



Implementasi Naïve Bayes Classifier pada Sistem Informasi untuk Klasifikasi Ketepatan Kelulusan Santri

Pramuditha Shinta Dewi Puspitasari*, Taufiq Rizaldy, Faisal Lutfi Afriansyah, Zhaqian Rouf Alfauzi

Teknologi Informasi, Manajemen Informatika, Politeknik Negeri Jember, Jember

Jl. Mastrip, Krajan Timur, Sumbersari, Kec. Sumbersari, Kabupaten Jember, Jawa Timur, Indonesia

Email: ¹*pramuditha@polije.ac.id, ²taufiq_r@polije.ac.id, ³faisal.lutfi@polije.ac.id, ⁴zroufalfauzi@gmail.com

Email Penulis Korespondensi: pramuditha@polije.ac.id

Submitted: 03/07/2026; Accepted: 20/07/2026; Published: 20/07/2026

Abstrak—Pondok Pesantren Mahasiswa merupakan model pendidikan yang mengintegrasikan pendidikan tinggi dengan sistem kepesantrenan untuk menghasilkan lulusan yang memiliki kompetensi akademik dan karakter religius. Namun, beban ganda dari tuntutan akademik dan kepesantrenan menyebabkan tingginya risiko santri mengalami keterlambatan kelulusan. Kondisi ini menuntut adanya intervensi dini yang krusial, di mana keterlambatan penanganan dapat berakibat pada kegagalan studi. Merespons urgensi tersebut, penelitian ini bertujuan mengembangkan sistem berbasis website yang mampu mengklasifikasikan ketepatan kelulusan santri menggunakan algoritma Naïve Bayes. Data penelitian diperoleh dari arsip PPM yang terdiri atas 211 data santri dengan atribut jenis kelamin, asal daerah, lama studi, capaian Al-Qur'an, dan capaian Al-Hadis. Sistem berhasil mengklasifikasikan kelulusan santri dengan hasil prediksi 172 santri berstatus tepat waktu dan 39 santri berstatus terlambat. Evaluasi menggunakan confusion matrix menghasilkan accuracy sebesar 86,79%, precision sebesar 93,41%, dan recall sebesar 91,40%, yang menunjukkan bahwa model memiliki kinerja klasifikasi yang baik. Selain itu, User Acceptance Testing memperoleh tingkat penerimaan sebesar 95,66% dari santri dan 84% dari dewan guru dengan kategori sangat baik. Kontribusi penelitian ini adalah mengintegrasikan atribut akademik dan kepesantrenan, khususnya capaian Al-Qur'an dan Al-Hadis, dalam model Naïve Bayes serta mengimplementasikannya ke dalam sistem berbasis web yang mendukung monitoring, evaluasi, dan pengambilan keputusan terkait ketepatan kelulusan santri.

Kata kunci: Naïve Bayes; Data Mining; Klasifikasi; Kelulusan Tepat Waktu; Monitoring; CRISP-DM; Siswa; Website

Abstract—Pondok Pesantren Mahasiswa (PPM) is an educational model that integrates higher education with the Islamic boarding school system to produce graduates with strong academic competence and religious character. However, the dual demands of academic and boarding school activities increase the risk of delayed graduation among students, highlighting the need for early intervention to prevent study failure. This study aims to develop a web-based system that classifies students' on-time graduation using the Naïve Bayes algorithm. The dataset was obtained from PPM archives and consisted of 211 student records with attributes including gender, region of origin, study duration, Al-Qur'an achievement, and Al-Hadith achievement. The proposed system classified 172 students as likely to graduate on time and 39 students as likely to graduate late. Performance evaluation using a confusion matrix achieved an accuracy of 86.79%, a precision of 93.41%, and a recall of 91.40%, indicating good classification performance. Furthermore, User Acceptance Testing (UAT) showed high user satisfaction, with acceptance rates of 95.66% from students and 84% from boarding school administrators, both categorized as excellent. The main contribution of this study is the integration of academic and Islamic boarding school attributes, particularly Al-Qur'an and Al-Hadith achievements, into a Naïve Bayes classification model and its implementation in a web-based system to support monitoring, evaluation, and decision-making related to students' on-time graduation.

Keywords: Naïve Bayes; Data Mining; Classification; Graduation Classification; Student Monitoring; CRISP-DM; Website

1. PENDAHULUAN

Dalam era globalisasi yang terus berkembang, pendidikan memiliki peran yang sangat penting sebagai sarana strategis dalam membentuk dan menyesuaikan masyarakat terhadap berbagai perubahan yang terjadi. Pendidikan tidak hanya berfungsi sebagai proses penyampaian ilmu pengetahuan, tetapi juga sebagai upaya komprehensif untuk mengembangkan pola pikir, meningkatkan kompetensi, serta menanamkan nilai-nilai positif yang mendukung kehidupan [1][2][3]. Dalam konteks pendidikan Islam di Indonesia, integrasi antara pendidikan tinggi dan sistem pendidikan pesantren melahirkan model Pondok Pesantren Mahasiswa (PPM)[4]. Model pendidikan ini dikembangkan untuk menghasilkan lulusan yang tidak hanya memiliki kompetensi akademik, tetapi juga berkarakter religius dan berakhlak mulia sehingga mampu menjawab tuntutan profesionalisme sekaligus kebutuhan pembinaan spiritual[5], [6].

Kendati demikian, beban ganda yang dipikul oleh mahasiswa yang sekaligus berstatus sebagai santri (mahasantri) sering kali memicu dinamika psikologis tersendiri, seperti tekanan akademik dan manajemen waktu yang ketat demi mencapai target kelulusan tepat waktu [7]. Oleh karena itu, setiap lembaga pendidikan dituntut untuk menghasilkan lulusan yang berkualitas, salah satunya melalui pencapaian tingkat kelulusan tepat waktu. Ketepatan kelulusan menjadi salah satu indikator mutu pembelajaran karena mencerminkan efektivitas proses pendidikan, kurikulum, serta sistem evaluasi yang diterapkan [8], [9].

Informasi mengenai potensi kelulusan tepat waktu diperlukan sebagai dasar evaluasi dan pengambilan keputusan bagi pengelola pesantren maupun santri. Beberapa penelitian terdahulu telah mengkaji prediksi kelulusan mahasiswa menggunakan berbagai algoritma klasifikasi[10], [11], [12], [13]. Salah satunya, penelitian yang dilakukan oleh Hozairi membandingkan kinerja tiga algoritma, yaitu k-Nearest Neighbor (K-NN), Naïve Bayes, dan Decision Tree, untuk menentukan metode yang memberikan performa terbaik dalam memprediksi

kelulusan mahasiswa [12]. Dari hasil penelitian tersebut Naïve Bayes memberikan performa terbaik dengan tingkat akurasi sebesar 89%, sementara K-NN dan Decision Tree masing-masing menghasilkan akurasi sebesar 77% dan 74%. Selain itu dari hasil penelitian sistematis oleh Setiyani menyimpulkan bahwa metode data mining naïve bayes dapat membuat suatu prediksi mengenai kelulusan bergantung pada atributnya [11]. Pada penelitian Sriani melakukan prediksi kelulusan santri tahfidz Qur'an dengan menggunakan teknik klasifikasi decision tree algoritma C4.5 [14]. Inputan yang digunakan berupa data santri yang meliputi dari prestasi, kedisiplinan, hafalan dan lafadz.

