BAB I
PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi kecerdasan buatan (*Artificial Intelligence* / AI) telah menjadi pendorong utama transformasi digital di berbagai sektor, termasuk industri, pemerintahan, dan layanan berbasis teknologi informasi. Pemanfaatan AI dalam otomatisasi proses, analisis data, serta peningkatan efisiensi sistem telah memberikan banyak manfaat dalam mendukung aktivitas manusia. Namun, di sisi lain, adopsi teknologi AI yang semakin luas juga memunculkan tantangan baru, khususnya dalam aspek keamanan digital.

Seiring dengan kemajuan tersebut, ancaman siber tidak lagi terbatas pada metode konvensional, melainkan telah berevolusi menjadi serangan yang lebih cerdas, adaptif, dan sulit dideteksi melalui pemanfaatan teknologi AI. Berbagai penelitian menunjukkan bahwa AI digunakan untuk mengotomatisasi serangan siber dengan skala yang lebih besar serta tingkat personalisasi yang tinggi, seperti pada serangan *phishing* berbasis *Natural Language Processing* dan manipulasi identitas digital menggunakan teknologi *deepfake* [1], [2]. Kondisi ini menyebabkan serangan siber menjadi semakin meyakinkan dan berpotensi menipu pengguna maupun sistem keamanan yang ada.

Teknologi *deepfake* yang memanfaatkan *deep learning*, khususnya *Generative adversarial networks* (GAN), mampu menghasilkan konten video dan audio palsu yang sangat *real*istis sehingga sulit dibedakan dari konten asli. Penyalahgunaan teknologi ini telah digunakan dalam berbagai bentuk penipuan, disinformasi, dan manipulasi identitas, yang secara langsung mengancam integritas informasi digital [2]. Di sisi lain, serangan *phishing* berbasis AI mampu meniru gaya bahasa dan pola komunikasi individu secara presisi, sehingga meningkatkan tingkat keberhasilan serangan dibandingkan metode *phishing* konvensional [3].

Berdasarkan kajian literatur dan pengamatan awal terhadap perkembangan keamanan siber, terlihat bahwa sistem keamanan tradisional yang masih mengandalkan pendekatan berbasis tanda tangan (*Signature-based*) dan aturan statis sering kali tidak mampu mengimbangi kecepatan evolusi ancaman siber

berbasis AI. Serangan yang bersifat adaptif dan dinamis ini berpotensi menimbulkan dampak yang signifikan terhadap tiga pilar utama keamanan informasi, yaitu kerahasiaan (*Confidentiality*), integritas (*Integrity*), dan ketersediaan (*Availability*) data [4].

Oleh karena itu, diperlukan suatu analisis yang komprehensif untuk memahami bentuk-bentuk penyalahgunaan AI dalam serangan siber serta dampaknya terhadap keamanan digital. Penelitian ini menjadi relevan karena berupaya menganalisis ancaman keamanan digital akibat penyalahgunaan AI secara sistematis, didukung oleh studi literatur, survei terhadap praktisi keamanan siber, dan eksperimen teknis terbatas. Hasil penelitian ini diharapkan dapat memberikan wawasan yang mendalam serta menjadi dasar dalam perumusan strategi mitigasi yang efektif guna menghadapi tantangan keamanan siber di era digital.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang masalah, maka perumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana bentuk-bentuk penyalahgunaan kecerdasan buatan dalam serangan siber memengaruhi keamanan digital?
2. Seberapa besar dampak yang ditimbulkan oleh serangan siber berbasis AI terhadap kerahasiaan, integritas, dan ketersediaan data?
3. Apa saja strategi mitigasi yang efektif untuk menghadapi ancaman keamanan digital akibat penyalahgunaan AI?

1.3 Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Penelitian ini merupakan studi literatur yang diperkuat dengan survei kualitatif dan eksperimen teknis terbatas, mengandalkan data sekunder dari publikasi ilmiah, artikel, dan laporan keamanan siber, serta data primer dari kuesioner dan hasil eksperimen.
2. Lingkup pembahasan difokuskan pada analisis ancaman siber yang menggunakan AI, dengan eksperimen teknis terbatas pada simulasi serangan *phishing* berbasis AI dan deteksi *deepfake* menggunakan tools yang tersedia

secara *open-source*.

