Isian Substansi Proposal

SKEMA PENELITIAN DASAR (PENELITIAN DOSEN PEMULA AFFIRMASI, PENELITIAN DOSEN PEMULA, PENELITIAN PASCASARJANA)

Pengusul hanya diperkenankan mengisi di tempat yang telah disediakan sesuai dengan petunjuk pengisian dan tidak diperkenankan melakukan modifikasi template atau penghapusan di setiap bagian.

A. JUDUL

Tuliskan judul usulan penelitian maksimal 20 kata

[Optimasi Deteksi Berbasis Signature pada Host-Based Intrusion Detection System Menggunakan Logistic Regression untuk Mendeteksi Serangan Web Defacement Judi Online]

B. RINGKASAN

Isian ringkasan penelitian tidak lebih dari 300 kata yang berisi urgensi, tujuan, metode, dan luaran yang ditargetkan

[Serangan *web defacement* bermuatan judi online terus meningkat dan mengancam integritas layanan web di Indonesia. Serangan ini sering memanfaatkan celah pada aplikasi web dan sistem operasi *server*. *Host-Based Intrusion Detection System* (HIDS) yang lazim digunakan masih mengandalkan *signature-based detection* berbasis kata kunci untuk mendeteksi perubahan pada file HTML. Pendekatan ini memiliki beberapa keterbatasan, seperti kurangnya pemahaman konteks teks, potensi terhadap *false positive*, serta bergantung pada daftar kata kunci yang tidak selalu mampu merepresentasikan variasi konten defacement yang terus berkembang. Keterbatasan ini mengakibatkan rendahnya akurasi deteksi dan efektivitas *active response*. Penelitian ini mengusulkan optimasi mekanisme deteksi signature pada HIDS dengan pendekatan machine learning menggunakan algoritma *Logistic Regression*. Model klasifikasi dibangun menggunakan dataset halaman *web defacement* yang dikumpulkan secara mandiri, kemudian mengintegrasikan model tersebut ke dalam alur HIDS sehingga *File Integrity Monitoring* (FIM) tidak hanya mendeteksi perubahan file, tetapi juga mengevaluasi konteks perubahan tersebut sebelum memicu *active response*. Evaluasi dilakukan dengan membandingkan performa HIDS berbasis signature murni dengan HIDS yang telah dioptimasi menggunakan *Logistic Regression*, dengan parameter pengukuran mencakup ketepatan deteksi, waktu deteksi, penggunaan sumber daya, dan ketepatan active response. Hasil penelitian diharapkan dapat menghasilkan sistem deteksi yang lebih adaptif, menurunkan tingkat kesalahan deteksi, serta memberikan kontribusi dalam pengembangan mekanisme deteksi intrusi berbasis signature yang lebih kontekstual dan efisien.]

C. KATA KUNCI

Isian 5 kata kunci yang dipisahkan dengan tanda titik koma (;)

[web defacement, HIDS, signature optimization, logistic regression, judi online]

D. PENDAHULUAN

Pendahuluan penelitian tidak lebih dari 1000 kata yang terdiri dari:

- *Latar belakang dan rumusan permasalahan yang akan diteliti*
- *Pendekatan pemecahan masalah*
- *State of the art dan kebaruan*
- *Peta jalan (road map) penelitian 5 tahun*

Sitasi disusun dan ditulis berdasarkan sistem nomor sesuai dengan urutan pengutipan.