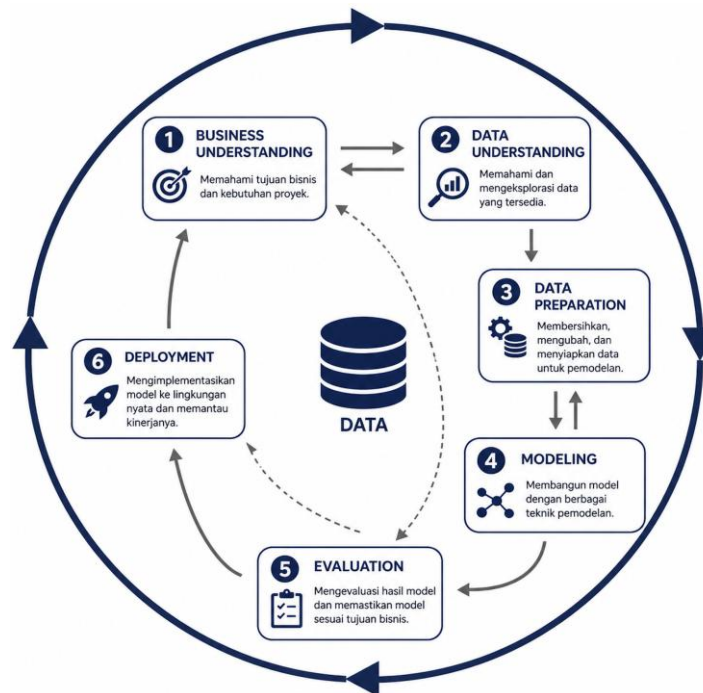
Beberapa penelitian telah memanfaatkan teknik data mining untuk memprediksi kelulusan mahasiswa maupun santri dengan berbagai pendekatan. Penelitian yang dilakukan oleh Hozairi membandingkan algoritma K-Nearest Neighbor (K-NN), Naïve Bayes, dan Decision Tree dalam memprediksi kelulusan mahasiswa [12]. Hasil penelitian menunjukkan bahwa Naïve Bayes memiliki performa terbaik dengan tingkat akurasi sebesar 89%. Meskipun demikian, penelitian tersebut hanya menggunakan atribut akademik mahasiswa sehingga belum mempertimbangkan karakteristik pendidikan berbasis pesantren yang memiliki indikator capaian kepesantrenan. Selanjutnya, Setiyani [11] melalui kajian sistematis menyimpulkan bahwa algoritma Naïve Bayes efektif digunakan untuk memprediksi kelulusan berdasarkan atribut yang dimiliki. Namun, penelitian tersebut masih bersifat umum karena hanya PPM. Penelitian lain yang dilakukan oleh Sriani [14] memprediksi kelulusan santri tahfidz Al-Qur'an menggunakan algoritma Decision Tree C4.5 dengan atribut prestasi, kedisiplinan, hafalan, dan lafaz. Walaupun telah mempertimbangkan karakteristik kepesantrenan, penelitian tersebut belum menerapkan pendekatan probabilistik Naïve Bayes serta belum mengintegrasikan indikator capaian akademik mahasiswa. Pada penelitian Junaidi [15] menerapkan algoritma Naïve Bayes untuk memprediksi kelulusan mahasiswa menggunakan atribut jenis kelamin, penghasilan orang tua, lama bimbingan, status bekerja, nilai semester, dan IPK. Penelitian tersebut menunjukkan bahwa atribut akademik berpengaruh terhadap prediksi kelulusan, namun belum memasukkan aspek kepesantrenan sebagai variabel prediksi. Selain itu pada penelitian oleh Kuncoro [16] menerapkan algoritma Naïve Bayes yang diintegrasikan dengan AdaBoost untuk mengklasifikasikan kualitas hafalan tematik santri berdasarkan empat atribut, yaitu jumlah ayat yang dihafal, nilai tajwid, waktu yang dibutuhkan dalam menghafal, dan nilai pemahaman tematik. Fokus penelitian tersebut adalah mengelompokkan kualitas hafalan santri ke dalam kategori tinggi, sedang, dan rendah berdasarkan hasil evaluasi hafalan. Meskipun berhasil meningkatkan performa klasifikasi, penelitian tersebut hanya memanfaatkan atribut yang berkaitan dengan kemampuan hafalan dan belum mempertimbangkan faktor akademik maupun pembinaan santri yang berpengaruh terhadap keberhasilan studi. Selain itu, luaran penelitian masih terbatas pada klasifikasi kualitas hafalan dan belum diarahkan untuk memprediksi kelulusan santri atau mendukung proses monitoring akademik.

Berdasarkan penelitian-penelitian tersebut, masih terdapat beberapa kesenjangan penelitian (research gap). Pertama, belum terdapat penelitian yang secara khusus menerapkan algoritma Naïve Bayes pada lingkungan Pondok Pesantren Mahasiswa yang memiliki karakteristik ganda, yaitu tuntutan akademik di perguruan tinggi dan pembinaan kepesantrenan. Kedua, klasifikasi prestasi akademik santri di lingkungan pesantren, yang hingga kini masih relatif jarang diteliti dibandingkan topik klasifikasi prestasi akademik mahasiswa. Pada penelitian sebelumnya belum mengintegrasikan atribut akademik dengan capaian materi tafsir Al-Qur'an dan Al-Hadis sebagai indikator pencapaian kepesantrenan yang berkaitan dengan kemampuan regulasi diri dan resiliensi akademik mahasiswa [15]. Ketiga, sebagian besar penelitian terdahulu hanya berfokus pada pembangunan model klasifikasi tanpa mengimplementasikannya ke dalam sistem yang dapat dimanfaatkan secara langsung oleh pengelola pesantren sebagai media monitoring dan evaluasi. Oleh karena itu, penelitian ini mengembangkan model prediksi ketepatan kelulusan santri menggunakan algoritma Naïve Bayes dengan mengintegrasikan atribut akademik dan kepesantrenan, kemudian mengimplementasikannya dalam sistem berbasis web untuk mendukung proses monitoring, evaluasi, dan pengambilan keputusan di Pondok Pesantren Mahasiswa.

Pada penelitian ini, atribut capaian tafsir Al-Qur'an dan Al-Hadis digunakan berdasarkan teori resiliensi akademik [17], yang menyatakan bahwa kemampuan mahasiswa dalam menyelesaikan target kepesantrenan mencerminkan regulasi diri dan ketahanan akademik yang berkontribusi terhadap keberhasilan studi. Selain mempertimbangkan atribut akademik, penelitian ini mengintegrasikan indikator capaian kepesantrenan sebagai variabel prediksi untuk merepresentasikan karakteristik khas lingkungan PPM. Pendekatan ini mengacu pada penelitian sebelumnya yang menunjukkan bahwa prestasi akademik santri dapat diklasifikasikan berdasarkan kombinasi atribut akademik dan nonakademik [18]. Oleh karena itu, penelitian ini menggunakan jumlah capaian materi tafsir Al-Qur'an dan Al-Hadis sebagai atribut utama dalam mengklasifikasikan potensi kelulusan tepat waktu santri. Algoritma Naïve Bayes dipilih karena mampu menghitung probabilitas setiap atribut terhadap kelas yang diprediksi sehingga menghasilkan proses klasifikasi yang sederhana dan efektif. Hasil klasifikasi kemudian diimplementasikan dalam sistem berbasis web untuk mendukung proses monitoring dan evaluasi capaian kelulusan santri di PPM. Kontribusi penelitian ini terletak pada integrasi atribut akademik dan kepesantrenan dalam model prediksi berbasis Naïve Bayes serta implementasinya ke dalam sistem berbasis web yang dapat dimanfaatkan sebagai alat bantu monitoring dan pengambilan keputusan bagi pengelola PPM. Sistem ini mendukung dewan guru dalam mendeteksi secara dini potensi keterlambatan kelulusan santri sehingga proses pembinaan, pendampingan, dan evaluasi dapat dilakukan secara lebih efektif, objektif, dan berbasis data.

2. METODOLOGI PENELITIAN

Penelitian ini menerapkan pendekatan analisis data untuk mengelompokkan kelulusan tepat waktu dengan memanfaatkan data historis. Proses penelitian mengacu pada metodologi Cross-Industry Standard Process for Data Mining (CRISP-DM) yang masih menjadi kerangka kerja paling banyak digunakan dalam pengembangan proyek data mining dan data science[20], [21].



