

BAB II

TINJAUAN PUSTAKA

2.1 Landasan Teori

Landasan teori merupakan kumpulan konsep, teori, dan hasil penelitian yang digunakan sebagai dasar dalam pelaksanaan penelitian. Landasan teori berfungsi untuk memberikan pemahaman mengenai konsep-konsep yang berkaitan dengan permasalahan penelitian sehingga dapat mendukung proses analisis, pemilihan metode, serta pembahasan hasil penelitian. Dalam penelitian ini, landasan teori mencakup konsep *data mining*, *Educational Data Mining (EDM)*, pertimbangan dalam memilih sekolah, teknik klasifikasi, algoritma *Naive Bayes*, perangkat lunak *RapidMiner*, metode *Cross Validation*, serta evaluasi model menggunakan *Confusion Matrix*.

2.1.1 Data Mining

Data mining merupakan proses penggalian informasi atau pengetahuan dari sekumpulan data dalam jumlah besar untuk menemukan pola, hubungan, maupun informasi yang sebelumnya tidak diketahui sehingga dapat dimanfaatkan sebagai dasar dalam pengambilan keputusan. Proses ini memanfaatkan berbagai teknik dari bidang statistik, *machine learning*, kecerdasan buatan, dan basis data untuk mengolah data menjadi informasi yang bernilai. Dengan kemampuan tersebut, *data mining* telah banyak diterapkan pada berbagai bidang, seperti bisnis, kesehatan, pemerintahan, dan pendidikan [1].

Selain berfungsi untuk menemukan pola dari data, *data mining* juga berperan dalam proses ekstraksi pengetahuan (*knowledge discovery*) dari data yang berukuran besar dan kompleks menjadi informasi yang lebih bermakna.

Informasi yang dihasilkan dapat dimanfaatkan untuk mendukung proses analisis, prediksi, klasifikasi, serta pengambilan keputusan pada berbagai bidang sesuai dengan kebutuhan pengguna [1], [10].

Perkembangan *data mining* juga mendorong penerapannya pada bidang pendidikan melalui konsep *Educational Data Mining (EDM)*. Konsep tersebut memanfaatkan teknik *data mining* untuk menganalisis data pendidikan sehingga mampu menghasilkan informasi yang mendukung peningkatan kualitas pembelajaran maupun pengambilan keputusan di lingkungan pendidikan [2].

2.1.2 Educational Data Mining

Educational Data Mining (EDM) merupakan penerapan teknik *data mining* pada bidang pendidikan untuk menganalisis data akademik maupun nonakademik sehingga menghasilkan informasi yang bermanfaat dalam mendukung pengambilan keputusan. Penerapan *Educational Data Mining* memanfaatkan berbagai teknik analisis data dan *machine learning* untuk menemukan pola, hubungan, maupun karakteristik yang terdapat pada data pendidikan sehingga dapat memberikan informasi yang mendukung peningkatan kualitas pembelajaran dan pengelolaan institusi pendidikan [2].

Dalam perkembangannya, *Educational Data Mining* telah banyak dimanfaatkan untuk menganalisis prestasi belajar peserta didik, memprediksi performa akademik, mengidentifikasi peserta didik yang berisiko mengalami kesulitan belajar, serta mendukung evaluasi proses pembelajaran [3]. Selain itu, hasil kajian selama sepuluh tahun terakhir menunjukkan bahwa penerapan *Educational Data Mining* terus berkembang dan menjadi salah satu pendekatan

yang banyak digunakan dalam penelitian pendidikan karena mampu menghasilkan informasi yang mendukung pengambilan keputusan berbasis data [3].

Pada penelitian ini, konsep *Educational Data Mining* diterapkan untuk menganalisis tingkat minat murid yang telah memilih MTs Mambaul Ulum berdasarkan atribut lokasi sekolah, kualitas sekolah, fasilitas sekolah, dan biaya pendidikan menggunakan algoritma *Naive Bayes*. Penerapan konsep tersebut diharapkan mampu menghasilkan informasi mengenai karakteristik minat murid yang telah memilih MTs Mambaul Ulum yang dapat dimanfaatkan sebagai bahan pertimbangan dalam pelaksanaan Sistem Penerimaan Murid Baru (SPMB).

