Journal of Applied Informatics and Computing (JAIC)
Vol.9, No.3, June 2025, pp. 651~657
e-ISSN: 2548-6861

OPTIMASI HYPERPARAMETER PADA MODEL REGRESI LOGISTIK UNTUK MENINGKATKAN AKURASI DETEKSI PHISHING BERBASIS KONTEN DAN METADATA

Aditya Putra Bahri[1], Antoni Haikal[2]
*Teknik Informatika, Politeknik Negeri Batam
Aditya.Putra.Bahri@student.polibatam.ac.id[1] Antoni@polibatam.ac.id[2]

Article Info

Article history:

Received 2025-03-14
Revised 2025-03-22
Accepted 2025-03-26

Keyword:

Phishing Detection,
Logistic Regression,
TF-IDF,
Hyperparameter Tuning,
Cybersecurity.

ABSTRACT

This study evaluates and optimizes the performance of the Logistic Regression algorithm for phishing email detection. The primary challenge lies in balancing the use of technical features (metadata) and textual features (content) to prevent overfitting. This research utilizes a large-scale combined dataset consisting of 102,486 emails, comprising the Phishing dataset (Naser Abdullah Alam) and the Valid dataset (Enron), processed using TF-IDF vectorization and metadata feature extraction techniques. Unlike previous studies, this research implements hyperparameter optimization (C regularization) to assess model stability. Experimental results demonstrate that the Content-Only model yields the most superior and stable performance, achieving an Area Under Curve (AUC) of 0.99 and an F1-Score exceeding 95.61%. In contrast, the incorporation of metadata features in the Hybrid model led to a decline in accuracy at high regularization values, indicating that metadata acts as noise. The study concludes that Logistic Regression utilizing content features alone is sufficiently robust and efficient for phishing detection, eliminating the need for the added complexity of metadata.

I. PENDAHULUAN

Perkembangan Kecerdasan Buatan (Artificial Intelligence/AI) pada dekade terakhir telah mendorong transformasi besar di berbagai bidang kehidupan, termasuk pada bidang keamanan informasi. Salah satu cabang dari AI yang memiliki pengaruh signifikan adalah Machine Learning (pembelajaran mesin), yaitu kemampuan komputer untuk mempelajari pola dari data historis dan kemudian membuat prediksi atau keputusan secara otomatis tanpa harus diprogram secara eksplisit. Melalui pendekatan ini, sistem dapat beradaptasi dan meningkatkan kinerjanya seiring bertambahnya data. Salah satu tugas utama dalam pembelajaran mesin adalah klasifikasi, yang bertujuan untuk mengelompokkan data ke dalam kategori tertentu berdasarkan karakteristik yang dimilikinya.

Di antara berbagai algoritma klasifikasi, Regresi Logistik (Logistic Regression) merupakan salah satu metode paling mendasar dan banyak digunakan karena kesederhanaannya, interpretabilitasnya, dan efektivitasnya dalam menangani permasalahan klasifikasi biner—situasi di mana keluaran hanya terdiri dari dua kemungkinan, misalnya “Ya” atau “Tidak”, “Benar” atau “Salah”, atau dalam konteks ini “Phishing” dan “Bukan Phishing”. Meskipun tergolong sederhana dibandingkan algoritma modern lainnya seperti Random Forest atau Neural Network, Regresi Logistik sering dijadikan baseline model karena kemampuannya dalam menghasilkan hasil yang stabil dan mudah dianalisis [1].

Salah satu penerapan yang sangat relevan bagi algoritma ini terdapat pada bidang keamanan siber (cyber security), khususnya dalam deteksi serangan phishing. Phishing merupakan salah satu bentuk serangan sosial (social engineering attack) yang paling sering terjadi di dunia digital. Serangan ini dilakukan dengan cara meniru komunikasi resmi, biasanya melalui email, untuk memancing korban agar memberikan informasi sensitif seperti kata sandi, nomor kartu kredit, atau data keuangan. Email phishing dirancang sedemikian rupa agar tampak meyakinkan, sehingga sulit dibedakan secara kasat mata dengan email sah. Karena itu, deteksi phishing menjadi tantangan penting yang membutuhkan

pendekatan otomatis berbasis pembelajaran mesin untuk membantu proses identifikasi.

