

BAB II

KAJIAN PUSTAKA

A. Kajian Teoritis

1. Sistem Pendukung Keputusan (SPK)

Sistem Pendukung Keputusan SPK merupakan suatu sistem informasi yang dirancang dengan bantuan komputer. Tujuan dari sistem ini adalah untuk mendukung proses pengambilan keputusan. SPK melakukan hal ini dengan memanfaatkan berbagai jenis data, teknologi informasi, serta model perhitungan dan analisis yang spesifik. Dengan demikian, SPK berperan penting dalam meningkatkan efektivitas dalam pengambilan keputusan (Hartono, 2025:156).

Sistem Pendukung Keputusan SPK merupakan salah satu solusi yang berpotensi, karena mampu menawarkan proses analisis yang terorganisir, konsisten, dan mudah digunakan dengan bantuan komputer. SPK adalah komponen dari sistem informasi berbasis komputer yang dikembangkan untuk membantu proses pengambilan keputusan dalam suatu organisasi dengan tampilan yang ramah pengguna (Veronika & Yusefin, 2025:876).

Dengan demikian, SPK dapat disimpulkan sebagai sistem yang membantu proses pengambilan keputusan secara efektif, terutama dalam situasi yang melibatkan banyak kriteria dan pertimbangan. Penerapan SPK berbasis website dengan metode *Naive Bayes* dalam penelitian ini diharapkan

mampu meningkatkan objektivitas, akurasi, dan efisiensi penentuan siswa berbakat di SDN 2 Ngadirejo.

2. Siswa Berbakat

Siswa berbakat merupakan siswa yang memiliki kemampuan atau potensi yang lebih tinggi dibandingkan dengan siswa lainnya dalam bidang akademik maupun non-akademik. Potensi tersebut dapat terlihat dari kemampuan berpikir, kreativitas, prestasi belajar, serta keaktifan dalam kegiatan sekolah (Kadek et al., 2023:154).

Identifikasi siswa berbakat sangat penting dilakukan oleh pihak sekolah agar potensi yang dimiliki oleh siswa dapat dikembangkan secara optimal. Dengan adanya proses identifikasi yang tepat, sekolah dapat memberikan program pembinaan yang sesuai dengan kemampuan dan minat siswa (Pardosi et al., 2025:1265).

Berdasarkan uraian tersebut, dapat disimpulkan bahwa siswa berbakat adalah individu yang memiliki potensi unggul yang perlu diidentifikasi dan dikembangkan secara tepat melalui proses yang sistematis. Identifikasi yang akurat tidak hanya membantu sekolah dalam memberikan pembinaan yang sesuai, tetapi juga memastikan bahwa setiap siswa mendapatkan kesempatan untuk mengembangkan kemampuan terbaiknya secara optimal. Oleh karena itu, diperlukan suatu metode atau sistem yang objektif dan terstruktur agar proses penentuan siswa berbakat dapat dilakukan secara adil, efektif, dan dapat dipertanggungjawabkan.

3. *Naive Bayes*

Naive Bayes merupakan metode klasifikasi pada *data mining* yang menerapkan pendekatan probabilitas dan statistik berdasarkan *Teorema Bayes*. Metode ini digunakan untuk mengklasifikasikan data ke dalam kelas tertentu sesuai atribut atau variabel yang dimiliki (Alfiah et al., 2026:88).

Metode *Naive Bayes* melakukan klasifikasi dengan menghitung probabilitas setiap kelas berdasarkan data pelatihan yang tersedia. Selanjutnya, kelas yang memiliki probabilitas terbesar dipilih sebagai hasil klasifikasi (Lubis et al., 2025:571). Dasar perhitungan pada metode *Naive Bayes* mengacu pada *Teorema Bayes* pada persamaan 2.1 di tunjukkan cara perumusannya.

$$P(C|X) = \frac{P(X|C) \times P(C)}{P(X)} \quad (2.1)$$

Metode *Naive Bayes* menggunakan rumus seperti persamaan diatas, X merupakan data atau atribut yang akan diklasifikasikan dan kelasnya belum diketahui. C merupakan hipotesis atau kelas tujuan yang akan diprediksi, misalnya kelas Berbakat atau Tidak Berbakat. $P(C)$ adalah probabilitas awal (prior probability), yaitu peluang suatu kelas sebelum mempertimbangkan data atau atribut yang dimiliki. $P(X)$ merupakan probabilitas kemunculan data atau atribut secara keseluruhan (*evidence*). Selanjutnya, $P(C|X)$ merupakan probabilitas posterior, yaitu peluang suatu data termasuk ke dalam kelas C setelah diketahui atribut X . Adapun $P(X|C)$ merupakan probabilitas likelihood, yaitu peluang munculnya atribut X dengan asumsi data tersebut berada pada kelas C . Berdasarkan komponen-komponen tersebut, metode

Naive Bayes menghitung probabilitas setiap kelas, kemudian memilih kelas yang memiliki nilai probabilitas posterior terbesar sebagai hasil klasifikasi.

Menurut Hamzanwadi et al., (2024:97) persamaan tersebut merupakan bentuk lain dari Teorema Bayes yang digunakan dalam proses perhitungan probabilitas pada metode *Naive Bayes*. Pada persamaan 2.2 ditunjukkan cara perumusannya.

$$Posterior = \frac{Likelihood \times Prior}{Evidence} \quad (2.2)$$

Pada persamaan di atas menunjukkan bahwa nilai posterior diperoleh dari hasil perkalian antara likelihood dan prior, kemudian dibagi dengan evidence. Nilai posterior merupakan probabilitas akhir yang digunakan sebagai dasar dalam menentukan kelas suatu data. Pada metode *Naive Bayes*, kelas yang memiliki nilai posterior terbesar akan dipilih sebagai hasil klasifikasi.