2.1.3 Pertimbangan dalam Memilih Sekolah

Minat calon murid dalam memilih sekolah berkaitan dengan berbagai pertimbangan yang berasal dari karakteristik sekolah maupun lingkungan sekitarnya. Berbagai pertimbangan tersebut menjadi bagian dari proses calon murid dan orang tua dalam menentukan sekolah yang sesuai dengan kebutuhan, harapan, serta tujuan pendidikan. Hasil penelitian menunjukkan bahwa keputusan memilih sekolah berkaitan dengan faktor internal dan faktor eksternal yang berhubungan dengan karakteristik individu maupun lingkungan pendidikan [5]. Selain itu, penelitian lain menunjukkan bahwa lingkungan keluarga, lingkungan sekolah, kompetensi lulusan, serta berbagai pertimbangan lainnya menjadi bagian dari proses calon murid dalam menentukan sekolah tujuan [6].

Hasil *Systematic Literature Review* menunjukkan bahwa pertimbangan yang sering digunakan dalam memilih sekolah meliputi lokasi sekolah, kualitas

sekolah, fasilitas sekolah, biaya pendidikan, reputasi sekolah, dan citra sekolah [7]. Berdasarkan hasil telaah tersebut, penelitian ini memfokuskan analisis pada empat atribut adalah sebagai berikut: lokasi sekolah, kualitas sekolah, fasilitas sekolah, dan biaya pendidikan. Keempat atribut tersebut dipilih karena relevan dengan berbagai pertimbangan dalam memilih sekolah yang telah dibahas dalam penelitian terdahulu serta sesuai dengan kondisi penelitian di MTs Mambaul Ulum [7].

1. Lokasi sekolah menjadi salah satu pertimbangan dalam memilih sekolah karena berkaitan dengan kemudahan akses menuju sekolah, efisiensi waktu tempuh, serta kenyamanan selama perjalanan. Penelitian menunjukkan bahwa lokasi sekolah berkaitan dengan keputusan memilih sekolah karena kemudahan akses menjadi salah satu pertimbangan calon murid dalam menentukan sekolah tujuan [8], [9].
2. Kualitas sekolah mencerminkan mutu layanan pendidikan yang diberikan kepada peserta didik, meliputi kualitas proses pembelajaran, kompetensi pendidik, serta prestasi sekolah. Sekolah yang memiliki kualitas pendidikan yang baik dapat menjadi pertimbangan bagi calon murid maupun orang tua dalam menentukan pilihan sekolah [10].
3. Fasilitas sekolah merupakan sarana dan prasarana yang mendukung pelaksanaan proses pembelajaran. Ketersediaan fasilitas yang memadai memberikan kenyamanan dalam kegiatan belajar mengajar sehingga menjadi salah satu pertimbangan dalam memilih sekolah [8], [10].

4. Biaya pendidikan merupakan besarnya biaya yang harus dipenuhi untuk memperoleh layanan pendidikan. Keterjangkauan biaya menjadi salah satu pertimbangan bagi calon murid maupun orang tua karena berkaitan dengan kemampuan ekonomi dalam memenuhi kebutuhan pendidikan. Penelitian menunjukkan bahwa biaya pendidikan berkaitan dengan keputusan memilih sekolah, terutama dalam pertimbangan kesesuaian antara biaya yang ditetapkan dan layanan pendidikan yang diberikan [9].

Berdasarkan uraian tersebut, lokasi sekolah, kualitas sekolah, fasilitas sekolah, dan biaya pendidikan digunakan sebagai atribut prediktor dalam penelitian ini. Keempat atribut tersebut selanjutnya digunakan dalam proses klasifikasi tingkat minat menggunakan algoritma *Naive Bayes*.

