

Penerapan Algoritma Naive Bayes untuk Klasifikasi Prediksi Penerimaan Siswa Baru SDN Inpres 1 Arso III

Yanik Utari^{1*}, Sartika Situmorang², Linarti³, Yanti Magrice Jewun⁴,
Heru Sutejo⁵

^{1,2,3,4,5} Universitas Sepuluh Nopember, Indonesia

* Correspondence e-mail; sartikasitumorang@gmail.com

Article history

Submitted: 2024/09/18; Revised: 2024/09/29; Accepted: 2024/10/11:

Abstract

Penerimaan peserta didik baru (PPDB) merupakan proses penting dalam sistem pendidikan yang mempengaruhi kualitas pendidikan di setiap sekolah. Di SDN Inpres 1 Arso III, proses seleksi penerimaan siswa baru masih dilakukan secara manual, yang berpotensi menyebabkan ketidakakuratan data dan memperlambat proses seleksi. Untuk mengatasi masalah ini, penelitian ini mengusulkan penerapan algoritma Naive Bayes untuk memprediksi penerimaan siswa baru secara lebih efisien dan akurat. Algoritma Naive Bayes, yang berbasis pada teori probabilitas, dapat digunakan untuk melakukan klasifikasi data dengan memperhitungkan kemungkinan suatu siswa memenuhi kriteria penerimaan berdasarkan variabel seperti usia, nilai ujian, dan faktor sosial ekonomi. Penelitian ini bertujuan untuk meningkatkan efisiensi, akurasi, dan objektivitas dalam proses seleksi PPDB di SDN Inpres 1 Arso III. Hasil penelitian diharapkan dapat memberikan solusi yang mengurangi kesalahan manusia dalam pengelolaan data dan membantu sekolah dalam membuat keputusan yang lebih tepat dalam menerima siswa baru.

Keywords

Naive Bayes; Klasifikasi; Prediksi; Siswa Baru



© 2024 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution 4.0 International (CC BY SA) license, <https://creativecommons.org/licenses/by-sa/4.0/>.

PENDAHULUAN

Penerimaan peserta didik baru merupakan proses penting yang harus dilaksanakan oleh sistem pendidikan setiap tahunnya. Proses seleksi ini melibatkan berbagai faktor yang dapat mempengaruhi keputusan penerimaan siswa, seperti usia, nilai ujian sebelumnya, lokasi tempat tinggal, dan beberapa faktor sosial ekonomi. Di SDN Inpres 1 Arso III, penerimaan siswa baru masih dilakukan secara manual sehingga dapat menyebabkan ketidakakuratan dalam pengelolaan data dan memperlambat proses seleksi.

Untuk meningkatkan efisiensi dan akurasi dalam proses penerimaan mahasiswa baru, diperlukan suatu sistem yang dapat memprediksi dengan lebih cepat dan akurat siapa saja yang memenuhi kriteria penerimaan siswa baru. Salah satu pendekatan

yang dapat diterapkan pada situasi ini adalah dengan menggunakan algoritma Naive Bayes. Algoritma ini merupakan salah satu metode dalam machine learning yang banyak digunakan dalam klasifikasi data khususnya masalah prediksi data yang memiliki banyak variabel atau fitur.

Algoritma Naive Bayes bekerja berdasarkan teori probabilitas, di mana prediksi klasifikasi dilakukan dengan menghitung kemungkinan suatu data termasuk dalam kategori tertentu, berdasarkan informasi yang tersedia. Meskipun disebut "naive", karena mengasumsikan bahwa setiap fitur bersifat independen satu sama lain, algoritma ini terbukti sangat efektif dalam berbagai masalah klasifikasi, termasuk pada prediksi penerimaan siswa baru.

Pada penelitian ini, penerapan Algoritma Naive Bayes diharapkan dapat memberikan solusi yang lebih efisien dalam memprediksi penerimaan siswa baru di SDN Inpres 1 Arso III. Melalui penerapan algoritma ini, diharapkan proses seleksi penerimaan siswa baru menjadi lebih objektif, cepat, dan mengurangi potensi kesalahan manusia dalam pengelolaan data. Selain itu, dengan penggunaan algoritma ini, sekolah dapat mengidentifikasi pola-pola penting yang mungkin tidak terlihat secara manual dalam data yang ada.

