

BAB I

PENDAHULUAN

1.1 Latar Belakang

Berdasarkan Undang-undang No 20 Tahun 2003 tentang Sistem Pendidikan Nasional Pasal 1, pendidikan formal adalah jalur pendidikan yang terstruktur dan berjenjang yang terdiri atas pendidikan dasar, pendidikan menengah, dan pendidikan tinggi. Sekolah Menengah Atas merupakan jenjang pendidikan menengah yang terakhir ditempuh sebelum memasuki jenjang pendidikan tinggi. Dalam pendidikan formal terdapat tahap penjurusan, yaitu pengelompokan siswa berdasarkan minat, bakat, dan kemampuan dalam bidang tertentu.

Siswa SMA yang duduk dibangku kelas 12 adalah siswa remaja yang sudah berada pada usia yang harus mengambil keputusan untuk masa depannya (Bahri, 2022). Pada tahap ini, penjurusan menjadi bagian penting dalam membantu siswa untuk memilih program studi yang sesuai dengan kemampuan, minat dan bakat. Penjurusan ini bertujuan untuk membantu siswa fokus pada bidang yang sesuai dengan potensi dan minat mereka, sehingga dapat lebih siap menghadapi tantangan di jenjang perguruan tinggi. Dengan penjurusan ini memungkinkan siswa menikmati proses belajar di kemudian hari tanpa ada rasa penyesalan di kemudian hari. Berdasarkan hasil penelitian Youthmanual ditemukan fakta yakni 92% siswa SMA/SMK sederajat bingung dan tidak tahu akan memilih jurusan apa dan 45% mahasiswa merasa salah mengambil jurusan.

SMA N 1 Kuningan merupakan salah satu sekolah negeri yang saat ini memiliki proses penjurusan di kelas 12 berdasarkan nilai rapor dan sesi konseling tanpa melakukan pengujian. Sebagian siswa kelas 12 sudah menentukan jurusan namun masih banyak juga yang merasa bingung dalam memilih jurusan yang sesuai. Siswa terpengaruh oleh orang sekitar seperti tekanan dari orang tua yang berbeda pandangan dengan keinginan siswa atau hanya sekedar mengikuti teman yang menyebabkan mereka memilih jurusan

tanpa mempertimbangkan kemampuan, minat dan bakatnya sendiri. Kesalahan yang dilakukan para siswa dalam memilih jurusan dapat menimbulkan akibat yang sangat besar. Siswa akan mudah stres karena tidak memahami mata kuliah pada jurusan yang diambil sehingga mengganggu proses belajar (Bahri, 2022).

Oleh karena itu, untuk membantu mengatasi permasalahan yang dihadapi siswa dan sekolah diperlukan sebuah pendekatan berbasis teknologi yang dapat memberikan rekomendasi jurusan secara objektif dan terarah. Salah satu metode yang dapat digunakan adalah membuat sistem klasifikasi penjurusan dengan mengimplementasikan data mining menggunakan naïve bayes. Data mining adalah sebuah proses menemukan hubungan yang memiliki sebuah arti, pola dan kebiasaan dengan memilih sebagian besar data yang disimpan dalam penyimpanan dengan menggunakan pola (Mafakhir & Solichin, 2020). Metode naïve bayes ini adalah salah satu metode pengklasifikasi dengan mencari nilai probabilitas dari setiap atribut. Naïve bayes adalah klasifikasi yang ditemukan oleh ilmuwan inggris Thomas Bayes yang menggunakan metode probabilitas dan *statistic* untuk memprediksi peluang masa depan berdasarkan pengalaman masa lalu (Naparín, 2016).

Penelitian terdahulu pernah mengangkat topik mengenai pemilihan jurusan menggunakan metode Naive Bayes oleh Husni Naparin pada tahun 2016 dengan judul “Klasifikasi Peminatan Siswa SMA Menggunakan Metode Naïve Bayes” dan hasil perbandingan dapat disimpulkan dengan menggunakan algoritma naïve bayes memiliki akurasi yang sangat baik yaitu sebesar 99.47% dalam mengklasifikasikan status peminatan siswa SMA (Naparín, 2016). Dan juga penelitian oleh Nur Khafifah pada tahun 2024 yang berjudul “Penerapan Metode Naïve Bayes Untuk Penentuan Jurusan Siswa – Siswi Pada Sman 9 Kendari” dengan analisis *data training* dan *data testing* yaitu dengan perbandingan 75% dan 25% dan hasil akurasinya sebesar 55,55% yang artinya performa dari hasil klasifikasi ini dapat dikatakan tepat karena melebihi 50% (Khafifah et al., 2024).

