

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Kecerdasan buatan merupakan salah satu cabang ilmu pengetahuan yang membantu manusia dalam menyelesaikan berbagai permasalahan. Salah satu pendekatan dalam kecerdasan buatan adalah *Machine Learning*, yang terdiri dari algoritma-algoritma komputer yang dapat meningkatkan kinerja program komputer berdasarkan kumpulan data yang sudah ada. *Machine Learning* memiliki tiga jenis pembelajaran, 1) *Supervised Learning*, digunakan secara luas untuk memprediksi dan mengklasifikasikan objek atau pola berdasarkan label atau kategori yang diharapkan. 2) *Unsupervised Learning*, mendeteksi pola pada data berdasarkan kemiripan atributnya dan tidak memiliki label pada outputnya. 3) *Reinforcement Learning*, menciptakan model pembelajaran dari tindakan yang diambil dalam lingkungan tertentu. Dalam konteks klasifikasi, terdapat beberapa algoritma yang dapat digunakan, seperti *C4.5*, *Naive Bayes*, *K-Nearest Neighbor*, *Artificial Neural Network*, dan lainnya.

Banyak peneliti telah memanfaatkan teknologi machine learning, seperti yang terlihat dalam berbagai studi berikut: 1) Penelitian berjudul "PREDIKSI PENERIMAAN MAHASISWA BARU DENGAN MULTI LAYER PERCEPTRON" menunjukkan bahwa Algoritma Multilayer Perceptron (MLP) menghasilkan prediksi yang cukup baik dengan tingkat akurasi sebesar 91%[1]. 2) Dalam studi yang berjudul "Analisis Perbandingan Algoritma C4.5 Dan Naïve Bayes Dalam Memprediksi Penyakit Cerebrovascular" oleh Kelvin Leonardi Kohsasih dan Zakarias Situmorang, algoritma Naïve Bayes mencapai tingkat akurasi sebesar 91%, sementara algoritma C4.5 memperoleh tingkat akurasi yang lebih tinggi, yaitu 95%[2]. 3) Penelitian yang dilakukan oleh Annisa Ulfa, Doni Winarso M.Kom, Edo Arribe M.MSi, pada tahun 2020 dengan judul "SISTEM REKOMENDASI JURUSAN KULIAH BAGI CALON MAHASISWA BARU MENGGUNAKAN ALGORITMA C4.5", mendapati hasil akurasi algoritma C4.5 sebesar 12%. Hal ini mengindikasikan bahwa sistem tersebut gagal melakukan

rekomendasi yang sesuai karena data dan variabel yang digunakan tidak cukup mendukung [3]. 4) Penelitian berjudul "Analisis Perbandingan Kinerja Algoritma Multilayer Perceptron dan K-Nearest Neighbor pada Klasifikasi Tipe Migrain" yang dilakukan oleh Yang Agita Rindri dan Agus Fitriyani pada tahun 2023 menunjukkan bahwa Algoritma Multilayer Perceptron (MLP) menghasilkan prediksi yang cukup baik dengan tingkat akurasi sebesar 91%. Di sisi lain, algoritma K-Nearest Neighbor menghasilkan prediksi dengan tingkat akurasi sebesar 72%[4]. Hasil-hasil ini mencerminkan perbedaan kinerja algoritma machine learning dalam berbagai konteks penelitian.

Meskipun terdapat banyak algoritma pembelajaran mesin yang tersedia, tidak ada satu algoritma yang dapat dianggap universal atau paling efektif untuk semua jenis data dan aplikasi. Peneliti sering kali dihadapkan pada dilema algoritma mana yang harus dipilih untuk mendapatkan hasil terbaik? Keputusan ini bergantung pada berbagai faktor, termasuk kompleksitas data, kecepatan komputasi yang diperlukan, dan ketersediaan sumber daya, ketidakpastian ini sering menyebabkan kebingungan dan kesulitan dalam membuat keputusan [5].

