

BAB I

PENDAHULUAN

1.1 Latar Belakang

Era revolusi 4.0 ditandai dengan kemajuan pada bidang teknologi dan komunikasi. Pada era revolusi 4.0 salah satu teknik pembelajaran yang diterapkan ialah *Collaborative* (kolaborasi). *Collaborative* (kolaborasi) adalah kemampuan berkolaborasi atau bekerja sama, saling bersinergi, beradaptasi dalam berbagai peran dan tanggungjawab [1]. Tentu rasa adaptif tersebut akan menjadi lebih efektif apabila didukung dengan rasa tanggung jawab. Rasa tanggung jawab tersebut dapat diperoleh salah satunya yaitu dengan memiliki minat belajar [2].

Minat merupakan salah satu faktor penentu dalam keberhasilan pendidikan. Dampak dari adanya minat belajar dapat lebih memfokuskan perhatian seseorang kepada apa yang sedang dipelajarinya. Dengan adanya minat belajar sangat berpengaruh sekali terhadap prestasi seseorang atas apa yang sudah dipelajarinya [3].

Ada beberapa faktor yang membuat seseorang ingin melanjutkan studi, diantaranya ialah meningkatnya kualitas sumber daya manusia sehingga memperkecil peluang pekerjaan, kebutuhan untuk ilmu melalui tingkat pendidikan yang lebih tinggi, sehingga setelah lulus mampu bersaing dalam memasuki lapangan kerja. Jadi minat melanjutkan studi merupakan pemusatan pikiran, perasaan senang, kemauan atau perhatian seseorang untuk melanjutkan studi [4]. Minat merupakan potensi diri yang dapat dikembangkan. Manusia mempunyai potensi diri yang dapat disadari bersama, diantaranya: (1) potensi berfikir, manusia memiliki potensi untuk belajar informasi-informasi baru, menghubungkan berbagai informasi, serta menghasilkan pemikiran baru; (2) potensi emosi, setiap manusia memiliki potensi cita-rasa, yang dengannya manusia dapat memahami perasaan orang lain, memperhatikan dan diperhatikan, menghargai dan dihargai; (3) manusia memiliki potensi yang luar biasa untuk membuat gerakan fisik yang efektif dan efisien serta memiliki kekuatan fisik yang tangguh. Orang yang berbakat dalam bidang fisik mampu mempelajari olahraga dengan cepat dan selalu menunjukkan permainan yang baik. Gerakan fisik yang mereka tunjukkan dilandasi oleh kecerdasan intelektual mereka, khususnya intelektualitas yang berkaitan dengan

fisik; (4) potensi sosial, pemilik potensi sosial yang besar memiliki kapasitas menyesuaikan diri dan mempengaruhi orang lain. Kemampuan menyesuaikan diri dan mempengaruhi orang lain didasari kemampuan belajar, baik dalam dataran pengetahuan maupun keterampilan [5].

Permasalahan utama yang masih sering saja terjadi ialah masih banyak orang yang memilih bidang studi tidak sesuai dengan minat mereka, sehingga yang terjadi ialah ketidaknyamanan dalam mempelajari apa yang bukan menjadi *passion* atau minat tersebut [6]. Selain itu juga, mempelajari sesuatu yang bukan atas dasar keinginan atau dengan paksaan dapat menyebabkan depresi pada diri seseorang yang dapat menyebabkan rasa putus asa [7]. Ada beberapa penelitian yang mendukung akar permasalahan diatas, diantaranya: (1) Penelitian yang dilakukan oleh Nurdin and Zainuddin n.d., pada tahun 2016 yang bertempat di Universitas Hasanuddin Makassar bahwa terdapat 115 mahasiswa dari 14 jurusan yang salah memilih jurusan [8]; (2) Penelitian yang dilakukan Rama Nida Siregar, Ahmad Nizar Rangkuti, dan Erna Ikawati pada tahun 2019 yang bertempat di Institut Agama Islam Negeri Padangsidimpuan dengan jumlah 150 responde menyatakan bahwa faktor minat menjadi faktor terbesar mahasiswa memilih Jurusan Tadris Matematika (TMM) dengan jumlah sebesar 24,08 % [9]; (3) Penelitian yang dilakukan oleh Putri Wahyu Wulandari, Sancka Stella, dan Indri Sarwilly pada tahun 2022 yang dilakukan di Program Studi Ilmu Keperawatan pada Sekolah Tinggi Ilmu Kesehatan (STIKes) Indonesia Maju yang berlokasi di Jakarta Selatan menyatakan bahwa terdapat 60% dari 75 responden mahasiswa mengalami ketidaksesuain program studi [10].

