BAB 1
PENDAHULUAN

1.1 Latar Belakang Masalah

Seiring berkembangnya zaman, teknologi semakin mempermudah kegiatan manusia. Termasuk dalam bidang komunikasi yang dulu sebelum *handphone* dapat mudah dimiliki oleh semua kalangan masyarakat, komunikasi dilakukan melalui berbagai cara seperti surat dan wartel. Meskipun dengan cara-cara ini orang dapat melakukan komunikasi jarak jauh, mereka memiliki kekurangan yang signifikan dibandingkan dengan kemudahan yang ditawarkan oleh *handphone*. Seperti surat yang memerlukan waktu yang lama untuk sampai ke tujuan. Wartel yang sering kali terletak jauh dari lokasi pengguna dan tidak selalu tersedia, membuatnya kurang praktis jika dibandingkan dengan *handphone* yang memungkinkan komunikasi instan, baik melalui panggilan suara maupun pesan teks, di mana saja dan kapan saja.

Sekarang, dengan hadirnya berbagai merk dan jenis *handphone* dengan harga yang terjangkau, makin banyak orang yang dapat melakukan komunikasi jarak jauh dengan mudah. Dengan semakin banyaknya pengguna *handphone*, maka risiko penipuan juga meningkat, salah satu jenis penipuan yang sering di temui adalah dengan pesan teks yang mengiming-imingi hadiah lalu mengarahkan pengguna ke link *website phishing* yang menyerupai perusahaan ternama atau menghubungi nomor telepon, setelah itu pengguna akan diarahkan untuk memberikan informasi rahasia seperti kode *One Time Passcode* (OTP) atau mengunduh aplikasi *malware* yang dapat mengambil data-data pribadi mereka yang kemudian digunakan oleh penipu untuk mengakses rekening pribadi atau melakukan pinjaman *online* atas nama korban itu. Penipuan yang dilakukan melalui pesan teks dapat diklasifikasikan sebagai serangan *social engineering*, bahkan 99% penipuan perbankan di Indonesia adalah akibat serangan *social engineering*[1].

Pada tahun 2022 hingga awal tahun 2024 serangan *social engineering* mengakibatkan kerugian sebesar Rp 2,5 T berdasarkan 155 ribu aduan konsumen dari 10 bank pada Otoritas Jasa Keuangan (OJK)[2]. Serangan *social engineering* masing sering kali terjadi dan berhasil di Indonesia disebabkan oleh kurangnya edukasi dan literasi di kalangan masyarakat tentang modus-modus penipuan yang sering beredar, sehingga mereka menjadi korban. Akibatnya, kepercayaan orang-

orang terhadap layanan perbankan menurun, dan mereka enggan untuk menyimpan uangnya di bank dan melakukan transaksi secara *online*. Hal ini dapat menyebabkan penurunan likuiditas di sektor perbankan, yang lalu mempengaruhi kemampuan bank untuk memberikan pinjaman dan mendukung pertumbuhan ekonomi. Selain itu, masyarakat yang memilih untuk menyimpan uang secara tunai atau di tempat yang tidak resmi berisiko lebih tinggi terhadap pencurian atau kehilangan.

Social engineering adalah teknik manipulasi psikologis yang dilakukan untuk menipu korban agar mengungkapkan informasi rahasia atau melakukan suatu tindakan yang tidak diinginkan. *Phishing* merupakan salah satu bentuk serangan *social engineering* yang paling umum terjadi. Berdasarkan laporan IDADX, total pengaduan serangan *phishing* di Indonesia mengalami peningkatan signifikan, dari 6.106 laporan pada kuartal IV tahun 2022 menjadi 26.675 laporan pada kuartal I tahun 2023. Artinya, terjadi lonjakan sebanyak 20.569 laporan dalam waktu yang relatif singkat [3]. Pesan teks menjadi media yang rentan terhadap teknik *phishing* ini, dan memerlukan pendekatan proaktif untuk deteksi dini. Meskipun ada beberapa langkah untuk mencegah dan melaporkan serangan *phishing*, deteksi otomatis dari pesan teks yang memiliki risiko penipuan masih kurang. Sistem deteksi saat ini masih sangat bergantung pada laporan pengguna, yang mungkin tidak selalu tepat waktu atau akurat. Oleh karena itu, penelitian ini bertujuan untuk mengatasi masalah tersebut dengan mengembangkan Implementasi Algoritma Naive Bayes untuk Klasifikasi Teks Phishing.