Naive Bayes adalah algoritma yang diusulkan pada penelitian ini, karena merupakan salah satu metode machine learning yang menggunakan perhitungan probabilitas (T.H. Apandi and C.A. Sugianto, 2018), (D.Y Utami,dkk, 2020), (A.R.Damanik, 2022). Algoritma Naive Bayes banyak digunakan untuk menyelesaikan masalah prediksi berupa klasifikasi (T.H. Apandi and C.A. Sugianto, 2019), (M.F. Rifai,dkk, 2019), (I.W. Saputro and B. W. Sari, 2020). Selain itu Naive Bayes memiliki kelebihan yaitu sederhana, cepat dan memiliki akurasi yang tinggi (E.Indrayuni,2019). Proses klasifikasi pada naive bayes biasanya dibagi menjadi dua fase yaitu *learning/training* dan *testing/classify*. Pada fase *learning*, sebagian data yang telah diketahui kelas data nya diumpankan untuk membentuk model perkiraan. Kemudian pada fase *testing* model yang sudah terbentuk diuji dengan sebagian data lainnya untuk mengetahui akurasi dari model tersebut (R. Bahtiar,dkk,2020).

Penelitian terkait yang telah dilakukan antara lain: Penelitian yang dilakukan untuk memprediksi potensi pendaftaran siswa pada SMK Taman Siswa Teluk Belitung menggunakan model Naive Bayes. Pada penelitian ini sistem yang dibangun berbasis web berupa sistem informasi prediksi potensi pendaftaran siswa dengan menggunakan model naive bayes yang memiliki kemampuan untuk

memprediksi potensi banyaknya siswa yang mendaftar dan tidak mendaftar di SMK Taman Siswa.

Aplikasi ini diuji menggunakan perhitungan akurasi dimana hasilnya didapatkan tingkat akurasi sebesar 86% untuk prediksi potensi pendaftaran siswa, sehingga dapat disimpulkan bahwa algoritma Naive Bayes cukup baik digunakan untuk melakukan prediksi potensi pendaftaran siswa (S.E.Y. Putri,dkk,2021). Berikutnya penelitian yang dilakukan untuk memprediksi jumlah siswa baru pada SD Negeri Sukosari No. 095126 menggunakan algoritma Naive Bayes dengan *tools Rapi Miner*. Penelitian ini menghasilkan nilai Akurasi sebesar 85.00 %, *class precision* pada prediksi Layak memiliki nilai 92.31%, sedangkan pada prediksi Tidak Layak memiliki nilai 55,56%. *Class recall* pada true Layak memiliki nilai 85,71%, sedangkan pada *true* Tidak Layak memiliki nilai 71,43 %.

Penerapan data mining dalam menentukan klasifikasi penerimaan siswa baru dengan Naive Bayes dapat digunakan untuk memprediksi layak atau tidak layak seorang siswa yang akan diterima (Amik Tunas Bangsa and S. Utara, 2021). Berikutnya penelitian dengan menerapkan algoritma Naive Bayes untuk klasifikasi dalam mendukung keputusan pada sekolah SMP Al Azhar 1 Bandar Lampung dalam melakukan penentuan penerima beasiswa. Sistem pendukung keputusan penerima beasiswa dibangun dengan menggunakan bahasa pemrograman Java dan database MySQL.

Hasil dari penelitian ini, bahwa dengan menerapkan metode naive bayes pada sistem pendukung keputusan penerima beasiswa dapat membantu pihak sekolah dalam melakukan penentuan penerima beasiswa dengan lebih cepat dan tepat (D.Alita,dkk,2021). Berdasarkan permasalahan dan uraian tersebut maka dilakukan penelitian ini.

METODE

2.1. Data Penelitian

Dataset yang digunakan pada penelitian ini dikumpulkan dengan menggunakan metode Observasi, melihat serta mempelajari permasalahan yang ada di lapangan yang berkaitan dengan objek yang diteliti, yaitu informasi mengenai Data siswa SDN Inpres 1 Arso III angkatan 2024 yang berjumlah 162 dengan menggunakan 4 variabel yaitu nilai Matematika, nilai IPA, nilai Bahasa Inggris dan nilai Bahasa Indonesia yang digunakan sebagai subjek dalam penelitian. Serta melaksanakan wawancara kepada pihak SDN Inpres 1 Arso III.

Tabel 1. Dataset Penelitian

No.	Nama Siswa	Nilai			
		MTK	IPA	B.Ingggris	B.Indonesia
1	Agustinus Meo Seran	70	60	70	60
2	Agustinus Nong Jon	75	65	80	65
3	Agustinus Reja S. Seran	70	60	75	60
4	Agustinus Weya	75	60	80	60
5	Ahmad Faris	70	60	80	60
6	Ahmad Faris Maulana	65	65	80	65
7	Aiyra Rafanda	70	70	80	70
8	Alessandra M. Rumwaropen	70	60	80	60
9	Alfredo Wonda	70	75	80	75
10	Aril	70	70	80	70
11	Arzenio Raja Muntazar	70	80	85	80
12	Aszril Thio Firmansyah	80	70	80	70
13	Aurelia Maulida	70	70	70	70
14	Belvania Aeirylin Octavilia	80	60	80	60
15	Chagiah Aldzamar A.	85	80	80	80
16	Dea Hayuningtyas Putri S. S.	80	80	60	80
17	Derry Valdo Wonda	70	75	85	75
18	Dheyra Afygha Q. F. Korwa	80	60	75	60
19	Dikison Wonda	80	65	70	65
20	Dorthea Towoli	85	65	85	65
21	Edison Sesewano	80	65	80	65
22	Ercina Towoli	70	70	80	70
23	Ferdinand M. Bukorpioper	80	70	80	70
24	Gamaliel A. Jhony Sihombing	70	70	85	70

