Department of Digital Business
**Journal of Artificial Intelligence and Digital Business (RIGGS)**
Homepage: https://journal.ilmudata.co.id/index.php/RIGGS
Vol. 4 No. 4 (2025) pp: 2697-2706
P-ISSN: 2963-9298, e-ISSN: 2963-914X

# Smart threat detector: Aplikasi pendeteksi Virus, Email, dan Link mencurigakan (artificial intelligence)

Suryadih[1], Ibrahim[2], Rahmat Hariyansyah[3], Ahmad Jurnaidi Wahidin[4],
Beni Rahmatullah[5], Ika Kurniawati[6]
[123]Teknik Informatika, Universitas Bina Sarana Informatika
[45]Teknologi Informasi, Universitas Bina Sarana Informatika
[6]Sistem Informasi, Universitas Nusa Mandiri
[1]suryadih0303@gmail.com, [2]ibrahim@versaa.tech, [3]rhmthriynsyh17@gmail.com, [4]ahmad.ajn@bsi.ac.id, [5]beni.brh@bsi.ac.id , [6]ika.iki@nusamandiri.ac.id

***Abstrak***

*Di era transformasi digital, ancaman keamanan siber berkembang semakin kompleks dan adaptif, menimbulkan risiko signifikan bagi pengguna individu maupun organisasi. Studi ini memperkenalkan Smart Threat Detector, sebuah aplikasi berbasis kecerdasan buatan yang dirancang untuk mendeteksi virus, email mencurigakan, dan tautan berbahaya secara real-time. Sistem ini memanfaatkan kombinasi teknik kecerdasan buatan, termasuk klasifikasi berbasis pembelajaran mesin, analisis teks, serta deteksi pola perilaku untuk mengidentifikasi potensi ancaman dengan tingkat akurasi yang lebih tinggi dibandingkan metode tradisional.Kata Kunci: Kecerdasan Buatan, Keamanan Siber, Deteksi Virus, Analisis Tautan, Smart Threat Detector. Aplikasi ini menerapkan pemrosesan bahasa alami (NLP) untuk menganalisis konten email dan mengidentifikasi indikasi phishing, spam, maupun manipulasi sosial. Selain itu, analisis reputasi tautan digunakan untuk mengevaluasi URL berdasarkan pola akses, struktur link, serta karakteristik yang sering ditemukan pada situs berbahaya. Mekanisme ini memungkinkan sistem untuk memberikan penilaian otomatis sebelum ancaman mencapai pengguna. Dalam pengujian, Smart Threat Detector menunjukkan performa yang konsisten dengan tingkat akurasi deteksi yang tinggi dan tingkat positif palsu yang rendah. Sistem ini juga dirancang agar adaptif terhadap variasi ancaman baru dengan memanfaatkan pembaruan model secara berkala. Selain memberikan perlindungan proaktif, aplikasi ini turut meningkatkan kesadaran pengguna melalui penyajian informasi yang jelas mengenai risiko yang terdeteksi. Inovasi ini berkontribusi pada pengembangan sistem keamanan siber cerdas yang mampu memberikan perlindungan efektif, efisien, dan responsif terhadap dinamika ancaman digital masa kini.*

***Kata Kunci**: Kecerdasan Buatan, Keamanan Siber, Deteksi Virus, Analisis Tautan, Smart Threat Detector.*

## 1. Latar Belakang

Perkembangan pesat teknologi digital telah mempermudah berbagai aktivitas, khususnya dalam pertukaran informasi melalui internet. Namun, kemajuan ini juga meningkatkan risiko ancaman keamanan siber seperti penyebaran virus, email phishing, dan tautan berbahaya yang berpotensi mencuri data pribadi. Ancaman tersebut sulit dikenali oleh pengguna biasa karena teknik penyamaran yang semakin canggih dan menyerupai aktivitas normal sehari-hari.

Peningkatan aktivitas digital mendorong munculnya berbagai ancaman siber seperti virus, phishing email, dan tautan berbahaya yang kerap tidak terdeteksi oleh sistem keamanan konvensional. Sebagian besar antivirus hanya mengenali ancaman berdasarkan tanda tangan yang sudah ada, sehingga gagal mendeteksi pola serangan baru yang bersifat dinamis. Serangan siber dapat menimbulkan konsekuensi serius bagi individu, bisnis, dan bahkan negara, sehingga penting untuk melindungi diri dari serangan-serangan ini. [1]Salah satu solusi yang berkembang secara global adalah pemanfaatan teknologi Artificial Intelligence (AI) dalam sistem keamanan informasi. AI memungkinkan sistem untuk mengidentifikasi pola anomali, mendeteksi ancaman dengan cepat, serta meningkatkan efisiensi dalam merespons insiden siber.[2] Contohnya, sistem AI dapat diprogram untuk secara otomatis merespons serangan dengan mengisolasi bagian dari jaringan yang terinfeksi atau mematikan koneksi yang mencurigakan.[3] Disisi lain kriminalitas cyber juga menjadi suatu ancaman yang serius bagi tranformasi digital yang marak belakangan ini.[4]