2.1.4 Klasifikasi

Klasifikasi merupakan salah satu teknik dalam data mining yang digunakan untuk mengelompokkan data ke dalam kelas atau kategori tertentu berdasarkan karakteristik yang dimiliki [1]. Teknik ini bekerja dengan membangun model klasifikasi menggunakan data yang telah diketahui kelasnya (*training data*), kemudian model tersebut digunakan untuk memprediksi kelas pada data baru yang belum memiliki label. Proses klasifikasi bertujuan menghasilkan model yang mampu mengidentifikasi pola hubungan antara atribut dengan kelas sehingga dapat dimanfaatkan sebagai dasar dalam proses prediksi maupun pengambilan keputusan [1].

Penerapan teknik klasifikasi telah banyak digunakan pada berbagai bidang, termasuk bidang pendidikan. Dalam konsep *Educational Data Mining*, teknik

klasifikasi dimanfaatkan untuk menganalisis karakteristik peserta didik, memprediksi prestasi akademik, mengidentifikasi peserta didik yang berisiko mengalami kesulitan belajar, serta mendukung berbagai proses pengambilan keputusan berbasis data [2], [3]. Pada penelitian ini, teknik klasifikasi digunakan untuk mengelompokkan tingkat minat murid yang telah memilih MTs Mambaul Ulum ke dalam kategori rendah, sedang, dan tinggi berdasarkan atribut lokasi sekolah, kualitas sekolah, fasilitas sekolah, dan biaya pendidikan. Atribut lokasi sekolah, kualitas sekolah, fasilitas sekolah, dan biaya pendidikan digunakan sebagai atribut prediktor, sedangkan minat digunakan sebagai kelas atau label yang diprediksi. Proses klasifikasi dilakukan menggunakan algoritma *Naive Bayes* untuk menghasilkan model yang mampu mengidentifikasi kategori minat murid berdasarkan data hasil kuesioner.

2.1.5 Algoritma *Naive Bayes*

Naive Bayes merupakan salah satu algoritma klasifikasi yang didasarkan pada *Teorema Bayes* untuk menghitung probabilitas suatu data termasuk ke dalam kelas tertentu berdasarkan atribut yang dimiliki. Algoritma ini mengasumsikan bahwa setiap atribut bersifat independen terhadap atribut lainnya sehingga proses perhitungan probabilitas menjadi lebih sederhana dan efisien [1].

Proses klasifikasi menggunakan algoritma *Naive Bayes* dilakukan dengan menghitung probabilitas setiap kelas berdasarkan data pelatihan (*training data*), kemudian menghitung probabilitas kemunculan setiap atribut pada masing-masing kelas. Nilai probabilitas akhir diperoleh dari hasil perhitungan seluruh

atribut yang dimiliki suatu data. Selanjutnya, data diklasifikasikan ke dalam kelas yang memiliki nilai probabilitas terbesar [1].

Teorema Bayes yang menjadi dasar algoritma *Naive Bayes* dirumuskan sebagai berikut:

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

Keterangan:

1. $P(H|X)$ = probabilitas hipotesis H berdasarkan kondisi X.
2. $P(X|H)$ = probabilitas munculnya X jika H terjadi.
3. $P(H)$ = probabilitas awal hipotesis H.
4. $P(X)$ = probabilitas munculnya X.
5. H = hipotesis atau kelas yang diprediksi.
6. X = data atau atribut yang digunakan sebagai dasar prediksi.

2.1.6 *RapidMiner*

RapidMiner merupakan perangkat lunak *data mining* yang digunakan untuk membantu proses pengolahan data, pemodelan, serta evaluasi hasil analisis. *RapidMiner* menyediakan antarmuka berbasis *Graphical User Interface (GUI)* sehingga pengguna dapat membangun model data mining tanpa harus menuliskan kode program secara langsung. Perangkat lunak ini mendukung berbagai teknik analisis data, meliputi proses data *preprocessing*, klasifikasi, klusterisasi, asosiasi, validasi model, serta evaluasi hasil analisis [11].

RapidMiner mendukung penerapan berbagai algoritma klasifikasi, termasuk algoritma *Naive Bayes*, serta menyediakan operator untuk melakukan *Cross Validation* dan evaluasi model menggunakan *Confusion Matrix*. Fitur-fitur tersebut memudahkan proses pembangunan model klasifikasi, pengujian model, dan pengukuran kinerja model sehingga *RapidMiner* banyak digunakan sebagai perangkat lunak pendukung dalam penelitian *data mining* [11]. Pada penelitian ini, *RapidMiner* digunakan untuk membangun model klasifikasi menggunakan algoritma *Naive Bayes*, melakukan proses validasi model, serta mengevaluasi hasil klasifikasi berdasarkan nilai *accuracy*, *precision*, dan *recall*.