Menurut Salsabila & Basri, (2026:245) perhitungan *Naive Bayes* dapat dilakukan melalui beberapa tahapan. Nilai prior digunakan untuk mengetahui peluang awal dari masing-masing kelas dengan rumus. Tahap pertama adalah menghitung nilai prior yang dimana pada persamaan 2.3 di tunjukkan cara perhitungannya.

$$P(C) = \frac{X}{A} \quad (2.3)$$

Pada persamaan diatas, $P(C)$ merupakan probabilitas awal (*prior*) dari suatu kelas sebelum mempertimbangkan atribut yang dimiliki oleh data. X menyatakan jumlah data pada masing-masing kelas, sedangkan A merupakan jumlah keseluruhan data yang digunakan sebagai data pelatihan. Nilai prior

digunakan untuk mengetahui peluang awal dari setiap kelas berdasarkan distribusi data yang tersedia.

Menurut Wisnu et al., (2023:200) tahap berikutnya adalah menghitung nilai likelihood. Nilai likelihood digunakan untuk mengetahui peluang kemunculan suatu atribut pada kelas tertentu. Pada persamaan 2.4 ditunjukkan cara perhitungannya.

$$P(X|C) = \frac{F}{B} \quad (2.4)$$

Pada persamaan diatas, $P(X|C)$ merupakan probabilitas *likelihood*, yaitu peluang munculnya suatu atribut X pada kelas C. F menyatakan banyaknya data fitur yang muncul pada setiap kelas, sedangkan B merupakan jumlah keseluruhan data fitur pada kelas tersebut. Nilai *likelihood* digunakan untuk mengetahui seberapa besar kemungkinan suatu atribut muncul pada kelas tertentu berdasarkan data pelatihan yang tersedia.

4. Evaluasi *Confusion Matrix*

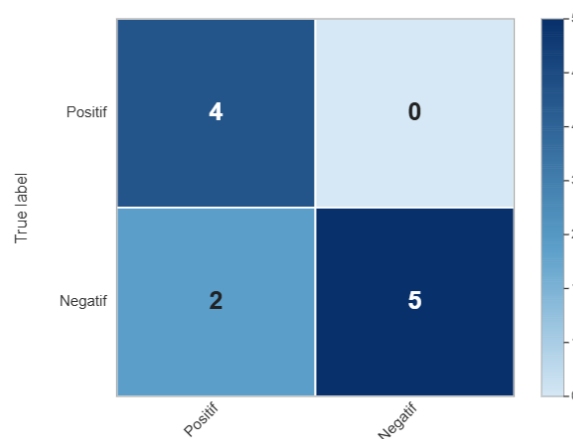
Confusion Matrix merupakan metode evaluasi yang digunakan untuk mengukur kinerja model klasifikasi dengan membandingkan hasil prediksi yang dilakukan sistem terhadap kondisi sebenarnya. Melalui *Confusion Matrix* dapat diketahui jumlah prediksi yang benar maupun prediksi yang salah sehingga dapat dihitung nilai akurasi, precision, recall, dan F1-Score. Metode ini banyak digunakan untuk mengevaluasi performa algoritma klasifikasi, termasuk metode *Naive Bayes* (Valerian et al., 2025:2244). Dapat dilihat seperti pada tabel 2.1 *Confusion matrix*

Tabel 2.1 *Confusion matrix*

Aktual / Prediksi	Positif	Negatif
Positif	True Positive (TP)	False Negative (FN)
Negatif	False Positive (FP)	True Negative (TN)

Sumber: (Valerian et al., 2025:2244)

Confusion Matrix divisualisasikan dalam bentuk heatmap yang menunjukkan jumlah prediksi benar dan prediksi salah pada setiap kelas. Nilai pada diagonal utama menunjukkan jumlah data yang berhasil diklasifikasikan dengan benar, sedangkan nilai di luar diagonal menunjukkan jumlah kesalahan klasifikasi. Visualisasi ini memudahkan dalam mengevaluasi kinerja model sebelum dilakukan perhitungan accuracy, precision, recall, dan F1-Score (Rininda et al., 2023:3341). Visualisasi *confusion matrix* dapat dilihat seperti pada gambar 2.1

Gambar 2.1 Visualisasi *confusion matrix*

Sumber: (Rininda et al., 2023:3341)

Nilai accuracy digunakan untuk mengetahui tingkat ketepatan model dalam mengklasifikasikan data berdasarkan jumlah prediksi yang benar

dibandingkan dengan seluruh data pengujian (Azmi & Voutama, 2024:116).

Pada persamaan 2.5 ditunjukkan cara perhitungannya.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (2.5)$$

Pada persamaan diatas, TP menunjukkan jumlah data positif yang diprediksi dengan benar, TN menunjukkan jumlah data negatif yang diprediksi dengan benar, FP menunjukkan jumlah data negatif yang diprediksi sebagai positif, sedangkan FN menunjukkan jumlah data positif yang diprediksi sebagai negatif. Nilai accuracy menunjukkan persentase keseluruhan prediksi yang berhasil dilakukan dengan benar oleh model klasifikasi.

