksi penyakit ikan arwana, yang menunjukkan kemampuan *Naive Bayes* dalam klasifikasi berdasarkan data terbatas. Selain itu, *Naive Bayes* juga diterapkan dalam diagnosis penyakit pada tanaman cabai oleh Efendi *et al.* (2021), yang menekankan keefektifan metode ini dalam klasifikasi data berbasis aturan. Studi yang dilakukan oleh Pasaribu dan Siani (2023) serta Tundo *et al.* (2024) pada analisis sentimen pengguna Shopee dan penggunaan rokok memperlihatkan bahwa *Naive Bayes* dapat diandalkan dalam analisis teks opini, baik di media sosial maupun dalam ulasan produk. Dalam konteks analisis sentimen terhadap platform live streaming dan isu nasional seperti naturalisasi pemain sepak bola dan pemindahan ibu kota, Nugraha dan Said (2024), Jaya dan Lestari (2024), serta Azis dan Wahyudi (2024) menunjukkan bahwa *Naive Bayes* memberikan akurasi yang tinggi dalam mengelompokkan sentimen positif dan negatif. Ini didukung oleh penelitian Julianto *et al.* (2024) yang menggunakan algoritma ini untuk menganalisis tanggapan publik terhadap program internet gratis di platform media sosial. Dalam aplikasi data mining, Lestari dan Sugiyono (2024) menunjukkan keandalan *Naive Bayes* dalam klasifikasi sentimen terhadap tumor jinak pada media sosial, sementara Fahmi dan Sutisna (2024) menerapkan algoritma ini dalam klasifikasi gejala penyakit TB, yang menunjukkan kegunaannya dalam sektor kesehatan. Azzahra dan Akbar (2024) membandingkan ketepatan algoritma C4.5 dengan *Naive Bayes* dalam pengiriman barang, menemukan bahwa *Naive Bayes* mampu bersaing dengan algoritma yang lebih kompleks dalam konteks logistik. *Naive Bayes* merupakan algoritma yang sederhana namun efektif, terutama dalam klasifikasi teks dan analisis sentimen. Meskipun algoritma ini membuat asumsi independensi yang sederhana, banyak penelitian menunjukkan bahwa *Naive Bayes* mampu memberikan akurasi yang kompetitif dalam berbagai aplikasi, seperti deteksi penyakit, analisis ulasan produk, dan klasifikasi data kesehatan. Keunggulannya terletak pada kecepatan dan efisiensi komputasi, terutama dalam menangani dataset besar.

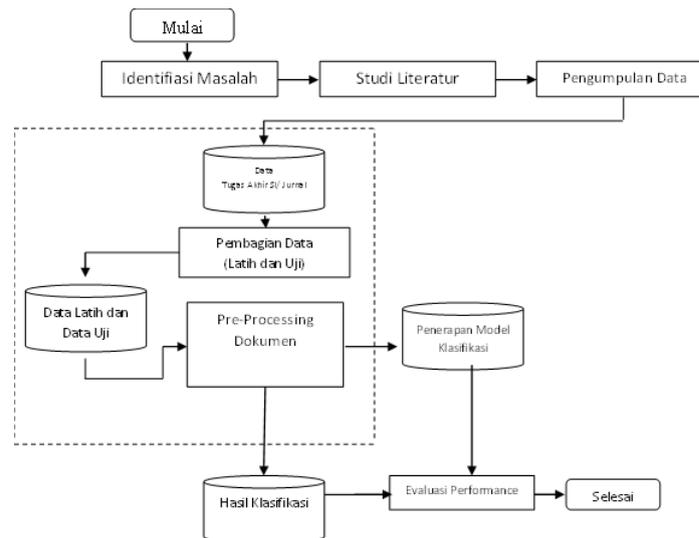
Decision Tree adalah salah satu metode pembelajaran mesin yang digunakan untuk tugas klasifikasi dan regresi. Metode ini bekerja dengan memecah dataset menjadi subset berdasarkan fitur-fitur tertentu, dengan tujuan akhir membentuk struktur pohon keputusan. Setiap simpul dalam pohon mewakili fitur data, cabang-cabangnya mewakili hasil dari tes pada fitur tersebut, dan daun akhir menunjukkan label atau klasifikasi yang dihasilkan. Salah satu keunggulan *Decision Tree* adalah interpretasi visualnya yang jelas, sehingga mempermudah pengambilan keputusan berbasis data. Lestari (2024) menerapkan *Decision Tree* dalam klasifikasi tingkat pemahaman siswa pada kegiatan belajar mengajar di SDN Malaka Jaya, menunjukkan bahwa metode ini mampu mengidentifikasi pola belajar siswa secara akurat. Penelitian ini membuktikan bahwa *Decision Tree* bisa digunakan dalam dunia pendidikan untuk menganalisis kinerja siswa berdasarkan berbagai kriteria. Afkar *et al.* (2024) menggunakan algoritma C4.5, varian dari *Decision Tree*, untuk memprediksi produksi cabai di Provinsi Aceh. Algoritma ini berhasil memberikan prediksi yang tepat, menyoroti kemampuan *Decision Tree* dalam menangani data yang terkait dengan hasil pertanian. Demikian pula, Limanauw dan Oetama (2024) menggunakan *Decision Tree* untuk menganalisis pandangan milenial tentang asuransi, menghasilkan informasi berguna yang dapat diterapkan pada industri asuransi. Dalam bidang mitigasi bencana, Mardiani *et al.* (2024) menerapkan *Decision Tree* untuk memprediksi potensi tsunami akibat gempa di Indonesia. Metode ini membantu mengidentifikasi daerah yang rentan terkena dampak bencana, menunjukkan pentingnya *Decision Tree* dalam pengambilan keputusan cepat di situasi kritis. Puteri *et al.* (2024) membandingkan *Decision Tree* dengan deep learning untuk memprediksi mahasiswa yang berpotensi drop out. Hasil penelitian menunjukkan bahwa *Decision Tree* memberikan hasil yang sebanding dengan deep learning, dengan proses yang lebih cepat. Ini menunjukkan bahwa *Decision Tree* tetap relevan dalam berbagai kasus di mana kecepatan dan interpretasi model sangat diperlukan. Sari Siregar *et al.* (2024) menerapkan algoritma Chaid, varian dari *Decision Tree*, untuk

menentukan pola penerimaan mahasiswa baru. Penelitian ini menunjukkan bahwa Chaid efektif dalam mengidentifikasi faktor-faktor yang paling mempengaruhi penerimaan mahasiswa, yang bisa membantu institusi pendidikan dalam merancang strategi penerimaan yang lebih baik. *Decision Tree* adalah metode yang sederhana namun efektif untuk tugas klasifikasi dan regresi. Kemampuannya untuk menghasilkan model yang mudah diinterpretasikan secara visual menjadikannya pilihan yang tepat dalam situasi yang memerlukan pemahaman langsung terhadap proses pengambilan keputusan. Metode ini diterapkan dalam berbagai bidang seperti pendidikan, pertanian, bencana, dan industri keuangan, menunjukkan fleksibilitasnya dalam menyelesaikan berbagai masalah berbasis data.