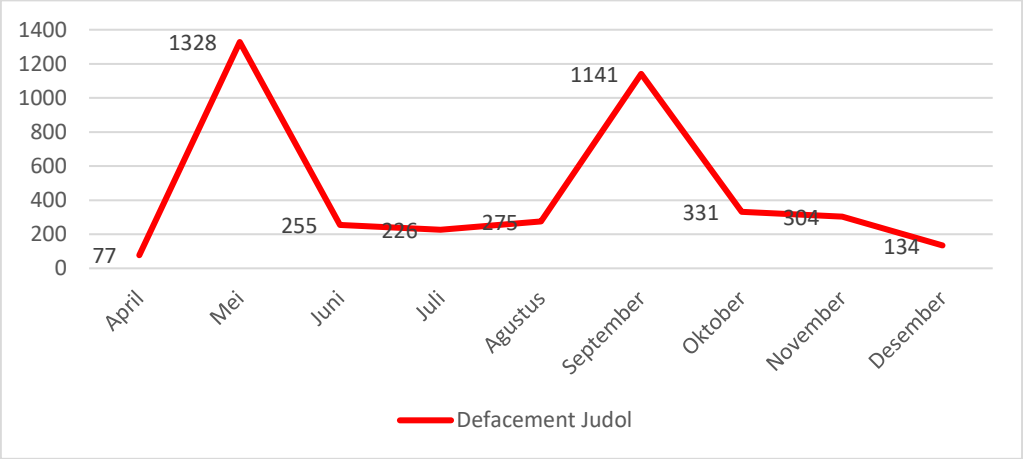
D.1. LATAR BELAKANG DAN RUMUSAN MASALAH

Tuliskan latar belakang penelitian dan rumusan permasalahan yang akan diteliti, serta urgensi dari dilakukannya penelitian ini

[Masifnya transformasi digital di berbagai sektor layanan publik dan pemerintahan diiringi dengan meningkatnya jumlah ancaman siber yang terjadi. Berdasarkan laporan Lanskap Keamanan Siber Indonesia 2024 yang diterbitkan oleh BSSN, sepanjang tahun 2024 tercatat lebih dari 330 juta anomali trafik siber di Indonesia(1). Beberapa aktivitas yang terdeteksi pada anomali tersebut mencakup berbagai bentuk serangan seperti *botnet*, *phishing*, *malware* serta *web defacement* yang berdampak pada gangguan layanan, pencurian data sensitif, dan penurunan kepercayaan publik terhadap sistem digital.

Menurut Holt, et al(2) *web defacement* adalah bentuk peretasan yang melibatkan perubahan konten pada sebuah situs web. Akibat perubahan halaman situs terjadi kehilangan pendapatan, penurunan produktivitas internal, serta kerusakan reputasi. Perkiraan terbaru menunjukkan bahwa 19,7% serangan siber yang terjadi tiap tahun di dunia merupakan *web defacement*.

Di Indonesia, sepanjang tahun 2024 telah terjadi 5.780 serangan *web defacement* dengan domain sch.id, ac.id dan go.id menjadi target utama serangan(1). Selain itu, situs milik pemerintah menjadi target dari 4.071 serangan *web defacement* yang terkait dengan promosi judi online pada tahun 2024(1). Gambar 1 menunjukkan sebaran serangan *web defacement* judi online sepanjang April hingga Desember tahun 2024.



Gambar 1. Serangan Web Defacement Judi Online Tahun 2024

Untuk mengatasi ancaman tersebut, banyak administrator sistem memanfaatkan *Host-Based Intrusion Detection System* (HIDS) yang dilengkapi dengan mekanisme *File Integrity Monitoring* (FIM) guna mendeteksi perubahan file secara real time. Pendekatan yang lazim digunakan pada HIDS adalah deteksi berbasis *signature*, yaitu dengan mencocokkan perubahan file terhadap pola atau kata kunci tertentu yang telah didefinisikan sebelumnya. Dari sisi analisis, pendekatan *signature based detection* hingga saat ini masih menjadi metode yang banyak digunakan dalam sistem deteksi intrusi karena keunggulannya dalam kecepatan deteksi, efisiensi sumber daya, dan kestabilan kinerja, sehingga ideal diterapkan pada lingkungan dengan keterbatasan komputasi seperti *production server*(3). Meskipun efektif, pendekatan signature-based detection memiliki sejumlah keterbatasan. Beberapa keterbatasan yakni ketidakmampuan memahami konteks teks secara menyeluruh, sehingga rentan menghasilkan *false positive* ketika terjadi perubahan isi file(4)(5). Selain itu, ketergantungan pada daftar kata kunci statis menyebabkan sistem sulit beradaptasi terhadap variasi konten web defacement judi online yang terus berkembang dan sering dimodifikasi oleh penyerang untuk menghindari deteksi. Keterbatasan ini berdampak langsung

pada menurunnya akurasi deteksi dan efektivitas mekanisme *active response* yang diterapkan oleh HIDS.