Untuk mengatasi masalah diatas tersebut akan di bangun sebuah sistem yang mampu memprediksi suatu saran terkait minat dan bakat seseorang yang berguna untuk melanjutkan studi ke jenjang yang lebih tinggi. Sebuah sistem *machine learning* yang mampu memprediksi minat studi lanjut berdasarkan hasil *questioner* yang diberikan oleh pengguna. Salah satu metode yang cocok digunakan dalam sistem prediksi tersebut adalah metode *K-Nearest Neighbor* (KNN). Metode tersebut digunakan karena menggunakan Algoritma Supervised dimana hasil diklasifikasikan berdasarkan data pembelajaran yang mengambil jarak paling dekat dengan suatu objek [11]. Dataset yang digunakan diunduh dari *website*

<https://www.kaggle.com/datasets/abtabm/hobby-prediction-basic>. Dataset tersebut berjudul “*Efficient Recommender System for Kid’s Hobby Using Machine Learning*”. Setelah dataset didapat, selanjutnya data akan dibagi menjadi data train dan data test. Sebelum dilakukan pemodelan dengan Algoritma K-Nearest Neighbor akan dilakukan terlebih dahulu 6 tahapan *CRISP-DM (Cross Industry Standar for Data Mining)*, selanjutnya pada tahap akhir akan dilakukan pembungkusan model dengan pickle yang akan dibahas lebih lanjut pada bab selanjutnya. Untuk pengujian kinerja Algoritma dengan menggunakan *confusion matrix* yang juga akan dibahas pada bab selanjutnya. Alasan memilih Algoritma K-Nearest Neighbor diantaranya ialah simpel, performa cukup baik, dan cukup efektif apabila menggunakan *data training* yang jumlahnya banyak.

Berdasarkan pada pemaparan diatas oleh karena itu, penelitian tugas akhir ini berjudul: **Prediksi Minat Studi Lanjut Menggunakan Algoritma K-Nearest Neighbor.**

1.2 Rumusan Masalah

1. Bagaimana membangun aplikasi dengan model Algoritma *K-Nearest Neighbor* terbaik untuk memprediksi jenis minat studi lanjut siswa?
2. Bagaimana kinerja Algoritma *K-Nearest Neighbor* pada sistem prediksi jenis minat studi lanjut siswa?

1.3 Tujuan Masalah

1. Membangun aplikasi dengan model Algoritma K-Nearest Neighbor terbaik untuk memprediksi jenis minat studi lanjut siswa.
2. Mengetahui kinerja Algoritma *K-Nearest Neighbor* pada sistem prediksi jenis minat studi lanjut siswa.

1.4 Batasan Masalah

Ditentukan beberapa batasan masalah, diantaranya ialah:

1. Menggunakan 1601 sampel data yang diunggah oleh Abduttayyeb Rampurawala pada tahun 2020 diunduh melalui *website* <https://www.kaggle.com/datasets/abtabm/hobby-prediction-basic> 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. Digunakan dataset tersebut dikarenakan dataset tersebut didapatkan oleh penelitian sebelumnya melalui survai langsung. Untuk tipe data pada dataset diantaranya ialah categorical, boolean, dan numerical.