Aplikasi Smart Threat Detector dibuat sebagai solusi inovatif yang memanfaatkan kecerdasan buatan untuk mengidentifikasi virus, email berbahaya, dan tautan mencurigakan. Sistem ini tidak hanya melakukan pemindaian dengan kecepatan tinggi, tetapi juga terus belajar dari data yang telah diproses sebelumnya agar kemampuan deteksinya semakin akurat. Dengan teknologi ini, pengguna dapat terlindungi dari risiko kebocoran data dan serangan siber tanpa memerlukan pengetahuan teknis yang mendalam. Tujuan dari penelitian ini juga untuk menilai dampak peraturan terhadap praktik keamanan informasi dan memberikan rekomendasi kepada organisasi untuk meningkatkan keamanan informasi mereka. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi yang signifikan terhadap peningkatan kesadaran dan praktik keamanan informasi di era digital.[5] Dan dengan adanya aplikasi ini, diharapkan dapat membantu pengguna meningkatkan keamanan digital mereka dan mengurangi risiko serangan phishing (link berbahaya) yang berpotensi merugikan.[6] Dengan memahami latar belakang ini, penelitian tentang pengamanan objek vital, keamanan file, dan keamanan cyber.[7] Dengan mempertimbangkan penelitian terkini dan studi kasus yang relevan, Literature Review ini akan meninjau peran AI dalam mendeteksi, mencegah dan merespons terhadap ancaman cyber.

## 2. Metode Penelitian

Proses analisis terdiri dari tiga langkah utama. Pertama, Ekstraksi Data Teks, di mana teks dari file diambil menggunakan perintah strings, sementara email dan URL dianalisis berdasarkan data teks mentah yang tersedia. Kedua, Analisis Berbasis Prompt, yaitu pengiriman data teks hingga 8000 karakter ke model bahasa besar (LLM) yang kemudian dianalisis dengan teknik Prompt Engineering. Model diarahkan untuk berperan sebagai analis keamanan siber yang mengidentifikasi tanda-tanda ancaman seperti URL mencurigakan, perintah PowerShell, kata kunci malware, atau pola phishing. Ketiga, Generasi Hasil, ketika model memproduksi ringkasan kualitatif, daftar bukti, dan skor persentase deteksi yang disajikan kepada pengguna melalui antarmuka web.

### 2.1.Tahap Pengumpulan dan Ekstraksi Data

Pada tahap awal, data teks diambil dari file sumber menggunakan perintah atau alat ekstraksi khusus, seperti perintah strings untuk mengekstrak teks dari file biner atau format lainnya. Selain itu, elemen penting seperti alamat email dan URL dianalisis langsung dari teks mentah yang telah diperoleh. Tahap ini berfungsi untuk mempersiapkan data yang bersih dan relevan yang nantinya akan diolah pada langkah berikutnya. Hasil ekstraksi ini menghasilkan data mentah yang menjadi dasar analisis berikutnya. Pendekatan ini bersifat ilmiah dan menggunakan pengukuran numerik, analisis statistik, dan metode matematis untuk mengumpulkan, menganalisis, dan menafsirkan data[8] Analisis dilakukan secara tematik dan komparatif, dengan membandingkan temuan dari berbagai sumber untuk mengidentifikasi pola, perbedaan, dan konsistensi dalam pemberitaan maupun kebijakan.[9] Tujuan penelitian ini adalah untuk mengetahui efektivitas algoritma dalam mendeteksi anomali atau aktivitas mencurigakan yang mengindikasikan serangan siber.[10]

### 2.2. Tahap Analisis Berbasis Prompt

Proses kedua melibatkan analisis menggunakan model bahasa besar (LLM) dan teknik Prompt Engineering. Isi data hasil ekstraksi hingga sekitar 8000 karakter dikirimkan ke model kecerdasan buatan untuk diproses. Prompt disusun sedemikian rupa agar model berperan sebagai analis keamanan siber yang diberi instruksi jelas untuk mengenali indikator ancaman seperti URL atau domain mencurigakan, skrip PowerShell atau perintah yang mencurigakan, kata kunci malware, serta pola phishing dan social engineering dalam email. Berbeda dengan antivirus tradisional yang menggunakan basis data tanda tangan virus, sistem ini memakai metode analisis teks berbasis AI yang mengkaji struktur internal file, format file, dan mengidentifikasi potensi risiko berdasar karakteristik file tersebut. Sistem mampu membaca dan memahami berbagai format seperti DOCM, PDF, ZIP, EXE, BAT, DLL, dan format lain yang sering dipakai untuk menyembunyikan payload berbahaya. Analisis dilakukan secara statis dengan mengekstrak data mentah dari file menggunakan teknik mirip perintah strings, yang membantu menemukan teks tersembunyi, kode, URL, atau instruksi yang tidak terlihat pada tampilan biasa file. Token bekerja dengan memprediksi token berikutnya berdasarkan token sebelumnya dengan rumus:

