

## BAB III

### METODE PELAKSANAAN

#### 3.1 Tahap Pelaksanaan Program

Berikut merupakan tahapan-tahapan yang dilakukan dalam penelitian ini:

a. Perumusan Masalah

Topik penelitian yang telah dipilih adalah pengembangan dan pembuatan ekstensi browser untuk deteksi situs phishing dengan menggunakan algoritma *Random Forest*.

b. Kajian Kepustakaan

Dengan merinci penelitian yang dipilih, langkah selanjutnya melibatkan kajian literatur dengan tujuan mendapatkan informasi dan teori-teori yang terkait dengan permasalahan tersebut. Ini mencakup *review* terhadap jurnal-jurnal yang membahas topik sejenis atau terkait yang telah dilakukan sebelumnya. Proses ini dilakukan untuk memperdalam pemahaman teoritis terkait algoritma *Random Forest* dan situs *phishing*, yang menjadi dasar teori bagi penelitian ini.

c. Mencari *Dataset*

Data yang digunakan dalam penelitian ini menggunakan *dataset* yang diperoleh dari laman Kaggle.com *Phishing Website Detector*. *Dataset* ini berisikan 11054 entitas data dengan 32 atribut. Setiap entitas data telah dilabeli dengan -1 yang berarti bukan *phishing* dan 1 yang berarti *phishing*.

d. Pembuatan Program Sistem

Setelah menentukan topik permasalahan yang akan diangkat, langkah berikutnya adalah melakukan analisis untuk memilih algoritma yang tepat untuk implementasi. Dalam penelitian ini, diputuskan untuk menggunakan algoritma *Random Forest*. Proses selanjutnya melibatkan pengembangan

sebuah *script Python* yang dirancang khusus untuk mengekstraksi fitur dari URL. Ekstraksi fitur ini merupakan langkah awal untuk memahami karakteristik dan informasi yang terkandung dalam setiap URL yang dianalisis. Setelah fitur-fitur berhasil diekstraksi, langkah berikutnya adalah menerapkan algoritma *Random Forest* untuk melakukan klasifikasi terhadap hasil ekstraksi fitur tersebut. Algoritma *Random Forest*, yang terkenal karena kemampuannya dalam menangani berbagai jenis data dan tugas klasifikasi, dipilih untuk menghasilkan model yang dapat memprediksi atau mengklasifikasikan data dengan akurasi yang tinggi. Dengan menggabungkan ekstraksi fitur dan model *Random Forest*, penelitian ini bertujuan untuk meningkatkan pemahaman dan kemampuan dalam mengidentifikasi pola atau karakteristik tertentu dari URL yang dapat dijadikan dasar untuk klasifikasi lebih lanjut. Hasil dari implementasi ini diharapkan dapat memberikan kontribusi dalam pemecahan masalah yang terkait dengan topik yang telah dipilih.

#### e. Implementasi Model ke Sistem

Pada tahapan ini peneliti melakukan implementasi model hasil algoritma *Random Forest* ke dalam sistem *browser extension*. Pemodelan tersebut diproses dan dilakukan oleh rekan peneliti yang berpartisipasi dalam pengerjaan penelitian ini secara keseluruhan, dimana rekan dari peneliti tersebut berfokus pada proses analisis data hingga pembuatan model menggunakan algoritma *Random Forest*. Pemodelan menggunakan algoritma ini akan menghasilkan satu *file* dengan tipe *pickle*, yang akan diimplementasikan ke dalam kodingan *website* sebagai model klasifikasi *link* yang diinput oleh *user* melalui GUI *website*.

