SENTRI: Jurnal Riset Ilmiah
https://ejournal.nusantaraglobal.ac.id/index.php/sentri
Email: sentri@nusantaraglobal.ac.id
DOI: https://doi.org/10.55681/sentri.v5i2.5650

Vol. 5, No. 2, February 2026
e-ISSN: 2963-1130
pp. 1523-1533

Deteksi Email Spam Menggunakan Multinomial Naive Bayes dengan Teknik Bag of Words

Widya Mulyaningtyas[1*], Kusrini[1]
[1] Program Studi PJJ Informatika, Fakultas Sains dan Teknologi, Universitas Amikom Yogyakarta, Yogyakarta, Indonesia

*Corresponding Author's e-mail: widyamulyaningtyas@students.amikom.ac.id

Article History:
Received: January 10, 2026
Revised: February 4, 2026
Accepted: February 19, 2026

Keywords:
Spam Classification, Spam Detection, Email Spam, Multinomial Naive Bayes, Bag Of Words

Abstract: *Email is a means of communication within internal networks and the internet for the exchange of information. Email is still used today because of its ease of use. However, with the increase in the number of incoming emails, the problem of spam has arisen, requiring effective methods for detecting spam so that users can manage their email more efficiently and avoid potential fraud and disruption. This study aims to analyze the thematic and linguistic patterns of email messages based on their content using text classification techniques with the Multinomial Naive Bayes algorithm, which is believed to have good accuracy in detecting spam emails. The research consists of collecting a dataset related to Indonesian-language spam emails, preprocessing the data, training the model by dividing it into two scenarios (with and without stemming), and evaluating the model. Features from the email text will be converted into numerical representations using the Bags-of-Words method. Classification performance evaluation is carried out using accuracy, precision, recall, F1-Score, and confusion matrix metrics. Experimental results demonstrate that the Multinomial Naive Bayes model without stemming achieved the highest performance with an Accuracy of 92.5%, Precision of 91.0%, and F1-Score of 91.7%. These findings indicate that stemming in short texts like spam emails eliminates crucial semantic features (affixes) characteristic of spam. This study contributes to providing optimal pre-processing recommendations for Indonesian short text classification.*

How to cite: Mulyaningtyas, W., & Kusrini, K. (2026). Deteksi Email Spam Menggunakan Multinomial Naive Bayes dengan Teknik Bag of Words. *SENTRI: Jurnal Riset Ilmiah*, *5*(2), 1523–1533. https://doi.org/10.55681/sentri.v5i2.5650

PENDAHULUAN

Surel merupakan sarana komunikasi dalam jaringan internal maupun internet untuk pertukaran informasi. Surel masih digunakan hingga saat ini karena kemudahan dalam hal penggunaannya. Saat ini, selain digunakan untuk komunikasi surel, juga digunakan untuk kebutuhan otentikasi aplikasi dan sinkronisasi media sosial seperti Instagram dan Facebook. Penggunaan surel yang tinggi bisa berdampak positif dan berdampak negatif karena tidak semua orang menggunakan surel dengan baik dan bahkan ada banyak sekali penyalahgunaan surel sehingga berpotensi merugikan pengguna surel lainnya. Surel yang disalahgunakan ini disebut sebagai spam atau surel sampah yang memiliki konten tentang iklan, penipuan, dan virus. S. N. D. Pratiwi dan B. S. S. Ulama pada tahun 2016 menyatakan dalam penelitian yang dilakukan oleh (Rahma et al., 2021).

Surel spam yang beredar di kalangan pengguna sebenarnya memiliki pola kesamaan tertentu, hanya saja banyak sekali pengguna tidak mengetahuinya. Biasanya, kasus yang terjadi adalah surel spam berjenis iklan penipuan yang memenuhi kotak masuk surel

pengguna, padahal surel tersebut tidak diinginkan. Pesan yang mengganggu dapat menyebabkan ketidakefisienan bandwidth karena merupakan kapasitas dari sebuah jaringan agar dapat dilewati oleh paket data. Bagi banyak orang, hal ini sangat mengganggu sehingga diperlukan metode yang efektif untuk deteksi pesan spam agar pengguna dapat mengelola pesan dengan lebih efisien dan terhindar dari potensi penipuan dan gangguan. Seperti penerapan *digital marketing* terbukti efektif meningkatkan penjualan produk UMKM, sebagaimana ditemukan oleh (Galuh et al., 2024). Sayangnya, strategi promosi via surel yang tidak terkontrol sering kali berkontribusi pada peningkatan volume pesan sampah (*spam*) di kotak masuk pengguna.