Nilai precision digunakan untuk mengetahui tingkat ketepatan prediksi terhadap data yang diprediksi sebagai kelas positif (Hidayah, 2024:11). Pada persamaan 2.6 ditunjukkan cara perhitungannya.

$$\text{Precision} = \frac{TP}{TP + FP} \times 100\% \quad (2.6)$$

Pada persamaan diatas, nilai precision diperoleh dari perbandingan antara jumlah True Positive (TP) dengan jumlah seluruh data yang diprediksi sebagai positif, yaitu TP dan FP. Semakin tinggi nilai precision, semakin kecil kemungkinan model menghasilkan prediksi positif yang salah.

Nilai recall digunakan untuk mengetahui kemampuan model dalam menemukan seluruh data yang benar-benar termasuk ke dalam kelas positif (Kamil et al., 2022:84). Pada persamaan 2.7 ditunjukkan cara perhitungannya.

$$\text{Recall} = \frac{TP}{TP + FN} \times 100\% \quad (2.7)$$

Pada persamaan diatas, nilai recall dihitung berdasarkan perbandingan antara jumlah True Positive (TP) dengan jumlah seluruh data yang sebenarnya termasuk kelas positif, yaitu TP dan FN. Nilai recall yang tinggi menunjukkan bahwa model mampu mengenali sebagian besar data positif dengan baik.

Nilai F1-Score digunakan untuk mengetahui keseimbangan antara nilai precision dan recall dalam mengevaluasi kinerja model klasifikasi (Normawati & Prayogi, 2021:702). Pada persamaan 2.8 ditunjukkan cara perhitungannya.

$$\text{F1-Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (2.8)$$

Pada persamaan diatas, nilai F1-Score diperoleh dari rata-rata harmonis antara nilai precision dan recall. Metrik ini digunakan untuk memberikan gambaran yang lebih seimbang terhadap performa model klasifikasi, terutama ketika jumlah data pada setiap kelas tidak seimbang.

5. *Extreme Programming (XP)*

Extreme Programming (XP) adalah metodologi pengembangan perangkat lunak yang ditujukan untuk meningkatkan kualitas perangkat lunak dan tanggap terhadap perubahan kebutuhan pelanggan. Jenis pengembangan perangkat lunak semacam ini dimaksudkan untuk meningkatkan produktivitas dan memperkenalkan pos pemeriksaan dimana persyaratan pelanggan baru dapat diadopsi. Tahapan *Extreme Programming* terdiri dari planning seperti memahami kriteria pengguna dan perencanaan pengembangan, designing seperti perancangan prototype dan tampilan, coding termasuk pengintegrasian, dan terakhir testing (Septilia, 2021:36).

Pada seperti gambar 2.7 Berikut adalah penjelasan dari gambar tahapan – tahapan metode *Extreme Programming*:

1. *Planning*

Pada tahap *Planning*, peneliti bersama dengan guru dan kepala sekolah SDN 2 Ngadirejo melakukan perencanaan kebutuhan sistem secara iteratif. Tahap ini meliputi pembuatan user stories, penentuan prioritas fitur, serta kesepakatan mengenai kriteria penilaian siswa berbakat (nilai akademik, keaktifan, kedisiplinan, dan prestasi). Perencanaan dilakukan secara fleksibel sehingga dapat menyesuaikan dengan masukan pengguna di setiap iterasi.

2. *Design*

Tahap *Design* dilakukan untuk merancang arsitektur sistem, basis data, antarmuka website, serta alur proses klasifikasi menggunakan metode *Naive Bayes*. Perancangan dilakukan menggunakan *Data Flow Diagram* (DFD), *Entity Relationship Diagram* (ERD), *Flowchart*, dan desain antarmuka (*user interface*) sebagai acuan dalam pengembangan sistem.

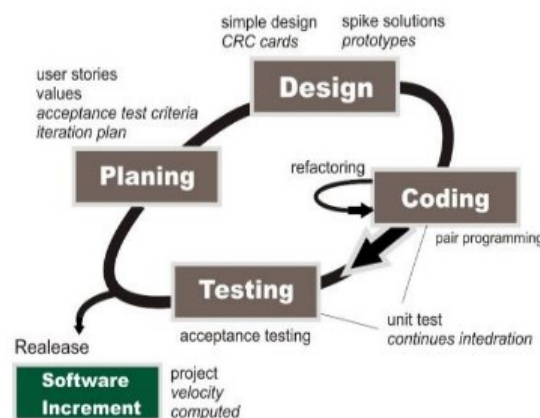
3. *Coding*

Pada tahap *Coding*, peneliti mengimplementasikan kode program berdasarkan desain yang telah disepakati. Sistem dikembangkan menggunakan bahasa pemrograman *PHP* dan database *MySQL*. Fitur yang dibuat meliputi login admin, pengelolaan data siswa, pengelolaan kriteria dan bobot, proses perhitungan probabilitas

Naive Bayes, serta tampilan hasil klasifikasi siswa berbakat atau tidak berbakat. Coding dilakukan secara incremental dan terus diuji pada setiap iterasi kecil.

