

# **BAB I**

## **PENDAHULUAN**

### **1.1. Latar Belakang**

Ketimpangan kinerja akademik mahasiswa masih menjadi isu strategis dalam pendidikan tinggi, baik di Indonesia maupun secara global. Capaian akademik yang direpresentasikan melalui Indeks Prestasi Kumulatif (IPK), ketepatan waktu kelulusan, serta konsistensi performa akademik antar semester menunjukkan variasi yang signifikan. Data Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi (Kemendikbudristek) tahun 2023 mencatat bahwa lebih dari 35% mahasiswa mengalami keterlambatan masa studi, sementara sekitar 20% mahasiswa belum mencapai IPK minimum 3,00 setelah memasuki tahun ketiga perkuliahan [1]. Kondisi ini menegaskan perlunya pendekatan analitik yang mampu mengidentifikasi risiko akademik secara dini dan objektif.

Fenomena tersebut juga tercermin pada tingkat institusi, khususnya di Universitas Multimedia Nusantara (UMN). Analisis awal terhadap data akademik mahasiswa UMN menunjukkan adanya ketimpangan performa yang cukup mencolok, baik dari sisi IPK, progres studi, maupun stabilitas capaian akademik. Berbagai penelitian terdahulu menunjukkan bahwa kinerja akademik mahasiswa dipengaruhi oleh kombinasi faktor akademik dan non-akademik, seperti latar belakang pendidikan, beban studi, keterlibatan dalam aktivitas organisasi, serta kondisi sosial ekonomi. Tanpa dukungan sistem prediksi berbasis data, proses identifikasi mahasiswa berisiko cenderung bersifat reaktif dan berpotensi menghambat intervensi akademik yang tepat waktu.

Dalam pemodelan prediktif kinerja akademik, Light Gradient Boosting Machine (LightGBM) banyak dilaporkan sebagai metode yang unggul dalam menangani data berdimensi tinggi dan berskala besar [2]. Beberapa penelitian menunjukkan bahwa LightGBM mampu menghasilkan akurasi yang lebih tinggi dibandingkan metode boosting konvensional dengan waktu komputasi yang lebih

efisien [3]. Kelebihan lain dari LightGBM adalah kemampuannya menangkap hubungan non-linear antar variabel akademik secara efektif. Namun demikian, penelitian sebelumnya juga mencatat bahwa LightGBM sensitif terhadap pemilihan hyperparameter dan berpotensi mengalami overfitting pada dataset berukuran kecil apabila tidak disertai proses tuning dan validasi yang memadai.

Metode pembandingan, Random Forest telah banyak digunakan dalam prediksi kinerja akademik karena stabilitasnya terhadap noise dan kemampuannya mengurangi variansi melalui mekanisme ensemble berbasis banyak pohon keputusan. Studi-studi terdahulu menunjukkan bahwa Random Forest mampu memberikan performa prediksi yang kompetitif serta relatif lebih tahan terhadap overfitting dibandingkan model pohon tunggal [4] [5]. Meskipun demikian, peningkatan jumlah pohon dan fitur dapat menyebabkan kompleksitas model yang tinggi, meningkatkan waktu komputasi, serta membatasi interpretabilitas global hasil prediksi, terutama dalam konteks pengambilan keputusan kebijakan akademik [6].

Multilayer Perceptron (MLP) juga banyak dimanfaatkan dalam penelitian kinerja akademik karena kemampuannya memodelkan pola non-linear yang kompleks melalui jaringan saraf berlapis. Penelitian sebelumnya melaporkan bahwa MLP mampu menangkap hubungan laten antar variabel akademik yang sulit dimodelkan oleh metode konvensional [7] [8]. Namun, MLP sering dikategorikan sebagai model kotak hitam karena rendahnya transparansi dalam proses pengambilan keputusan. Selain itu, performanya sangat bergantung pada kualitas data, pemilihan arsitektur jaringan, serta jumlah data latih, sehingga berpotensi mengalami overfitting apabila tidak disertai regularisasi dan validasi yang tepat [9].

Berdasarkan kelebihan dan keterbatasan masing-masing metode tersebut, penelitian ini bertujuan untuk melakukan komparasi sistematis antara LightGBM, Random Forest, dan Multilayer Perceptron guna menentukan model terbaik dalam memprediksi kinerja akademik mahasiswa UMN. Proses penelitian meliputi tahap

prapemrosesan data, rekayasa fitur, pelatihan dan evaluasi model menggunakan metrik kinerja yang relevan, serta analisis interpretabilitas menggunakan SHAP. Beberapa penelitian terkini menunjukkan bahwa SHAP efektif dalam menjelaskan kontribusi fitur pada model machine learning yang kompleks dan meningkatkan kepercayaan pengguna terhadap hasil prediksi [10] [11]. Dengan demikian, penelitian ini tidak hanya berfokus pada akurasi prediksi, tetapi juga pada penyediaan wawasan interpretatif yang dapat mendukung pengambilan kebijakan akademik yang lebih adaptif, transparan, dan berkelanjutan.