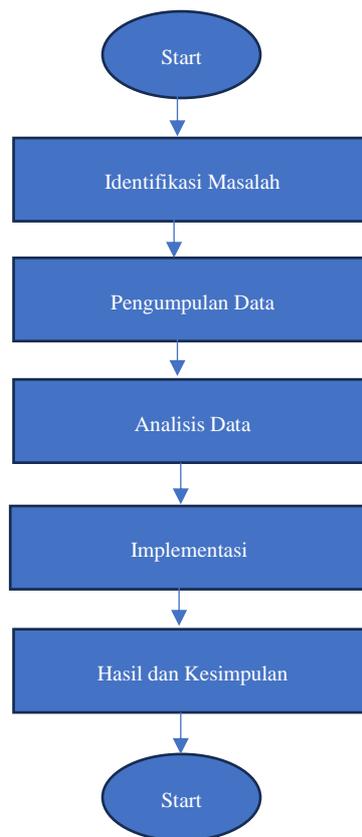
No.	Nama Siswa	Nilai			
		MTK	IPA	B.Ingggris	B.Indonesia
25	Gratia Lien Sihombing	80	60	80	60
26	Ifan Towoli	70	70	80	70
27	Ikbali Salim	75	70	75	70

Tamilis Synex: Multidimensional Collaboration

28	Immanuel Mindi Wonda	80	70	80	70
...
161	Zahra Nur Layla	80	65	80	65
162	Zhaky Ramadhan Shanny	80	70	80	70

2.2. Tahapan Penelitian

Tahapan Penelitian digunakan untuk menyelesaikan permasalahan dalam penelitian. Adapun tahapan-tahapan dalam penelitian ini dapat dilihat pada gambar 1.



Gambar 1. Tahapan Penelitian

a. Identifikasi Masalah

Hal ini bertujuan untuk memahami permasalahan dengan jelas serta dapat mengatasi masalah tersebut dengan baik. Dengan demikian dapat ditentukan rencana kerja serta hal apa saja yang dibutuhkan dalam penelitian ini

b. Pengumpulan Data

Dataset yang digunakan pada penelitian ini dikumpulkan dengan menggunakan metode Observasi, melihat serta mempelajari permasalahan yang ada di lapangan yang berkaitan dengan objek yang diteliti, yaitu informasi mengenai Data siswa SDN Inpres 1 Arso III

c. Analisis Data

Tahap ini berupa tahap seleksi data yang bertujuan untuk mendapatkan data yang bersih dan siap untuk digunakan dalam penelitian. Melakukan seleksi dengan memperhatikan konsistensi data, missing value dan redundant data. Kemudian data diolah menggunakan *Rapid Miner* menggunakan Performance yang berfungsi sebagai validasi dan reabilitas data untuk mencari keakuratan data. Berikut ini sampel data yang digunakan dalam penelitian

d. Implementasi

Selanjutnya proses tahapan implementasi dengan menggunakan *Tools Rapid Miner*

e. Hasil dan Kesimpulan

Setelah implementasi menggunakan *tools Rapid Miner* dilakukan, maka selanjutnya adalah melihat hasil dari penggunaan rapid miner, dilanjutkan dengan membuat kesimpulan berdasarkan hasil yang telah diperoleh.

HASIL DAN PEMBAHASAN

Dalam menggunakan Algoritma Naive bayes hal yang perlu dilakukan yaitu menentukan *input training* serta target *testing* yang ingin dihasilkan. Dalam penelitian ini data yang digunakan adalah data penerimaan siswa dengan total 110. Kemudian data dibagi menjadi 2 data set yaitu 80 data training dan 30 sebagai data testing. Data tersebut akan digunakan untuk melakukan proses perhitungan data probabilitas baik itu training dan testing serta proses menghitung probabilitas akhir. Setelah mendapatkan probabilitas akhir proses terakhir untuk menentukan prediksi penerimaan siswa baru yaitu membandingkan antara diterima dan tidak diterima. Data penelitian yang digunakan terlebih dahulu akan dikonversi menjadi huruf agar lebih mudah dalam proses perhitungan di *tools Rapid miner* yang dapat di lihat pada tabel 2.

