TUGAS AKHIR

PERBANDINGAN KINERJA ALGORITMA KLASIFIKASI DAN IMPLEMENTASI MODEL TERBAIK PADA WEBSITE DETEKSI LINK PHISHING

M DICKY ALFANSYAH

183510468

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS ISLAM RIAU
PEKANBARU
2024

HALAMAN PENGESAHAN TUGAS AKHIR

Nama : M Dicky Alfansyah

NPM : 183510468

Kelompok Keahlian : Artificial Intelegence

Program Studi : Teknik Informatika

Jenjang Pendidikan : Strata Satu (S1)

Judul TA : Perbandingan Kinerja Algoritma Klasifikasi dan Implementasi Model Terbaik Pada Website Deteksi Link Phishing

Format sistematika dan pembahasan materi pada masing-masing bab dan sub bab dalam tugas akhir ini telah dipelajari dan dinilai relatif telah memenuhi ketentuan-ketentuan dan kriteria- kriteria dalam metode penelitian ilmiah. Oleh karena itu tugas akhir ini dinilai layak dapat disetujui untuk disidangkan dalam ujian **Seminar Tugas Akhir.**

Pekanbaru, 25 Januari 2024

Di sahkan oleh :

Penguji I	Penguji II
Ause Labellapansa, S.T., M.Cs., M.Kom	**Mutia Fadhilla, S.ST., M.Sc**
NIDN 1018088102	**NIDN 1025059401**

Ketua Program Studi Teknik Informatika	Dosen Pembimbing
Apri Siswanto, S.Kom., M.Kom., Ph.D	**Dr. Arbi Haza Nasution, B.IT., M.IT**
NIDN 1016048502	**NIDN 1023048901**

HALAMAN PENGESAHAN
DEWAN PENGUJI TUGAS AKHIR

Nama : M Dicky Alfansyah

NPM : 183510468

Kelompok Keahlian : Artificial Intelegence

Program Studi : Teknik Informatika

Jenjang Pendidikan : Strata Satu (S1)

Judul TA : Perbandingan Kinerja Algoritma Klasifikasi dan Implementasi Model Terbaik Pada Website Deteksi Link Phishing

Tugas Akhir ini secara keseluruhan dinilai telah memenuhi ketentuan-ketentuan dan kaidah-kaidah dalam penulisan penelitian ilmiah serta telah diuji dan dapat dipertahankan dihadapan dewan penguji. Oleh karena itu, Tim Penguji Ujian Tugas Akhir Fakultas Teknik Universitas Islam Riau menyatakan bahwa mahasiswa yang bersangkutan dinyatakan **Telah Lulus Mengikuti Ujian Tugas Akhir Pada Tanggal 27 Maret 2024** dan disetujui serta diterima untuk memenuhi salah satu syarat guna memperoleh gelar Sarjana Strata Satu Bidang Ilmu Teknik Informatika.

Pekanbaru, 27 Maret 2024

Dewan Penguji

1. Pembimbing : Dr. Arbi Haza Nasution, B.IT., M.IT (.....)
2. Penguji 1 : Ause Labellapansa, S.T., M.Cs., M.Kom (.....)
3. Penguji 2 : Mutia Fadhilla, S.ST., M.Sc (.....)

Disahkan Oleh :

Ketua Program Studi

Teknik Informatika

Apri Siswanto, S.Kom., M.Kom., Ph.D

NIDN 1016048502

PERNYATAAN KEASLIAN TUGAS AKHIR

Dengan ini saya menyatakan bahwa tugas akhir ini merupakan karya saya sendiri dan semua sumber yang tercantum didalamnya baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar sesuai ketentuan. Jika terdapat unsur penipuan atau pemalsuan data maka saya bersedia dicabut gelar yang telah saya peroleh.

Pekanbaru, 27 Maret 2024

Penulis

M Dicky Alfansyah

NPM. 183510468

KATA PENGANTAR

Assalamu'alaikum warahmatullahi wabarakatuh.

Alhamdulillah, segala puji bagi Allah Subhanna wa Ta'ala yang telah memberikan rahmat dan karunia-Nya sehingga tugas akhir ini dapat terselesaikan. Judul tugas akhir ini adalah **"Perbandingan Kinerja Algoritma Klasifikasi dan Implementasi Model Terbaik pada Website Deteksi Link Phishing"**, yang merupakan syarat untuk memperoleh gelar Sarjana Teknik dari Program Studi Teknik Informatika Universitas Islam Riau. Penulis menyadari bahwa banyak pihak telah memberikan dukungan dan bantuan selama proses penulisan tugas akhir ini, serta memberikan wawasan ilmu selama perkuliahan. Tanpa bantuan mereka, tentu akan sulit rasanya untuk mendapatkan gelar Sarjana Teknik ini. Oleh karena itu penulis ingin mengucapkan terima kasih kepada:

1. Bapak Dr. Apri Siswanto, M. Kom selaku Kepala Prodi Teknik Informatika.
2. Bapak Dr. Arbi Haza Nasution, B.IT, M.I.T selaku dosen pembimbing, yang telah memberikan motivasi, arahan dan nasihat yang berharga selama penyusunan tugas akhir ini.
3. Bapak Yudhi Arta, S.T, M.Kom selaku dosen pembimbing akademik, yang telah memberikan arahan dan nasihat selama menjalani perkuliahan.
4. Bapak/Ibu dosen atas bimbingan dan ilmu yang telah diberikan selama proses belajar mengajar di bangku perkuliahan.
5. Seluruh staff dan pegawai Fakultas Teknik Universitas Islam Riau, yang telah memberikan kemudahan dalam pelayanan administrasi perkuliahan.
6. Kedua orang tua yang senantiasa bersabar dan memberikan dukungan penuh baik material, moral maupun spiritual selama menjalani perkuliahan hingga menyelesaikan tugas akhir ini.
7. Adik kandung yang selalu mengingatkan penulis untuk terus semangat menyelesaikan tugas akhir ini.
8. Kekasih Nurul Azizah yang telah memberikan dukungan dan semangat kepada penulis selama menyelesaikan tugas akhir ini.

9. Kedua orang tua dan keluarga besar kekasih yang memberikan dukungan kepada penulis dengan mengajak berlibur sehingga penulis tidak terlalu merasa terbebani dalam menyelesaikan tugas akhir ini.
10. Sahabat terbaik Alfan Syaid, Muhammad Hari Mukti, Riski Kriswono, Muhammad Fachryan Heraldy, Bayu Prabowo, Angga Wisnu Wardana dan teman-teman seperjuangan yang telah memberikan penulis dukungan untuk dapat menyelesaikan tugas akhir ini.

Teriring doa penulis ucapkan, semoga Allah memberikan balasan atas segala kebaikan semua pihak yang telah membantu. Semoga tugas akhir ini membawa manfaat bagi pengembangan ilmu pengetahuan.

Pekanbaru, Maret 2024

Penulis

M Dicky Alfansyah

NPM. 183510468

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

PERBANDINGAN KINERJA ALGORITMA KLASIFIKASI DAN IMPLEMENTASI MODEL TERBAIK PADA WEBSITE DETEKSI LINK PHISHING

M Dicky Alfansyah
Jurusan Teknik Informatika Fakultas Teknik Universitas Islam Riau
Email: dicky.alfansyah1510@student.uir.ac.id

ABSTRAK

Perkembangan teknologi informasi telah mengubah cara praktis dalam berinteraksi, berkomunikasi, dan melakukan transaksi secara online. Namun, hal ini juga membawa risiko keamanan, seperti meningkatnya serangan phishing. Indonesia Anti-Phishing Data Exchange (IDADX) mencatat pada kuartal pertama tahun 2023 terdapat sebanyak 26.675 laporan phishing. Penelitian ini bertujuan untuk mengidentifikasi link phishing dengan membangun sistem yang menerapkan algoritma klasifikasi menggunakan pendekatan *machine learning* untuk menghasilkan model terbaik, dengan GPT (Generative Pre-trained Transformer) sebagai salah satu fitur ekstraksi. Berdasarkan hasil penelitian, model *neural network* yang dikembangkan mencapai kinerja tertinggi, dengan akurasi, presisi, recall, dan f1-score mencapai 92.11%, menunjukkan kemampuan yang sangat baik dalam mengklasifikasikan URL phishing.

Kata Kunci: risiko keamanan, *machine learning*, *neural network*, klasifikasi URL

PERFORMANCE COMPARISON OF CLASSIFICATION ALGORITHMS AND BEST MODEL IMPLEMENTATION ON PHISHING LINK DETECTION WEBSITE

M Dicky Alfansyah
Department of Informatics Engineering Faculty of Engineering
Riau Islamic University
Email: dicky.alfansyah1510@student.uir.ac.id

ABSTRACT

The development of information technology has changed the practical way of interacting, communicating and conducting transactions online. However, this also brings security risks, such as the increase in phishing attacks. Indonesia Anti-Phishing Data Exchange (IDADX) noted that in the first quarter of 2023 there were 26,675 phishing reports. This research aims to identify phishing links by building a system that applies classification algorithms using a machine learning approach to produce the best model, with GPT (Generative Pre-trained Transformer) as one of the extraction features. Based on the research results, the developed neural network model achieved the highest performance, with accuracy, precision, recall, and f1-score reaching 92.11%, showing excellent ability in classifying phishing URLs.

Keywords: security risk, machine learning, neural network, URL classification

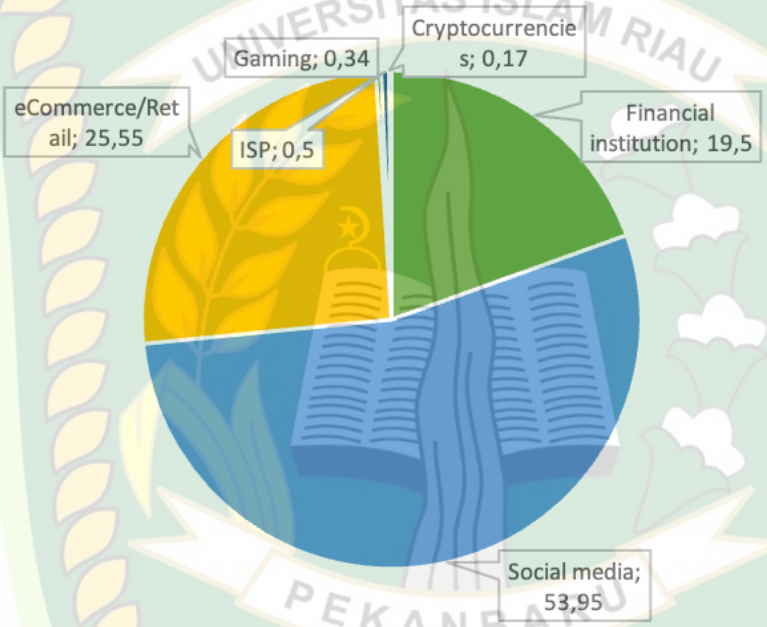
BAB I
PENDAHULUAN

1.1. Latar Belakang

Pada era digital saat ini perkembangan teknologi informasi yang sangat cepat dan terus berkembang seiring waktu, yang membawa perubahan dan kemudahan dalam berbagai aspek kehidupan. Hal ini memungkinkan semua orang dapat dengan mudah dan cepat berinteraksi, berkomunikasi, mendapatkan informasi, serta melakukan berbagai transaksi secara online. Namun, hal ini dapat menimbulkan berbagai masalah terkait dengan keamanan informasi dan privasi. Salah satu dampak negatif dari pesatnya perkembangan teknologi di era digital adalah dengan maraknya modus pencurian data pribadi dan informasi, seperti yang terjadi pada penipuan dalam internet banking yang menggunakan pemalsuan situs web, yang dikenal dengan istilah phishing. Phishing adalah teknik penipuan melalui internet yang bertujuan untuk mencuri informasi pribadi pengguna seperti password, kartu kredit, atau data sensitif lainnya dengan memperdaya pengguna untuk memasukkan informasi sensitif ke situs palsu yang menyerupai situs asli.

Jumlah laporan phishing yang diterima oleh Indonesia Anti-Phishing Data Exchange (IDADX) pada kuartal pertama tahun 2023 tercatat sebanyak 26.675 laporan dibandingkan dengan kuartal keempat tahun 2022 sebanyak 6.106 laporan. Hal ini menunjukkan kenaikan yang signifikan sebanyak 20.569 laporan dari kuartal sebelumnya , sedangkan pada kuartal kedua tahun 2023 tercatat sebanyak 20.330 laporan, dimana terdapat beberapa ccLTD yang menjadi sasaran phishing yaitu .id, .biz.id, dan .my.id. Ketiga ccLTD tersebut banyak digunakan untuk aktivitas phishing dikarenakan syarat pendaftarannya yang lebih mudah,

dan harga yang lebih terjangkau dibandingkan ccLTD lainnya. Sektor industri yang paling sering menjadi target serangan phishing di posisi pertama adalah media sosial sebesar 53,95%, sedangkan *eCommerce/Retail* menempati posisi kedua sebesar 25,55% dan lembaga keuangan menempati posisi ketiga sebesar 19,5%. (Indonesia Anti-Phishing Data Exchange (IDADX), 2023b, 2023a)



Gambar 1. 1 Persentase industri target phishing.
(Sumber: Indonesia Anti-Phishing Data Exchange (IDADX), 2023b)

Pesatnya perkembangan teknologi saat ini, terbukti dari banyaknya inovasi dan perkembangan baru dalam berbagai bidang seperti *Artificial Intelligence, Machine Learning, Internet of Things, Cloud Computing, Big Data, Virtual Reality,* dan *Blockchain.* Dengan adanya perkembangan teknologi baru salah satunya dapat bermanfaat dalam mencegah kejahatan phishing sebagai contoh adalah penerapan teknologi *machine learning* dalam website deteksi link phising. *Machine learning* dapat digunakan untuk mempelajari dan mengenali pola dari link phishing dengan menggunakan algoritma klasifikasi. Hal ini dapat membantu

sistem membuat prediksi dan mengidentifikasi link phishing dengan lebih akurat daripada pendekatan manusia yang biasanya lebih subjektif.

Dari latar belakang permasalahan tersebut membuat penulis tertarik untuk melakukan penelitian mengenai **"PERBANDINGAN KINERJA ALGORITMA KLASIFIKASI DAN IMPLEMENTASI MODEL TERBAIK PADA WEBSITE DETEKSI LINK PHISHING"**. Penelitian ini dilakukan untuk mempelajari serta memahami cara mengidentifikasi link phishing dengan melibatkan perbandingan kinerja algoritma klasifikasi *Support Vector Machine (SVM), Decision Tree, Logistic Regression, Random Forest, K-Nearest Neighbors, Neural Network,* dan *Naïve Bayes*. Hasil akhir dari penelitian ini adalah model terbaik yang dapat diterapkan pada website deteksi link phishing.

1.2. Identifikasi Masalah

Berdasarkan uraian pada latar belakang dapat disimpulkan identifikasi masalah yang terjadi yaitu, sebagai berikut:

1. Perkembangan teknologi informasi yang pesat dan terus berkembang memiliki dampak signifikan terhadap keamanan informasi dan privasi.
2. Maraknya modus pencurian data pribadi dan informasi melalui praktik phishing telah mengakibatkan kerugian bagi pengguna internet.
3. Tampilan situs web phishing seringkali menyerupai dengan situs web asli, begitu juga dengan URL yang cenderung mirip dengan URL web asli, sehingga pengguna kesulitan membedakannya.
4. Sejumlah sektor industri menjadi target serangan phishing, di antaranya adalah media sosial, *eCommerce/Retail*, dan lembaga keuangan, yang rentan terhadap jenis serangan ini.

1.3. Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, dapat diambil sebuah rumusan masalah. Rumusan masalah tersebut yaitu "Bagaimana cara mengidentifikasi link phishing secara efektif dengan membandingkan kinerja algoritma klasifikasi *Support Vector Machine (SVM), Decision Tree, Logistic Regression, Random Forest, K-Nearest Neighbors, Neural Network,* dan *Naïve Bayes*, dan bagaimana mengimplementasikan model terbaik pada website deteksi link phishing?"

1.4. Batasan Masalah

Untuk memfokuskan permasalahan dalam penelitian ini, peneliti mengidentifikasi batasan masalah sebagai berikut:

1. Fokus penelitian ini adalah untuk membandingkan kinerja berbagai algoritma klasifikasi dalam mengidentifikasi URL phishing.
2. Data yang digunakan dalam penelitian ini mencakup sampel URL phishing dan URL legitimate yang telah dikumpulkan dan diolah untuk melatih serta menguji berbagai algoritma klasifikasi.
3. Penelitian ini akan mempertimbangkan berbagai fitur ekstraksi yang digunakan untuk meningkatkan kinerja algoritma klasifikasi.
4. Model terbaik yang dihasilkan dari perbandingan algoritma klasifikasi akan diimplementasikanpada website deteksi link phishing.

1.5. Tujuan Penelitian

Tujuan yang ingin dicapai dari penelitian ini adalah sebagai berikut:

1. Aspek Pengguna

a. Meningkatkan pemahaman pengguna tentang ancaman phishing dalam lingkungan digital.
b. Menyediakan sistem yang efektif bagi pengguna untuk mengidentifikasi link phishing.

2. Aspek Penelitian
   a. Membandingkan kinerja berbagai algoritma klasifikasi dalam mengidentifikasi link phishing.
   b. Mengembangkan dan mengimplementasikan model terbaik pada website deteksi link phishing.

