ABSTRAK

ANALISIS RISIKO *DEEPFAKE ARTIFICIAL INTELLIGENCE* TERHADAP KEMAMPUAN MASYARAKAT MEMBEDAKAN KONTEN NYATA DAN BUATAN *AI* DENGAN METODE DESKRIPTIF KUANTITATIF

Nama : Muhammad Rizki Akbar

NPM : 2255201140

Pembimbing : Pahrizal, S. Kom, M. Kom.

Perkembangan teknologi *Artificial Intelligence* (*AI*) melahirkan teknologi *Deepfake* yang mampu menghasilkan konten visual menyerupai kenyataan, sehingga menimbulkan risiko disinformasi dan menyulitkan masyarakat membedakan konten nyata dan buatan *AI*. Penelitian ini bertujuan menganalisis tingkat kemampuan masyarakat membedakan konten tersebut serta hubungan antara tingkat pengetahuan masyarakat tentang *Deepfake AI* (X1) dengan kemampuan mendeteksi konten *Deepfake* (Y1). Penelitian menggunakan metode deskriptif kuantitatif dengan 100 responden pengguna media sosial usia 17-30 tahun, dikumpulkan daring melalui *Google Forms* pada 24 Juni-24 Juli 2026. Instrumen terdiri atas lima pernyataan Likert untuk mengukur X1 dan sepuluh pasangan foto asli-*Deepfake* untuk mengukur Y1. Hasil menunjukkan kemampuan deteksi tinggi (rata-rata 77,1%) dan pengetahuan tinggi (rata-rata 3,87 skala 1-5). Kedua instrumen valid dan reliabel (Cronbach Alpha X1=0,834; KR-20 Y1=0,712). Uji korelasi Pearson menunjukkan tidak terdapat hubungan signifikan antara X1 dan Y1 ($r=-0,028$; $p=0,784>0,05$), dikonfirmasi regresi linear sederhana (R Square=0,001), sehingga H0 diterima; pengetahuan konseptual terbukti tidak otomatis meningkatkan kemampuan deteksi. Penelitian merekomendasikan peningkatan literasi digital, regulasi watermarking konten *AI*, dan penyempurnaan instrumen deteksi pada penelitian selanjutnya.

Kata Kunci: Deepfake, Artificial Intelligence, Kemampuan Deteksi, Pengetahuan, Deskriptif Kuantitatif

ABSTRACT

Analysis of Artificial Intelligence Deepfake Risks Regarding the Public's Ability to Distinguish Between Real and AI-Generated Content Using a Quantitative Descriptive Method

Name : Muhammad Rizki Akbar

Student Id : 2255201140

Supervisor : Pahrizal,S. Kom.,M.Kom.

The development of Artificial Intelligence (AI) technology has given rise to Deepfake technology capable of producing visual content resembling reality, creating a risk of disinformation and making it difficult for the public to distinguish real from AI-generated content. This study aims to analyze the public's ability to distinguish such content and the relationship between the public's knowledge of Deepfake AI (X1) and their ability to detect Deepfake content (Y1). This study uses a quantitative descriptive method with 100 social media user respondents aged 17-30 years, collected online via Google Forms from June 24 to July 24, 2026. The instrument consisted of five Likert-scale statements measuring X1 and ten pairs of real-Deepfake photos measuring Y1. Results show that detection ability is high (average 77.1%) and knowledge level is also high (average 3.87 on a 1-5 scale). Both instruments were valid and reliable (Cronbach's Alpha for X1 = 0.834; KR-20 for Y1 = 0.712). A Pearson correlation test showed no significant relationship between X1 and Y1 ($r = -0.028$; $p = 0.784 > 0.05$), confirmed by simple linear regression (R Square = 0.001), so the null hypothesis (H0) is accepted; conceptual knowledge was shown not to automatically improve actual detection ability. This study recommends improving digital literacy, AI content watermarking regulations, and refining the detection instrument in future research.

Keywords: Deepfake, Artificial Intelligence, Detection Ability, Knowledge, Quantitative Descriptive

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi kecerdasan buatan atau Artificial Intelligence (AI) telah membawa perubahan besar dalam berbagai aspek kehidupan manusia, terutama pada bidang komunikasi digital dan media informasi. Salah satu perkembangan AI yang saat ini banyak dibahas adalah teknologi Deepfake. Teknologi Deepfake merupakan teknologi berbasis Deep learning yang mampu memanipulasi gambar, audio, maupun video sehingga terlihat seperti asli. Teknologi ini dapat digunakan untuk membuat konten palsu yang sangat realistis dengan cara mengganti wajah, suara, atau gerakan seseorang menggunakan algoritma AI (Payne, 2026).