Penelitian ini bertujuan untuk mengelompokkan pesan Email berdasarkan isi atau kontennya dengan menggunakan teknik klasifikasi Email menggunakan algoritma Multinomial Naive Bayes yang diyakini memiliki keakuratan yang baik dalam deteksi pesan spam. Penelitian terdiri dari pengumpulan dataset terkait Email Spam, preprocessing data, melatih model, dan evaluasi model. Fitur-fitur dari pesan Email akan dikonversi menjadi representasi angka menggunakan metode Bags-of-Words. Evaluasi kinerja klasifikasi dilakukan menggunakan metrik akurasi, presisi, recall, F1-Score, dan confusion matrix. Hasil dari beberapa studi literatur, seperti penelitian yang dilakukan oleh (Hanif et al., 2024) menunjukkan bahwa metode klasifikasi Multinomial Naive Bayes memberikan hasil yang optimal dalam pengelompokan pesan email spam Hasil penelitian ini diharapkan dapat memberikan wawasan terkait deteksi pesan spam pada surel (Email) dengan menghasilkan model yang efisien dan efektif.

Mayoritas penelitian deteksi spam yang ada berfokus pada dataset bahasa Inggis (Aleisa et al., 2024).Penerapan metode klasifikasi teks pada Bahasa Indonesia memiliki tantangan tersendiri karena karakteristik morfologi yang kaya akan imbuhan (afiksasi). Penelitian terdahulu seperti (Rahma et al., 2021) telah menerapkan Naive Bayes, namun belum secara mendalam menganalisis pengaruh hilangnya fitur linguistik akibat proses *stemming* pada akurasi deteksi spam dan menurut (Amin et al., 2024) telah menerapkan model BERT untuk deteksi spam berbahasa Indonesia. Sedangkan menurut Penelitian (Dharrao et al., 2024) membuktikan bahwa teknik NLP dan *Machine Learning* efektif mendeteksi spam pada SMS. Pendekatan serupa relevan untuk diadaptasi dalam meningkatkan akurasi deteksi spam pada surel. Menurut peneliti sebelumnya (Rustam et al., 2024) menerapkan *Random Forest* dengan *Continuous Bag-of-Words* untuk deteksi spam. Sebagai alternatif yang lebih efisien secara komputasi, sedangkan penelitian ini menggunakan algoritma Multinomial Naive Bayes.Di sisi lain, pendekatan berbasis *Deep Learning* seperti yang dikembangkan oleh (Nasreen et al., 2024) memanfaatkan model BERT dan teknik seleksi fitur baru untuk mencapai akurasi tinggi. Namun, mengingat kompleksitas komputasi model tersebut, algoritma Multinomial Naive Bayes tetap menjadi pilihan menarik karena menawarkan keseimbangan antara efisiensi dan performa yang optimal untuk klasifikasi teks.

Oleh karena itu, kesenjangan penelitian (*research gap*) yang diangkat dalam studi ini adalah komparasi efektivitas pra-pemrosesan morfologi pada algoritma Multinomial Naive Bayes. Kebaruan (*novelty*) penelitian ini terletak pada pembuktian empiris bahwa untuk kasus *spam* berbahasa Indonesia, mempertahankan bentuk kata asli (tanpa *stemming*) lebih efektif daripada melakukan reduksi ke kata dasar, berbeda dengan praktik umum pada pemrosesan teks bahasa Inggris.

Berdasarkan uraian diatas penelitian ini akan mengukur metode tersebut seberapa baik performa metode Naive Bayes untuk menangani permasalahan surel spam yang

berbahasa Indonesia karena Algoritma Multinomial Naive Bayes (MNB) dipilih sebagai metode utama dalam penelitian ini karena beberapa keunggulan fundamentalnya dalam konteks klasifikasi teks MNB secara inheren dirancang untuk menangani data fitur diskrit, seperti frekuensi kemunculan kata (term frequency) yang dihasilkan oleh teknik Bag of Words, MNB dikenal memiliki efisiensi komputasi yang tinggi, baik dalam proses pelatihan maupun klasifikasi, sehingga sangat cocok untuk dataset teks berskala besar, meskipun asumsi independensi antar fitur ("naive") seringkali tidak sepenuhnya benar, MNB terbukti memberikan performa yang sangat kompetitif dan bahkan optimal pada banyak kasus deteksi spam, seperti yang ditunjukkan oleh penelitian sebelumnya (Prasad & Christy, 2022). Selain itu, pemilihan algoritma ini juga didukung oleh studi komparasi yang dilakukan oleh (Wijaya et al., 2025). Penelitian tersebut membandingkan Naive Bayes dengan *Random Forest* dan menyimpulkan bahwa meskipun algoritma *ensemble* seperti *Random Forest* tangguh, Naive Bayes menawarkan alternatif yang lebih efisien dengan performa yang tetap kompetitif, terutama untuk klasifikasi probabilistik sederhana. Terakhir, MNB menawarkan yang baik, di mana kita dapat dengan mudah mengekstraksi probabilitas setiap kata terhadap kelas spam dan non-spam, yang berguna untuk menjawab pertanyaan penelitian mengenai fitur-fitur penting dengan ini peneliti membuat judul "Deteksi Email Spam Menggunakan Multinomial Naive Bayes Dengan Teknik Bag Of Words".