Gambar 1. Tahapan dalam Metode CRISP-DM

Metodologi CRISP-DM tetap menjadi kerangka kerja yang relevan dalam proyek Data Science karena bersifat kuat, fleksibel, dan mudah disesuaikan dengan berbagai kebutuhan organisasi [22]. Seperti yang ditampilkan pada Gambar 1, metodologi ini terdiri atas enam tahapan utama. Tahapan tersebut dimulai dari Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, dan Deployment[23]. Keenam tahapan tersebut membentuk alur kerja yang sistematis, mulai dari identifikasi permasalahan, pemahaman serta persiapan data, pembangunan model, evaluasi kinerja model, hingga implementasi hasil, sehingga mampu menghasilkan model yang sesuai dengan tujuan penelitian[24]. Berikut merupakan penjelasan dari tahapan CRISP-DM:

2.1 Business Understanding

Tahap Business Understanding bertujuan untuk memahami tujuan, kebutuhan, dan permasalahan bisnis yang menjadi dasar pelaksanaan proyek [24]. Pada penelitian ini, proses dilakukan dengan memahami kebutuhan PPM dalam monitoring dan evaluasi proses pembelajaran. Proses pengumpulan data dan informasi ini dilakukan dengan beberapa tahap. Sehingga dalam prosesnya melibatkan pihak yang memiliki informasi dan paham mengenai data yang diperlukan serta yang dibutuhkan sebagai tolak ukur keputusan tepat waktu kelulusan santri. Berdasarkan hasil wawancara dan studi literatur, parameter kelulusan santri ditentukan oleh pencapaian materi Al-Qur'an dan Al-Hadis yang telah divalidasi oleh dewan guru melalui proses munaqosah. Oleh karena itu, sistem yang dikembangkan dilengkapi dengan fitur munaqosah untuk mendukung proses validasi tersebut. Selain itu, diperlukan data sebagai dasar klasifikasi ketepatan kelulusan santri berdasarkan tahun angkatan, jenis kelamin, dan asal daerah. Data penelitian diperoleh dari arsip PPM meliputi data santri angkatan 2020–2023 yang berasal dari arsip munaqosah dan data administrasi santri. Berdasarkan kebutuhan tersebut, dibangun sebuah sistem klasifikasi kelulusan tepat waktu yang mengacu pada metode yang telah dibuktikan pada penelitian terdahulu.

Pada Tabel 1 merupakan atribut-atribut hasil studi literatur dari penelitian-penelitian terdahulu yang menggunakan berbagai atribut dalam proses klasifikasi dan prediksi kelulusan. Suardika pada penelitiannya [25], memanfaatkan atribut nama program studi, nama fakultas, angkatan, tanggal wisuda, IPK, dan masa studi untuk memprediksi kelulusan mahasiswa. Penelitian oleh Kuncoro menerapkan algoritma Naïve Bayes yang diintegrasikan dengan AdaBoost untuk mengklasifikasikan kualitas hafalan tematik santri berdasarkan empat atribut, yaitu jumlah ayat yang dihafal, nilai tajwid, waktu yang dibutuhkan dalam menghafal, dan nilai pemahaman tematik[16]. Penelitian Hasibuan dan Mahdiana memprediksi kelulusan tepat waktu mahasiswa menggunakan atribut akademik dan demografi, yaitu jenis kelamin, fakultas, program studi, jalur masuk, IPS

semester 1–8, dan IPK [25]. Sementara itu, Junaidi menerapkan beberapa algoritma klasifikasi, yaitu Naïve Bayes, Random Forest, Support Vector Machine (SVM), dan Artificial Neural Network (ANN) untuk memprediksi tingkat prestasi akademik mahasiswa. Penelitian tersebut menggunakan sejumlah atribut, yaitu jenis kelamin, penghasilan orang tua, lama bimbingan, status mahasiswa bekerja atau tidak, nilai semester 1 hingga semester 8, serta Indeks Prestasi Kumulatif (IPK) [15].

Pada konteks pendidikan pesantren, Sriani menggunakan atribut prestasi, kedisiplinan, hafalan, dan lafadz dalam memprediksi kelulusan santri tahfidz Al-Qur'an[14]. Penelitian oleh Syafrizal menjelaskan bahwa pembelajaran tahfidz Al-Qur'an dan Al-Hadis di pesantren dilaksanakan secara sistematis melalui proses pembinaan, setoran hafalan, murojaah, dan evaluasi berkala untuk mengukur ketercapaian target hafalan santri. Capaian Al-Qur'an dan Al-Hadis tidak hanya mencerminkan kemampuan hafalan, tetapi juga menjadi indikator keberhasilan santri dalam mengikuti proses pembinaan yang telah ditetapkan oleh pesantren[18]. Oleh karena itu, capaian Al-Qur'an dan Al-Hadis dapat dijadikan sebagai atribut yang relevan dalam memprediksi kelulusan tepat waktu mahasantri, khususnya pada pesantren yang menetapkan target capaian hafalan sebagai bagian dari persyaratan penyelesaian program pembinaan. Berdasarkan penelitian-penelitian tersebut, dapat disimpulkan bahwa pemilihan atribut disesuaikan dengan karakteristik objek penelitian. Berbeda dengan penelitian sebelumnya, penelitian ini berfokus pada atribut yang secara langsung menjadi indikator kelulusan santri di PPM, yaitu capaian materi Al-Qur'an dan Al-Hadis, serta didukung oleh atribut tahun angkatan, jenis kelamin, dan asal daerah sebagai faktor pendukung dalam proses klasifikasi kelulusan tepat waktu.

Tabel 1. Atribut-atribut yang digunakan pada penelitian terdahulu

Peneliti	Atribut yang Digunakan	Metode
Suardika	Nama Program Studi, Nama Fakultas, Angkatan, Tanggal Wisuda, IPK, Masa Studi [25].	Naïve Bayes
Kuncoro	Jumlah ayat yang dihafal, Nilai tajwid, waktu yang dibutuhkan dalam menghafal, dan nilai pemahaman tematik [16].	Naïve Bayes
Hasibuan dan Mahdiana	Jenis kelamin, fakultas, program studi, jalur masuk, IPS semester 1–8, dan IPK [26].	Algoritma C4.5
Junaidi	Jenis kelamin, penghasilan orang tua, lama bimbingan, status mahasiswa bekerja atau tidak, nilai semester 1 hingga semester 8, serta Indeks Prestasi Kumulatif (IPK) [15].	Naïve Bayes, Random Forest, Support Vector Machine, dan Artificial Neural Network
Sriani	Prestasi, kedisiplinan, hafalan, dan lafadz [14].	Decision Tree (C4.5)

2.2 Data Understanding

Tahap Data Understanding berfokus pada proses pengumpulan dan eksplorasi data awal untuk memperoleh pemahaman mengenai karakteristik dataset. Melalui tahap ini dilakukan identifikasi struktur, kualitas, dan pola data sehingga dapat diperoleh wawasan awal serta hipotesis yang akan mendukung proses analisis selanjutnya.

Pemahaman yang diambil dari proses wawancara mendapatkan bahwa diperlukan data tentang klasifikasi kelulusan berdasarkan capaian yang diraih tiap santri. Kemudian tahap Data Understanding, dilakukan proses identifikasi dan pemahaman terhadap dataset yang digunakan dalam penelitian. Tahap ini bertujuan untuk memahami karakteristik, struktur, dan kualitas data sebelum dilakukan proses pengolahan lebih lanjut. Dataset yang digunakan berasal dari data yang tersedia pada arsip sekretaris PPM. Data yang digunakan dan diolah dalam penelitian ini merupakan data munaqosah santri serta data santri yang diperoleh dari arsip sekretaris PPM. Sehingga atribut yang digunakan pada penelitian ini yaitu Nama, NIS, Jenis Kelamin, Tahun Angkatan, Asal Daerah, Halaman tafsir Al-Quran terisi, Halaman tafsir Al Hadis terisi, dan Status Capaian.

2.2.1 Data Preparation

Tahap Data Preparation bertujuan menyiapkan dataset yang akan digunakan pada proses pemodelan. Kegiatan yang dilakukan meliputi transformasi data arsip fisik yang dikelola PPM menjadi data digital. Kemudian dilakukan pemilihan atribut yang relevan, transformasi data, serta pembersihan data dari nilai yang tidak valid atau tidak lengkap. Tahap ini bersifat iteratif sehingga dapat dilakukan berulang kali sesuai kebutuhan hingga diperoleh dataset yang siap digunakan.

