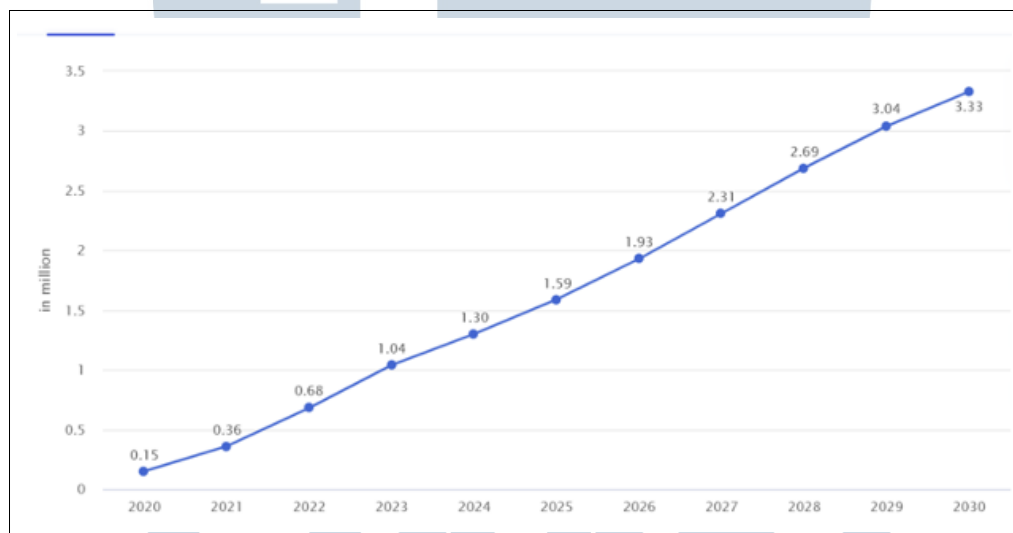


BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Dalam beberapa tahun ke belakang perkembangan *artificial intelligence* (AI) mengalami peningkatan yang cukup signifikan, khususnya semenjak munculnya *Large Language Model System* dan teknologi *generative AI*, seperti ChatGPT atau DALL-E [1]. Grafik peningkatan pengguna AI ditunjukkan dalam Gambar 1.1.



Gambar 1.1. Jumlah pengguna *artificial intelligence* di Indonesia
Sumber: [2]

Penelitian mengenai manfaat penggunaan AI dalam pendidikan Indonesia menunjukkan bahwa AI dapat memberikan kontribusi positif dalam personalisasi pembelajaran, peningkatan efisiensi, eksplorasi ide, dan peningkatan kreativitas mahasiswa dalam pemecahan masalah [3]. Akan tetapi, tingkat penerimaan mahasiswa terhadap penggunaan AI tidak bersifat homogen. Sebagian mahasiswa menunjukkan antusiasme tinggi terhadap penggunaan AI, sementara sebagian lainnya masih ragu karena faktor kepercayaan, manfaat, dan kemudahan penggunaan [4]. Di sisi lain, muncul kekhawatiran terhadap ketergantungan mahasiswa dalam penggunaan AI yang dapat mempengaruhi kompetensi serta perilaku dalam pembelajaran mahasiswa, karena dapat menghilangkan kemampuan berpikir kreatif dan kritis [5].

Penggunaan *Technology Acceptance Model* (TAM) merupakan model teoritis yang cocok digunakan untuk memahami faktor-faktor yang mempengaruhi penerimaan teknologi. TAM secara orisinal dikemukakan oleh Fred D. Davis pada tahun 1989, terdiri dari *perceived usefulness* dan *perceived ease of use* sebagai konstruk utama, Davis juga menjelaskan bahwa *perceived usefulness* dan *perceived ease of use* mempengaruhi penerimaan teknologi yang kemudian berdampak pada sikap serta minat perilaku dalam penggunaan teknologi [6]. Namun, dalam konteks *artificial intelligence* yang sensitif terhadap isu keamanan dan privasi data, penelitian terkini menunjukkan penambahan konstruk *perceived trust* dan *attitude toward using* dalam meningkatkan eksplanatori model [7]. *Attitude toward using* menggambarkan sikap pengguna terhadap teknologi secara keseluruhan, sedangkan *perceived trust* menggambarkan tingkat kepercayaan pengguna terhadap keamanan data dan keandalan teknologi, khususnya pada teknologi kecerdasan buatan yang melibatkan privasi data pengguna. Dengan demikian, penelitian ini menggunakan *extended TAM* dengan empat konstruk sebagai *predictor* untuk melihat minat mahasiswa terhadap penggunaan teknologi *artificial intelligence* dalam pembelajaran, sehingga model ini lebih komprehensif dibandingkan TAM original dan berpotensi memberikan prediksi yang lebih akurat.