1.6. Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah sebagai berikut:

1. Melindungi data pribadi pengguna dengan upaya pencegahan untuk meminimalisir kerugian akibat serangan phishing.
2. Meningkatkan kesadaran pengguna tentang bahaya phishing.
3. Membantu pengguna dalam mengenali link phishing dengan menerapkan model terbaik pada website deteksi link phishing.
4. Berkontribusi dalam upaya antisipasi terhadap kejahatan *cyber*, menciptakan lingkungan online yang lebih aman, dan mengurangi risiko serangan phishing.

BAB II
LANDASAN TEORI

2.1. Tinjauan Pustaka

Dalam sebuah proses penelitian, tinjauan studi dari penelitian terdahulu merupakan keharusan. Tinjauan studi tersebut berfungsi sebagai bahan acuan pembanding dengan penelitian yang dilakukan. Oleh karena itu, penulis melakukan tinjauan studi dari beberapa jurnal yang berkaitan dengan kebutuhan penelitian.

Penelitian pertama dilakukan oleh Anthony Chandra, Gregorius, John Immanuel M. S, Alexander Agung Santoso Gunawan, dan Anderies (2022) dengan judul penelitian ***"Accuracy Comparison of Different Machine Learning Models in Phishing Detection"***. Pada penelitian tersebut, diusulkan pendeteksian serangan phishing dengan menggunakan algoritma klasifikasi seperti *Naive Bayes, Support Vector Machine, Decision Tree, Random Forest, K-Nearest-Neighbor,* dan *Logistic Regression*. Penelitian tersebut menggunakan fitur-fitur untuk klasifikasi dataset, yakni fitur berbasis URL (Keberadaan alamat IP, Panjang URL, Jumlah simbol), fitur berbasis konten (Jumlah *hyperlink*, Rasio *hyperlink internal eksternal*, Jumlah CSS *eksternal*), dan fitur berbasis *eksternal* (Umur domain, Peringkat halaman). Hasil penelitian tersebut menunjukkan bahwa dalam uji coba pemodelan *machine learning, Random Forest* memiliki akurasi tertinggi untuk mendeteksi serangan phishing, dengan nilai akurasi mencapai 98,04% (Chandra dkk., 2022).

Penelitian kedua dilakukan oleh Sanaa Kaddoura (2021) dengan judul penelitian ***"Classification of Malicious and Benign Websites by Network***

Features Using Supervised Machine Learning Algorithms". Pada penelitian tersebut, diusulkan pendeteksian situs web berbahaya menggunakan *supervised machine learning* dengan karakteristik algoritma yang dipilih, yakni menggunakan algoritma *Random Forest, Gradient Boosted, Logistic Regression, Extreme Gradient Boost, Decision Tree, K Nearest Neighbors (k=5),Neural Network, Support Vector Machine,* dan *Stochastic Gradient Descent*. Penelitian ini juga menggunakan dataset yang berisi URL phishing dan URL legitimate, yang terdiri dari 1.782 URL dengan pembagian kumpulan data, di mana 1.565 merupakan data yang berisi kelas label legitimate dan 216 dengan kelas label phishing. Pengkategorian pada kumpulan data atau dataset terdiri atas parameter, yang di antaranya adalah *URL, URL Length, Number Special Characters, Charset, Server, Content Length, Whois Country, Whois Statepro, App Bytes, Sourceapp Packets, Remoteapp Packets, App Packets, Dns Query Times,* dan *Type*. Hasil dari penelitian ini menunjukkan bahwa klasifikasi URL phishing dan legitimate mendapatkan hasil eksperimen dari f1-score, dimana untuk *Support Vector Machine* mendapatkan nilai 92%, hal ini merupakan nilai tertinggi dibandingkan algoritma lainnya. Pada ROC-AUC, algoritma *Random Forest* mendapatkan nilai 0,993, yang menunjukkan kinerja yang signifikan terhadap algoritma lainnya (Kaddoura, 2021).

Penelitian ketiga dilakukan oleh Wendy Sarasjati, Supriadi Rustad, Purwanto, Heru Agus Santoso, Muljono, Abdul Syukur, Fauzi Adi Rafrastara, dan De Rosal Ignatius Moses Setiadi (2022) dengan judul penelitian ***"Comparative Study of Classification Algorithms for Website Phishing Detection on Multiple Datasets"***. Pada penelitian tersebut, diusulkan deteksi situs phishing dari

kumpulan dataset dengan membandingkan algoritma klasifikasi *Naïve Bayes*, *Support Vector Machine, Neural Network,* dan *Random Forest*. Penelitian tersebut mengambil dataset yang disediakan oleh UCI dan Mendeley. Dataset yang disediakan oleh UCI diperoleh dari *University of California, Irvine Machine Learning Repository* yang berisi website transaksi pembayaran seperti *e-banking* dan *e-commerce*. Data tersebut terdiri dari 1353 kumpulan data dengan pembagian kumpulan data, di mana 548 merupakan URL legitimate, 702 merupakan URL phishing, dan 103 merupakan URL yang mencurigakan. Data ini merupakan dataset *multiclass* yang memiliki tiga kelas label yang berbeda, dengan ketentuan situs phishing direpresentasikan sebagai -1, situs mencurigakan direpresentasikan sebagai 0, dan untuk situs legitimate direpresentasikan sebagai 1. Dataset yang disediakan oleh Mendeley diperoleh dari departemen ilmu komputer dan teknologi informasi di Universitas Malaysia Sarawak yang terdiri dari 10.000 kumpulan data dengan pembagian kumpulan data, di mana 5.000 merupakan URL legitimate dan 5.000 merupakan URL phishing. Untuk dataset URL legitimate direpresentasikan sebagai 0 dan dataset URL phishing direpresentasikan sebagai 1. Hasil dari pengujian pertama menggunakan dataset UCI menunjukkan bahwa *Random Forest* mencapai nilai akurasi tertinggi dengan nilai 88,92%, sedangkan *Support Vector Machine* dengan *Polynomial Kernel* mendapatkan nilai terendah, yaitu 67,03%. Kemudian hasil dari pengujian kedua menggunakan dataset Mendeley menunjukkan bahwa *Random Forest* dan *Neural Network* dengan fungsi aktivasi ReLu mencapai akurasi tertinggi dengan nilai 97,50% dan 97,37%, sedangkan *Support Vector Machine* dengan *Polynomial Kernel* memiliki skor terendah yaitu 50,55% (Sarasjati dkk., 2022).

Penelitian keempat dilakukan oleh Diki Wahyudi, Muhammad Niswar, dan A. Ais Prayogi Alimuddin (2022) dengan judul penelitian ***"Website Phishing Detection Application Using Support Vector Machine"***. Pada penelitian tersebut, diusulkan pendeteksi website phishing menggunakan pendekatan algoritma klasifikasi *Support Vector Machine* dengan *Decision Tree dan K-Nearest Neighbors* sebagai pembanding kinerja sistem. Untuk hasil penelitian tersebut akan diimplementasikan pada aplikasi Python. Penelitian ini menggunakan ekstraksi fitur yang terdiri atas pengecekan URL, memeriksa data *Whois, Record DNS, Alexa database, PhishTank website,* dan *StopBadware.* Proses rangkaian penelitian tersebut dilakukan dengan melakukan studi literatur, persiapan data, ekstraksi fitur, pemrosesan data, perancangan sistem, implementasi sistem, uji coba dan analisis hasil, dan pembuatan laporan. Penggunaan data yang digunakan terdiri dari 11055 data yang dibagi dengan metode *10-fold cross-validation*. Hasil pengujian menggunakan *10-fold cross-validation* dengan algoritma klasifikasi *Support Vector Machine, Decision Tree, dan K-Nearest Neighbor,* serta proses optimasi parameter untuk masing-masing algoritma, menunjukkan bahwa akurasi terbaik dari masing-masing algoritma adalah *Support Vector Machine kernel linear* 77,09%, *Polinomial Kernel* 85,71%, *Kernel RBF* 85,32%, *Decision Tree* 85,40%, dan *K-Nearest Neighbors* 84,43%. Maka diperoleh nilai akurasi terbaik dicapai oleh algoritma *Support Vector Machine Polinomial Kernel* dengan nilai 85,71% (Wahyudi dkk., 2022).

Penelitian kelima dilakukan oleh Habiba BOUIJIJ dan Amine BERQIA (2021) dengan judul penelitian ***"Machine Learning Algorithms Evaluation for Phishing URLs Classification"***. Pada penelitian tersebut, dilakukan evaluasi

machine learning dengan melakukan klasifikasi pada URL phishing. Penelitian ini menggunakan algoritma *Decision Tree, K-Nearest Neighbors, Gradient Tree Boosting, Logistic Regression, Naive Bayes, Random Forest, Support Vector Machine, Neural Network, Extra-Tree,* dan *AdaBoost.* Untuk dataset, diperoleh dari phishtank.com dan openphish.org dengan mengumpulkan sebanyak 11.338 dataset yang terdiri atas 8.495 URL phishing dan 2.843 URL legitimate, yang kumpulan data tersebut disimpan dalam format CSV. Penelitian tersebut mengadopsi pendekatan analisis leksikal untuk mengekstrak fitur URL dan menghitung metrik performa akurasi untuk setiap algoritma. Hasil yang diperoleh adalah algoritma *Extra-Tree* dan *K-Nearest Neighbors* mencapai skor akurasi terbaik, yakni mencapai 91% (BOUIJIJ & BERQIA, 2021).

Penelitian keenam dilakukan oleh Norzaidah Binti Md Noh dan M. Nazmi Bin M. Basri (2021) dengan judul penelitian ***"Phishing Website Detection Using Random Forest and Support Vector Machine: A Comparison"***. Pada penelitian tersebut, dilakukan perbandingan antara algoritma klasifikasi *Random Forest* dan *Support Vector Machine*. Penelitian ini menggunakan dataset yang diperoleh dari website Fakultas Ilmu Komputer dan Teknologi Informasi (FSIT) di bawah University Malaysia Sarawak (UNIMAS). Dataset terdiri dari 30.000 kumpulan data dengan pembagian data 15.000 merupakan URL phishing dan 15.000 merupakan URL legitimate, pada kumpulan data memiliki label kelas *RAW-HTML (hyperlink), SCREENSHOT* (Antarmuka situs web), URL (Nama URL), *WEBPAGE* (Ikon di situs web), dan *WHOIS*. Dari dataset yang telah terkumpul, penelitian ini hanya menggunakan 20.000 kumpulan data dari total URL phishing dan URL legitimate. Untuk melatih model klasifikasi, dataset diproses dalam

tahap pra-pemrosesan data dengan tokenisasi dataset kode HTML menggunakan algoritma *Byte Pair Encoding (BPE)*. Setelah itu, dilakukan perhitungan skor *Term Frequency Inverse Document Frequency (TFIDF)* untuk mendapatkan representasi fitur dari setiap file HTML. Kemudian, dilakukan klasifikasi dengan algoritma *Random Forest* dan *Support Vector Machine*. Dataset tersegmentasi menjadi 90% dan 10% untuk pelatihan dan pengujian dataset pada setiap model. Dari hasil uji coba yang dilakukan, menunjukkan bahwa tingkat akurasi dari *Support Vector Machine* mencapai 84,73% dan pada *Random Forest* mendapatkan tingkat nilai akurasi mencapai 99,98%. hal ini menunjukkan bahwa algoritma *Random Forest* sebagai klasifikasi model dengan tingkat akurasi yang lebih tinggi dibandingkan dengan *Support Vector Machine* (Binti Md Noh & Bin M. Basri, 2021).

Berdasarkan tinjauan pustaka penelitian terdahulu yang telah dijelaskan, penulis memilih untuk membandingkan berbagai algoritma guna menentukan model terbaik yang dapat diterapkan pada website deteksi link phishing. Algoritma yang digunakan meliputi *Support Vector Machine (SVM), Decision Tree, Logistic Regression, Random Forest, K-Nearest Neighbors, Neural Network,* dan *Naïve Bayes.*

2.2. Dasar Teori

2.2.1. Dataset

Dataset merupakan kumpulan data yang dapat digunakan untuk analisis atau penelitian. Dataset dapat berisi berbagai jenis data, baik terstruktur maupun tidak terstruktur, dan dapat mencakup informasi seperti gambar, teks, angka, atau atribut lainnya yang dapat diekstraksi dan dianalisis.

2.2.2. Machine Learning

Machine learning adalah sebuah cabang dari ilmu komputer yang memungkinkan sistem untuk belajar dan meningkatkan kinerjanya secara otomatis melalui pengalaman. Model *machine learning* dapat digunakan untuk melakukan berbagai tugas seperti pengenalan gambar, pengenalan suara, dan sebagainya. Dalam membangun model *machine learning*, dataset digunakan untuk melatih model agar dapat menghasilkan prediksi yang akurat pada data yang belum pernah dilihat sebelumnya.

2.2.3. Fitur

Fitur merujuk pada atribut atau karakteristik dari data yang digunakan sebagai input pada model *machine learning*. Fitur dapat berupa berbagai jenis data, seperti angka, teks, dan gambar, tergantung pada jenis masalah yang ingin diselesaikan.

Pemilihan fitur sangat penting untuk memperoleh performa dan keakuratan model. Fitur harus memiliki informasi yang relevan dan signifikan. Fitur juga dapat diubah atau diolah sebelum dimasukkan ke dalam model *machine learning*, misalnya dengan normalisasi atau pengkodean ulang *(encoding)*.

2.2.4. Algoritma

Algoritma adalah serangkaian instruksi atau prosedur matematis yang digunakan oleh komputer untuk mempelajari pola atau struktur dalam data dan kemudian membuat prediksi atau mengambil tindakan berdasarkan pola tersebut. Algoritma ini digunakan untuk mempelajari bagaimana data berkaitan dan berinteraksi satu sama lain dan kemudian menerapkan pola yang dipelajari pada data baru untuk menghasilkan prediksi atau tindakan.

2.2.4.1. Support Vector Machine (SVM)

SVM merupakan salah satu algoritma yang dapat diterapkan dalam deteksi URL phishing. Pendekatan *SVM* memusatkan perhatian pada pemisahan dua kelas (klasifikasi biner) secara optimal dengan tujuan untuk mengklasifikasikan apakah suatu URL termasuk kategori legitimate atau phishing. *SVM* mempelajari dan memahami struktur serta pola yang rumit di balik URL phishing. Mekanisme kerja *SVM* dengan berupaya menemukan *hyperplane* yang optimal sebagai batas keputusan terbaik dengan tujuan meminimalkan kesalahan klasifikasi. Ini dilakukan dengan memetakan data ke dalam ruang berdimensi tinggi dan memisahkannya ke dalam kelas-kelas yang berbeda. Hal ini disebut dengan proses memaksimalkan *margin* yang bertujuan agar performa model *SVM* dapat ditingkatkan, dan dapat menghasilkan klasifikasi yang lebih akurat. (Gupta dkk., 2022).

2.2.4.2. Random Forest

Random Forest merupakan algoritma klasifikasi yang menggunakan pendekatan ensemble learning dengan cara melatih banyak pohon keputusan secara acak. Setiap pohon menghasilkan klasifikasi atau prediksi, dan hasil dari semua pohon kemudian digabungkan untuk membentuk prediksi akhir. Dengan menggunakan *Random Forest*, prediksi dapat menjadi lebih akurat dan fitur-fitur penting dalam data dapat diidentifikasi lebih baik, sehingga membedakan antara URL phishing dan URL legitimate menjadi lebih mudah. Pendekatan ini memiliki keuntungan tersendiri dalam menangani *overfitting* dan meningkatkan ketahanan terhadap variasi data yang tidak terstruktur, dengan kombinasi hasil dari setiap pohon. (Fayoumi dkk., 2022; Sindhu dkk., 2020).

2.2.4.3. Decision Tree

Decision Tree dapat digunakan dalam deteksi URL phishing. Algoritma ini melibatkan pembangunan model prediksi berdasarkan aturan keputusan yang memecah data menjadi beberapa bagian. Struktur pohon yang digunakan dalam *Decision Tree* memainkan peran penting dalam klasifikasi data dengan mengurutkannya berdasarkan nilai atribut yang relevan (Mandadi dkk., 2022). Dengan demikian, *Decision Tree* dapat memberikan keputusan akhir yang dapat memprediksi apakah suatu URL termasuk dalam kategori phishing atau legitimate. *Decision Tree* dapat memberikan cara yang efektif dan jelas untuk mengidentifikasi URL yang berpotensi membahayakan dan membantu dalam upaya perlindungan terhadap serangan phishing.

2.2.4.4. K-Nearest Neighbors (KNN)

KNN bekerja dengan mencari sekelompok objek dalam data latih yang paling mirip dengan objek dalam data baru atau data pengujian. *KNN* menunjukkan kinerja yang baik dalam analisis deteksi URL phishing (Bhoj dkk., 2021). Kemampuan analisis *KNN* sering kali lebih unggul dibandingkan algoritma lain dalam beberapa situasi atau kasus pada dataset tertentu. *KNN* memprediksi nilai target dengan menggunakan kelas mayoritas dari tetangga terdekat sebagai prediksi (Okfalisa dkk., 2017).

2.2.4.5. Logistic Regression

Logistic Regression merupakan teknik klasifikasi dalam pembelajaran terkontrol, efektif dalam menghitung *probabilitas* variabel target (Rambabu dkk., 2022). *Regresi logistik* memungkinkan kita meningkatkan akurasi deteksi URL berbahaya, berkontribusi dalam perlindungan informasi pengguna dan

mengurangi resiko pelanggaran privasi (Soumya dkk., 2022). Selain itu, *Logistic Regression* menunjukkan performa baik ketika data yang digunakan menunjukkan hubungan *linear* antara variabel input dan outputnya. Keunggulan lain dari *Logistic Regression* adalah kemudahan penggunaan dan interpretabilitas yang sangat baik, yang mendorong penerapannya secara luas (Ahammad dkk., 2022).