4. *Testing*

Tahap *Testing* dilakukan secara terus-menerus dan intensif setelah setiap bagian kode selesai dibuat. Pengujian meliputi unit testing, integration testing, pengujian fungsional sistem, pengujian akurasi metode *Naive Bayes* menggunakan *confusion matrix*, serta user acceptance testing yang melibatkan guru dan kepala sekolah SDN 2 Ngadirejo. Hasil pengujian digunakan untuk segera memperbaiki kekurangan sehingga kualitas sistem terus meningkat di setiap iterasi. Dapat dilihat pada gambar 2.2.



Gambar 2.2 *Extreme Programming*

Sumber:(Septilia, 2021:36)

6. Website

Website merupakan sekumpulan laman *web* yang terkoneksi satu sama lain dan dapat diakses melalui browser dengan memanfaatkan jaringan internet. Dalam penelitian ini, sistem dikembangkan menggunakan *PHP*

Native sebagai bahasa pemrograman untuk mengelola logika aplikasi (*backend*) serta *MySQL* untuk sistem manajemen berbasis data (*database*) yang digunakan sebagai penyimpanan data klasifikasi (Rudjiono & Saputro, 2022:58).

Terdapat berbagai macam aplikasi yang dikenal sebagai browser web, yang berfungsi untuk mempermudah akses ke *World Wide Web*. Contohnya termasuk Firefox, Microsoft Internet Explorer, Chrome, dan lainnya. Melalui browser ini, pengguna dapat memanfaatkan berbagai layanan dan informasi yang tersedia di internet (Asmara, 2022:3).

Berdasarkan uraian tersebut, dapat disimpulkan bahwa *website* merupakan platform digital yang berfungsi sebagai media penyampaian informasi yang terkoneksi satu sama lain dan mudah dibuka oleh pengguna secara daring(*online*). *Website* tidak hanya menyajikan informasi secara statis, tetapi juga mampu memberikan layanan interaktif yang mendukung berbagai kebutuhan pengguna. Karena sebab itu, *website* sangat efektif digunakan untuk sarana implementasi sistem informasi, termasuk dalam pengembangan sistem penentuan siswa berbakat berbasis web.

7. *PHP*

PHP adalah skrip yang terintegrasi dengan *HTML* dan dijalankan di server (pemrograman skrip *HTML* yang tertanam di sisi server) *PHP* berfungsi sebagai skrip untuk pemrograman server web, yang memungkinkan pembuatan dokumen *HTML* secara langsung, yang

berarti dokumen *HTML* tersebut dihasilkan oleh aplikasi dan bukan dibuat menggunakan editor teks atau editor *HTML* (Pasaribu & Susanti, 2021:31).

PHP dikenal sebagai bahasa pemrograman skrip yang terintegrasi dengan tag *HTML*, dijalankan di server, dan digunakan untuk menghasilkan halaman web yang dinamis, serupa dengan *Active Server Pages* (ASP) atau *Java Server Pages* (JSP). *PHP* adalah perangkat lunak dengan lisensi Open Source (Hermiati et al., 2021:55).

Berdasarkan uraian tersebut, dapat disimpulkan bahwa *PHP* adalah bahasa pemrograman yang efektif dalam pengembangan aplikasi berbasis web karena memiliki kemampuan dalam membangun sistem yang dinamis dan terintegrasi dengan basis data. Selain itu, sifatnya yang open-source dan fleksibel menjadikan *PHP* mudah digunakan serta banyak diterapkan dalam berbagai jenis sistem informasi. Oleh karena itu, *PHP* sangat sesuai digunakan sebagai teknologi dalam pengembangan sistem penentuan siswa berbakat berbasis *website*.

8. *MySQL*

MySQL adalah perangkat lunak yang termasuk dalam kategori DBMS (Sistem Manajemen Basis Data) yang bersifat Open Source. Open source menunjukkan bahwa perangkat lunak ini dilengkapi dengan kode sumber (*source code*) yang digunakan untuk membuat *MySQL*, selain juga tersedia dalam bentuk executable atau kode yang dapat langsung dijalankan di sistem operasi (Utami, 2022:155).

MySQL dipilih karena kemampuan mengelola dan memproses data secara efisien, yang sangat penting untuk kebutuhan industri. Dalam kegiatan ini, digunakan pendekatan kualitatif dengan metode studi kasus, yang berfokus pada penerapan *MySQL* dalam pengolahan data siswa. Data dikumpulkan melalui wawancara, observasi, dan dokumentasi untuk menganalisis keefektifan *MySQL* dalam operasional sekolah (Siregar et al., 2024:229).

Berdasarkan uraian tersebut, dapat disimpulkan bahwa *MySQL* merupakan sistem manajemen basis data yang handal dalam mengelola data secara terstruktur dan efisien. Kemampuannya dalam mendukung pengolahan data dalam jumlah besar serta akses multi-user menjadikan *MySQL* sangat cocok digunakan dalam pengembangan sistem berbasis web. Oleh karena itu, *MySQL* dapat diandalkan sebagai media penyimpanan data dalam sistem penentuan siswa berbakat agar proses pengolahan dan pengambilan data dapat berjalan dengan cepat dan akurat.

