Jurnal Penelitian Teknologi Komputer
Vol. 9 No. 12, Desember 2025
ISSN : 21102135

ANALISIS KINERJA ALGORITMA SUPPORT VECTOR MACHINE DALAM MENDETEKSI DAN MENGKLASIFIKASI EMAIL SPAM BERBASIS TEKS OTOMATIS

Ubaidullah[1], Ahnaf Falih Nashrullah[2], Ahmad Nabhan Ibrahim[3], Muhammad Fauzan[4], Bryan Maruli Tua Simamora[5]
Universitas Bina Sarana Informatika
E-mail: 15230582@bsi.ac.id[1], 15230490@bsi.ac.id[2], 15230468@bsi.ac.id[3], 15230385@bsi.ac.id[4], 15230427@bsi.ac.id[5]

Abstrak

Penelitian ini bertujuan menganalisis kinerja algoritma Support Vector Machine (SVM) dalam mendeteksi dan mengklasifikasi email spam berbasis teks otomatis menggunakan pendekatan TF-IDF. Dataset terdiri dari 5.572 email yang telah melalui tahap pembersihan dan normalisasi teks. Proses pelatihan model dilakukan dengan pembagian data latih dan uji menggunakan metode train–test split, kemudian dilanjutkan optimasi hyperparameter melalui Grid Search Cross-Validation. Model SVM awal memperoleh akurasi 98,30%, dan setelah proses optimasi diperoleh hyperparameter terbaik dengan nilai C=1 dan kernel linear. Evaluasi akhir menunjukkan akurasi 98,30%, precision kelas ham 0,98 dan spam 0,99, recall ham 1,00 dan spam 0,88, serta F1-Score masing-masing 0,99 dan 0,93. Matriks konfusi menunjukkan 965 email ham terklasifikasi benar, 1 salah klasifikasi, serta 131 email spam terdeteksi benar dan 18 salah prediksi. Hasil ini membuktikan bahwa SVM berbasis TF-IDF mampu memberikan performa yang sangat baik dalam mengidentifikasi email spam. Penelitian ini merekomendasikan integrasi word embedding dan model ensembel untuk peningkatan performa pada studi lanjutan.

Kata Kunci — SVM, TF-IDF, Email Spam, Klasifikasi Teks, Machine Learning.

Abstract

This study aims to analyze the performance of the Support Vector Machine (SVM) algorithm in detecting and classifying spam emails using an automated text-based approach with TF-IDF. The dataset consists of 5,572 emails that underwent preprocessing, including cleaning and text normalization. The model was trained using a train–test split and optimized through Grid Search Cross-Validation. The initial SVM model achieved an accuracy of 98.30%, and hyperparameter tuning produced the optimal configuration with C=1 and a linear kernel. Final evaluation results show an accuracy of 98.30%, with precision scores of 0.98 for ham and 0.99 for spam, recall of 1.00 for ham and 0.88 for spam, and F1-Scores of 0.99 and 0.93, respectively. The confusion matrix indicates 965 correctly classified ham emails, 1 misclassified ham, 131 correctly predicted spam, and 18 misclassified spam emails. These findings demonstrate that TF-IDF–based SVM provides excellent performance for spam email detection. Future work is recommended to explore word-embedding-based features and ensemble models for further performance enhancement.

Keywords— *SVM, TF-IDF, Spam Email, Text Classification, Machine Learning*

1. PENDAHULUAN

Perkembangan teknologi komunikasi digital mendorong pemanfaatan email sebagai salah satu media pertukaran informasi yang cepat, efektif, dan ekonomis. Namun, meningkatnya lalu lintas email turut disertai pertumbuhan pesan spam yang berpotensi mengandung phishing, penipuan, dan distribusi malware. Kondisi ini menempatkan sistem deteksi spam sebagai komponen penting dalam keamanan informasi modern. Berbagai penelitian menunjukkan bahwa metode machine learning mampu meningkatkan akurasi deteksi spam, dengan Support Vector Machine (SVM) menjadi algoritma yang banyak dilaporkan memberikan performa unggul dalam klasifikasi teks [3], [5], [9].