Saat ini belum ada penelitian yang secara menyeluruh membandingkan algoritma *Multilayer Perceptron* dan *Naïve Bayes* ini dalam hal prediksi minat studi. Oleh karena itu, penelitian ini bertujuan untuk membandingkan kinerja algoritma *Naïve Bayes* dan *MLP*, dalam hal prediksi minat studi calon siswa. Penulis juga menggunakan metode klasifikasi *data mining* yang disebut *Knowledge Discovery in Database* (KDD) untuk mendukung penelitian ini. KDD bertujuan untuk mengumpulkan dan menggunakan data historis untuk menemukan pola keteraturan dalam kumpulan data yang sangat besar [5]. Penelitian ini diharapkan akan memberikan wawasan yang bermanfaat tentang algoritma yang paling efektif untuk klasifikasi minat studi. Hasilnya juga diharapkan akan membantu mengembangkan sistem klasifikasi minat studi yang lebih efektif dan dapat digunakan secara luas.

## 1.2 Rumusan Masalah

Dari latar belakang yang sudah diuraikan, dapat dilakukan penelitian dengan rumusan masalah sebagai berikut:

1. Bagaimana kinerja Algoritma *Multilayer Perceptron* dan *Naïve Bayes* untuk memprediksi minat studi calon Mahasiswa?
2. Bagaimana perbandingan kinerja antara algoritma *Naïve Bayes* dan *Multilayer Perceptron* dalam prediksi minat studi calon Mahasiswa?

## 1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk melakukan perbandingan komprehensif antara algoritma MLP dan Naive Bayes, mengidentifikasi keunggulan dan keterbatasan masing-masing dalam berbagai skenario aplikasi. Evaluasi akan mencakup metrik kinerja utama seperti akurasi, presisi, *recall*, *F1-score*, dan waktu komputasi. Dengan ini, penelitian diharapkan dapat mengetahui performa dari algoritma *Multilayer Perceptron* (MLP) dan *Naïve Bayes* dalam prediksi minat studi calon Mahasiswa.

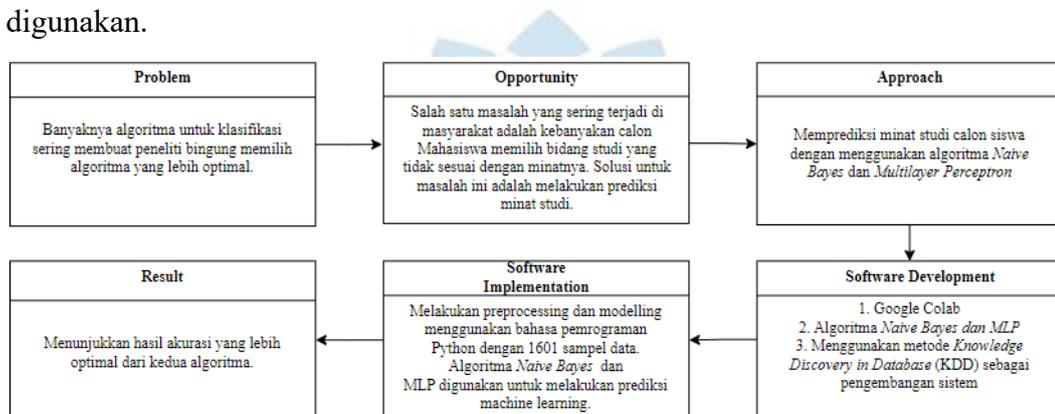
## 1.4 Batasan Masalah

Dalam penelitian ini terdapat beberapa batasan masalah agar tidak menyimpang dari pokok pembahasan, diantaranya:

1. Sampel data yang digunakan diunduh melalui website Kaggle <https://www.kaggle.com/datasets/abtabm/hobby-prediction-basic> diunggah oleh Abduttayyeb Rampurawala pada tahun 2020 dengan nama “Kids Hobby Prediction Dataset”. Dataset terdiri dari 13 attributes dan 3 label dimana 13 attributes tersebut dimodelkan untuk menghasilkan salah satu dari 3 label,
2. Algoritma yang diuji adalah algoritma *Multilayer Perceptron* dan algoritma *Naïve Bayes*,
3. Hasil akhir dari penelitian ini adalah perbandingan hasil akurasi dari kedua Algoritma yang digunakan.