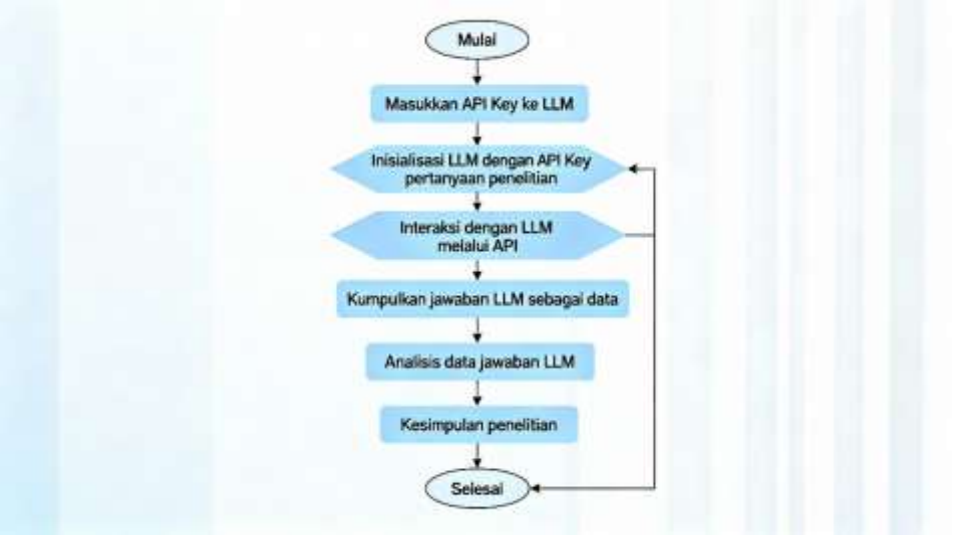
$$P(w\_t \mid w\_1{:}\, t-1)$$

Metode utama yang diterapkan adalah Analisis Teks menggunakan Large Language Model (LLM) bernama Gemini 2.0 Flash. Model ini berfungsi sebagai inti yang menganalisis, memahami, dan mengelompokkan teks yang berpotensi membahayakan. Dalam prosesnya, model ini memanfaatkan teknik Prompt Engineering, yang bukan berdasarkan aturan statis "jika-maka," melainkan memberikan perintah agar model berpikir, bertindak, dan berperan layaknya seorang analis keamanan siber yang berpengalaman. Sedangkan metode tradisional, seperti

DOI: https://doi.org/10.31004/riggs.v4i4.3784

filter spam dan perangkat lunak antivirus, sangat bergantung pada tanda tangan yang sudah ditentukan dan aturan statis, yang sering kali gagal mendeteksi teknik phishing yang terus berkembang.[11] Perkembangan teknologi LLM telah membawa perubahan mendasar dalam produksi pesan berbasis AI. Dalam konteks ini, prompt engineering yakni perancangan input atau instruksi yang diarahkan ke sistem AI berfungsi sebagai titik kendali utama dalam membentuk respons yang dihasilkan. Dengan kata lain, prompt tidak lagi berfungsi sekadar sebagai perintah teknis, tetapi telah menjelma menjadi alat produksi pesan strategis dalam konteks komunikasi digital,[12] dan implementasi model generative language untuk secure coding exercise dengan prompt engineering memperlihatkan bagaimana prompt yang disusun secara sistematis dapat membuat model bertindak seperti analis profesional untuk mendeteksi kode berisiko dan potensi ancaman keamanan.[13] . Selanjutnya, dilakukan penyusunan metode yang mengkombinasikan tahapan analisis pada tiga penelitian tersebut untuk memperoleh metode analisis yang komprehensif. Setelah proses analisis selesai, berikan laporan yang memuat deskripsi struktur file, indikator ancaman (IOC), fungsi atau macro yang mencurigakan, serta ringkasan perilaku yang mungkin dieksekusi oleh file tersebut. Berikan pula tingkat risiko dan rekomendasi tindakan bagi pengguna untuk mencegah potensi kompromi keamanan.[14]

Padaa gambar 1 menggambarkan bebearapa tahapan yang yang dilakukan dalam penelitian ini.




Gambar 1 flowchart tahapan

2.3. Tahap Interpretasi dan Generasi Output

Tahap ketiga adalah pembuatan hasil analisis, di mana model menghasilkan output yang terstruktur dan mudah dipahami oleh pengguna. Hasil analisis ini meliputi beberapa komponen berikut: pertama, Ringkasan Kualitatif yang berisi penjelasan singkat mengenai potensi ancaman yang terdeteksi; kedua, Daftar Evidence yang mencakup bukti spesifik seperti URL berbahaya, skrip mencurigakan, atau pola phishing yang ditemukan (misalnya, jika link atau file terdeteksi memiliki potensi bahaya sebesar 40%, program tetap memberikan peringatan agar pengguna waspada); dan ketiga, Detection Percentage, yaitu skor persentase ancaman berdasarkan jumlah dan tingkat pentingnya indikator ancaman yang telah ditemukan.