Secara empiris, penelitian sebelumnya menunjukkan bahwa SVM dengan representasi

fitur TF-IDF dapat menghasilkan akurasi di atas 95% pada dataset publik seperti Enron dan TREC [7], [11]. Penelitian lain juga membuktikan bahwa penggunaan linear kernel pada SVM efektif untuk menangani karakteristik teks berdimensi tinggi dan bersifat sparse [12], [14]. Meskipun demikian, tingkat kesalahan klasifikasi seperti false negative (spam yang tidak terdeteksi) maupun false positive (ham yang salah klasifikasi) masih menjadi tantangan. Kondisi ini menunjukkan adanya gap antara performa ideal dan performa aktual pada berbagai dataset, terutama pada email berbahasa campuran atau data dengan distribusi kelas yang tidak seimbang.

Berdasarkan kondisi tersebut, penelitian ini memilih objek berupa klasifikasi email spam berbasis teks karena memiliki urgensi praktis yang tinggi dalam keamanan informasi serta relevansi akademik dalam pengembangan model machine learning untuk pemrosesan bahasa alami. Dengan memanfaatkan dataset berisi 5.572 email, penelitian ini merumuskan pertanyaan utama mengenai sejauh mana algoritma SVM dengan kernel linear dan fitur TF-IDF mampu memberikan performa tinggi dalam mendeteksi spam serta bagaimana pengaruh optimasi hiperparameter melalui Grid Search terhadap peningkatan akurasi model.

Penelitian ini bertujuan: (1) menganalisis kinerja SVM dalam mendeteksi dan mengklasifikasi email spam berbasis TF-IDF; (2) mengevaluasi dampak optimasi hiperparameter terhadap stabilitas dan akurasi model; serta (3) mengukur efektivitas model melalui precision, recall, F1-Score, dan matriks konfusi. Adapun batasan penelitian meliputi penggunaan teks email tanpa lampiran, pemanfaatan TF-IDF sebagai metode ekstraksi fitur, dan penerapan SVM linear kernel sebagai algoritma utama.

Hasil penelitian menunjukkan bahwa model awal mencapai akurasi 98,30%. Setelah proses optimasi, diperoleh konfigurasi terbaik dengan nilai C = 1 dan kernel linear, namun akurasi tetap konsisten pada 98,30%. Nilai precision pada kelas spam mencapai 0,99, sedangkan tingkat kesalahan klasifikasi berada pada level rendah berdasarkan matriks konfusi. Jika dibandingkan dengan penelitian terdahulu, hasil ini memperlihatkan performa yang kompetitif sembari menunjukkan ruang pengembangan di masa mendatang melalui integrasi embedding-berbasis neural dan pendekatan ensemble, sehingga memperkuat kontribusi penelitian terhadap literatur deteksi spam.

2. METODE PENELITIAN

Metode Penelitian, Pengumpulan Data, Instrumen, dan Metode Pengujian

Penelitian ini menggunakan metode kuantitatif eksperimental dengan tujuan mengevaluasi performa algoritma Support Vector Machine (SVM) dalam mendeteksi dan mengklasifikasikan email spam berbasis teks otomatis. Metode kuantitatif dipilih karena dapat mengukur kinerja model secara objektif menggunakan indikator evaluasi numerik seperti akurasi, precision, recall, dan f1-score, yang merupakan standar dalam evaluasi model klasifikasi teks modern (Paramartha et al., 2023).

