

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Secara global, transformasi teknologi informasi telah mengubah peta kebutuhan tenaga kerja, di mana penguasaan spesialisasi teknis yang mendalam kini menjadi standar utama dalam rekrutmen industri. Tren internasional saat ini menunjukkan bahwa sekadar memiliki literasi digital dasar tidak lagi memadai. Industri membutuhkan talenta yang memiliki kemahiran teknologi (*advanced digital skills*) untuk menjalankan ekosistem digital yang semakin kompleks. Di Indonesia, urgensi penyediaan SDM berkualitas ini menjadi semakin krusial seiring dengan pesatnya pertumbuhan sektor digital yang diprediksi akan terus mendominasi penggerak ekonomi nasional, dengan proyeksi nilai pasar mencapai USD 146 miliar pada tahun 2025 (Gayatri et al., 2022).

Meskipun peluang di sektor digital semakin besar, terdapat tantangan serius yang dikenal sebagai *digital talent gap*. Berdasarkan penelitian (Gayatri et al., 2022), ditemukan adanya kondisi yang cukup kontradiktif dalam dunia ketenagakerjaan di Indonesia. Di satu sisi, jumlah lulusan di bidang teknologi diperkirakan mencapai surplus hingga 600.000 orang setiap tahunnya. Namun di sisi lain, industri masih mengalami kesulitan dalam menemukan tenaga kerja yang memiliki keahlian spesialis, khususnya pada level tinggi (level 7-9 KKNI). Hal ini menunjukkan bahwa permasalahan utama bukan terletak pada jumlah lulusan, melainkan pada ketidaksesuaian kompetensi atau *skills mismatch*. Istilah ini

merujuk pada adanya perbedaan antara pengetahuan yang diperoleh selama perkuliahan dengan keterampilan praktis yang dibutuhkan di dunia kerja.

Temuan ini juga diperkuat oleh penelitian (Tackbir Abubakar et al., 2024), yang menunjukkan bahwa kualitas talenta digital Indonesia masih berada pada peringkat 111 dari 176 negara di tingkat global. Kondisi tersebut dipengaruhi oleh masih rendahnya jumlah lulusan di bidang STEM serta belum meratanya penguasaan keterampilan spesifik yang dibutuhkan oleh industri. Tantangan ketidaksesuaian kompetensi tersebut berkaitan erat dengan karakteristik bidang Sistem Informasi yang secara fundamental menjembatani tata kelola bisnis dan solusi teknologi. Sebagaimana dijelaskan oleh (McKenzie & Bennett, 2022), banyak mahasiswa Teknologi Informasi memasuki perguruan tinggi tanpa memiliki gambaran yang jelas mengenai jalur karier yang akan ditempuh. Meskipun memiliki minat terhadap bidang teknologi, sebagian besar mahasiswa masih memiliki aspirasi karier yang bersifat umum dan belum mampu menentukan bidang yang ingin ditekuni secara spesifik.

Kondisi tersebut menyebabkan proses pengembangan kompetensi selama masa perkuliahan menjadi kurang terarah sehingga diperlukan dukungan yang dapat membantu mahasiswa mengubah minat menjadi pilihan karier yang jelas. Dalam konteks Program Studi Sistem Informasi yang memiliki cakupan keilmuan luas, kondisi tersebut sering diilustrasikan dengan ungkapan "*knows a bit of everything, but a master of none*", yaitu memahami berbagai bidang secara umum, tetapi belum memiliki keahlian yang mendalam pada salah satu bidang tertentu. Oleh karena itu, diperlukan suatu mekanisme yang mampu memetakan kecenderungan minat

mahasiswa secara lebih objektif sejak masa perkuliahan agar pengembangan kompetensi dapat dilakukan secara lebih terarah.