Selanjutnya penelitian yang dilakukan oleh Detrinal Putra dan Arief Wibowo pada tahun 2020 dengan judul “Prediksi Keputusan Minat Penjurusan

Siswa SMA Yadika 5 Menggunakan Algoritma Naïve Bayes” dengan hasil prediksi penentuan peminatan jurusan dapat dilakukan dengan metode naïve bayes karena diperoleh nilai akurasi model naïve bayes sebesar 93.75%, *precision* sebesar 83.33% dan *recall* 100% (D. Putra & Wibowo, 2020).

Berdasarkan uraian di atas, penulis tertarik untuk melakukan penelitian guna membantu siswa Sekolah Menengah Atas Negeri 1 Kuningan dalam menentukan jurusan yang sesuai untuk melanjutkan ke perguruan tinggi. Penelitian ini dibuat dengan mengambil judul: **“Implementasi Data Mining dengan Algoritma Naïve Bayes untuk Klasifikasi Siswa Sekolah Menengah Atas dalam Penentuan Jurusan Perguruan Tinggi.”**

1.2 Identifikasi Masalah

Identifikasi masalah adalah proses penting untuk menemukan masalah yang berada pada objek penelitian yang diteliti, penulis mengidentifikasi beberapa permasalahan sebagai berikut:

1. Banyak siswa mengalami kebingungan saat mengambil keputusan karena kurangnya pemahaman siswa terhadap kemampuan, minat, dan bakat mereka sehingga menimbulkan penyesalan atas keputusan yang diambil.
2. Proses penjurusan di sekolah sering hanya didasarkan pada nilai akademik dan sesi konseling tanpa adanya pengujian.
3. Keputusan siswa sering dipengaruhi oleh faktor eksternal seperti tekanan dari orang tua atau teman, yang tidak selalu sesuai dengan kemampuan atau minat siswa.
4. Kesalahan dalam pengambilan jurusan menyebabkan terganggunya proses belajar dikarenakan siswa mengalami stres karena tidak memahami mata kuliah pada jurusan yang diambil.

1.3 Rumusan Masalah

Dari uraian yang telah dipaparkan pada latar belakang, maka permasalahannya dirumuskan sebagai berikut:

1. Bagaimana merancang dan membangun sistem berbasis web untuk membantu siswa SMAN 1 Kuningan dalam memilih jurusan di perguruan tinggi?

2. Bagaimana mengimplementasikan data mining dan algoritma naive bayes untuk mengklasifikasikan siswa SMAN 1 Kuningan dalam menentukan jurusan di perguruan tinggi?

1.4 Batasan Masalah

Agar pembahasan lebih jelas dan terarah maka diperlukan adanya batasan masalah. Adapun batasan masalah yaitu sebagai berikut:

1. Metode yang digunakan yaitu data mining dengan menggunakan algoritma naïve bayes.
2. Sistem yang dibangun adalah aplikasi berbasis web dengan bahasa pemrograman PHP, *framework* bootstrap dan MySQL sebagai *database* nya.
3. Sistem digunakan oleh guru BK dan siswa.
4. Data yang digunakan adalah data kelas 12.
5. Fitur aplikasi terdiri dari input dataset yang dilakukan oleh guru BK, proses klasifikasi yang dilakukan oleh sistem, dan hasil dari klasifikasi berserta probabilitasnya.
6. Kriteria klasifikasi penjurusan terdiri dari asal jurusan, nilai mata pelajaran, minat siswa dan 7 bidang keilmuan.
7. Kriteria nilai mata pelajaran yang digunakan dikelompokkan menjadi dua kategori, yaitu “Tinggi” dan “Rendah” berdasarkan pada nilai yang dihasilkan dari data historis siswa.
8. Bidang keilmuan yang akan di klasifikasikan yaitu, Kesehatan, Teknik dan Teknologi, Manajemen dan Ekonomi, Pendidikan, Seni dan Desain, Hukum dan Ilmu Sosial, Ilmu alam dan Sains.

1.5 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan, maka tujuan dalam penelitian ini adalah sebagai berikut:

1. Merancang dan membangun sistem berbasis web untuk membantu siswa SMAN 1 Kuningan dalam memilih jurusan di perguruan tinggi.
2. Mengimplementasikan data mining dan algoritma naive bayes untuk mengklasifikasikan siswa SMAN 1 Kuningan dalam menentukan jurusan di perguruan tinggi.