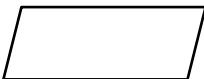
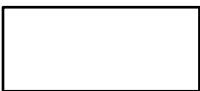
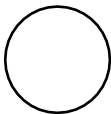
9. *Flowchart*

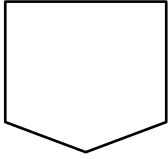
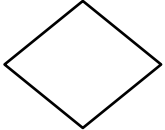
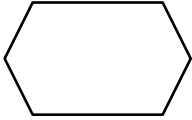
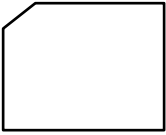
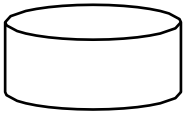
Flowchart merupakan diagram alur yang digunakan untuk menggambarkan urutan proses suatu sistem secara logis dan terstruktur menggunakan simbol-simbol tertentu. *Flowchart* berfungsi untuk memvisualisasikan langkah-langkah proses sehingga memudahkan dalam memahami alur kerja sistem yang akan dibangun. Dalam penelitian sebelumnya dijelaskan bahwa *flowchart* berperan penting dalam menerjemahkan proses berjalannya sebuah program atau sistem, mulai dari

input data, proses pengolahan, hingga menghasilkan output yang diinginkan (Tuasamu et al., 2023:499).

Dalam konteks sistem penentuan siswa berbakat menggunakan metode *Naive Bayes*, *flowchart* difungsikan sebagai penjabar alur sistem mulai dari proses login pengguna (*admin* atau *user*), penginputan data siswa dan kriteria, proses perhitungan menggunakan metode *Naive Bayes*, hingga menghasilkan keputusan berupa klasifikasi siswa berbakat atau tidak. Berdasarkan penelitian yang ada, alur sistem dimulai dari input data, dilanjutkan dengan proses pengolahan dan klasifikasi, kemudian sistem akan menampilkan hasil keputusan sebagai output (Listyoningrum et al., 2023:103). Penggunaan *flowchart* dalam perancangan sistem membantu menggambarkan alur kerja secara jelas sehingga mudah dipahami serta memiliki susunan proses yang sistematis dan terstruktur bagi pengembang maupun pengguna seperti pada tabel 2.2 simbol-simbol *flowchart*.

Tabel 2.2 Simbol-simbol *flowchart*

SIMBOL	NAMA	KETERANGAN
	<i>Input/Output</i>	Menggambarkan data masukan, data keluaran, maupun informasi yang diproses dalam sistem.
	Proses	Merepresentasikan Pekerjaan
	<i>Connector</i>	Menghubungkan alur proses dengan bagian lain pada <i>flowchart</i> di halaman yang sama.

	<i>Off-page Connector</i>	Menghubungkan alur proses dengan bagian lain pada <i>flowchart</i> yang berada di halaman berbeda.
	Anak Panah	Merepresentasikan alur kerja
	Keputusan	Menandai titik pengambilan keputusan pada proses
	<i>Preparation</i>	Menginisialisasi nilai awal
	<i>Terminal Point</i>	Mulai dan berakhirnya <i>flowchart</i>
	<i>Punched Card</i>	Menunjukkan proses masukan atau keluaran data melalui kartu berlubang (<i>punched card</i>).
	Dokumen	Menunjukkan proses masukan atau keluaran dengan bentuk hasil cetak.
	Manual Input	Menunjukkan proses pemasukan data secara manual melalui keyboard.
	Database	Memasukkan ke base data

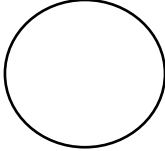
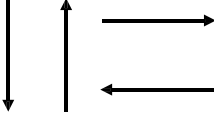
Sumber: (Yulianeu & Oktamala, 2022:128).

10. *Data Flow Diagram (DFD)*

Data Flow Diagram (DFD) merupakan suatu model yang digunakan untuk menggambarkan aliran data pada sebuah sistem informasi secara logis. DFD menunjukkan bagaimana data mengalir dari entitas luar menuju sistem, bagaimana data diproses, disimpan, dan dihasilkan kembali sebagai informasi. Diagram ini membantu pengembang maupun pengguna dalam memahami proses bisnis yang berlangsung pada sistem tanpa memperhatikan implementasi program atau perangkat keras yang digunakan (Made et al., 2023:237).

Menurut Rahmadan & Gunawan, (2024:4), *Data Flow Diagram (DFD)* digunakan untuk memodelkan aliran data pada sistem penentuan siswa berbakat berbasis website menggunakan metode *Naive Bayes*. DFD menggambarkan hubungan antara pengguna dengan sistem, proses pengolahan data, penyimpanan data pada basis data, hingga menghasilkan informasi berupa hasil klasifikasi siswa berbakat. Dengan adanya DFD, alur proses sistem dapat dipahami secara lebih jelas sehingga memudahkan proses analisis, perancangan, dan pengembangan sistem. Selain itu, DFD juga memberikan gambaran mengenai proses masuknya data, pengolahan data, penyimpanan data, hingga keluaran yang dihasilkan oleh sistem secara terstruktur (Sari et al., 2023:127). Dalam penyusunannya, DFD menggunakan beberapa simbol yang memiliki fungsi berbeda untuk menggambarkan komponen-komponen sistem. Simbol-simbol yang digunakan dalam *Data Flow Diagram (DFD)*. Pada penelitian ini dapat dilihat pada tabel 2.3

Tabel 2.3 Simbol – simbol *data flow diagram* (DFD)

SIMBOL	NAMA	KETERANGAN
	Proses	Menunjukkan proses atau aktivitas pengolahan data yang dilakukan oleh sistem sehingga menghasilkan keluaran berupa informasi.
	Terminator	Menunjukkan entitas luar yang berinteraksi dengan sistem, baik sebagai pemberi data input maupun penerima informasi output.
	Arus Data	Menunjukkan arah perpindahan atau aliran data antar proses, entitas, dan penyimpanan data dalam sistem.
	Data Store	Menunjukkan tempat penyimpanan data berupa file atau database yang digunakan oleh sistem untuk menyimpan dan mengelola data.