Pendekatan machine learning banyak dimanfaatkan untuk meningkatkan kemampuan sistem deteksi intrusi, khususnya dalam menganalisis pola data yang bersifat kompleks dan kontekstual. Salah satu algoritma yang relatif sederhana namun efektif untuk klasifikasi teks adalah *Logistic Regression*. Algoritma ini memiliki keunggulan dalam hal interpretabilitas model, efisiensi komputasi, serta kemampuannya dalam menghasilkan keputusan klasifikasi yang stabil(6).

Berdasarkan permasalahan tersebut, penelitian ini berfokus pada upaya mengoptimasi mekanisme deteksi berbasis signature pada HIDS dengan mengintegrasikan algoritma *Logistic Regression* sebagai pendukung evaluasi konteks perubahan file web. Integrasi ini diharapkan mampu meningkatkan ketepatan deteksi serangan web defacement judi online dengan cara tidak hanya mendeteksi adanya perubahan file, tetapi juga menganalisis makna dan konteks perubahan tersebut sebelum memicu *active response*. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan sistem deteksi intrusi berbasis host yang lebih adaptif, akurat, dan efisien, serta relevan untuk diterapkan dalam lingkungan server web di Indonesia.]

D.2. PENDEKATAN PEMECAHAN MASALAH

Tuliskan pendekatan dan strategi pemecahan masalah yang telah dirumuskan

[Penelitian ini menerapkan pendekatan optimasi deteksi berbasis signature pada *Host-Based Intrusion Detection System* (HIDS) dengan mengintegrasikan algoritma *Logistic Regression* sebagai mekanisme evaluasi konteks terhadap perubahan file yang terdeteksi oleh *File Integrity Monitoring* (FIM). Konten halaman web yang mengalami perubahan dianalisis secara kontekstual menggunakan representasi fitur teks untuk mengklasifikasikan apakah perubahan tersebut mengindikasikan serangan web defacement bermuatan judi online. Model klasifikasi ini berfungsi sebagai lapisan pendukung dalam pengambilan keputusan deteksi dan pemicu *active response*, tanpa menggantikan mekanisme signature yang sudah ada. Pendekatan ini dievaluasi dengan membandingkan kinerja HIDS sebelum dan sesudah optimasi berdasarkan ketepatan deteksi, tingkat *false positive*, waktu deteksi, dan penggunaan sumber daya sistem.]

D.3. STATE OF THE ART DAN KEBARUAN

Tuliskan keunggulan dari pemecahan masalah yang ditawarkan pengusul dibandingkan dengan penelitian pengusul sebelumnya atau peneliti lainnya dalam konteks permasalahan yang sama, serta kebaruan usulan dari aspek pendekatan, metode, dsb

[Penelitian-penelitian sebelumnya yang berfokus pada deteksi serangan *web defacement* serta deteksi judi online umumnya masih mengandalkan pendekatan *Intrusion Detection System* (IDS) berbasis jaringan (*network-based)* dengan metode *anomaly detection* dan *passive response*. Hal ini dikarenakan metode ini mampu melakukan pemantauan dalam skala besar secara berulang tanpa memerlukan akses langsung ke infrastruktur server, sehingga dinilai efektif untuk implementasi sistem pengawasan yang bersifat berkelanjutan(7). Meskipun pendekatan tersebut mampu menghasilkan akurasi deteksi yang baik, sistem yang diusulkan masih bersifat observasional dan belum mampu memberikan respons otomatis terhadap serangan yang terdeteksi.