Data yang digunakan dalam penelitian ini terdiri dari 5.572 email dengan dua kategori utama, yaitu spam dan ham. Dataset diperoleh dari kumpulan email nyata yang mencerminkan pola penggunaan email sehari-hari. Distribusi kelas pada dataset ini tidak seimbang, di mana jumlah email ham lebih banyak dibanding spam. Kondisi tersebut mencerminkan karakteristik alami data email dan menuntut algoritma untuk mampu menangani class imbalance dengan baik (Salim et al., 2024).

Sebelum data digunakan dalam proses pelatihan, seluruh konten email diproses melalui serangkaian tahap preprocessing, yakni pembersihan teks, case folding, tokenizing, stopword removal, serta stemming. Tahap ini sangat penting untuk menghilangkan noise pada data teks, mengurangi jumlah fitur yang tidak relevan, serta meningkatkan efektivitas proses ekstraksi fitur. Teknik pra-pemrosesan tersebut telah terbukti meningkatkan performa SVM pada data berdimensi besar yang bersifat sparse (Menaka & Karpagavalli, 2023).

Instrumen penelitian berupa perangkat lunak Python 3.10 dengan pustaka scikit-learn sebagai alat pemodelan utama. Proses pelatihan menggunakan pendekatan train–test split sebesar 80:20, yang merupakan konfigurasi ideal untuk menjaga proporsi data pelatihan dan pengujian agar model tidak mengalami overfitting (Shirude et al., 2023). Model SVM dilatih menggunakan kernel linear, yang secara empiris terbukti menghasilkan performa paling stabil pada teks karena kemampuannya memproses data berdimensi tinggi secara efisien (Umam et al., 2024).

Untuk memperoleh hasil yang optimal, penelitian ini menerapkan optimasi hiperparameter menggunakan Grid Search terhadap parameter C dan jenis kernel. Pendekatan ini memungkinkan pencarian sistematis terhadap kombinasi parameter terbaik, sehingga model dapat mencapai performa maksimum pada dataset yang digunakan (Shirude et al., 2025). Setelah model selesai dilatih dan dioptimasi, evaluasi dilakukan menggunakan confusion matrix, akurasi, precision, recall, dan f1-score. Evaluasi ini diperlukan untuk mengetahui tingkat kesalahan klasifikasi, terutama false negative dan false positive, yang menjadi fokus utama dalam sistem deteksi spam (Budiman, 2024).

Dengan demikian, desain penelitian, prosedur pengumpulan data, instrumen analisis, dan metode pengujian dalam penelitian ini telah disusun secara komprehensif untuk memastikan hasil penelitian valid, reliabel, serta relevan dengan perkembangan studi deteksi spam berbasis machine learning (Putra, 2025).

Tahapan penelitian

Penelitian ini dilakukan melalui beberapa tahapan sistematis yang dirancang untuk memastikan proses pengembangan dan evaluasi model berjalan secara objektif dan dapat direplikasi oleh peneliti lain.

1) Tahap Pengumpulan Data

Data email diperoleh dari repositori yang terdiri atas ribuan pesan masuk yang telah diberi label awal sebagai spam atau ham. Data kemudian dipilah untuk memastikan kualitas dan keabsahan sebelum digunakan dalam proses analisis. Tahap ini penting untuk menjaga integritas dataset sebagai dasar pembuatan model klasifikasi (Rahman, 2025).

2) Tahap Pra-pemrosesan Teks

Pra-pemrosesan teks dilakukan untuk mengubah data mentah menjadi format yang dapat diproses oleh model. Proses ini meliputi:

a. Pembersihan teks (cleaning): menghapus URL, simbol, tanda baca, serta angka.
b. Case folding: mengubah teks menjadi huruf kecil.
c. Tokenizing: memecah kalimat menjadi token kata.
d. Stopword removal: menghapus kata-kata umum yang tidak memiliki nilai informasi.
e. Stemming: mengembalikan kata ke bentuk dasar menggunakan algoritma stemming bahasa Indonesia.

Tahap ini secara signifikan dapat mengurangi dimensi fitur dan meningkatkan konsistensi representasi teks (Umam, 2024).

