BAB I
PENDAHULUAN

1.1 Latar Belakang Penelitian

Perkembangan teknologi digital di Indonesia tidak terlepas dari bagaimana seluruh dunia menghadapi pandemi Covid-19 pada tahun 2020. Dilansir dari *website* resmi Asosiasi Penyelenggara Jasa Internet Indonesia (APJII), yang mengutip platform berita Detik, Sebelum terjadinya pandemi, tingkat penetrasi internet di Indonesia pada tahun 2018 – 2020 hanya bernilai 64,8 persen. Namun, tingkat persentase penetrasi internet mengalami peningkatan hingga 10% pada tahun 2020, dan terus berlanjut hingga tahun 2024, di mana tingkat penetrasi internet di Indonesia telah mencapai 79,5%. Peningkatan jumlah yang signifikan pada tahun 2020 menunjukkan bahwa pandemi memaksa seluruh lapisan masyarakat untuk mulai beradaptasi dan lebih banyak menggunakan jaringan internet dalam aktivitas sehari – hari. Kebijakan pemerintah yang meminta masyarakat untuk mengubah kegiatan tatap muka secara langsung menjadi kegiatan tatap muka secara daring, mengharuskan masyarakat untuk menggunakan jaringan internet secara ekstensif dibanding sebelum terjadinya pandemi.

Integrasi pada kegiatan sehari – hari masyarakat dengan berbagai macam teknologi dan jaringan internet tentu memiliki dampak yang mengubah cara hidup masyarakat Indonesia. Namun, seiring berkembangnya integrasi kehidupan masyarakat dengan penggunaan teknologi yang masif, Perkembangan ini tidak dibarengi dengan sikap dan etika yang baik saat menggunakan jaringan internet. Topik keamanan jaringan internet dan literasi digital yang penting untuk diketahui oleh semua pengguna jaringan internet, menjadi sebuah pengetahuan yang jarang dimiliki. Padahal, di zaman yang semuanya serba digital, data yang bersikulasi di dalam aktivitas sehari – hari masyarakat merupakan sebuah komoditas berharga.

Peningkatan angka penetrasi internet di Indonesia, yang tidak dibarengi dengan pengetahuan dasar penggunaan internet, menyebabkan terjadinya kejahatan melalui jaringan internet yang dikenal sebagai *Cyber Crime*. Menurut Laporan BSSN pada tahun 2021, Jumlah serangan siber yang terjadi pada tahun 2020 – 2021 melompat jauh hingga menyentuh angka 621.167.829 serangan dibandingkan dengan serangan yang terjadi pada tahun 2019 yang hanya menyentuh angka

99.808.361 (Chirzah, 2023). Temuan ini juga diperkuat oleh pernyataan dari APWG, *The Anti-Phising Working Group,* mengenai jumlah serangan yang meningkat selama pandemi Covid-19. Jumlah serangan melalui jaringan internet yang sangat masif tersebut terdiri dari berbagai macam jenis dan bentuk serangan, di antaranya adalah Spam dan *Phising*.

Phising menurut APWG adalah kegiatan mencuri informasi – informasi yang rahasia dan penting untuk seseorang menggunakan metode, teknik, dan alat yang mutakhir. *Phising* dapat dilakukan melalui *email (email phising)*, *online social media*, dan aplikasi *mobile* (Catal C dkk, 2022). *P*ara pelaku kriminal siber menggunakan kamuflase identitas palsu yang menyerupai orang di sekeliling calon korban atau menggunakan metode teknis seperti membuat *website* palsu yang menyerupai *website* resmi dengan tujuan memancing calon korban untuk memasukkan data – data penting. Para penyerang bahkan menggunakan SSL (*Secure Sockets Layer)* untuk menipu penggunanya.