2. Metode yang digunakan adalah Algoritma *K-Nearest Neighbor*.
3. Dalam sistem terdapat beberapa *questioner* yang akan di jawab sehingga dari hasil jawaban tersebut akan dihasilkan prediksi minat studi yang diharapkan.
4. Hasil akhir sistem menampilkan hasil prediksi minat studi.

1.5 Kerangka Pemikiran

Berikut merupakan kerangka pemikiran tugas akhir ini, diantaranya ialah:

1. *Problem*

Banyak orang yang melanjutkan studi tidak sesuai dengan minat dan bakat dan masih banyak juga yang belum mengetahui minat dan bakat mereka masing-masing.

2. *Opportunity*

Adanya teknologi *machine learning* yang dapat memprediksi minat studi lanjut berdasarkan *questioner* yang diberikan kepada *user*.

3. *Approach*

Menggunakan Algoritma *K-Nearest Neighbor* pada sistem prediksi

4. *Software Development*

Pengembangan sistem yang digunakan ialah *jupyter notebook*, untuk Bahasa Pemograman yang digunakan ialah *python*, dan model pengembangan menggunakan *Cross-Industry Standard Process for Data Mining* (CRISP-DM).

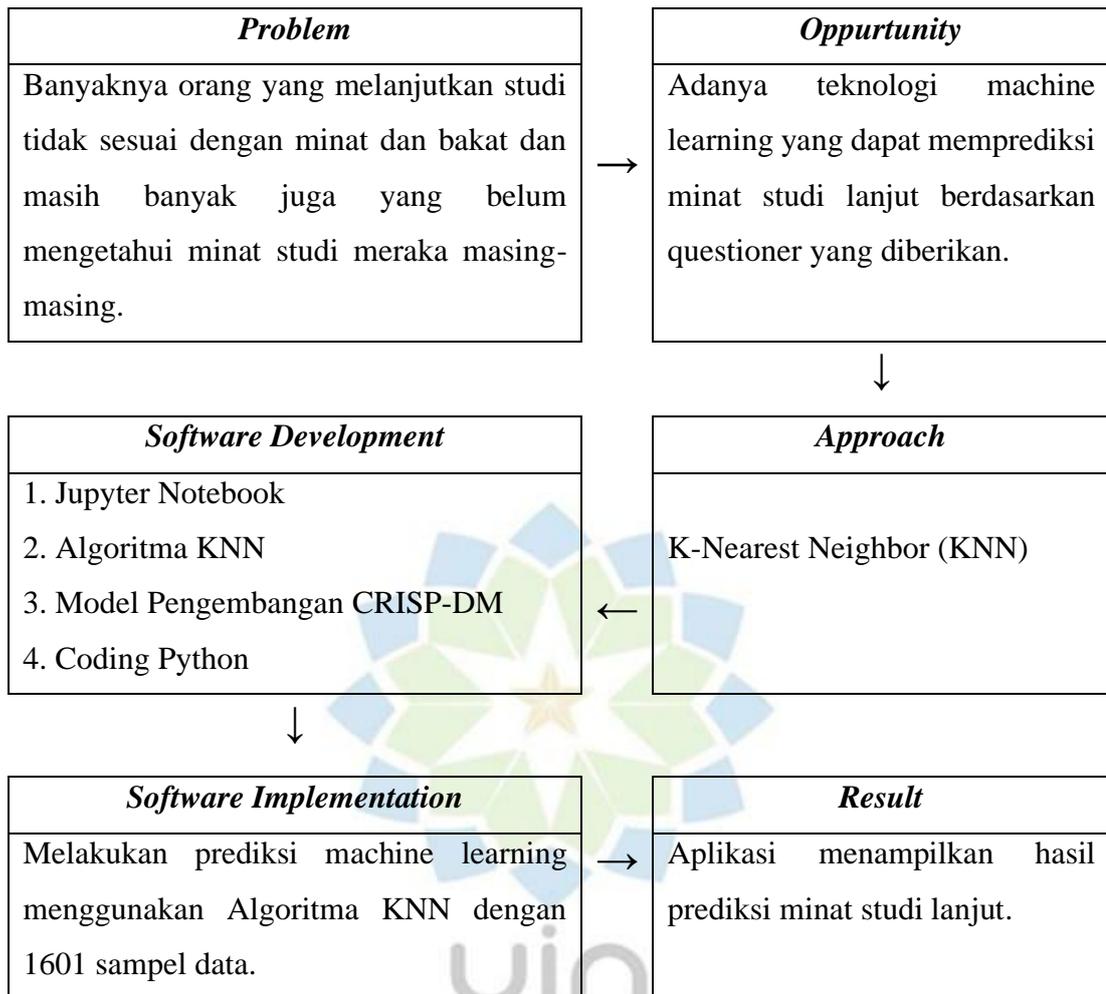
5. *Software Implementation*

Melakukan prediksi *machine learning* dengan Algoritma *K-Nearest Neighbor* menggunakan 1601 sampel data.

6. *Result*

Sistem menampilkan hasil prediksi minat studi lanjut.

Tabel 1.1 Kerangka Pemikiran



1.6 Metodologi Penelitian

1.6.1 Teknik Pengumpulan Data

1. Menggunakan data yang diunggah oleh Abduttayyeb Rampurawala melalui website <https://www.kaggle.com/datasets/abtabm/hobby-prediction-basic> pada tahun 2020 dengan nama “*Kids Hobby Prediction Dataset*”. Digunakan dataset tersebut dikarenakan dataset berformat *.csv* yang mudah untuk diolah machine learning, serta dataset tersebut didapatkan oleh penelitian sebelumnya melalui survei langsung dan juga diperuntukan untuk rentang usia remaja. Dataset tersebut berjumlah 1601 sampel data yang relevan dengan optimasi Algoritma K-Nearest Neighbor apabila menggunakan sampel data yang cukup banyak. Dataset terdiri dari 13 attributes dan 3 label dimana 13 attributes tersebut dimodelkan untuk