2.2.4.6. Neural Network

Neural Network adalah model komputasi yang terinspirasi oleh struktur dan fungsi jaringan saraf biologis manusia. Ini adalah salah satu jenis algoritma klasifikasi dalam *machine learning* yang digunakan untuk memproses informasi dan membuat prediksi atau keputusan. Cara kerja *Neural Network* melibatkan pemrosesan urutan karakter dari URL yang ditargetkan untuk deteksi. Urutan karakter ini kemudian dianalisis dan konteksnya dihasilkan melalui penggunaan metode *embedding* karakter dan *Bi-LSTM* dalam model *sequence-to-sequence* (Resiandi dkk., 2022). Setiap karakter dalam URL diubah menjadi vektor *embedding* karakter yang mengkonversi karakter tersebut menjadi nilai numerik. Hal ini memungkinkan fungsi *encoder* untuk menghasilkan representasi URL yang lebih akurat, mempermudah dalam mendeteksi serangan phishing (Novakovic & Markovic, 2022).

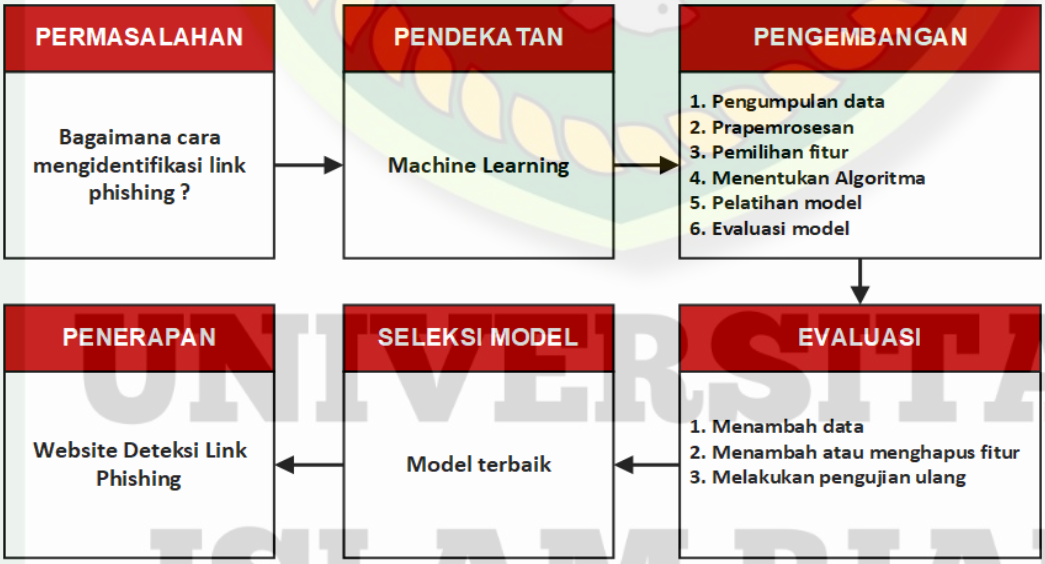
2.2.4.7. Naïve Bayes

Naïve Bayes merupakan algoritma klasifikasi *machine learning* yang dapat digunakan dalam model prediksi untuk deteksi phishing (Orunsolu dkk., 2022). Algoritma ini beroperasi berdasarkan *teorema bayes*, dengan asumsi bahwa setiap fitur yang digunakan bersifat independen satu sama lain (Fayoumi dkk., 2022). Fitur-fitur tersebut, yang diekstrak dari URL, mencakup berbagai elemen seperti

Panjang URL, Jumlah karakter numerik, Penggunaan simbol khusus, dan lainnya. Dengan memperhitungkan probabilitas bahwa URL tersebut termasuk dalam kategori phishing atau legitimate berdasarkan fitur-fitur tersebut, *Naïve Bayes* dapat menentukan kategori dengan probabilitas tertinggi sebagai hasil klasifikasinya (Mustafa & Karabatak, 2023).

2.3. Kerangka Pemikiran

Kerangka pemikiran merupakan suatu pendekatan sistematik yang digunakan dalam perancangan, pengembangan, dan implementasi solusi untuk memecahkan suatu masalah. Kerangka pemikiran ini mencakup konsep-konsep kunci, teori-teori, atau model-model yang membentuk landasan analisis terhadap masalah yang tengah dihadapi. Dengan menerapkan kerangka pemikiran, peneliti dapat mengorganisir informasi dengan lebih terstruktur, membantu dalam identifikasi variabel-variabel yang relevan, pengembangan hipotesis, serta perencanaan metodologi penelitian. Adapun kerangka pemikiran yang akan diterapkan pada penelitian ini dapat dilihat pada gambar dibawah sebagai berikut:



Gambar 2. 1 Kerangka pemikiran

BAB III
METODOLOGI PENELITIAN

3.1.Alat dan Bahan

3.1.1. Alat

Dalam konteks penelitian ini, istilah alat atau perangkat merujuk pada berbagai komponen, perangkat keras, dan perangkat lunak yang digunakan untuk melakukan eksperimen, analisis, dan pembuatan laporan. Alat-alat ini membantu dalam berbagai tahap penelitian, termasuk pengumpulan dan persiapan data dan analisis hasil. Adapun komponen yang digunakan sebagai berikut:

1. Perangkat Keras *(hardware)*

   Perangkat keras yang digunakan dalam penelitian ini dapat dilihat pada tabel dibawah ini sebagai berikut :

Tabel 3. 1 Spesifikasi perangkat keras

NO	Perangkat Keras	Spesifikasi
1	Laptop Acer Swift 3 Intel Evo	- Processor Intel® Core™ i5-1135G7 Gen 11. - GPU Intel(R) Iris(R) Xe Graphics - Ram Memory 8GB - SSD 512GB

2. Perangkat Lunak *(software)*

   Perangkat lunak yang digunakan dalam penelitian ini dapat dilihat pada tabel dibawah ini sebagai berikut :

Tabel 3. 2 Spesifikasi perangkat lunak

Sistem Operasi	Windows 11 Education
Bahasa Pemrograman	Python, Html, JavaScript
Web Browser	Brave

Tools	Visual Studio Code, Microsoft Word, Microsoft Exel
Framework	Flask
Library	sklearn, pandas, numpy, time, pickle, joblib, urllib, flask, json, urlparse, unidecode, re, BeautifulSoup, requests, whois, matplotlib, os, openai

Tabel 3. 3 Daftar model dan parameter default library sklearn

No	Algoritma	Model	Parameter	Nilai
1	Support Vector Machine	SVC	C	1.0
			break_ties	False
			cache_size	200
			class_weight	None
			coef0	0.0
			decision_function_shape	ovr
			degree	3
			gamma	scale
			kernel	rbf
			max_iter	-1
			probability	False
			random_state	None
			shrinking	True
			tol	0.001
			verbose	False
2	Random Forest	Random Forest Classifier	bootstrap	True
			ccp_alpha	0.0
			class_weight	None
			criterion	gini
			max_depth	None
			max_features	sqrt
			max_leaf_nodes	None
			max_samples	None
			min_impurity_decrease	0.0
			min_samples_leaf	1
			min_samples_split	2
			min_weight_fraction_leaf	0.0
			n_estimators	100
			n_jobs	None

			oob_score	False
			random_state	None
			verbose	0
			warm_start	False
3	Decision Tree	Decision Tree Classifier	ccp_alpha	0.0
			class_weight	None
			criterion	gini
			max_depth	None
			max_features	None
			max_leaf_nodes	None
			min_impurity_decrease	0.0
			min_samples_leaf	1
			min_samples_split	2
			min_weight_fraction_leaf	0.0
			random_state	None
			splitter	best
4	K-Nearest Neighbors	K neighbors Classifier	algorithm	auto
			leaf_size	30
			metric	minkowski
			metric_params	None
			n_jobs	None
			n_neighbors	5
			p	2
			weights	uniform
5	Logistic Regression	Logistic Regression	C	1.0
			class_weight	None
			dual	False
			fit_intercept	True
			intercept_scaling	1
			l1_ratio	None
			max_iter	100
			multi_class	auto
			n_jobs	None
			penalty	12
			random_state	None
			solver	lbfgs
			tol	0.0001

			verbose	0
			warm_start	False
6	Neural Network	MLP Classifier	activation	relu
			alpha	0.0001
			batch_size	auto
			beta_1	0.9
			beta_2	0.999
			early_stopping	False
			epsilon	0.0
			hidden_layer_sizes	(100,)
			learning_rate	constant
			learning_rate_init	0.001
			max_fun	15000
			max_iter	1000
			momentum	0.9
			n_iter_no_change	10
			nesterovs_momentum	True
			power_t	0.5
			random_state	None
			shuffle	True
			solver	adam
			tol	0.0001
			validation_fraction	0.1
			verbose	False
			warm_start	False
7	Naïve Bayes	Gaussian NB	priors	None
			var_smoothing	0.0

3.1.2. Bahan

3.1.2.1. Pengumpulan Data

Pengumpulan data memegang peran penting dalam sebuah penelitian. Dalam penelitian ini, proses pengumpulan data dimulai dengan pengambilan data dari berbagai sumber yang relevan, kemudian data tersebut disimpan dalam satu file dengan format CSV untuk analisis lebih lanjut.

Dataset yang telah dikumpulkan sebanyak 786 data yang dibagi menjadi dua kelas kategori, yakni kategori URL phishing dengan jumlah 393 data dan kategori URL legitimate dengan jumlah 393 data. Pengumpulan data ini diambil dari berbagai sumber yang dapat dilihat pada tabel berikut.

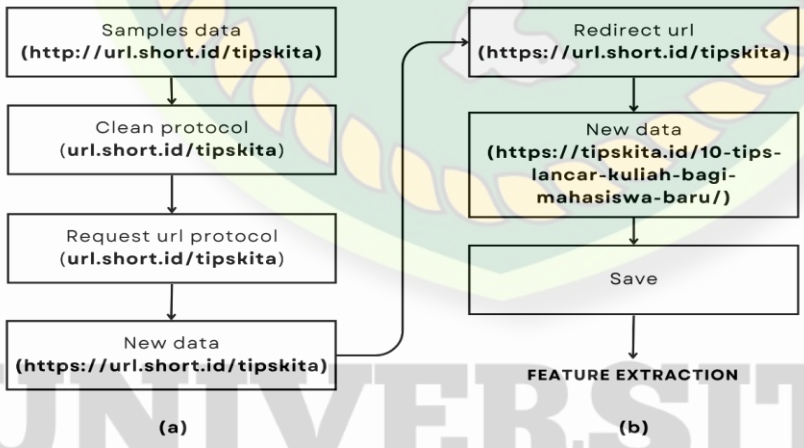
Tabel 3. 4 Pengumpulan data

No	Sumber	Jumlah
1	openphish.com	291
2	phishtank.org	93
3	facebook.com	15
4	Pengambilan data manual	387
Total		**786**

3.2. Pra-Pemrosesan

3.2.1. Normalisasi Data

Pada tahap ini melakukan normalisasi keseluruhan dari dataset. Tujuan dari normalisasi ini untuk memudahkan dalam pra-pemrosesan tahap ekstraksi fitur. Berikut adalah bentuk normalisasi yang dilakukan:



Gambar 3. 1 Normalisasi protokol; (b) Normalisasi *redirect*

1. Normalisasi Protokol: Melakukan normalisasi inputan URL dengan menghapus penggunaan protokol *http://* dan *https://*. Langkah selanjutnya

adalah membuat permintaan ke URL yang penggunaan protokol yang telah dihapus. Jika permintaan berhasil, maka URL akan dikembalikan dengan menggunakan protokol *https://* namun, jika permintaan tidak berhasil, maka akan dikembalikan dengan menggunakan protokol http://. Fungsi normalisasi ini dapat meminimalkan kesalahan selama input URL yang berdampak pada proses ekstraksi fitur (Mustafa & Karabatak, 2023; Sindhu dkk., 2020).

2. Normalisasi *Redirect*: dalam kasusnya terdapat beberapa URL yang di inputkan berisi singkatan atau pengalihan ke URL lain, untuk itu diperlukan teknik untuk mengetahui URL yang sebenarnya dengan cara melakukan permintaan HEAD ke URL dengan mengikuti pengalihan. Jika permintaan berhasil dan mengarah ke URL baru, URL baru ini akan digunakan dalam proses ekstraksi fitur. Jika permintaan tidak berhasil, proses ekstraksi fitur akan tetap menggunakan URL sebelumnya (Jain & Gupta, 2019; Mustafa & Karabatak, 2023; Pandey dkk., 2023).

3.2.2. Ekstraksi Fitur

Ekstraksi fitur pada dataset dilakukan untuk mempermudah dalam memahami karakteristik data yang diharapkan untuk dapat meningkatkan kualitas efisien dalam pengolahan data. Bentuk dari fitur-fitur yang akan digunakan pada penelitian ini adalah sebagai berikut:

1. Fitur Berbasis Internal/Karakteristik URL
   a. Panjang Domain: adalah fitur penting dalam deteksi URL phishing karena dapat mengindikasikan kemungkinan upaya penipuan berdasarkan karakteristiknya. Dari data yang telah dikumpulkan,

rata-rata panjang domain pada URL adalah 8 karakter, sehingga pada fitur ini panjang domain akan dihitung maksimal 8 karakter. Untuk mengurai URL, maka digunakan URLparse untuk mendapatkan bagian domain. Bagian domain akan dipisahkan dengan titik, dan akhiran (TLD/ccLTD) akan dihilangkan. Jika panjang domain melebihi 8 karakter, maka fitur ini akan diberi nilai 0. Namun , jika panjang domain kurang dari atau sama dengan 8 karakter, maka fitur ini akan diberi nilai 1. (Jalil dkk., 2023; Mandadi dkk., 2022; Pandey dkk., 2023).

b. Tanda Hubung Domain: URL phishing seringkali memanfaatkan tanda hubung dalam domain untuk mengelabui pengguna. Sebagai contoh, dalam URL *"https://reward-off-garena.ru"* terdapat tanda hubung yang mencurigakan, yaitu *"reward-off-garena."* Penggunaan tanda hubung dengan pola yang tidak umum dapat menjadi indikator potensial untuk URL phishing. Fitur ini digunakan untuk memeriksa apakah terdapat tanda hubung (-) pada domain yang digunakan. Jika tanda hubung ditemukan, maka fitur diberi nilai 0. Namun, jika tidak ditemukan tanda hubung, maka fitur diberi nilai 1 (Ahammad dkk., 2022; Jalil dkk., 2023; Mandadi dkk., 2022; Orunsolu dkk., 2022).

c. Protokol HTTPS: situs phishing cenderung memanfaatkan protokol HTTP biasa, yang tidak mengenkripsi data dengan baik. Dalam proses analisis, fitur ini digunakan untuk memeriksa apakah URL menggunakan protokol *https://* atau *http://* hasil dari normalisasi data. Jika URL menggunakan protokol *https://*, fitur ini akan memberikan nilai 1, menunjukkan tingkat keamanan yang baik. Sebaliknya, jika URL

menggunakan protokol *http://*, fitur ini akan memberikan nilai 0, mengindikasikan bahwa URL tersebut tidak aman dan berpotensi menjadi ancaman serangan phishing. (Ahammad dkk., 2022; BOUIJIJ & BERQIA, 2021; Chiew dkk., 2019; Jalil dkk., 2023; Mandadi dkk., 2022; Soumya dkk., 2022).

d. Pengecualian Domain: data yang dirangkum dari laman situs cybercrimeinfocenter.org pada Mei-Juli 2023 menunjukkan aktivitas TLD yang sering dimanfaatkan dalam kegiatan phishing. TLD yang sering dimanfaatkan sebagai phishing meliputi "cn", "ml", 'tk', "top", "info", "tk", "ga", "cf", "cyou", "xyz", "gq", "buzz", "shop", "online", "ru", "live", "site", "click", "cfd", "pw", "link", "support", "store", "co", "space", "net", dan "co.uk.". Indonesia juga memiliki domain yang rawan digunakan dalam praktik phishing karena harganya terjangkau dan persyaratan pendaftaran yang relatif mudah, yaitu ccLTD my.id. Untuk mengatasi hal ini, diperlukan pemfilteran dengan menggunakan fitur pengecualian domain sebagai metode analisis data, dimana jika TLD/ccLTD yang terdaftar dalam daftar pengecualian ditemukan dalam URL, maka fitur akan diberi nilai 0, Namun, jika tidak ditemukan, maka fitur akan diberi nilai 1. (Cybercrime Information Center, 2023)

e. Karakter Terlarang: URL phishing seringkali mengandung karakter yang tidak sah atau jarang terlihat dalam URL yang sah, seperti '@', '!', ';', '*', atau simbol-simbol aneh lainnya. Penggunaan fitur ini membantu mengidentifikasi URL yang mencurigakan sebagai potensi ancaman serangan phishing. Jika URL mengandung karakter terlarang, fitur ini

akan diberi nilai 0. Namun, jika tidak mengandung karakter terlarang, fitur akan diberi nilai 1. (Ahammad dkk., 2022).

f. Jumlah Karakter Khusus: mengukur jumlah karakter khusus dalam path dari sebuah URL sebagai salah satu metode untuk mengklasifikasikan URL yang meliputi karakter #, +, ?, -, _, =, &, $, dan %. Jumlah karakter khusus ini merupakan indikator potensial untuk menilai sebuah URL. Jika jumlah karakter khusus dalam path kurang dari 10, maka URL akan dianggap valid dengan nilai 1. Namun, jika jumlahnya mencapai atau melebihi 10, maka URL akan dianggap tidak valid dengan nilai 0. Sebagai contoh URL *"https://www.example.com/contohfiturkarakterkhususdari-url?user=12345&token=abcdefg"*, memiliki 3 karakter khusus dalam path, yaitu "?", "=", dan "&", maka jumlah karakter khususnya adalah 3, yang kurang dari 10, sehingga URL tersebut dianggap valid dengan nilai 1. (BOUIJIJ & BERQIA, 2021; Chandra dkk., 2022; Jalil dkk., 2023; Kaddoura, 2021; Sahingoz dkk., 2019).

g. Jumlah Path: menghitung jumlah path pada sebuah URL guna membantu dalam mengidentifikasi potensi serangan phishing, dapat digambarkan menggunakan URL *"https://grow.google/intl/id_id/"* sebagai contoh. URL tersebut memiliki jumlah path sebanyak dua, yaitu *"/intl/id_id/"*. Dalam skenario ini, jika jumlah path kurang dari atau sama dengan 3, fitur akan memberikan nilai 1 yang mengindikasikan bahwa URL tersebut tidak berpotensi sebagai serangan phishing. Namun, jika jumlah path melebihi 3, fitur akan memberikan nilai 0 untuk mengindikasikan potensi serangan phishing.