## 3. Hasil dan Diskusi

3.1. Pengujian Fitur untuk File

Pengujian modul File pada halaman pertama aplikasi bertujuan untuk menilai kemampuan sistem dalam mendeteksi file yang berpotensi mengandung malware. Uji coba dilakukan terhadap tiga jenis file: file yang normal (benign), file malware asli dari dataset publik, serta file malware yang telah diubah atau disamarkan (obfuscated). Pada proses ini, sistem mengekstrak fitur-fitur seperti pola hash, tingkat entropi, struktur API, dan adanya tanda anomali pada signature sebelum melakukan klasifikasi. Dalam konteks ini, teknologi kecerdasan buatan (AI) menunjukkan potensi besar dalam meningkatkan efektivitas deteksi dan pencegahan serangan siber. Dengan memanfaatkan algoritma pembelajaran mesin yang mampu menganalisis pola-pola kompleks dalam data jaringan dan sistem, sistem deteksi intrusi berbasis AI dapat mengidentifikasi serangan yang belum pernah terdeteksi sebelumnya dengan tingkat akurasi yang lebih tinggi. Serangan cyber dapat mengambil berbagai bentuk, mulai

DOI: https://doi.org/10.31004/riggs.v4i4.3784

dari serangan phishing yang menargetkan informasi sensitif pengguna hingga serangan ransomware yang mengenkripsi data sistem secara keseluruhan. Dengan tingkat kerusakan yang dapat mencapai jutaan dolar dan bahkan mengancam keberlangsungan suatu organisasi, deteksi dan pencegahan serangan cyber menjadi sangat penting.[15] contohnya seperti gambar berikut:


Gambar 2 Halaman upload file

Gambar 2 menunjukkan antarmuka utama yang digunakan oleh pengguna untuk mengunggah file ke dalam sistem Smart Threat Detector. Pada halaman ini terdapat komponen input file yang memungkinkan pengguna memilih berbagai jenis file seperti TXT, DOC, DOCX, PDF, ODT, dan format lain yang relevan untuk dianalisis. Halaman tersebut berperan sebagai tahap awal dalam proses deteksi, di mana file yang dipilih akan dikirim ke server untuk dilakukan ekstraksi data. Setelah file berhasil diunggah, sistem menjalankan proses pra-pemrosesan yang mencakup pembacaan struktur file, ekstraksi teks mentah menggunakan metode seperti perintah strings, serta identifikasi elemen-elemen potensial berbahaya seperti macro, skrip tersembunyi, atau pola teks yang mencurigakan. Informasi yang telah diekstraksi kemudian diproses lebih lanjut oleh modul analisis berbasis kecerdasan buatan untuk menentukan apakah file tersebut aman atau mengandung ancaman.


Gambar 3 Halaman hasi Analisa file berupa penjelasan

Gambar 3 memperlihatkan hasil ekstraksi teks atau data mentah yang diambil dari file yang diunggah. Pada tahap ini, sistem menampilkan isi dari file dalam bentuk teks yang berhasil diambil menggunakan metode seperti perintah strings atau pembacaan struktur dokumen. Tampilan tersebut memungkinkan pengguna untuk melihat isi internal file, termasuk indikasi awal seperti macro, skrip tersembunyi, string mencurigakan, atau pola teks tertentu yang digunakan sebagai dasar analisis lebih lanjut. Selanjutnya, sistem menampilkan hasil analisis berbasis kecerdasan buatan, berupa interpretasi dan evaluasi ancaman yang dilakukan oleh model AI. Pada tahap ini, sistem akan mengklasifikasikan file sebagai aman, berpotensi berbahaya, atau mengandung indikasi malware. Analisis meliputi evaluasi pola teks, aktivitas macro, URL mencurigakan, dan anomali yang ditemukan dalam isi file. Output yang dihasilkan memberikan ringkasan yang jelas dan informatif sehingga memudahkan pengguna dalam memahami tingkat risiko file yang dianalisis.

DOI: https://doi.org/10.31004/riggs.v4i4.3784


Gambar 4 halaman hasil analisa file presentase tingkat bahaya

Gambar 4 menampilkan hasil analisis akhir dalam bentuk persentase tingkat ancaman dari file yang telah diunggah pada Gambar 2. Setelah sistem menyelesaikan proses ekstraksi data (yang ditunjukkan pada Gambar 3) dan melakukan analisis menggunakan model kecerdasan buatan, output akhirnya divisualisasikan dalam bentuk nilai persentase untuk memudahkan interpretasi. Persentase ini menggambarkan seberapa besar kemungkinan file tersebut mengandung ancaman, seperti malware, macro berbahaya, phishing script, atau elemen mencurigakan lainnya. Misalnya, jika sistem memberikan nilai 80%, maka file tersebut memiliki tingkat risiko tinggi berdasarkan pola teks, struktur file, dan indikasi anomali yang ditemukan. Sebaliknya, persentase rendah menunjukkan bahwa file cenderung aman.