Penelitian ini berfokus pada penerapan algoritma Regresi Logistik dalam membangun sistem klasifikasi yang mampu membedakan antara email phishing dan email sah (non-phishing) secara otomatis. Melalui proses pelatihan model menggunakan dataset email modern yang berisi contoh phishing dan non-phishing, penelitian ini bertujuan untuk mengevaluasi tingkat akurasi, presisi, recall, serta efektivitas model dalam melakukan deteksi. Dengan demikian, penelitian ini tidak hanya memberikan pemahaman mengenai penerapan algoritma Regresi Logistik pada domain keamanan siber, tetapi juga berkontribusi dalam pengembangan metode deteksi phishing yang lebih efisien, terukur, dan dapat digunakan sebagai dasar pengembangan sistem keamanan email cerdas di masa mendatang.

Dalam penelitian ini, pemilihan algoritma *Logistic Regression* (LR) didasarkan pada tiga justifikasi utama yang relevan dengan karakteristik data teks.

Pertama, efisiensi komputasi pada data *sparse*. Proses vektorisasi teks menggunakan TF-IDF menghasilkan matriks berdimensi tinggi yang sangat jarang (*sparse matrix*). Literatur menunjukkan bahwa LR memiliki efisiensi tinggi dalam memproses tipe data ini dibandingkan algoritma non-linear yang kompleks [2]

Kedua, aspek interpretabilitas (*Explainable AI*). Dalam keamanan siber, kemampuan menjelaskan alasan deteksi sangat krusial. LR memungkinkan analisis bobot fitur (*coefficients*) untuk mengidentifikasi kata kunci spesifik yang menjadi indikator serangan, berbeda dengan model *Deep Learning* yang bersifat *black-box* [3]. LR memiliki efisiensi tinggi dalam memproses tipe data ini dibandingkan algoritma non-linear yang kompleks..

Ketiga, kemampuan kontrol *overfitting*. Melalui mekanisme regularisasi (parameter *C*), LR memungkinkan penyetelan untuk menyeimbangkan bias dan varians, sehingga model tetap stabil meski dilatih pada dataset yang besar [4].

II. LANDASAN TEORI

A. Phishing

Phishing merupakan bentuk kejahatan siber yang bertujuan untuk menipu korban agar memberikan informasi sensitif seperti kata sandi, data finansial, maupun identitas pribadi. Serangan ini umumnya dilakukan melalui email, pesan teks, atau situs web palsu yang meniru tampilan layanan resmi. Dampak yang ditimbulkan mencakup kerugian finansial, pencurian data, serta hilangnya kepercayaan terhadap sistem digital. Jenis-jenis phishing yang sering ditemui antara lain spear phishing yang menargetkan individu secara spesifik dengan pesan yang dipersonalisasi, whaling yang menyasar eksekutif tingkat tinggi, clone phishing yang menduplikasi email sah untuk menjebak korban, serta smishing yang dilakukan melalui pesan teks. Pola serangan tersebut sering kali memanfaatkan urgensi palsu dan manipulasi psikologis untuk memancing pengguna agar mengklik tautan berbahaya. Masalah utama yang menjadi fokus penelitian ini adalah meningkatnya kompleksitas serangan phishing yang sulit dikenali secara manual, sehingga dibutuhkan pendekatan berbasis machine learning yang mampu mendeteksi email phishing secara otomatis dan akurat. Penelitian ini secara spesifik berupaya merancang dan mengevaluasi efektivitas model Regresi Logistik dalam membedakan email phishing dan email sah berdasarkan analisis fitur metadata dan konten.

B. Machine Learning untuk Keamanan Siber

Machine Learning (ML) memiliki peran penting dalam sistem keamanan siber modern karena kemampuannya mengenali pola dan mendeteksi anomali dari data dalam jumlah besar secara otomatis. Teknologi ini mampu mengidentifikasi serangan baru yang belum pernah muncul sebelumnya, melampaui pendekatan konvensional berbasis tanda tangan. Algoritma yang umum digunakan dalam deteksi ancaman mencakup Regresi Logistik, Support Vector Machine (SVM) [5] [6] , Naive Bayes, Decision Tree, dan Random Forest. Untuk data tidak terstruktur seperti teks atau gambar, pendekatan Deep Learning berbasis Convolutional Neural Network (CNN) dan Transformer seperti BERT atau GPT menjadi pilihan utama karena kemampuannya memproses representasi kompleks. Dalam konteks deteksi phishing, alur kerja machine learning umumnya meliputi pengumpulan dan pembersihan data email, ekstraksi fitur menggunakan metode seperti TF-IDF, pembagian dataset menjadi 80% data latih dan 20% data uji, pelatihan model dengan algoritma klasifikasi, serta evaluasi performa model menggunakan metrik seperti akurasi, presisi, recall, dan F1-score.

