BAB I

PENDAHULUAN

1.1 Latar Belakang

Dalam era digital yang terus berkembang, email telah menjadi salah satu sarana komunikasi yang paling penting dan efisien. Namun, kemajuan ini juga diiringi dengan peningkatan serangan spam email yang semakin meresahkan. Spam email, yang merujuk pada pesan-pesan tidak diinginkan yang sering kali berisi penawaran palsu, tautan berbahaya, atau konten tidak senonoh, telah menjadi ancaman serius bagi produktivitas pengguna dan keamanan informasi. Dengan semakin kompleksnya teknik spamming yang digunakan oleh para penyerang, deteksi spam email menjadi semakin menantang.(Ahmed et al., 2022)

Meskipun sudah ada upaya untuk mengatasi masalah spam email melalui filter berbasis aturan dan metode deteksi heuristik, keberhasilannya masih terbatas. Dalam konteks ini, penggunaan machine learning telah menjadi fokus utama dalam mengembangkan solusi yang lebih adaptif dan responsif terhadap evolusi spam email. Dengan memanfaatkan kemampuan machine learning untuk mengenali pola-pola data yang kompleks, metode deteksi spam email dapat ditingkatkan secara signifikan.(Mashaleh et al., 2022)

Algoritma Naïve Bayes dan algoritma C45 telah menjadi dua pendekatan yang sering digunakan dalam konteks deteksi spam email. Algoritma Naïve Bayes, yang berdasarkan pada teorema probabilitas sederhana, telah terbukti efektif dalam berbagai aplikasi klasifikasi teks, termasuk dalam deteksi spam email. (Douzi et al., 2020)Di sisi lain, algoritma C45, yang mengandalkan pendekatan decision tree, mampu menghasilkan aturan keputusan yang dapat digunakan untuk mengklasifikasikan data. Perbandingan kinerja antara kedua algoritma ini diharapkan dapat memberikan wawasan yang lebih mendalam tentang

kemampuan masing-masing dalam mengenali pola-pola spam email dengan akurat.(Hartono et al., 2023)

Mengingat pentingnya keamanan informasi dalam era digital yang semakin kompleks, penelitian ini menjadi relevan dan penting dalam upaya untuk mengembangkan sistem keamanan yang lebih efektif dalam menghadapi tantangan spam email. Dengan memahami perbedaan antara algoritma Naïve Bayes dan algoritma C45, diharapkan penelitian ini akan memberikan wawasan yang lebih dalam tentang kelebihan dan kelemahan masing-masing algoritma dalam konteks deteksi spam email.(Rahma et al., 2021)

Selain itu, penelitian ini juga memiliki implikasi yang signifikan dalam pengembangan metodologi deteksi spam email yang lebih mutakhir. Diharapkan bahwa hasil penelitian ini dapat memberikan landasan bagi pengembangan strategi yang lebih adaptif dan responsif terhadap perubahan pola spam email. Dengan demikian, penelitian ini memiliki potensi untuk memperkaya pemahaman kita tentang penggunaan machine learning dalam konteks keamanan informasi dan aplikasi deteksi spam email.(Pane & Ikram, 2023)

Melalui penelitian ini, kami berharap dapat memberikan kontribusi yang signifikan bagi pengembangan teknologi deteksi spam email yang lebih canggih. Dengan membandingkan kinerja antara algoritma Naïve Bayes dan algoritma C45, diharapkan penelitian ini akan memberikan panduan yang berguna bagi para praktisi dan peneliti dalam memilih pendekatan yang paling sesuai dan efektif dalam memerangi spam email.(Azhari et al., 2021)

Dalam konteks implementasi, perlu diakui bahwa ada beberapa tantangan teknis yang perlu diatasi dalam pengembangan model deteksi spam email. Salah satu tantangan tersebut adalah dalam memastikan keandalan dan ketepatan model dalam mengenali spam email dari variasi besar pola dan konten yang digunakan oleh para pengirim spam. Oleh karena itu, penggunaan dataset yang representatif dan teknik validasi yang

cermat akan menjadi kunci kesuksesan dalam pengembangan model deteksi yang andal.