Hal ini dikarenakan serangan phishing seringkali menggunakan URL dengan path yang lebih panjang, dengan tujuan untuk mengelabui pengguna. (Chiew dkk., 2019; Jalil dkk., 2023; Pandey dkk., 2023).

h. Jumlah Single Slash: berdasarkan analisis data yang dikumpulkan, rata-rata jumlah single slash pada URL adalah 2, dengan kemungkinan adanya 3 slash. Oleh karena itu, ini dapat dianggap sebagai batasan dalam mengevaluasi fitur single slash. Sebagai contoh, dapat mempertimbangkan URL *"https://www.linkedin.com/jobs/"* yang memiliki 2 single slash, dimulai setelah domain yaitu "/jobs/". Dengan batasan yang telah ditetapkan, URL yang memiliki single slash lebih dari 3 akan diberi nilai 0, yang mengindikasikan potensi sebagai URL phishing. Namun, jika URL memiliki 3 single slash atau kurang, maka akan diberi nilai 1, menunjukkan bahwa URL tersebut cenderung tidak mencurigakan dari segi struktur dan jumlah slashnya. (BOUIJIJ & BERQIA, 2021; Jalil dkk., 2023; Orunsolu dkk., 2022; Pandey dkk., 2023).

i. Jumlah Titik: berdasarkan analisis data yang dikumpulkan, rata-rata jumlah tanda titik pada URL adalah sebanyak 2, dengan kemungkinan adanya 3 tanda titik yang digunakan. Jumlah rata-rata tanda titik ini dapat digunakan sebagai acuan batasan dalam fitur penghitungan jumlah tanda titik yang digunakan pada URL. Sebagai contoh, URL *"https://185.60.218.35/login.php"* akan berpotensi dianggap sebagai phishing karena memiliki 4 tanda titik, melebihi batasan yang telah ditetapkan. Dapat dinyatakan bahwa jika sebuah URL memiliki lebih dari

3 tanda titik, maka akan diberi nilai 0, mengindikasikan potensi sebagai URL phishing. Namun, jika jumlah tanda titik pada URL kurang dari atau sama dengan 3, maka akan diberi nilai 1, sebagai indikasi bahwa URL tersebut tidak berpotensi sebagai phishing. (BOUIJIJ & BERQIA, 2021; Chiew dkk., 2019; Jalil dkk., 2023; Pandey dkk., 2023).

j. Panjang URL: berdasarkan data yang telah dikumpulkan, rata-rata panjang domain pada data adalah 65 karakter, sehingga pada fitur ini panjang domain akan dihitung maksimal 65 karakter. Jika panjang URL melebihi 65 karakter, fitur diberi nilai 0. Namun, jika panjang URL kurang dari atau sama dengan 65 karakter, fitur diberi nilai 1 (Ahammad dkk., 2022; BOUIJIJ & BERQIA, 2021; Chandra dkk., 2022; Chiew dkk., 2019; Jalil dkk., 2023; Kaddoura, 2021; Mandadi dkk., 2022; Pandey dkk., 2023; Sindhu dkk., 2020).

2. Fitur Berbasis External/Luar Karakteristik URL

a. Isi Konten: dari analisis data yang telah dikumpulkan, terdapat situs URL yang mengandung kata-kata sensitif yang dapat mengindikasikan tindakan phishing. Beberapa kata sensitif yang dimaksud adalah "klaim," "reward," "ambil," "hadiahmu," "kepo," "kepoin," "melihat," "stalking," "ngepoin," "fb," "ig," "tarif," "dana," dan "facebookkepo." Untuk mengidentifikasi potensi phishing, fitur ini akan memproses semua elemen HTML atau XML pada halaman website dan mencari apakah terdapat isi konten yang mengandung kata-kata sensitif tersebut. Jika isi konten mengandung kata-kata sensitif tersebut, maka fitur akan memberikan nilai 0. Namun, jika konten URL tidak mengandung kata-

kata sensitif, maka fitur akan memberikan nilai 1. (Chandra dkk., 2022).

b. Favicon: favicon yang tidak sesuai dengan domain situs web dapat menjadi indikasi potensial untuk URL phishing. Sebagai contoh, jika favicon menggambarkan logo perusahaan A, tetapi URL mengarah ke situs web perusahaan B, hal ini dapat menjadi tanda potensi adanya URL phishing. Selain itu, penting untuk memeriksa format favicon yang digunakan dalam sumber daya HTML yang terdapat di dalam tag <head></head>. Format favicon yang diharapkan adalah ICO dan PNG. Jika favicon ditemukan sesuai dengan format yang ditentukan, fitur ini akan diberi nilai 1. Namun, jika favicon tidak ditemukan atau tidak sesuai dengan format yang telah ditentukan, maka fitur ini akan diberi nilai 0. (Jain & Gupta, 2019; Mustafa & Karabatak, 2023; Pandey dkk., 2023).

c. Whois: fitur ini berguna untuk memeriksa apakah sebuah domain yang ada pada URL aktif atau tidak. Jika domain aktif, maka akan mengembalikan nilai 1. Namun, jika tidak aktif atau ada masalah dalam pengambilan informasi Whois, maka akan mengembalikan nilai 0. Ini dapat digunakan sebagai salah satu langkah awal dalam analisis keamanan untuk mengidentifikasi domain-domain yang mencurigakan. (Ahammad dkk., 2022; Kaddoura, 2021; Soumya dkk., 2022; Wahyudi dkk., 2022).

3. Fitur Berbasis OpenAI

Generative Pre-Training Transformer (GPT) yang memiliki kemampuan untuk menganalisis berbagai skenario masalah. Dalam penelitian ini, API key GPT akan digunakan untuk ekstraksi fitur. Model GPT yang akan

digunakan mencakup GPT-3.5 dan GPT-4. Model ini akan membantu dalam menganalisis URL yang diinputkan. Jika hasil analisis URL menunjukkan adanya upaya phishing, fitur akan diberi nilai 0. Namun, jika hasil analisis URL tidak menunjukkan tanda-tanda upaya phishing, maka fitur akan diberi nilai 1. Berikut adalah tabel penggunaan model GPT:

Tabel 3. 5 Model GPT

No	Nama	Model	Parameter	Nilai
1	GPT-3.5	gpt-3.5-turbo	temperature	0.2
			top_p	0.1
			max_tokens	1
2	GPT-4	gpt-4	temperature	0.2
			top_p	0.1
			max_tokens	1

3.3. Optimasi Model

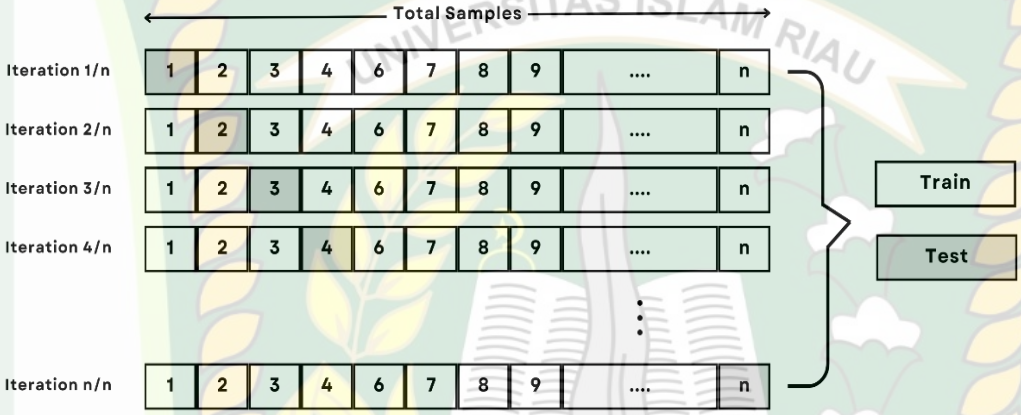
3.3.1. *Feature Importance*

Setelah menyelesaikan proses ekstraksi fitur, langkah selanjutnya adalah melakukan analisis hasil ekstraksi fitur guna menentukan fitur-fitur yang memiliki tingkat kepentingan yang tinggi di antara semua fitur yang ada dalam dataset. Tujuan dari proses ini adalah untuk mengevaluasi kontribusi setiap fitur terhadap kinerja atau prediksi model. Hasil dari analisis *feature importance* ini akan menjadi panduan untuk memilih fitur-fitur yang akan digunakan.

3.3.2. *Leave One Out Cross Validation (LOOCV)*

Merupakan teknik validasi silang *(cross-validation)* yang digunakan dalam pengembangan dan evaluasi model. Tujuan utama dari LOOCV adalah untuk mengukur sejauh mana model machine learning dapat digeneralisasikan ke data yang belum pernah dilihat sebelumnya, sehingga mengukur seberapa baik model

tersebut dalam mengatasi overfitting. Teknik ini sangat berguna ketika memiliki dataset yang terbatas. Teknik LOOCV bekerja dengan cara mengambil satu data sebagai data uji pada setiap iterasi, sementara data lainnya digunakan sebagai data latih. Proses ini berulang hingga semua data dalam dataset telah diuji secara keseluruhan.



Gambar 3. 2 Metode *leave one out cross validation*

3.3.3. Metrik Evaluasi

Metrik evaluasi digunakan untuk mengukur kinerja suatu model dalam menyelesaikan tugas tertentu, seperti klasifikasi, regresi, atau tugas lainnya. Metrik yang digunakan dalam penelitian ini untuk evaluasi adalah akurasi, presisi, recall, dan f1-score. Perhitungan metrik didasarkan pada elemen-elemen yang terdapat dalam *confusion matrix*, yaitu *True Positive (TP), True Negative (TN), False Positive (FP),* dan *False Negative (FN)* (Tang & Mahmoud, 2021, 2022). Rumus yang digunakan untuk menghitung hasil evaluasi dari setiap metrik adalah sebagai berikut:

1. Akurasi: mengukur seberapa baik model dalam mengklasifikasikan semua kelas dengan benar.

$$akurasi = \frac{TP+TN}{TP+TN+FN+FP} \quad (3.1)$$

2. Presisi: mengukur seberapa banyak dari yang diklasifikasikan sebagai positif oleh model yang benar-benar positif.

$$presisi = \frac{TP}{TP+FP} \quad (3.2)$$

3. Recall: mengukur seberapa banyak dari keseluruhan positif yang telah diidentifikasi dengan benar oleh model.

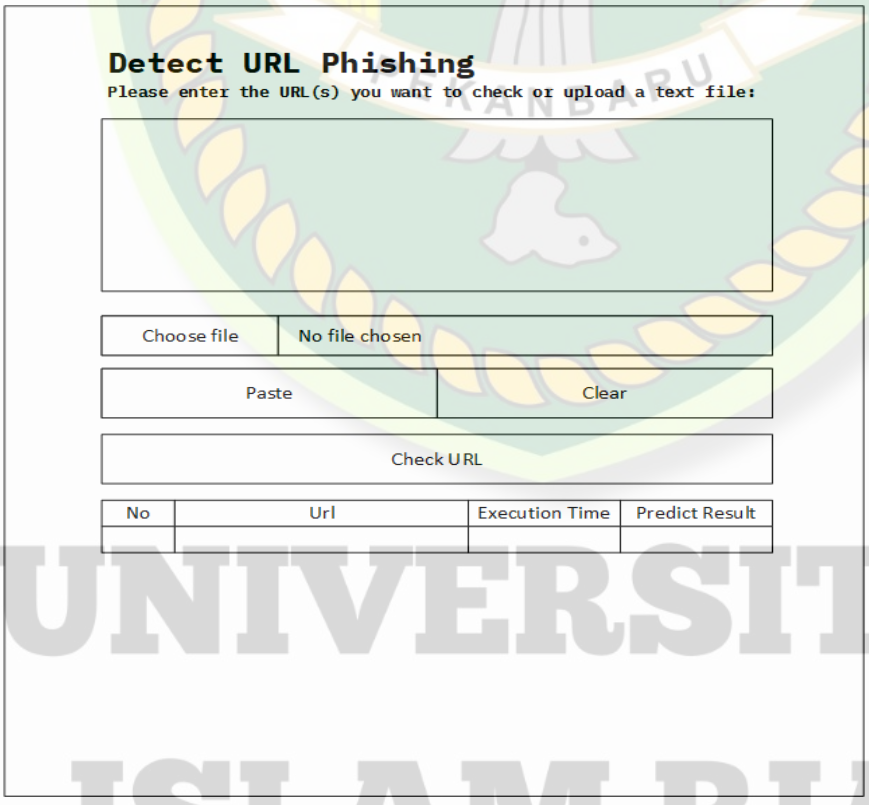
$$recall = \frac{TP}{TP+FN} \quad (3.3)$$

4. F1-Score: merupakan gabungan antara presisi dan recall, memberikan gambaran yang lebih lengkap tentang kinerja model.

$$f1 - score = 2x \frac{Presisi\ x\ Recall}{Presisi + Recall} \quad (3.4)$$

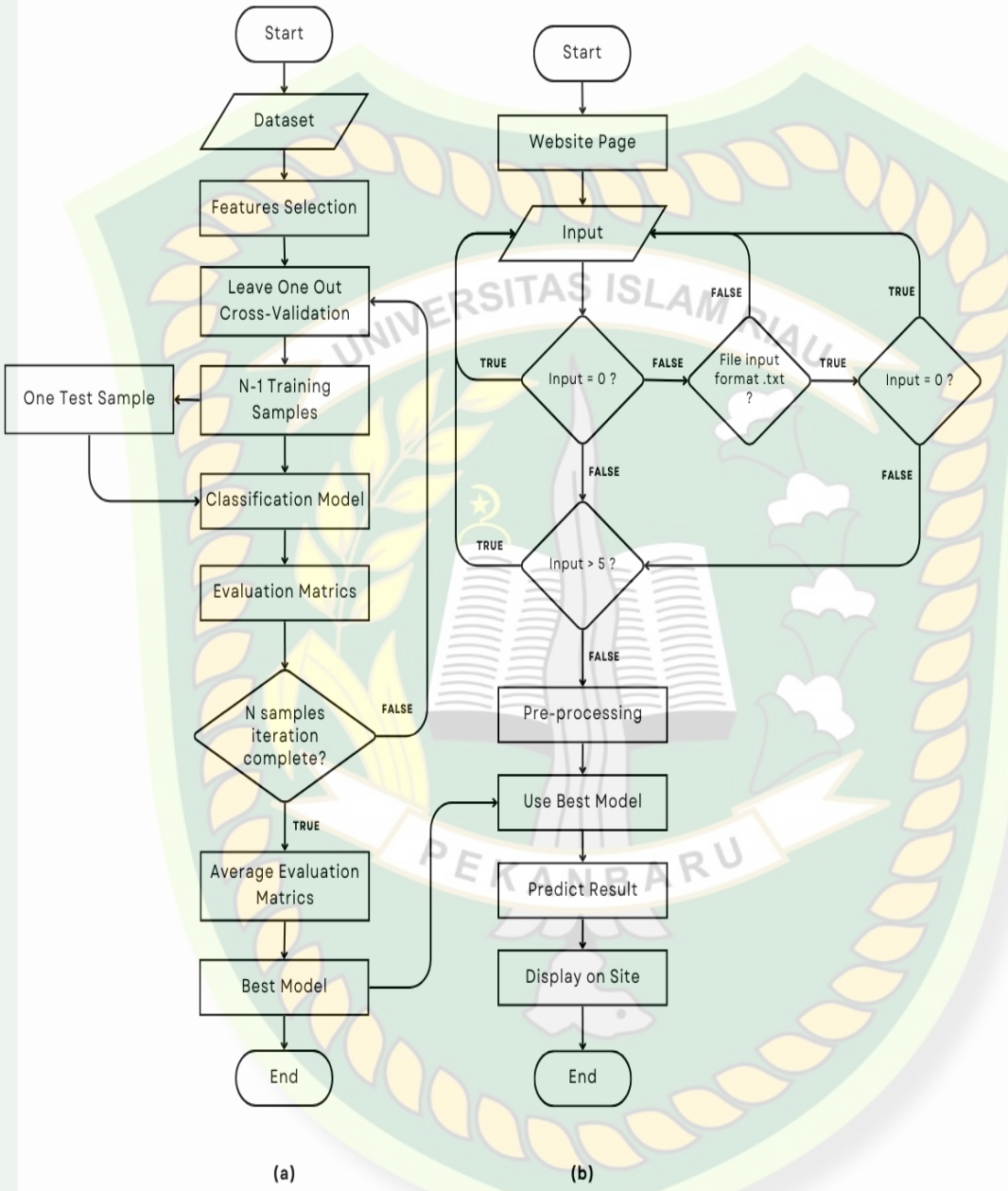
3.4. Pengembangan Sistem

3.4.1. Rancangan Tampilan



Gambar 3. 3 Tampilan rancangan website

3.4.2. *Flowchart modeling* dan implementasi model



Gambar 3. 4 (a) *Flowchart modeling*; (b) *Flowchart* implementasi model

Keterangan :

1. *Flowchart Modeling*

   Merupakan alur untuk membuat model prediksi dengan hasil akhir adalah model terbaik. Berikut penjelasan alur tersebut:

a. Menyiapkan dataset yang sudah di pra-pemrosesan.
b. Melakukan pemilihan fitur yang akan digunakan.
c. Menggunakan teknik LOOCV untuk melatih model dan mengevaluasi model dengan satu sampel sebagai data uji, pada setiap iterasi, sementara data lainnya digunakan sebagai data latih. Proses ini berulang hingga semua data dalam dataset telah diuji secara keseluruhan.
d. Hasil setiap iterasi data dihitung evaluasi metriknya.
e. Jika semua data dalam dataset telah diuji, maka lanjut untuk menghitung rata-rata metrik evaluasi dari keseluruhan iterasi.
f. Menyimpan model dan memilih model terbaik.