Meskipun banyak penelitian menerapkan TAM dalam analisis korelasional, penelitian yang mengintegrasikan TAM dengan pendekatan *predictive modeling* masih terbatas. Penelitian ini mengisi *gap* tersebut dengan mengintegrasikan *logistic regression* sebagai metode utama karena interpretabilitasnya yang tetap mempertahankan fokus TAM, serta menggunakan *random forest* sebagai perbandingan performa. Dengan demikian, penelitian ini mengintegrasikan aspek akurasi prediktif model (evaluasi performa) dan pemahaman faktor-faktor TAM yang mempengaruhi minat (interpretasi).

Untuk menjawab *gap* tersebut, penelitian ini menerapkan pendekatan *predictive modeling* melalui tugas *binary classification*, yaitu memprediksi kecenderungan *behavioral intention* mahasiswa terhadap penggunaan *artificial intelligence* ke dalam dua kategori, yaitu tinggi dan rendah. Variabel-variabel yang berasal dari konstruk TAM, yaitu *perceived usefulness*, *perceived ease of use*, *attitude toward using*, serta *perceived trust*, digunakan sebagai fitur (*features*) untuk membangun model prediksi. Dalam konteks ini, *predictive modeling* merupakan pendekatan umum untuk membangun model prediksi, sedangkan *binary classification* merupakan bentuk implementasi *machine learning* yang digunakan dalam penelitian ini. Oleh karena itu, algoritma *logistic regression*

dan *random forest* dipilih karena keduanya merupakan algoritma klasifikasi yang sesuai untuk memodelkan hubungan antara fitur-fitur TAM dengan kategori *behavioral intention*. Selain menghasilkan prediksi, kedua algoritma tersebut juga memungkinkan evaluasi performa model menggunakan metrik klasifikasi seperti *accuracy*, *precision*, *recall*, *F1-score*, dan *ROC-AUC*.

Logistic regression adalah metode klasifikasi yang banyak digunakan untuk memodelkan permasalahan biner [8]. Kemampuannya dalam memodelkan probabilitas suatu kejadian berdasarkan variabel independen membuatnya sangat sesuai. Metode ini memiliki interpretabilitas yang baik, di mana koefisien model dapat diinterpretasikan dalam bentuk *odds ratio* sehingga memungkinkan identifikasi faktor-faktor yang paling berpengaruh terhadap minat penggunaan dalam kerangka *Technology Acceptance Model*. Selain itu, dengan jumlah sampel sebanyak 535 responden dan jumlah variabel independen yang relatif terbatas, penggunaan *machine learning* konvensional dinilai sudah efisien dan memadai, sejalan dengan *rule of thumb* yang merekomendasikan minimal 10 observasi untuk setiap variabel independen [9]. Data yang digunakan berupa kuesioner dengan skala *Likert* yang bersifat terstruktur, sehingga lebih sesuai dianalisis menggunakan model yang tidak memerlukan kompleksitas tinggi.

Di samping itu, *random forest* digunakan sebagai model pembanding karena kemampuannya dalam mengungkap hubungan yang tidak berbentuk linier antar variabel serta menghasilkan performa yang lebih *robust* melalui pendekatan *ensemble learning*. Penggunaan kedua model ini memungkinkan analisis perbandingan antara model yang bersifat *interpretable* dan model yang berorientasi pada performa prediktif. Dengan hal tersebut diharapkan dapat memberikan pemahaman yang lebih menyeluruh terhadap hasil temuan penelitian.