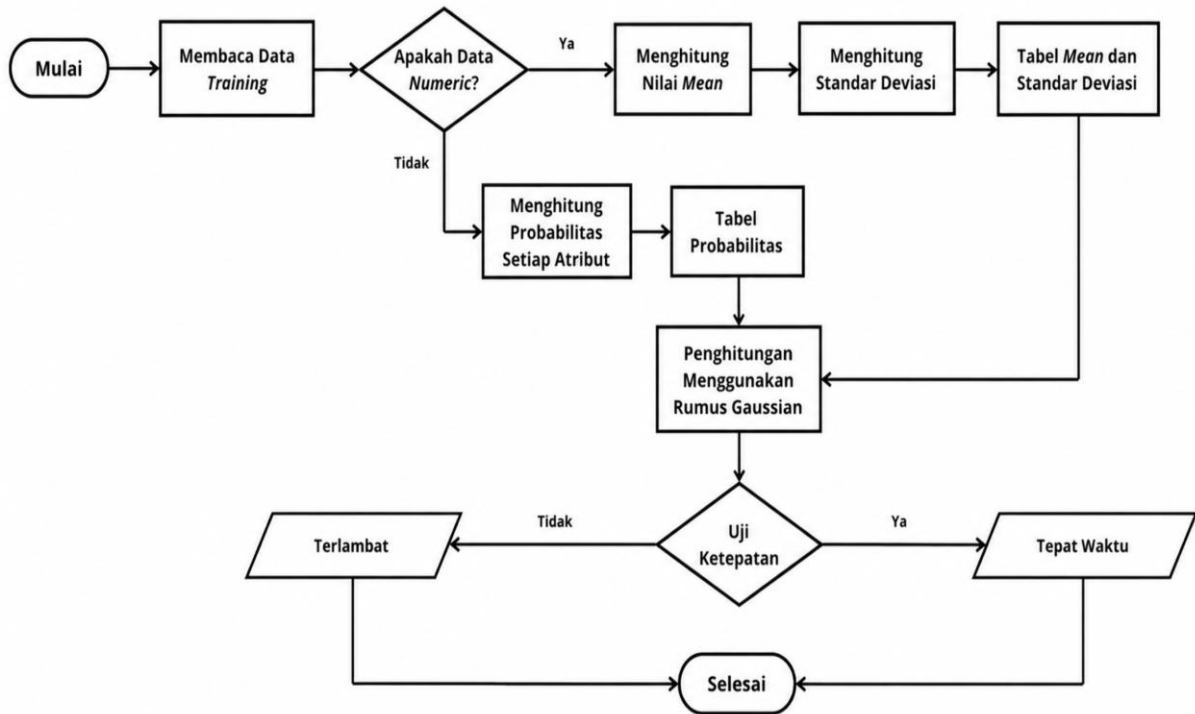
Pada tahap ini dilakukan proses data cleaning untuk meningkatkan kualitas data dengan mengidentifikasi dan menangani data yang tidak lengkap (missing value). Nilai yang kosong atau null pada atribut tertentu dihapus sehingga dataset yang digunakan menjadi lebih konsisten dan dapat menghasilkan analisis yang lebih akurat. Pada penelitian ini, proses pembersihan data dilakukan pada atribut yang masih memiliki beberapa nilai kosong.

2.2.2 Modeling

Tahap Modeling merupakan proses pembangunan model menggunakan teknik data mining yang sesuai dengan tujuan penelitian. Pada tahap ini dilakukan pemilihan algoritma Naïve Bayes, penentuan parameter (Jenis Kelamin,

Asal Daerah, Status Capaian, dan Tahun Angkatan). Apabila selama proses ditemukan permasalahan pada dataset, peneliti dapat kembali ke tahap sebelumnya untuk melakukan perbaikan atau penyesuaian data.

Naïve Bayes merupakan algoritma klasifikasi berbasis probabilitas yang menerapkan Teorema Bayes dengan asumsi bahwa setiap atribut bersifat independen terhadap atribut lainnya [27]. Algoritma ini digunakan untuk memperkirakan probabilitas suatu data termasuk ke dalam kelas tertentu berdasarkan nilai atribut yang dimilikinya. Metode ini memiliki kemampuan klasifikasi yang kompetitif dibandingkan algoritma lain, sehingga banyak digunakan dalam berbagai permasalahan klasifikasi, seperti prediksi, analisis sentimen, diagnosis, dan data mining.



Gambar 1. Flowchart Penerapan Algoritma Naïve Bayes Classifier

Proses modeling diawali dengan menggunakan data hasil praproses sebagai data latih yang terdiri atas atribut kategorikal dan numerikal. Gambar 2 merupakan tahap penghitungan menggunakan algoritma Naïve Bayes Classifier.

2.3 Evaluation

Tahap Evaluation bertujuan mengevaluasi hasil model yang telah dibangun untuk memastikan bahwa model mampu memenuhi tujuan bisnis maupun tujuan penelitian. Evaluasi dilakukan dengan meninjau kembali seluruh proses yang telah dilaksanakan serta mengukur kinerja model menggunakan metrik evaluasi yang telah ditentukan.

2.4 Deployment

Tahap Deployment merupakan proses penerapan model agar dapat dimanfaatkan oleh pengguna atau organisasi. Implementasi model berupa penyajian laporan hasil analisis, pengembangan system yang mengintegrasikan model ke dalam sistem informasi berbasis web sehingga dapat digunakan secara berkelanjutan dalam mendukung pengambilan Keputusan PPM. Seluruh fitur diimplementasikan menggunakan framework Laravel dengan bahasa pemrograman PHP, sedangkan algoritma Naïve Bayes Classifier diterapkan melalui library PHP-ML.

3. HASIL DAN PEMBAHASAN

3.1 Hasil

Data yang diperoleh dari data santri, data pribadi munaqosah, dan data rekapitulasi munaqosah kemudian disusun sesuai kebutuhan penelitian. Selanjutnya dilakukan transformasi atribut, yaitu Asal Daerah, Status Capaian, dan Tahun Angkatan. Atribut Asal Daerah dikelompokkan menjadi Dalam Provinsi (Jawa Timur) dan Luar Provinsi (selain Jawa Timur). Selanjutnya, atribut Status Capaian ditransformasikan menjadi data kategorikal berdasarkan persentase pencapaian target yang ditetapkan oleh PPM. Santri dikategorikan Tepat apabila mampu menyelesaikan target 2.603 halaman, yang terdiri atas 606 halaman Al-Qur'an dan 1.997 halaman Al-Hadis, dalam waktu 3 tahun (1.095 hari); sedangkan capaian di bawah target dikategorikan Terlambat. Persentase capaian dihitung

menggunakan rumus 1, dengan ketentuan bahwa n merupakan nilai persentase pencapaian target, x menyatakan jumlah halaman tafsir yang telah diselesaikan hingga saat ini, dan y menyatakan jumlah hari yang telah dilalui selama periode pencapaian target.

$$n = \left(\frac{x}{y \times \frac{2780}{1095}} \right) \times 100 \quad (1)$$

Apabila nilai n mendapatkan hasil $\geq 100\%$ maka atribut Status Capaian diklasifikasikan menjadi Tepat dan Terlambat apabila capaian $< 100\%$. Setelah nilai Status Capaian diperoleh, atribut tersebut digabungkan dengan atribut lainnya untuk membentuk data latih yang akan digunakan pada proses klasifikasi menggunakan algoritma Naïve Bayes Classifier pada tahap modeling. Sementara itu, atribut Tahun Angkatan dikonversi menjadi lama studi santri yang dihitung sejak tahun masuk hingga waktu penelitian. Setelah proses data cleaning selesai, dipilih atribut-atribut yang digunakan seperti Tabel 2 merupakan hasil dari data preparation.

Tabel 2. Hasil data preparation

No	Inisial Nama	Jenis Kelamin	NIS	Asal Daerah	Tahun Angkatan	Capaian Al Qur'an	Capaian Hadis	Status Capaian
1	ABE	Laki-laki	A21001	Dalam Prov	2021	306	1200	Tidak Tercapai
2	AIA	Laki-laki	A24016	Dalam Prov	2021	186	1075	Tidak Tercapai
3	ARI	Laki-laki	A22002	Dalam Prov	2022	159	1511	Tidak Tercapai
4	ASD	Laki-laki	A22021	Luar Prov	2022	606	2174	Tercapai
5	ARF	Laki-laki	A23001	Dalam Prov	2023	117	212	Tidak Tercapai
6	AFH	Laki-laki	A23002	Luar Prov	2023	606	555	Tercapai
s/d	...	...	...	...	...	...	...	...
211	AD	Laki-laki	A21201	Dalam Pro	2021	606	2174	Tercapai

Tahap pertama dalam proses Naïve Bayes Classifier adalah menghitung probabilitas prior pada atribut Status Capaian. Nilai probabilitas diperoleh dari perbandingan jumlah data pada setiap kelas dengan total data latih, sehingga diperoleh $P(\text{Tercapai}) = 186/211 = 0,88$ dan $P(\text{Tidak Tercapai}) = 25/211 = 0,12$.

Setelah probabilitas prior setiap kelas diperoleh, tahap berikutnya adalah menghitung probabilitas kondisional untuk setiap atribut, yaitu jenis kelamin, asal daerah, tahun angkatan, capaian Al-Qur'an, dan capaian Al-Hadis. Perhitungan dilakukan berdasarkan tipe data atribut. Atribut tahun angkatan, capaian Al-Qur'an, dan capaian Al-Hadis diproses sebagai data numerik, sedangkan jenis kelamin dan asal daerah diproses sebagai data kategorikal.