2. *Flowchart* Implementasi Model

Merupakan alur implementasi model terbaik pada website deteksi link phishing. Berikut penjelasan alur tersebut:

a. Tampilan halaman website.
b. Input data.
c. Jika inputan kosong (input = 0), maka akan kembali ke halaman inputan data dan menampilkan notifikasi *"Please input URL or upload a .txt file!"*.
d. Jika inputan URL lebih dari 5 (inputan > 5), maka akan kembali ke halaman inputan data dan menampilkan notifikasi *"Maximum 5 URLs can be inputted!"*.
e. Jika inputan berupa file dengan format unggahan tidak sesuai ketentuan (.txt), maka akan kembali ke halaman inputan data dan menampilkan notifikasi *"Invalid file format! Please upload a .txt file"*.

f. Jika format file inputan valid (.txt) tetapi isi file kosong (input = 0), maka akan kembali ke halaman inputan data dan menampilkan notifikasi *"The uploaded file is empty!"*.

g. Jika semua ketentuan terpenuhi, maka tahap selanjutnya melakukan pra-pemrosesan (normalisasi URL, ekstrak fitur URL).

h. Selanjutnya melakukan prediksi dengan menggunakan model terbaik yang dipilih.

i. Kemudian mendapatkan hasil prediksi dan menampilkan hasil prediksi pada halaman website.

3.5. Jadwal kegiatan penelitian

Jadwal kegiatan pada penelitian ini dilakukan dalam jangka waktu kurang lebih lima bulan. Rincian rencana kegiatan pada penelitian ini dapat dilihat pada tabel di bawah ini :

Tabel 3. 6 Jadwal kegiatan penelitian

Kegiatan	Waktu Kegiatan (per Minggu)																			
	Maret 2023				April 2023				Juni 2023				Juli 2023				September 2023			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Identifikasi Permasalahan, Rumusan Masalah, Objek Permasalahan																				
Studi Literatur dan Teori Dasar																				
Rancangan Penelitian																				
Pengumpulan Data																				
Pengolahan Data																				
Analisis Hasil																				
Pembuatan Laporan Hasil Penelitian																				

BAB IV
HASIL DAN PEMBAHASAN

Bab ini akan menjelaskan hasil dan analisis dari penelitian yang telah dilakukan, termasuk eksplorasi yang mendalam terhadap data dan evaluasi terhadap temuan yang diperoleh. Hasil dari setiap tahap metodologi yang dilakukan akan diuraikan secara sistematis untuk memberikan pemahaman yang komprehensif.

4.1. Hasil Normaliasi Protokol dan *Redirect*

Normalisasi protokol dan *redirect* dari dataset ini memiliki peran yang signifikan untuk mempermudah tahap ekstraksi fitur, Hasil normalisasi protokol dan *redirect* dari dataset membantu dalam melakukan pengkategorian klasifikasi fitur dengan lebih efektif. Adapaun hasil dari proses normalisasi protokol dan *redirect* dapat dilihat dibawah ini.

	status	url	normalisasi_protokol	normalisasi_redirect
0	Legitimate	bit.ly/SatuKato-Siak	http://bit.ly/SatuKato-Siak	https://satukato.siakkab.go.id/login
1	Phishing	https://mlcro-out-look-verify.pages.dev/robots...	https://mlcro-out-look-verify.pages.dev/robots...	https://mlcro-out-look-verify.pages.dev/robots...
2	Legitimate	s.id/GooglePlayGames	https://s.id/GooglePlayGames	https://play.google.com/googleplaygames?utm_so...
3	Phishing	https://claim-dagethadiah.eventt.asia/	http://claim-dagethadiah.eventt.asia/	http://claim-dagethadiah.eventt.asia/
4	Legitimate	web.sipa-ftuir.com/seminar/index.php/register	https://web.sipa-ftuir.com/seminar/index.php/r...	https://web.sipa-ftuir.com/seminar/index.php/r...
5	Phishing	https://apple.appleidqs.xyz/	https://apple.appleidqs.xyz/	https://apple.appleidqs.xyz/
6	Legitimate	https://account.kompas.com/login/a29tcGFz/aHR0...	https://account.kompas.com/login/a29tcGFz/aHR0...	https://account.kompas.com/login/a29tcGFz/aHR0...
7	Phishing	https://x.g-code.co.id/	https://x.g-code.co.id/	https://x.g-code.co.id/
8	Legitimate	https://openphish.com/whatsnew.html	https://openphish.com/whatsnew.html	https://openphish.com/whatsnew.html
9	Phishing	http://demo.servernokos.com/	http://demo.servernokos.com/	http://demo.servernokos.com/
10	Legitimate	https://t.me/joinchat/UwdjdlV7TbUp5noLPsK49w	https://t.me/joinchat/UwdjdlV7TbUp5noLPsK49w	https://t.me/joinchat/UwdjdlV7TbUp5noLPsK49w
11	Phishing	https://misty-sun-d394.cuknulikno.workers.dev/	https://misty-sun-d394.cuknulikno.workers.dev/	https://misty-sun-d394.cuknulikno.workers.dev/
12	Legitimate	https://s.id/1YLgh	https://s.id/1YLgh	https://www.vidio.com/users/login
13	Phishing	http://openseapro-nftbox.web.app/	http://openseapro-nftbox.web.app/	https://openseapro-nftbox.web.app/
14	Legitimate	https://react.dev/community/conferences	https://react.dev/community/conferences	https://react.dev/community/conferences
15	Phishing	http://e-devletaidelerim.com/	http://e-devletaidelerim.com/	http://e-devletaidelerim.com/
16	Legitimate	https://police.govt.nz/contact-us/calling-emer...	https://police.govt.nz/contact-us/calling-emer...	https://www.police.govt.nz/contact-us/calling-...
17	Phishing	http://smuanskannyatajs7wjprimespprtmail.dynna...	http://smuanskannyatajs7wjprimespprtmail.dynna...	http://smuanskannyatajs7wjprimespprtmail.dynna...
18	Legitimate	https://pintu.co.id/blog/apa-itu-metaverse-di-...	https://pintu.co.id/blog/apa-itu-metaverse-di-...	https://pintu.co.id/academy/post/apa-itu-metav...
19	Phishing	https://usps.mytrackinggw.com/information	http://usps.mytrackinggw.com/information	http://usps.mytrackinggw.com/information

Gambar 4. 1 Hasil normalisasi protokol dan *redirect*

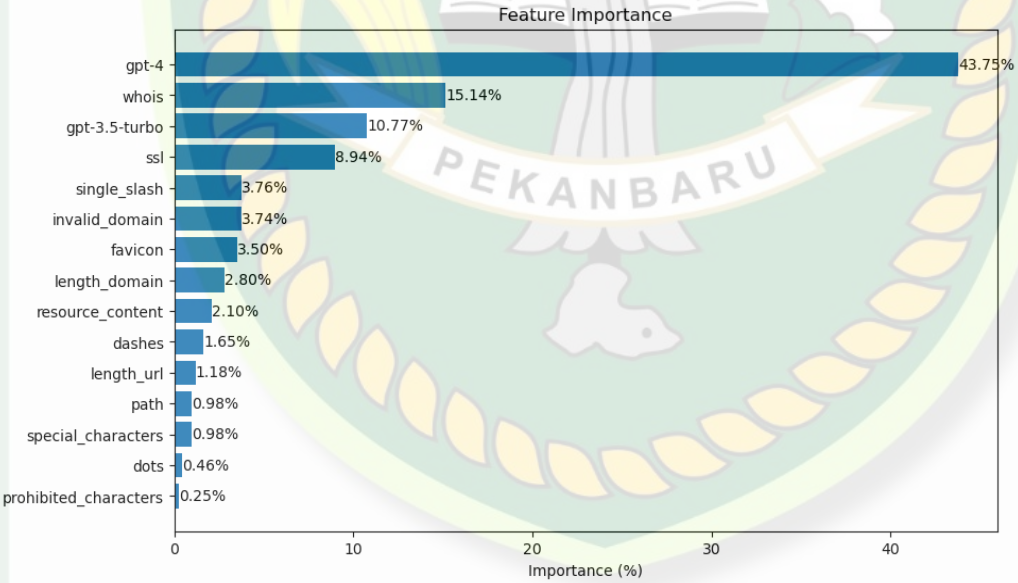
4.2. Hasil Ekstraksi Fitur

Berikut adalah hasil ekstraksi fitur URL berbasis Internal/Karakteristik URL, External/Luar Karakteristik URL, dan OpenAI dari dataset yang dapat dilihat pada gambar dibawah ini.

	url	status	length_domain	dashes	ssl	invalid_domain	prohibited_characters	special_characters	path	single_slash	dots	length_url	resource_content	favicon	whois	gpt-3.5-turbo	gpt-4
0	https://satukato.siakkab.go.id/login	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
1	https://mlcro-out-look-verify.pages.dev/robots...	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
2	https://play.google.com/googleplaygames?utm_so...	1	1	1	1	1	1	0	1	1	1	0	1	0	1	1	1
3	http://claim-dagethadiah.eventt.asia/	0	1	1	0	1	1	1	1	1	1	1	1	0	1	0	0
4	https://web.sipa-ftuir.com/seminar/index.php/r...	1	0	0	1	1	1	1	1	0	1	1	1	0	1	1	1
5	https://apple.appleidqs.xyz/	0	1	1	1	0	1	1	1	1	1	1	1	0	1	0	0
6	https://account.kompas.com/login/a29tcGFz/aHR0...	1	1	1	1	1	1	1	1	0	1	0	1	1	1	0	1
7	https://x.g-code.co.id/	0	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1
8	https://openphish.com/whatsnew.html	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1
9	http://demo.servernokos.com/	0	0	1	0	1	1	1	1	1	1	1	1	0	1	1	0
10	https://t.me/joinchat/UwdjdlV7TbUp5noLPsK49w	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
11	https://misty-sun-d394.cuknulikno.workers.dev/	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
12	https://www.vidio.com/users/login	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
13	https://openseapro-nftbox.web.app/	0	0	0	1	1	1	1	1	1	1	1	1	0	0	1	0
14	https://react.dev/community/conferences	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
15	http://e-devletaidelerim.com/	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0
16	https://www.police.govt.nz/contact-us/calling-...	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
17	http://smuanskannyatajs7wjprimespprtmail.dynna...	0	1	1	0	0	1	1	1	1	1	0	1	0	1	0	0
18	https://pintu.co.id/academy/post/apa-itu-metav...	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1
19	http://usps.mytrackinggw.com/information	0	0	1	0	1	1	1	1	1	1	1	1	0	1	0	0

Gambar 4. 2 Hasil ekstraksi fitur

Dapat dilihat pada gambar 4.2, diperoleh hasil ekstraksi fitur dari Panjang Domain, Tanda Hubung Domain, Protokol HTTPS, Pengecualian Domain, Karakter Terlarang, Jumlah Karakter Khusus, Jumlah Path, Jumlah Single Slash, Jumlah Titik, Panjang URL, Isi Konten, Favicon, Whois, GPT-3.5-turbo, dan GPT-4 dengan nilai 0 sebagai kelas phishing dan 1 sebagai kelas legitimate. Fitur juga dikategorikan sebagai variabel (X) dan status dikategorikan sebagai variabel (Y) yang merupakan acuan atau target dalam melakukan prediksi dari fitur. Fitur-fitur ini akan digunakan dalam pembuatan model evaluasi dengan melakukan pemilihan fitur berdasarkan nilai tertinggi dari hasil tingkat kepentingan fitur *(feature importance)*. Fitur yang dianggap relevan akan digunakan dalam proses pelatihan dan pengujian model dengan metode LOOVC.



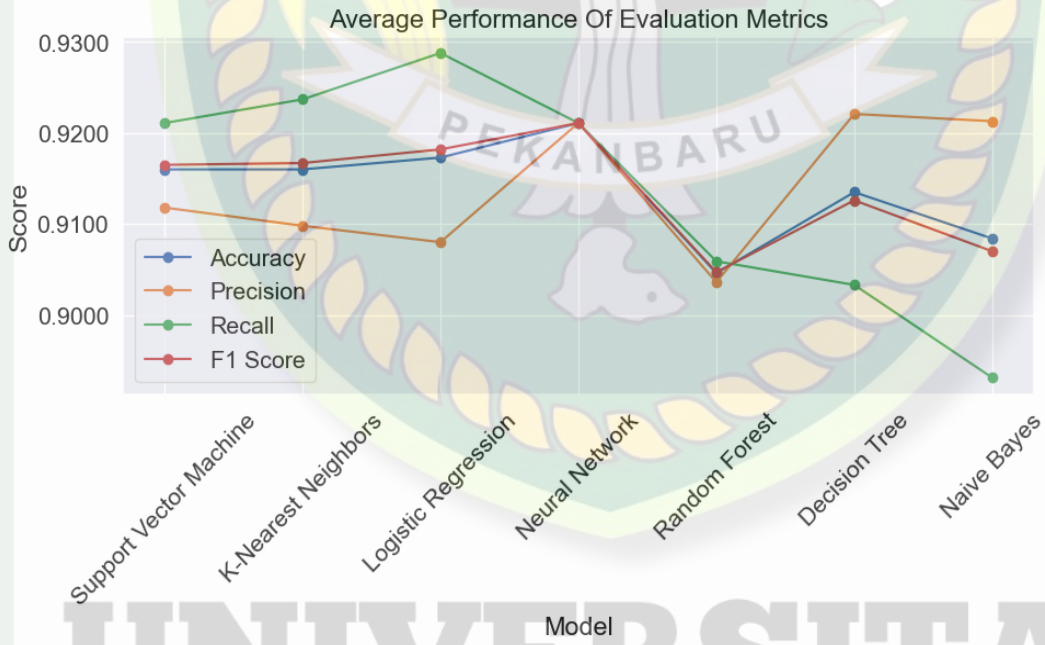
Gambar 4. 3 *Feature importance*

Pada gambar 4.3 terlihat bahwa fitur berbasis OpenAI GPT-4 memperoleh tingkat kepentingan yang sangat tinggi sebesar 43.75%, menunjukkan kinerja yang sangat baik. Kemudian, diurutan kedua, fitur berbasis eksternal Whois mendapatkan nilai sebesar 15.14%.

Pada proses pemilihan fitur, peneliti menggunakan 9 fitur yang dianggap relevan dan efisien untuk pelatihan dan pengujian model. Fitur yang dipilih yaitu, GPT-4, Whois, Protokol HTTPS, Jumlah Single Slash, Pengecualian Domain, Favicon, Panjang Domain, Isi Konten, dan Tanda Hubung Domain. Fitur GPT-3.5-turbo tidak dipilih dalam pemilihan fitur karena pada fitur berbasis OpenAI hanya membandingkan antara fitur GPT-3.5-turbo dan GPT-4, sehingga fitur yang dipilih hanya fitur yang memiliki tingkat kepentingan paling tinggi.