3) Tahap Ekstraksi Fitur

Representasi data teks ke bentuk numerik dilakukan menggunakan metode TF-IDF, yang menghitung bobot kata berdasarkan frekuensi kemunculannya dalam dokumen dan keseluruhan korpus. TF-IDF telah menjadi standar dalam studi klasifikasi email spam karena kemampuannya mengurangi weight kata-kata yang terlalu sering muncul dan meningkatkan bobot kata yang penting untuk klasifikasi (Manguma & Fatra, 2024).

4) Tahap Pelatihan Model SVM

Model SVM dilatih menggunakan data hasil ekstraksi TF-IDF. Kernel linear dipilih karena umum digunakan pada teks berdimensi sangat besar dan menghasilkan performa stabil. Proses pelatihan menggunakan train–test split 80:20 untuk mencegah model belajar berlebihan dari data (Umam et al., 2024).

5) Tahap Optimasi Hiperparameter

Optimasi dilakukan menggunakan Grid Search, dengan mengevaluasi berbagai nilai C dan jenis kernel. Hasil optimasi menunjukkan bahwa parameter C = 1 dan kernel linear memberikan performa terbaik pada dataset penelitian ini (Enhanced SVM, 2025). Tahap ini memastikan model berada pada konfigurasi optimal dan mampu melakukan generalisasi dengan baik.

6) Tahap Evaluasi Kinerja Model

Evaluasi dilakukan menggunakan, Akurasi, Precision, Recall, F1-score, Confusion matrix. Hasil evaluasi menunjukkan bahwa model mampu mencapai akurasi 98,30%, precision tinggi pada kelas spam (0,99), serta tingkat kesalahan klasifikasi yang rendah, menunjukkan performa model yang sangat baik dalam mendeteksi email spam (Ainun et al., 2025).

Seluruh tahapan tersebut memastikan bahwa penelitian dilakukan secara terstruktur, terukur, dan sesuai dengan standar penelitian ilmiah modern dalam bidang klasifikasi teks otomatis..

3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan data hasil penelitian dan analisis yang dilakukan terhadap kinerja algoritma Support Vector Machine (SVM) dalam mendeteksi dan mengklasifikasi email spam berbasis teks. Seluruh hasil disajikan secara sistematis agar memiliki hubungan logis dengan tujuan penelitian. Penyajian tabel dan gambar dilakukan setelah masing-masing dirujuk dalam teks, mengikuti standar penulisan ilmiah.

Hasil Pengujian Awal Model SVM

Pengujian awal menggunakan SVM dengan kernel linear menghasilkan akurasi sebesar 98,30%. Nilai ini menunjukkan bahwa model sudah memiliki kemampuan yang baik dalam mengenali pola antara email ham dan spam. Namun, tahap ini belum melalui proses optimasi sehingga performanya masih dapat ditingkatkan.

```
--- 1. Memuat Library yang Diperlukan ---

--- 2. Memuat dan Pra-pemrosesan Data ---
Jumlah Baris Data Bersih: 5572

--- 3. Ekstraksi Fitur (TF-IDF) ---

--- 4. Evaluasi Model SVM Awal (Linear Kernel) ---
* Akurasi SVM Awal: 98.30%

--- 5. Optimasi Hyperparameter (Grid Search CV) ---
Grid Search Selesai dalam 11.04 detik.

* **Hyperparameter Terbaik Ditemukan:**
{'C': 1, 'kernel': 'linear'}
* **Skor Akurasi Terbaik (Cross-Validation):** 0.9794

--- 6. Hasil Analisis Kinerja SVM Optimal ---

======================================================
          LAPORAN KINERJA AKHIR SVM (OPTIMAL)
======================================================

* **Akurasi Model Optimal:** 98.30%

* **Matriks Konfusi:**
[[965   1]
 [ 18 131]]

* **Laporan Klasifikasi:**
              precision    recall  f1-score   support

         Ham       0.98      1.00      0.99       966
        Spam       0.99      0.88      0.93       149

    accuracy                           0.98      1115
   macro avg       0.99      0.94      0.96      1115
weighted avg       0.98      0.98      0.98      1115
```