C. Regresi Logistik (Logistic Regression)

Regresi Logistik merupakan algoritma klasifikasi biner yang memodelkan probabilitas suatu data termasuk ke dalam kelas tertentu. Model ini bekerja dengan menggabungkan fitur masukan secara linear dan mengubahnya menjadi nilai probabilitas melalui fungsi sigmoid yang bernilai antara 0 dan 1. Keunggulan utama Regresi Logistik adalah kemampuannya untuk memberikan interpretasi yang jelas terhadap pengaruh masing-masing fitur terhadap hasil prediksi. Model ini juga efisien secara komputasi, mudah diimplementasikan, dan sering dijadikan baseline dalam penelitian machine learning karena sifatnya yang sederhana namun stabil. Walaupun tidak selalu menghasilkan akurasi tertinggi dibanding model yang lebih kompleks, Regresi Logistik tetap menjadi alat fundamental yang efektif untuk dataset berstruktur dan sangat berguna dalam menjelaskan hubungan antar fitur pada deteksi phishing.

D. Fitur Email untuk Deteksi Phishing

Untuk mendeteksi phishing secara akurat, model machine learning dilatih menggunakan berbagai sinyal yang diekstraksi dari metadata dan konten email. Fitur metadata mencakup elemen-elemen teknis seperti alamat IP pengirim [7], hasil autentikasi SPF, DKIM, dan DMARC, kesesuaian domain antara pengirim dan penerima, serta analisis subjek email yang sering kali memuat kata-kata manipulatif seperti “Urgent”, “Verify”, atau “Account”.

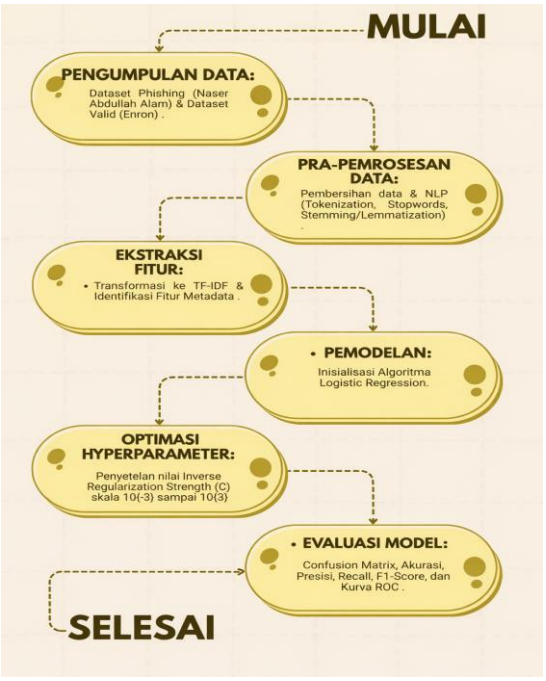
Fitur-fitur ini membantu model mengenali karakteristik teknis yang sulit dipalsukan oleh penyerang. Sementara itu, fitur konten berfokus pada analisis teks dan struktur isi email [8] , termasuk frekuensi kata mencurigakan seperti "password" dan "login", penggunaan salam generik tanpa personalisasi, serta analisis tautan (URL) untuk mendeteksi keberadaan alamat IP, layanan pemendek tautan, atau ketidaksesuaian antara teks tautan dan alamat sebenarnya. Dari sisi struktur HTML, elemen seperti formulir input data (<form>) dan skrip JavaScript mencurigakan juga dianalisis karena sering digunakan dalam email berbahaya. Seluruh fitur ini kemudian diubah menjadi bentuk numerik melalui teknik vektorisasi seperti TF-IDF agar dapat diproses oleh model Regresi Logistik.

E. Evaluasi Model Klasifikasi

Evaluasi model klasifikasi dilakukan untuk mengukur sejauh mana model mampu mendeteksi phishing secara efektif pada data uji yang belum pernah dilihat sebelumnya. Pengukuran ini dilakukan menggunakan Confusion Matrix yang merangkum empat kategori hasil prediksi, yaitu True Positive (TP) untuk email phishing yang berhasil dideteksi dengan benar, True Negative (TN) untuk email sah yang diklasifikasikan benar, False Positive (FP) untuk email sah yang salah dikategorikan sebagai phishing, dan False Negative (FN) untuk email phishing yang gagal terdeteksi. Berdasarkan nilai-nilai tersebut, sejumlah metrik evaluasi digunakan, yaitu Akurasi, Presisi, Recall, dan F1-Score. Akurasi mengukur proporsi prediksi yang benar secara keseluruhan, Presisi menilai sejauh mana model benar dalam menandai email phishing, Recall mengukur kemampuan model menangkap seluruh email phishing yang ada, dan F1-Score menggambarkan keseimbangan antara Presisi dan Recall. Dalam konteks keamanan siber, khususnya deteksi phishing, prioritas utama adalah meminimalkan False Negative karena kesalahan jenis ini memungkinkan email berbahaya lolos dan mencapai pengguna akhir.