Tabel 2. Dataset Penelitian

No	Nilai	Konversi
1	80-100	Sangat Baik
2	70-79	Baik
3	60-69	Cukup

Tamilis Synex: Multidimensional Collaboration

4	55-59	Kurang
5	50-54	Sangat Kurang

Tujuan pembagian data ini untuk menentukan data yang akan digunakan sebagai data *training* dan data *testing* dalam proses menentukan hasil menggunakan Algoritma Naive Bayes, dapat dilihat pada tabel 3 dan 4.

Tabel 3. Data Training

No.	Nama	Kriteria				Hasil
		A1	A2	A3	A4	
1	Siswa 1	B	B	SB	B	Diterima
2	Siswa 2	SB	B	SB	B	Diterima
3	Siswa 3	SB	SB	SB	SB	Diterima
4	Siswa 4	SB	SB	SB	B	Diterima
5	Siswa 5	SB	B	SB	B	Diterima
6	Siswa 6	SB	K	B	SB	Diterima
7	Siswa 7	C	C	C	K	Tidak
8	Siswa 8	K	K	K	K	Tidak
9	Siswa 9	B	C	C	K	Diterima
10	Siswa 10	B	B	C	K	Tidak
11	Siswa 11	B	C	B	B	Tidak
12	Siswa 12	B	B	SB	B	Diterima
13	Siswa 13	C	C	C	C	Diterima
14	Siswa 14	SB	SB	SB	B	Diterima
15	Siswa 15	B	B	B	B	Diterima
16	Siswa 16	B	C	C	K	Tidak
17	Siswa 17	B	B	B	B	Diterima
18	Siswa 18	SB	B	SB	B	Diterima
19	Siswa 19	C	C	K	K	Diterima
..
80	Siswa 80	B	SB	B	B	Diterima

Keterangan :

SB = Sangat Baik, B = Baik, C = Cukup, K = Kurang, SK = Sangat Kurang

Tabel 4. Data Testing

No.	Nama	Kriteria				Hasil
		A1	A2	A3	A4	

1	Siswa 1	B	C	C	C	?
2	Siswa 2	B	B	B	B	?
3	Siswa 3	B	B	B	B	?
4	Siswa 4	SB	B	B	C	?
5	Siswa 5	B	B	B	B	?
6	Siswa 6	B	B	B	B	?
7	Siswa 7	B	B	B	C	?
8	Siswa 8	B	B	B	B	?
9	Siswa 9	SB	B	B	B	?
10	Siswa 10	SB	B	B	B	?
11	Siswa 11	B	B	SB	B	?
12	Siswa 12	B	C	B	B	?
13	Siswa 13	B	B	B	B	?
14	Siswa 14	B	B	B	B	?
15	Siswa 15	B	B	B	B	?
16	Siswa 16	SB	B	B	C	?
17	Siswa 17	B	C	C	K	?
18	Siswa 18	B	C	B	K	?
19	Siswa 19	B	C	C	SK	?
20	Siswa 20	B	B	B	C	?
21	Siswa 21	B	B	B	C	?
22	Siswa 22	SB	B	K	K	?
23	Siswa 23	B	B	B	C	?
24	Siswa 24	B	B	B	C	?
25	Siswa 25	B	B	B	K	?
26	Siswa 26	C	C	B	B	?
27	Siswa 27	B	B	B	B	?
28	Siswa 28	B	C	C	B	?
29	Siswa 29	B	B	B	B	?
30	Siswa 30	B	C	C	B	?

Kemudian data diolah dengan menggunakan Microsoft Excel lalu diproses secara perhitungan manual dengan Algoritma Naive Bayes dan proses pengujian dilakukan dengan menggunakan *Tools Rapid Miner 5.3*.

3.1 Penerapan Algoritma naive Bayes

Setelah data ditentukan, langkah awal adalah melakukan perhitungan diterima dan tidak. Dari 80 data set yang digunakan, diketahui diterima sebanyak 62 data, dan tidak sebanyak 18 data. Adapun tahapan proses perhitungan menggunakan algoritma Naive Bayes untuk prediksi penerimaan siswa sebagai berikut :

3.1.1 Baca Data Training

Seperti yang tertera pada tabel 3 dataset yang digunakan bukanlah data numerik, maka proses selanjutnya menghitung jumlah dan probabilitas dari setiap kategori.