Kemajuan yang sangat pesat dalam bidang kecerdasan buatan (Artificial Intelligence/AI) saat ini telah berhasil menghadirkan suatu teknologi digital yang semakin canggih dan sulit untuk dibedakan dari konten-konten asli (Raharjo, 2023). Teknologi digital ini kemudian dikenal dengan istilah Deepfake. Perkembangan teknologi AI yang kian maju ini telah memicu munculnya ancaman-ancaman baru yang sangat serius bagi keamanan digital secara global, khususnya dalam konteks permasalahan penipuan identitas digital. Keberadaan Deepfake yang semakin sulit untuk dideteksi dan dibedakan dari konten asli telah menjadi salah satu tantangan utama yang harus dihadapi dalam upaya menjaga keamanan serta integritas identitas

digital di tengah era teknologi yang terus berkembang pesat saat ini. penyalah gunaan teknologi Deepfake ini telah menimbulkan berbagai macam tantangan hukum dan etika yang cukup serius. Sehingga memunculkan banyak pertanyaan mengenai batasan-batasan hukum yang harus diterapkan dalam menangani kasus-kasus penipuan identitas digital yang disebabkan oleh kehadiran teknologi Deepfake. (Puspita, L., & Firdaus, H. S. 2025).

Dalam konteks media sosial ini, berita palsu semakin meluas karena kemajuan teknologi informasi yang memungkinkan siapa pun untuk dengan mudah berbagi informasi. Media sosial menyediakan platform terbuka untuk pertukaran informasi, tetapi tidak ada pengawasan atau penyaringan yang ketat. Oleh karena itu, media sosial telah menjadi saluran utama penyebaran berita palsu di dalam komunitas digital. (Sabela, I. N. D., Saputra, A., & Kuncoro, W.,2026).

Menurut konsep dasarnya, Deepfake merupakan media sintetis yang dihasilkan menggunakan teknologi Artificial Intelligence (AI) untuk menampilkan sesuatu yang sebenarnya tidak pernah terjadi. Teknologi ini umumnya dibuat menggunakan metode Generative Adversarial Network (GAN), yaitu dua model AI yang saling bekerja untuk menghasilkan konten palsu yang semakin realistis dan sulit dibedakan dari konten asli. Perkembangan teknologi Deepfake membuat manipulasi gambar, audio, dan video menjadi semakin canggih sehingga berpotensi menimbulkan penyebaran disinformasi dan hoaks di media digital (Altuncu, Franqueira, dan Li, 2024).

AI juga merupakan bidang ilmu komputer yang bertujuan untuk menciptakan mesin atau program komputer yang dapat melakukan tugas yang memerlukan kecerdasan manusia. (Juditha, C. .,2025).

Teknologi *Deepfake* memungkinkan manipulasi konten visual dan audio untuk menciptakan media yang sangat realistis namun palsu, seperti video atau gambar yang menampilkan seseorang melakukan atau mengatakan sesuatu yang sebenarnya tidak pernah terjadi. Teknologi ini memanfaatkan algoritma *AI* untuk menggantikan wajah atau suara seseorang dengan orang lain, menghasilkan konten yang sulit dibedakan dari yang asli oleh pengamat manusia. Meskipun awalnya digunakan dalam konteks hiburan dan seni, *Deepfake* kini menjadi perhatian serius karena potensi penyalah gunaannya dalam menyebarkan informasi palsu dan propaganda (Darmawan, Junaidi, dan Khaerudin, 2025).

Deepfake memanfaatkan *Generative Adversarial Network (GAN)*, yang bekerja dengan jaringan saraf untuk menganalisis sejumlah besar data guna meniru ekspresi wajah, perilaku, suara, manusia. Istilah *Deepfake* berasal dari gabungan kata *Deep learning*, yang merujuk pada teknologi pembelajaran mesin yang mendalam, dan *fake*, yang berarti palsu. Kemunculan *Deepfake* pertama kali digunakan untuk memberikan kemudahan pada manusia dalam menciptakan hasil foto atau video yang lebih nyata dalam konten hiburan seperti film ataupun game. Namun seiring berjalannya waktu, teknologi ini mulai disalah gunakan untuk berbagai kepentingan negatif (Seveney, Wicaksono, dan Soetijono, 2025).