Upaya untuk memetakan potensi mahasiswa sebenarnya telah banyak dilakukan melalui berbagai pendekatan penelitian. Namun, hasil yang diperoleh masih menunjukkan adanya ruang untuk pengembangan lebih lanjut. Penelitian oleh (Rahmallia, 2024) mengembangkan sistem identifikasi minat bakat menggunakan kombinasi algoritma *Random Forest* dan *Decision Tree*. Sistem tersebut mampu mencapai tingkat akurasi sebesar 84%. Akan tetapi, penelitian ini masih berfokus pada tingkat Sekolah Menengah Atas (SMA). Selain itu, variabel yang digunakan masih relatif sederhana, sehingga belum cukup untuk menggambarkan kompleksitas kebutuhan pemilihan jalur karier di perguruan tinggi yang menuntut pemetaan kemampuan akademik dan pengalaman teknis yang lebih mendalam.

Selanjutnya, penelitian dari (Hamidani et al., 2025) mencoba mengatasi permasalahan identifikasi minat bakat di tingkat universitas dengan menerapkan algoritma *Naïve Bayes Classifier*. Meskipun demikian, metode ini memiliki keterbatasan mendasar, yaitu asumsi *class conditional independence*, yang berarti setiap variabel dianggap berdiri sendiri tanpa mempertimbangkan keterkaitan antar pola perilaku belajar mahasiswa secara keseluruhan. Di samping itu, pendekatan yang digunakan bersifat satu tahap, di mana sistem langsung menghasilkan prediksi jalur peminatan berdasarkan data yang bersifat statis. Hal ini membuat metode tersebut kurang mampu menangkap perubahan dan kompleksitas karakter individu mahasiswa yang sebenarnya lebih dinamis.

Celah serupa juga ditemukan pada riset (Divya & Manaswini, 2025) yang membandingkan performa algoritma klasifikasi seperti *Support Vector Machines (SVM)* hingga *Neural Networks* untuk rekomendasi karier. Kelemahan utama dari penelitian ini adalah tidak adanya tahap awal berupa pengelompokan atau *clustering* untuk mengenali pola karakter mahasiswa. Tanpa proses klasterisasi, sistem tidak terlebih dahulu memahami profil dasar mahasiswa sebelum memberikan klasifikasi ke dalam kategori tertentu. Akibatnya, hasil rekomendasi yang dihasilkan cenderung bersifat umum dan kurang memberikan personalisasi yang mendalam.

Penelitian ini bertujuan untuk membantu mahasiswa dalam menentukan minat akademik yang sesuai melalui pengembangan Sistem Rekomendasi Minat Akademik berbasis Machine Learning. Sistem dibangun menggunakan data yang diperoleh dari 13 indikator kuesioner, yang mencakup aspek perilaku belajar, motivasi, gaya belajar, pengalaman akademik, pengalaman profesional, serta preferensi terhadap mata kuliah tertentu. Data tersebut kemudian diproses menggunakan tahapan Knowledge Discovery in Database (KDD) untuk memperoleh informasi yang relevan secara sistematis.

Metode yang digunakan merupakan pendekatan hybrid yang menggabungkan algoritma K-Means Clustering dan Decision Tree CART. Pemilihan metode ini didasarkan pada kondisi data yang belum memiliki label kategori minat. Pada tahap awal, algoritma K-Means digunakan untuk mengelompokkan mahasiswa berdasarkan kemiripan karakteristik yang dimiliki. Proses ini memungkinkan

terbentuknya kelompok-kelompok mahasiswa dengan pola karakteristik yang serupa tanpa memerlukan label awal.

Selanjutnya, setiap klaster dianalisis melalui nilai centroid untuk mengidentifikasi karakteristik dominan pada masing-masing kelompok. Berdasarkan hasil analisis tersebut, setiap klaster diberikan label yang merepresentasikan profil mahasiswa yang terbentuk. Label hasil interpretasi ini kemudian dimanfaatkan sebagai data target pada tahap klasifikasi menggunakan algoritma Decision Tree CART.