III. METODE

Penelitian ini bertujuan untuk melakukan evaluasi mendalam terhadap model Regresi Logistik dalam kemampuannya mendeteksi email phishing. Alur penelitian ini disusun secara sistematis, mencakup tahap pengumpulan data dari sumber publik, pra-pemrosesan data, ekstraksi fitur metadata dan konten email, serta pemodelan dan evaluasi kinerja secara kuantitatif.



Gambar 1 Diagram Alir Metodologi Penelitian dan Optimasi Hyperparameter

A. Pengumpulan Data (Data Collection)

Untuk memastikan validitas model dan mencegah bias akibat dataset yang kecil, penelitian ini menggunakan dataset gabungan dari dua sumber publik terverifikasi:

- Dataset Phishing: Diambil dari repositori *Phishing Email Dataset* oleh Naser Abdullah Alam [9] , yang berisi koleksi email serangan siber terkini.
- Dataset Valid (Ham): Diambil dari *Enron Email Dataset* [10] , yang merupakan standar industri untuk korpus email korporat yang sah.

Data dimuat secara iteratif menggunakan pustaka Pandas untuk menangani volume data yang besar, kemudian digabungkan dan diberi label biner (1 untuk Phishing, 0 untuk Valid/Ham).

B. Pra-pemrosesan Data (Data Preprocessing)

Data mentah yang telah dikumpulkan melalui tahap pembersihan dan persiapan untuk meningkatkan kualitas dan menghilangkan 'spam' (gangguan). Aktivitas utama pada tahap ini meliputi:

1. Pembersihan Data: Menghapus data atau email duplikat, memperbaiki kesalahan format, dan menghilangkan informasi yang tidak relevan.
2. Pra-pemrosesan Teks:
   Pada tahap ini, konten email diolah secara mendalam menggunakan teknik *Natural Language Processing* (NLP) dengan parameter spesifik agar data siap diproses oleh algoritma Regresi Logistik. Tahapan ini dilakukan menggunakan pustaka **NLTK**

(*Natural Language Toolkit*) pada **Python 3.12** dengan rincian sebagai berikut:

- **Tokenisasi (*Tokenization*):** Memecah teks menjadi unit kata (*token*) dengan menghapus tanda baca dan karakter khusus. Parameter yang digunakan adalah pola *regex* `r'\b\w\w+\b'` untuk memastikan hanya kata dengan panjang minimal dua karakter yang diambil.
- **Pembersihan Kata Henti (*Stopwords Removal*):** Menghapus kata-kata umum yang tidak memiliki bobot informasi penting (seperti: *"the"*, *"and"*, *"is"*) menggunakan daftar *stopwords* standar dari library NLTK.
- **Stemming:** Mengubah kata berimbuhan menjadi kata dasar menggunakan PorterStemmer (untuk dataset bahasa Inggris) atau Sastrawi (untuk bahasa Indonesia). Proses ini memotong akhiran kata (seperti: *"verified"* menjadi *"verify"*) agar frekuensi kata menjadi lebih konsisten.
- **Lemmatization:** Berbeda dengan *stemming*, tahap ini menggunakan WordNetLemmatizer untuk mengembalikan kata ke bentuk kamusnya berdasarkan konteks linguistik, sehingga menjaga makna kata tetap akurat sebelum dilakukan pembobotan TF-IDF.

C. Ekstraksi Fitur (Feature Extraction)

Sesuai dengan judul penelitian, data yang sudah bersih diubah menjadi format numerik (vektor fitur) dengan fokus pada dua kategori fitur:

1. Identifikasi Fitur Metadata:
   Mengekstrak informasi dari *header* email (pengirim, subjek, status autentikasi seperti SPF, DKIM, DMARC). Secara teknis, **SPF (Sender Policy Framework)** bekerja dengan memvalidasi alamat IP pengirim terhadap daftar IP yang diizinkan oleh pemilik domain. **DKIM (DomainKeys Identified Mail)** menambahkan tanda tangan digital terenkripsi pada *header* email untuk memastikan isi pesan tidak dimanipulasi. Sementara itu, **DMARC** bertindak sebagai pengatur kebijakan yang menggunakan hasil dari SPF dan DKIM untuk menentukan apakah sebuah email harus diterima, dikarantina, atau ditolak jika autentikasi gagal.
2. Identifikasi Fitur Konten: Mengekstrak informasi langsung dari badan (*body*) email. Bagian ini dibagi menjadi beberapa aspek penting:
   - **Fitur Berbasis Teks (Textual Features):** Analisis ini berfokus pada penggunaan bahasa dan kata-kata di dalam email. Teknik TF-IDF (Term Frequency-Inverse Document Frequency) digunakan untuk memberikan bobot numerik pada setiap kata. Kata-kata yang merupakan ciri khas *phishing* (seperti *"verify"*, *"urgent"*, *"suspended"*, atau *"password"*) akan mendapatkan skor tinggi karena kemunculannya yang spesifik pada email berbahaya, sehingga membantu model Regresi Logistik mengenali pola penipuan secara akurat.
   - **Fitur Berbasis Tautan (URL Features):** Fitur ini menganalisis keamanan tautan yang disematkan. Hal yang paling krusial adalah membandingkan antara Anchor Text (teks yang terlihat, misalnya: *"www.bankanda.com"*) dengan alamat asli (*href*) yang dituju (misalnya: *"bit.ly/curi-data"*). Jika terdapat ketidakcocokan (*mismatch*), ini merupakan indikator utama serangan *phishing*.
   - **Fitur Struktural**: Menganalisis elemen teknis seperti keberadaan formulir HTML (`<form>`) yang meminta *input* data sensitif secara langsung atau penggunaan JavaScript yang mencurigakan untuk pengalihan otomatis (*redirect*).
3. Representasi Teks:
   Menggunakan teknik seperti TF-IDF (*Term Frequency-Inverse Document Frequency*) untuk merepresentasikan konten teks email secara numerik.

D. Pemodelan

Model dibangun menggunakan algoritma Logistic Regression dengan solver liblinear yang dioptimalkan untuk dataset berukuran besar. Berbeda dengan pendekatan standar, penelitian ini menerapkan skema Hyperparameter Tuning menggunakan metode Grid Search pada parameter tunggal. Proses ini dilakukan dengan melatih ulang model menggunakan variasi nilai parameter Inverse Regularization Strength C pada skala logaritmik, mulai dari $10^{\{-3\}}$ $hingga$ 10^3
Parameter C ini merupakan teknik regularisasi, bukan normalisasi, yang berfungsi mengontrol keseimbangan antara kompleksitas model dan kemampuannya untuk melakukan generalisasi. Nilai C yang lebih kecil $10^{\{-3\}}$ memberikan regularisasi yang kuat (penalti besar) untuk mencegah overfitting, sedangkan nilai C yang besar 10^3 memberikan regularisasi yang lemah agar model dapat mempelajari data latih lebih detail . Optimasi ini krusial untuk menemukan konfigurasi parameter yang menghasilkan akurasi terbaik sekaligus memastikan model tetap tangguh (robust) saat menghadapi data email baru di luar dataset.

- Rumus Regularisasi L_2

$$J(w) = \Sigma_{i=1}^{n} \log\left(1 + e^{-yi(w^t xi)}\right) + \frac{1}{2}||w||^2$$

Penjelasan : C adalah *Inverse Regularization Strength* dan $\frac{1}{2}||w||^2$ adalah penalti L_2 untuk mencegah *overfitting*.

IV. PEMBAHASAN

A. Karakteristik Dataset

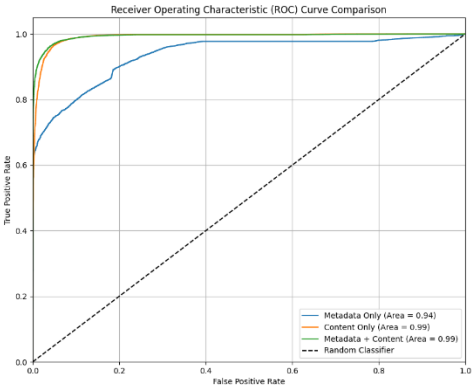
```
===== STATISTIK DATASET =====
Jumlah email phishing : 52719
Jumlah email valid    : 49767
Total email           : 102486
=============================
```

Gambar 2 Rincian total data

Sebelum melakukan pengujian model, dilakukan analisis terhadap karakteristik dataset yang digunakan untuk memastikan model memiliki kemampuan generalisasi yang kuat pada data skala besar. Dataset yang digunakan merupakan gabungan dari dua sumber publik dengan rincian sebagai berikut:

- **Total Data**: Penelitian ini menggunakan total **102.486 email**.
- **Rincian Kelas**: Terdiri dari **52.719 email phishing** (label 1) dan **49.767 email valid atau ham** (label 0).
- **Sumber Data**: Data dikumpulkan secara iteratif dari repositori *Naser Abdullah Alam* untuk kategori phishing dan *Enron Email Dataset* untuk kategori email sah .
- **Tujuan**: Penggunaan dataset yang seimbang (*balanced*) ini bertujuan untuk menghindari bias pada model sehingga metrik evaluasi seperti Akurasi dapat memberikan hasil yang valid dan tepercaya.