Disisi yang lain, Spam adalah sebuah bentuk interaksi antar pengguna jaringan internet yang melibatkan pertukaran pesan. Umumnya, interaksi ini melibatkan seorang pengguna dengan sebuah perusahaan yang sedang melakukan *advertising* perusahaannya. Sayangnya, pesan – pesan yang dikirimkan kepada korban merupakan pesan yang diklasifikasikan sebagai *unwanted messages*, di mana jasa dan barang yang ditawarkan kepada penerima cenderung memaksa atau merupakan sesuatu yang tidak dibutuhkan oleh sang penerima. Bahkan, pesan Spam cenderung mengandung konten – konten tidak pantas, seperti pornografi, pesan palsu, berita palsu, dan *malware* berbahaya untuk sang penerima pesan Spam (Sanjeev Rao, 2021).

Xiaoxu Liu (2021) menyebutkan dalam penelitiannya, bahwa Spam banyak terjadi pada *SMS* atau aplikasi yang ada di dalam aplikasi *mobile*. *Drunk message*, sebutan dari pesan – pesan Spam, dapat diartikan sebagai pesan – pesan yang berisi konten yang tidak relevan dengan penerima pesan yang dikirim melalui *SMS*. Hal ini disebabkan oleh tingginya jumlah pengguna layanan *SMS* atau aplikasi yang terhubung dengan jaringan internet lainnya. Selain itu, murahnya harga dan sumber untuk melakukan pesan Spam menyebabkan mudahnya para penyerang untuk menggunakan teknik serangan Spam. Walaupun terdapat beberapa layanan khusus

yang disediakan oleh manufaktur *mobile phone*, limitasi sumber daya komputasi di kebanyakan perangkat pengguna gagal untuk melakukan filtrasi secara tepat akurat.

Kejahatan Spam dan *Phising* berkembang ke dalam bentuk yang lebih berbahaya dan mulai tidak dapat dibedakan oleh pengguna. Serangan Spam dan atau *Phising* terjadi tidak hanya melalui *SMS* dan *Email,* Tetapi juga melalui sosial media *Whatsapp*. Penyerang berkamuflase sebagai saudara atau teman yang menjadi kurir paket, mengirimkan dokumen yang jika ditelusuri lebih lanjut merupakan sebuah *file* dengan ekstensi *.apk* yang dibuat dengan tujuan untuk melakukan *Phising* terhadap pengguna.

Sayangnya, perilaku para pengguna internet di Indonesia seakan – akan mengindahkan bahaya dari Spam dan *Phising*. Menurut penelitian yang dilakukan oleh Akmal Hidayat pada tahun 2023, Sebagian besar mahasiswa di Indonesia yang telah memiliki perangkat komputer tidak melakukan praktik – praktik sederhana untuk mengamankan data – data penting responden dan tidak tahu bahwa data yang disimpan di dalam komputer mereka dapat dicuri. Responden pada penelitian tersebut juga menyatakan bahwa responden merasa kesulitan untuk membedakan bentuk serangan yang datang terhadap mereka. Sebagian responden akan membuka *e-mail* dari pengirim yang tidak dikenal, apabila *e-mail* tersebut terlihat menarik.

Walaupun begitu, kajian terhadap *Cyber Law,* hukum yang diberlakukan pada jaringan dan teknologi, menjadi kabar baik untuk penanggulangan dan perlindungan pengguna internet di Indonesia. Namun, penerapan *Cyber Law* di Indonesia masih terkendala beberapa faktor yang menyebabkan kesalahan dan kebingungan dalam penegakkan hukum Indonesia pada bidang teknologi (Riko Nugraha, 2021).

Faktor pertama adalah kejahatan yang dilakukan melalui jaringan internet memerlukan adanya alat khusus untuk melacak keberadaan pelaku kejahatan untuk kemudian ditangkap dan diadili secara sah di depan hukum yang berlaku. Namun, penggunaan terhadap alat khusus itu juga perlu ilmu pengetahuan yang cukup dan memadai agar prosesnya dapat berjalan dengan baik. Faktor kedua adalah belum jelasnya batasan daerah yurisdiksi tempat penyerang melakukan kejahatannya. Faktor ketiga adalah adanya norma – norma hukum terkait yang perlu dikaji kembali lebih jauh untuk menyesuaikan dengan hukum KUHP.