Peningkatan jumlah mahasiswa yang mengambil mata kuliah tugas akhir menimbulkan kesulitan dalam pengelompokan tema oleh pihak akademik, seperti yang terjadi di STMIK Abulyatama. Teknik data mining diperlukan untuk mengelola data yang besar dan beragam. *Support Vector Machine (SVM)* merupakan salah satu algoritma unggulan yang efektif dalam mengklasifikasikan data berdimensi tinggi, termasuk dalam analisis sentimen, prediksi medis, dan klasifikasi gambar. SVM terbukti memberikan akurasi tinggi dan efisien dalam berbagai studi. *Naive Bayes*, meskipun sederhana, juga sangat efektif dalam klasifikasi teks dan analisis sentimen dengan kecepatan tinggi dalam komputasi. Berbagai penelitian menunjukkan bahwa *Naive Bayes* memberikan hasil yang kompetitif dalam klasifikasi ulasan produk, analisis kesehatan, dan prediksi logistik. *Decision Tree* adalah metode klasifikasi dan regresi yang efektif dengan visualisasi yang mudah dipahami, sehingga banyak digunakan di bidang pendidikan, prediksi pertanian, dan mitigasi bencana. Meskipun sederhana, *Decision Tree* tetap relevan untuk kasus yang membutuhkan interpretasi langsung dan kecepatan. Berdasarkan latar belakang tersebut, dapat disimpulkan bahwa penerapan data mining dalam pengelolaan data berjumlah besar sangat membantu dan terbukti lebih efektif dan efisien. Oleh karena itu, penelitian ini dilakukan dengan judul Optimasi *Support Vector Machine (SVM)* Menggunakan *Naive Bayes* dan *Decision Tree* untuk Klasifikasi Tema Tugas Akhir Mahasiswa Sistem Informasi.

2. Metode

Suatu penelitian dikatakan berhasil apabila telah tercapainya tujuan penelitian tersebut. Namun sebelum mencapai tujuannya harus melalui banyak proses atau tahapan-tahapan. Adapun tahapan-tahapan tersebut penulis tuangkan dalam diagram penelitian berikut ini.

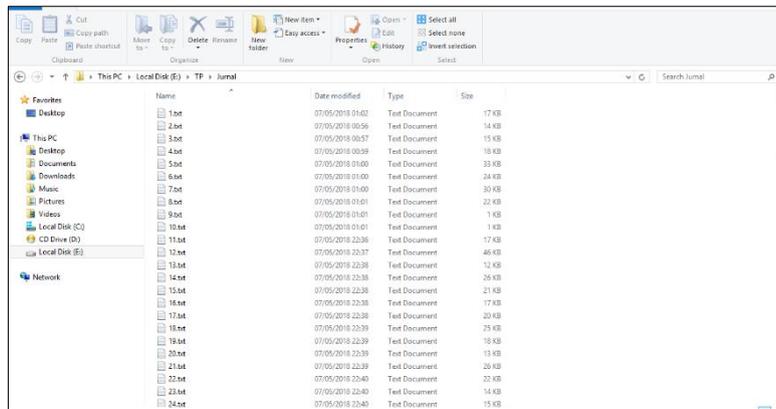


Gambar 1. Tahapan Penelitian

Berdasarkan diagram alir penelitian diatas dapat dijelaskan bahwa untuk mencapai tujuan penelitian yang diajukan pada proposal ini terdapat beberapa proses atau tahapan yang harus dilalui. Pertama mulai selanjutnya indentifikasi masalah, lalu studi literatur dan pengumpulan data, data tugas akhir mahasiswa sistem informasi disimpan dalam database, setelah data terkumpul data akan di bagi menjadi 2 (dua) kelompok yaitu data latih dan data uji. Hasil pengelompokan data latih dan uji pun disimpan dalam database. Adapun data penelitian ini dikumpulkan dari internet dalam bentuk jurnal penelitian jurusan komputer dari berbagai kampus

dan penerbit jurnal di Indonesia yang telah dipublikasi, diantaranya adalah Universitas Indonesia, Politeknik Harapan Bersama Tegal, Universitas Ahmad Dahlan Yogyakarta, ITB Bandung, IPB Bogor dan lain-lain. Semua jurnal ini akan dikelompokkan dalam 4 (empat) pengelompokan tema, yaitu tema data mining, jaringan, kecerdasan buatan, dan website. Suatu penelitian akan memerlukan perangkat untuk membantu tercapainya suatu tujuan penelitian. Adapun alat dan bahan yang diperlukan dalam penelitian ini berupa *hardware* (perangkat keras) dan *software* (perangkat lunak).

Dataset diperoleh dari data berbagai sumber jurnal di Indonesia pada tahun terbit antara 2015 sampai dengan 2017, sebanyak 80 jurnal dengan berbagai macam jenis tema terdiri dari data mining, jaringan, kecerdasan buatan, dan website. Alat atau tools yang digunakan dalam penelitian ini adalah dengan menggunakan Software Rapid Miner 8.1, sebagai pendukung pengolahan data menggunakan Microsoft Excel 2016.



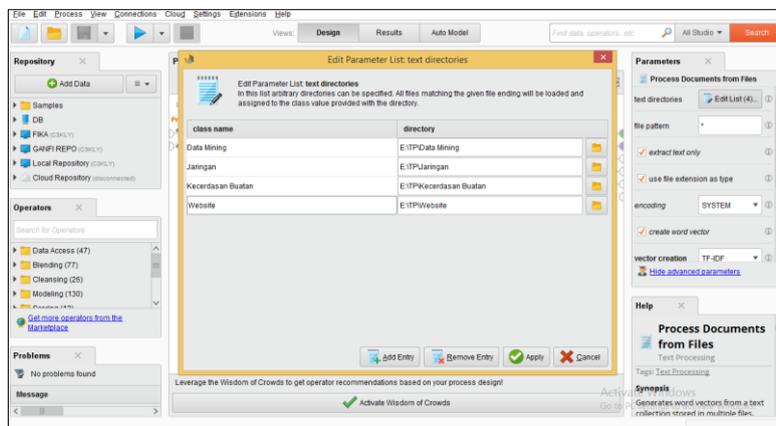
Gambar 2. File Jurnal Yang Telah dikonversi menjadi Format Text

Data jurnal berjumlah 80 jika dirincikan seperti pada tabel berikut 1.