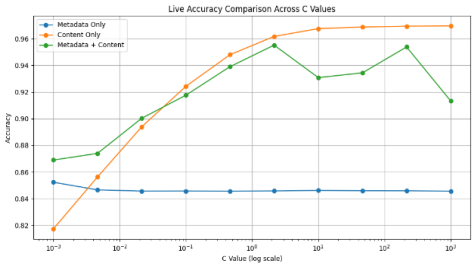
B. Analisis Kurva ROC



Gambar 3 Perbandingan Kurva ROC (Receiver Operating Characteristic)

Berdasarkan Gambar 3, terlihat perbedaan kinerja yang signifikan antara model. Model berbasis Konten (*Content-Only*) dan model Kombinasi (*Hybrid*) menunjukkan kinerja hampir sempurna dengan nilai AUC sebesar **0.99**. Garis kurva yang mendekati sudut kiri atas menandakan tingkat *True Positive Rate* yang tinggi dengan *False Positive Rate* yang sangat rendah. Sebaliknya, model *Metadata-Only* tertinggal dengan AUC **0.94**, membuktikan bahwa fitur teknis saja kurang diskriminatif dibandingkan fitur teks.

C. Analisis Stabilitas & Noise



Gambar 4 Perbandingan Akurasi Terhadap Nilai Parameter C

Hasil uji stabilitas pada Gambar 4 memperlihatkan temuan menarik terkait peran fitur metadata. Model *Content-Only* (Garis Oranye) menunjukkan tren performa yang stabil dan terus meningkat seiring bertambahnya nilai *C*, mencapai akurasi puncak di atas **95.5%**..
Namun, model *Hybrid* (Garis Hijau) justru mengalami penurunan akurasi yang tajam ketika nilai *C* melebihi angka 10. Penurunan ini mengindikasikan bahwa pada kompleksitas model yang tinggi, fitur metadata bertindak sebagai **'noise'** atau gangguan yang mengaburkan pola fitur utama. Hal ini menjawab mengapa penggabungan fitur tidak selalu menghasilkan performa terbaik; dalam kasus ini, model berbasis

konten murni terbukti lebih *robust* (tangguh) dibandingkan model hybrid.

D. Tabel Performa Lengkap Head-to-Head.

- Metrik Evaluasi Performa
  Untuk mengukur efektivitas model Regresi Logistik dalam mendeteksi email *phishing*, penelitian ini menggunakan metrik evaluasi standar yang dihitung berdasarkan nilai *Confusion Matrix*. Metrik yang digunakan meliputi Presisi (*Precision*), *Recall*, dan *F1-Score* dengan rumus sebagai berikut:
  1. ***Precision***: Mengukur tingkat keakuratan model dalam menandai email sebagai *phishing*.

$$PRESISION = \frac{TP}{TP + FP}$$

  2. **Recall**: Mengukur kemampuan model dalam menangkap seluruh email *phishing* yang ada dalam dataset.

$$RECALL = \frac{TP}{TP + FN}$$

  3. **F1-Score**: Nilai rata-rata harmonis yang menyeimbangkan antara Presisi dan *Recall*.

$$F1 - SCORE = 2 \times \frac{PRESISION \times RECALL}{PRESISION + RECALL}$$

Metrik Evaluasi	Skenario A (Metadata)	Skenario B (Content)	Skenario C (Hybrid)
Akurasi	84.56%	95.51%	95.63%
Presisi	91.50%	96.22%	97.58%
Recall	77.15%	95.00%	93.83%
F1-Score	83.71%	95.61%	95.67%
ROC AUC	0.94	0.99	0.99

Tabel 1 Tabel Performa Lengkap Head-to-Head.

Berdasarkan Tabel 1, terlihat bahwa penggabungan fitur pada Skenario C (Hybrid) menghasilkan akurasi tertinggi sebesar 95,63%. Namun, jika ditinjau dari aspek keamanan siber yang mengutamakan deteksi ancaman (Recall), Skenario B (Content Only) menunjukkan performa yang lebih efektif dengan tingkat Recall 95,00%. Hal ini mengindikasikan bahwa fitur konten teks memiliki bobot pengaruh yang lebih dominan dibandingkan fitur metadata dalam mendeteksi pola email phishing pada dataset ini.