3. Eksperimen dilakukan dalam lingkungan terkontrol (*controlled environment*) menggunakan *dataset* publik dan tidak melibatkan serangan terhadap sistem nyata atau data sensitif untuk mematuhi etika penelitian.
4. Contoh kasus yang dianalisis dibatasi pada kasus yang telah dipublikasikan dan relevan dengan topik, seperti *deepfake*, *phishing* otomatis, dan *Malware* adaptif.
5. Responden kuesioner adalah praktisi keamanan siber, profesional IT, dan akademisi di bidang keamanan informasi dengan minimal 1 tahun pengalaman.
6. Eksperimen teknis dibatasi pada *proof-of-concept* untuk memvalidasi temuan literatur dan tidak bertujuan untuk mengembangkan sistem keamanan lengkap.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Menganalisis berbagai bentuk penyalahgunaan kecerdasan buatan (AI) yang digunakan dalam serangan siber, seperti *deepfake*, *phishing* otomatis, dan *Malware* adaptif.
2. Mengidentifikasi dan menganalisis dampak serta potensi ancaman yang ditimbulkan oleh serangan siber berbasis AI terhadap keamanan digital, khususnya terkait dengan kerahasiaan (*Confidentiality*), integritas (*Integrity*), dan ketersediaan (*Availability*) data.
3. Melakukan eksperimen teknis untuk memvalidasi temuan literatur mengenai efektivitas serangan berbasis AI dan kemampuan deteksi sistem keamanan yang ada.
4. Memvalidasi temuan literatur melalui survei terhadap praktisi keamanan siber mengenai pengalaman dan persepsi mereka terhadap ancaman berbasis AI.
5. Memberikan rekomendasi strategi mitigasi dan pencegahan yang dapat diterapkan oleh individu dan organisasi untuk mengurangi risiko yang muncul dari ancaman siber berbasis AI, berdasarkan hasil analisis literatur, survei, dan eksperimen.

1.5 Manfaat Penelitian

Hasil dari penelitian ini diharapkan dapat memberikan manfaat baik secara teoritis

maupun praktis:

1. Manfaat Teoritis
    a. Penelitian ini diharapkan dapat memperkaya literatur dan menjadi tambahan referensi ilmiah di bidang keamanan digital, khususnya terkait dengan ancaman siber yang menggunakan teknologi kecerdasan buatan.
    b. Penelitian ini dapat menjadi dasar bagi penelitian lanjutan yang lebih mendalam, baik dalam bentuk studi kasus maupun pengembangan model pertahanan siber berbasis kecerdasan buatan.
    c. Kontribusi metodologis melalui kombinasi studi literatur dan survei kualitatif untuk analisis ancaman siber berbasis AI.
2. Manfaat Praktis
    a. Bagi individu, penelitian ini dapat meningkatkan kesadaran akan berbagai bentuk serangan siber berbasis AI dan memberikan wawasan tentang cara-cara praktis untuk melindungi data pribadi mereka.
    b. Bagi organisasi, temuan dari penelitian ini dapat menjadi bahan pertimbangan dalam merancang dan memperkuat strategi keamanan siber, serta membantu dalam mengidentifikasi area-area yang rentan terhadap serangan AI.
    c. Bagi pengambil kebijakan, penelitian ini dapat memberikan data dan analisis yang relevan sebagai dasar untuk perumusan regulasi dan kebijakan yang lebih responsif terhadap dinamika ancaman siber di era digital.
    d. Bagi praktisi keamanan siber, penelitian ini memberikan panduan praktis dalam mengidentifikasi dan merespons ancaman berbasis AI.

1.6 Sistematika Penulisan

Penelitian ini akan disusun dalam lima bab dengan sistematika sebagai berikut:

BAB I PENDAHULUAN, berisi latar belakang masalah yang menguraikan fenomena dan urgensi penelitian, rumusan masalah yang menjadi pertanyaan utama penelitian, batasan masalah untuk memfokuskan ruang lingkup, tujuan dan manfaat

penelitian, serta sistematika penulisan secara keseluruhan.

BAB II TINJAUAN PUSTAKA, berisi landasan teori dan kajian literatur yang relevan dengan topik penelitian. Pembahasan akan mencakup konsep dasar keamanan digital, definisi dan prinsip *Artificial Intelligence* (AI), serta tinjauan terhadap penelitian terdahulu yang berhubungan dengan penyalahgunaan AI dalam serangan siber.

BAB III METODE PENELITIAN, menjelaskan jenis penelitian yang digunakan, yaitu studi literatur yang diperkuat dengan survei kualitatif. Selain itu, bab ini akan memaparkan sumber data yang digunakan (data sekunder dan primer), teknik pengumpulan data melalui studi literatur dan kuesioner, serta metode analisis data kualitatif yang akan diterapkan untuk memecahkan rumusan masalah.

BAB IV HASIL DAN PEMBAHASAN, menyajikan hasil analisis dari penelitian yang telah dilakukan. Pembahasan akan mencakup identifikasi bentuk-bentuk penyalahgunaan AI dalam serangan siber, analisis ancaman yang ditimbulkannya, hasil survei dari praktisi keamanan siber, serta perumusan strategi mitigasi dan pencegahan yang relevan.

BAB V PENUTUP, berisi kesimpulan dari seluruh rangkaian penelitian yang merangkum jawaban atas rumusan masalah, serta saran-saran praktis dan teoritis yang dapat menjadi masukan bagi berbagai pihak terkait, termasuk pemerintah, organisasi, dan masyarakat.