Pada data numerik, probabilitas untuk tahun angkatan yang tercapai memiliki nilai mean 2022 dengan standar deviasi 0,65, sementara probabilitas yang tidak tercapai memiliki nilai rata-rata 2024 dengan standar deviasi yang lebih tinggi, yaitu 1,29.

Tabel 3. Tabel Probabilitas Capaian Al Quran

Probabilitas	Mean	Standar Deviasi
Tercapai	523,68	128,45
Tidak Tercapai	286,28	192,62

Berdasarkan Tabel 3 dan Tabel 4, kelas Tercapai memiliki nilai rata-rata capaian Al-Qur'an sebesar 523,68 dan capaian Al-Hadis sebesar 1.923,70, sedangkan kelas Tidak Tercapai memiliki rata-rata masing-masing sebesar 286,28 dan 638,82. Nilai standar deviasi menunjukkan bahwa variasi data pada kelas Tidak Tercapai lebih besar dibandingkan kelas Tercapai, terutama pada atribut capaian Al-Qur'an.

Tabel 4. Tabel Probabilitas Capaian Al Hadis

Probabilitas	Mean	Standar Deviasi
Tercapai	1923,70	361,60
Tidak Tercapai	638,82	390,85

Tabel 5. Tabel Probabilitas Jenis Kelamin

Probabilitas	Laki-laki	Perempuan
Tercapai	$61/186 = 0,32$	$125/186 = 0,67$
Tidak Tercapai	$12/25 = 0,48$	$13/25 = 0,52$

Pada data kategorik berdasarkan Tabel 5, atribut Jenis Kelamin menunjukkan bahwa probabilitas kelas Tercapai sebesar 0,32 untuk laki-laki dan 0,67 untuk perempuan. Sementara itu, pada kelas Tidak Tercapai, probabilitas laki-laki sebesar 0,48 dan perempuan sebesar 0,52.

Berdasarkan Tabel 6, atribut Asal Daerah menunjukkan bahwa pada kelas Tercapai, probabilitas santri yang berasal dari dalam provinsi sebesar 0,56, sedangkan luar provinsi sebesar 0,43. Adapun pada kelas Tidak Tercapai, probabilitas santri dari dalam provinsi sebesar 0,68, lebih tinggi dibandingkan luar provinsi yang sebesar 0,32.

Tabel 6. Tabel Probabilitas Asal Daerah

Probabilitas	Dalam Provinsi	Luar Provinsi
Tercapai	$105/186 = 0,56$	$81/186 = 0,43$
Tidak Tercapai	$17/25 = 0,68$	$8/25 = 0,32$

Dalam asal daerah yang dihitung pada Tabel 6, yang tercapai memiliki probabilitas 0,56 untuk dalam provinsi dan 0,43 untuk luar provinsi, sedangkan yang tidak tercapai memiliki probabilitas lebih tinggi untuk dalam provinsi (0,68) dan lebih rendah untuk luar provinsi (0,32).

Setelah probabilitas prior dan probabilitas kondisional setiap atribut diperoleh, langkah selanjutnya adalah menentukan hasil klasifikasi pada data santri menggunakan algoritma Naïve Bayes. Sebagai contoh, untuk data santri inisial ABE diperoleh nilai $P(\text{Tercapai}) = 0,000009$ dan $P(\text{Tidak Tercapai}) = 0,214$. Hasil klasifikasi ditentukan berdasarkan nilai probabilitas terbesar, yaitu apabila $P(\text{Tercapai}) > P(\text{Tidak Tercapai})$ maka santri diklasifikasikan Tepat Waktu, sedangkan apabila $P(\text{Tercapai}) < P(\text{Tidak Tercapai})$ maka diklasifikasikan Terlambat. Berdasarkan hasil tersebut, inisial ABE termasuk dalam kategori Terlambat. Proses perhitungan yang sama diterapkan pada seluruh data santri hingga diperoleh hasil klasifikasi yang disajikan pada Tabel 7.

Tabel 7. Hasil Klasifikasi

No	Inisial Nama	NIS	Status Capaian	Hasil Klasifikasi
1	ABE	A21001	Tidak Tercapai	Terlambat
2	AIA	A24016	Tidak Tercapai	Terlambat
3	ARI	A22002	Tidak Tercapai	Terlambat
4	ASD	A22021	Tercapai	Tepat
5	ARF	A23001	Tidak Tercapai	Terlambat
6	AFH	A23002	Tercapai	Tepat
..s/d	...	...	...	
211	AD	A21201	Tercapai	Tepat

Hasil klasifikasi kemudian dievaluasi menggunakan confusion matrix untuk mengukur kinerja model. Pengujian dilakukan dengan menggunakan 50% dari total dataset sebagai data uji, yaitu 106 data dari 211 data santri. Hasil evaluasi menggunakan confusion matrix disajikan pada Tabel 8.

Tabel 8. Tabel Penghitungan Akurasi data dengan Confusion Matrix

Correct Classification		Kelas Klasifikasi	
		Tepat	Terlambat
Kelas Aktual	Tercapai	85 (TP)	7 (TN)
	Tidak Tercapai	6 (FP)	8 (FN)

Kemudian diambil nilai accuracy dengan rumus :

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{Total Data Uji}} = \frac{85 + 7}{106} = 0,8679$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} = \frac{85}{85 + 6} = 0,934$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{85}{85 + 8} = 0,914$$

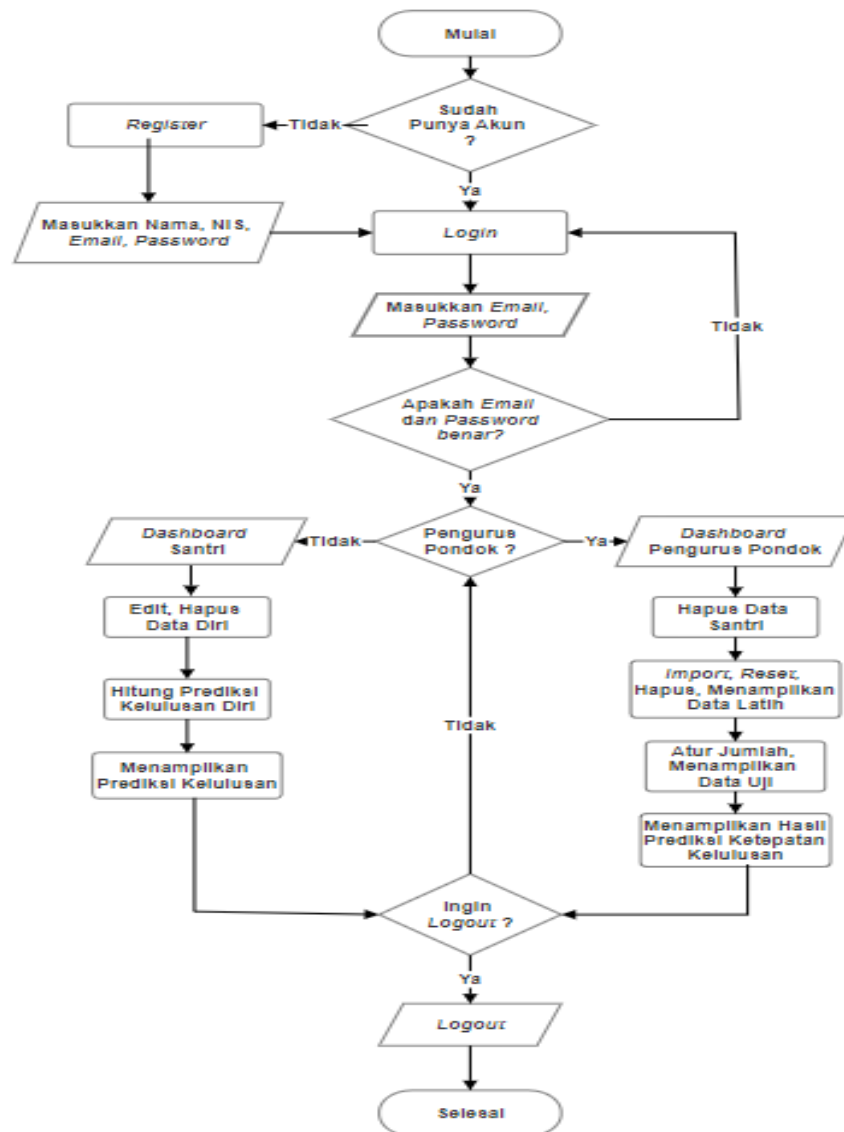
Berdasarkan hasil perhitungan nilai accuracy yaitu 86,79%, Precision 93,4%, dan Recall 91,40%, model klasifikasi Naïve Bayes yang dikembangkan menunjukkan kinerja yang sangat baik dan dapat diandalkan. Hal ini dibuktikan dengan tingkat akurasi sebesar 86,79%, yang berarti model secara umum mampu memprediksi kelas data santri (Tepat/Terlambat) dengan benar pada persentase tersebut. Selain itu, model ini memiliki tingkat presisi yang sangat tinggi, yaitu 93,4%, sehingga ketika model memprediksi seorang santri masuk dalam kategori capaian "Tepat", tebakan tersebut sangat terpercaya dengan tingkat salah prediksi yang sangat minim. Kinerja ini juga didukung oleh nilai recall sebesar 91,40%, yang mengindikasikan bahwa sensitivitas model sudah sangat baik dalam mendeteksi dan mengenali hampir seluruh data santri yang aktualnya memang berstatus "Tepat". Secara keseluruhan, tingginya ketiga indikator metrik ini menegaskan bahwa model tersebut sangat layak dan efektif untuk diimplementasikan sebagai sistem pendukung keputusan.