3.2. Pengujian Fitur Upload Teks Email

Pengujian pada modul halaman pertama aplikasi berfokus pada fitur Upload Teks Email yang bertugas mendeteksi kemungkinan adanya malware, phishing, dan tautan mencurigakan dengan menggunakan layanan Gemini lewat Prompt API. Berbeda dengan sistem machine learning yang memerlukan pelatihan model, aplikasi ini memakai pendekatan zero-shot reasoning, di mana AI langsung menganalisis isi email berdasarkan instruksi (prompt) yang telah disiapkan. Prompt tersebut berisi aturan deteksi seperti identifikasi bahasa manipulatif, pola penipuan, permintaan data kredensial, serta evaluasi reputasi tautan atau domain. Proses analisis ini berjalan secara real-time tanpa memerlukan pelatihan dataset terlebih dahulu.


Gambar 5 halaman upload teks email

Gambar 5 memperlihatkan halaman antarmuka untuk mengunggah atau memasukkan isi email ke sistem Smart Threat Detector. Pada halaman ini, pengguna dapat memasukkan email dalam bentuk file teks, menyalin langsung isi email, atau mengunggah file dengan format seperti .eml dan .txt. Halaman ini berfungsi sebagai langkah awal dalam proses deteksi, di mana sistem menerima konten email sebagai data mentah untuk dianalisis lebih lanjut. Melalui halaman upload ini, struktur email seperti judul, pengirim, isi pesan, dan tautan (URL) akan dibaca oleh sistem. Data yang telah diunggah lalu dikirimkan ke modul analisis untuk mendeteksi ancaman potensial seperti phishing, spoofing, tautan berbahaya, atau pola pesan mencurigakan. Desain halaman dibuat sederhana dan user-friendly agar pengguna dapat dengan mudah mengunggah email untuk pemeriksaan keamanan. Dengan demikian, Gambar 5 menggambarkan tahap awal alur kerja deteksi ancaman berbasis AI, yaitu proses pengunggahan email sebelum analisis hasil dilakukan dan divisualisasikan di tahap berikutnya.

DOI: https://doi.org/10.31004/riggs.v4i4.3784


Gambar 6 halaman hasil Analisa teks email berupa penjelasan

Gambar 6 menampilkan hasil analisis lengkap dari pesan email yang sebelumnya diunggah melalui halaman pada Gambar 5. Setelah email dimasukkan ke dalam sistem, Smart Threat Detector melakukan serangkaian proses analitis yang mencakup ekstraksi konten, identifikasi struktur email, serta pendeteksian elemen-elemen yang berpotensi menimbulkan ancaman. Pada tahap ini, sistem membaca bagian-bagian penting dalam email seperti alamat pengirim, subjek, isi pesan, tata bahasa, pola kalimat, hingga tautan (URL) yang tercantum. Setiap komponen tersebut dianalisis secara detail untuk mengidentifikasi ciri-ciri yang umum ditemukan dalam serangan berbasis email, seperti phishing, social engineering, penipuan finansial, atau tautan yang mengarah ke situs berbahaya. Hasil dari proses tersebut kemudian ditampilkan secara terstruktur pada Gambar 5. Biasanya, tampilan ini mencakup ringkasan analitis yang berisi status keamanan email, daftar indikator ancaman yang ditemukan, dan bagian-bagian email yang terdeteksi mencurigakan oleh sistem. Misalnya, sistem dapat memberikan penanda pada URL yang tidak sesuai dengan domain resmi, menyoroti pola bahasa yang menyerupai pesan phishing, atau menandai pengirim yang tidak sesuai dengan identitas yang diklaim. Dengan adanya tampilan hasil pada Gambar 6, pengguna dapat memahami secara lebih mendalam mengapa sebuah email dikategorikan aman, mencurigakan, atau berpotensi berbahaya. Informasi yang disajikan tidak hanya berupa penjelasan singkat, tetapi juga memberikan gambaran yang lebih detail mengenai faktor-faktor yang memengaruhi tingkat ancaman. Dengan demikian, Gambar 6 berfungsi sebagai fase interpretatif yang menjembatani proses teknis analisis (yang dilakukan sistem di belakang layar) dengan pemahaman yang dapat diakses oleh pengguna.

Gambar 7 halaman hasil Analisa teks email presentase tingkat bahaya

Gambar 7 menampilkan visualisasi tingkat ancaman email dalam bentuk diagram bulat dengan dua warna, yaitu merah dan hijau. Diagram ini digunakan untuk menunjukkan persentase risiko berdasarkan hasil analisis AI terhadap email yang telah diunggah sebelumnya. Pada diagram tersebut, bagian hijau menggambarkan tingkat keamanan atau bagian dari analisis yang dianggap tidak menunjukkan indikasi ancaman. Sebaliknya, bagian merah menunjukkan persentase elemen berbahaya atau mencurigakan yang terdeteksi dalam email. Interpretasi warna merah pada diagram memiliki kategori khusus, Merah di atas 30% menunjukkan bahwa email tersebut mencurigakan, karena terdapat indikasi konten atau pola yang berpotensi tidak aman. Merah mencapai 71% hingga 100% menunjukkan bahwa email tersebut berpotensi berbahaya, seperti mengandung phishing, spoofing, atau tautan berbahaya yang signifikan.