#### f. Evaluasi User

Setelah *prototype* sistem selesai dibuat, maka tahap selanjutnya adalah melakukan evaluasi sistem terhadap *user*. Untuk *user*, peneliti akan memilih 3 orang yang akan berperan sebagai *user*. Ketiga orang tersebut nantinya akan diberikan *form* terkait beberapa hal yang sekiranya dapat dievaluasi terkait sistem yang telah dibuat oleh peneliti. Adapun beberapa poin evaluasi yang akan diberikan berkaitan dengan *user experience* ketika *user* menggunakan *browser extension phishing detection* ini. Berikut beberapa hal yang akan mendapatkan evaluasi dari pihak *user*, yaitu:

- a. Bagaimana kemudahan dalam menggunakan *browser extension* untuk keperluan deteksi *phishing*?
- b. Bagaimana kejelasan informasi terkait hasil dari deteksi URL?
- c. Bagaimana kecepatan *browser extension* dalam memproses deteksi URL?

Ketiga pertanyaan evaluasi tersebut akan dibuat dalam bentuk *form* yang dapat dijawab dengan menggunakan parameter angka dengan rentang 1 sampai dengan 5. Poin akhir yang didapat berdasarkan hasil evaluasi akan menentukan apakah perlu ada perbaikan pada *browser extension* supaya lebih informatif dan *user-friendly*.

#### g. Penyusunan Laporan Penelitian

Langkah terakhir dalam penelitian ini adalah menyusun laporan yang menggambarkan hasil keseluruhan dari penelitian yang telah dilakukan. Laporan ini berfungsi sebagai dokumentasi lengkap dari seluruh proses penelitian. Penyusunan laporan dilakukan secara terstruktur sesuai dengan pedoman penyusunan dan penulisan laporan ilmiah yang telah ditetapkan sebelumnya. Laporan ini mencakup semua tahapan penelitian, metode yang digunakan, hasil eksperimen, interpretasi data, dan temuan utama yang ditemukan selama penelitian. Dengan menyusun laporan ini, diharapkan

informasi yang diperoleh dari penelitian dapat diakses dan dipahami dengan baik oleh pembaca, serta dapat memberikan kontribusi dalam konteks ilmiah atau praktis sesuai dengan tujuan penelitian yang telah ditetapkan.

### 3.2 Fase Akhir yang Akan Dicapai

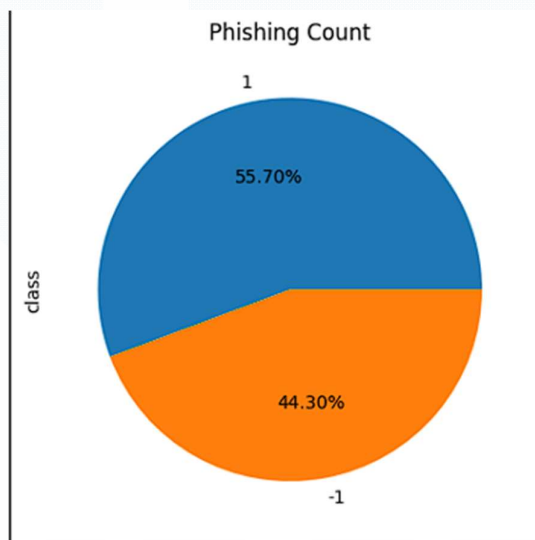
Hasil dari penelitian ini adalah implementasi sebuah website inovatif yang mampu mendeteksi tautan *phishing* secara efektif. *Website* ini memberikan pengguna kemampuan untuk memilih metode deteksi yang sesuai dengan preferensi mereka. Dengan pendekatan yang *user-friendly*, pengguna dapat memilih metode tertentu, seperti analisis URL, analisis konten halaman, atau penggunaan algoritma *machine learning* khusus, untuk mengidentifikasi potensi tautan *phishing*. Pengembangan tersebut memberikan keleluasaan kepada pengguna untuk menyesuaikan tingkat keamanan sesuai dengan kebutuhan mereka. Dengan demikian, *website* ini tidak hanya memitigasi risiko tautan *phishing* tetapi juga memberikan pengalaman pengguna yang disesuaikan dan dapat diandalkan dalam menjelajahi internet dengan aman.