Perilaku pengguna jaringan internet yang belum tepat dan penegakkan hukum jaringan (*Cyber Law*) di Indonesia tentunya menjadi sebuah permasalahan yang tidak dapat diabaikan. Sehingga, diperlukan adanya aplikasi yang dapat membantu para pengguna jaringan internet untuk membantu para pengguna untuk mengamankan data – data selama pengguna menggunakan jaringan internet.

Proses pengembangan dan penelitian untuk aplikasi – aplikasi yang membantu pengguna meningkatkan keamanan dan kenyamanan pengguna dalam menggunakan jaringan internet telah banyak dilakukan. Salah satu topik penelitian pada bidang keamanan adalah dengan memanfaatkan bidang keilmuan *Artificial Intelligence* (AI) dalam menanggulangi, mengatasi, dan memperbaiki dampak serangan siber. Salah satu penelitian yang mengangkat topik *machine learning* sebagai alat pertahanan untuk mengatasi serangan siber adalah penelitian yang dilakukan pada tahun 2020 oleh Hongli Yuan.

Hongli Yuan dalam penelitiannya, mengimplementasikan metode pra – processing data TF – IDF dan beberapa model algoritma *machine learning* dengan tujuan untuk melakukan pemeriksaan terhadap *file* berformat .apk untuk sistem operasi android. Penelitian ini dilakukan berdasarkan temuan bahwa Sistem operasi android dan perangkat kerasnya adalah perangkat dengan jumlah pengguna yang masif. Dengan banyaknya pilihan aplikasi yang dapat diunduh oleh pengguna, hal ini juga berdampak pada pertumbuhan aplikasi *malware* yang berbahaya dan dapat mengancam keamanan perangkat pengguna.

Peneliti Hongli Yuan membandingkan beberapa algoritma klasifikasi *machine learning* yaitu, Naive Bayes (NB), Bayesian Network (BN), Random Tree (RT), J48, Random Forest (RF), K-Nearest Neighbor (K-NN), dan K-fold. Metode yang digunakan peneliti dalam memeriksa aplikasi yang dicurigai sebagai *malware* adalah dengan menggunakan model yang telah dilatih menggunakan dataset yang diproses dengan TF-IDF untuk memeriksa berkas *androidmanifest.xml* yang tersimpan di dalam setiap *file* berformat .apk. Dengan parameter *required permission* yang ada pada berkas manifes aplikasi, maka aplikasi dapat dikategorikan menjadi *malware* atau aplikasi aman menggunakan model yang telah dibuat. Pada penelitian tersebut, disimpulkan bahwa menggunakan metode evaluasi model *percentage split* dan *K-fold cross validation*, algoritma klasifikasi yang

memiliki nilai evaluasi tertinggi berdasarkan nilai TPR, FPR, Presisi, Akurasi, *Recall*, F-M, dan waktu adalah model J48, dengan Random Forest sebagai model kedua yang memiliki nilai evaluasi tertinggi setelah J48.

Pemanfaatan algoritma *machine learning* dan metode *pre-processing* data TF - IDF pada keamanan jaringan tidak terbatas pada identifikasi aplikasi *malware*. Penelitian juga dilakukan untuk membandingkan metode *pre-processing* TF – IDF dan berbagai algoritma *machine learning* untuk mengklasifikasi *Short Message Services* (SMS) berdasarkan parameter apakah pesan yang dikirim tersebut merupakan pesan Spam atau bukan (Nilam N, 2019). Di dalam penelitian tersebut, kesimpulan yang dihasilkan menunjukkan bahwa Metode TF-IDF saat dipadukan dengan algoritma model *classifier* Random Forest memberikan nilai evaluasi yang sangat tinggi, berdasarkan akurasi, presisi, dan F-Measure. Saat dibandingkan dengan algoritma klasifikasi Multinomial Naïve Bayes (MNB), K-Nearest Neighbour (KNN), Support Vector Machine (SVM), Decision Tree (DT), Algoritma klasifikasi Random Forest menjadi algoritma yang paling akurat dalam mengklasifikasi apakah sms yang diterima pengguna merupakan pesan Spam atau pesan yang aman.