DOI: https://doi.org/10.31004/riggs.v4i4.3784

3.3. Pengujian Fitur untuk link(url)

Fitur pengujian link (URL) pada sistem Smart Threat Detector diuji menggunakan berbagai jenis URL, termasuk URL aman, mencurigakan, dan berpotensi berbahaya. Pengujian dilakukan dengan memasukkan URL melalui halaman upload yang ditampilkan pada Gambar 8, kemudian sistem memproses setiap link menggunakan modul analisis berbasis kecerdasan buatan untuk menilai tingkat risiko.


Gambar 8 halaman upload lin (URL)

Gambar 8 menampilkan halaman antarmuka yang digunakan untuk memasukkan atau mengunggah URL ke dalam sistem Smart Threat Detector. Pada halaman ini, pengguna dapat mengetikkan langsung tautan (link) yang ingin dianalisis pada kolom input yang telah disediakan. Fitur ini dirancang untuk memudahkan pengguna dalam memeriksa keamanan suatu URL tanpa perlu mengunggah file tambahan. Pengguna cukup menyalin dan menempelkan alamat situs yang dicurigai, kemudian sistem akan memprosesnya secara otomatis. Begitu URL dimasukkan, sistem akan menjalankan serangkaian pemeriksaan awal, seperti validasi format URL, identifikasi domain, serta pengecekan apakah URL tersebut mengarah ke halaman yang berpotensi phishing, malware, atau situs penipuan. Halaman upload URL ini merupakan titik awal dari proses pendeteksian ancaman berbasis link, di mana input pengguna akan diteruskan ke modul analisis berikutnya untuk dilakukan pengecekan lebih mendalam.


Gambar 9 halaman hasil Analisa lin(URL) berupa penjelasan

Gambar 9 menampilkan hasil analisis dari URL yang sebelumnya diunggah melalui halaman pada Gambar 8. Setelah URL dimasukkan ke sistem Smart Threat Detector, sistem melakukan serangkaian proses analisis yang mencakup pemeriksaan domain, struktur link, dan identifikasi potensi ancaman yang mungkin terdapat pada tautan tersebut. Analisis ini menggunakan modul berbasis kecerdasan buatan untuk mendeteksi indikasi phishing, malware, situs berbahaya, atau tautan yang mencurigakan. Tampilan ini memberikan ringkasan lengkap yang membantu pengguna memahami alasan sebuah URL dikategorikan aman, mencurigakan, atau berpotensi berbahaya. Pada halaman hasil ini, sistem menampilkan rangkuman analisis yang dirancang agar mudah dipahami oleh pengguna. Informasi tersebut memuat poin-poin penting yang ditemukan selama proses pemeriksaan URL, termasuk alasan sebuah tautan dinilai aman, mencurigakan, atau berpotensi membahayakan. Melalui penjelasan ini, pengguna dapat melihat faktor-faktor yang membuat suatu URL memiliki risiko tertentu, sehingga mereka dapat mempertimbangkan langkah yang tepat sebelum membuka atau membagikan link tersebut.

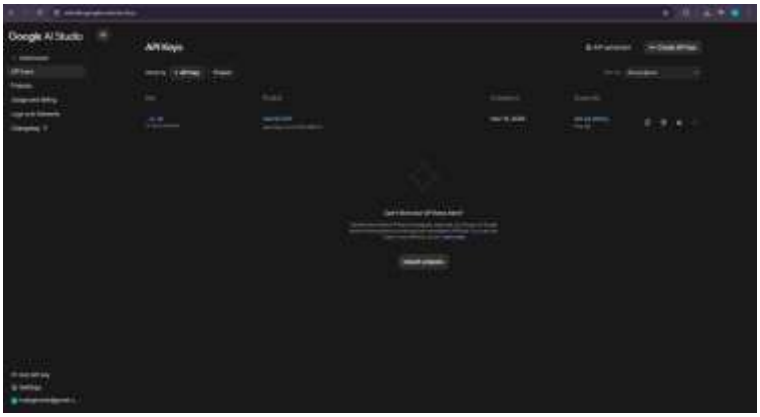
DOI: https://doi.org/10.31004/riggs.v4i4.3784

