



Klasifikasi Email Spam Menggunakan Algoritma Naive Bayes

Muhammad Hanif Zaky Ubaidillah¹, Zaehol Fatah²

¹ Teknologi Informasi, Sains dan Teknologi, Universitas Ibrahimy

² Sistem Informasi, Sains dan Teknologi Universitas Ibrahimy

muhammadhanifubaidillah00@email.com, zaeholfatah@email.com,

Abstrak

Email spam merupakan salah satu masalah signifikan dalam komunikasi digital yang dapat mengganggu produktivitas pengguna. Untuk mengembangkan model klasifikasi email spam maka digunakan Algoritma Naive Bayes, yang dikenal karena kesederhanaan dan efektivitasnya dalam menangani data teks. Data email yang terdiri dari spam dan non-spam diambil dari sumber yang dapat diakses secara publik dan diproses menggunakan aplikasi RapidMiner. Langkah-langkah preprocessing, termasuk penghapusan stopwords, stemming, dan vektorisasi teks, dilakukan untuk meningkatkan akurasi model. Model Naive Bayes dilatih dan diuji menggunakan teknik cross-validation untuk memastikan keandalan hasil. Hasil evaluasi menunjukkan bahwa model yang dikembangkan memiliki tingkat akurasi yang tinggi dengan hasil 60%, dengan nilai presisi 40% dan recall yang memadai. Hasil ini menunjukkan potensi penggunaan Algoritma Naive Bayes dalam klasifikasi email spam dan memberikan kontribusi terhadap pengembangan sistem filtrasi email yang lebih efektif.

Kata Kunci: Naive Bayes, Email, Teks

PENDAHULUAN

Dalam era digital saat ini, email menjadi salah satu sarana komunikasi yang paling umum digunakan. Namun, meningkatnya volume email juga disertai dengan masalah signifikan berupa spam, yaitu pesan yang tidak diinginkan yang dapat mengganggu pengguna dan mengurangi produktivitas. Diperkirakan bahwa 55% dari total email yang dikirim setiap hari adalah spam (Pratiwi & Ulama, 2016). Hal ini menunjukkan perlunya sistem yang efektif untuk mengidentifikasi dan memfilter email spam guna melindungi pengguna dari potensi risiko dan kerugian.

Email umumnya berjalan di Internet, walau bisa juga di jaringan komputer. Beberapa sistem email jaman dahulu mensyaratkan agar penerima dan pengirim online pada saat yang bersamaan. Sistem email sekarang menggunakan model store-and-forward. Di mana server email menerima, mem-forward, dan menyimpan pesan. Sehingga user dan komputernya tidak perlu online di saat yang bersamaan. Prngguna hanya perlu terhubung ke mail server sebentar saja, ketika akan mengirim atau menerima pesan saja. (Ali Zaki, Edy Winarno ST, n.d.)

Naive Bayes merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai Teorema Bayes. Teorema tersebut dikombinasikan dengan Naive di mana diasumsikan kondisi antar atribut saling bebas. Algoritma Naive Bayes merupakan salah satu algoritma yang terdapat pada teknik klasifikasi. Klasifikasi Naive Bayes diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lainnya. (Dr.Rina Fitriana, ST., MM. et al., n.d.)

Beberapa penelitian sebelumnya telah menunjukkan keberhasilan penggunaan Algoritma Naive Bayes dalam klasifikasi email spam. Misalnya (Classifier, 2024) menemukan bahwa model Naive Bayes dapat mencapai akurasi di atas 97% dalam pengenalan email spam. Penelitian lain oleh (Fitriyah et al., 2020) dapat mencapai akurasi di atas 84.8% dalam pengenalan email spam.

Pada penelitian sebelumnya yang menggunakan Naive Bayes bagaimana membantu pihak terkait memudahkan dalam menentukan kelayakan penerima Kartu Indonesia Pintar, karena perhitungan yang kurang tepat akan menimbulkan kecemburuan sosial, dengan itu maka membutuhkan perhitungan yang mudah diimplementasikan dengan memanfaatkan salah satu algoritma metode data mining yaitu Algoritma Naive Bayes guna menjawab permasalahan. (Homaidi et al., 2024)

Sebagai solusi, klasifikasi email spam menggunakan teknik pembelajaran mesin telah terbukti efektif. Salah satu algoritma yang banyak digunakan adalah Naive Bayes, yang didasarkan pada teorema Bayes dan asumsi independensi antar fitur. Dalam penelitian ini, kami menggunakan aplikasi RapidMiner sebagai platform untuk implementasi algoritma Naive Bayes, yang memungkinkan pengguna untuk melakukan analisis data dan pengembangan model dengan antarmuka yang intuitif.