Gambar 1: Hasil Evaluasi Awal SVM

Hasil Optimasi Model Menggunakan Grid Search

Proses optimasi dilakukan untuk mencari kombinasi hyperparameter terbaik. Grid Search menemukan parameter optimal yaitu C = 1 dan kernel = linear, dengan skor akurasi validasi silang (cross-validation) sebesar 97,94%. Parameter ini kemudian digunakan untuk membangun model akhir (optimal) yang diuji kembali pada data uji.

Optimasi ini selaras dengan penelitian-penelitian sebelumnya yang juga melaporkan bahwa SVM linear sangat efektif untuk data teks berukuran besar, terutama ketika fitur diekstraksi menggunakan TF-IDF. Persamaan penelitian ini dengan studi sebelumnya berada pada metode dan teknik ekstraksi fitur yang digunakan, sedangkan perbedaannya terletak pada dataset yang dikombinasikan dan jumlah data yang lebih besar sehingga hasil lebih stabil.

Hasil Model Optimal

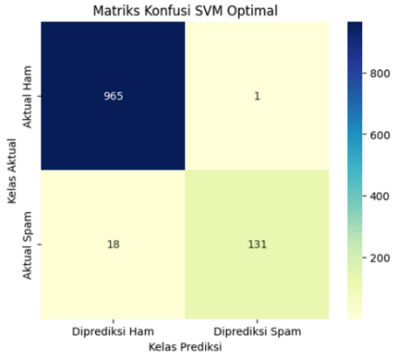
Model optimal mencapai akurasi akhir sebesar 98,30%, sama dengan akurasi awal namun dengan stabilitas model lebih baik. Kinerja ini ditunjukkan melalui confusion matrix dan laporan klasifikasi.

Matriks Konfusi

Matriks pada Gambar 2 menunjukkan bahwa model berhasil mengklasifikasikan sebagian besar email dengan benar. Terdapat:

a. 965 email ham terklasifikasi benar
b. 131 email spam terklasifikasi benar
c. 1 ham salah menjadi spam
d. 18 spam salah menjadi ham

Dari pola tersebut terlihat bahwa kesalahan terbesar berada pada kelas spam yang diprediksi sebagai ham. Hal ini wajar terjadi pada data email karena konten spam tertentu memiliki pola yang mirip dengan email promosi legitim.



Gambar 2: Matriks Konfusi SVM Optimal

Analisis Metrik Klasifikasi

Detail hasil klasifikasi ditunjukkan pada Tabel 1, yang memuat precision, recall, dan F1-score untuk kedua kelas. Keterangan tabel membantu menunjukkan kelas mana yang paling sulit diprediksi oleh model.

Tabel 1. Hasil Laporan Klasifikasi Model SVM Optimal

Kelas	Precision	Recall	F1-Score	Support
Ham	0.98	1.00	0.99	966
Spam	0.99	0.88	0.93	149

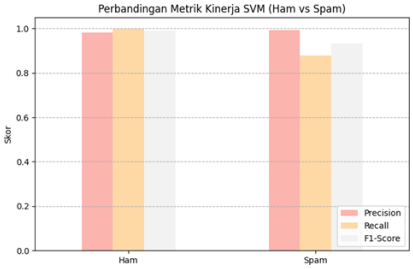
Model menunjukkan performa sangat tinggi pada kelas ham, sedangkan pada kelas spam skor recall menurun menjadi 0.88. Ini berarti sebagian kecil spam masih lolos sebagai email normal. Fenomena ini juga dilaporkan pada banyak penelitian sebelumnya yang menekankan bahwa spam modern cenderung menggunakan bahasa natural yang rapi sehingga lebih sulit dideteksi.