1.6 Manfaat Penelitian

Adapun hasil penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Manfaat Teoritis

Hasil penelitian ini diharapkan dapat menambah referensi ilmiah tentang implementasi data mining dengan metode naive bayes dalam bidang pendidikan, khususnya dalam hal klasifikasi, bermanfaat juga dalam pengembangan aplikasi berbasis web dan dapat dijadikan referensi bagi penelitian-penelitian selanjutnya.

2. Manfaat Praktis

a. Bagi Peneliti

Meningkatkan pengetahuan dan keterampilan mengenai Data Mining dan Algoritma Naive Bayes dalam implementasinya pada klasifikasi siswa Sekolah Menengah Atas dalam penentuan jurusan di perguruan tinggi.

b. Bagi Sekolah

Mendukung tugas guru Bimbingan dan Konseling (BK) dalam memberikan arahan penjurusan yang lebih akurat dan berbasis data. Sistem ini berfungsi sebagai alat bantu dalam proses pengambilan keputusan yang lebih efisien, objektif, dan terukur.

c. Bagi Siswa

Membantu siswa memilih jurusan dengan sistem rekomendasi yang terarah karena berdasarkan data yang relevan dan terukur serta mengurangi kebingungan dalam memilih jurusan.

1.7 Pertanyaan Penelitian

Berdasarkan uraian dari rumusan masalah, tujuan serta manfaat diatas, maka pertanyaan dari penelitian ini diantaranya:

1. Apakah sistem yang dibuat dapat membantu siswa SMAN 1 Kuningan dalam menentukan jurusan perguruan tinggi?
2. Apakah algoritma naive bayes dapat digunakan untuk mengklasifikasikan siswa SMAN 1 Kuningan dalam menentukan jurusan di perguruan tinggi ?

1.8 Hipotesis Penelitian

Adapun hipotesis dalam penelitian ini adalah sebagai berikut:

1. Sistem yang dibuat dapat membantu siswa SMAN 1 Kuningan dalam menentukan jurusan perguruan tinggi.
2. Implementasi algoritma naïve bayes dapat digunakan untuk mengklasifikasikan siswa SMAN 1 Kuningan dalam menentukan jurusan di perguruan tinggi.

1.9 Metodologi Penelitian

1.9.1 Metode Pengumpulan Data

Pada penelitian ini terdapat beberapa metode yang digunakan untuk memperoleh data yang dibutuhkan, antara lain: observasi, wawancara dan studi pustaka. Penjelasan dari metode-metode tersebut akan diuraikan sebagai berikut:

1. Observasi

Pada metode observasi, dilakukan pengumpulan data dengan pengamatan langsung di SMA N 1 Kuningan untuk mengetahui lebih jelas lagi mengenai sistem yang sedang berjalan.

2. Wawancara

Pada metode wawancara, dilakukan wawancara dengan Ibu Yuliza dan Ibu Ina selaku guru BK. Wawancara dilakukan untuk mengetahui lebih jauh mengenai proses dan permasalahan yang dihadapi mengenai penjurusan untuk kelas 12.

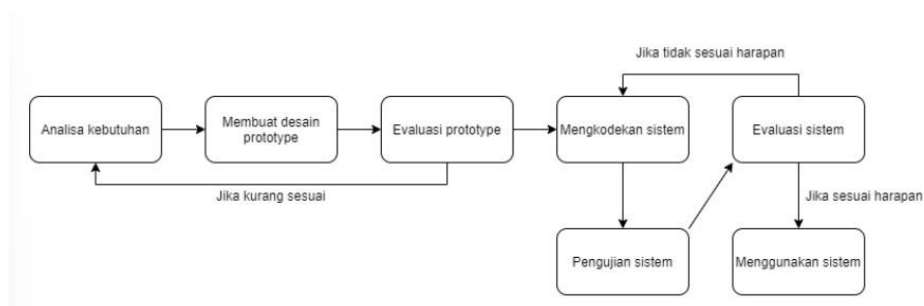
3. Studi Pustaka

Pada metode studi pustaka, menggunakan sumber-sumber seperti buku, jurnal, dan internet untuk memperoleh data penelitian. Studi litelatur ini berguna untuk mengetahui landasan teori, pengetahuan dan informasi pada penelitian ini dengan jurnal referensi yang relevan dengan topik penelitian yaitu penentuan atau rekomendasi jurusan dengan metode Naïve Bayes.