Teknologi ini dapat digunakan untuk membuat video palsu yang memperlihatkan tokoh mengatakan sesuatu yang sebenarnya tidak pernah mereka ucapkan. Pengguna dapat dengan mudah menciptakan video palsu yang membuat tokoh terkenal terlihat seperti berkata atau melakukan sesuatu yang sebenarnya tidak pernah terjadi. Selain itu, teknologi ini juga bisa dimanfaatkan untuk menyebarkan disinformasi di media sosial seperti Facebook, Instagram, dan WhatsApp, serta digunakan oleh oknum-oknum untuk mempengaruhi atau memanipulasi konten yang tersebar secara meluas sehingga dapat mengakibatkan ancaman kepada masyarakat, mulai dari tindakan kriminalitas seperti penipuan, pencemaran nama baik, menimbulkan perpecahan, dan lain-lain (Fadhilah dan Retnoningsih, 2024).

Berdasarkan uraian di atas, sebagian besar penelitian terdahulu lebih banyak menyoroti aspek hukum, regulasi, maupun metode deteksi *Deepfake* berbasis *machine learning*, sementara kajian mengenai

hubungan antara tingkat pengetahuan masyarakat tentang teknologi *Deepfake AI* dengan kemampuan mereka secara langsung membedakan konten nyata dan konten buatan *AI* masih terbatas, khususnya di Indonesia. Kesenjangan inilah yang mendasari penelitian ini untuk memfokuskan kajian pada dua variabel utama, yaitu tingkat pengetahuan masyarakat tentang *Deepfake AI* dan kemampuan mereka dalam mendeteksi konten *Deepfake*, guna memberikan gambaran empiris mengenai sejauh mana pengetahuan berperan terhadap kemampuan deteksi tersebut.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana tingkat kemampuan masyarakat dalam membedakan konten nyata dan konten hasil *Deepfake AI*?
2. Bagaimana hubungan antara tingkat pengetahuan masyarakat tentang teknologi *Deepfake AI* dengan kemampuan mereka dalam mendeteksi konten *Deepfake AI*?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah di atas, tujuan penelitian ini adalah sebagai berikut:

1. Menganalisis tingkat kemampuan masyarakat dalam membedakan konten nyata dan konten hasil *Deepfake AI.*
2. Menganalisis hubungan antara tingkat pengetahuan masyarakat tentang teknologi *Deepfake AI* dengan kemampuan mereka dalam mendeteksi konten *Deepfake AI.*

1.4 Manfaat Penelitian

1. Manfaat Teoritis Penelitian ini diharapkan dapat memberikan kontribusi terhadap pengembangan kajian keilmuan di bidang teknik informatika, khususnya terkait literasi digital masyarakat dan risiko penyalahgunaan teknologi kecerdasan buatan berbasis *Deepfake.*

2. Manfaat Praktis

   a. Bagi masyarakat, hasil penelitian ini diharapkan dapat meningkatkan kewaspadaan dan kemampuan dalam mengenali konten *Deepfake* di media digital.

   b. Bagi pemerintah dan pembuat kebijakan, hasil penelitian ini dapat menjadi bahan pertimbangan dalam merumuskan kebijakan penanganan risiko *Deepfake* di Indonesia.

   c. Bagi akademisi dan peneliti selanjutnya, hasil penelitian ini dapat menjadi referensi untuk penelitian lanjutan terkait deteksi dan mitigasi risiko *Deepfake.*

1.5 Batasan Masalah

Agar penelitian ini lebih terarah dan fokus, maka ditetapkan batasan masalah sebagai berikut:

a. Penelitian ini menggunakan metode deskriptif kuantitatif dengan instrumen kuesioner daring melalui *Google Forms.*

b. Responden penelitian adalah masyarakat pengguna media sosial yang berada pada rentang usia 17–30 tahun.

c. Materi stimulus yang digunakan dalam kuesioner berupa sepuluh foto yang terdiri atas konten asli dan konten hasil *Deepfake.*

d. Penelitian ini tidak membangun sistem atau model deteksi *Deepfake* berbasis *machine learning*, melainkan berfokus pada analisis kemampuan persepsi dan literasi digital masyarakat.

e. Teknik pengambilan sampel menggunakan purposive sampling dengan mempertimbangkan keterwakilan demografis responden.