Algoritma *Naïve Bayes* dipilih karena berdasarkan penelitian-penelitan terdahulu yang juga melakukan klasifikasi teks, seperti yang dilakukan oleh Michael Chen pada tahun 2023 [4] melakukan klasifikasi teks *scam* menggunakan metode *Federated Learning* dengan model *gated recurrent unit* (GRU), jumlah dataset 1,780 dengan pembagian 70% untuk *training* dan *validation*, serta 30% untuk *testing* mendapatkan akurasi sebesar 55,05%. Penelitian lainnya oleh Ahmad Zamsuri pada tahun 2023 [5] melakukan klasifikasi jenis emosi dalam teks menggunakan algoritma *K-Nearest Neighbor* (KNN) 2 Label dengan TF-IDF, jumlah dataset sekitar 1,600 dengan dengan pembagian data *training* dan *testing* 80%:20% mendapatkan akurasi sebesar 77%. Sedangkan penelitian terdahulu oleh Ade Clinton Sitepu pada tahun 2021 [6] melakukan klasifikasi teks *bullying* menggunakan algoritma *Naïve Bayes*, jumlah dataset 1,000 dengan pembagian data *training* dan *testing* 60%:40% mendapatkan akurasi sebesar 88%. Serta penelitian oleh Alrico Rizki Wibowo pada tahun 2024 [7] melakukan analisis sentimen teks berita menggunakan algoritma *Naïve Bayes* dengan jumlah dataset

1,886 dengan pembagian data *training* dan *testing* 90%:10% mendapatkan akurasi sebesar 94.64%. Maka dari itu, untuk mengembangkan Implementasi Algoritma Naive Bayes untuk Klasifikasi Teks Phishing ini saya memilih untuk menggunakan algoritma *Naïve Bayes*.

1.2 Rumusan Masalah

1. Bagaimana cara mengimplementasi algoritma *Naïve Bayes* dalam melakukan klasifikasi teks *phishing* pada pesan teks?
2. Berapa jumlah akurasi penerapan algoritma *Naïve Bayes* dalam melakukan klasifikasi teks *phishing* pada pesan teks?

1.3 Batasan Permasalahan

1. *Dataset* yang digunakan diambil dari penelitian sebelumnya yang berisi pesan teks dari *Short Message Service* (SMS) dan dimodifikasi dengan *dataset* tambahan tersebut dari *repository* di *Mendeley Data* dan *GitHub*.
2. Model *Naïve Bayes* yang digunakan merupakan model yang telah dilatih dengan dataset berbahasa Indonesia dan Inggris yang berasal dari pesan teks.

1.4 Tujuan Penelitian

1. Mengetahui cara mengimplementasikan algoritma *Naïve Bayes* dalam melakukan klasifikasi teks *phishing* pada pesan teks.
2. Mengetahui jumlah akurasi algoritma *Naïve Bayes* dalam melakukan klasifikasi teks *phishing* pada pesan teks.

1.5 Manfaat Penelitian

1. Hasil dari penelitian ini dapat membantu meningkatkan keamanan platform komunikasi digital dengan memberikan metode yang efektif untuk mendeteksi serangan *phishing* pada pesan teks.
2. Hasil dari penelitian ini dapat membantu mengurangi risiko karena dapat mengidentifikasi karakteristik dan pola teks yang mencirikan serangan *phishing*

1.6 Sistematika Penulisan

Berisikan uraian singkat mengenai struktur isi penulisan laporan penelitian, dimulai dari Pendahuluan hingga Simpulan dan Saran.

Sistematika penulisan laporan adalah sebagai berikut:

- Bab 1 PENDAHULUAN
  Bab ini berisi tentang latar belakang masalah, rumusan masalah, batasan masalah, tujuan dan manfaat penelitian, serta cara melakukan penulisan sistematika pada laporan skripsi akan dijelaskan dalam bab ini.
- Bab 2 LANDASAN TEORI
  Bab ini berisi tentang teori dan studi yang digunakan dalam penelitian yaitu algoritma *Naïve Bayes* dan *phishing* dalam pesan teks.
- Bab 3 METODOLOGI PENELITIAN
  Bab ini berisi tentang metodologi dan rancangan yang digunakan dalam pembuatan model algoritma *Naïve Bayes*.
- Bab 4 HASIL DAN DISKUSI
  Bab ini berisi tentang hasil dari implementasi yang telah dilakukan berdasarkan metodologi dan rancangan yang telah dituliskan.
- Bab 5 KESIMPULAN DAN SARAN
  Bab ini berisi tentang kesimpulan akhir dari semua rangkaian penelitian yang telah dilakukan serta memberikan saran-saran untuk penelitian yang akan datang.