Tabel 1. Pengelompokan Manual Jurnal

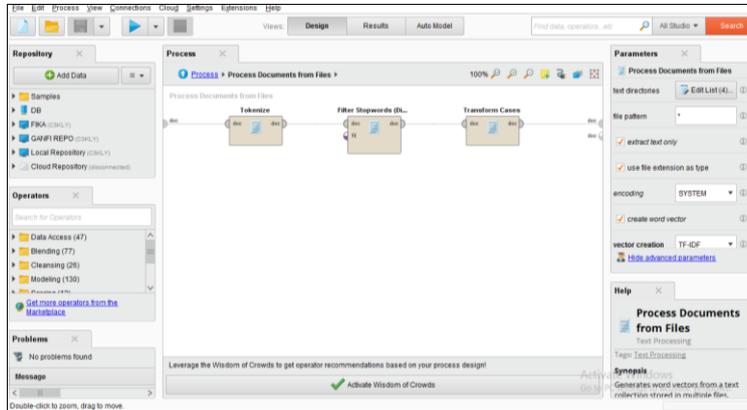
No	Tema Jurnal	Jumlah Jurnal
1	Data Mining	20
2	Jaringan	20
3	Kecerdasan Buatan	20
4	Website	20
Total		80

Pada processing tekx maka akan dilakukan tahap seperti tokenizing, stemming, *stopwords*, dan n-gram. Adapun langkah-langkah sebagai pengujian processing text akan dilakukan tahap sebagai berikut. Pada proses *Process Documents From File*, bertujuan untuk memproses dokumen agar mendapatkan daftar kata-kata yang berbeda dan jumlah individu dari setiap dokumen. Pada bagian ini juga nantinya akan memilih file jurnal yang telah dikonversi sebelumnya dan memilih folder target jurnal.



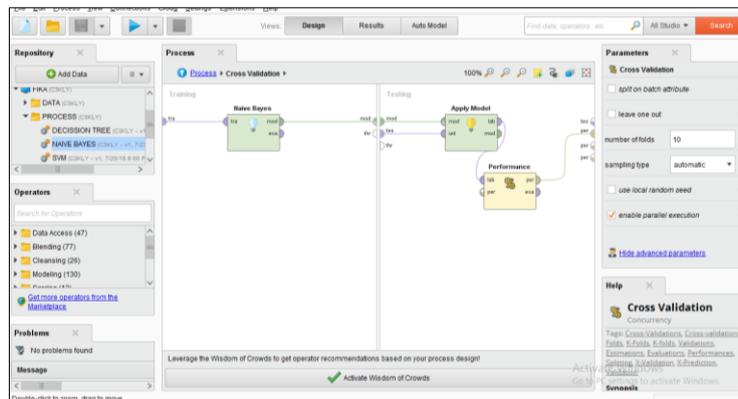
Gambar 3. Parameter *List Document*

Setelah data jurnal dipilih, tahap pemrosesan teks dilakukan melalui beberapa subproses agar dokumen dapat digunakan untuk proses pengelompokan. Pertama, dilakukan subproses Documents From File, yang terdiri dari beberapa langkah. Langkah pertama adalah Tokenize, yaitu proses memisahkan teks menjadi beberapa token berdasarkan pembatas seperti spasi atau tanda baca. Selanjutnya, dilakukan Transform Case, di mana seluruh karakter dalam dokumen diubah menjadi huruf kecil atau huruf besar secara konsisten. Terakhir, dilakukan Filter Stopwords, yaitu menghapus setiap token yang sesuai dengan daftar kata umum (common words) yang biasanya muncul dalam jumlah besar dan dianggap tidak memiliki makna penting dalam proses analisis. Proses-proses ini memastikan bahwa dokumen yang diproses siap untuk pengelompokan berdasarkan pola teks yang lebih relevan.

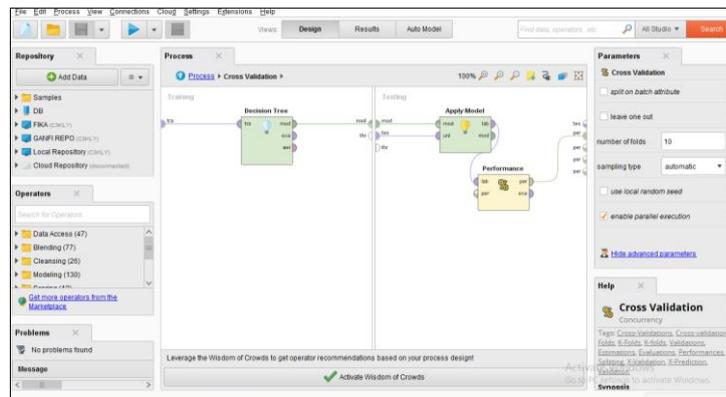


Gambar 4. Sub Proses Documents From File

Proses penerapan model dilakukan terhadap hasil sebelum proses yang merupakan representasi data dalam bentuk model ruang vector. Metode pertama ialah penerapan model dokumen yang ada dengan *naïve bayes* dan *decision tree*.

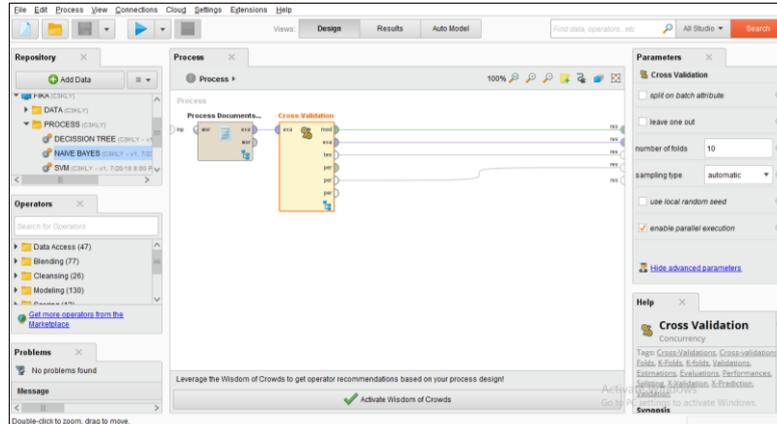


Gambar 5. Penerapan Model Naïve Bayes

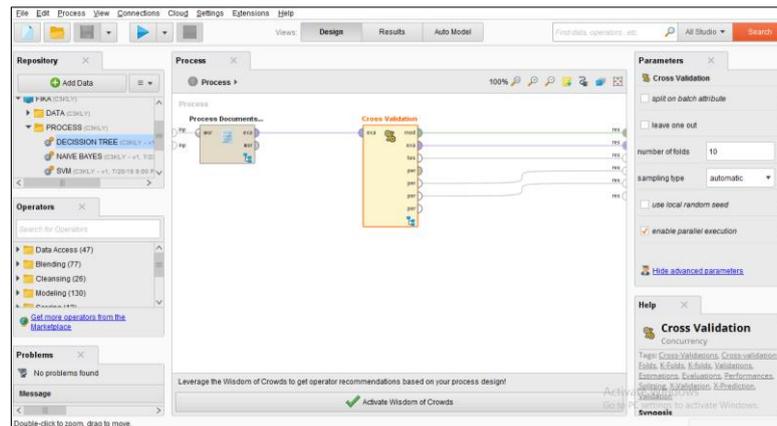


Gambar 6. Penerapan Model Decision Tree

Sebagai evaluasi dari model yang diusulkan, yaitu dengan menggunakan *Kfolds Cross Validations* untuk mencari nilai akurasi yang kemudian hasil dari akurasi tersebut dievaluasi dengan cara membandingkan tingkat akurasi yang dihasilkan oleh model *naive bayes* dan *decision tree*.