3.1. Implementasi

Selain menghasilkan model klasifikasi, penelitian ini juga mengimplementasikan model Naïve Bayes ke dalam aplikasi berbasis web yang mendukung proses prediksi ketepatan kelulusan, monitoring, evaluasi, dan

pengambilan keputusan bagi dewan guru. Sistem menyediakan fitur pengelolaan data latih, data uji, proses munaqosah, visualisasi hasil klasifikasi, serta pemantauan perkembangan capaian santri.

Pada tahap implementasi, rancangan sistem diwujudkan ke dalam aplikasi berbasis web. Berdasarkan hasil wawancara, sistem dilengkapi dengan fitur munaqosah santri yang berfungsi memvalidasi capaian Al-Qur'an dan Al-Hadis oleh dewan guru. Sistem dikembangkan dengan dua peran pengguna, yaitu santri dan admin. Pada Gambar 3 ditunjukkan alur penggunaan sistem, di mana santri dapat melakukan registrasi, login, melihat dashboard, mengubah profil, melakukan klasifikasi, dan melihat riwayat klasifikasi. Sementara itu, admin memiliki fitur dashboard, ubah password, munaqosah santri, kelola admin, kelola santri, data latih, data uji, dan hasil klasifikasi. Hasil implementasi antarmuka sistem ditampilkan pada Gambar 4 sampai dengan Gambar 3.



Gambar 3. Flowchart sistem dengan user santri dan dewan guru

Halaman dashboard admin pada Gambar 3 menyajikan visualisasi hasil klasifikasi kelulusan santri berdasarkan data hasil pengujian dalam bentuk kartu informasi, diagram, grafik, dan tabel untuk memudahkan proses monitoring serta analisis capaian santri. Dashboard menampilkan ringkasan jumlah santri, jumlah santri yang diprediksi lulus tepat waktu, tanggal dan waktu saat ini, distribusi hasil klasifikasi, grafik berdasarkan jenis kelamin dan asal daerah, serta rekapitulasi hasil klasifikasi tiap angkatan. Informasi tersebut membantu admin memantau perkembangan capaian santri dan mendukung proses evaluasi serta pengambilan keputusan.

Selain itu terdapat grafik perkembangan persentase capaian, jumlah capaian Al-Qur'an, dan capaian Al-Hadis berdasarkan tanggal munaqosah sehingga memudahkan admin dalam memantau perkembangan santri dari waktu ke waktu. Selain itu, tersedia tabel riwayat yang memuat informasi tanggal munaqosah, tahun angkatan, jumlah capaian Al-Qur'an dan Al-Hadis, nilai akhir, status klasifikasi kelulusan, hasil prediksi, serta aksi untuk mengelola data. Tampilan ini memberikan informasi yang lebih terstruktur sehingga proses monitoring, evaluasi perkembangan, dan verifikasi hasil klasifikasi dapat dilakukan secara lebih efektif.



Gambar 4. Halaman Dashboard Dewan Guru

Halaman Data Latih pada Gambar 4 digunakan untuk mengelola data yang akan digunakan sebagai data pelatihan (training data) pada proses klasifikasi menggunakan algoritma Naïve Bayes. Halaman ini menyediakan fitur untuk mengimpor data berformat .xlsx, mengeksport template data sebagai acuan pengisian, serta melakukan reset guna mengembalikan data latih sesuai dengan data munaqosah santri yang tersimpan pada sistem. Selain itu, halaman ini menampilkan tabel data latih yang berisi atribut-atribut yang digunakan dalam proses klasifikasi, meliputi nama santri, jenis kelamin, asal daerah, tahun angkatan, capaian Al-Hadis, capaian Al-Qur'an, status target, dan status kelulusan. Tabel juga dilengkapi dengan fitur pencarian, pengaturan jumlah data yang ditampilkan, serta aksi untuk menghapus data tertentu sehingga memudahkan admin dalam melakukan validasi, pembaruan, dan pengelolaan data sebelum proses pelatihan model dilakukan.

Data Latih
Admin / Data Latih

Tindakan

Choose File No file chosen Import Export Reset Data

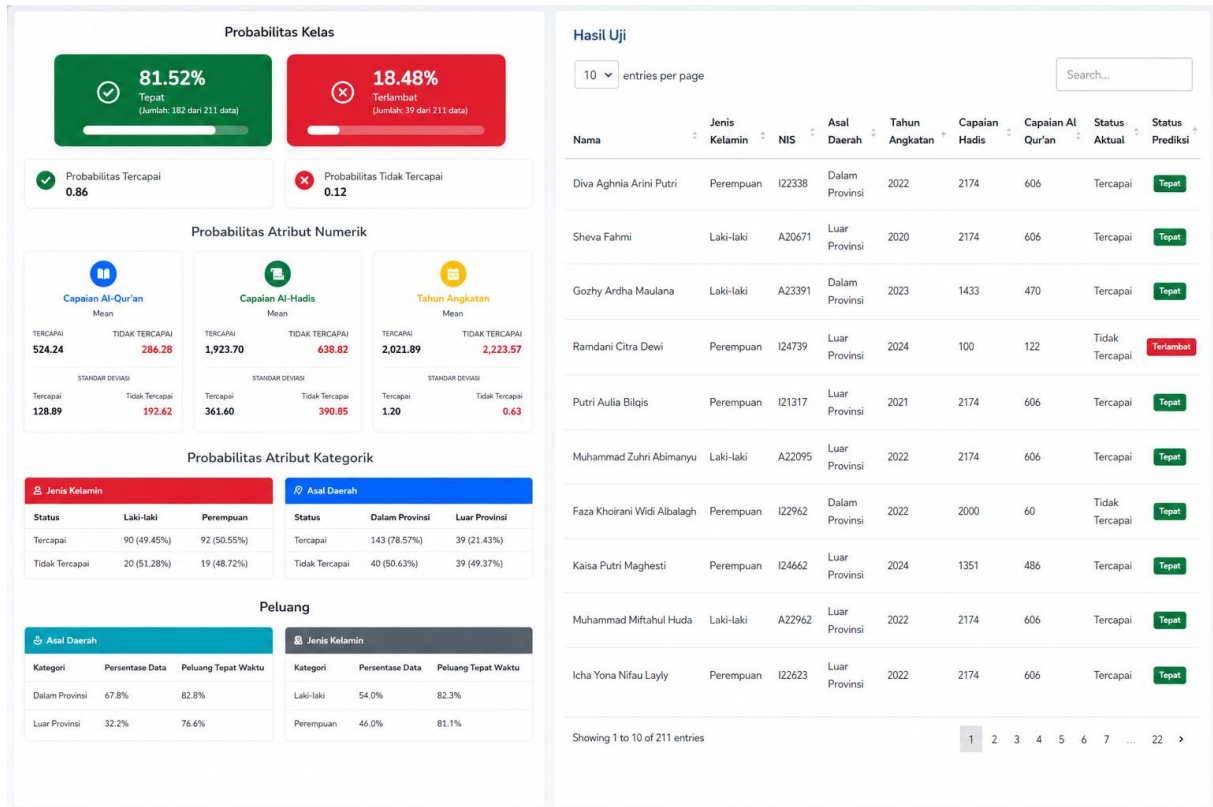
Data Latih

Tampilkan 10 data Cari:

Nama	Jenis Kelamin	NIS	Asal Daerah	Tahun Angkatan	Capaian Al-Hadis	Capaian Al-Qur'an	Target	Aksi
Achmad Bayu Dwi Y. A	Laki-laki	A21915	Dalam Provinsi	2021	Dalam Khasam	Khasam	Tidak Tercapai	Delete
Achmad Ivan Khaegar	Laki-laki	A24230	Dalam Provinsi	2021	Khasam	Khasam	Tercapai	Delete
Adam Rizkia Imawan	Laki-laki	A22246	Dalam Provinsi	2022	Khasam	Khasam	Tercapai	Delete

Gambar 5. Halaman Data Latih

Pada halaman Data Uji (Gambar 5) menampilkan hasil klasifikasi data uji menggunakan algoritma Naïve Bayes, meliputi data santri, status aktual, dan status prediksi. Selain itu, halaman ini menyajikan hasil perhitungan probabilitas kelas beserta jumlah dan persentase data, serta probabilitas atribut kategorikal (jenis kelamin dan asal daerah) dan numerik (tahun angkatan, capaian Al-Qur'an, dan capaian Al-Hadis). Informasi tersebut membantu pengguna dalam menganalisis hasil klasifikasi dan mengevaluasi performa model. Tampilan halaman ditunjukkan pada Gambar 6.



Gambar 6. Halaman Data Uji bagian Probabilitas

3.2. Pengujian Sistem

Pengujian User Acceptance Testing (UAT) dilakukan untuk mengevaluasi tingkat penerimaan pengguna terhadap aplikasi menggunakan kuesioner berbasis skala Likert [28]. Aspek yang diuji meliputi usability, efficiency, memorability, dan satisfaction. Responden terdiri atas 15 santri sebagai pengguna utama dan 1 dewan guru sebagai admin. Pengujian dilakukan dengan meminta responden menggunakan seluruh fitur sesuai perannya, kemudian mengisi kuesioner evaluasi.

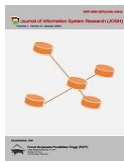
Tabel 11. Rekapitulasi Hasil User Acceptance Testing (UAT)

Responden	Jumlah Responden	Jumlah Pertanyaan	Skor Maksimal	Skor Diperoleh	Persentase	Kategori
Santri	15	16	1200	1148	95,66%	Sangat Baik
Admin (Dewan Guru)	1	15	75	63	84,00%	Sangat Baik

Tabel 11 merupakan hasil pengujian, dimana kelompok santri memperoleh tingkat penerimaan sebesar 95,66%, sedangkan admin (dewan guru) memperoleh 84,00%. Kedua nilai tersebut termasuk dalam kategori sangat baik, sehingga dapat disimpulkan bahwa aplikasi telah memenuhi kebutuhan pengguna baik dari sisi kemudahan penggunaan, efisiensi, kemudahan diingat, maupun kepuasan pengguna. Selain itu, aplikasi juga telah diuji secara langsung melalui implementasi di PPM dan dapat digunakan sesuai kebutuhan operasional.

3.3. Pembahasan

Hasil penelitian menunjukkan bahwa algoritma Naïve Bayes mampu memberikan performa klasifikasi yang baik pada klasifikasi ketepatan kelulusan santri. Temuan ini sejalan dengan penelitian Hozairi [12] dan Setiyani [11] yang menunjukkan bahwa Naïve Bayes efektif digunakan untuk klasifikasi kelulusan. Namun, penelitian tersebut berfokus pada konteks mahasiswa dan belum mempertimbangkan karakteristik pendidikan berbasis pesantren maupun implementasi model ke dalam sistem yang mendukung proses monitoring.



Berbeda dengan penelitian Sriani [14], Junaidi [15], dan Kuncoro [16] yang hanya memanfaatkan atribut akademik atau kepesantrenan secara terpisah, penelitian ini mengintegrasikan atribut akademik dengan capaian Al-Qur'an dan Al-Hadis ke dalam model klasifikasi berbasis Naïve Bayes yang diimplementasikan pada sistem berbasis web. Hasil penelitian menunjukkan bahwa integrasi atribut tersebut mampu menghasilkan performa klasifikasi yang baik sekaligus menyediakan media monitoring dan evaluasi bagi dewan guru. Temuan ini memperluas penerapan Naïve Bayes pada konteks Pondok Pesantren Mahasiswa sebagai pendukung pengambilan keputusan terkait ketepatan kelulusan santri.

4. KESIMPULAN

Penelitian ini berhasil mengembangkan sistem berbasis website yang mengimplementasikan algoritma Naïve Bayes untuk mengklasifikasikan kelulusan santri di Pondok Pesantren Mahasiswa (PPM). Evaluasi model menunjukkan performa yang baik dengan accuracy sebesar 86,79%, precision sebesar 93,41%, dan recall sebesar 91,40%. Selain itu, sistem memperoleh tingkat penerimaan pengguna (User Acceptance Testing) yang tinggi, yaitu 95,66% dari santri dan 84% dari dewan guru, sehingga terbukti tidak hanya memiliki kinerja klasifikasi yang baik, tetapi juga mudah digunakan dan bermanfaat dalam mendukung proses monitoring serta evaluasi capaian santri. Keterbaruan penelitian ini terletak pada integrasi atribut akademik dan kepesantrenan, khususnya capaian tafsir Al-Qur'an dan Al-Hadis, ke dalam model klasifikasi berbasis Naïve Bayes pada konteks Pondok Pesantren Mahasiswa, yang selanjutnya diimplementasikan dalam sistem berbasis website sebagai pendukung monitoring, evaluasi, dan pengambilan keputusan terkait ketepatan kelulusan santri. Kontribusi penelitian ini adalah menyediakan sistem prediksi yang mampu membantu dewan guru dalam mengidentifikasi secara dini santri yang berpotensi mengalami keterlambatan kelulusan, sehingga proses pembinaan, pendampingan, dan evaluasi dapat dilakukan secara lebih terarah, objektif, dan berbasis data. Selain itu, penelitian ini memperluas penerapan algoritma Naïve Bayes pada lingkungan pendidikan berbasis pesantren dengan memanfaatkan atribut kepesantrenan sebagai faktor prediksi kelulusan. Meskipun demikian, komposisi dataset yang digunakan masih belum seimbang, yaitu 186 data "Tercapai" dan 25 data "Tidak Tercapai", sehingga berpotensi memengaruhi kemampuan model dalam mengenali kelas minoritas. Oleh karena itu, penelitian selanjutnya disarankan menerapkan teknik penyeimbangan data, seperti SMOTE atau cost-sensitive learning, melaporkan metrik evaluasi untuk setiap kelas, membandingkan performa Naïve Bayes dengan algoritma lain seperti Random Forest dan Support Vector Machine (SVM), menambahkan atribut yang telah diuji signifikansinya, seperti kehadiran dan nilai sikap, serta melakukan validasi pada data dari angkatan yang berbeda agar sistem menjadi lebih akurat, representatif, dan andal secara berkelanjutan.