3.1.2 Hitung Jumlah Probabilitas

a. Perhitungan probabilitas prior ($P(C_i)$)

Berdasarkan dari data set jumlah data yang digunakan sebanyak 80 data dimana terdapat 62 data yang diterima dan 18 data tidak diterima. Proses perhitungan probabilitas adalah sebagai berikut :

$$P(C_{Diterima}) = 80/62 = 0,775$$

$$P(C_{Tidak}) = 80/18 = 0,225$$

b. Selanjutnya menghitung probabilitas pada masing-masing kategori

1. Parameter Nilai Matematika (A1)

$$P(\text{Sangat Baik} |_{Diterima}) = 17/82 = 0.2742$$

$$P(\text{Baik} |_{Tidak}) = 1/18 = 0,0556$$

$$P(\text{Baik} |_{Puas}) = 37/62 = 0,5968$$

$$P(\text{Baik} |_{Tidak}) = 8/18 = 0,4444$$

$$P(\text{Cukup} |_{Diterima}) = 8/62 = 0,1290$$

$$P(\text{Cukup} |_{Tidak}) = 4/18 = 0,2222$$

$$P(\text{Kurang} |_{Diterima}) = 0/62 = 0$$

$$P(\text{Kurang} |_{Tidak}) = 3/18 = 0,1667$$

$$P(\text{Sangat Kurang} |_{Diterima}) = 0/62 = 0$$

$$P(\text{Sangat Kurang} |_{Tidak}) = 2/18 = 0,1111$$

Hasil perhitungan probabilitas pada parameter Nilai Matematika dengan kategori A1 dapat dilihat pada tabel 5 berikut.

Tabel 5. Hasil Perhitungan Probabilitas Kategori A1

Kategori : A1	Diterima	Tidak	P. Diterima	P. Tidak
Sangat Baik	17	1	0,2742	0,0556

Yanik Utari, et al.

Baik	37	8	0,5968	0,4444
Cukup	8	4	0,1290	0,2222
Kurang	0	3	0	0,1667
Sangat Kurang	0	2	0	0,1111
Jumlah	62	18	1	1
Total Data	80			

2. Paramater Nilai IPA (A2)

$$P(\text{Sangat Baik} \mid \text{Diterima}) = 14/62 = 0,2258$$

$$P(\text{Baik} \mid \text{Tidak}) = 1/18 = 0,0556$$

$$P(\text{Baik} \mid \text{Puas}) = 39/62 = 0,6290$$

$$P(\text{Baik} \mid \text{Tidak}) = 7/18 = 0,3889$$

$$P(\text{Cukup} \mid \text{Diterima}) = 8/62 = 0,1290$$

$$P(\text{Cukup} \mid \text{Tidak}) = 5/18 = 0,2778$$

$$P(\text{Kurang} \mid \text{Diterima}) = 1/62 = 0,0161$$

$$P(\text{Kurang} \mid \text{Tidak}) = 3/18 = 0,1667$$

$$P(\text{Sangat Kurang} \mid \text{Diterima}) = 0/62 = 0$$

$$P(\text{Sangat Kurang} \mid \text{Tidak}) = 2/18 = 0,1111$$

Hasil perhitungan probabilitas pada parameter Nilai Matematika dengan kategori A2 dapat dilihat pada tabel 6 berikut.

Tabel 6. Hasil Perhitungan Probabilitas Kategori A2

Kategori : A2	Diterima	Tidak	P. Diterima	P. Tidak
Sangat Baik	14	1	0,2258	0,0556
Baik	39	7	0,6290	0,3889
Cukup	8	5	0,1290	0,2778
Kurang	1	3	0,0161	0,1667
Sangat Kurang	0	2	0,0000	0,1111
Jumlah	62	18	1	1
Total Data	80			

3. Paramater Nilai Bahasa Inggris (A3)

$$P(\text{Sangat Baik} \mid \text{Diterima}) = 17/62 = 0,2742$$

$$P(\text{Baik} \mid \text{Tidak}) = 1/18 = 0,0556$$

$$P(\text{Baik} \mid \text{Puas}) = 28/62 = 0,4516$$

$$P(\text{Baik} \mid \text{Tidak}) = 2/18 = 0,1111$$

Tamilis Synex: Multidimensional Collaboration

$$P(\text{Cukup} | \text{Diterima}) = 18/62 = 0,2581$$

$$P(\text{Cukup} | \text{Tidak}) = 8/18 = 0,4444$$

$$P(\text{Kurang} | \text{Diterima}) = 1/62 = 0,0161$$

$$P(\text{Kurang} | \text{Tidak}) = 5/18 = 0,2778$$

$$P(\text{Sangat Kurang} | \text{Diterima}) = 0/62 = 0$$

$$P(\text{Sangat Kurang} | \text{Tidak}) = 2/18 = 0,1111$$

Hasil perhitungan probabilitas pada parameter Nilai Matematika dengan kategori A1 dapat dilihat pada tabel 7 berikut.

