BAB 1 PENDAHULUAN

1.1 Latar Belakang

Saat ini setiap orang sangat bergantung pada internet. Semua orang banyak melakukan aktivitas online seperti Bank *online*, pemesanan *online*, isi ulang *online* dan lainnya di internet [1]. Namun *web* juga telah menjadi platform untuk membantu berbagai kriminal seperti *spam*, penipuan keuangan, dan menyebarkan *malware*. Supaya orang mengunjunginya dimungkinkan menggunakan *email* atau halaman situs lain. Untuk mengatasi hal ini, komunitas keamanan merespons dengan mengembangkan layanan daftar hitam yang memberikan peringatan atau peringatan dengan umpan balik yang akurat. Namun ada situs yang terlalu baru, tidak diklasifikasikan, atau salah diklasifikasikan, sehingga banyak situs berbahaya tidak masuk daftar hitam [2].

Analogi serangan *phishing* berasal dari "memancing" korban teknik untuk penyerang (juga disebut sebagai *phisher*) yang membuka situs web palsu, yang memiliki desain yang persis sama dengan situs populer dan legal di internet. Meskipun halaman ini memiliki antarmuka pengguna grafis yang serupa situs *phishing* memiliki *Uniform Resource Locator* (URL) yang berbeda dengan situs *legitimate*. Karena banyak melakukan aktivitas online maka ancaman *phising* meningkat. *Phishing* adalah jenis ancaman untuk mengambil informasi seperti login, id, password dan informasi kartu kredit. *Phishing* bisa mengambil Informasi pribadi pengguna yang seharusnya rahasia jadi diambil oleh pihak lain, dan disalahgunakan oleh pihak lain. Informasi kartu kredit diambil maka bisa dipakai transaksi sesuai keinginan yang mengambil [1].

Perkembangan terakhir yang telah dilakukan untuk menangani *phishing* ada pada jurnal yang berjudul *phishing detection using machine learning technique*, Junaid Rashid dan yang lainnya tahun 2020 melakukan riset dengan membandingkan teknik *machine learning* yang satu dengan yang lain dengan metode *random forest* dan *support vector machine* [2]. *Machine learning based phishing detection from URLs*, Ozgur Koray Sahingoz dan yang lainnya tahun 2018 melakukan riset dengan membandingkan teknik *machine learning* yang satu dengan yang lain dengan metode *decision tree, adaboost, kstar, kNN (k = 3), random forest* [1]. Avira merupakan *software* keamanan komputer buatan jerman yang juga memiliki *anti phising* [3]. ESET *internet security* merupakan *software* keamanan komputer buatan perusahaan Slovak *internet security* yang juga memiliki *anti phising* [4].

Permasalahan *phising* harus ditangani karena *phising* mencuri informasi pribadi dari pengguna dan membuat pegguna melakukan transaksi bodong yang merugikan pengguna.

Karena itu untuk mencegah pencurian informasi melalui web phising dan untuk mencegah pengguna internet dalam melakukan transaksi bodong agar tidak diperdaya sehingga tidak dirugikan secara materil.

Alternatif solusi yang diusulkan dalam penelitian adalah *machine learning* sudah terbukti ampuh untuk mengklasifikasikan aktivitas berbahaya [1]. Menggunakan teknik *deep learning* dengan menggunakan *Convolutional Neural Network* (CNN) karena menunjukkan kinerja yang sangat baik dalam analisis klasifikasi topik [5].

Hasil akhir dari penelitian ini adalah model *machine learning*, algoritme dari *model machine learning* yang didapat dari melatih model *machine learning* menggunakan data yang tersedia dan *user interface* untuk pengujian URL. Manfaat yang bisa diperoleh dari hasil akhir dari penelitian ini adalah pengguna dapat terhindar dari pencurian informasi secara ilegal di situs phising dan pengguna dapat menghindari transaksi bodong yang ada di situs *phising.*

1.2 Rumusan Masalah

Berdasarkan latar belakang yang sudah diuraikan, maka dapat ditetapkan rumusan masalah yaitu bagaimana membuat model *deep learning* untuk mendeteksi situs *phising*?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dipaparkan, maka tujuan penelitian ini adalah membuat model *deep learning* dengan menggunakan *Convolutional Neural Network* untuk mendeteksi situs *phising.*

1.4 Batasan Masalah

Dalam riset ini, untuk mencegah perancangan sistem di luar kebutuhan maka dibuatlah batasan-batasan masalah seperti:

1. Sistem ini belum terintegrasi dengan *browser*
2. Memblok situs *phising* tidak termasuk dalam penelitian ini
3. URL *Shortener* tidak termasuk dalam penelitian ini. URL *Shortener* adalah alat untuk mengurangi panjang URL contohnya *Short URL* yang bisa mengubah "https://www.google.com/" menjadi "shorturl.at/jqCEO" sehingga situs *phishing* menjadi lebih sulit untuk dideteksi.

1.5 Hasil dan Manfaat

Hasil akhir dari penelitian berupa model algoritme *machine learning* dari model *machine learning* dan *user interface* untuk pengujian URL. Manfaat yang bisa diperoleh

pengguna dapat terhindar dari pencurian informasi secara ilegal di situs *phising*, pengguna dapat menghindari transaksi bodong yang ada di situs *phising*.

1.6 Metodologi Penelitian

1. Studi literatur

   Tahap ini penulis memulai penelitian dengan melalui online jurnal, artikel, dan *course* yang berkaitan tentang perkembangan terakhir dan *machine learning*.

2. Analisis masalah

   Tahap ini menganalisis masalah dari jurnal dan *software* yang telah berkembang saat ini, serta menentukan Batasan masalah untuk penelitian ini.

3. Data *preprocesing*

   Pada tahap ini adalah membuang data duplikat dari dataset. Kemudian data yang digunakan akan dibagi menjadi tiga, yaitu data *training*, data *validation*, dan data *testing*.

4. *Training* model *deep learning*

   Pada tahap ini akan dilakukan pelatihan model untuk mendeteksi *website phishing.* Tahap ini akan menggunakan data *training*.

5. Menguji model

   Tahap ini model yang ada akan di *testing* menggunakan data lainnya.

6. Evaluasi

   Pada tahap ini akan dievaluasi hasil dari model untuk mendeteksi website phishing.

7. Membuat *software* untuk input URL

   Pada tahap ini akan dibuat *software* untuk input URL, kemudian input URL tersebut akan dilakukan *feature extraction* dan hasil pelatihan model akan di gunakan untuk memprediksi URL.