METODE PENELITIAN

Pada bagian ini akan memberikan informasi terkait proses pengumpulan dataset, preprocessing dataset, model yang digunakan, serta performance metrics yang digunakan dalam melakukan evaluasi keakuratan model dalam deteksi spam pada Email.

Dataset Collection

Untuk melakukan percobaan ini dataset yang digunakan adalah dataset Email berbahasa Indonesia dari Kaggle. Kumpulan data terdiri dari 2638 Surel teks yang belum dilakukan preprocessing, dengan 2 kategori Surel, yaitu ham sebesar 47.8% (1252) dan spam sebesar 52.2% (1368) yang belum dilakukan preprocessing. Hanya data berlabel yang akan digunakan untuk menguji kinerja model.

Dataset Preprocessing

Penelitian ini menggunakan seluruh dataset untuk percobaan, namun dataset mengalami preprocessing terlebih dahulu sebelum masuk ke tahap implementasi pada metode yang diusulkan. Langkah preprocessing dilakukan untuk membersihkan dan menyiapkan dataset untuk tahap selanjutnya. Berikut adalah langkah-langkah preprocessing dataset pada penelitian ini (Figure 1) :

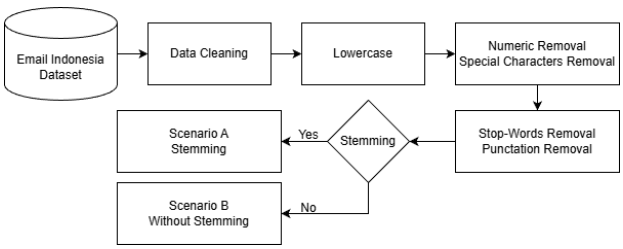

Figure 1. Urutan langkah dalam preprocessing Email Indonesia Dataset

Cleaning Data: Melakukan pemeriksaan awal terhadap dataset untuk memastikan tidak ada data yang hilang atau tidak lengkap, dan memahami struktur dataset seperti

jumlah kolom dan baris. Kemudian, melakukan pembersihan data dengan langkah-langkah berikut :

Proses pra-pemrosesan data dilakukan melalui beberapa tahap penting. Pertama, **penghapusan baris duplikat** dilakukan untuk membersihkan dataset dari baris-baris yang berulang. Selanjutnya, tahap **mengatasi missing values** dilakukan dengan menangani nilai yang hilang atau kosong, baik dengan menggantinya menggunakan nilai yang sesuai seperti rata-rata atau median, maupun dengan menghapus baris atau kolom yang mengandung nilai yang hilang. Kemudian, dilakukan **penghapusan kolom yang tidak diperlukan** untuk menyingkirkan kolom-kolom yang tidak relevan atau tidak digunakan dalam analisis. Selain itu, dilakukan **labeling atribut**, yaitu mengubah atribut target menjadi numeric, dengan spam diberi nilai 1 dan ham diberi nilai 0.

Tahap berikutnya adalah **normalisasi teks**, yang mencakup beberapa langkah. Pertama, teks diubah menjadi **lowercase** agar konsisten. Selanjutnya, dilakukan **penghapusan angka**, karena konten teks surel dapat mengandung angka yang tidak berguna dalam model pelatihan, sehingga penghapusan angka ini membantu mengurangi kompleksitas rangkaian fitur. Selain itu, dilakukan **penghapusan karakter khusus**, termasuk spasi yang berlebihan. Terakhir, dilakukan **penghapusan stop words dan punctuation**, untuk menyingkirkan kata-kata umum yang tidak memberikan nilai tambah dalam analisis serta menghapus tanda baca.