Penelitian – penelitian yang telah disebutkan sebelumnya menunjukkan bahwa pemanfaatan metode pemrosesan TF – IDF dan model algoritma klasifikasi Random Forest memiliki potensi yang sangat luas untuk dimanfaatkan di dalam bidang keamanan jaringan. Sehingga diperlukan adanya aplikasi yang dapat melakukan filtrasi terhadap pesan yang masuk menggunakan metode TF-IDF dan Random Forest yang telah menjadi topik penelitian pada penelitian terdahulu. Berfokus pada permasalahan pesan Spam dan *Phising* yang terjadi pada aplikasi pesan Whatsapp, Peneliti akan menggunakan metode TF-IDF untuk memproses data pesan yang telah ada untuk selanjutnya diklasifikasikan dengan model klasifikasi Random Forest, yang dinilai menjadi model yang akurat, dan presisi.

Aplikasi yang dibuat pada sistem operasi android diharapkan dapat membantu para pengguna perangkat *mobile* android yang menggunakan aplikasi Whatsapp untuk terhindar dari pesan – pesan yang mengandung konten – konten berjenis Spam dan mengandung *Phising* yang membahayakan keamanan data pengguna. Aplikasi yang dibuat akan mengintegrasikan model *machine learning*

dengan metode *pre-processing* data menggunakan TF-IDF dan algoritma klasifikasi Random Forest.

1.2 Rumusan Masalah Penelitian

Berdasarkan latar belakang yang telah dipaparkan sebelumnya, Peneliti merumuskan beberapa rumusan masalah dalam penelitian ini di antaranya:

1. Bagaimana metode TF – IDF dan algoritma *machine learning* Random Forest dapat melakukan Filtrasi terhadap pesan Spam dan *Phising*?
2. Bagaimana mengintegrasikan aplikasi android dengan model algoritma *machine learning* untuk melakukan filtrasi pesan Spam dan *Phising* Whatsapp?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dirumuskan sebelumnya, maka penelitian ini bertujuan untuk memenuhi beberapa tujuan sebagai berikut:

1. Mengembangkan *supervised machine learning* model menggunakan metode *pre-processing* TF-IDF dan algoritma *machine learning* Random Forest melakukan filtrasi terhadap Spam dan *Phising* menggunakan dataset SMS, *e-mail*, dan teks yang telah dilabeli sebagai pesan Spam*, Phising,* dan pesan aman.
2. Membangun sebuah aplikasi android yang mengintegrasikan *machine learning* sebagai platform pengembangan aplikasi – aplikasi lainnya dalam menghadapi kejahatan siber berbasiskan TF-IDF dan/atau *machine learning* lainnya.

1.4 Manfaat Penelitian

Berdasarkan tujuan penelitian yang telah dipaparkan sebelumnya, diharapkan penelitian ini dapat bermanfaat bagi perkembangan teknologi terutama di bidang *machine learning* dan *network security*. Berikut beberapa manfaat dari penelitian ini di antaranya:

1.4.1. Manfaat Teoritis

Diharapkan manfaat secara teoritis dalam penelitian ini sebagai berikut:

1. Memberikan referensi yang baru bagi penelitian dengan tema dan topik yang sama dengan harapan dapat mengembangkan penelitian ini menjadi lebih baik dan lebih bermanfaat lagi, serta dapat dikaji lebih baik lagi untuk penelitian – penelitian ke depannya

2. Menciptakan inovasi dalam mengintegrasikan *network security* dan *machine learning* dalam menciptakan aplikasi pertahanan serangan siber sehingga menjadi ruang gerak baru bagi para peneliti selanjutnya untuk mengembangkan atau menciptakan aplikasi – aplikasi lainnya.

1.4.2. Manfaat Praktis

Selain manfaat teoritis, terdapat manfaat praktis dalam penelitian ini, di antaranya:

1. Bagi pengguna aplikasi *chat* Whatsapp, aplikasi ini akan sangat bermanfaat untuk secara langsung memfiltrasi pesan – pesan yang masuk ke dalam aplikasi Whatsapp, agar pengguna tidak perlu membuka pesan mencurigakan yang diterima.
2. Bagi pengembang aplikasi *chat*, Aplikasi ini bisa menjadi dasar bagi para pengembang untuk mengintegrasikan secara langsung fitur filtrasi yang didasarkan pada metode TF-IDF dan algoritma *machine learning* ke dalam aplikasinya.
3. Bagi peneliti, peneliti dapat mengimplementasikan pembelajaran dan pengalaman yang telah didapatkan selama masa perkuliahan berlangsung. Selain itu, Peneliti dapat mengembangkan jiwa kreatif dan inovatif peneliti dengan memadukan beberapa keilmuan secara sekaligus, yaitu *machine learning* dan *network security* yang tentunya akan bermanfaat bagi banyak pihak.