Gambar 7. Penerapan Cross Validation Model Naïve Bayes



Gambar 8. Penerapan Cross Validation Model Decision Tree

3. Hasil dan Pembahasan

3.1 Hasil

Masih sulitnya dalam menentukan klasifikasi tema tugas akhir mahasiswa sering dialami oleh setiap perguruan tinggi. Tujuan penelitian ini adalah memberikan sebuah penunjang keputusan bagi pengambil kebijakan di program study khususnya pada STMIK Abulyatama agar setiap judul tugas akhir mahasiswa yang diajukan sesuai dengan tema yang telah ditentukan. Penggunaan algoritma *Naïve bayes* dan *Decision Tree* yang digunakan akan membuktikan tingkat klasifikasi judul tugas akhir yang lebih baik. Untuk lebih rinci tahapan perancangan tersebut dapat dilihat pada penjelasan pada sub bab berikut. Dalam metode *Naïve bayes*, metode ini bertujuan untuk mengetahui tingkat akurasi klasifikasi. Data yang tidak terklasifikasi secara benar jika hasil sentimennya berbeda dari sentimen yang seharusnya. Model distribusi untuk atribut label data mining, jaringan, kecerdasan buatan, dan website adalah sebagai berikut:

Kelas Data Mining	: 4.187 distribusi
Kelas Jaringan	: 4.187 distribusi
Kelas Kecerdasan Buatan	: 4.187 distribusi
Kelas Website	: 4.187 distribusi

Dari hasil *cross validation* maka didapatkan hasil akurasi dan klasifikasi error dari penggunaan model *naive bayes* dan *decision tree* sebagai berikut. Hasil akurasi dengan menggunakan algoritma *naive bayes* sebesar 82.50% seperti ditunjukkan pada gambar 9 berikut:

accuracy: 82.50% +/- 10.00% (mikro: 82.50%)

	true Data Mining	true Jaringan	true Kecerdasan Bu...	true Website	class precision
pred. Data Mining	16	0	2	2	80.00%
pred. Jaringan	0	18	0	0	100.00%
pred. Kecerdasan B...	3	0	16	2	76.19%
pred. Website	1	2	2	16	76.19%
class recall	80.00%	90.00%	80.00%	80.00%	

Gambar 9. Hasil Akurasi Model *Naive bayes*

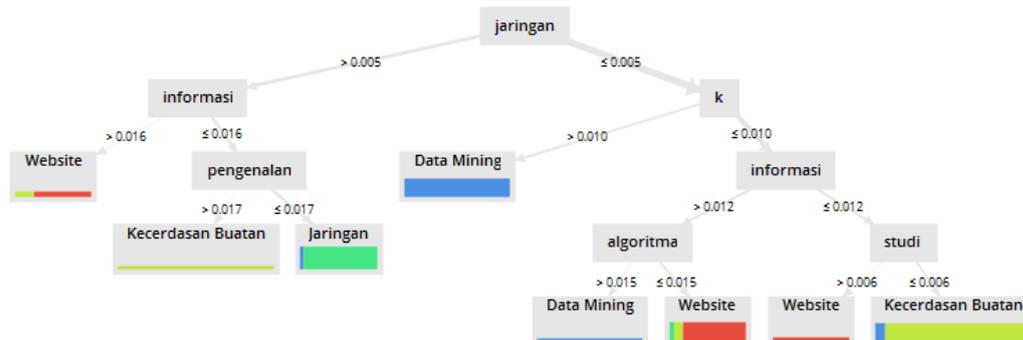
Pada evaluasi akurasi ini untuk mengetahui keakuratan pada setiap model dan mengetahui nilai akurasi yang paling tinggi dengan membandingkan akurasi pada model *naive bayes*. Berdasarkan hasil akurasi dengan menggunakan model *naive bayes* itu menunjukkan tingkat akurasinya 82,50% yang artinya model klasifikasi tema tugas akhir menggunakan *naive bayes* memiliki tingkat keakuratan yang tinggi sehingga hasilnya dapat memprediksi cukup akurat. Lalu, pada Class precision pred Data Mining sebesar 80,00%, Class Precision pred Jaringan sebesar 100,00%, Class Precision pred Kecerdasan Buatan sebesar 76,19%, dan Class Precision pred Website sebesar 76,19%. Sedangkan pada Class recall true Data Mining sebesar 80,00%, Class recall true Jaringan sebesar 90,00%, Class recall true Kecerdasan Buatan sebesar 80,00%, dan Class recall true Website sebesar 80,00%. Hasil akurasi dengan menggunakan *decision tree* sebesar 67.50% seperti terlihat pada gambar 10 berikut:

accuracy: 67.50% +/- 17.85% (mikro: 67.50%)

	true Data Mining	true Jaringan	true Kecerdasan Bu...	true Website	class precision
pred. Data Mining	13	0	0	0	100.00%
pred. Jaringan	1	18	2	2	78.26%
pred. Kecerdasan B...	1	1	12	7	57.14%
pred. Website	5	1	6	11	47.83%
class recall	65.00%	90.00%	60.00%	55.00%	

Gambar 10. Hasil Akurasi Model *Decision Tree*

Pada evaluasi akurasi ini untuk mengetahui keakuratan pada model dan mengetahui nilai akurasi yang paling tinggi dengan membandingkan akurasi pada model *decision tree*. Berdasarkan hasil akurasi dengan menggunakan model *decision tree* itu menunjukkan tingkat akurasinya 67,50% yang artinya model klasifikasi tema tugas akhir menggunakan *decision tree* lebih rendah daripada dengan *naive bayes*, tetapi *decision tree* juga memiliki tingkat keakuratan yang tinggi sehingga hasilnya dapat memprediksi cukup akurat. Pada Class precision pred Data Mining sebesar 100,00%, Class Precision pred Jaringan sebesar 78,26%, Class Precision pred Kecerdasan Buatan sebesar 57,14%, dan Class Precision pred Website sebesar 47,83%. Sedangkan pada Class recall true Data Mining sebesar 65,00%, Class recall true Jaringan sebesar 90,00%, Class recall true Kecerdasan Buatan sebesar 60,00%, dan Class recall true Website sebesar 55,00%.