Tujuan dari penelitian ini adalah untuk mengembangkan dan mengevaluasi model klasifikasi email spam menggunakan Algoritma Naive Bayes di aplikasi RapidMiner, serta untuk mengidentifikasi faktor-faktor yang berkontribusi terhadap kinerja model. Diharapkan hasil penelitian ini dapat memberikan kontribusi dalam

pengembangan sistem filtrasi email yang lebih efektif dan efisien. Selain itu, penelitian ini juga bertujuan untuk memberikan wawasan tentang aplikasi praktis dari Algoritma Naive Bayes dalam konteks klasifikasi teks, khususnya dalam dunia email.

METODE

Metode Pengambilan Data

Bagian ini menjelaskan metode yang digunakan untuk mengumpulkan data terkait penelitian yang sedang dilakukan. Peneliti mengumpulkan data dengan metode studi literatur menggunakan data testing dan data training, yaitu data spam email yang diambil dari situs Kaggle. Data ini kemudian disaring dan akan diolah untuk di klasifikasi status spam dan tidak spam menggunakan aplikasi RapidMiner. Beberapa atribut yang digunakan meliputi jumlah kata promosi, link external, Panjang email, jumlah kata unik, subjek berisi promosi. Dataset tersebut ditampilkan dalam gambar berikut:

Tabel 1. Dataset Training

Jumla_kata_promosi	Ada_Link_External	Panjang_Email	Jumlah_Kata_Unik	Subjek_berisi_promosi	Label (spam / tidak spam)
3	1	140	82	1	Spam
2	0	50	43	0	Tidak spam
6	1	160	78	1	Spam
1	0	55	50	0	Tidak spam
4	1	135	70	1	Spam
0	0	45	42	0	Tidak spam
3	1	115	68	0	Spam
5	1	180	80	1	Spam
2	0	90	66	0	Tidak spam
4	1	145	85	1	spam

Tabel 2. Dataset Testing

Jumla_kata_promosi	Ada_Link_External	Panjang_Email	Jumlah_Kata_Unik	Subjek_berisi_promosi
3	1	140	82	1
2	0	50	43	0
6	1	160	78	1
1	0	55	50	0
4	1	135	70	1
0	0	45	42	0
3	1	115	68	0
5	1	180	80	1
2	0	90	66	0
4	1	145	85	1

Email

Email spam, yang sering dikenal sebagai email sampah, surat spam, atau hanya spam, mengacu pada pengiriman massal iklan yang tidak diinginkan melalui email (spamming). Spam menjadi menjengkelkan karena terus-menerus muncul di kotak masuk anda. Karena penerima biasanya harus membayar untuk spam tersebut, Hukum dan litigasi belum terlalu efektif dalam menghentikan spam di mana pun, meskipun definisi dan status spam bervariasi dari satu yurisdiksi ke yurisdiksi lainnya. Sebagian besar email spam berisi iklan. Banyak email, baik komersial atau lainnya, tidak hanya mengganggu karena mencuri waktu dan fokus yang berharga, namun juga berbahaya karena dapat menyertakan link ke situs web berbahaya, seperti halaman phishing atau situs yang menghosting malware atau memiliki malware yang dilampirkan ke email itu sendiri. (Sabry, n.d.)

Naive Bayes

Naive Bayes merupakan suatu algoritma yang dapat mengklasifikasikan suatu variable tertentu dengan menggunakan metode probabilitas dan statistic. Naive Bayes menggunakan sebuah ilmu cabang matematika yang dikenal juga dengan teori probabilitas untuk mencari peluang terbesar dari kemungkinan klasifikasi, dengan cara melihat frekuensi tiap klasifikasi pada data training. Keuntungan dari penggunaan metode Naive Bayes adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (Training Data) yang kecil saja untuk menentukan estimasi parameter yang

diperlukan dalam proses pengklasifikasian dan dapat bekerja jauh lebih baik.(Ema Ainun Novia, Woro Iisti Rahayu, 2020)

Rumus dasar Naive Bayes sebagai berikut:

$$P(\text{spam} | X) = \frac{P(\text{spam}) \cdot P(X | \text{spam})}{P(X)}$$

Di mana:

- $P(\text{spam}|X)$ adalah probabilitas bahwa email tersebut adalah spam dengan fitur X.
- $P(\text{spam})$ adalah proir atau probabilitas awal bahwa sebuah email adalah spam.
- $P(X|\text{spam})$ adalah probabilitas kemunculan fitur X jika email adalah spam.
- $P(X)$ adalah probabilitas total fitur X yang digunakan sebagai pengali dalam Naive Bayes.