1.9.2 Metode Pengembangan Sistem

Metode *prototype* adalah metode yang memungkinkan pengguna atau *user* memiliki gambaran awal tentang perangkat lunak yang akan dikembangkan, serta pengguna dapat melakukan pengujian di awal sebelum perangkat lunak dirilis.

Tujuan dari metode *prototype* adalah mengumpulkan data informasi dari pengguna sehingga sistem dapat merespon keluhan pengguna. Pengembangan sistem informasi ini meliputi perancangan prototipe proses, perancangan basis data dan antarmuka (Al Muhtadi & Junaedi, 2021).



Gambar 1. 1 Metode Prototype (Al Muhtadi & Junaedi, 2021)

Pada gambar 1.1 terdapat 7 tahapan metode prototype, penjelasan masing-masing tahapan akan dijelaskan sebagai berikut:

1. Analisa kebutuhan

Di tahap ini, pengembang mengidentifikasi perangkat lunak yang akan dibuat dan persyaratan sistem. Dalam pengembangan perangkat lunak, diadakan pertemuan dengan para pemangku kepentingan, terutama dengan bimbingan dan saran dari guru dan siswa, untuk menentukan tujuan keseluruhan dari perangkat lunak yang akan dikembangkan dan untuk meminta semua spesifikasi kebutuhan yang diketahui saat ini.

2. Membuat *prototype*

Membuat rancangan sementara yang berfokus pada alur program kepada pengguna.

3. Evaluasi *prototype*

Evaluasi dilakukan untuk mengetahui apakah model *prototype* sudah sesuai dengan harapan. Evaluasi dilakukan dengan pemangku kepentingan agar mendapatkan umpan balik yang sesuai dengan kebutuhan pengguna.

4. Mengkodekan sistem

Jika model *prototype* telah disetujui maka akan diimplementasikan ke dalam bahasa pemrograman yang sesuai. Bahasa pemrograman yang digunakan yaitu PHP dengan *framework* bootstrap.

5. Pengujian sistem

Setelah perangkat lunak sudah siap, perangkat lunak harus melewati pengujian. Pengujian ini biasanya dilakukan dengan *White Box Testing*, *Black Box Testing*, dan lain-lain.

6. Evaluasi sistem

Pengguna melakukan evaluasi apakah perangkat lunak sudah sesuai dengan apa yang diharapkan atau tidak. Jika ya, lakukan tahap selanjutnya. Jika tidak, ulangi tahap mengkodekan sistem dan pengujian sistem.

7. Menggunakan sistem

Perangkat lunak yang telah diuji dan disetujui siap untuk digunakan oleh pengguna.

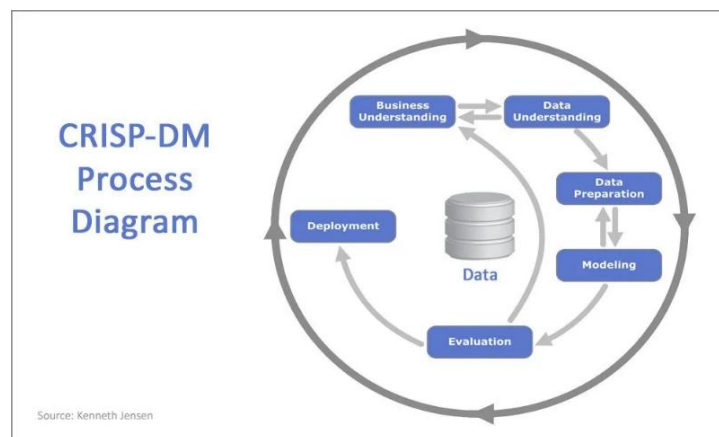
1.9.3 Metode Penyelesaian Masalah

Metode penyelesaian masalah yang digunakan dalam penelitian ini adalah Metode Data Mining dengan menggunakan Teknik klasifikasi. Adapun Teknik klasifikasi yang digunakan yaitu metode naïve bayes.

1. Data Mining

Data mining adalah metode pemrosesan data dalam jumlah besar untuk menemukan pola dan pengetahuan tersembunyi untuk menginformasikan proses pengambilan keputusan (M. Y. Putra & Putri, 2022).