## **1.2. Rumusan Masalah**

Rumusan masalah dalam penelitian ini adalah sebagai berikut:

- 1) Bagaimana kinerja model Light Gradient Boosting Machine (LightGBM), Random Forest, dan Multilayer Perceptron (MLP) dalam memprediksi kinerja akademik mahasiswa Universitas Multimedia Nusantara berdasarkan data akademik yang tersedia?
- 2) Model manakah di antara LightGBM, Random Forest, dan Multilayer Perceptron yang memberikan performa terbaik dalam memprediksi kinerja akademik mahasiswa Universitas Multimedia Nusantara berdasarkan metrik evaluasi yang digunakan, serta bagaimana perbandingan keunggulan dan keterbatasan masing-masing model dalam konteks prediksi kinerja akademik?
- 3) Sejauh mana hasil prediksi dan interpretasi model yang diperoleh melalui pendekatan Explainable Machine Learning (SHAP) dapat dimanfaatkan sebagai dasar pendukung pengambilan keputusan akademik yang lebih adaptif, objektif, dan berbasis data di Universitas Multimedia Nusantara, khususnya dalam mengidentifikasi mahasiswa berisiko?

### 1.3. Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diidentifikasi, tujuan utama dari penelitian ini antara lain:

- 1) Menganalisis kinerja model Light Gradient Boosting Machine (LightGBM), Random Forest, dan Multilayer Perceptron (MLP) dalam memprediksi kinerja akademik mahasiswa Universitas Multimedia Nusantara berdasarkan data akademik yang tersedia.
- 2) Membandingkan performa model LightGBM, Random Forest, dan Multilayer Perceptron serta menentukan model terbaik dalam memprediksi kinerja akademik mahasiswa Universitas Multimedia Nusantara berdasarkan metrik evaluasi yang digunakan, sekaligus mengkaji keunggulan dan keterbatasan masing-masing model.
- 3) Mengevaluasi pemanfaatan hasil prediksi dan interpretasi model berbasis Explainable Machine Learning menggunakan SHAP sebagai dasar pendukung pengambilan keputusan akademik yang lebih adaptif, objektif, dan berbasis data di Universitas Multimedia Nusantara, khususnya dalam mengidentifikasi mahasiswa yang berisiko mengalami penurunan kinerja akademik.

### 1.4. Urgensi Penelitian

Urgensi dari penelitian ini dapat diidentifikasi secara sistematis dengan mempertimbangkan beberapa aspek berikut:

- 1) Urgensi Akademik

Ketimpangan kinerja akademik mahasiswa merupakan permasalahan fundamental yang berdampak langsung pada kualitas lulusan dan efektivitas sistem pendidikan tinggi. Secara akademik, masih terdapat keterbatasan penelitian yang secara komprehensif membandingkan berbagai model machine

learning modern dalam konteks prediksi kinerja mahasiswa, khususnya dengan mengintegrasikan aspek interpretabilitas model. Sebagian besar studi sebelumnya cenderung berfokus pada pencapaian akurasi prediksi tanpa memberikan pemahaman yang memadai mengenai faktor-faktor yang memengaruhi hasil prediksi tersebut. Oleh karena itu, penelitian ini menjadi penting untuk memperkaya khazanah keilmuan dengan menghadirkan analisis komparatif antara LightGBM, Random Forest, dan Multilayer Perceptron yang tidak hanya menilai performa prediktif, tetapi juga mengevaluasi transparansi dan keterjelasan model melalui pendekatan Explainable Machine Learning. Hasil penelitian diharapkan dapat menjadi referensi akademik dalam pengembangan model prediksi kinerja mahasiswa yang lebih akurat, interpretable, dan relevan dengan kebutuhan institusi pendidikan tinggi.

## 2) Urgensi Teknologi

Dari sisi teknologi, meningkatnya ketersediaan data akademik menuntut pemanfaatan metode analitik cerdas yang mampu mengolah data secara efisien dan menghasilkan informasi yang bernilai guna. Namun, penerapan model machine learning di lingkungan pendidikan masih menghadapi tantangan terkait kepercayaan pengguna dan keterbatasan interpretasi hasil model, terutama ketika digunakan sebagai dasar pengambilan keputusan strategis. Penelitian ini memiliki urgensi teknologi karena mengadopsi dan membandingkan model machine learning yang representatif dari pendekatan ensemble tree-based dan neural network, serta mengintegrasikan teknik SHAP untuk menjelaskan kontribusi setiap fitur terhadap hasil prediksi. Pendekatan ini diharapkan dapat menghasilkan solusi prediktif yang tidak hanya unggul secara teknis, tetapi juga transparan dan dapat diimplementasikan sebagai sistem pendukung keputusan akademik yang adaptif, objektif, dan berbasis data di Universitas Multimedia Nusantara.