Zagi juga mencatat bahwa rata-rata waktu yang dibutuhkan untuk melakukan proses pemulihan setelah insiden *defacement* mencapai sekitar tiga hari(7). Hal ini menunjukkan bahwa mekanisme deteksi dan mitigasi yang ada belum efisien serta masih kurang tanggap terhadap serangan yang terjadi. Sementara itu, hanya sebagian kecil penelitian yang mulai mengeksplorasi pendekatan *host-based* IDS dengan dukungan *active response*, dimana sistem dapat mendeteksi sekaligus

menanggulangi serangan secara otomatis. Pendekatan *signature based detection* banyak digunakan karena kecepatan deteksi, efisiensi sumber daya, dan kestabilan kinerja, sehingga ideal diterapkan pada lingkungan dengan keterbatasan komputasi(3). Namun, di balik keunggulannya, metode ini memiliki keterbatasan mendasar dalam kemampuan analisis kontekstual. Sistem hanya mencocokkan pola atau kata kunci tertentu sehingga tidak dapat membedakan apakah kemunculan kata kunci terkait judi online merupakan bagian dari serangan *web defacement* atau sekadar konten legal yang membahas topik perjudian secara informatif, edukatif, atau jurnalistik.

Pada penelitan ini digunakan pendekatan yang relevan untuk menjawab kebutuhan tersebut adalah penerapan *Logistic Regression* (LR) yang mampu memberikan perfoma kuat dengan kompleksitas komputasi rendah dan efisien serta stabil dalam pemodelan data berbasis teks(8)(9).

]

D.4. PETA JALAN PENELITIAN

Tuliskan peta jalan penelitian dari tahapan yang telah dicapai, tahapan yang akan dilakukan selama jangka waktu penelitian, dan tahapan yang direncanakan.

[Peta jalan penelitian 5 tahun

Tahun 1 (TKT 4)

Pengembangan dan validasi prototipe HIDS berbasis *signature* yang dioptimasi menggunakan *Logistic Regression* untuk deteksi web defacement judi *online* (tesis magister).

Tahun 2 (TKT 4–5)

Peningkatan akurasi dan generalisasi model melalui perluasan dataset serta pengujian pada berbagai skenario serangan.

Tahun 3 (TKT 5)

Integrasi sistem dengan lingkungan operasional terbatas dan evaluasi performa deteksi serta *active response*.

Tahun 4 (TKT 6)

Implementasi pada lingkungan institusi mitra dan penyempurnaan sistem berdasarkan hasil uji lapangan.

Tahun 5 (TKT 7)

Standarisasi, dokumentasi, dan pengembangan menuju adopsi atau hilirisasi teknologi.]

E. METODE

Isian metode atau cara untuk mencapai tujuan yang telah ditetapkan tidak lebih dari 1000 kata. Pada bagian metoda wajib dilengkapi dengan:

- *Diagram alir penelitian yang menggambarkan apa yang sudah dilaksanakan dan yang akan dikerjakan selama waktu yang diusulkan. Format diagram alir dapat berupa file JPG/PNG.*
- *Metode penelitian harus memuat, sekurang-kurangnya proses, luaran, indikator capaian yang ditargetkan, serta anggota tim/mitra yang bertanggung jawab pada setiap tahapan penelitian.*
- *Metode penelitian harus sejalan dengen Rencana Anggaran Biaya (RAB)*

1. [**Tahapan Penelitian**

   Tahap 1 – Studi Literatur dan Analisis Kebutuhan

   Tahap ini meliputi kajian literatur terkait web defacement, HIDS berbasis signature, *File Integrity Monitoring* (FIM), serta penerapan *machine learning* untuk klasifikasi teks. Selain itu, dilakukan analisis kebutuhan sistem untuk menentukan arsitektur HIDS, sumber data log, serta skema integrasi model *Logistic Regression*.