Dalam data mining terdapat model teknik analisis data yang bisa digunakan, salah satunya yaitu model CRISP-DM (*Cross Industry Standard Process for Data Mining*). CRISP-DM (*Cross-Industry Standard Process for Data Mining*) diperkenalkan pada akhir 1990-an oleh empat perusahaan: Integral Solutions Ltd. (penyedia solusi data mining), NCR (penyedia *database*), DaimlerChrysler (produsen mobil), dan OHRA (perusahaan asuransi). Model ini ditetapkan sebagai standar proses untuk menyusun strategi pemecahan masalah dalam data mining (Pambudi, 2023).



Gambar 1. 2 Model CRISP-DM (Kenneth Jensen)

2. Teknik Klasifikasi

Klasifikasi adalah metode pengelompokan objek berdasarkan ciri-ciri yang dimiliki objek tersebut. Klasifikasi dapat dilakukan dengan berbagai cara, secara manual atau dengan bantuan teknologi. Klasifikasi secara manual adalah klasifikasi yang dilakukan oleh manusia tanpa bantuan algoritma komputer cerdas, sedangkan klasifikasi dilakukan dengan menggunakan teknologi mempunyai beberapa algoritma antara lain Naive Bayes, Support Vector Machines, Decision Trees, Fuzzy dan Artificial Neural Networks (Wibawa et al., 2018).

3. Naive Bayes Classifier

Naive Bayes Classifier adalah suatu teknik yang mengklasifikasikan berdasarkan persamaan dan perbedaan dengan menggunakan statistik yang dapat memprediksi kemungkinan sebuah kelas (Lestari et al., 2020).

Metode pengklasifikasian naive bayes dengan menggunakan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes. Ciri utama dari *Naive Bayes Classifier* ini adalah asumsi yang sangat kuat (naif) akan independensi dari masing-masing kondisi / kejadian.

Naive Bayes Classifier merupakan metode klasifikasi berdasarkan probabilitas dan teorema bayes dengan menggunakan asumsi bahwa setiap variabel bebas (independen) dan asumsi bahwa sebuah fitur tidak ada kaitannya dengan fitur lain. Dengan asumsi keindependenan atribut menghilangkan kebutuhan akan banyaknya jumlah data latih dari perkalian kartesius seluruh atribut yang dibutuhkan untuk mengklasifikasikan sebuah data.

Metode klasifikasi Naïve Bayes termasuk metode klasik yang menggunakan konsep statistik probabilitas dalam melakukan klasifikasi data. Naïve Bayes menggunakan teorema probabilitas dengan formula dasar menggunakan persamaan sebagai berikut :

$$P(C|X) = \frac{P(X|C)P(C)}{P(X)} \dots(1)$$

Keterangan:

- **P (C|X)**: Probabilitas bahwa suatu data X termasuk ke dalam kelas C (disebut *posterior probability*).
- **P (X|C)**: Probabilitas mendapatkan data X jika sudah diketahui bahwa data tersebut berasal dari kelas C (disebut *likelihood*).
- **P(C)**: Probabilitas awal (sebelum melihat data X) bahwa kelas tersebut adalah C (disebut *prior probability*).
- **P(X)**: Probabilitas terjadinya data X di seluruh kemungkinan kelas.

1.10 Sistematika Penelitian

Sistematika penulisan tugas akhir disusun untuk memberikan gambaran umum tentang penelitian yang dijalankan. Sistematika penulisan akhir sebagai berikut:

BAB I : PENDAHULUAN

Pada bab ini menjelaskan mengenai latar belakang, identifikasi masalah, rumusan masalah, Batasan masalah, tujuan penelitian, manfaat penelitian, pertanyaan penelitian, metodologi penelitian, dan sistematika penulisan.

BAB II : LANDASAN TEORITIS

Pada bab ini menjelaskan mengenai teori-teori yang berkaitan dengan pengetahuan dasar yang dibutuhkan dalam pembahasan pembuatan sistem dengan menggunakan PHP dan MySQL.

BAB III : ANALISA DAN PERANCANGAN

Pada bab ini menjelaskan mengenai perancangan yang digunakan dalam membangun sistem mulai dari analisis sistem yang sedang berjalan, analisis permasalahan dan kebutuhan sistem, perencanaan sistem, serta rancangan input dan output pada program.

BAB IV : IMPLEMENTASI DAN PENGUJIAN

Pada bab ini menjelaskan mengenai tampilan input dan output sistem, implementasi dan pengujian sistem yang telah dibangun.

BAB V : KESIMPULAN DAN SARAN

Pada bab ini berisi kesimpulan bagi sistem yang sudah dibangun secara keseluruhan dan dikemukakan saran-saran untuk perbaikan dan pengembangan sistem.