Metodologi yang akan digunakan dalam penelitian ini akan didasarkan pada pendekatan eksperimental yang menggabungkan pengumpulan data, implementasi algoritma, dan evaluasi kinerja model. Data yang digunakan akan berupa kumpulan email yang telah di-annotate secara manual, mencakup contoh-contoh spam email dan non-spam email. Kedua algoritma, Naïve Bayes dan C45, akan diimplementasikan dan dievaluasi menggunakan berbagai metrik performa, termasuk akurasi, presisi, recall, dan F1-score. (Azhari et al., 2021)Penggunaan teknik validasi silang akan dilakukan untuk memastikan keandalan dan generalisasi model yang dihasilkan.

Melalui penggunaan pendekatan eksperimental yang cermat dan komprehensif, diharapkan penelitian ini akan menghasilkan temuan yang signifikan dalam pengembangan sistem deteksi spam email yang lebih canggih dan handal. Dengan demikian, penelitian ini tidak hanya akan memberikan kontribusi pada pengetahuan akademis dalam bidang keamanan informasi, tetapi juga akan memiliki implikasi praktis yang signifikan bagi pengembangan sistem keamanan informasi di berbagai sektor. Diharapkan bahwa hasil penelitian ini akan memberikan panduan yang berguna bagi praktisi keamanan informasi dalam memilih dan mengimplementasikan metode deteksi spam email yang paling efektif dan andal dalam lingkungan yang terus berkembang dan berubah. Dengan demikian, penelitian ini diharapkan dapat menjadi langkah maju yang penting dalam upaya global untuk mengatasi masalah spam email dan meningkatkan keamanan informasi dalam dunia digital yang terus berkembang.(Hidayat, 2022)

1.2 Identifikasi Masalah

- beberapa identifikasi masalah yang mungkin dapat diidentifikasi adalah sebagai berikut:
- Bagaimana cara mengimplementasi dan mengevaluasi kinerja dari algoritma Naïve Bayes dan algoritma C45 dalam konteks deteksi spam email?
- Apakah penelitian ini mempertimbangkan faktor – faktor eksternal yang mungkin mempengaruhi kinerja algoritma?
- Apakah aspek implementasi teknis seperti waktu komputasi dan penggunaan sumber daya menjadi tujuan utama dalam penelitian ini?

1.3 Tujuan Penelitian

beberapa tujuan yang mungkin ingin dicapai adalah sebagai berikut:

1. Mengevaluasi Efektivitas Kinerja Algoritma
2. Membandingkan Akurasi dan Ketepatan dari Kedua Algoritma
3. Menganalisis Kelebihan dan Kekurangan Masing-Masing Algoritma
4. Mengidentifikasi Algoritma yang Paling Optimal untuk Deteksi Spam Email
5. Memberikan Panduan Praktis bagi Praktisi Keamanan Informasi

1.4 Batasan Masalah

Dalam melakukan penelitian ini, ada beberapa batasan yang perlu dipertimbangkan sebagai berikut

1. Penelitian ini akan memfokuskan pada implementasi dan evaluasi kinerja dari algoritma Naïve Bayes dan algoritma C45 dalam konteks deteksi spam email.
2. Penelitian tidak akan mempertimbangkan faktor-faktor eksternal yang mungkin mempengaruhi kinerja algoritma, seperti perubahan

tren spam email dari waktu ke waktu atau variabilitas dari sumber data yang digunakan

3. Aspek implementasi teknis seperti waktu komputasi dan penggunaan sumber daya tidak akan menjadi fokus utama dalam penelitian ini.

1.5 Kontribusi Penelitian

Diharapkan bahwa hasil dari penelitian ini akan memberikan kontribusi yang signifikan dalam pengembangan teknik deteksi spam email yang lebih canggih dan efektif. Melalui perbandingan antara algoritma Naïve Bayes dan algoritma C45, penelitian ini diharapkan dapat memberikan wawasan yang lebih mendalam mengenai kinerja kedua algoritma dalam konteks deteksi spam email. Selain itu, hasil penelitian ini diharapkan dapat memberikan panduan praktis bagi pengembang sistem keamanan informasi dalam memilih algoritma yang paling sesuai dan efektif dalam memerangi spam email. Dengan demikian, penelitian ini memiliki potensi untuk memberikan kontribusi yang signifikan bagi perkembangan ilmu pengetahuan dan teknologi di bidang keamanan informasi.