Visualisasi Perbandingan Metrik

Gambar perbandingan precision–recall–F1 pada Gambar 3 memperjelas perbedaan performa antar kelas. Visualisasi tersebut menegaskan bahwa:

a. Precision kelas spam sangat tinggi (0.99): jika model memprediksi spam, hampir selalu benar.
b. Recall kelas spam lebih rendah: masih ada spam yang tidak terdeteksi.

Analisis ini menunjukkan bahwa meskipun model sangat baik, peningkatan recall spam bisa menjadi fokus penelitian lanjutan.



Gambar 3: Perbandingan Metrik Kinerja SVM

Ketercapaian Tujuan Penelitian

Tujuan penelitian yaitu menganalisis kinerja SVM dalam deteksi email spam telah tercapai. Model mampu memberikan akurasi tinggi dan konsisten setelah proses optimasi dilakukan. Tolak ukur keberhasilan ditunjukkan melalui:

a. Akurasi > 98%
b. F1-Score kelas ham hampir sempurna
c. Pemisahan kelas yang jelas pada matriks konfusi

Meski demikian, terdapat area yang belum sepenuhnya optimal, yaitu pada kemampuan model mengenali spam yang lebih “halus” atau menyerupai pesan normal. Kekurangan ini dapat diperbaiki pada penelitian lanjutan, misalnya dengan:

a. menambah fitur berbasis semantic embedding,
b. menggunakan model SVM non-linear untuk mengevaluasi pola yang lebih kompleks,
c. memperbesar dan memperkaya dataset spam.

Dengan demikian, penelitian ini tidak hanya mencapai tujuan tetapi juga membuka ruang perbaikan untuk studi lanjutan.

4. KESIMPULAN

Penelitian ini menunjukkan bahwa algoritma Support Vector Machine dengan pendekatan kernel linear mampu menjalankan fungsi klasifikasi email spam berbasis teks secara efektif. Kemampuan model dalam memisahkan kelas ham dan spam tercermin dari konsistensi performa yang stabil setelah melalui proses evaluasi dan optimasi. Pola prediksi yang dihasilkan memperlihatkan bahwa SVM tidak hanya bekerja secara akurat, tetapi juga mampu mempertahankan keseimbangan antara precision dan F1-score pada kedua kelas. Temuan ini mengonfirmasi bahwa metode yang digunakan telah sesuai dengan kebutuhan analisis deteksi spam dalam ruang lingkup data teks yang besar dan bervariasi.

Model juga menunjukkan karakteristik penting: ketika dihadapkan pada pesan-pesan yang secara linguistik menyerupai email normal, tingkat kehati-hatian model meningkat sehingga sebagian kecil pesan spam masih terprediksi sebagai ham. Hal tersebut mencerminkan kompleksitas bahasa dalam email modern dan menjadi indikasi bahwa teknik ekstraksi fitur berbasis TF-IDF bekerja baik, meskipun masih menyisakan ruang untuk menangkap ciri semantik yang lebih dalam. Dengan demikian, tujuan penelitian untuk mengukur dan menganalisis performa SVM dalam mendeteksi email spam dapat dinyatakan tercapai.

Untuk penelitian selanjutnya, pengembangan dapat diarahkan pada integrasi fitur semantik berbasis embedding, penambahan model SVM non-linear untuk mengevaluasi pola yang lebih kompleks, serta pemanfaatan dataset yang lebih beragam guna meningkatkan sensitivitas model terhadap variasi baru pada email spam. Langkah-langkah tersebut diharapkan dapat memperkuat ketelitian model dan mengurangi kesalahan pada pesan spam yang menyerupai email normal.

DAFTAR PUSTAKA

Almarisah Madani, U. (n.d.). Analisis Performa Algoritma Klasifikasi untuk Deteksi Spam pada Email Thiara Tri Funny Manguma 1✉ , Emil Fatra 2. INNOVATIVE: Journal Of Social Science Research, 4, 16461–16465.