## 1.5 Kerangka Pemikiran

Gambar 1.1 mewakili kerangka pemikiran yang akan digunakan dalam penelitian ini. Permasalahan yang diangkat adalah kebingungan yang sering muncul akibat banyaknya pilihan algoritma yang tersedia untuk tujuan klasifikasi[6]. Untuk mengatasi hal ini, pendekatan yang diusulkan melibatkan perbandingan antara algoritma-algoritma tersebut. Penelitian ini selanjutnya menerapkan dua algoritma berbeda untuk klasifikasi, dengan menggunakan metode *Knowledge Discovery in Database* (KDD) dalam pengembangan sistem. Tahapan metode ini akan diimplementasikan menggunakan bahasa pemrograman *Python*, dengan tujuan akhirnya untuk mencapai akurasi yang paling optimal diantara algoritma yang digunakan.



Gambar 1.1 Kerangka Pemikiran

## 1.6 Sistematika Penulisan

Laporan tugas akhir kemudian disusun menggunakan data dan informasi yang diperoleh melalui metode yang telah dijelaskan sebelumnya. Untuk memberikan gambaran dan sistematika yang cukup jelas, penulis membagi laporan menjadi lima bab, yang diurutkan sebagai berikut:

### 1. BAB I : PENDAHULUAN

Pada bab ini berisi uraian mengenai latar belakang masalah, batasan masalah, tujuan penelitian, kerangka pemikiran, metodologi penelitian dan sistematika penulisan.

### 2. BAB II : KAJIAN LITERATUR

Pada bab ini membahas tentang perkembangan paling mutakhir dalam dunia keilmuan dan penelitian atau sering disebut dengan state of the art dari teori

yang sedang dikaji dan kedudukan masalah penelitian dalam bidang informatika yang diteliti.

### **3. BAB III : METODOLOGI PENELITIAN**

Pada bab ini menjelaskan langkah-langkah dan teknik yang dilakukan dalam penelitian, dijelaskan secara kronologis dan sistematis. Umumnya metode penelitian mengacu pada model proses pengembangan perangkat lunak, atau model lain yang sesuai dengan kebutuhan dan karakteristik penelitian yang dilakukan.

### **4. BAB IV : HASIL DAN PEMBAHASAN**

Pada bab ini dipaparkan dua hal utama, pertama pemaparan tentang temuan atau hasil penelitian berdasarkan tahapan penelitian yang dilakukan. Dipaparkan hasil penelitian dalam bentuk kualitatif atau kuantitatif berdasarkan hasil pengelolaan dan analisis data. Pemaparan hasil penelitian disesuaikan dengan urutan rumusan masalah penelitian. Kedua pembahasan hasil atau temuan penelitian untuk menjawab rumusan penelitian.

### **5. BAB V : KESIMPULAN DAN SARAN**

Pada bab ini berisi simpulan hasil penelitian serta saran penulis terhadap pembaca. Bab ini berisi kesimpulan komprehensif dari penelitian dan analisis, termasuk jawaban terhadap rumusan masalah dan pencapaian tujuan penelitian. Disertai dengan saran untuk penelitian lanjutan atau pengembangan lebih lanjut terkait topik yang diteliti.

### **6. DAFTAR PUSTAKA**

Daftar Pustaka berisi seluruh sumber yang ada pada laporan penulisan dan digunakan dalam penelitian

### **7. LAMPIRAN**

Lampiran berisikan dokumen yang digunakan dalam proses penyusunan dan perancangan dalam penelitian.