Melalui proses klasifikasi, sistem mempelajari hubungan antaratribut dan menghasilkan aturan keputusan dalam bentuk pohon keputusan yang mudah dipahami. Aturan tersebut digunakan sebagai dasar dalam memberikan rekomendasi minat akademik kepada mahasiswa. Selain menghasilkan model klasifikasi, penelitian ini juga mengembangkan aplikasi berbasis web menggunakan framework Flask yang berfungsi sebagai media implementasi sistem rekomendasi.

Kinerja model dievaluasi menggunakan Confusion Matrix dengan pengukuran akurasi, presisi, dan recall untuk memastikan kualitas hasil klasifikasi yang diperoleh. Dengan adanya sistem ini, proses penentuan minat akademik diharapkan dapat dilakukan secara lebih objektif dan terarah sehingga membantu mahasiswa dalam mengembangkan kompetensi yang sesuai dengan kebutuhan dunia kerja.

B. Rumusan Masalah

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

1. Bagaimana proses pengelompokan karakteristik mahasiswa menggunakan algoritma K-Means untuk membentuk dasar klasifikasi minat akademik?
2. Bagaimana membangun model klasifikasi menggunakan algoritma Decision Tree berdasarkan hasil clustering untuk menghasilkan rekomendasi minat akademik mahasiswa?
3. Bagaimana mengevaluasi tingkat akurasi dan relevansi rekomendasi minat akademik yang dihasilkan oleh model?

C. Batasan Masalah

Agar penelitian ini memiliki ruang lingkup yang jelas dan tetap fokus pada tujuan yang telah ditetapkan, maka diberikan batasan-batasan masalah sebagai berikut:

1. Data yang digunakan dalam penelitian ini dibatasi pada hasil kuesioner mahasiswa Program Studi Sistem Informasi Universitas PGRI Madiun. Pemilihan responden tersebut didasarkan pada tujuan penelitian, yaitu mengembangkan sistem rekomendasi minat akademik yang disesuaikan dengan karakteristik mahasiswa, kurikulum, serta bidang spesialisasi yang terdapat pada Program Studi Sistem Informasi Universitas PGRI Madiun. Oleh karena itu, hasil penelitian ini tidak dimaksudkan untuk digeneralisasikan pada program studi lain.

2. Penentuan rekomendasi minat akademik didasarkan pada tiga belas indikator yang meliputi tingkat kehadiran, keaktifan di kelas, jam belajar, motivasi belajar, gaya belajar, keaktifan dalam kegiatan di luar perkuliahan, pengalaman mengikuti proyek IT, topik proyek IT, peran dalam proyek, pengalaman pekerjaan freelance, keaktifan organisasi, serta preferensi terhadap mata kuliah yang diminati. Penelitian ini tidak memasukkan aspek psikometri, asesmen profesional, maupun faktor psikologis lainnya sebagai variabel dalam proses analisis.
3. Responden penelitian dipilih menggunakan teknik purposive sampling dengan kriteria mahasiswa aktif Program Studi Sistem Informasi Universitas PGRI Madiun yang mengisi kuesioner secara lengkap dan memenuhi seluruh variabel yang digunakan dalam pembentukan model rekomendasi minat akademik.
4. Proses pengelompokan data dilakukan menggunakan algoritma K-Means untuk mengidentifikasi kesamaan karakteristik antar mahasiswa tanpa menggunakan label awal. Hasil dari proses clustering tersebut berupa label cluster yang kemudian digunakan sebagai dasar dalam tahap klasifikasi menggunakan algoritma Decision Tree CART.
5. Proses klasifikasi dilakukan menggunakan algoritma Classification and Regression Tree (CART) untuk menghasilkan aturan keputusan berdasarkan pola yang terbentuk dari data.
6. Kinerja model yang digunakan dalam penelitian ini dievaluasi untuk mengetahui tingkat keakuratan dan konsistensi hasil yang diperoleh.