Gambar 11. Hasil Grafik Model *Decision Tree*

Dapat kita lihat pada gambar 11 *decision tree* di atas, atribut yang terbukti prediktif menggunakan accuracy merupakan predictor yang utama ketika menggunakan algoritma *information_gain* sebagai parameter untuk pohon keputusan menjadi sangat sederhana, menghasilkan sedikit node dan leaf. Berdasarkan hasil model *decision tree* diatas dapat diketahui sebagai berikut :

- 1) Jika jaringan lebih dari 0.005 maka yang dipilih tema tugas akhir adalah informasi, tetapi jika kurang dari 0.005 maka yang dipilih adalah klasifikasi.
- 2) Jika Informasi yang dipilih lebih besar dari 0.016 maka yang akan digunakan adalah website tetapi jika kurang dari 0.016 maka dipilih pengenalan.
- 3) Jika kata pengenalan yang banyak digunakan sebesar 0.017 maka akan dipilih tema kecerdasan buatan jika kurang dari 0.017 maka akan dipilih tema jaringan.
- 4) Jika klasifikasi lebih dari 0.010 maka akan dipilih tema data mining tetapi jika kurang dari 0.010 maka akan dipilih tema informasi
- 5) Hal lain menunjukkan jika informasi lebih dari algoritma sebesar 0.012 maka akan dipilih tema data mining dan website tetapi jika kurang dari 0.012 maka akan dipilih tema website dan kecerdasan buatan.

Adapun penjelasan dapat dilihat dari deskripsi grafik dibawah ini:

```

jaringan > 0.005
| informasi > 0.016: Website {Data Mining=0, Jaringan=0, Kecerdasan Buatan=1, Website=3}
| informasi ≤ 0.016
| | pengenalan > 0.017: Kecerdasan Buatan {Data Mining=0, Jaringan=0, Kecerdasan Buatan=2, Website=0}
| | pengenalan ≤ 0.017: Jaringan {Data Mining=1, Jaringan=19, Kecerdasan Buatan=0, Website=0}
jaringan ≤ 0.005
| k > 0.010: Data Mining {Data Mining=16, Jaringan=0, Kecerdasan Buatan=0, Website=0}
| k ≤ 0.010
| | informasi > 0.012
| | | algoritma > 0.015: Data Mining {Data Mining=2, Jaringan=0, Kecerdasan Buatan=0, Website=0}
| | | algoritma ≤ 0.015: Website {Data Mining=0, Jaringan=1, Kecerdasan Buatan=2, Website=14}
| | informasi ≤ 0.012
| | | studi > 0.006: Website {Data Mining=0, Jaringan=0, Kecerdasan Buatan=0, Website=3}
| | | studi ≤ 0.006: Kecerdasan Buatan {Data Mining=1, Jaringan=0, Kecerdasan Buatan=15, Website=0}
    
```

classification_error: 17.50% +/- 10.00% (mikro: 17.50%)

	true Data Mining	true Jaringan	true Kecerdasan Bu...	true Website	class precision
pred. Data Mining	16	0	2	2	80.00%
pred. Jaringan	0	18	0	0	100.00%
pred. Kecerdasan B...	3	0	16	2	76.19%
pred. Website	1	2	2	16	76.19%
class recall	80.00%	90.00%	80.00%	80.00%	

Gambar 12. Hasil Klasifikasi Error Model *Naïve bayes*

Berdasarkan hasil klasifikasi error seperti ditunjukkan pada gambar 12 dengan menggunakan model *naïve bayes* itu menunjukkan tingkat klasifikasi error sebesar 17,50% yang artinya model klasifikasi tema tugas akhir menggunakan *naïve bayes* memiliki tingkat klasifikasi error lebih rendah sehingga hasilnya dapat memprediksi cukup akurat. Class precision pred Data Mining sebesar 80,00%, Class Precision pred Jaringan sebesar 100,00%, Class Precision pred Kecerdasan Buatan sebesar 76,19%, dan Class Precision pred Website sebesar 76,19%. Sedangkan pada Class recall true Data Mining sebesar 80,00%, Class recall true Jaringan sebesar 90,00%, Class recall true Kecerdasan Buatan sebesar 80,00%, dan Class recall true Website sebesar 80,00%.

classification_error: 32.50% +/- 17.85% (mikro: 32.50%)

	true Data Mining	true Jaringan	true Kecerdasan Bu...	true Website	class precision
pred. Data Mining	13	0	0	0	100.00%
pred. Jaringan	1	18	2	2	78.26%
pred. Kecerdasan B...	1	1	12	7	57.14%
pred. Website	5	1	6	11	47.83%
class recall	65.00%	90.00%	60.00%	55.00%	

Gambar 13. Hasil Klasifikasi Error Model Decision Tree

Sedangkan klasifikasi error menggunakan algoritma decision tree memiliki tingkat error sebesar 32,50% lebih tinggi daripada menggunakan *naive bayes*, tetapi decision tree juga memiliki tingkat keakuratan yang tinggi sehingga hasilnya dapat memprediksi cukup akurat. Pada Class precision pred Data Mining sebesar 100,00%, Class Precision pred Jaringan sebesar 78,26%, Class Precision pred Kecerdasan Buatan sebesar 57,14%, dan Class Precision pred Website sebesar 47,83%. Sedangkan pada Class recall true Data Mining sebesar 65,00%, Class recall true Jaringan sebesar 90,00%, Class recall true Kecerdasan Buatan sebesar 60,00%, dan Class recall true Website sebesar 55,00%.