4.3. Evaluasi Model

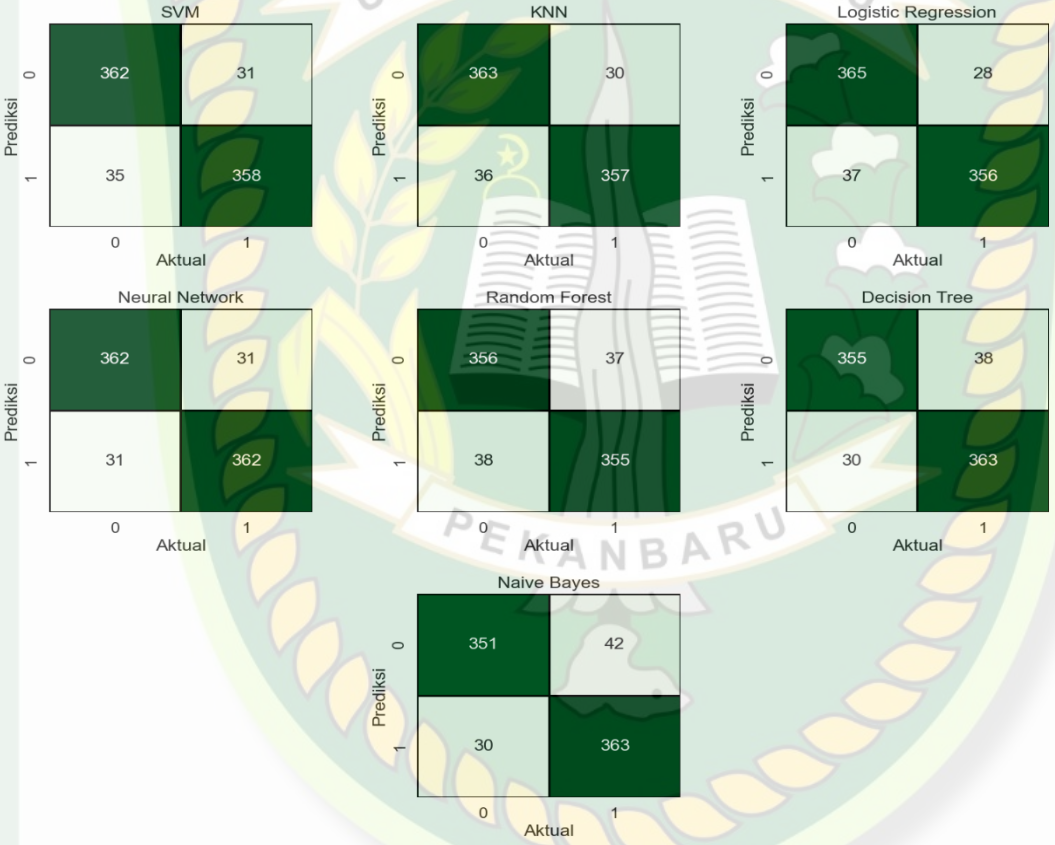
Evaluasi model diperoleh melalui metode LOOCV yang telah dilakukan dengan melibatkan fitur yang dipilih dengan menghasilkan metrik evaluasi akurasi, presisi, recall, dan f1-score.



Gambar 4. 4 Rata-rata evaluasi metrik

Dapat dilihat pada gambar 4.4, kinerja rata-rata metrik evaluasi akurasi, presisi, recall, dan f1-score dari model *Neural Network* terlihat sangat baik,

konsisten dan seimbang. Model *Naïve Bayes* memiliki nilai rata-rata presisi yang tinggi namun recall yang sangat rendah dibandingkan dengan model lainnya. Sementara itu, model *Random Forest* juga menunjukkan konsistensi hampir di setiap hasil metrik evaluasi, namun nilai rata-ratanya lebih rendah dibandingkan dengan model *Support Vector Machine, K-Nearest Neighbors, Logistic Regression* dan *Neural Network.*



Gambar 4. 5 *Confusion Matrix*

Terlihat pada gambar 4.5, kumpulan *confusion matrix* dari masing-masing model yang merupakan hasil rata-rata setiap iterasi evaluasi metode LOOCV. *confusion matrix* ini dapat dihitung nilai akurasi, presisi, recall, dan skor f1-score dengan menggunakan rumus (3.1), (3.2), (3.3), dan (3.4). Hasil dari evaluasi metrik dapat dilihat pada rincian tabel berikut:

Tabel 4. 1 Evaluasi metrik

No	Model	Accuracy (average)	Precision (average)	Recall (average)	F1-Score (average)
1	Support Vector Machine	91.60%	91.18%	92.11%	91.65%
2	K-Nearest Neighbors	91.60%	90.98%	92.37%	91.67%
3	Logistic Regression	91.73%	90.80%	**92.88%**	91.82%
4	Neural Network	**92.11%**	**92.11%**	**92.11%**	**92.11%**
5	Random Forest	90.46%	90.36%	90.59%	90.47%
6	Decision Tree	91.35%	**92.21%**	90.33%	91.26%
7	Naïve Bayes	90.84%	92.13%	89.31%	90.70%

Hasil evaluasi metrik pada tabel 4.1, model *Neural Network* memiliki kinerja terbaik dengan akurasi rata-rata sebesar 92.11%, serta presisi, recall, dan f1-score yang konsisten sebesar 92.11%. Diikuti oleh model *Logistic Regression* yang mencapai akurasi rata-rata sebesar 91.73%, presisi 90.80%, f1-score 91.82%, dan recall tertinggi dibandingkan model lain, yakni 92.88%. Model *K-Nearest Neighbor* dan *Support Vector Machine* menunjukkan hasil serupa dengan akurasi rata-rata masing-masing sebesar 91.60%. *K-Nearest Neighbor* memiliki sedikit keunggulan dalam presisi, recall, dan f1-score dibandingkan *Support Vector Machine*, yakni masing-masing 90.98%, 92.37%, dan 91.67%. Sementara itu, model *Decision Tree* memiliki akurasi rata-rata 91.35%, presisi 92.21%, recall 90.33%, dan f1-score 91.26%. Meskipun memiliki presisi tinggi, model ini sedikit kurang dalam recall, yang berdampak pada nilai f1-score. Model *Random Forest* dan *Naive Bayes* menunjukkan performa yang cukup baik dengan akurasi rata-rata masing-masing sebesar 90.46% dan 90.84%. *Random Forest* memiliki nilai presisi, recall, dan f1-score yang hampir sama, yakni masing-masing 90.36%, 90.59%, dan 90.47%, sedangkan *Naïve Bayes* memiliki recall yang sangat rendah dibandingkan rata-rata model lainnya, yakni 89.31%. Secara keseluruhan model

Neural Network menjadi pilihan terbaik untuk dapat diimplementasi pada Website Deteksi Link Phishing karena memiliki kinerja yang dianggap stabil dan konsisten di semua hasil metrik evaluasi.

4.4. Uji Model dan Implementasi Model Terbaik

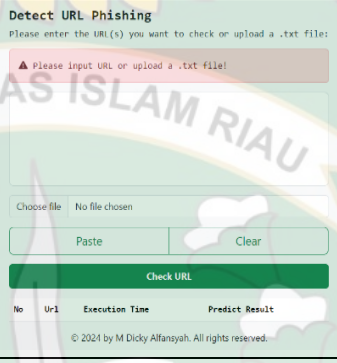
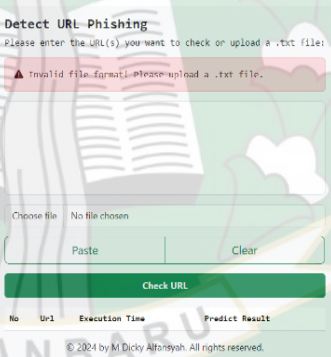
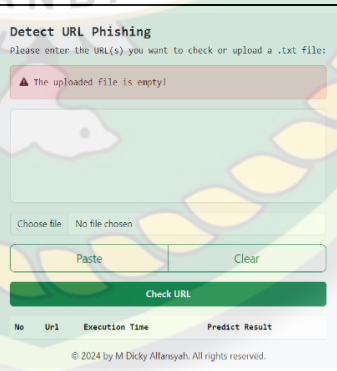
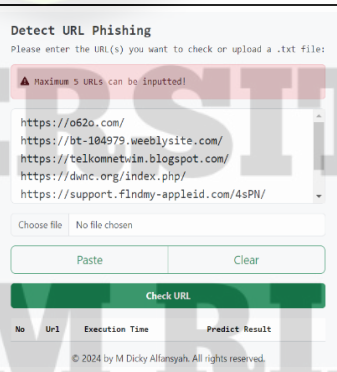
Implementasi model terbaik pada Website Deteksi Link Phishing dilakukan menggunakan *framework Flask*. *Framework* ini merupakan salah satu yang paling populer dan mudah digunakan dalam pengembangan aplikasi web berbasis Python. Dengan *Flask*, pembuatan aplikasi web yang efisien dan tangguh menjadi lebih mudah, serta memungkinkan pemanfaatan berbagai fitur yang disediakan oleh *framework* ini, seperti *routing*, *template rendering*, dan dukungan untuk pengembangan API.

4.4.1. Pengujian *Black Box*

Black Box merupakan metode pengujian untuk melihat fungsi-fungsi yang ada pada sistem tanpa memperhatikan bagaimana fungsi tersebut dibuat. Dalam konteks ini, skenario pengujian merujuk pada serangkaian langkah atau situasi yang dirancang untuk menguji respon sistem dengan output yang diharapkan. Hasil dari skenario yang dilakukan berjalan dengan baik tanpa terjadi masalah atau *error*. Rincian hasil pengujian dapat dilihat pada tabel dibawah ini.

Tabel 4. 2 Skenario pengujian sistem website deteksi link phishing

No	Skenario Pengujian	Hasil Yang Diharapkan	Hasil Pengujian	Ket
1	Pengujian dengan melakukan paste pada inputan URL dengan klik button paste	Sistem akan mempaste URL pada inputan URL	Sesuai Harapan	Berhasil

2	Pengujian dengan menghapus inputan dengan klik button clear	Sistem akan menghapus inputan yang ada pada inputan URL atau inputan file	Sesuai Harapan	Berhasil
3	Pengujian dengan mengosongkan inputan, kemudian klik button check URL	Sistem akan menolak akses dan menampilkan pesan notifikasi *"Please input URL or upload a .txt file!"*		Berhasil
4	Pengujian dengan menginputkan file yang tidak sesuai dengan ketentuan format file	Sistem akan menolak akses dan menampilkan pesan notifikasi *"Invalid file format! Please upload a .txt file"*		Berhasil
5	Pengujian dengan menginputkan file yang sesuai dengan ketentuan format file namun isi file kosong	Sistem akan menolak akses dan menampilkan pesan notifikasi *"The uploaded file is empty!"*		Berhasil
6	Pengujian dengan menginputkan inputan lebih dari ketentuan yang ditetapkan yaitu maksimal inputan sebanyak 5	Sistem akan menolak akses dan menampilkan pesan notifikasi *"Maximum 5 URLs can be inputted!"*		Berhasil

4.4.2. Uji Model

Pada tahap ini, dilakukan pengujian menggunakan model klasifikasi yang telah dilatih untuk membandingkan kinerja masing-masing model. Berikut adalah hasil pengujian perbandingan kinerja masing-masing model dengan menggunakan sampel 100 URL yang dapat dilihat pada tabel dibawah ini.

Tabel 4. 3 Perbandingan model

No	Url	Aktual	Prediksi						
			Support Vactor Machine	k-Nearest Neighbors	Logistic Regression	Neural Network	Random Forest	Decision Tree	Naïve Bayes
1	http://bet895202400.com/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
2	https://gdhtb.blogspot.md/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
3	https://ussp.usspaoo.top/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
4	http://cbxupkftqm.duckdns.org/en/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
5	http://cjonkhdqsp.duckdns.org/en/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
6	https://wxmmtls.onflashdrive.app/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
7	http://discord.taojay315.workers.dev/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
8	https://bmqehg.blogspot.is/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
9	https://fgmhvn.blogspot.hr/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
10	https://apps-auth-coinbase.webflow.io/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
11	http://vdbnh.blogspot.si/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
12	http://uspz.usspaoh.top/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
13	http://www-tokenpoket.com/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing

No	Url	Aktual	Prediksi						
			Support Vactor Machine	k-Nearest Neighbors	Logistic Regression	Neural Network	Random Forest	Decision Tree	Naïve Bayes
14	http://deutschebank-aktualisierung.com/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
15	https://pagina.pro/AI-Facebook	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
16	https://bq0--ut-b-j.nedizl.my.id/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
17	http://124.181.42.106/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
18	https://wwe.iolitrdse.cloudns.biz/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
19	http://boxmint-project.web.app/	Phishing	Phishing	Legitimate	Phishing	Phishing	Phishing	Phishing	Phishing
20	https://supportfr.pages.dev/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
21	https://singaporevideoasia.tele-groups.com/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
22	https://saiteja7065.github.io/Netflix-Clone	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
23	https://3as6sd5.xyz/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
24	https://danaxid-costumerr.webnew.my.id/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
25	https://usps.postbv.com/update	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
26	https://arma.52ker.my.id/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
27	https://danaa-indonesia.bantuan-official.biz.id/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
28	https://uniswapywj5.pages.dev/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
29	https://tgadminuser.webaab.vip/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
30	https://tgadminuser.webaab.pw/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
31	https://tgadminuser.webaab.top/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
32	http://ingresar4.hstn.me/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
33	http://imtoken-ae.net/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
34	https://dbmobile-aktivieren.app/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
35	https://dana-resmii.webnet.my.id/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing

No	Url	Aktual	Prediksi						
			Support Vactor Machine	k-Nearest Neighbors	Logistic Regression	Neural Network	Random Forest	Decision Tree	Naïve Bayes
36	https://west.linktain.skin/handle	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
37	https://servicesclients.online/KA/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
38	http://www.verifybyclubhouse.com/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
39	http://cliente.deletrica.com.br/ex/rebrand/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Legitimate
40	https://elite-kyiv.site/ukrain	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
41	https://vesna.cartina-deti.homes/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
42	https://fajwever.store/vote	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
43	http://facebook-com-2ym.pages.dev/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
44	https://postlogisticstrack.com/	Phishing	Legitimate	Phishing	Legitimate	Legitimate	Phishing	Phishing	Phishing
45	http://iusman00.github.io/netflix	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
46	http://bhthth.blogspot.fi/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
47	https://tinyurl.com/mr44th5e	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
48	http://east.linkfaq.info/1284	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
49	http://east.linkttags.click/1284	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
50	https://kosiarki-sikorski.pl/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
51	https://cntralem.ydns.eu/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
52	http://uua.pr8hm7.top/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
53	https://virallnnn39.click-me44.biz.id/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
54	https://steamcomunai.com/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
55	https://steamcomunai.com/gifts/activate	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
56	https://pool-nftbox.web.app/	Phishing	Phishing	Legitimate	Phishing	Phishing	Phishing	Phishing	Phishing
57	https://1pld.hoangducdung.com/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing

No	Url	Aktual	Prediksi						
			Support Vactor Machine	k-Nearest Neighbors	Logistic Regression	Neural Network	Random Forest	Decision Tree	Naïve Bayes
58	http://k9nq6.shop/	Phishing	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
59	https://tp.usdtwebs.top/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
60	https://loginfacebook.com.tr/	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
61	https://bit.ly/3U5KUv9	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
62	https://id.wikipedia.org/wiki/Log_masuk	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
63	https://secure.login.gov/	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
64	https://bit.ly/3VLLwas	Legitimate	Legitimate	Phishing	Phishing	Legitimate	Legitimate	Legitimate	Phishing
65	https://www.ustraveldocs.com/	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
66	https://www.netflix.com/id-en/	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
67	https://cloud.digitalocean.com/login	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Phishing	Phishing	Phishing
68	https://bit.ly/3VOoy2p	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
69	https://www.icloud.com/	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
70	https://developer.apple.com/sign-in-with-apple/	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
71	https://bit.ly/3J4j8Jg	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Phishing
72	https://music.apple.com/us/login	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
73	https://soundcloud.com/djlogin	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
74	https://music.youtube.com/history	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
75	https://amzn.to/3VN9GkP	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Phishing
76	https://bit.ly/3vycuYH	Legitimate	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing	Phishing
77	https://ces.du.ac.in/index.php/site/login	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
78	https://journal.uir.ac.id/index.php/ITJRD/login	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Phishing	Legitimate
79	https://cabi.scienceconnect.io/login	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate

No	Url	Aktual	Prediksi						
			Support Vactor Machine	k-Nearest Neighbors	Logistic Regression	Neural Network	Random Forest	Decision Tree	Naïve Bayes
80	https://www.codashop.com/id-id/free-fire	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
81	https://www.lapakgaming.com/id-id/free-fire	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
82	https://www.unipin.com/id/garena/free-fire	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
83	https://itemku.com/daftar	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
84	https://colab.research.google.com/	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
85	https://id.aliexpress.com/?gatewayAdapt=glo2idn	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
86	https://bit.ly/3J6y1ur	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
87	https://bit.ly/4alxLnj	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
88	https://www.bankofamerica.com/	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
89	https://digital.bankofsingapore.com/clfeweb/	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
90	https://bit.ly/43MWNJC	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	**Phishing**	Legitimate
91	https://bit.ly/uob-malaysia-login	Legitimate	**Phishing**	Legitimate	**Phishing**	**Phishing**	**Phishing**	**Phishing**	**Phishing**
92	https://sign.peruri.co.id/account/auth/login/	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	**Phishing**	Legitimate
93	https://www.telkomsel.com/login/auth	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
94	https://my.signinapp.com/login	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
95	https://bit.ly/3PU0Y0G	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
96	https://bit.ly/3xsso7u	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	**Phishing**
97	https://fhcibumn.com/login	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
98	https://music.tiktok.com/login	Legitimate	**Phishing**	**Phishing**	**Phishing**	Legitimate	**Phishing**	Legitimate	Legitimate
99	https://www.pinterest.com/login/	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate	Legitimate
100	https://login.beritajakarta.id/account/login	Legitimate	**Phishing**	**Phishing**	**Phishing**	**Phishing**	**Phishing**	**Phishing**	**Phishing**

Hasil yang disajikan pada tabel 4.3 menunjukkan kinerja perbandingan model klasifikasi dengan kemampuan yang bervariasi dalam memprediksi kelas phishing dan legitimate, meskipun masih terdapat beberapa prediksi yang tidak sesuai harapan. Hasil pengujian yang dilakukan menggunakan 100 URL yang terdiri dari 60 sampel URL phishing dan 40 sampel URL legitimate, didapatkan *confusion matrix* dengan jumlah prediksi yang benar yaitu *True Positive* (prediksi benar URL phishing) dan *True Negative* (prediksi benar URL legitimate) serta prediksi yang salah yaitu *False Positive* (prediksi salah URL phishing) dan *False Negative* (prediksi salah URL legitimate) dari masing-masing model. Berikut adalah rincian *confusion matrix* yang dapat dilihat pada tabel dibawah ini:

Tabel 4. 4 Confusion matrix hasil uji model

No	Model	True Positive (TP)	False Positive (FP)	False Negative (FN)	True Negative (TN)
1	Support Vactor Machine	58	2	4	36
2	K-Nearest Neighbors	57	3	4	36
3	Logistic Regression	58	2	5	35
4	Neural Network	58	2	3	37
5	Random Forest	59	1	5	35
6	Decision Tree	59	1	7	33
7	Naïve Bayes	58	2	8	32

Confusion matrix yang disajikan pada tabel 4.4 dapat dihitung menggunakan rumus (3.1), (3.2), (3.3), dan (3.4), untuk mendapatkan nilai metrik akurasi, presisi, recall, dan f1-score dari masing-masing model, Berikut adalah hasil dari perhitungan tersebut:

Tabel 4. 5 Evaluasi metrik uji model

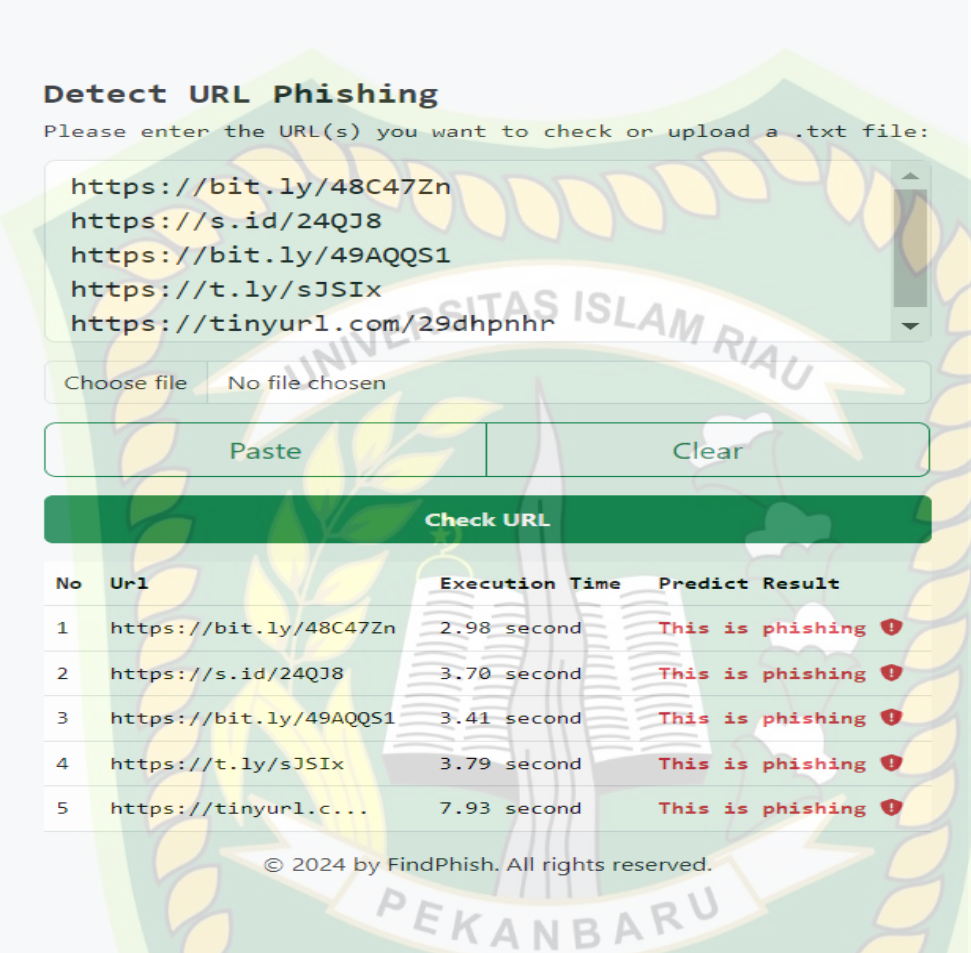
No	Model	Accuracy (average)	Precision (average)	Recall (average)	F1-Score (average)
1	Support Vector Machine	94.00%	96.67%	93.55%	95.08%
2	K-Nearest Neighbors	93.00%	95.0%	93.44%	94.21%
3	Logistic Regression	93.00%	96.67%	92.06%	94.31%
4	Neural Network	**95.00%**	**96.67%**	**95.08%**	**95.87%**
5	Random Forest	94.00%	**98.33%**	92.19%	95.16%
6	Decision Tree	92.00%	**98.33%**	89.39%	93.65%
7	Naïve Bayes	90.00%	96.67%	87.88%	92.07%

Dapat dilihat pada Tabel 4.5, hasil perhitungan akurasi, presisi, recall, dan f1-score menunjukkan bahwa *Neural Network* memiliki kinerja yang sangat baik dalam memprediksi 100 URL dibandingkan kinerja model klasifikasi lainnya.

4.4.3. Implementasi Model Terbaik

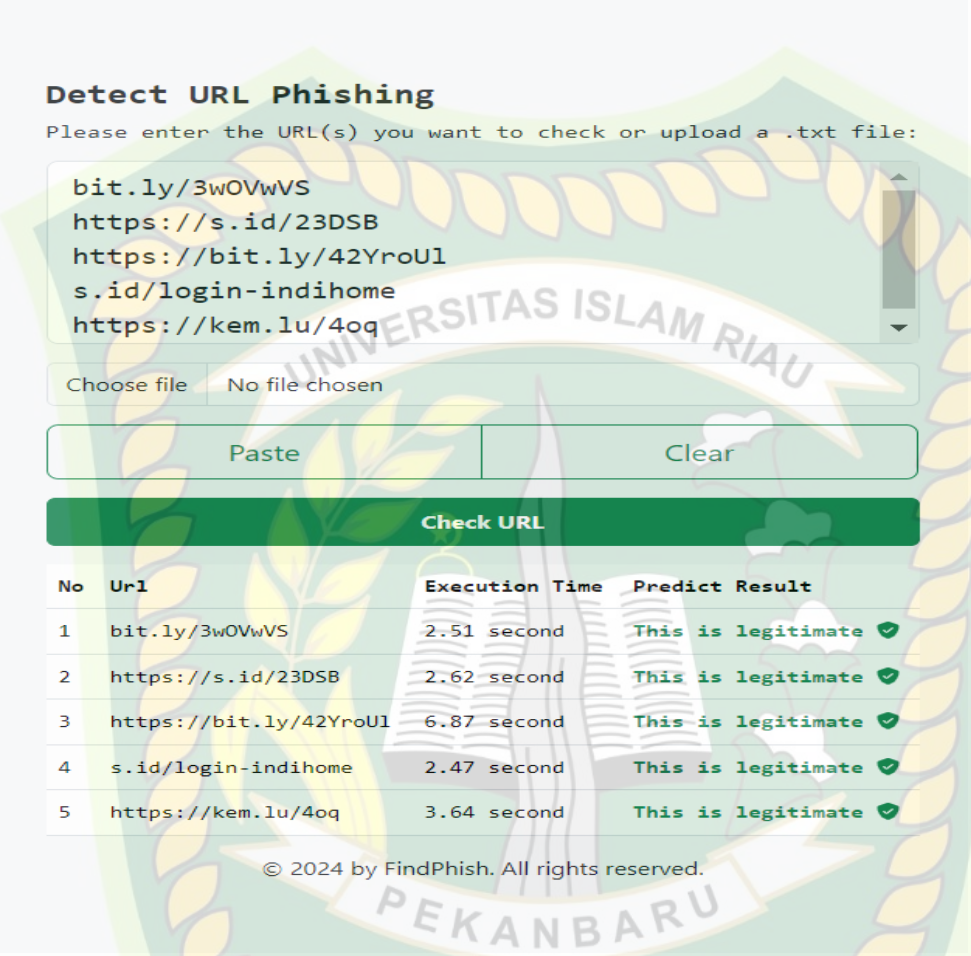
Pada tahap ini, pengujian model terbaik, yaitu model *Neural Network*, dilakukan pada Website Deteksi Link Phishing dengan menggunakan 3 bentuk simulasi sebagai berikut:

1. Melakukan pengujian model terbaik yang diimplementasiklan pada Website Deteksi Link Phishing dengan menggunakan inputan sebanyak 5 URL yang merupakan kelas phishing. Pengujian ini dilakukan dengan cara memprediksi URL phishing yang disembunyikan pada URL penyingkat *(short URL)*, URL yang di input diproses pada tahap normalisasi (normalisasi protokol dan *redirect*) selanjutnya dilakukan ekstraksi fitur pada setiap inputan URL, kemudian hasil ekstraksi fitur diproses menggunakan model terbaik untuk mendapatkan hasil prediksi. Hasil prediksi yang diharapkan adalah kelas phishing secara keseluruhan.



Gambar 4. 6 Hasil prediksi short url phishing

2. Melakukan pengujian model terbaik yang diimplementasiklan pada Website Deteksi Link Phishing dengan menggunakan inputan sebanyak 5 URL yang merupakan kelas legitimate. Pengujian ini dilakukan dengan cara memprediksi URL phishing yang disembunyikan pada URL penyingkat *(short URL)*, URL yang di input diproses pada tahap normalisasi (normalisasi protokol dan *redirect*) selanjutnya dilakukan ekstraksi fitur pada setiap inputan URL, kemudian hasil ekstraksi fitur diproses menggunakan model terbaik untuk mendapatkan hasil prediksi. Hasil prediksi yang diharapkan adalah kelas legitimate secara keseluruhan.



Gambar 4. 7 Hasil prediksi short url legitimate

3. Melakukan pengujian model terbaik yang diimplementasiklan pada Website Deteksi Link Phishing dengan menggunakan inputan kelas phishing sebanyak 3 URL dan kelas legitimate sebanyak 2 URL. Pengujian ini dilakukan dengan cara memprediksi URL tanpa disembuntikan pada URL pentingkat *(short URL)*, URL yang di input diproses pada tahap normalisasi (normalisasi protokol dan *redirect*) selanjutnya dilakukan ekstraksi fitur pada setiap inputan URL, kemudian hasil ekstraksi fitur diproses menggunakan model terbaik untuk mendapatkan hasil prediksi. Hasil prediksi yang diharapkan adalah 3 phishing dan 2 legitimate.

Gambar 4. 8 Hasil prediksi dengan url asli

Implementasi model *Neural Network* pada Website Deteksi Link Phishing dalam pengujian terhadap 15 url berhasil dilakukan, terbukti mendapatkan hasil yang sangat baik, sehingga model ini dapat diimplementasikan.

BAB V
KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan hasil analisis dan implementasi model terbaik pada Website Deteksi Link Phishing maka dapat disimpulkan:

1. Fitur berbasis OpenAI memiliki tingkat kepentingan *(feature importance)* yang sangat baik, dengan fitur GPT-4 mendapatkan nilai sebesar 43,75%. Hal ini menunjukkan bahwa fitur berbasis OpenAI sangat efektif digunakan untuk meningkatkan kinerja model.
2. Model *Neural Network* mendapatkan tingkat kinerja tertinggi dengan akurasi, presisi, recall, dan f1-score sebesar 92.11%, menunjukkan kemampuan yang sangat baik dalam mengklasifikasikan data secara konsisten dan akurat.
3. Hasil skenario pengujian *black box* diketahui bahwa aplikasi deteksi link phishing berhasil dibangun tanpa terjadi masalah atau *error*. Proses pengujian dilakukan dengan cermat untuk memastikan fungsi-fungsi bekerja dengan baik.
4. Pengujian model klasifikasi menggunakan sampel 100 URL mendapatkan model *Neural Network* bekerja dengan sangat baik dalam melakukan deteksi link phishing dibandingkan dengan model lainnya.
5. Implementasi model terbaik *Neural Network* pada Website Deteksi Link Phishing menggunakan 15 url yang diuji berhasil memprediksi dengan akurat, ini menunjukkan bahwa model yang diimplementasikan efektif dalam mengidentifikasi link phishing.

5.2. Saran

Dari kesimpulan yang telah diambil, dapat diusulkan beberapa saran rekomendasi yang akan mendukung perkembangan sistem ini:

1. Meningkatkan kinerja model dengan menambahkan variasi data lebih banyak dan fitur-fitur ekstraksi yang digunakan untuk mencapai model yang optimal.
2. Aplikasi ini dapat dikembangkan lagi dengan meningkatkan kualitas *user interface* dan *user experience* untuk memastikan pengguna memiliki pengalaman yang lebih intuitif dan memuaskan.

DAFTAR PUSTAKA

Ahammad, S. K. H., Kale, S. D., Upadhye, G. D., Pande, S. D., Babu, E. V., Dhumane, A. V, & Bahadur, Mr. D. K. J. (2022). Phishing URL detection using machine learning methods. *Advances in Engineering Software*, *173*, 103288. https://doi.org/https://doi.org/10.1016/j.advengsoft.2022.103288

Bhoj, N., Bawari, R., Tripathi, A., & Sahai, N. (2021). Naive and Neighbour Approach for Phishing Detection. *2021 10th IEEE International Conference on Communication Systems and Network Technologies (CSNT)*, 171–175. https://doi.org/10.1109/CSNT51715.2021.9509566

Binti Md Noh, N., & Bin M. Basri, M. N. (2021). Phishing Website Detection Using Random Forest and Support Vector Machine: A Comparison. *2021 2nd International Conference on Artificial Intelligence and Data Sciences (AiDAS)*, 1–5. https://doi.org/10.1109/AiDAS53897.2021.9574282

BOUIJIJ, H., & BERQIA, A. (2021). Machine Learning Algorithms Evaluation for Phishing URLs Classification. *2021 4th International Symposium on Advanced Electrical and Communication Technologies (ISAECT)*, 1–5. https://doi.org/10.1109/ISAECT53699.2021.9668489

Chandra, A., Gregorius, Immanuel, M. S. J., Gunawan, A. A. S., & Anderies. (2022). Accuracy Comparison of Different Machine Learning Models in Phishing Detection. *2022 5th International Conference on Information and Communications Technology (ICOIACT)*, 24–29. https://doi.org/10.1109/ICOIACT55506.2022.9972107

Chiew, K. L., Tan, C. L., Wong, K., Yong, K. S. C., & Tiong, W. K. (2019). A new hybrid ensemble feature selection framework for machine learning-based phishing detection system. *Information Sciences*, *484*, 153–166. https://doi.org/https://doi.org/10.1016/j.ins.2019.01.064

Cybercrime Information Center. (2023). *Phishing Activity in Top-level Domains (TLDs) May 1, 2023 - July 31, 2023*. https://www.cybercrimeinfocenter.org/phishing-activity-in-tlds-may-july-2023

Fayoumi, M. Al, Odeh, A., Keshta, I., Aboshgifa, A., AlHajahjeh, T., &

Abdulraheem, R. (2022). Email phishing detection based on naïve Bayes, Random Forests, and SVM classifications: A comparative study. *2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC)*, 7–11. https://doi.org/10.1109/CCWC54503.2022.9720757

Gupta, V., Mishra, V. K., Singhal, P., & Kumar, A. (2022). An Overview of Supervised Machine Learning Algorithm. *2022 11th International Conference on System Modeling & Advancement in Research Trends (SMART)*, 87–92. https://doi.org/10.1109/SMART55829.2022.10047618

Indonesia Anti-Phishing Data Exchange (IDADX). (2023a). *LAPORAN AKTIVITAS PHISHING Q1 2023*. https://api.idadx.id/documents/uploads/1689234933_Laporan%20Q1%202023.pdf.pdf

Indonesia Anti-Phishing Data Exchange (IDADX). (2023b). *LAPORAN AKTIVITAS PHISHING Q2 2023*. https://api.idadx.id/documents/uploads/1697785994_Laporan%20Q2%202023.pdf.pdf

Jain, A. K., & Gupta, B. B. (2019). A machine learning based approach for phishing detection using hyperlinks information. *Journal of Ambient Intelligence and Humanized Computing*, *10*(5), 2015–2028. https://doi.org/10.1007/s12652-018-0798-z

Jalil, S., Usman, M., & Fong, A. (2023). Highly accurate phishing URL detection based on machine learning. *Journal of Ambient Intelligence and Humanized Computing*, *14*(7), 9233–9251. https://doi.org/10.1007/s12652-022-04426-3

Kaddoura, S. (2021). Classification of malicious and benign websites by network features using supervised machine learning algorithms. *2021 5th Cyber Security in Networking Conference (CSNet)*, 36–40. https://doi.org/10.1109/CSNet52717.2021.9614273

Mandadi, A., Boppana, S., Ravella, V., & Kavitha, R. (2022). Phishing Website Detection Using Machine Learning. *2022 IEEE 7th International conference for Convergence in Technology (I2CT)*, 1–4. https://doi.org/10.1109/I2CT54291.2022.9824801

Mustafa, T., & Karabatak, M. (2023). Feature Selection for Phishing Website by

Using Naive Bayes Classifier. *2023 11th International Symposium on Digital Forensics and Security (ISDFS)*, 1–4. https://doi.org/10.1109/ISDFS58141.2023.10131884

Novakovic, J., & Markovic, S. (2022). Detection of URL-based Phishing Attacks Using Neural Networks. *2022 International Conference on Theoretical and Applied Computer Science and Engineering (ICTASCE)*, 132–136. https://doi.org/10.1109/ICTACSE50438.2022.10009645

Okfalisa, Gazalba, I., Mustakim, & Reza, N. G. I. (2017). Comparative analysis of k-nearest neighbor and modified k-nearest neighbor algorithm for data classification. *2017 2nd International conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, 294–298. https://doi.org/10.1109/ICITISEE.2017.8285514

Orunsolu, A. A., Sodiya, A. S., & Akinwale, A. T. (2022). A predictive model for phishing detection. *Journal of King Saud University - Computer and Information Sciences*, *34*(2), 232–247. https://doi.org/https://doi.org/10.1016/j.jksuci.2019.12.005