7. Sistem yang dibangun dalam penelitian ini berbasis web dengan memanfaatkan Bootstrap sebagai antarmuka pengguna serta Python dengan framework Flask sebagai bagian backend dalam pengolahan data dan implementasi model.
8. Hasil analisis sistem berfungsi sebagai alat pendukung keputusan (*Decision Support Tool*) atau bahan pertimbangan (*second opinion*) bagi mahasiswa dalam pengembangan akademik, bukan merupakan keputusan mutlak.
9. Hasil sistem berperan sebagai alat bantu dalam pengambilan keputusan dan tidak digunakan sebagai dasar keputusan yang bersifat mutlak.
10. Pengujian sistem dibatasi pada pengujian fungsional untuk memastikan setiap fitur, proses pengolahan data, serta mekanisme rekomendasi dapat berjalan sesuai dengan rancangan sistem yang telah dikembangkan.
11. Sistem rekomendasi yang dikembangkan hanya mencakup enam kategori minat akademik berdasarkan bidang spesialisasi pada kurikulum Program Studi Sistem Informasi Universitas PGRI Madiun, yaitu Pengembangan Aplikasi, Data Science, UI/UX Design, Digital Business, Cyber Security & Jaringan, serta Sistem Informasi (Manajemen & Analisis). Sistem tidak mencakup rekomendasi profesi maupun bidang keahlian lain di luar kategori spesialisasi yang telah ditentukan.

D. Tujuan Penelitian

Berdasarkan rumusan masalah yang telah ditetapkan, maka tujuan dari penelitian ini adalah sebagai berikut:

1. Mengelompokkan karakteristik mahasiswa berdasarkan data kebiasaan belajar, minat, dan kompetensi menggunakan algoritma K-Means sebagai dasar pembentukan klasifikasi minat akademik.
2. Membangun model klasifikasi rekomendasi minat akademik menggunakan algoritma Decision Tree CART dengan memanfaatkan hasil pengelompokan K-Means sebagai salah satu fitur masukan untuk menghasilkan rekomendasi spesialisasi akademik mahasiswa.
3. Mengevaluasi kinerja model rekomendasi minat akademik berdasarkan tingkat akurasi klasifikasi dan relevansi hasil rekomendasi, serta mengimplementasikan model ke dalam sistem berbasis web sebagai media penyajian hasil analisis kepada pengguna.

E. Manfaat Penelitian

Melalui hasil yang dicapai, penelitian ini diharapkan dapat memberikan nilai guna yang nyata dan kontribusi positif bagi berbagai pihak sebagai berikut:

1. Manfaat bagi Peneliti

- a. Menjadi wadah bagi peneliti untuk mempraktikkan langsung teori kecerdasan buatan (*Machine Learning*) dan pembuatan website (menggunakan *framework* Flask dan Bootstrap) yang dipelajari di bangku kuliah untuk menyelesaikan masalah nyata di lingkungan kampus.
- b. Menjadi referensi atau contoh dasar bagi peneliti lain yang ingin mengembangkan sistem serupa dengan cakupan data yang lebih besar atau menggunakan metode teknologi yang lebih canggih lagi

2. Manfaat bagi Mahasiswa

- a. Memberikan pandangan baru atau "pendapat kedua" (*second opinion*) yang objektif bagi mahasiswa agar lebih memahami minat akademik mereka melalui hasil pengolahan data kuesioner yang akurat.
- b. Membantu mahasiswa menyelaraskan antara minat akademik dengan pengalaman praktis yang telah dimiliki, sehingga perencanaan masa depan profesional menjadi lebih terukur, terarah, dan sesuai dengan kapasitas diri yang sesungguhnya.

3. Manfaat bagi Program Studi dan Institusi

- a. Menyediakan data yang tervalidasi mengenai kecenderungan minat dan rekam jejak teknis mahasiswa, yang dapat dimanfaatkan oleh Program Studi Sistem Informasi UNIPMA sebagai bahan pertimbangan dalam menyempurnakan kurikulum atau mata kuliah pilihan.
- b. Menjadi alat bantu bagi Dosen Pembimbing Akademik (DPA) dalam memberikan arahan yang lebih personal dan akurat kepada mahasiswa mengenai fokus studi yang paling sesuai dengan profil data mereka.