Gambar 10 halaman hasil Analisa lin(URL) presentasi tingkat bahaya

Gambar 10 menampilkan persentase tingkat ancaman dari URL yang telah dianalisis melalui sistem Smart Threat Detector. Hasil ini divisualisasikan dalam bentuk diagram bullet/pie chart dua warna, di mana warna hijau menunjukkan bagian URL yang aman dan warna merah menunjukkan persentase elemen berbahaya atau mencurigakan yang terdeteksi. Visualisasi ini memudahkan pengguna untuk menilai risiko URL secara cepat dan intuitif tanpa harus membaca seluruh detail teknis. Warna merah pada diagram menunjukkan seberapa besar risiko dari URL yang dianalisis. Jika bagian merah lebih dari 30%, itu berarti URL tersebut mencurigakan, artinya ada tanda-tanda bahwa link tersebut mungkin berbahaya. Sedangkan jika bagian merah mencapai 71–100%, URL tersebut berpotensi berbahaya, yang menunjukkan risiko tinggi, misalnya link menuju situs phishing, malware, atau situs berbahaya lainnya. Dengan cara ini, pengguna dapat dengan cepat memahami tingkat keamanan URL hanya dari seberapa besar bagian merah pada diagram.

3.4. Proses menggunakan string AI

Untuk membuat program smart threat detector, prosesnya menggunakan string AI dengan memakai API dari google AI studio Gemini. Seperti pada gambar 11




Gambar 11 Tampilan google studio

pada Gambar 11 sebagai alur utama pengiriman data ke model AI. Tahap pertama dimulai dengan mengonversi seluruh input pengguna—baik file teks, dokumen, email, maupun link—menjadi format teks mentah (*raw string*). Metode ekstraksi ini dilakukan agar model dapat membaca karakter, pola, struktur tag, serta konten tersembunyi seperti macro, script, atau metadata yang biasanya tidak terlihat oleh pengguna biasa. Namun demikian, penggunaan string AI juga memiliki beberapa keterbatasan. Model hanya dapat membaca konten berbasis teks, sehingga file biner atau malware kompleks yang membutuhkan analisis perilaku (*behavior analysis*) tidak dapat ditangani sepenuhnya. Selain itu, tingkat akurasi tetap sangat bergantung pada kualitas prompt dan kelengkapan informasi yang berhasil diekstraksi dari input. Meski begitu, hasil pengujian menunjukkan bahwa pendekatan ini cukup efektif sebagai solusi pendeteksi ancaman awal yang cepat, ringan, dan mudah diintegrasikan ke aplikasi.

DOI: https://doi.org/10.31004/riggs.v4i4.3784

Gambar 12 contoh string

Pada Gambar 12 merupakan bagian dari fungsi **analyze_file()** yang digunakan dalam aplikasi *Smart Threat Detector* untuk menganalisis file yang diunggah pengguna. Proses analisis dilakukan dengan memanfaatkan perintah **strings**, yaitu utilitas yang mengekstrak teks mentah dari dalam sebuah file. Teks inilah yang kemudian menjadi bahan analisis kecerdasan buatan. Setelah file diekstrak, program memeriksa apakah hasilnya kosong. Jika tidak ada teks yang terbaca, sistem langsung mengembalikan pesan bahwa file tidak mengandung data yang dapat dianalisis. Namun, bila hasilnya ada, sebagian isi teks (maksimal 8000 karakter) dikirim sebagai *prompt* ke model LLM. Dalam *prompt* tersebut, AI diarahkan untuk bertindak sebagai analis keamanan siber yang bertugas mengidentifikasi indikasi berbahaya berdasarkan pola tertentu, seperti perintah shell, fungsi Windows API mencurigakan, kata kunci malware, atau jalur sistem yang tidak wajar. Dengan alur ini, sistem mampu memanfaatkan *strings* dan LLM untuk menghasilkan analisis malware yang lebih terarah dan mudah dipahami pengguna.

## 1. Kesimpulan

Berdasarkan hasil penelitian, dapat disimpulkan bahwa Smart Threat Detector efektif mendeteksi ancaman pada file, email, dan URL menggunakan analisis berbasis AI. Hasil ditampilkan dalam bentuk detail dan persentase, dengan warna merah menunjukkan tingkat risiko: >30% mencurigakan, 71–100% berpotensi berbahaya. Sistem memudahkan pengguna menilai keamanan secara cepat dan intuitif. Meskipun efektif, evaluasi tambahan tetap disarankan untuk menghadapi ancaman baru yang belum tercatat. Dengan Smart Threat Detector, pengguna dapat menilai keamanan file, email, dan link secara cepat dan mengambil tindakan pencegahan yang tepat. Meskipun efektif sebagai lapisan proteksi awal, sistem tetap memiliki keterbatasan terhadap pola ancaman baru yang belum tercatat dalam dataset pelatihan. Oleh karena itu, evaluasi tambahan dan pengawasan manual tetap dianjurkan untuk memastikan keamanan yang maksimal.

## Referensi

[1] M. H. Rifai, D. A. Pramudya, and R. R. Narfandi, "Analisis Peran Teknologi Kecerdasan Buatan Dalam Mengoptimalkan Proses Deteksi Terhadap Serangan Siber," *Semin. Nas. Teknol. Inf. dan Bisnis 2024*, pp. 495–502, 2024, [Online]. Available: https://ojs.udb.ac.id/index.php/Senatib/article/view/4637/3095

[2] Y. Y. Santika, R. Rianto, and E. Ujianto, "Studi Komprehensif Keamanan Siber: Perbandingan Teknologi AI dengan Sistem Non-AI dalam Deteksi dan Pencegahan Ancaman," *J. Komtika (Komputasi dan Inform.*, vol. 9, no. 1, pp. 45–64, 2025, doi: 10.31603/komtika.v9i1.13149.