2.1.7 Cross Validation

Cross Validation merupakan metode pengujian model yang digunakan untuk mengetahui kemampuan model dalam melakukan klasifikasi secara lebih objektif. Metode ini dilakukan dengan membagi dataset ke dalam beberapa bagian (*fold*), kemudian melakukan proses pelatihan dan pengujian secara bergantian hingga seluruh data memperoleh kesempatan menjadi data uji. Teknik ini digunakan untuk mengurangi bias dalam pengujian model dan menghasilkan estimasi performa model yang lebih stabil.

Dalam penelitian ini digunakan metode *10-Fold Cross Validation*. Metode tersebut dilakukan dengan membagi dataset menjadi sepuluh bagian (*fold*). Sembilan bagian digunakan sebagai data latih (*training data*), sedangkan satu bagian digunakan sebagai data uji (*testing data*). Proses tersebut dilakukan secara bergantian hingga seluruh bagian dataset memperoleh kesempatan sebagai data uji sehingga dihasilkan sepuluh kali proses pengujian.

Pada penelitian ini, metode *10-Fold Cross Validation* digunakan untuk mengevaluasi model klasifikasi yang dibangun menggunakan algoritma *Naive Bayes*. Hasil evaluasi kemudian digunakan untuk mengetahui kinerja model berdasarkan nilai *accuracy*, *precision*, dan *recall*.

2.1.8 *Confusion Matrix*

Confusion Matrix merupakan metode evaluasi yang digunakan untuk mengukur kinerja model klasifikasi dengan membandingkan hasil prediksi model terhadap kelas sebenarnya. Metode ini menyajikan hasil klasifikasi dalam bentuk matriks sehingga dapat menunjukkan jumlah data yang diklasifikasikan dengan benar maupun data yang mengalami kesalahan klasifikasi [12].

Hasil evaluasi menggunakan *Confusion Matrix* menghasilkan beberapa metrik pengukuran. Metrik tersebut adalah sebagai berikut:

1. *Accuracy* merupakan persentase jumlah data yang berhasil diklasifikasikan dengan benar terhadap seluruh data yang digunakan dalam pengujian.
2. *Precision* merupakan tingkat ketepatan model dalam memprediksi suatu kelas berdasarkan jumlah prediksi yang benar pada kelas tersebut.
3. *Recall* merupakan kemampuan model dalam mengidentifikasi seluruh data yang termasuk ke dalam suatu kelas berdasarkan data sebenarnya [12].

Pada penelitian ini, *Confusion Matrix* digunakan untuk mengevaluasi kinerja algoritma *Naive Bayes*. Hasil evaluasi ditunjukkan melalui nilai *accuracy*, *precision*, dan *recall* sehingga dapat diketahui kinerja model dalam mengklasifikasikan tingkat minat murid yang telah memilih MTs Mambaul Ulum

berdasarkan atribut lokasi sekolah, kualitas sekolah, fasilitas sekolah, dan biaya pendidikan.