Sumber: (Djauhari & Irfandi, 2022:23).

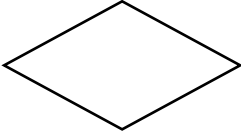
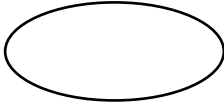
11. *Entity Relationship Diagram* (ERD)

Entity Relationship Diagram (ERD) merupakan diagram yang digunakan untuk memodelkan struktur basis data dengan menggambarkan hubungan antar entitas beserta atribut yang dimilikinya. ERD berfungsi sebagai acuan dalam merancang database sehingga hubungan antar tabel dapat disusun secara sistematis, mengurangi redundansi data, serta menjaga integritas data dalam sistem informasi (Afiifah et al., 2022:18).

Pada penelitian ini, *Entity Relationship Diagram* (ERD) digunakan untuk merancang struktur basis data pada sistem penentuan siswa berbakat menggunakan metode *Naive Bayes* berbasis website. ERD menggambarkan hubungan antar entitas seperti pengguna, data siswa, kriteria, nilai kriteria,

proses klasifikasi, dan hasil klasifikasi. Dengan adanya ERD, proses implementasi database menjadi lebih terstruktur, memudahkan pengelolaan data, serta mendukung proses pengolahan data dan penyajian informasi secara efektif (Akbar & Haryanti, 2021:29). Pada penelitian ini dapat dilihat pada tabel 2.4.

Tabel 2.4 Simbol – simbol *entity relationship diagram* (ERD)

SIMBOL	NAMA	KETERANGAN
	Entity	Menunjukkan objek atau himpunan data yang memiliki informasi dan akan disimpan dalam basis data. Setiap entitas memiliki atribut sebagai identitas atau karakteristiknya..
	Garis	Menunjukkan hubungan atau penghubung antara entitas dengan relasi maupun atribut sehingga menggambarkan keterkaitan antar komponen pada ERD..
	Relationship	Menunjukkan hubungan atau keterkaitan yang terjadi antara dua atau lebih entitas dalam suatu basis data..
	Atribut	Menunjukkan karakteristik atau informasi yang dimiliki oleh suatu entitas sebagai penjelas data yang disimpan..

Sumber: (Triaulia & Rusydi, 2023:22)

B. Kajian Empiris

Dalam penelitian ini, terdapat lima kajian empiris yang digunakan sebagai referensi relevan dengan topik sistem pendukung keputusan berbasis

Naïve Bayes untuk penentuan bakat, rekomendasi jurusan, serta klasifikasi kemampuan akademik.

Penelitian pertama dilakukan oleh Dewi et al., (2021:161) berjudul “Implementasi Metode *Forward Chaining* Dan *Naïve Bayes* Terhadap Sistem Pakar Penentuan Bakat”. Penelitian ini bertujuan agar seseorang dapat memahami bakat yang dimiliki serta cara mengembangkannya. Metode yang digunakan adalah pengumpulan data langsung dari pakar psikologi (Siti Marifah S, S.Psi.) berupa 50 data training dan 20 data testing, kemudian diterapkan sistem pakar dengan aturan *Forward Chaining* dan klasifikasi probabilitas menggunakan *Naïve Bayes*. Hasil penelitian menunjukkan tingkat akurasi sistem mencapai 95%.

Penelitian kedua dilakukan oleh Bagus et al., (2023:361) berjudul “Sistem Rekomendasi Jurusan Menggunakan Algoritma *Naïve Bayes Gaussian* Berbasis Web”. Penelitian ini bertujuan mengatasi kebingungan siswa lulusan SMA/K dalam memilih jurusan di STIKI Malang. Metode yang digunakan adalah pengumpulan data survey dari 107 mahasiswa aktif dan lulusan sebagai data latih, diikuti analisis data dengan 8 atribut (jenis kelamin, jurusan SMA, skill, hobi, alasan kuliah, dll.) serta klasifikasi menggunakan *Naïve Bayes Gaussian*. Hasil penelitian menunjukkan akurasi klasifikasi mencapai 87,85%.

Penelitian ketiga dilakukan oleh Walangare, (2022:81) dengan judul “Penerapan Algoritma *Naive Bayes* Dalam Sistem Pendukung Keputusan Pemilihan Peminatan Konsentrasi Berdasarkan Nilai Akademik Berbasis Web Pada Program Studi S1 Pendidikan Teknologi Informasi”. Studi ini

mempertimbangan objektif bagi mahasiswa semester 5 dalam memilih konsentrasi (RPL, MM, atau TKJ) berdasarkan nilai akademik mata kuliah dasar. Metode yang digunakan adalah *Naïve Bayes* dengan data nilai mahasiswa angkatan 2018 dan 2019, diuji menggunakan *black-box* testing dan validasi ahli media. Hasil penelitian menunjukkan akurasi 100%, dengan validasi ahli media sebesar 90% dan 82%.