Tabel 7. Hasil Perhitungan Probabilitas Kategori A3

Kategori : A3	Diterima	Tidak	P. Diterima	P. Tidak
Sangat Baik	17	1	0,2742	0,0556
Baik	28	2	0,4516	0,1111
Cukup	16	8	0,2581	0,4444
Kurang	1	5	0,0161	0,2778
Sangat Kurang	0	2	0,0000	0,1111
Jumlah	62	18	1	1
Total Data	80			

4. Paramater Nilai Bahasa Indonesia (A4)

$$P(\text{Sangat Baik} | \text{Diterima}) = 7/62 = 0,1129$$

$$P(\text{Baik} | \text{Tidak}) = 1/18 = 0,0556$$

$$P(\text{Baik} | \text{Puas}) = 31/62 = 0,5$$

$$P(\text{Baik} | \text{Tidak}) = 5/18 = 0,2778$$

$$P(\text{Cukup} | \text{Diterima}) = 14/62 = 0,2258$$

$$P(\text{Cukup} | \text{Tidak}) = 0/18 = 0$$

$$P(\text{Kurang} | \text{Diterima}) = 10/62 = 0,1613$$

$$P(\text{Kurang} | \text{Tidak}) = 8/18 = 0,4444$$

$$P(\text{Sangat Kurang} | \text{Diterima}) = 0/62 = 0$$

$$P(\text{Sangat Kurang} | \text{Tidak}) = 4/18 = 0,2222$$

Hasil perhitungan probabilitas pada parameter Nilai Matematika dengan kategori A4 dapat dilihat pada tabel 8 berikut.

Tabel 8. Hasil Perhitungan Probabilitas Kategori A4

Kategori : A3	Diterima	Tidak	P. Diterima	P. Tidak
Sangat Baik	7	1	0,1129	0,0556
Baik	31	5	0,5000	0,2778
Cukup	14	0	0,2258	0,0000
Kurang	10	8	0,1613	0,4444
Sangat Kurang	0	4	0,0000	0,2222
Jumlah	62	18	1	1
Total Data	80			

3.1.3 Baca Data Testing

Data uji yang digunakan dalam penelitian ini terdiri dari 30 data responden. Berikut adalah data uji yang digunakan dapat dilihat pada tabel 9.

Tabel 9. Data Testing

No.	Nama	Kriteria				
		A1	A2	A3	A4	Hasil
1	Siswa 1	B	C	C	C	?
2	Siswa 2	B	B	B	B	?
3	Siswa 3	B	B	B	B	?
4	Siswa 4	SB	B	B	C	?
5	Siswa 5	B	B	B	B	?
6	Siswa 6	B	B	B	B	?
7	Siswa 7	B	B	B	C	?
8	Siswa 8	B	B	B	B	?
9	Siswa 9	SB	B	B	B	?
10	Siswa 10	SB	B	B	B	?
...
27	Siswa 27	B	B	B	B	?
28	Siswa 28	B	C	C	B	?
29	Siswa 29	B	B	B	B	?
30	Siswa 30	B	C	C	B	?

3.1.4 Menghitung Probabilitas Data Testing

Berdasarkan data *training* berikut merupakan perhitungan nilai Diterima dari data 81 sampai 110 dengan jumlah 30 mahasiswa. Sehingga untuk mendapatkan nilai dilakukan proses perhitungan sebagai berikut.

Tamilis Synex: Multidimensional Collaboration

$$\begin{aligned}P(M1 | Puas) &= P(XA1=Baik | Diterima) \times P(XA2=Cukup | Diterima) \times P(XA3=Cukup | \\ & \text{Diterima}) \times P(XA4=Cukup | \text{Diterima}) \\ &= 0,5968 \times 0,1290 \times 0,2581 \times 0,2258 = 0,00448673\end{aligned}$$

$$\begin{aligned}P(M2 | PDiterima) &= P(XA1=Baik | Diterima) \times P(XA2=Baik | Diterima) \times P(XA3=Baik | \\ & \text{Diterima}) \times P(XA4=Baik | \text{Diterima}) \\ &= 0,5968 \times 0,6290 \times 0,4516 \times 0,5 = 0,08476\end{aligned}$$

Lakukan Perhitungan samapai dengan M30.

Sedangkan untuk menghitung nilai Tidak diterima data 80 sampai ke 110. Rumus yang digunakan sama dengan rumus untuk menentukan nilai tidak diterima. Sehingga untuk mendapatkan nilai dilakukan sebagai berikut.

$$\begin{aligned}P(R1 | Tidak Puas) &= P(XA1=Baik | Tidak) \times P(XA2=Cukup | Tidak) \times \\ & P(XA3=Cukup | Tidak) \times P(XA4=Cukup | Tidak) \\ &= 0,4444 \times 0,0000 \times 0,3889 \times 0,2778 = 0\end{aligned}$$

$$\begin{aligned}P(R2 | Tidak Puas) &= P(XA1=Baik | Tidak) \times P(XA2=Baik | Tidak) \times P(XA3=Baik | Tidak) \\ & \times P(XA4=Baik | Tidak) \\ &= 0,4444 \times 0,3889 \times 0,1111 \times 0,2778 = 0,005334\end{aligned}$$