2.2 Penelitian Terdahulu

Tabel 2.1 Matriks literatur review dan posisi penelitian

ANALISIS MINAT CALON MURID DALAM MEMILIH MTS MAMBAUL ULUM MENGGUNAKAN ALGORITMA NAIVE
BAYES

No	Judul Penelitian	Nama Peneliti, Tahun, Index	Metode Penelitian	Hasil	Keunggulan dan Kelemahan	Perbandingan
1	Faktor Penentu Keputusan Siswa dalam Memilih Sekolah Menengah Kejuruan [5]	M. K. Arifai, 2023 (Jurnal Pemasaran Kompetitif)	Menggunakan pendekatan kuantitatif untuk menganalisis faktor-faktor yang memengaruhi keputusan siswa dalam memilih Sekolah Menengah Kejuruan (SMK).	Hasil penelitian menunjukkan bahwa keputusan siswa dalam memilih sekolah dipengaruhi oleh beberapa faktor, meliputi lokasi sekolah, kualitas pendidikan, fasilitas sekolah,	Keunggulan: mampu mengidentifikasi faktor-faktor yang memengaruhi keputusan siswa dalam memilih sekolah. Kelemahan: tidak menerapkan algoritma klasifikasi atau metode data mining.	Penelitian tersebut berfokus pada analisis faktor-faktor yang memengaruhi keputusan memilih sekolah, sedangkan penelitian ini menerapkan algoritma Naive Bayes untuk mengklasifikasikan tingkat minat murid berdasarkan atribut

				biaya pendidikan, dan citra sekolah.		lokasi sekolah, kualitas sekolah, fasilitas sekolah, dan biaya pendidikan.
2	Analisis Faktor-Faktor yang Mempengaruhi Minat Siswa Luar Kota dalam Memilih Sekolah Menengah Kejuruan di Kota Surakarta Tahun Ajaran 2023/2024 (Studi Kasus di SMK Negeri 6 Surakarta) [6]	A. Widiananda dan T. Murwaningsih, 2024 (Jurnal Pendidikan dan Pengajaran Cendekia)	Menggunakan pendekatan kuantitatif untuk menganalisis faktor-faktor yang memengaruhi minat siswa luar kota dalam memilih Sekolah Menengah Kejuruan (SMK).	Hasil penelitian menunjukkan bahwa minat siswa dalam memilih sekolah dipengaruhi oleh berbagai faktor, meliputi lingkungan keluarga, lingkungan sekolah, kompetensi lulusan, serta pertimbangan lain sebelum	Keunggulan: mampu mengidentifikasi berbagai faktor yang memengaruhi minat siswa dalam memilih sekolah. Kelemahan: tidak menerapkan metode klasifikasi atau algoritma data mining.	Penelitian tersebut berfokus pada analisis faktor-faktor yang memengaruhi minat siswa dalam memilih sekolah, sedangkan penelitian ini menerapkan algoritma Naive Bayes untuk mengklasifikasikan tingkat minat murid berdasarkan atribut lokasi sekolah, kualitas sekolah, fasilitas sekolah, dan biaya pendidikan.

				menentukan sekolah tujuan.		
3	Faktor-Faktor yang Mempengaruhi Orang Tua terhadap Minat Memilih Sekolah Dasar Swasta: Systematic Literature Review [7]	M. D. Restarie, Nurhattati, dan Kamaludin, 2025 (Journal on Education)	Menggunakan metode Systematic Literature Review (SLR) untuk mengidentifikasi faktor-faktor yang memengaruhi orang tua dalam memilih sekolah dasar swasta berdasarkan berbagai penelitian terdahulu.	Hasil penelitian menunjukkan bahwa keputusan memilih sekolah dipengaruhi oleh kualitas pendidikan, fasilitas sekolah, lokasi sekolah, biaya pendidikan, reputasi sekolah, dan citra sekolah.	Keunggulan: menyajikan hasil sintesis dari berbagai penelitian sehingga memberikan gambaran yang komprehensif mengenai faktor-faktor pemilihan sekolah. Kelemahan: tidak melakukan pengujian menggunakan data primer maupun metode klasifikasi.	Penelitian tersebut mengidentifikasi berbagai pertimbangan dalam pemilihan sekolah melalui kajian literatur, sedangkan penelitian ini mengolah data hasil kuesioner dan menerapkan algoritma Naive Bayes untuk mengklasifikasikan tingkat minat murid berdasarkan atribut lokasi sekolah, kualitas sekolah, fasilitas sekolah, dan biaya pendidikan.
4	Pengaruh Lokasi, Fasilitas Pendidikan	Y. M. Simamora, H.	Menggunakan pendekatan	Hasil penelitian menunjukkan	Keunggulan: menggunakan variabel yang relevan	Penelitian tersebut menganalisis pengaruh