1.5 Batasan Penelitian

Batasan pada penelitian ini ditetapkan berdasarkan landasan – landasan berikut :

1. Penelitian dilakukan menggunakan dataset yang didapatkan dari penelitian – penelitian terdahulu yang dilakukan terhadap topik penelitian sejenis. Data dapat berupa pesan – pesan yang dikirim melalui SMS dan *e-mail*. Data juga datang dari isi pesan aplikasi pengguna dan pesan yang tersebar luas dan ditandai sebagai pesan dengan konten *Phising*.
2. Penelitian akan difokuskan pada pengembangan model filtrasi pesan Spam dan *Phising* menggunakan metode *pre-processing* TF – IDF dan algoritma Random Forest.
3. Proses klasifikasi pesan yang dianggap *Phising* dilakukan berdasarkan teks yang ditulis di dalam pesan yang diterima bukan isi pesan yang diterima (aplikasi, dokumen, gambar, dan video).

4. Pengujian akan dilakukan terhadap keberhasilan integrasi model yang telah dikembangkan dengan aplikasi berbasis sistem operasi android. Aplikasi tidak akan melakukan *intervensi* pada lapisan aplikasi Whatsapp melainkan mengakses *item* pada notifikasi yang dikirim oleh Whatsapp
5. Keberhasilan pengujian aplikasi diukur berdasarkan metode *black-box* yang mengevaluasi apakah *output* yang diberikan aplikasi yang telah terintegrasi dengan model sesuai dengan apa yang diharapkan oleh peneliti berdasarkan data sampel uji yang disiapkan oleh peneliti sebelumnya.

1.6 Struktur Organisasi Skripsi

Berdasarkan pada Peraturan Rektor UPI (Universitas Pendidikan Indonesia) Nomor.7867/UN40/HK/2021 tentang Pedoman Penulisan Karya Ilmiah Universitas Pendidikan Indonesia Tahun Akademik 2021, maka, Penelitian yang berjudul "Aplikasi Filtrasi Spam dan *Phising* Pesan Whatsapp Dengan Metode Term Frequency Inverse Document Frequency Dan Machine Learning" terdiri dari lima bab. Adapun gambaran mengenai penjelasan kelima bab tersebut dijelaskan ke dalam sistematika sebagai berikut:

A. PENDAHULUAN

Pada bab I menjelaskan tentang latar belakang penelitian, rumusan masalah, tujuan penelitian, manfaat penelitian, dan struktur organisasi penelitian.

B. KAJIAN PUSTAKA

Pada bab II menjelaskan mengenai kajian pustaka, kerangka pemikiran, dan penelitian terdahulu. Kajian pustaka terdiri dari *machine learning*, pemanfaatan metode *feature extraction* TF – IDF data teks, implementasi Random Forest sebagai algoritma klasifikasi, Spam, *Phising* dan aplikasi Whatsapp,

C. METODE PENELITIAN

Pada bab III menjelaskan mengenai uraian metode penelitian yang digunakan. Meliputi objek penelitian, metode penelitian, sumber data, dan metode pengujian sistem.

D. TEMUAN DAN PEMBAHASAN

Pada bab IV ini menjelaskan mengenai hasil dan pembahasan dari

penelitian yang dilakukan yang merujuk kepada rumusan masalah.

E. SIMPULAN, IMPLIKASI, DAN REKOMENDASI

Pada bab V ini menjelaskan terkait kesimpulan dari penelitian yang telah dilakukan oleh peneliti serta implikasi dan rekomendasi. Peneliti juga memberikan saran sebagai bentuk rekomendasi dari temuan yang ada di lapangan.