Lakukan Perhitungan sampai dengan M30

3.1.5 Menghitung Probabilitas Akhir

Probabilitas akhir merupakan pemaksimalan klasifikasi Diterima dan tidak:

$$\begin{aligned}P(\text{Diterima} | X) &= P(X | \text{Diterima}) * P(C | \text{Diterima}) \\ &= 0,00448673 \times 0,775 \\ &= 0,00347721575\end{aligned}$$

$$\begin{aligned}P(\text{Tidak} | X) &= P(X | \text{Tidak}) * P(C | \text{Tidak}) \\ &= 0 \times 0,225 \\ &= 0\end{aligned}$$

3.2 Implementasi

Setelah mengetahui jumlah responden prediksi puas dan tidak puas berikutnya adalah proses pembuktian nilai probabilitas di atas akan diuji dengan data sebanyak 30 data dan diselesaikan dengan menggunakan *tools Rapid Miner 5.3* sehingga

Yanik Utari, et al.

diperoleh hasil pengujian probabilitas *confidence* klasifikasi seperti yang terlihat pada gambar 2.

Row No.	PROSES	A1	A2	A3	A4
1	Diterima	Baik	Cukup	Cukup	Cukup
2	Diterima	Baik	Baik	Baik	Baik
3	Diterima	Baik	Baik	Baik	Baik
4	Diterima	Sangat Baik	Baik	Baik	Cukup
5	Tidak	Baik	Baik	Baik	Baik
6	Tidak	Baik	Baik	Baik	Baik
7	Diterima	Baik	Baik	Baik	Cukup
8	Tidak	Baik	Baik	Baik	Baik
9	Diterima	Sangat Baik	Baik	Baik	Baik
10	Diterima	Sangat Baik	Baik	Baik	Baik
11	Diterima	Baik	Baik	Sangat Baik	Baik
12	Tidak	Baik	Cukup	Baik	Baik
13	Diterima	Baik	Baik	Baik	Baik

Gambar 2. Hasil Pengujian Nilai Probabilitas Data Testing dengan *Rapid Miner*

Berdasarkan gambar 2 terdapat empat siswa dengan prediksi tidak diterima dan 26 siswa dinyatakan diterima terhadap proses yang telah diuji menggunakan *tools*. Setelah melihat bukti pengujian hasil nilai probabilitas selanjutnya melihat seberapa besar akurasi yang didapatkan pada proses prediksi yang telah dilakukan. Akurasi Hasil pengujian Model Algoritma Naive Bayes *Classifier* dapat dilihat pada gambar 3.

accuracy: 76.67%				
	true Diterima	true Tidak	true Tidak	class precision
pred. Diterima	23	4	0	85.19%
pred. Tidak	3	0	0	0.00%
pred. Tidak	0	0	0	0.00%
class recall	88.46%	0.00%	0.00%	

Gambar 3. Nilai *AccuracyPerformance*

Keterangan :

- Pada gambar di atas Nilai *Accuracy* sebesar 76,6 %. Dengan Demikian Algoritma Naive bayes dapat diterapkan pada prediksi penerimaan murid baru dengan *Accuracy* mendekati 80 %.
- Data pengujian sebanyak 30 data *testing* yang diolah di dalam *RapidMiner 5.3*. mendapatkan hasil pengujian dengan akurasi sebesar 76,6% yaitu 26 siswa diterima dan sebanyak 4 tidak diterima. Sehingga proses dalam melakukan prediksi lebih cepat dan akurat serta dapat menghasilkan tingkat akurasi dengan nilai yang didapatkan.

3.3 Hasil Akhir

Tamilis Synex: Multidimensional Collaboration

Hasil akhir dari proses prediksi berupa klasifikasi secara keseluruhan berdasarkan data yang telah valid dan diuji menggunakan pemodelan *tools Rapid Miner 5.3* dapat dilihat pada tabel 10.