Setelah normalisasi teks, terdapat dua skenario dalam pemrosesan lebih lanjut. **Skenario A (Dengan Stemming)**, dimana teks email dibersihkan melalui case folding, penghapusan tanda baca, dan angka, kemudian dilakukan penghapusan stopword, dan diakhiri dengan proses stemming untuk mengubah setiap kata ke bentuk dasarnya. Pendekatan ini sejalan dengan penelitian (Jaiswal et al., 2021) yang menunjukkan bahwa kombinasi teknik stemming dan stop word removal bersama algoritma Naive Bayes terbukti efektif meningkatkan akurasi deteksi spam. Sedangkan pada **Skenario B (Tanpa Stemming)**, teks email menjalani proses yang sama persis dengan Skenario A, namun tidak melalui tahap stemming, sehingga kata-kata dibiarkan dalam bentuk aslinya setelah stopword dihilangkan.

Feature Extraction

Pendekatan berbasis konten melibatkan penggunaan kata atau frekuensi karakter. Teknik Bag of Words (BOW) dapat digunakan untuk merepresentasikan setiap kata yang tersisa sebagai fitur dokumen. Langkah-langkah preprocessing menyederhanakan dokumen tetapi juga berpotensi menurunkan makna. Nilai fitur dalam BOW dapat dimodifikasi dengan menggunakan teknik seperti term count, term frequency,and term frequency–inverse document frequency (Krishna Juluru et al., 2021). Dalam pendekatan berbasis Bow, urutan kata dan hubungan semantiknya tidak dipertimbangkan. Teknik Bag-of-Words dapat mengkonversi teks dengan panjang variabel menjadi teks dengan panjang tetap atau disebut vector. Jadi untuk lebih spesifiknya, dengan menggunakan teknik bag-of-words (BoW), teks akan diubah menjadi vektor angka yang setara. Hal ini diperlukan karena metode *machine learning* bekerja dengan memproses informasi numerik, bukan informasi tekstual secara langsung (Raharjo, 2021). Setelah melakukan feature extraction, maka data dapat dibagi menjadi dua dua subset, yaitu data latih (training data) dan data uji (testing data) untuk digunakan pada pelatihan dan evaluasi model.

Algoritma Multinomial Naive Bayes

Metode Multinomial Naïve Bayes didasarkan pada Teorema Bayes dan cocok untuk data yang diwakili sebagai representasi vektor (Hassan et al., 2022). Metode ini merupakan

variasi dari Naive Bayes yang bekerja dengan asumsi bahwa fitur-fitur input adalah independen dan didistribusikan multinomial. Dalam konteks deteksi Email spam, fitur-fiturnya yang digunakan ialah kata-kata dalam Surel, dengan kelas 'spam' atau 'ham'. Metode ini digunakan untuk membuat model klasifikasi atau deteksi spam pada Email, dengan cara melatih dan menguji model menggunakan dataset. Pembagian dataset dilakukan dengan rasion 0.2, yang berarti 20% (525) untuk data uji dan 80% (2097) untuk data latih.

Performance Evaluasi

Tahapan ini merupakan proses untuk mengukur kinerja model dalam memprediksi apakah suatu Email termasuk dalam kategori spam atau bukan. Evaluasi model ini penting untuk memastikan bahwa model yang dikembangkan dapat memberikan prediksi yang akurat dan dapat diandalkan.

Confusion Matrix

Confusion matrix adalah tabel yang menggambarkan performa model klasifikasi dengan membandingkan prediksi yang dibuat oleh model dengan nilai sebenarnya dari data uji.

Tabel 1. Confusion Matrix

		Prediksi Class	
Actual Class		**Negative**	**Positive**
	Negative	TN	FN
	Positive	FP	TP

Keterangan :

True Negative = Nilai prediksi spam dan nilai sebenarnya juga spam

False Negative = Nilai prediksi ham, tetapi nilai sebenarnya spam

False Positive = Nilai prediksi spam, tetapi nilai sebenarnya ham

True Positive = Nilai prediksi ham dan nilai sebenarnya juga ham

Akurasi mengukur seberapa sering model memberikan prediksi yang benar secara keseluruhan dan dinyatakan sebagai berikut :

$$Accuracy = (TN+TP) / (TN+FP+FN+TP)$$

Presisi mengukur seberapa sering model memberikan prediksi positif yang benar dari semua prediksi positif yang diberikan dan dinyatakan sebagai berikut :

$$Precision = TP / (TP+FP)$$

Recall mengukur seberapa sering model dapat menemukan semua instance yang benar positif dan dinyatakan sebagai berikut :

$$Recall = TP / (TP+FN)$$

F1-score merupakan rata-rata harmonis antara presisi dan recall, memberikan pengukuran keseluruhan tentang kinerja model, dan dinyatakan sebagai berikut :

$$F1\text{-}Score = 2\times((Precision \times Recall) / (Precision+Recall))$$

HASIL DAN PEMBAHASAN

Preprocessing Data

Tahapan ini menghasilkan data yang bersih dan siap untuk membangun model. Setelah data bersih (Table 2), total data Surel teks yang tersisa sebanyak 2638 dengan total Surel ham sebesar 47.8% dan Surel spam sebesar 52.2% (Table 3).