Penelitian keempat dilakukan oleh Hartati & Apriliani, (2025:52) berjudul “Rekomendasi Pemilihan Jurusan Bagi Siswa Baru di SMK XYZ Menggunakan Metode *Naive Bayes* dan *Simple Additive Weighting* (SAW)”. Penelitian ini bertujuan menggantikan proses pemilihan jurusan yang masih konvensional dan subjektif di SMK XYZ. Metode yang digunakan meliputi pengumpulan data primer (nilai rapor SMP: Matematika, IPA, Bahasa Inggris, serta hasil wawancara) dan sekunder, penentuan 4 kriteria dengan bobot, serta klasifikasi menggunakan *Naïve Bayes* yang dikombinasikan dengan SAW untuk perankingan multikriteria. Hasil penelitian menunjukkan bahwa kombinasi kedua metode menghasilkan rekomendasi jurusan yang lebih objektif, akurat, efisien, dan minim subjektivitas sesuai karakteristik siswa.

Penelitian kelima dilakukan oleh Iskandar et al., (2025:1609) berjudul “*Naïve Bayes Classifier-Based Intelligent System for Student Academic Performance Assessment*”. Penelitian ini bertujuan merancang sistem penilaian kemampuan akademik siswa secara cerdas yang mampu mengklasifikasikan performa siswa menjadi kategori rendah, sedang, dan tinggi. Metode yang digunakan adalah *Naïve Bayes Classifier* dengan data indikator

akademik (nilai tes, kehadiran, perilaku) yang diimplementasikan menggunakan *PHP* dan *MySQL*. Hasil pengujian selama sepuluh iterasi menunjukkan akurasi rata-rata 96,67%, serta tingkat kepuasan pengguna (230 responden) mencapai 86,8%.

Berdasarkan kelima penelitian terdahulu tersebut, dapat disimpulkan bahwa metode *Naïve Bayes* memiliki kemampuan yang sangat baik dalam melakukan klasifikasi dan pengambilan keputusan berbasis data, baik secara mandiri maupun dikombinasikan dengan metode lain (*Forward Chaining*, *Gaussian*, atau SAW).. Oleh karena itu, metode *Naïve Bayes* sangat relevan dan dapat diadopsi dalam penelitian ini untuk membangun sistem penentuan siswa berbakat berbasis website.

C. Kerangka Berpikir

Penelitian pertama dilakukan oleh Dewi et al., (2021:161) berjudul “Implementasi Metode *Forward Chaining* Dan *Naïve Bayes* Terhadap Sistem Pakar Penentuan Bakat”. Penelitian ini bertujuan agar seseorang dapat memahami bakat yang dimiliki serta cara mengembangkannya. Metode yang digunakan adalah pengumpulan data langsung dari pakar psikologi (Siti Marifah S, S.Psi.) berupa 50 data training dan 20 data testing, kemudian diterapkan sistem pakar dengan aturan *Forward Chaining* dan klasifikasi probabilitas menggunakan *Naïve Bayes*. Hasil penelitian menunjukkan tingkat akurasi sistem mencapai 95%.

Penelitian kedua dilakukan oleh Bagus et al., (2023:361) berjudul “Sistem Rekomendasi Jurusan Menggunakan Algoritma *Naïve Bayes Gaussian*

Berbasis Web”. Penelitian ini bertujuan mengatasi kebingungan siswa lulusan SMA/K dalam memilih jurusan di STIKI Malang. Metode yang digunakan adalah pengumpulan data survey dari 107 mahasiswa aktif dan lulusan sebagai data latih, diikuti analisis data dengan 8 atribut (jenis kelamin, jurusan SMA, skill, hobi, alasan kuliah, dll.) serta klasifikasi menggunakan *Naïve Bayes* Gaussian. Hasil penelitian menunjukkan akurasi klasifikasi mencapai 87,85%.