Luaran: Dokumen kebutuhan sistem dan kerangka konseptual penelitian
Indikator capaian: Tersusunnya spesifikasi sistem dan rancangan metode
Penanggung jawab: Mahasiswa (dibimbing oleh dosen pembimbing)

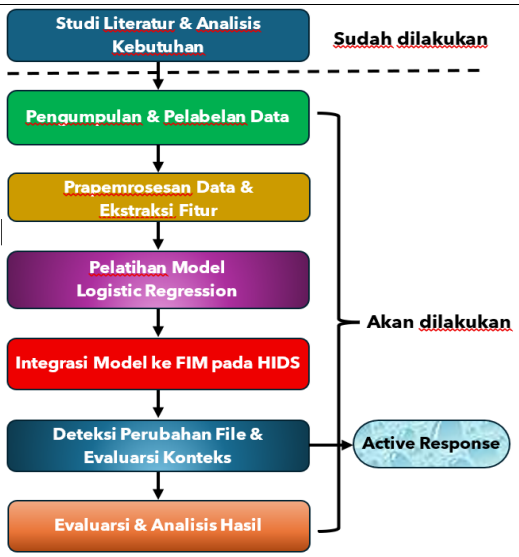
Tahap 2 – Pengumpulan dan Prapemrosesan Data
Data penelitian berupa halaman web defacement dan halaman web normal yang dikumpulkan secara mandiri. Data kemudian dilabeli dan dipraproses melalui proses pembersihan data, seperti penghapusan tag HTML yang tidak relevan, karakter khusus, dan elemen halaman yang tidak berkaitan dengan konten, sehingga dataset siap digunakan dalam proses pembelajaran mesin.
Luaran: Dataset terlabel *web defacement* judi *online* dan *non-defacement*.
Indikator capaian: Dataset siap latih dan uji
Penanggung jawab: Mahasiswa

Tahap 3 – Ekstraksi Fitur dan Pembuatan Model
Pada tahap ini, konten halaman web direpresentasikan dalam bentuk fitur teks menggunakan pendekatan *term weighting (TF-IDF)*. Selanjutnya, model klasifikasi dibangun menggunakan algoritma *Logistic Regression* untuk membedakan perubahan file normal dan serangan web defacement judi online.
Luaran: Model klasifikasi berbasis *Logistic Regression*
Indikator capaian: Model mampu melakukan klasifikasi dengan akurasi yang terukur
Penanggung jawab: Mahasiswa

Tahap 4 - Integrasi Model ke dalam HIDS
Model klasifikasi yang telah dibangun diintegrasikan ke dalam alur kerja HIDS pada tahap *File Integrity Monitoring*. Setiap perubahan file yang terdeteksi oleh FIM akan dievaluasi oleh model sebelum sistem memicu *active response*.
Luaran: Prototipe HIDS teroptimasi
Indikator capaian: Sistem mampu melakukan deteksi berbasis *signature* dan evaluasi konteks
Penanggung jawab: Mahasiswa dan Dosen Pembimbing

Tahap 5 - Pengujian dan Evaluasi Sistem
Pengujian dilakukan dengan membandingkan kinerja HIDS berbasis *signature* murni dan HIDS yang telah dioptimasi. Parameter evaluasi meliputi ketepatan deteksi, tingkat *false positive*, waktu deteksi, penggunaan sumber daya sistem, dan ketepatan *active response*.
Luaran: Hasil evaluasi kinerja system
Indikator capaian: Terjadi peningkatan performa deteksi setelah optimasi
Penanggung jawab: Mahasiswa dan Dosen Pembimbing

2. Diagram Alir Penelitian
Diagram alir penelitian terdiri atas alur berikut



3. Keterkaitan Metode dengan RAB

Metode penelitian ini disusun sejalan dengan Rencana Anggaran Biaya (RAB), di mana:

- Tahap pengumpulan data dan eksperimen mendukung kebutuhan biaya komputasi dan perangkat lunak.
- Tahap pengujian dan evaluasi mendukung penggunaan sumber daya server dan analisis hasil.
- Tahap pelaporan dan publikasi mendukung penyusunan tesis dan luaran ilmiah.

Dengan demikian, setiap tahapan penelitian memiliki keterkaitan langsung dengan komponen pembiayaan yang diusulkan.]