### 3.3 Koleksi Data

Pada bab sebelumnya sudah dijelaskan bahwa dalam penelitian rancang bangun ini diperlukan data dalam bentuk *dataset*. Dimana *dataset* tersebut berisikan hasil klasifikasi dari beberapa *sample URL phishing* dan *non-phishing* yang dikumpulkan oleh pembuat *dataset* tersebut. Data ini diperlukan sebagai sumber data dalam pembuatan model deteksi menggunakan algoritma Random Forest. Model yang dihasilkan pada proses pengerjaan modeling diimplementasikan ke *server backend* sebagai bagian dari keseluruhan sistem deteksi *phishing*. Hasil pemodelan di-*export* dalam bentuk *file Pickle* yang penulis panggil pada bagian *server*.

Gambar 3.1 merupakan gambar pembagaian klasifikasi variabel *class* pada dataset. *Dataset* yang digunakan dalam penelitian ini diambil dari sumber laman

Kaggle.com. Meskipun sudah berumur 4 *dataset* ini, berdasarkan *update* terakhir di laman Kaggle, *dataset* ini masih dinilai cukup relevan lantaran belum berumur lebih dari 5 tahun. Didalam *dataset* sudah termasuk satu *file dataset* dengan tipe data csv yang berisikan lebih dari 11000 entitas data dengan 32 atribut variabel. Entitas data tersebut merupakan kumpulan dari *website URL* yang terdiri dari 55.70% merupakan URL *phising* dan 44.30% merupakan URL yang aman untuk dikunjungi. Pembagian data klasifikasi *Class* pada *dataset* tersebut diketahui dari hasil *data preparation* menggunakan bahasa pemrograman *Python*.



Gambar 3. 1 Pembagian Klasifikasi variabel *Class* pada *Dataset*

### 3.4 Penyusunan Desain Teknis

Dalam proses penyusunan desain teknis, digunakan metode *prototyping* dalam pembuatan *website* deteksi *phishing*. *Prototyping* merupakan pendekatan dalam pengembangan perangkat lunak yang melibatkan pembuatan versi awal dari sistem untuk tujuan demonstrasi konsep, uji coba desain, serta identifikasi masalah dan solusi potensial. Dengan menggunakan model *prototype*, pengguna dapat mengalami secara langsung bagaimana sistem beroperasi. Penelitian ini

mengadopsi metode *prototyping* dengan tujuan merancang aplikasi melalui pembuatan *prototipe* terlebih dahulu, yang nantinya akan dievaluasi oleh pengguna. Setelah aplikasi *prototipe* dinilai oleh pengguna, informasi tersebut akan menjadi landasan untuk mengembangkan aplikasi menjadi produk akhir, menjadi *output* utama dari penelitian ini.

Metode *prototyping* dimulai dengan pengumpulan kebutuhan, melibatkan pengembang dan pengguna sistem untuk menentukan tujuan, fungsi dan kebutuhan operasional [10]. Berikut merupakan alur kerja dalam metode *prototyping*:

- a. Pengumpulan kebutuhan
- b. Proses desain
- c. Membangun *prototype*
- d. Evaluasi dan perbaikan

Mengumpulkan kebutuhan melibatkan pertemuan antara pengembang dan pelanggan guna menetapkan tujuan umum pembuatan perangkat lunak dan mengidentifikasi garis besar kebutuhan dasar sistem yang akan dikembangkan. Proses desain difokuskan pada representasi aspek perangkat lunak dari perspektif pengguna, mencakup *input*, proses, dan format *output*. Desain yang cepat mengarah ke pembangunan prototipe, yang dievaluasi oleh pengguna dan analisis desain untuk menyesuaikan kebutuhan perangkat lunak yang akan dikembangkan. Prototipe diatur agar memenuhi kebutuhan pengguna, memberikan pemahaman yang lebih jelas dan detail kepada pengembang mengenai langkah-langkah yang perlu diambil. Setelah menjalankan keempat langkah *prototyping*, langkah berikutnya adalah pembuatan atau perancangan produk yang sesungguhnya.