Untuk klasifikasi, dapat mengabaikan $P(X)$ karena nilai ini sama untuk kedua kelas (spam dan tidak spam). Jadi rumusnya disederhanakan menjadi:

$$P(\text{spam}|X) \propto P(\text{spam}) \cdot P(X|\text{spam})$$

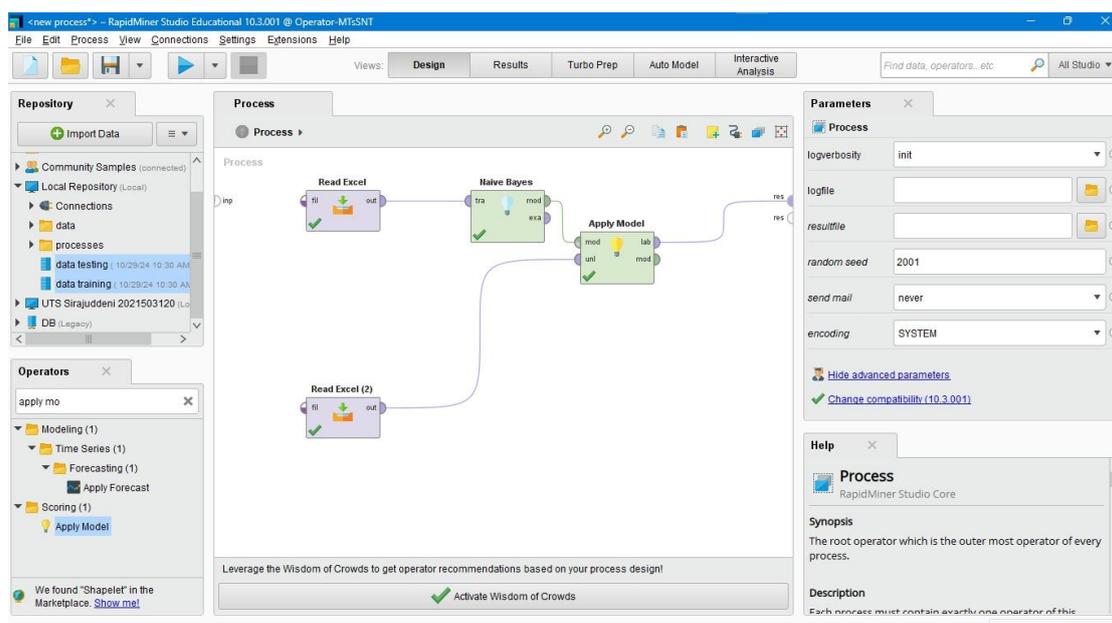
$$P(\text{tidak spam}|X) \propto P(\text{tidak spam}) \cdot P(X|\text{tidak spam})$$

Data Mining

Data mining merupakan sebuah teknologi yang dapat memproses data dalam volume besar yang digunakan oleh perusahaan untuk mengubah data mentah menjadi informasi yang berguna untuk membuat suatu keputusan bisnis yang sangat penting. Pada dasarnya data mining mempunyai 7 fungsi yaitu Description, Classification, Clustering, Association, Sequencing, Forecasting, dan Prediction. Data mining mempunyai tujuan sebagai Explanatory yaitu untuk menjelaskan beberapa kondisi terkait dengan suatu penelitian, Confirmatory digunakan sebagai mempertegas hipotesis, dan Exploratory yang berguna dalam menganalisis data untuk hubungan yang baru yang tidak diharapkan.(Mustika, Yunita Ardilla, n.d.)

HASIL DAN PEMBAHASAN

Pada Gambar 1 menjelaskan tentang penginputan proses data melalui Rapid Miner penghitungan tools rapid miner Read Excel untuk menginput data yang akan di proses, Naive Bayes untuk menghitung data yang akan di proses, Apply Model untuk menguji data hasil operator Naive Bayes model ini mendapatkan presiksi pada data yang tidak terlihat atau untuk mengubah data dengan menerapkan model preprocessing.



Gambar 1. Proses Penghitungan Tools Rapidminer dengan Metode Nive Bayes

Pada Gambar 2 merupakan hasil klasifikasi data mining melalui Rapid Miner hasil dari penghitungan pada gambar 1. Dapat disimpulkan bahwa probabilitas awal untuk email berlabel spam adalah 0.6, yang berarti ada kemungkinan 60% bahwa sebuah email akan berlabel spam. Untuk probabilitas awal untuk email berlabel tidak spam adalah 0.4, yang berarti ada kemungkinan 40% bahwa sebuah email akan berlabel tidak spam.