F. HASIL YANG DIHARAPKAN

Jelaskan hasil yang diharapkan atau luaran yang dijanjikan dari penelitian

[Penelitian ini diharapkan menghasilkan prototipe *Host-Based Intrusion Detection System* (HIDS) yang dioptimasi menggunakan algoritma *Logistic Regression* untuk mendeteksi serangan web defacement bermuatan judi *online* secara lebih akurat. Sistem yang dihasilkan tidak hanya mampu mendeteksi perubahan file melalui mekanisme *File Integrity Monitoring* (FIM), tetapi juga mengevaluasi konteks perubahan konten sehingga dapat menurunkan tingkat *false positive* dan meningkatkan ketepatan *active response*.

Hasil penelitian juga diharapkan memberikan kontribusi ilmiah berupa rekomendasi metode optimasi deteksi berbasis signature pada HIDS serta luaran akademik berupa tesis magister dan publikasi ilmiah sesuai ketentuan skema PTM.]

G. JADWAL PENELITIAN

Jadwal penelitian disusun berdasarkan pelaksanaan penelitian, harap disesuaikan berdasarkan lama tahun pelaksanaan penelitian

[Tahun ke-1

No	Nama Kegiatan	Bulan											
		1	2	3	4	5	6	7	8	9	10	11	12
1	Studi literatur & analisis kebutuhan	X	X										
2	Pengumpulan & pelabelan data		X	X									
3.	Prapemrosesan data & ekstraksi fitur (TF-IDF)			X	X								

4.	Pelatihan model logistic regression				X	X							
5.	Integrasi Model ke FIM pada HIDS					X	X						
6.	Deteksi perubahan file & evaluasi konteks						X	X					
7.	Evaluasi & analisis hasil							X	X				
8.	Penyusunan tesis magister								X	X	X		
9.	Penyusunan artikel ilmiah dan laporan akhir										X	X	X

]

H. DAFTAR PUSTAKA

Sitasi disusun dan ditulis berdasarkan sistem nomor sesuai dengan urutan pengutipan. Hanya pustaka yang disitasi pada usulan penelitian yang dicantumkan dalam Daftar Pustaka.

[

1. Holt TJ, Lee JR, Freilich JD, Chermak SM, Bauer JM, Shillair R, et al. An Exploratory Analysis of the Characteristics of Ideologically Motivated Cyberattacks. Terrorism and Political Violence. 2022 Oct 3;34(7):1305–20.

2. Direktorat Operasi Keamanan Siber BSSN. Lanskap Keamanan Siber Indonesia 2024. Jakarta; 2024.

3. Erawan E, Satria RP, Nastiti AD, Juniardi W, Anbiya DR, Go Reinnamah DA. A Lightweight Design Approach to Detect Malicious Injections in Government Websites: Combating Hidden Online Gambling Pages. In: 2024 Beyond Technology Summit on Informatics International Conference (BTS-I2C). IEEE; 2024. p. 365–9.

4. Hoang XD, Nguyen NT. Detecting website defacements based on machine learning techniques and attack signatures. Computers. 2019 Jun 1;8(2).

5. Uddin I. DETECTION APPROACHES FOR WEB DEFACEMENT ATTACKS. GSJ [Internet]. 2025;13(6). Available from: www.globalscientificjournal.com

6. Vasudev Karthik Ravindran, Sharad Shyam Ojha, Arvind Kamboj. A Comparative Analysis of Signature-Based and Anomaly-Based Intrusion Detection Systems. International Journal of Latest Technology in Engineering Management & Applied Science. 2025 Jun 3;14(5):209–14.

7. Zagi LM, Digdo GP, Shalannanda W. Just Dork and Crawl: Measuring Illegal Online Gambling Defacement in Indonesian Websites. 2025 Aug 26; Available from: http://arxiv.org/abs/2508.19368

8. Kholwal R. TEXT-CLASSIFY: A COMPREHENSIVE COMPARATIVE STUDY OF LOGISTIC REGRESSION, RANDOM FOREST, AND KNN MODELS FOR ENHANCED TEXT CLASSIFICATION PERFORMANCE. Int J Adv Eng Technol. 2023 Oct;16:415–33.

9. Chen Y, Liu P, Teo CP. Regularised Text Logistic Regression: Key Word Detection and Sentiment Classification for Online Reviews. 2020 Sep 9; Available from: http://arxiv.org/abs/2009.04591