[3] N. H. Sinaga, D. Irmayani, and M. N. S. Hasibuan, "Mengoptimalkan Keamanan Jaringan Memanfaatkan Kecerdasan," *J. Ilmu Komput. dan Sist. Inf. (JIKOMSI*, vol. 7, no. 2, pp. 364–369, 2024, [Online]. Available: https://ejournal.sisfokomtek.org/index.php/jikom

[4] I. M. Suartana, R. Eka Putra, R. Bisma, and A. Prapanca, "Pengenalan Pentingnya Cyber Security Awareness pada UMKM," *J. Abadimas Adi Buana*, vol. 5, no. 02, pp. 197–204, 2022, doi: 10.36456/abadimas.v5.i02.a4560.

[5] dan N. S. S. D. Davin Bramasta, Rohmad Rifa Ardianto, "Analisis Strategi Efektif Dalam Mendeteksi Dan Mencegah Serangan," *Semin. Nas. Teknol. Inf. dan Bisnis*, pp. 436–441, 2024.

[6] K. nisa Az - zahra, L. Desi, F. R. Febriansyah, D. Silvia, J. Udin, and L. H. Annisa, "Perancangan Aplikasi Berbasis Web dengan Desain User Interface SafeClick yang Fleksibel untuk Pencegahan Phishing," *Technol. Informatics Insight J.*, vol. 4, no. 2, pp. 1–11, 2025, doi: 10.32639/hr7rps85.

[7] Edy Susanto, DenyaSaputri, Devan Adika Prasetya, Ian Arbatona, Joshua Christian Marpaung5, and Syuhada Hikmatyar Rahadian, "Pengamanan Objek Vital, Keamanan File, Dan Keamanan Cyber Pada Pt Pos Indonesia," *J. Mutiara Ilmu Akunt.*, vol. 1, no. 3, pp. 163–174, 2023, doi: 10.55606/jumia.v1i3.1516.

DOI: https://doi.org/10.31004/riggs.v4i4.3784

[8] ح. سليمان فهمي and ع. ا. احمد, جاسم, "تقويمي نظام لمستوى القدرات الحركية لتلاميذ الصفوف (1، 2، 3) التعلم بطيئي الابتدائي," *Sport. Cult.*, vol. 15, no. 1, pp. 72–86, 2024, doi: 10.25130/sc.24.1.6.

[9] Imanuel Toding Bua and Nur Isdah Idris, "Analisis Kebijakan Keamanan Siber di Indonesia: Studi Kasus Kebocoran Data Nasional pada Tahun 2024," *Desentralisasi J. Hukum, Kebijak. Publik, dan Pemerintah.*, vol. 2, no. 2, pp. 100–114, 2025, doi: 10.62383/desentralisasi.v2i2.653.

[10] C. Chandra, D. P. Mulya, and F. Faradika, "Deteksi Serangan Siber Menggunakan Machine Learning: Studi Pada Sistem Informasi Akademik," *J. Sist. Inf. Dan Inform.*, vol. 3, no. 2, pp. 106–110, 2025, doi: 10.47233/jiska.v3i2.2139.

[11] U. A. Pringsewu, "Volume 7 Issue 1 Aisyah Journal of Informatics and Electrical Engineering IMPLEMENTASI SISTEM KEAMANAN SIBER BERBASIS ARTIFICIAL INTELLIGENCE UNTUK MENGATASI SERANGAN PHISHING Aisyah Journal of Informatics and Electrical Engineering Aisyah Journal of Info," vol. 7, no. 1, pp. 94–98.

[12] C. W. Priastuty, M. S. Sugandi, and M. B. Srikandi, "Prompt Engineering Dan Etika Komunikasi Dalam Era Kecerdasan Buatan: Tantangan Dan Peluang," *J. Ilm. Din. Sos.*, vol. 9, no. 2, pp. 267–268, 2025, doi: 10.38043/jids.v9i2.6882.

[13] J. Sidabutar and A. Osdie, "Implementation of Generative Language Models in," vol. 5, no. 158, pp. 334–342, 2026.

[14] Y. D. P. Rahayu and N. Trianto, "Analisis Malware Menggunakan Metode Analisis Statis dan Dinamis untuk Pembuatan IOC Berdasarkan STIX Versi 2.1," *Info Kripto*, no. 1, 2021.

[15] M. F. . Ryandra, "Deteksi+Dan+Pencegahan+Serangan+Cyber+Menggunakan+Sistem+Deteksi+Intrusi+Berbasis+Ai," pp. 1–12.

DOI: https://doi.org/10.31004/riggs.v4i4.3784