menghasilkan salah satu dari 3 label (*academics*, *arts*, dan *sports*). Untuk tipe data pada dataset diantaranya ialah categorical, boolean, dan numerical.

2. Studi Pustaka

Studi pustaka merupakan metode pencarian referensi dari berbagai sumber seperti buku, *e-book*, jurnal dan yang lainnya sebagai referensi dalam melakukan penelitian.

1.6.2 Metode Pengembangan Sistem

Metode pengembangan sistem yang digunakan adalah CRISP-DM (*Cross-Industry Standard Procces for Data Mining*) yang mana terdapat 6 fase pada metode tersebut, diantaranya ialah:

1. *Business Understanding Phase* (Fase Pemahaman Bisnis)
Penentuan tujuan dan kebutuhan sistem yang akan dibuat secara detail.
2. *Data Understanding Phase* (Fase Pemahaman Data)
Pengumpulan data, penganalisisan data dan pengevaluasian kualitas data.
3. *Data Preparation* (Fase Pengolahan Data)
Mempersiapkan data yang akan digunakan.
4. *Modeling Phase* (Fase Modeling)
Pengimplementasian algoritma untuk mengolah data.
5. *Impementation Phase* (Fase Penerapan)
Menerapkan sistem sesuai dengan rancangan.
6. *Evaluation Phase* (Fase Evaluasi)
Menganalisa dan menetapkan apakah hasil yang diperoleh sudah sesuai dengan tujuan pada fase awal dan menentukan apakah terdapat permasalahan yang tidak tertangani dengan baik.

1.7 Sistematika Penulisan

Untuk menyajikan gambaran serta sistematika yang cukup jelas, disusun menjadi 5 (Lima) bab dengan urutan 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 : Simpulan dan Saran

Pada bab ini berisi simpulan hasil penelitian serta saran penulis terhadap pembaca.

6. REFERENSI

7. LAMPIRAN