E. Skenario 1: Metadata Only

```
Training Content Only model...

============================
MODEL CONTENT ONLY
============================
Accuracy : 0.9551175724460923
Precision: 0.9622478386167147
Recall   : 0.9500189681335357
F1 Score : 0.9560943018039515
ROC AUC  : 0.9911397483431372
Confusion Matrix:
[[ 9561   393]
 [  527 10017]]
```

Gambar 5 Hasil Evaluasi Model Metadata Only

Berdasarkan Gambar 5, model yang hanya menggunakan fitur metadata teknis menghasilkan akurasi sebesar **84,56%** dengan *Recall* sebesar **77,15%**. Hal ini menunjukkan bahwa informasi *header* saja belum cukup kuat untuk mendeteksi seluruh ancaman *phishing*.

F. Skenario 2: Content Only

```
Training Content Only model...

============================
MODEL CONTENT ONLY
============================
Accuracy : 0.9551175724460923
Precision: 0.9622478386167147
Recall   : 0.9500189681335357
F1 Score : 0.9560943018039515
ROC AUC  : 0.9911397483431372
Confusion Matrix:
[[ 9561   393]
 [  527 10017]]
```

Gambar 6 Hasil Evaluasi Model Content Only

Pada Gambar 6, terlihat peningkatan performa yang signifikan dengan akurasi **95,51%** dan *Recall* tertinggi sebesar **95,00%**. Ini membuktikan bahwa fitur teks (konten) adalah

prediktor yang paling dominan dalam deteksi *phishing*.

G. Skenario 3: Metadata + Content (Hybrid)

```
============================
MODEL METADATA + CONTENT
============================
Accuracy : 0.9562884183822812
Precision: 0.9758334977313079
Recall   : 0.9382587253414264
F1 Score : 0.9566773039357895
ROC AUC  : 0.9934115462996247
Confusion Matrix:
[[9709  245]
 [ 651 9893]]
```

Gambar 7 Hasil Evaluasi Model Hybrid

Hasil pada Gambar 7, menunjukkan akurasi tertinggi sebesar **95,63%**. Meskipun akurasinya paling tinggi, model ini memiliki *Recall* (**93,83%**) yang sedikit lebih rendah dibandingkan Skenario 2, yang mengindikasikan adanya pengaruh *noise* dari fitur metadata pada beberapa kasus tertentu.

H. Pembahasan dan Perbandingan Komparatif

Meskipun hasil pengujian disajikan secara terpisah pada gambar di atas, analisis perbandingan menunjukkan bahwa penggabungan fitur (Hybrid) mampu meningkatkan **Presisi (97,58%)** dibandingkan skenario lainnya. Hal ini berarti model Hybrid sangat akurat dalam memastikan bahwa email yang ditandai sebagai *phishing* benar-benar merupakan ancaman, sehingga meminimalkan kesalahan *False Positive*.

Namun, untuk perlindungan yang lebih luas, Skenario 2 (*Content Only*) tetap unggul dalam aspek *Recall*, yang berarti lebih efektif dalam menangkap variasi email *phishing* agar tidak lolos ke kotak masuk pengguna

V. KESIMPULAN DAN SARAN

A. Kesimpulan

Berdasarkan hasil implementasi dan evaluasi komparatif terhadap model Regresi Logistik, diperoleh beberapa kesimpulan penting sebagai berikut:

1. Model Regresi Logistik telah berhasil diimplementasikan dan diuji pada tiga skenario berbeda, yaitu: Hanya Metadata, Hanya Konten, dan Hybrid (Metadata + Konten).
2. Skenario Hanya Konten (Skenario 2) terbukti memberikan performa paling optimal, dengan Akurasi 95.51% dan Recall 95.00%. Meskipun terdapat sedikit penurunan dibandingkan data latih, model ini berhasil menekan False Negative menjadi 527 kasus dari total 20.498 data uji, yang menunjukkan kemampuan generalisasi yang baik pada dataset skala besar.
3. Penambahan fitur metadata dalam model Hybrid (Skenario 3) tidak meningkatkan kinerja secara signifikan. Justru hasilnya menurun ke tingkat yang sama dengan model Hanya Metadata (Skenario 1). Hal ini menunjukkan bahwa pada dataset ini, fitur metadata berperan sebagai noise, yakni informasi tambahan yang justru mengaburkan pola utama yang sudah berhasil ditangkap oleh fitur konten.
4. Meskipun demikian, pengujian terhadap email sampel dengan indikasi phishing kuat menunjukkan bahwa model Hybrid mampu memberikan tingkat keyakinan prediksi tertinggi (99.97%). Temuan ini menandakan bahwa kombinasi metadata dan konten tetap berguna dalam memperkuat keyakinan model terhadap deteksi ancaman yang jelas.