Row No. ↑	prediction(la...	confid...	confi...	m lah_k...	ada_link_e...	panjang_em...	jumlah_kata...	subjek_beri...
1	spam	1	0	3	1	140	82	1
2	tidak spam	0	1	2	0	50	43	0
3	spam	1	0	6	1	160	78	1
4	tidak spam	0	1	1	0	55	50	0
5	spam	1	0	4	1	135	70	1
6	tidak spam	0	1	0	0	45	42	0
7	spam	1	0	3	1	115	68	0
8	spam	1	0	5	1	180	80	1
9	tidak spam	0	1	2	0	90	66	0
10	spam	1	0	4	1	145	85	1

Gambar 2. Hasil Klasifikasi

KESIMPULAN

Dari hasil klasifikasi menggunakan metode Naive Bayes dengan data training, lebih banyak email yang diberi label spam dengan akurasi 60% dibandingkan dengan yang diberi label tidak spam 40%. Prioritas ini memberikan dasar bagi Naive Bayes untuk memprediksi label berdasarkan fitur yang diamati pada data testing. Probabilitas prior ini nantinya dikalikan dengan probabilitas kemunculan fitur spesifik untuk memutuskan apakah sebuah email termasuk spam atau tidak spam.

UCAPAN TERIMA KASIH

Saya ingin menyampaikan rasa terimakasih yang sebesar-besarnya kepada semua pihak yang telah mendukung dan membantu dalam penyusunan jurnal ini. Terutama, saya ucapkan terimakasih kepada (bapak zaehol fatah M.Kom) yang telah memberikan bimbingan, masukan berharga, serta arahan yang sangat membantu dalam setiap penelitian ini. Terimakasih juga kepada rekan-rekan dan sahabat yang senantiasa memberikan dukungan, semangat, dan motivasi selama proses penulisan jurnal ini. Tidak lupa, terimakasih kepada keluarga tercinta yang selalu mendoakan dan memberikan dukungan tanpa henti. Terimakasih juga kepada semua pihak yang tidak dapat saya sebutkan satu persatu, namun telah memberikan kontribusi dan dukungan yang berarti. Semoga jurnal ini dapat memberikan manfaat dan kontribusi positif bagi perkembangan ilmu pengetahuan di bidang Teknologi Informasi.

DAFTAR PUSTAKA

- Ali Zaki, Edy Winarno ST, M. (n.d.). *email. elex media komputindo*.
- Classifier, N. B. (2024). *NAIVE BAYES CLASSIFIER UNTUK DETEKSI EMAIL SPAM - Google Scholar*. 15(4), 675–680.
https://scholar.google.com/scholar?hl=id&as_sdt=0%2C5&q=NAIVE+BAYES+CLASSIFIER+UNTUK+DETEKSI+EMAIL+SPAM&btnG=
- Dr.Rina Fitriana, ST., MM., I., Anik Nur Habyba, STP., M. S., & Elfira Febriani, STP., M. S. (n.d.). *naive bayes* (N. Wahid (ed.)). Wawasan Ilmu.
- Ema Ainun Novia, Woro Iisti Rahayu, C. P. (2020). *Naive Bayes Classifier* (R. M. Awangga (ed.)). Kreatif Industri Nusantara.
- Fitriyah, N. Q., Oktavianto, H., & Hasbullah, H. (2020). Deteksi Spam Pada Email Berbasis Fitur Konten Menggunakan Naïve Bayes. *JUSTINDO (Jurnal Sistem Dan Teknologi Informasi Indonesia)*, 5(1), 1–7.
<https://doi.org/10.32528/justindo.v5i1.3414>
- Homaidi, A., Fatah, Z., & others. (2024). Penerapan Metode Naïve Bayes Untuk Menentukan Penerima Kartu Indonesia Pintar (KIP). *G-Tech: Jurnal Teknologi Terapan*, 8(3), 1806–1815.
- Mustika, Yunita Ardilla, A. M. (n.d.). *Data Mining dan Aplikasinya* (N.Rismawati (ed.)). WIDINA BHAKTI PERSADA BANDUNG.
- Pratiwi, S. N. D., & Ulama, B. S. S. (2016). Klasifikasi Email Spam dengan Menggunakan Metode Support Vector Machine dan k-Nearest Neighbor. *Jurnal Sains Dan Seni ITS*, 5(2), 2337–3520.
- Sabry, F. (n.d.). *email spam*. One Billion Knowledgeable.