Tabel 2. Sample Dataset

Pesan	Skenario A	Skenario B
Secara alami tak tertahankan identitas perusahaan	alami tak tahan identitas usaha sangat sulit	alami tak tertahankan identitas perusahaan
Fanny Gunslinger Perdagangan Saham adalah merrill muzo	fanny gunslinger dagang saham merrill muzo	fanny gunslinger perdagangan saham merrill muzo
Rumah - rumah baru yang luar biasa menjadi mudah	rumah rumah baru luar biasa jadi mudah ingin	rumah rumah baru luar biasa menjadi mudah ingin
4 Permintaan Khusus Pencetakan Warna Informasi	minta khusus cetak warna informasi	permintaan khusus pencetakan warna informasi

Tabel 3. Distribuasi Kelas Dataset

Nama Kelas	Presentase (%)	Total
Spam (1)	52.2	1368
Ham (0)	47.8	1252

Pembagian Dataset

Pembagian dataset (Table 4) dilakukan menjadi dua bagian, yaitu data uji dan data latih. Dataset dialokasikan 80% untuk data latih (2097) dan 20% (525) untuk data uji.

Tabel 4. Pembagian Dataset

Dataset	Total	Spam	Ham
Data Latih	2097	1094	1001
Data Data Uji	525	274	251

Tahap Ekstraksi Fitur dan Pemodelan

Ekstraksi Fitur Bag of Words (BoW): Teks yang telah bersih dari kedua skenario (A dan B) kemudian diubah menjadi representasi vektor numerik menggunakan teknik Bag of Words. Hasilnya adalah matriks yang merepresentasikan frekuensi kemunculan setiap kata dalam setiap Email.

Model Multinomial Naïve bayes tanpa stemming (Tabel 6) menunjukkan performa yang cukup baik dalam melakukan klasifikasi Email spam atau ham berbahasa indonesia. Secara keseluruhan model ini berhasil mencapai tingkat akurasi sebesar 99%. Hasil evaluasi (Tabel 5) menunjukkan bahwa model dengan memiliki tangkat precision, recall, dan F1-Score sebesar 98% evaluasi yang dilakukan dengan stemming. Sebagai pembanding, penelitian ini juga menguji algoritma *Support Vector Machine* (SVM) yang dikenal handal dalam klasifikasi teks, sebagaimana juga diterapkan dalam studi (Given Putra, 2025) **h**asil evaluasi (Tabel 7) menunjukkan bahwa model dengan memiliki tangkat precision, recall, dan F1-Score sebesar 96% evaluasi SVM yang dilakukan tanpa stemming.

Tabel 5. Hasil Evaluasi Model Multinomial Naïve Bayes Dengan Stemming

Kelas	Precision (%)	Recall (%)	F1-Score (%)
Spam	98	99	99
Ham	99	98	98
Akurasi (%)	98%		

Tabel 6. Hasil Evaluasi Model Multinomial Naïve Bayes Tanpa Stemming

Kelas	Precision (%)	Recall (%)	F1-Score (%)
Spam	99	99	99
Ham	99	99	99
Akurasi (%)	99%		

Tabel 7. Hasil Evaluasi Model SVM Tanpa Stemming

Kelas	Precision (%)	Recall (%)	F1-Score (%)
Spam	96	97	96
Ham	96	96	96
Akurasi (%)	96%		

Performa Model MNB-A dibandingkan secara langsung dengan Model MNB-B terlihat pada (Figure 2). Terlihat hasil perbandingan ini menyimpulkan bagaimana proses stemming sangat memengaruhi efektivitas deteksi spam. Visualisasi Analisis Kesalahan Model juga ditunjukkan oleh confusion matrix (Figure 3).