B. Saran

Walaupun hasil penelitian ini menunjukkan performa yang sangat baik, terdapat beberapa peluang pengembangan yang dapat dilakukan pada penelitian selanjutnya:

1. Studi Komparatif dengan Algoritma Lain
   Disarankan untuk membandingkan Regresi Logistik dengan algoritma lain seperti Random Forest, XGBoost, atau SVM guna menilai apakah model lain dapat mengatasi efek "noise" dari metadata dan menghasilkan performa hybrid yang lebih optimal.
2. Rekayasa dan Seleksi Fitur
   Diperlukan penerapan teknik feature selection yang lebih mendalam untuk mengidentifikasi fitur metadata yang benar-benar berkontribusi positif serta mengeliminasi fitur yang tidak relevan atau menyebabkan gangguan dalam proses pembelajaran model.
3. Analisis Threshold Keputusan
   Eksperimen terhadap penyesuaian ambang batas klasifikasi (threshold) perlu dilakukan untuk meningkatkan Recall tanpa mengorbankan Presisi yang tinggi. Pendekatan ini berpotensi menurunkan jumlah False Negative pada kasus email phishing yang kompleks.
4. Implementasi pada Sistem Real-Time
   Model Hanya Konten (Skenario 2) sebagai model terbaik dapat dikembangkan lebih lanjut menjadi prototipe fungsional, misalnya dalam bentuk API deteksi phishing atau add-on email, sehingga dapat divalidasi kinerjanya terhadap data real-time dalam lingkungan produksi.

DAFTAR PUSTAKA

[1] A. Dutta, "Comparative Investigation of Traditional Machine-Learning Models and Transformer Models for Phishing Email Detection," *MDPI Electronics*, vol. 13, no. 24, p. 4877, 2024. [Online]. Available: https://www.mdpi.com/2079-9292/13/24/4877.

[2] S. Mishra and D. Soni, "Detection of Phishing URLs Using a Term Frequency Inverse Document Frequency (TF-IDF)," *International Journal for Multidisciplinary Research (IJFMR)*, vol. 6, no. 2, 2024. [Online]. Available: https://www.ijfmr.com/papers/2024/3/21435.pdf.

[3] R. S. Mohammed and R. Abdulhammed, "Comparative Analysis of Machine Learning Algorithms for Phishing Email Detection," *NTU Journal of Engineering and Technology*, vol. 4, no. 3, 2025. [Online]. Available: https://journals.ntu.edu.iq/index.php/NTU-JET/article/view/894.

[4] A. Khan, M. Ahmed, and A. Fathima, "Enhanced Phishing Detection Using Machine Learning Algorithms: A Comparative Study of Random Forest, SVM, and Logistic Regression Models," in *Proceedings of the International Conference on Innovative Computing & Communication (ICICC 2024)*, 2025. Available: https://ssrn.com/abstract=5191566.

[5] Z. Alkhalil, C. Hewage, L. Nawaf, and I. Khan, "In-Depth Analysis of Phishing Email Detection: Evaluating Machine Learning Models Across Multiple Datasets," *MDPI Applied Sciences*, vol. 15, no. 6, p. 3396, 2025. [Online]. Available: https://www.mdpi.com/2076-3417/15/6/3396.

[6] M. Gallo, A. Botta, and G. Ventre, "Machine Learning Algorithms for Phishing Email Detection: A Comparative Study," *AASMR Journal*, vol. 10, no. 2, 2023. [Online]. Available: http://www.aasmr.org/liss/Vol.10/No.2%202023/Vol.10%20No.2.17.pdf.

[7] I. AbdulNabi and Q. Yaseen, "Dual-Path Phishing Detection: Integrating Transformer-Based NLP with Structural URL Analysis," *arXiv preprint arXiv:2509.20972*, 2025. [Online]. Available: https://arxiv.org/abs/2509.20972.

[8] H. Bacha, "Detecting Phishing Attacks Using URL and Content-Based Features: A Data-Driven Classification Model," *ResearchGate*, 2024. [Online]. Available: https://www.researchgate.net/publication/397418477_Detecting_Phishing_Attacks_Using_URL_and_Content-Based_Features_A_Data-Driven_Classification_Model.

[9] N. A. Alam, "Phishing Email Dataset," *Kaggle*, 2023. [Online]. Available: https://www.kaggle.com/datasets/naserabdullahalam/phishing-email-dataset. [Accessed: 27-Nov-2025].

[10] W. Cukierski, "Enron Email Dataset," *Kaggle*, 2015. [Online]. Available: https://www.kaggle.com/datasets/wcukierski/enron-email-dataset. [Accessed: 27-Nov-2025].