### 1.5. Luaran Penelitian

Luaran utama dari penelitian ini adalah artikel ilmiah yang disusun oleh mahasiswa dan direncanakan untuk dipublikasikan pada *Journal of Informatics and Visualization (JOIV)*, yaitu jurnal internasional yang terindeks Scopus pada peringkat Q3. Publikasi ini diharapkan menjadi sarana pengembangan kompetensi akademik mahasiswa, khususnya dalam bidang penelitian, penulisan ilmiah, dan penerapan metode analisis berbasis data. Kontribusi yang diharapkan dari penelitian ini meliputi:

- 1) Memberikan pengalaman akademik kepada mahasiswa dalam menerapkan konsep machine learning dan explainable artificial intelligence pada permasalahan nyata, sehingga memperkuat pemahaman teoritis dan keterampilan praktis di bidang analitik data.
- 2) Menghasilkan kontribusi ilmiah berupa kajian komparatif model prediksi yang dapat dijadikan referensi bagi mahasiswa dan peneliti lain dalam pengembangan penelitian lanjutan, khususnya pada bidang analisis kinerja akademik dan educational data mining.
- 3) Meningkatkan kemampuan mahasiswa dalam seluruh tahapan penelitian, mulai dari pengolahan dan analisis data, pemodelan dan evaluasi, hingga interpretasi hasil penelitian secara ilmiah dan sistematis.
- 4) Menjadi bentuk luaran akademik mahasiswa yang mencerminkan ketercapaian pembelajaran berbasis penelitian serta mendukung budaya publikasi ilmiah di lingkungan Universitas Multimedia Nusantara melalui jurnal internasional bereputasi.

### 1.6. Manfaat Penelitian

Penelitian ini diharapkan memberikan manfaat yang lebih luas, baik dari sisi pengembangan ilmu pengetahuan maupun penerapannya dalam praktik di perguruan tinggi. Adapun manfaat yang dapat diperoleh adalah sebagai berikut:

### 1.6.1 Manfaat Teoritis

Manfaat Teoritis dalam penelitian ini antara lain:

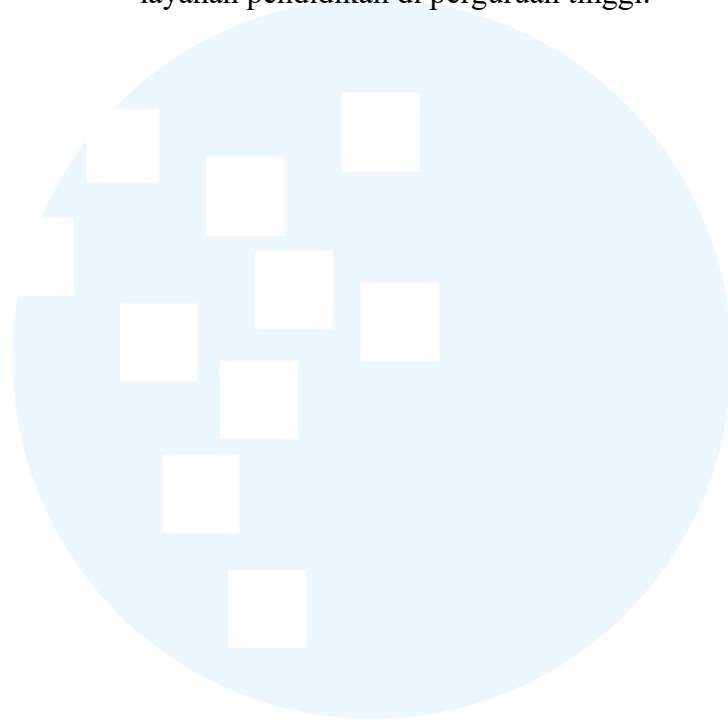
- 1) Memberikan sumbangan keilmuan dalam bidang educational data mining dan explainable artificial intelligence, khususnya terkait pengembangan model prediksi kinerja akademik mahasiswa yang berbasis data.
- 2) Memperkaya literatur ilmiah mengenai pemanfaatan algoritma machine learning, seperti LightGBM, Random Forest, dan Multilayer Perceptron, dalam membangun model prediktif yang tidak hanya menekankan akurasi, tetapi juga memiliki tingkat keterjelasan dan transparansi yang baik.
- 3) Menjadi landasan konseptual bagi penelitian selanjutnya yang berfokus pada pengembangan sistem prediksi kinerja akademik yang terintegrasi dengan strategi intervensi akademik yang lebih adaptif dan berkelanjutan.

### 1.6.2 Manfaat Praktis

Manfaat Praktis dalam penelitian ini antara lain:

- 1) Membantu pihak perguruan tinggi dalam melakukan identifikasi dini terhadap mahasiswa yang berpotensi mengalami penurunan kinerja akademik melalui pendekatan prediktif yang sistematis dan mudah dipahami.
- 2) Menyediakan hasil analisis dan visualisasi interpretatif yang dapat dimanfaatkan oleh dosen, pembimbing akademik, serta pengambil kebijakan sebagai dasar dalam merancang program pembinaan dan intervensi akademik yang lebih tepat sasaran.
- 3) Mendukung pengembangan awal sistem pendukung keputusan berbasis Explainable AI yang dapat

diimplementasikan dalam pengelolaan akademik, guna meningkatkan kualitas proses pembelajaran dan efektivitas layanan pendidikan di perguruan tinggi.



UMN  
UNIVERSITAS  
MULTIMEDIA  
NUSANTARA