REFERENCES

- [1] L. Sinta, Y. M. L. Malaikosa, and D. H. Supriyanto, "Implementasi Penguatan Pendidikan Karakter pada Siswa Kelas Rendah di Sekolah Dasar," *Jurnal Obsesi : Jurnal Pendidikan Anak Usia Dini*, vol. 6, no. 4, pp. 3193–3202, Feb. 2022, doi: 10.31004/obsesi.v6i4.2326.
- [2] O. Ramadhani, A. Marsanda, P. D. Damayanti, A. Suriyansyah, and C. Cinantya, "Pentingnya Pendidikan Karakter di Sekolah Dasar untuk Membangun Generasi Berkualitas," *MARAS : Jurnal Penelitian Multidisiplin*, vol. 3, no. 1, pp. 151–160, Jan. 2025, doi: 10.60126/maras.v3i1.659.
- [3] A. Rahmulyana et al., "Peran Lembaga Pendidikan Dalam Menghasilkan Lulusan Berkualitas Dan Berdaya Saing Tinggi," *Bhinneka Multidisiplin Journal*, vol. 1, no. 2, pp. 55–64, Apr. 2024, doi: 10.53067/bmj.v1i2.7.
- [4] T. Hidayat, A. S. Rizal, F. Fahrudin, and I. Istianah, "Islamic Education Program Approach to Islamic Personality Development," *Al Ulya: Jurnal Pendidikan Islam*, vol. 9, no. 2, pp. 224–244, Oct. 2024, doi: 10.32665/alulya.v9i2.3404.
- [5] Muhammad Alif Al Raihan, Anisa Ananda, Muthmainah, Muhammad Reza Syahbudi, and Arga Sabda Wiguna, "Adaptation of Curriculum and Technology in Responding to the Flow of Globalization: Case Study of PPM Al-Azhar Muncar Banyuwangi," *Journal of Research in Islamic Education*, vol. 7, no. 1, pp. 256–269, Jun. 2025, doi: 10.25217/jrie.v7i1.6308.
- [6] N. M. Jalal, "Resiliensi Akademik Siswa Pondok Pesantren," *Indonesian Journal of Social Science and Education (IJOSSE)*, vol. 1, no. 3, pp. 1107–1116, Dec. 2025, doi: 10.62567/ijosse.v1i3.1643.
- [7] E. Norman, F. D. Rahmawati, A. Paramansyah, and T. Alawiyah, "Integrated Islamic Time Management Model for Tahfidz Students based on Surah Al-'Ashr as a Productivity Framework," *Managere: Indonesian Journal of Educational Management*, vol. 6, no. 3, pp. 407–419, Dec. 2024, doi: 10.52627/managere.v6i3.750.
- [8] I. Riadi, R. Umar, and R. Anggara, "Prediksi Kelulusan Tepat Waktu Berdasarkan Riwayat Akademik Menggunakan Metode K-Nearest Neighbor," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 2, pp. 249–256, Apr. 2024, doi: 10.25126/jtiik.20241127330.
- [9] A. Wahyudi, "Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Decision Tree Dan Naïve Bayes," *Jurnal Permata Indonesia*, vol. 14, no. 2, pp. 132–138, Nov. 2023, doi: 10.59737/jpi.v14i2.276.
- [10] I. N. Rudy Hendrawan, I. M. A. Budhi Saputra, G. A. P. Cahya Dewi, I. G. S. Adi Pranata, and N. L. N. Wedasari, "Klasifikasi Lama Studi dan Predikat Kelulusan Mahasiswa menggunakan Metode Naïve Bayes," *Jurnal Eksplorasi Informatika*, vol. 11, no. 1, pp. 50–56, Jan. 2022, doi: 10.30864/eksplorasi.v11i1.606.
- [11] L. Setiyani, M. Wahidin, D. Awaludin, and S. Purwani, "Analisis Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Data Mining Naïve Bayes : Systematic Review," *Faktor Exacta*, vol. 13, no. 1, p. 35, Jun. 2020, doi: 10.30998/faktorexacta.v13i1.5548.



- [12] H. Hozairi, A. Anwari, and S. Alim, "Implementasi Orange Data Mining Untuk Klasifikasi Kelulusan Mahasiswa Dengan Model K-Nearest Neighbor, Decision Tree Serta Naive Bayes," *Network Engineering Research Operation*, vol. 6, no. 2, p. 133, Nov. 2021, doi: 10.21107/nero.v6i2.237.
- [13] Imam Riadi, Rusydi Umar, and Rio Anggara, "Prediksi Kelulusan Tepat Waktu Berdasarkan Riwayat Akademik Menggunakan Metode Naive Bayes," *Decode: Jurnal Pendidikan Teknologi Informasi*, vol. 4, no. 1, pp. 191–203, Jan. 2024, doi: 10.51454/decode.v4i1.308.
- [14] Sriani and J. Juraidah, "Sistem Prediksi Kelulusan Santri Tahfidz Qur'an Menggunakan Algoritma C4.5," *The Indonesian Journal of Computer Science*, vol. 12, no. 4, Aug. 2023, doi: 10.33022/ijcs.v12i4.3325.
- [15] Satrio Junaidi, R. Valicia Anggela, and D. Kariman, "Klasifikasi Metode Data Mining untuk Prediksi Kelulusan Tepat Waktu Mahasiswa dengan Algoritma Naive Bayes, Random Forest, Support Vector Machine (SVM) dan Artificial Neural Network (ANN)," *Journal of Applied Computer Science and Technology*, vol. 5, no. 1, pp. 109–119, Jun. 2024, doi: 10.52158/jacost.v5i1.489.
- [16] A. Kuncoro, "Integrasi AdaBoost dengan Naive Bayes untuk Klasifikasi Kualitas Hafalan Tematik Santri PPICeLKISI," *Journal of Information System and Application Development*, vol. 3, no. 2, pp. 96–102, Sep. 2025, doi: 10.26905/jisad.v3i2.16138.
- [17] M. Affandi and A. S. Mubarak, "Resiliensi Mahasiswa Santri Tahfidz ditinjau dari Kecerdasan Spiritual dan Religiusitas," *Jurnal Psikologi Islam*, vol. 9, no. 2, pp. 2549–9297, 2022, doi: 10.47399/jpi.v9i2.204.
- [18] S. Syafrizal, R. Trisoni, and I. Satriadi, "Dinamika Pembelajaran Tahfidz Al-Qur'an dan Hadis: Dari Proses Implementasi hingga Capaian Santri," *Instructional Development Journal*, no. vol. 9, 1, Apr. 2026, Accessed: Jul. 17, 2026. Available: <https://ejournal.uin-suska.ac.id/index.php/IDJ/article/view/39538>
- [19] S. Tripathi, D. Muhr, M. Brunner, H. Jodlbauer, M. Dehmer, and F. Emmert-Streib, "Ensuring the Robustness and Reliability of Data-Driven Knowledge Discovery Models in Production and Manufacturing," *Front. Artif. Intell.*, vol. 4, Jun. 2021, doi: 10.3389/frai.2021.576892.
- [20] I. P. A. V. Vidyantanta and K. T. D. Dermawan, "Perbandingan Algoritma K-Means Dan K-Medoids Dalam Pengelompokan Provinsi Di Indonesia Berdasarkan Indikator Keadaan Sekolah Dasar," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 3S1, Oct. 2025, doi: 10.23960/jitet.v13i3S1.8145.
- [21] A. M. Shimaoka, R. C. Ferreira, and A. Goldman, "The evolution of CRISP-DM for Data Science: Methods, Processes and Frameworks," *SBC Reviews on Computer Science*, vol. 4, no. 1, pp. 28–43, Oct. 2024, doi: 10.5753/reviews.2024.3757.
- [22] R. H. Bemthuis, R. R. Govers, and A. Asadi, "A CRISP-DM-based methodology for assessing agent-based simulation models using process mining," *Journal of Simulation*, pp. 1–22, Jun. 2025, doi: 10.1080/17477778.2025.2508245.
- [23] A. M. Shimaoka, R. C. Ferreira, and A. Goldman, "The evolution of CRISP-DM for Data Science: Methods, Processes and Frameworks," *SBC Reviews on Computer Science*, vol. 4, no. 1, pp. 28–43, Oct. 2024, doi: 10.5753/reviews.2024.3757.
- [24] I. G. I. Suardika, "Prediksi Tingkat Kelulusan Mahasiswa Tepat Waktu Menggunakan Naive Bayes: Studi Kasus Fakultas Ekonomi Dan Bisnis Universitas Pendidikan Nasional," *Jurnal Ilmu Komputer Indonesia*, vol. 4, no. 2, pp. 37–44, Sep. 2019, doi: 10.23887/jik.v4i2.2775.
- [25] T. H. Hasibuan and D. Mahdiana, "Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Algoritma C4.5 Pada Uin Syarif Hidayatullah Jakarta," *SKANIKA*, vol. 6, no. 1, pp. 61–74, Jan. 2023, doi: 10.36080/skanika.v6i1.2976.
- [26] B. Ravinder, S. K. Seenii, V. S. Prabhu, P. Asha, S. P. Maniraj, and C. Srinivasan, "Web Data Mining with Organized Contents Using Naive Bayes Algorithm," in *2024 2nd International Conference on Computer, Communication and Control (IC4)*, IEEE, Feb. 2024, pp. 1–6. doi: 10.1109/IC457434.2024.10486403.
- [27] P. S. D. Puspitasari, B. Etikasari, T. D. Puspitasari, R. C. Kartika, L. Perdanasari, and A. A. Kurniasari, "Android-Based Application for Children's Growth Monitoring as a Complement for Child Development Card," *Jurnal Teknokes*, vol. 15, no. 1, pp. 44–50, Mar. 2022, doi: 10.35882/teknokes.v15i1.7.