3.2 Pembahasan

Hasil penelitian menunjukkan bahwa Naive Bayes memiliki kinerja yang lebih baik dengan akurasi sebesar 82,50%, sedangkan Decision Tree mencapai akurasi 67,50%. Hasil ini sejalan dengan penelitian yang dilakukan oleh Jimly Hanif *et al.* (2023), yang menunjukkan bahwa pendekatan probabilistik seperti Naive Bayes efektif dalam mengklasifikasikan judul tugas akhir melalui Natural Language Processing (NLP). Naive Bayes, dengan prinsip independensi antar fitur, menyederhanakan perhitungan dan menghasilkan prediksi yang akurat meskipun dengan asumsi sederhana. Di sisi lain, Decision Tree, meskipun mudah diinterpretasikan, menunjukkan performa yang lebih rendah. Salah satu kekurangan Decision Tree adalah risiko overfitting, terutama ketika model menjadi terlalu spesifik dengan menghasilkan terlalu banyak node dan cabang. Hal ini juga diungkapkan oleh Wibiyanto dan Wibowo (2023), yang menyebutkan bahwa model Decision Tree cenderung overfitting pada dataset yang kompleks, seperti pada klasifikasi tema tugas akhir. Akibatnya, tingkat kesalahan klasifikasi dengan Decision Tree lebih tinggi dibandingkan dengan Naive Bayes, dengan error sebesar 32,50%. Yesisca *et al.* (2022) juga menemukan bahwa metode klasifikasi berbasis pembelajaran mesin yang lebih akurat, seperti Support Vector Machine (SVM), cenderung lebih unggul dalam tugas klasifikasi dibandingkan metode lain, termasuk Decision Tree. Meskipun penelitian ini menggunakan Naive Bayes alih-alih SVM, hasil menunjukkan bahwa pendekatan probabilistik seperti Naive Bayes lebih efektif dalam menangani variasi data yang ada pada tugas akhir. Selain itu, penelitian oleh Genius Zentrato *et al.* (2022) menggarisbawahi bahwa metode yang lebih sederhana, seperti Naive Bayes atau K-Means, tetap memberikan hasil yang baik ketika digunakan pada dataset yang beragam. Dalam penerapan algoritma untuk proses akademik, penelitian Salam *et al.* (2022) tentang sistem rekomendasi tugas akhir menekankan pentingnya algoritma yang efisien dalam mendukung pengambilan keputusan. Naive Bayes, dengan tingkat akurasi yang lebih tinggi, menjadi metode yang lebih diandalkan dalam klasifikasi tema tugas akhir, mengingat kecepatan dan kemudahan penerapannya dibandingkan Decision Tree. Dengan mempertimbangkan hasil dan penelitian terkait, Naive Bayes lebih unggul dibandingkan Decision Tree dalam klasifikasi tema tugas akhir. Akurasi yang lebih tinggi dan kesederhanaan dalam penerapannya menjadikan Naive Bayes sebagai metode yang lebih tepat digunakan dalam klasifikasi ini. Meskipun Decision Tree masih berguna untuk interpretasi visual, Naive Bayes lebih unggul dalam hal akurasi dan efisiensi.

4. Kesimpulan dan Saran

Berdasarkan hasil analisis dalam penelitian ini mengenai optimasi Support Vector Machine (SVM) dengan menggunakan algoritma Naive Bayes dan Decision Tree untuk klasifikasi tema tugas akhir mahasiswa, beberapa kesimpulan dapat diambil. Pertama, hasil analisis menunjukkan bahwa baik model Naive Bayes maupun Decision Tree dapat digunakan oleh pengambil kebijakan dalam proses klasifikasi tugas akhir. Model ini dapat membantu dalam penentuan tema tugas akhir secara lebih akurat dan efisien, mendukung keputusan akademik yang lebih baik. Kedua, penerapan algoritma Naive Bayes dan Decision Tree berhasil meningkatkan tingkat akurasi SVM dalam proses klasifikasi tema tugas akhir. Hal ini menunjukkan bahwa kedua model ini berfungsi sebagai alat optimalisasi yang efektif. Ketiga, Naive Bayes terbukti memiliki akurasi tertinggi, yaitu sebesar 82,50%, dibandingkan dengan Decision Tree yang memiliki akurasi 67,50%. Dengan demikian, Naive Bayes lebih direkomendasikan untuk digunakan dalam klasifikasi tema tugas akhir karena tingkat akurasinya yang lebih tinggi.

Beberapa saran yang dapat diberikan berdasarkan hasil penelitian ini adalah sebagai berikut. Pertama, dalam penerapan algoritma Naive Bayes dan Decision Tree, pengaturan parameter yang lebih bervariasi sangat dianjurkan untuk mendapatkan hasil yang lebih optimal. Dengan pengaturan parameter yang tepat, akurasi dari model klasifikasi dapat ditingkatkan lebih lanjut, sehingga menghasilkan model

yang lebih presisi dan sesuai dengan kebutuhan klasifikasi data tugas akhir. Kedua, untuk penelitian di masa depan, model yang telah diterapkan dalam studi ini sebaiknya diimplementasikan dalam bentuk aplikasi berbasis web atau desktop. Penggunaan bahasa pemrograman tertentu, seperti Python atau Java, dapat mempermudah pengembangan aplikasi yang akan memproses data tugas akhir secara otomatis dan efisien. Dengan aplikasi yang terintegrasi, proses pengelompokan dan klasifikasi tema tugas akhir akan semakin mudah dan dapat langsung diaplikasikan dalam skala institusi pendidikan.

Referensi

- Afkar, M. K., Wali, M., & Imilda. (2024). Aplikasi prediksi produksi cabai dengan algoritma C.45 untuk dinas pertanian provinsi Aceh berbasis web. *Jurnal Ilmu Komputer Dan Teknologi Informasi*, 1(1), 1–13. <https://doi.org/10.35870/jikti.v1i1.732>
- Amini, T. A., & Setiawan, K. (2024). Application of the Naive Bayes algorithm in Twitter sentiment analysis of 2024 vice presidential candidate Gibran Rakabuming Raka using Rapidminer. *International Journal Software Engineering and Computer Science (IJSECS)*, 4(1), 234–246. <https://doi.org/10.35870/ijsecs.v4i1.2236>
- Apriansyah, F. A., Hermawan, A., & Avianto, D. (2024). Optimization of K value in KNN algorithm for spam and HAM classification in SMS texts. *International Journal Software Engineering and Computer Science (IJSECS)*, 4(2), 767–779. <https://doi.org/10.35870/ijsecs.v4i2.2681>
- Azis, I. M. A., & Wahyudi, T. (2024). Analisa sentimen rencana pemindahan ibu kota Nusantara dari Jakarta ke Kalimantan Timur menggunakan algoritma Naïve Bayes. *Jurnal Indonesia : Manajemen Informatika Dan Komunikasi*, 5(3), 2668-2679. <https://doi.org/10.35870/jimik.v5i3.969>
- Azzahra, Y. A., & Akbar, Y. (2024). Komparasi penerapan algoritma C4.5 dan Naïve Bayes untuk ketepatan waktu pengiriman barang pada PT. Rtrans Logistik Artamandiri. *Jurnal Indonesia : Manajemen Informatika Dan Komunikasi*, 5(3), 2768-2780. <https://doi.org/10.35870/jimik.v5i3.1003>
- Baidowi, A., & Sutisna. (2024). Implementasi data mining klasifikasi fuel surcharge menggunakan algoritma Naive Bayes studi kasus PT Pelabuhan Indonesia (Persero) Regional 2 Tanjung Priok. *Jurnal Indonesia : Manajemen Informatika Dan Komunikasi*, 5(3), 2854-2863. <https://doi.org/10.35870/jimik.v5i3.960>
- Efendi, A. N., Triayudi, A., & Winarsih, W. (2022). Sistem pakar mendeteksi penyakit ikan arwana Asia menggunakan metode Naïve Bayes. *Jurnal JTIC (Jurnal Teknologi Informasi Dan Komunikasi)*, 6(2), 243–249. <https://doi.org/10.35870/jtik.v6i2.415>
- Efendi, R., Fauziah, F., & Gunaryati, A. (2021). Diagnosa penyakit tanaman cabai menggunakan metode forward chaining dan Naïve Bayes. *Jurnal JTIC (Jurnal Teknologi Informasi Dan Komunikasi)*, 5(2), 164–172. <https://doi.org/10.35870/jtik.v5i2.208>
- Fahmi, H., & Sutisna. (2024). Implementasi data mining klasifikasi gejala penyakit TB menggunakan algoritma Naive Bayes pada studi kasus Puskesmas Pegangsaan Dua B. *Jurnal Indonesia : Manajemen Informatika Dan Komunikasi*, 5(3), 2888-2898. <https://doi.org/10.35870/jimik.v5i3.970>
- Farhani, A., & Sutisna. (2024). Analisis sentimen terhadap kendaraan listrik di Indonesia menggunakan metode klasifikasi Naïve Bayes. *Jurnal Indonesia : Manajemen Informatika Dan Komunikasi*, 5(3), 2680-2690. <https://doi.org/10.35870/jimik.v5i3.983>
- Fatmasari, R., Widodo, A. Z. P., Hakim, V. F., Gata, W., & Saputra, D. D. (2023). Pengkategorian komentar Instagram terhadap layanan akademik dan non-akademik Universitas Terbuka. *Jurnal JTIC (Jurnal Teknologi Informasi Dan Komunikasi)*, 7(1), 59–70. <https://doi.org/10.35870/jtik.v7i1.669>