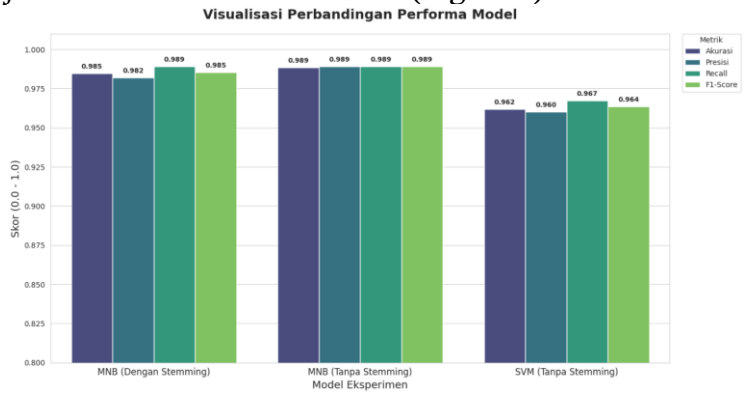

Figure 2. Visualisasi Perbandingan Performa Model

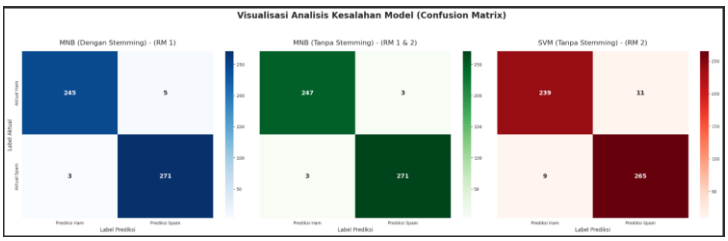

Figure 3. Confusion Matrix

Visualisasi Pola Kata

Melakukan visualisasi terhadap pola kata yang sering muncul pada masing-masing kelas, dapat memberikan gambaran tentang isi konten pada Surel ham dan spam.

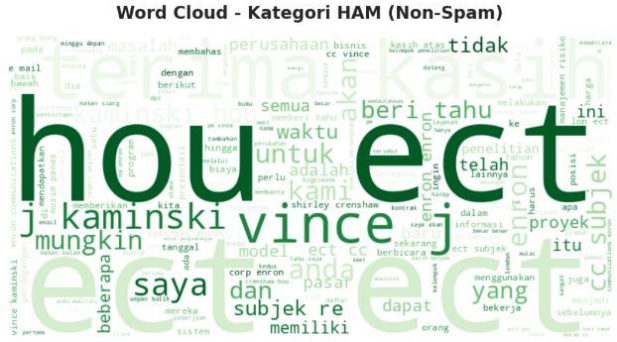

Figure 4. Visualisasi Surel Ham

Figure 5. Visualisasi Surel Spam

Kata-kata seperti “Mungkin”, “Saya”, ”dan”, ”beri”, ”tahu”, ”tidak”, ”tanggal”, “subjek”, “memiliki”, “yang”, dan “itu” cenderung muncul lebih sering pada kelas ham. Hal ini menunjukkan bahwa, Surel teks pada kelas ham memiliki karakteristik bahasa yang digunakan dalam komunikasi sehari-hari. Kata-kata tersebut termasuk dalam kategori umum untuk digunakan dalam percakapan normal dan surel non-spam. Sementara itu, pada kelas spam kata-kata seperti “anda”, “situs”, “web”, “perusahaan”, “perangkat”, “lunak”, “gratis”, “lebih” dan “lanjut” cenderung muncul lebih sering. Hal ini mencerminkan karakteristik umum dari Surel-Surel spam, dimana kata-kata tersebut seringkali digunakan untuk mempromosikan sesuatu atau menarik perhaitan pengguna.

Analisis Pola Tematik

Analisis Pola Linguistik & Tematik. Model MNB-B (dipilih karena interpretasi probabilistiknya) akan dianalisis secara mendalam. Fitur (kata) dengan nilai probabilitas tertinggi untuk setiap kelas (spam dan ham) akan diekstraksi. Analisis ini bukan sekadar mendaftar kata, melainkan menginterpretasikan pola linguistik dan tema yang muncul dari kata-kata tersebut untuk memahami mengapa model mengklasifikasikan Email dengan cara tertentu. Hal ini menunjukkan bahwa model mampu mengenali pola dan ciri khas dari surel spam dengan akurat. Dengan demikian proses ini tidak hanya memvalidasi performa model, tetapi juga memberikan wawasan berharga dalam pengelolaan dan peningkatan kualitas model yang digunakan.