	dan Citra Sekolah Terhadap Keputusan Memilih Sekolah [8]	Siagian, dan P. Pelawi, 2023 (Jurnal Wira Ekonomi Mikroskil)	kuantitatif untuk menganalisis pengaruh lokasi sekolah, fasilitas pendidikan, dan citra sekolah terhadap keputusan siswa dalam memilih sekolah.	bahwa lokasi sekolah, fasilitas pendidikan, dan citra sekolah berpengaruh positif terhadap keputusan siswa dalam memilih sekolah.	dengan keputusan memilih sekolah. Kelemahan: tidak menerapkan algoritma klasifikasi atau metode data mining.	lokasi, fasilitas pendidikan, dan citra sekolah terhadap keputusan memilih sekolah, sedangkan penelitian ini menerapkan algoritma Naive Bayes untuk mengklasifikasikan tingkat minat murid berdasarkan atribut lokasi sekolah, kualitas sekolah, fasilitas sekolah, dan biaya pendidikan.
5	Pengaruh Lokasi dan Biaya Pendidikan Terhadap Keputusan Pemilihan Sekolah yang Dimediasi Reputasi Sekolah	D. M. Attabik Bahta dan R. A. Santoso, 2025 (Paradoks:	Menggunakan pendekatan kuantitatif untuk menganalisis pengaruh lokasi dan biaya	Hasil penelitian menunjukkan bahwa lokasi sekolah dan biaya pendidikan berpengaruh	Keunggulan: mengkaji hubungan antarvariabel secara lebih mendalam melalui variabel mediasi. Kelemahan: tidak menggunakan metode	Penelitian tersebut menganalisis pengaruh lokasi dan biaya pendidikan terhadap keputusan memilih sekolah, sedangkan

	pada SMP Muhammadiyah 4 Kebomas [9]	Jurnal Ilmu Ekonomi)	pendidikan terhadap keputusan memilih sekolah dengan reputasi sekolah sebagai variabel mediasi.	terhadap keputusan memilih sekolah, baik secara langsung maupun melalui reputasi sekolah sebagai variabel mediasi.	klasifikasi atau algoritma data mining.	penelitian ini menggunakan algoritma Naive Bayes untuk mengklasifikasikan tingkat minat murid berdasarkan atribut lokasi sekolah, kualitas sekolah, fasilitas sekolah, dan biaya pendidikan.
6	The Influence of Educational Quality, School Facilities, and Brand Image on Students' School Choice Decisions [10]	R. L. Tambunan, M. Bukhori, dan T. A. Karnawati, 2026 (Journal of English Language and Education)	Menggunakan pendekatan kuantitatif untuk menganalisis pengaruh kualitas pendidikan, fasilitas sekolah, dan brand image terhadap keputusan siswa	Hasil penelitian menunjukkan bahwa kualitas pendidikan, fasilitas sekolah, dan brand image berpengaruh positif terhadap keputusan siswa dalam memilih sekolah.	Keunggulan: menggunakan variabel yang relevan dengan keputusan memilih sekolah. Kelemahan: tidak menerapkan algoritma data mining untuk proses klasifikasi.	Penelitian tersebut menganalisis pengaruh kualitas pendidikan, fasilitas sekolah, dan brand image terhadap keputusan memilih sekolah, sedangkan penelitian ini menerapkan algoritma Naive Bayes untuk mengklasifikasikan

			dalam memilih sekolah.			tingkat minat murid berdasarkan atribut lokasi sekolah, kualitas sekolah, fasilitas sekolah, dan biaya pendidikan.
7	Implementasi Algoritma Naive Bayes untuk Sistem Pendukung Keputusan Penerima Bantuan Pendidikan [1]	C. W. Nugroho, S. Tamriesfatno, dan Sulkifli, 2025 (Jurnal Publikasi Ilmu Komputer dan Multimedia)	Menggunakan algoritma Naive Bayes untuk mengklasifikasikan penerima bantuan pendidikan berdasarkan data yang telah ditentukan sebagai dasar pengambilan keputusan.	Hasil penelitian menunjukkan bahwa algoritma Naive Bayes mampu membantu proses klasifikasi dan menghasilkan keputusan yang efektif dalam menentukan penerima bantuan pendidikan.	Keunggulan: menerapkan algoritma Naive Bayes pada bidang pendidikan sehingga mendukung proses pengambilan keputusan. Kelemahan: objek penelitian berfokus pada penentuan penerima bantuan pendidikan, bukan analisis minat memilih sekolah.	Penelitian tersebut menerapkan algoritma Naive Bayes untuk klasifikasi penerima bantuan pendidikan, sedangkan penelitian ini menggunakan algoritma Naive Bayes untuk mengklasifikasikan tingkat minat murid dalam memilih MTs Mambaul Ulum berdasarkan atribut lokasi sekolah, kualitas