Berbeda dengan metode konvensional yang telah disebutkan, metode berbasis *deep learning* meskipun menunjukkan performa yang tinggi dalam permasalahan klasifikasi, penggunaannya memiliki beberapa keterbatasan, terutama terkait kebutuhan data bervolume besar untuk mencapai performa optimal karena kompleksitas parameter yang tinggi. Penelitian [10] menunjukkan bahwa model *deep learning* cenderung membutuhkan data yang besar untuk mencapai performa optimal karena kompleksitas parameter yang tinggi. Dalam konteks *machine learning* konvensional, dataset dengan 535 sampel masih tergolong memadai untuk metode seperti *logistic regression* dan *random forest*. Namun, untuk konteks *deep learning*, jumlah tersebut termasuk dalam kategori dataset kecil yang berpotensi menyebabkan *overfitting* dan penurunan performa model. Di sisi lain, model *deep*

learning memiliki sifat *black-box* yang memiliki keterbatasan untuk menjelaskan bagaimana hubungan antara variabel *input* dan *output* terbentuk dalam proses prediksi, sehingga mengurangi tingkat transparansi dan kepercayaan terhadap model.

Berdasarkan uraian tersebut, terdapat celah antara kajian penerimaan teknologi berbasis *Technology Acceptance Model* (TAM), yang umumnya bersifat analisis korelasional dengan pendekatan prediktif berbasis *machine learning*, yang berfokus pada kemampuan klasifikasi. Selain celah tersebut, terdapat pula celah metodologis dari perspektif *machine learning*. Meskipun berbagai algoritma klasifikasi telah banyak digunakan dalam penelitian, masih terbatas studi komprehensif yang membandingkan model yang bersifat dapat diinterpretasikan dengan model yang bersifat *ensemble non-linear* khususnya dalam konteks prediksi penerimaan teknologi. Perbandingan ini penting karena kedua pendekatan memiliki karakteristik yang berbeda.

Oleh karena itu, analisis terhadap *trade-off* antara interpretabilitas model dan kinerja prediktif menjadi bagian penting dari penelitian ini. Dengan pendekatan ini penelitian diharapkan dapat menghasilkan model prediksi yang memiliki tingkat keandalan tinggi dan mudah diinterpretasikan, serta memberikan kontribusi dalam pendekatan berbasis data untuk analisis penerimaan teknologi di bidang pendidikan.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijabarkan, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana performa model *logistic regression* dalam memprediksi niat mahasiswa terhadap penggunaan *artificial intelligence* dalam kegiatan akademik, berdasarkan variabel kerangka *Technology Acceptance Model*?
2. Variabel *Technology Acceptance Model* mana yang paling berpengaruh terhadap niat mahasiswa dalam menggunakan *artificial intelligence*, ditinjau dari perspektif inferensial melalui interpretasi koefisien pada *logistic regression*, serta dari perspektif prediksi melalui *feature importance* pada *random forest*?
3. Bagaimana perbandingan performa *logistic regression* dan *random forest* saat memprediksi niat mahasiswa, serta bagaimana *trade-off* antara interpretabilitas dan akurasi prediktif dari kedua metode tersebut?

1.3 Batasan Masalah

Agar penelitian tetap terarah dan sesuai dengan ruang lingkup yang telah ditentukan, batasan masalah penelitian ini adalah sebagai berikut.

1. Dataset yang digunakan merupakan dataset publik yang bersumber dari Mendeley Data [11] sehingga memungkinkan adanya data tidak lengkap (*missing values*) atau keterbatasan dalam kualitas pengukuran.
2. Penelitian ini menggunakan data *cross-sectional*, sehingga tidak dapat menangkap perubahan minat mahasiswa terhadap penggunaan *artificial intelligence* dari waktu ke waktu.
3. Data yang digunakan bersumber dari *self-report survey*, sehingga memiliki potensi terjadinya *social desirability* dan *response bias* dalam pengisian kuesioner oleh responden.
4. Sampel penelitian terbatas pada mahasiswa di Indonesia yang terdapat dalam dataset, sehingga hasil penelitian memiliki keterbatasan dalam segi generalisasi terhadap seluruh populasi mahasiswa dengan karakteristik beragam.
5. Penelitian ini tidak memasukkan dan mengontrol variabel lain di luar kerangka TAM, seperti pengalaman sebelumnya dalam menggunakan AI, latar belakang secara teknis, atau motivasi belajar mahasiswa.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang dijabarkan, tujuan dari penelitian ini adalah sebagai berikut:

1. Menganalisis performa model *logistic regression* dalam memprediksi niat mahasiswa terhadap penggunaan *artificial intelligence* dalam kegiatan akademik berdasarkan variabel kerangka *Technology Acceptance Model*.
2. Mengidentifikasi variabel-variabel konstruk *Technology Acceptance Model* yang paling signifikan dalam mempengaruhi niat mahasiswa terhadap penggunaan *artificial intelligence*, ditinjau dari perspektif inferensial melalui *logistic regression*, serta dari perspektif prediktif melalui *random forest*.

3. Membandingkan hasil prediksi *logistic regression* dan *random forest* dalam memprediksi niat mahasiswa, dan menganalisis *trade-off* antara interpretabilitas model dengan akurasi prediktif dari kedua metode.

1.5 Manfaat Penelitian

Manfaat dalam penelitian ini adalah sebagai berikut:

1. Berkontribusi terhadap pengembangan studi penerimaan teknologi berbasis *Technology Acceptance Model* dengan pendekatan prediktif.
2. Mengimplementasikan teori perilaku adopsi teknologi dengan metode klasifikasi berbasis *machine learning*.
3. Memberikan pandangan kepada institusi pendidikan tinggi mengenai pola minat mahasiswa terhadap penggunaan *artificial intelligence* dalam pembelajaran, dan diharapkan dapat menjadi perumusan strategi implementasi teknologi yang lebih efektif.
4. Hasil prediksi model dapat dimanfaatkan dalam mengidentifikasi kelompok mahasiswa dengan tingkat minat yang rendah terhadap penggunaan *artificial intelligence*, sehingga institusi dapat merancang program atau pendampingan yang sesuai dengan kebutuhan mahasiswa.

1.6 Sistematika Penulisan

1. BAB 1 Pendahuluan
Bab ini menjelaskan latar belakang penelitian, rumusan masalah, batasan penelitian, tujuan penelitian, manfaat penelitian, dan sistematika penulisan laporan.
2. BAB 2 Landasan Teori
Bab ini membahas teori-teori yang mendukung penelitian, mencakup teori tentang *Technology Acceptance Model* (TAM), *artificial intelligence* dalam pembelajaran, *machine learning*, *logistic regression*, *random forest*, evaluasi model klasifikasi dan penelitian terdahulu.
3. BAB 3 Metodologi Penelitian
Bab ini menguraikan sumber data, definisi variabel, serta pengukuran variabel

yang digunakan dalam penelitian. Selain itu bab ini juga menjelaskan tahapan persiapan dan validasi data, yang mencakup pengumpulan dan persiapan data, data *preprocessing*, analisis statistik, serta pengembangan model penelitian. Masing-masing tahapan dilakukan untuk memastikan data yang digunakan sudah sesuai dan layak untuk dianalisis, yang nantinya akan ditampilkan pada bab berikutnya.

4. BAB 4 Hasil dan Diskusi

Bab ini menyajikan hasil penelitian yang mencakup penujian model dan analisis variabel pada model. Analisis tersebut dilakukan melalui interpretasi koefisien pada model *logistic regression* dan *feature importance* pada model *random forest*, untuk mengetahui kontribusi masing-masing variabel terhadap hasil prediksi. Selain itu, pada bab ini juga membahas evaluasi performa model menggunakan *confusion matrix*, analisis performa berdasarkan kelas, dan analisis misklasifikasi. Terakhir, pada bab ini menjelaskan kesesuaian hasil penelitian dengan kerangka *Technology Acceptance Model* untuk melihat sejauh mana temuan penelitian mendukung teori yang digunakan.

5. BAB 5 Simpulan dan Saran

Bab ini berisi kesimpulan dan saran dari hasil penelitian yang dilakukan. Kesimpulan disusun untuk menjawab rumusan masalah berdasarkan hasil analisis yang diperoleh. Pada bab ini juga menjabarkan saran yang ditunjukkan kepada pihak institusi dan peneliti selanjutnya guna mendukung perkembangan penelitian dimasa mendatang.

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A