Pandey, M. K., Singh, M. K., Pal, S., & Tiwari, B. B. (2023). Prediction of phishing websites using machine learning. *Spatial Information Research*, *31*(2), 157–166. https://doi.org/10.1007/s41324-022-00489-8

Rambabu, V., Malathi, K., & Mahaveerakannan, R. (2022). An Innovative Method to Predict the Accuracy of Phishing Websites by Comparing Logistic Regression Algorithm with Support Vector Machine Algorithm. *2022 6th International Conference on Electronics, Communication and Aerospace Technology*, 646–650. https://doi.org/10.1109/ICECA55336.2022.10009351

Resiandi, K., Murakami, Y., & Nasution, A. H. (2022). A Neural Network Approach to Create Minangkabau-Indonesia Bilingual Dictionary. Dalam M. Melero, S. Sakti, & C. Soria (Ed.), *Proceedings of the 1st Annual Meeting of the ELRA/ISCA Special Interest Group on Under-Resourced Languages* (hlm. 122–128). European Language Resources Association. https://aclanthology.org/2022.sigul-1.16

Sahingoz, O. K., Buber, E., Demir, O., & Diri, B. (2019). Machine learning based phishing detection from URLs. *Expert Systems with Applications*, *117*, 345–

357. https://doi.org/https://doi.org/10.1016/j.eswa.2018.09.029

Sarasjati, W., Rustad, S., Purwanto, Santoso, H. A., Muljono, Syukur, A., Rafrastara, F. A., & Ignatius Moses Setiadi, D. R. (2022). Comparative Study of Classification Algorithms for Website Phishing Detection on Multiple Datasets. *2022 International Seminar on Application for Technology of Information and Communication (iSemantic)*, 448–452. https://doi.org/10.1109/iSemantic55962.2022.9920475

Sindhu, S., Patil, S. P., Sreevalsan, A., Rahman, F., & N., Ms. S. A. (2020). Phishing Detection using Random Forest, SVM and Neural Network with Backpropagation. *2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE)*, 391–394. https://doi.org/10.1109/ICSTCEE49637.2020.9277256

Soumya, T. R., P, R., Rosy, N. A., Pughazendi, N., Padmapriya, S., & Khilar, R. (2022). Logistic Regression based Machine Learning Technique for Phishing Website Detection. *2022 4th International Conference on Inventive Research in Computing Applications (ICIRCA)*, 683–686. https://doi.org/10.1109/ICIRCA54612.2022.9985643

Wahyudi, D., Niswar, M., & Alimuddin, A. A. P. (2022). WEBSITE PHISING DETECTION APPLICATION USING SUPPORT VECTOR MACHINE (SVM). *Journal of Information Technology and Its Utilization*, *5*(1), 18–24. https://doi.org/10.56873/jitu.5.1.4836

SURAT KEPUTUSAN DEKAN FAKULTAS TEKNIK UNIVERSITAS ISLAM RIAU
NOMOR : 0174/KPTS/FT-UIR/2024
TENTANG PENGANGKATAN TIM PEMBIMBING PENELITIAN DAN PENYUSUNAN SKRIPSI

DEKAN FAKULTAS TEKNIK

Membaca : Surat Ketua Program Studi Teknik Informatika Nomor : 17/TA-TI/FT/2023 tentang persetujuan dan usulan pengangkatan Tim Pembimbing penelitian dan penyusunan Skripsi.

Menimbang : 1. Bahwa untuk menyelesaikan perkuliahan bagi mahasiswa Fakultas Teknik perlu membuat Skripsi.

2. Untuk itu perlu ditunjuk Tim Pembimbing penelitian dan penyusunan Skripsi yang diangkat dengan Surat Keputusan Dekan.

Mengingat : 1. Undang - Undang Nomor 12 Tahun 2012 Tentang Pendidikan Tinggi
2. Peraturan Presiden Republik Indonesia Nomor 8 Tahun 2012 Tentang Kerangka Kualifikasi Nasional Indonesia
3. Peraturan Pemerintah Republik Indonesia Nomor 37 Tahun 2009 Tentang Dosen
4. Peraturan Pemerintah Republik Indonesia Nomor 66 Tahun 2010 Tentang Pengelolaan dan Penyelenggaraan Pendidikan
5. Peraturan Menteri Pendidikan Nasional Nomor 63 Tahun 2009 Tentang Sistem Penjaminan Mutu Pendidikan
6. Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 49 Tahun 2014 Tentang Standar Nasional Pendidikan Tinggi
7. Statuta Universitas Islam Riau Tahun 2018
8. Peraturan Universitas Islam Riau Nomor 001 Tahun 2018 Tentang Ketentuan Akademik Bidang Pendidikan Universitas Islam Riau

MEMUTUSKAN

Menetapkan : 1. Mengangkat saudara-saudara yang namanya tersebut dibawah ini sebagai Tim Pembimbing Penelitian & penyusunan Skripsi Mahasiswa Fak. Teknik Program Studi Teknik Informatika.

No	Nama	Pangkat	Jabatan
1.	Dr. Arbi Haza Nasution, B.IT., M.IT.	Lektor Kepala	Pembimbing

2. Mahasiswa yang akan dibimbing :

Nama : M Dicky Alfansyah
NPM : 183510468
Program Studi : Teknik Informatika
Jenjang Pendidikan : Strata Satu (S1)
Judul Skripsi : Perbandingan Kinerja Algoritma Klasifikasi Dan Implementasi Model Terbaik Pada Website Deteksi Link Phishing

3. Keputusan ini mulai berlaku pada tanggal ditetapkannya dengan ketentuan bila terdapat kekeliruan dikemudian hari segera ditinjau kembali.

Ditetapkan di : Pekanbaru
Pada Tanggal : 12 Sya'ban 1445 H
22 Februari 2024 M

Dekan,

Prof. Dr. Eng. Ir. Muslim.,ST.,MT.,IPU
NPK : 1016047901

Tembusan disampaikan :
1. Yth. Bapak Rektor UIR di Pekanbaru.
2. Yth. Sdr. Ketua Program Studi Teknik Informatika FT-UIR
3. Arsip

**Surat ini ditandatangani secara elektronik*

YAYASAN LEMBAGA PENDIDIKAN ISLAM (YLPI) RIAU
UNIVERSITAS ISLAM RIAU
FAKULTAS TEKNIK
PROGRAM STUDI TEKNIK INFORMATIKA
Jalan Kaharuddin Nasution No. 113 P. Marpoyan Pekanbaru Riau Indonesia – Kode Pos: 28284
Telp. +62 761 674674 Fax. +62 761 674834 *Website*: www.uir.ac.id *Email*: info@uir.ac.id

F.A.3.10

KARTU BIMBINGAN TUGAS AKHIR
SEMESTER GENAP TA 2023/2024

NPM : 183510468
Nama Mahasiswa : M Dicky Alfansyah
Dosen Pembimbing : Dr. Arbi Haza Nasution, B.IT (Hons)., M.IT
Program Studi : Teknik Informatika
Judul Tugas Akhir : Perbandingan Kinerja Algoritma Klasifikasi Dan Implementasi Model Terbaik Pada Website Deteksi Link Phishing

Lembar Ke : 1/2

NO.	Hari/Tanggal Bimbingan	Materi Bimbingan	Hasil/Saran Bimbingan	Paraf Dosen Pembimbing
1	Selasa, 28 Maret 2023	Bab 1	Lanjut bab 2.	
2	Rabu, 12 April 2023	Bab 2	Perbaiki hasil kesimpulan tinjauan pustaka dan tulisan kerangka pemikiran ke bahasa indonesia.	
3	Jum'at 26 Mei 2023	Bab 3	Perbaiki karakteristik data, dan parameter fitur yang digunakan.	
4	Selasa 13 Juni 2023	Bab 3	Menambahkan flowchart normalisasi data, perbaiki bentuk sampel kode menjadi pseudocode dan perbaiki penulisan sumber pengambilan data.	

YAYASAN LEMBAGA PENDIDIKAN ISLAM (YLPI) RIAU
UNIVERSITAS ISLAM RIAU
FAKULTAS TEKNIK
PROGRAM STUDI TEKNIK INFORMATIKA
Jalan Kaharuddin Nasution No. 113 P. Marpoyan Pekanbaru Riau Indonesia – Kode Pos: 28284
Telp. +62 761 674674 Fax. +62 761 674834 *Website*: www.uir.ac.id *Email*: info@uir.ac.id

F.A.3.10

5	Selasa 20 Juni 2023	Jurnal	Perbaikan jurnal.	
6	Kamis, 21 September 2023	Bab 1, 2, 3	Perbaikan margin bab 2,dan subbab penjelasan algoritma pada bab 3 pindah ke bab 2.	
7	Selasa, 03 Oktober 2023	Bab 2, dan 3	Acc.	
8	Rabu, 13 Maret 2024	Bab 4, dan 5	Acc. Lanjut Kompre	

Pekanbaru, 2 April 2024

Wakil Dekan I

MTYZNTEWMZA5

Dr. Mursyidah, M.Sc.
NIDN 1013056902

Catatan :

1. Lama bimbingan Tugas Akhir/ Skripsi maksimal 2 semester sejak TMT SK Pembimbingditerbitkan
2. Kartu ini harus dibawa setiap kali berkonsultasi dengan pembimbing dan HARUS dicetak kembali setiap memasuki semester baru melalui SIKAD
3. Saran dan koreksi dari pembimbing harus ditulis dan diparaf oleh pembimbing
4. Setelah skripsi disetujui (ACC) oleh pembimbing, kartu ini harus ditandatangani oleh Wakil Dekan I/ Kepala departemen/Ketua prodi 5. Kartu kendali bimbingan asli yang telah ditandatangani diserahkan kepada Ketua Program Studi dan kopiannya dilampirkan pada skripsi.
6. Jika jumlah pertemuan pada kartu bimbingan tidak cukup dalam satu halaman, kartu bimbingan ini dapat di download kembali melalui SIKAD

SURAT KEPUTUSAN DEKAN FAKULTAS TEKNIK UNIVERSITAS ISLAM RIAU
NOMOR : 0409/KPTS/FT-UIR/2024
TENTANG PENETAPAN DOSEN PENGUJI SKRIPSI MAHASISWA FAK. TEKNIK UNIV. ISLAM RIAU

DEKAN FAKULTAS TEKNIK

Menimbang : 1. Bahwa untuk menyelesaikan studi S.1 bagi mahasiswa Fakultas Teknik Univ. Islam Riau dilaksanakan Ujian Skripsi/Komprehensif sebagai tugas akhir. Untuk itu perlu ditetapkan mahasiswa yang telah memenuhi syarat untuk ujian dimaksud serta dosen penguji.

2. Bahwa penetapan mahasiswa yang memenuhi syarat dan dosen penguji yang bersangkutan perlu ditetapkan dengan Surat Keputusan Dekan.

Mengingat : 1. Undang - Undang Nomor 12 Tahun 2012 Tentang Pendidikan Tinggi
2. Peraturan Presiden Republik Indonesia Nomor 8 Tahun 2012 Tentang Kerangka Kualifikasi Nasional Indonesia
3. Peraturan Pemerintah Republik Indonesia Nomor 37 Tahun 2009 Tentang Dosen
4. Peraturan Pemerintah Republik Indonesia Nomor 66 Tahun 2010 Tentang Pengelolaan dan Penyelenggaraan Pendidikan
5. Peraturan Menteri Pendidikan Nasional Nomor 63 Tahun 2009 Tentang Sistem Penjaminan Mutu Pendidikan
6. Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 49 Tahun 2014 Tentang Standar Nasional Pendidikan Tinggi
7. Statuta Universitas Islam Riau Tahun 2018
8. Peraturan Universitas Islam Riau Nomor 001 Tahun 2018 Tentang Ketentuan Akademik Bidang Pendidikan Universitas Islam Riau

MEMUTUSKAN

Menetapkan : 1. Mahasiswa Fakultas Teknik Universitas Islam Riau yang tersebut namanya dibawah ini :

Nama	: M Dicky Alfansyah
NPM	: 183510468
Program Studi	: Teknik Informatika
Jenjang Pendidikan	: Strata Satu (S1)
Judul Skripsi	: Analisis dan Implementasi Machine Learning Pada Website Deteksi Link Phising Berbasis Web

2. Penguji Skripsi/Komprehensif mahasiswa tersebut terdiri dari :

1. Dr. Arbi Haza Nasution, B.IT., M.IT..	Sebagai Ketua Merangkap Penguji
2. Ause Labellapansa, S.T., M.Cs., M.Kom.	Sebagai Anggota Merangkap Penguji
3. Mutia Fadhillah, S.ST., M.Sc.	Sebagai Anggota Merangkap Penguji

3. Laporan hasil ujian serta berita acara telah sampai kepada Pimpinan Fakultas selambat-lambatnya 1(satu) bulan setelah ujian dilaksanakan.

4. Keputusan ini mulai berlaku pada tanggal ditetapkannya dengan ketentuan bila terdapat kekeliruan dikemudian hari segera ditinjau kembali.

KUTIPAN : Disampaikan kepada yang bersangkutan untuk dapat dilaksanakan dengan sebaik-baiknya.

Ditetapkan di : Pekanbaru
Pada Tanggal : 24 Ramadhan 1445 H
03 April 2024 M

Dekan,

Dr. Deddy Purnomo Retno, S.T., M.T.
NPK : 1005057702

Tembusan disampaikan :
1. Yth. Rektor UIR di Pekanbaru.
2. Yth. Ketua Program Studi Teknik Informatika FT-UIR
3. Yth. Pembimbing dan Penguji Skripsi
3. Mahasiswa yang bersangkutan
5. Arsip

**Surat ini ditandatangani secara elektronik*

YAYASAN LEMBAGA PENDIDIKAN ISLAM (YLPI) RIAU
UNIVERSITAS ISLAM RIAU
FAKULTAS TEKNIK
PROGRAM STUDI TEKNIK INFORMATIKA
Jalan Kaharuddin Nasution No. 113 P. Marpoyan Pekanbaru Riau Indonesia – Kode Pos: 28284
Telp. +62 761 674674 *Website*: www.eng.uir.ac.id *Email*: fakultas_teknik@uir.ac.id

BERITA ACARA UJIAN SKRIPSI

Berdasarkan Surat Keputusan Dekan Fakultas Teknik Universitas Islam Riau, Pekanbaru, tanggal 01 April 2024, Nomor: 0409 /KPTS/FT-UIR/2024, maka pada hari Rabu, tanggal 27 Maret 2024, telah dilaksanakan Ujian Skripsi Program Studi Teknik Informatika Fakultas Teknik Universitas Islam Riau, Jenjang Studi S1, Tahun Akademik 2023/2024 berikut ini.

1. Nama : M Dicky Alfansyah
2. NPM : 183510468
3. Judul Skripsi : Analisis dan Implementasi Machine Learning Pada Website Deteksi Link Phising Berbasis Web
4. Waktu Ujian : 14.00 WIB s.d. Selesai
5. Tempat Pelaksanaan Ujian : Ruang Sidang Fakultas Teknik UIR

Dengan keputusan Hasil Ujian Skripsi:

Lulus*/ Lulus dengan Perbaikan*/ Tidak Lulus*

** Coret yang tidak perlu.*

Nilai Ujian:

Nilai Ujian Angka =.....82..... Nilai Huruf = ...A.....

Tim Penguji Skripsi.

No	Nama	Jabatan	Tanda Tangan
1	Dr. Arbi Haza Nasution, B.IT., M.IT..	Ketua	1.
2	Ause Labellapansa, S.T., M.Cs., M.Kom.	Anggota	2.
3	Mutia Fadhilla , S.ST., M.Sc.	Anggota	3.

Panitia Ujian
Ketua,

Dr. Arbi Haza Nasution, B.IT., M.IT..
NIDN. 1023048901

Pekanbaru, 27 Maret 2024
Mengetahui,
Dekan Fakultas Teknik

Dr. Deddy Purnomo Retno, S.T.,M.T.,GP.A-Utama.
NIDN. 090602372

UNIVERSITAS ISLAM RIAU
FAKULTAS TEKNIK
الجامعة الإسلامية الريوية
Alamat: Jalan Kaharuddin Nasution No.113, Marpoyan, Pekanbaru, Riau, Indonesia - 28284
Telp. +62 761 674674 Email: fakultas_teknik@uir.ac.id Website: www.eng.uir.ac.id

SURAT KETERANGAN BEBAS PLAGIAT

Nomor: 104/A-UIR/5-T/2024

Fakultas Teknik Universitas Islam Riau menerangkan bahwa Mahasiswa/i dengan identitas berikut:

Nama	:	**M DICKY ALFANSYAH**
NPM	:	183510468
Program Studi	:	Teknik Informatika
Jenjang Pendidikan	:	Strata Satu (S1)
Judul Skripsi TA	:	PERBANDINGAN KINERJA ALGORITMA KLASIFIKASI DAN IMPLEMENTASI MODEL TERBAIK PADA WEBSITE DETEKSI LINK PHISHING

Dinyatakan **Bebas Plagiat,** berdasarkan hasil pengecekan pada Turnitin menunjukkan angka **Similarity Index < 30%** sesuai dengan peraturan Universitas Islam Riau yang berlaku.

Demikian surat keterangan ini dibuat untuk dapat dipergunakan sebagaimana mestinya.

Mengetahui,

Kaprodi. Teknik Informatika

Apri Siswanto, S.Kom., M.Kom., Ph.D

Pekanbaru, 19 March 2024 M
9 Romadhōn 1445 H

Staff Pemeriksa

Khezi Triandini Dafan, S.E