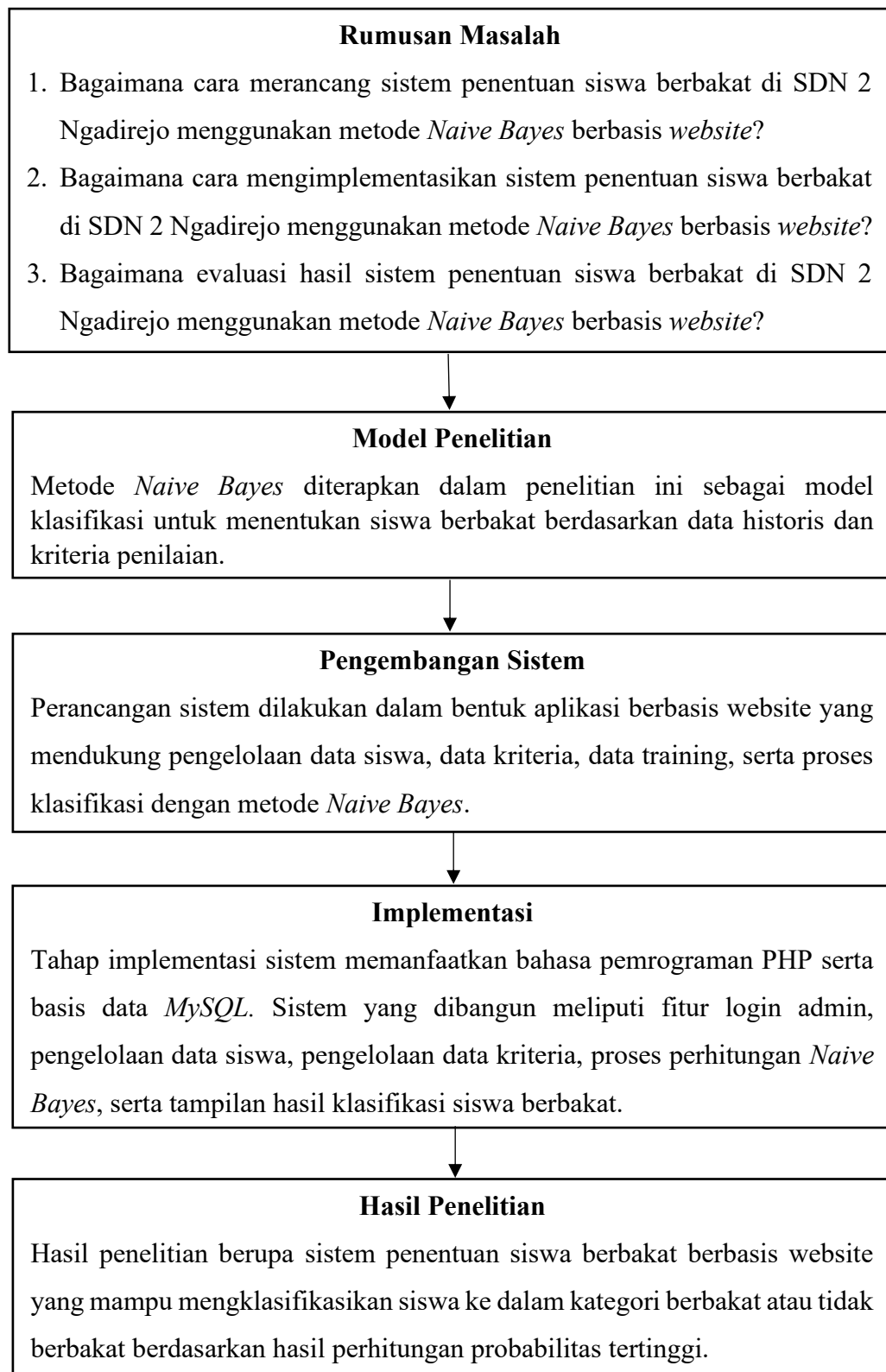
Penelitian ketiga dilakukan oleh Walangare, (2022:81) dengan judul “Penerapan Algoritma *Naive Bayes* Dalam Sistem Pendukung Keputusan Pemilihan Peminatan Konsentrasi Berdasarkan Nilai Akademik Berbasis Web Pada Program Studi S1 Pendidikan Teknologi Informasi”. Studi ini mempertimbangan objektif bagi mahasiswa semester 5 dalam memilih konsentrasi (RPL, MM, atau TKJ) berdasarkan nilai akademik mata kuliah dasar. Metode yang digunakan adalah *Naïve Bayes* dengan data nilai mahasiswa angkatan 2018 dan 2019, diuji menggunakan *black-box* testing dan validasi ahli media. Hasil penelitian menunjukkan akurasi 100%, dengan validasi ahli media sebesar 90% dan 82%.

Penelitian keempat dilakukan oleh Hartati & Apriliani, (2025:52) berjudul “Rekomendasi Pemilihan Jurusan Bagi Siswa Baru di SMK XYZ Menggunakan Metode *Naive Bayes* dan *Simple Additive Weighting* (SAW)”. Penelitian ini bertujuan menggantikan proses pemilihan jurusan yang masih konvensional dan subjektif di SMK XYZ. Metode yang digunakan meliputi pengumpulan data primer (nilai rapor SMP: Matematika, IPA, Bahasa Inggris, serta hasil wawancara) dan sekunder, penentuan 4 kriteria dengan bobot, serta

klasifikasi menggunakan *Naïve Bayes* yang dikombinasikan dengan SAW untuk perankingan multikriteria. Hasil penelitian menunjukkan bahwa kombinasi kedua metode menghasilkan rekomendasi jurusan yang lebih objektif, akurat, efisien, dan minim subjektivitas sesuai karakteristik siswa.

Penelitian kelima dilakukan oleh Iskandar et al., (2025:1609) berjudul “*Naïve Bayes Classifier-Based Intelligent System for Student Academic Performance Assessment*”. Penelitian ini bertujuan merancang sistem penilaian kemampuan akademik siswa secara cerdas yang mampu mengklasifikasikan performa siswa menjadi kategori rendah, sedang, dan tinggi. Metode yang digunakan adalah *Naïve Bayes Classifier* dengan data indikator akademik (nilai tes, kehadiran, perilaku) yang diimplementasikan menggunakan *PHP* dan *MySQL*. Hasil pengujian selama sepuluh iterasi menunjukkan akurasi rata-rata 96,67%, serta tingkat kepuasan pengguna (230 responden) mencapai 86,8%.

Pengembangan potensi peserta didik sejak jenjang sekolah dasar memerlukan proses identifikasi yang dilakukan secara tepat. Salah satu bentuk implementasinya adalah penentuan siswa berbakat sebagai dasar pemberian pembinaan yang sesuai dengan kemampuan yang dimiliki, baik pada aspek akademik maupun non-akademik. Akan tetapi, pelaksanaannya di sekolah masih banyak mengandalkan penilaian secara manual sehingga tingkat objektivitas dan transparansi belum sepenuhnya terpenuhi. Kondisi tersebut dapat memengaruhi ketepatan hasil penilaian serta keputusan yang diambil. Berikut kerangka berpikir dapat dilihat pada gambar 2.3.



Gambar 2.3 Kerangka Berpikir.