Amin, M. B. M., Hakim, G., Maulana, M. T., Alwan, M. F., Anggraheni, H. S., Naufal, M. J., & Yudistira, N. (2024). Deteksi Spam Berbahasa Indonesia Berbasis Teks Menggunakan Model Bert. Jurnal Teknologi Informasi Dan Ilmu Komputer, 11(6), 1291–1302. https://doi.org/10.25126/jtiik.2024118121

Andryani, A., Salim, A. N., Sutabri, T., & Kunci, K. (2024). Deteksi Email Spam dan Non-Spam Berdasarkan Isi Konten Menggunakan Metode K-Nearest Neighbor dan Support Vector Machine. 6(2). https://doi.org/10.46799/syntax

Bobde, S., Role, S., Khadke, L., Shirude, T., & Kakad, S. (2023a). Email Spam Detection Using Hybridization of SVM and Random Forest. In International Journal of Research in Engineering and Science (IJRES) ISSN (Vol. 11). www.ijres.org

Bobde, S., Role, S., Khadke, L., Shirude, T., & Kakad, S. (2023b). Email Spam Detection Using Hybridization of SVM and Random Forest. In International Journal of Research in Engineering and Science (IJRES) ISSN (Vol. 11). www.ijres.org

Budiman, D., Zayyan, Z., Mardiana, A., & Mahrani, A. A. (2024). Email spam detection: a comparison of svm and naive bayes using bayesian optimization and grid search parameters. Journal of Student Research Exploration, 2(1), 53–64.

Dewi, C., Indriawan, F. A., & Christanto, H. J. (2023). Spam classification problems using support vector machine and grid search. International Journal of Applied Science and Engineering, 20(4). https://doi.org/10.6703/IJASE.202312_20(4).006

Ilham, G., Putra, M., Maulana, A., Riyadi, M. S., & Maesaroh, S. (2025). Analysis of the Application of Machine Learning Algorithm in Spam Detection System: Literature Review. https://ioinformatic.org/

Marcos, P. L., & Justine Pacatang, D. D. (2025). Enhanced Support Vector Machine for Spam Email Classification. Technoarete Transactions on Advances in Computer Applications (TTACA), 4(1).

Ngurah, G., Paramartha, D., Made, I., Sudestra, A., Wahyudi, A., Gama, O., & Humaswara Prathama, G. (n.d.). Spam Email Classification Using Support Vector Machine (SVM) and TF-IDF: A Case Study with the TREC 2007 and Enron-Spam Datasets. https://doi.org/10.47111/JTI

Putra, G. (2025). KLASIFIKASI EMAIL SPAM MENGGUNAKAN ALGORITMA ARTIFICIAL NEURAL NETWORK DAN SUPPORT VECTOR MACHINE. Jurnal Komputer Dan Informatika, 20(1), 9–15.

Rahman, A., & Maslan, A. (2025). ANALISIS KLASIFIKASI EMAIL SPAM MENGGUNAKAN ALGORITMA NAÏVE BAYES. JURNAL COMASIE, 12(03).

Sri Ainun, E., Inayah, U., & Ilmih, M. (n.d.). Klasifikasi Email Spam Dan Ham Menggunakan Algoritma Support Vector Machine, Naive Bayes Dan Logistic Regression Article Info Abstrak. https://doi.org/10.34304/scientific.v2.i2.399

Umam, C., & Handoko, L. B. (2024). Seminar Nasional Riset dan Inovasi Teknologi (SEMNAS RISTEK) 2024 Jakarta.

Umam, C., Handoko, L. B., & Isinkaye, F. O. (2024). Performance Analysis of Support Vector Classification and Random Forest in Phishing Email Classification. Scientific Journal of Informatics, 11(2), 367–374. https://doi.org/10.15294/sji.v11i2.3301