Terlihat pada (Figure 6) yaitu Pola Tematik SPAM yang memiliki tema: Urgensi, keuangan, dan Ajakan. Dimana kata kunci yang paling banyak yaitu “Hari”, “Memiliki”, “sini”, “informasi” dan “yang”. Kemudian pada (Figure 7) yaitu Pola Tematik HAM yang memiliki tema: Profesional, Relasional, dan Informatif. Kata kunci yang paling banyak yaitu “Risiko”, “Untuk”, “Tahu”, dan “Lebih”.

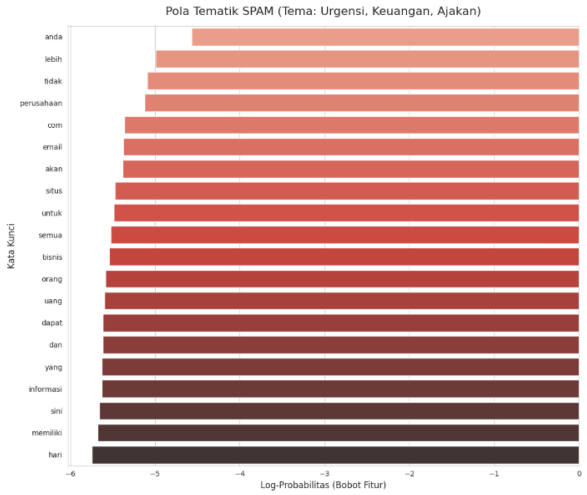

Figure 6. Pola Tematik SPAM

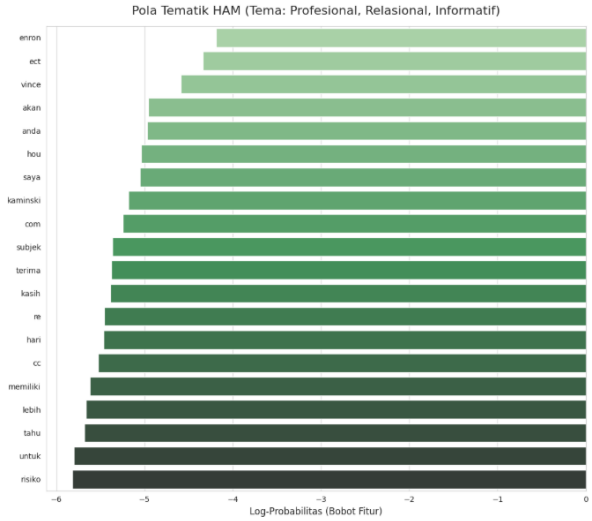

Figure 7. Figure 7 Pola Tematik HAM

Berdasarkan hasil eksperimen di atas, terlihat bahwa skenario tanpa stemming lebih unggul. Temuan ini sejalan dengan teori Information Retrieval yang dikemukakan oleh (Christopher D. Manning et al., 2008), yang menyatakan bahwa stemming dapat meningkatkan recall namun seringkali menurunkan precision karena terjadinya over-stemming (dua kata berbeda makna direduksi menjadi akar yang sama). Dalam konteks deteksi spam bahasa Indonesia, kata-kata persuasif yang menggunakan partikel atau akhiran (misalnya: 'dapatkan', 'menangkan') adalah fitur diskriminatif yang kuat. Penghilangan imbuhan tersebut melalui proses stemming menyebabkan hilangnya informasi semantik spesifik yang membedakan email promosi (spam) dan email normal, sehingga akurasi model menurun.

KESIMPULAN

Penelitian ini dirancang sebagai studi eksperimental yang sistematis untuk menganalisis efektivitas dan karakteristik model machine learning dalam mendeteksi email spam berbahasa Indonesia. Dengan berfokus pada algoritma Multinomial Naive Bayes (MNB) dengan representasi Bag of Words (BoW), penelitian ini tidak hanya bertujuan

untuk mengukur performa, tetapi juga untuk menghasilkan pengetahuan yang lebih mendalam tentang kapan dan mengapa sebuah metode efektif digunakan.

Dalam pelaksanaannya, penelitian ini menguraikan tiga alur analisis yang saling terkait. Pertama, penelitian akan menguji secara kuantitatif dampak preprocessing, khususnya stemming, terhadap trade-off performa model antara presisi dan recall. Kedua, penelitian akan membandingkan efektivitas dan efisiensi MNB dengan algoritma baseline Support Vector Machine (SVM) untuk memberikan rekomendasi berbasis bukti dalam konteks efisiensi komputasi. Ketiga, dan yang paling fundamental, penelitian akan beralih dari sekadar angka performa menuju analisis interpretatif, dengan mengekstraksi fitur probabilitas tertinggi dari model MNB untuk mengidentifikasi dan menganalisis pola linguistik serta tematik yang paling fundamental dalam membedakan email spam dan non-spam berbahasa Indonesia.