						sekolah, fasilitas sekolah, dan biaya pendidikan.
8	Analisis Kelayakan Pemilihan Jurusan Siswa-Siswi SMA/Sederajat di Era Merdeka Belajar Menggunakan Algoritma Naïve Bayes [15]	N. K. Utami, R. I. Borman, A. Wantoro, et al., 2025 (ZONasi: Jurnal Sistem Informasi)	Menggunakan algoritma Naive Bayes untuk mengklasifikasikan kelayakan pemilihan jurusan siswa berdasarkan data akademik dan data pendukung lainnya.	Hasil penelitian menunjukkan bahwa algoritma Naive Bayes dapat digunakan untuk membantu proses klasifikasi pemilihan jurusan dengan tingkat akurasi sebesar 65%.	Keunggulan: menerapkan algoritma Naive Bayes pada bidang pendidikan untuk membantu pengambilan keputusan pemilihan jurusan. Kelemahan: objek penelitian berfokus pada pemilihan jurusan siswa, bukan minat memilih sekolah.	Penelitian tersebut menggunakan algoritma Naive Bayes untuk mengklasifikasikan kelayakan pemilihan jurusan siswa, sedangkan penelitian ini menggunakan algoritma Naive Bayes untuk mengklasifikasikan tingkat minat murid dalam memilih MTs Mambaul Ulum berdasarkan atribut lokasi sekolah, kualitas sekolah, fasilitas sekolah, dan biaya pendidikan.

Berdasarkan hasil perbandingan penelitian terdahulu yang telah disajikan, penelitian ini memiliki perbedaan pada konteks penerapan algoritma *Naive Bayes*. Berdasarkan literatur yang ditelaah dalam penelitian ini, penelitian terdahulu yang berkaitan dengan pemilihan sekolah sebagian besar membahas berbagai pertimbangan dan keputusan dalam memilih sekolah menggunakan pendekatan kuantitatif, sedangkan penelitian lainnya menerapkan algoritma *Naive Bayes* pada bidang pendidikan untuk klasifikasi penerima bantuan pendidikan dan pemilihan jurusan. Penelitian ini menerapkan algoritma *Naive Bayes* untuk mengklasifikasikan tingkat minat murid yang telah memilih MTs Mambaul Ulum berdasarkan data hasil kuesioner. Atribut yang digunakan meliputi lokasi sekolah, kualitas sekolah, fasilitas sekolah, dan biaya pendidikan sebagai atribut prediktor, sedangkan minat digunakan sebagai label atau kelas yang diklasifikasikan.

Penelitian ini berfokus pada penerapan klasifikasi tingkat minat murid dalam konteks pemilihan sekolah pada jenjang Madrasah Tsanawiyah, khususnya di MTs Mambaul Ulum. Perbedaan tersebut terletak pada konteks objek penelitian, atribut prediktor yang digunakan, serta penerapan algoritma *Naive Bayes* untuk mengklasifikasikan tingkat minat berdasarkan persepsi responden terhadap atribut lokasi sekolah, kualitas sekolah, fasilitas sekolah, dan biaya pendidikan. Hasil penelitian diharapkan dapat memberikan informasi mengenai karakteristik tingkat minat murid yang telah memilih MTs Mambaul Ulum serta menjadi bahan pertimbangan bagi pihak sekolah dalam menyusun strategi promosi, meningkatkan kualitas layanan pendidikan, dan mendukung pengambilan keputusan pada pelaksanaan Sistem Penerimaan Murid Baru (SPMB).