- Firmansyah, D. R., & Lestariningsih, E. (2024). Analisis sentimen ulasan aplikasi Smart Campus Unisbank di Google Playstore menggunakan algoritma Naive Bayes. *Jurnal JTIK (Jurnal Teknologi Informasi Dan Komunikasi)*, 8(2), 498–507. <https://doi.org/10.35870/jtik.v8i2.1882>
- Fuad, M., Wattimena, F. Y., Rizani, A., & Yuswardi. (2023). Investment decision making in digital business using Tsukamoto fuzzy logic. *International Journal Software Engineering and Computer Science (IJSECS)*, 3(2), 144–150. <https://doi.org/10.35870/ijsecs.v3i2.1525>
- Hafid, I., Gata, W., Hilyati, K., Hakim, V. F., & Rahayu, S. (2023). Sentimen analisis masyarakat Indonesia terhadap presiden Rusia pada komentar media berita online. *Jurnal JTIK (Jurnal Teknologi Informasi Dan Komunikasi)*, 7(1), 172–178. <https://doi.org/10.35870/jtik.v7i1.698>
- Hanif, J., Farid, M. N., & Hasanah, B. (2023). Penerapan Natural Language Processing untuk klasifikasi bidang minat berdasarkan judul tugas akhir. *Jurnal Sistim Informasi Dan Teknologi*, 5(1), 41–49. <https://doi.org/10.37034/jsisfotek.v5i1.196>
- Jaya, D. R. P., & Lestari, S. (2024). Analisis sentimen naturalisasi tim nasional Indonesia U-23 di era Shin Tae-yong menggunakan algoritma Naive Bayes dan K-Nearest Neighbors. *Jurnal Indonesia : Manajemen Informatika Dan Komunikasi*, 5(3), 3262-3277. <https://doi.org/10.35870/jimik.v5i3.1024>
- Julianto, M. R., Akbar, Y., & Wahyudi, T. (2024). Analisis sentimen respon publik terhadap program internet gratis di platform X melalui pendekatan algoritma Naive Bayes. *Jurnal Indonesia : Manajemen Informatika Dan Komunikasi*, 5(3), 2940-2950. <https://doi.org/10.35870/jimik.v5i3.981>
- Legito, R., Nuraini, R., Judijanto, L., & Lubis, A. I. (2023). The application of *Convolutional Neural Networks* in floristic recognition. *International Journal Software Engineering and Computer Science (IJSECS)*, 3(3), 520–528. <https://doi.org/10.35870/ijsecs.v3i3.1827>
- Lestari, D., & Lestari, S. (2024). Penerapan data mining klasifikasi tingkat pemahaman siswa pada kegiatan belajar mengajar dengan metode decision tree (Studi kasus SDN Malaka Jaya 11 Duren Sawit). *Jurnal Indonesia: Manajemen Informatika Dan Komunikasi*, 5(2), 1260–1268. <https://doi.org/10.35870/jimik.v5i2.662>
- Lestari, S. A., & Sugiyono. (2024). Analisis sentimen pada media sosial X (Twitter) terhadap tumor jinak payudara menggunakan metode Naive Bayes. *Jurnal Indonesia : Manajemen Informatika Dan Komunikasi*, 5(3), 3336-3348. <https://doi.org/10.35870/jimik.v5i3.1015>
- Limanauw, G. C. T., & Oetama, R. (2024). Analysis of the millennial view on insurance as a key financial priority using K-Means and Decision Tree. *Jurnal JTIK (Jurnal Teknologi Informasi Dan Komunikasi)*, 8(1), 202–209. <https://doi.org/10.35870/jtik.v8i1.1445>
- Mardiani, E., Rahmansyah, N., Ningsih, S., Lantana, D. A., Wulandana, N. P., Lombu, A. A., & Budyarti, S. (2024). Classification of potential tsunami disaster due to earthquakes in Indonesia based on machine learning. *International Journal Software Engineering and Computer Science (IJSECS)*, 4(1), 13–23. <https://doi.org/10.35870/ijsecs.v4i1.2084>
- Nofiar Am, A., & Nasari, F. (2023). Android smartphone damage diagnosis expert system by web-based forward chain method. *International Journal Software Engineering and Computer Science (IJSECS)*, 3(1), 11–18. <https://doi.org/10.35870/ijsecs.v3i1.1019>
- Nugraha, D., & Said, F. (2024). Implementasi algoritma C4.5 dan Naive Bayes untuk analisis sentimen publik terhadap platform live streaming Dukov di Indonesia. *Jurnal Indonesia : Manajemen Informatika Dan Komunikasi*, 5(3), 3326-3335. <https://doi.org/10.35870/jimik.v5i3.1011>
- Pasaribu, N. A., & Sriani. (2023). The Shopee application user reviews sentiment analysis employing Naive Bayes algorithm. *International Journal Software Engineering and Computer Science (IJSECS)*, 3(3), 194–204. <https://doi.org/10.35870/ijsecs.v3i3.1699>