DAFTAR REFERENSI

Aleisa, M. A., Alsuwit, M. H., & Haq, M. A. (2024). Advancing Email Spam Classification using Machine Learning and Deep Learning Techniques. *Engineering, Technology and Applied Science Research*, *14*(4), 14994–15001. https://doi.org/10.48084/etasr.7631

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

Christopher D. Manning, Prabhakar Raghavan, & Hinrich Schütze. (2008). Introduction to Information Retrieval. *Cambridge University Press*, *11, 239–250*.

Dharrao, D., Gaikwad, P., Gawai, S. V., Bongale, A. M., Patel, K., & Singh, A. (2024). Classifying SMS as Spam or Ham: Leveraging NLP and Machine Learning Techniques. *International Journal of Safety and Security Engineering*, *14*(1), 289–296. https://doi.org/10.18280/ijsse.140128

Galuh, U., Kasus, S., Rasa Galendo, P. D., Cigembor, K., Ciamis, Kecamatan, Ciamis, Kabupaten, Barat, J., Deassy,), Sari, R. J., & Nurhayaty, M. (2024). PENERAPAN GO DIGITAL MARKETING UNTUK MENINGKATKAN PENJUALAN PRODUK PADA UMKM GALENDO. *Jurnal Media Teknologi*, *11*(01).

Given Putra. (2025). 09-15. *Jurnal Komputer Dan Informatika Vol 20 No 1, April 2025: Hlm 09- 15*.

Hanif, M., Ubaidillah, Z., & Fatah, Z. (2024). *Volume 2 ; Nomor 10*. 129–132. https://doi.org/10.59435/gjmi.v2i11.536

Hassan, S. U., Ahamed, J., & Ahmad, K. (2022). Analytics of machine learning-based algorithms for text classification. *Sustainable Operations and Computers*, *3*, 238–248. https://doi.org/10.1016/j.susoc.2022.03.001

Jaiswal, M., Das, S., & Khushboo. (2021). Detecting spam e-mails using stop word TF-IDF and stemming algorithm with Naïve Bayes classifier on the multicore GPU. *International Journal of Electrical and Computer Engineering*, *11*(4), 3168–3175. https://doi.org/10.11591/ijece.v11i4.pp3168-3175

Krishna Juluru, MD Hao-Hsin Shih, MS Krishna Nand Keshava Murthy, MS Pierre Elnajjar, & MS. (2021). *Bag-of-Words Technique in Natural Language Processing_ A Primer for Radiologists*.

Nasreen, G., Murad Khan, M., Younus, M., Zafar, B., & Kashif Hanif, M. (2024). Email spam detection by deep learning models using novel feature selection technique and BERT. *Egyptian Informatics Journal*, *26*. https://doi.org/10.1016/j.eij.2024.100473

Prasad, J. K., & Christy, S. (2022). SMS Spam Detection Using Multinational Naive Bayes Algorithm Compared with Decision Tree Algorithm. *BALTIC JOURNAL OF LAW & POLITICS A Journal of Vytautas Magnus University*, *15*(4), 349–356. https://doi.org/10.2478/bjlp-2022-004037

Raharjo, B. (2021). *P Y YAYASAN PRIMA AGUS TEKNIK YAYASAN PRIMA AGUS TEKNIK YAYASAN PRIMA AGUS TEKNIK Pembelajaran Mesin (Machine Learning)*.

Rahma, F., Farmadiansyah, A. Z., & Hidayatullah, A. F. (2021). *Deteksi Surel Spam dan Non Spam Bahasa Indonesia Menggunakan Metode Naïve Bayes*.

Rustam, M., Brotokuncoro, A., & Roestam, R. (2024). DETEKSI EMAIL SPAM DENGAN CONTINUOUS BAG-OF-WORDS DAN RANDOM FOREST. In *Jurnal Ilmiah Sain dan Teknologi* (Vol. 2, Number 4).

Wijaya, A. P., Penulis, *, & Diajukan, K. (2025). Perbandingan Algoritma Klasifikasi Random Foresst dengan Naïve Bayes Classifier pada Studi Penyakit Berdasarkan Pola Nutrisi. *Remik: Riset Dan E-Jurnal Manajemen Informatika Komputer*, *9*(1). https://doi.org/10.33395/remik.v9i1.14652