- Pernama, B., & Purnomo, H. D. (2023). Analisis risiko pinjaman dengan metode Support Vector Machine, Artificial Neural Network, dan Naïve Bayes. *Jurnal JTIC (Jurnal Teknologi Informasi Dan Komunikasi)*, 7(1), 92–99. <https://doi.org/10.35870/jtik.v7i1.693>
- Puteri, D. W., Buana, P. W., & Sukarsa, I. M. (2024). Komparasi metode decision tree dan deep learning dalam meramalkan jumlah mahasiswa drop out berdasarkan nilai akademik. *Journal of Internet and Software Engineering*, 1(2), 12. <https://doi.org/10.47134/pjise.v1i2.2327>
- Qubra, R., & Saputra, R. A. (2024). Classification of hoax news using the Naïve Bayes method. *International Journal Software Engineering and Computer Science (IJSECS)*, 4(1), 40–48. <https://doi.org/10.35870/ijsecs.v4i1.2068>
- Salam, A., Albahri, F. P., & Fathurrahmad. (2022). Sistem rekomendasi tugas akhir mahasiswa pada AMIK Indonesia untuk mendukung Merdeka Belajar-Kampus Merdeka menggunakan metode Collaborative Filtering (CF). *Jurnal JTIC (Jurnal Teknologi Informasi Dan Komunikasi)*, 6(2), 281–288. <https://doi.org/10.35870/jtik.v6i2.420>
- Samantri, M., & Afiyati. (2024). Perbandingan algoritma Support Vector Machine dan *Random Forest* untuk analisis sentimen terhadap kebijakan pemerintah Indonesia terkait kenaikan harga BBM tahun 2022. *Jurnal JTIC (Jurnal Teknologi Informasi Dan Komunikasi)*, 8(1), 1–9. <https://doi.org/10.35870/jtik.v8i1.1202>
- Sari Siregar, Y., Handoko, D., Khairani, M., Syahputri, N. I., & Harahap, H. (2024). Implementasi data mining klasifikasi algoritma Chaid dalam menentukan pola penerima mahasiswa baru. *Digital Transformation Technology*, 3(2), 978–989. <https://doi.org/10.47709/digitech.v3i2.3612>
- Setiawan, K., & Saputry, Y. Y. A. (2024). Clustering data calon siswa baru menggunakan metode K-Means di pusat pengembangan anak Fajar Baru Cengkareng. *Jurnal JTIC (Jurnal Teknologi Informasi Dan Komunikasi)*, 8(1), 75–83. <https://doi.org/10.35870/jtik.v8i1.1426>
- Sugiyono, Haryati, Sarimole, F. M., & Tundo. (2024). Data mining modeling using the K-Means algorithm to analyze the impact of new media on early childhood psychology at Bimba Rainbow Kids Sukmajaya. *International Journal Software Engineering and Computer Science (IJSECS)*, 4(2), 647–655. <https://doi.org/10.35870/ijsecs.v4i2.2874>
- Supendi, Kumala, D., & Yulianti, M. L. (2024). Implications of deep learning for stock market forecasting. *International Journal Software Engineering and Computer Science (IJSECS)*, 4(1), 68–80. <https://doi.org/10.35870/ijsecs.v4i1.2281>
- Syah, F., Fajrin, H., Afif, A. N., Saeputra, M. R., Mirranty, D., & Saputra, D. D. (2023). Analisa sentimen terhadap Twitter IndihomeCare menggunakan perbandingan algoritma Smote, Support Vector Machine, AdaBoost dan Particle Swarm Optimization. *Jurnal JTIC (Jurnal Teknologi Informasi Dan Komunikasi)*, 7(1), 53–58. <https://doi.org/10.35870/jtik.v7i1.686>
- Tundo, & Rachmawati, D. N. (2024). Implementasi algoritma Naive Bayes untuk analisis sentimen terhadap program makan siang gratis. *Jurnal Indonesia: Manajemen Informatika Dan Komunikasi*, 5(3), 2925–2939. <https://doi.org/10.35870/jimik.v5i3.978>
- Tundo, Eldina, R., Setiawan, K., & Fajri, R. (2024). Sentiment analysis of cigarette use based on opinions from X using Naive Bayes and SVM. *Jurnal Indonesia : Manajemen Informatika Dan Komunikasi*, 5(3), 2561–2569. <https://doi.org/10.35870/jimik.v5i3.947>
- Tyar, F., & Wahyuddin, M. I. (2022). Sistem pakar menggunakan metode Naïve Bayes dan certainty factor untuk mendeteksi hama pada tanaman alpukat berbasis web. *Jurnal JTIC (Jurnal Teknologi Informasi Dan Komunikasi)*, 6(4), 488–496. <https://doi.org/10.35870/jtik.v6i4.519>

- Wibiyanto, A., & Wibowo, A. (2023). Penerapan algoritma Multiclass Support Vector Machine dan TF-IDF untuk klasifikasi topik tugas akhir. *SKANIKA: Sistem Komputer dan Teknik Informatika*, 6(1), 42–50.
- Widyawati, D., Faradibah, A., & Belluano, L. L. (2023). Comparison analysis of classification model performance in lung cancer prediction using Decision Tree, Naive Bayes, and Support Vector Machine. *Indonesian Journal of Data and Science*, 4(2), 78-86. <https://doi.org/10.56705/ijodas.v4i2.76>
- Yansyah, H., Fauziah, S., & Maulana, D. (2023). Classification of production machine spare part stock data request needs using the K-Nearest Neighbor method. *International Journal Software Engineering and Computer Science (IJSECS)*, 3(3), 457–466. <https://doi.org/10.35870/ijsecs.v3i3.1878>
- Yesisca, F., Ratnawati, D. E., & Rahayudi, B. (2022). Analisis perbandingan klasifikasi topik skripsi mahasiswa menggunakan K-Nearest Neighbor dan Support Vector Machine (Studi Kasus: Jurusan Sistem Informasi, Fakultas Ilmu Komputer, Universitas Brawijaya). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 6(5), 2328–2335. Diambil dari <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/11045>
- Yumansyah, Q., Fatchan, M., & Turmudi Zy, A. (2023). Design and development of an information system for indemnity claim box recapitulation using SDLC method at Mandiri Inhealth Insurance. *International Journal Software Engineering and Computer Science (IJSECS)*, 3(3), 553–559. <https://doi.org/10.35870/ijsecs.v3i3.1970>
- Zendrato, G. F. S., Triayudi, A., & E, E. T. (2022). Analisis clustering dokumen tugas akhir mahasiswa sistem informasi Universitas Nasional menggunakan metode K-Means clustering. *Jurnal JTik (Jurnal Teknologi Informasi Dan Komunikasi)*, 6(1), 70–76. <https://doi.org/10.35870/jtik.v6i1.389>.

How Cites

Arfika. (2024). Optimasi Support Vector Machine (SVM) Menggunakan Naive Bayes dan Decision Tree untuk Klasifikasi Tema Tugas Akhir Mahasiswa Sistem Informasi. *Computer Journal*, 2(2), 91–104. <https://doi.org/10.58477/cj.v2i2.179>.

Publisher's Note

Yayasan Pendidikan Mitra Mandiri Aceh (YPPMA) remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Submit your manuscript to YPMMA Journal and benefit from: <https://journal.ypmma.org/index.php/cj>.