Tabel 10. Hasil Akhir

Nama Siswa	A1	A2	A3	A4	Hasil
Jecklin I. Wanimbo	Baik	Cukup	Cukup	Cukup	Diterima
Juano Brandley Wonda	Baik	Baik	Baik	Baik	Diterima
Keinar Kaleo Keno	Baik	Baik	Baik	Baik	Diterima
Khalifah Artanti Arsa	Sangat Baik	Baik	Baik	Cukup	Diterima
Khumaira Naoko P.	Baik	Baik	Baik	Baik	Diterima
Marselinus Seran	Baik	Baik	Baik	Baik	Diterima
Martha Maria Erari	Baik	Baik	Baik	Cukup	Diterima
Maulana Ahmad F.	Baik	Baik	Baik	Baik	Diterima
Mikael S. Randongkir	Sangat Baik	Baik	Baik	Baik	Diterima
Mohammad Hafiz A.	Sangat Baik	Baik	Baik	Baik	Diterima
Muhamad Adriyan	Baik	Baik	Sangat Baik	Baik	Diterima
Muhammad Arohim	Baik	Cukup	Baik	Baik	Diterima
Muhammad Fadil S.	Baik	Baik	Baik	Baik	Diterima
Muhammad Rafif	Baik	Baik	Baik	Baik	Diterima
Muhammad Riskyiah	Baik	Baik	Baik	Baik	Diterima
Muhammad Shabri P.	Sangat Baik	Baik	Baik	Cukup	Diterima
Muh. Zatyatullah	Baik	Cukup	Cukup	Kurang	Tidak
Nathania Latifah Putra	Baik	Cukup	Baik	Kurang	Tidak
Nela	Baik	Cukup	Cukup	Sangat Kurang	Tidak
Nona Inseri Ayamiseba	Baik	Baik	Baik	Cukup	Diterima
Novan	Baik	Baik	Baik	Cukup	Diterima
Nur'Ainun Saparuddin	Sangat Baik	Baik	Kurang	Kurang	Tidak
Nur Cantika Aprilia	Baik	Baik	Baik	Cukup	Diterima
Nur Indah Ramadhani	Baik	Baik	Baik	Cukup	Diterima
Nur Zyafara	Baik	Baik	Baik	Kurang	Diterima
Nur Zyafira	Cukup	Cukup	Baik	Baik	Diterima
Ovan	Baik	Baik	Baik	Baik	Diterima
Paskalina Balingga	Baik	Cukup	Cukup	Baik	Diterima
Paulus Kmur	Baik	Baik	Baik	Baik	Diterima
Penende Wonda	Baik	Cukup	Cukup	Baik	Diterima

KESIMPULAN

Berdasarkan pengujian sebanyak 30 data testing yang diolah menggunakan Rapid Miner, diperoleh tingkat akurasi sebesar 86,6% yaitu 26 siswa diterima dan sebanyak 4 tidak diterima. Sehingga dapat disimpulkan bahwa proses prediksi berupa klasifikasi dengan menggunakan Naive Bayes dapat lebih cepat dan akurat serta menghasilkan tingkat akurasi yang tinggi bila diterapkan untuk mengatasi masalah klasifikasi prediksi jumlah siswa baru pada SDN Inpres 1 Arso III.

REFERENSI

- T. H. Apandi and C. A. Sugianto, "Analisis Komparasi Machine Learning Pada Data Spam Sms," *Jurnal TEDC*, vol. 12, no. 1, pp. 58–62, 2018.
- D. Y. Utami, E. Nurlelah, and N. Hikmah, "Attribute Selection in Naive Bayes Algorithm Using Genetic Algorithms and Bagging for Prediction of Liver Disease," *Journal of Informatics and Telecommunication Engineering*, vol. 4, no. 1, pp. 76–85, 2020.
- A. R. Damanik et al., "Prediksi Tingkat Kematian di Indonesia Akibat Covid-19 Menggunakan Algoritma Naive Bayes," *SISTEMASI: Jurnal Sistem Informasi*, vol. 11, no. 1, pp. 73–83, 2022.
- T. H. Apandi and C. A. Sugianto, "Algoritma Naive Bayes untuk Prediksi Kepuasan Pelayanan Perekaman e-KTP (Naive Bayes Algorithm for Satisfaction Prediction of e-ID," *JUITA (Jurnal Informatika) UMP*, vol. 7, no. 2, pp. 125– 128, 2019.
- M. F. Rifai, H. Jatnika, and B. Valentino, "Penerapan Algoritma Naive Bayes Pada Sistem Prediksi Tingkat Kelulusan Peserta Sertifikasi Microsoft Office Specialist (MOS)," *PETIR: Jurnal Pengkajian dan Penerapan Teknik Informatika*, vol. 12, no. 2, pp. 131–144, 2019.
- I. W. Saputro and B. W. Sari, "Uji Performa Algoritma Naive Bayes untuk Prediksi Masa Studi Mahasiswa," *Creative Information Technology Journal*, vol. 6, no. 1, pp. 1–11, 2020.
- E. Indrayuni, "Klasifikasi Text Mining Review Produk Kosmetik Untuk Teks Bahasa Indonesia Menggunakan Algoritma Naive Bayes," *Jurnal Khatulistiwa Informatika*, vol. 7, no. 1, pp. 29–36, 2019.
- R. Bahtiar, M. D. S. T, A. Setiawan, and P. Rosyani, "Analisis Perbandingan Detection Traffic Anomaly Dengan Metode Naive Bayes Dan Support Vector Machine (Svm)," *Jurnal Kreativitas Mahasiswa Informatika*, vol. 1, no. 2, pp.99–103, 2020.
- S. E. Y. Putri, Saniati, and A. Surahman, "Penerapan Model Naive Bayes Untuk Memprediksi Potensi Pendaftaran Siswa di SMK Taman Siswa Teluk Belitung Berbasis Web," *Jurnal JECST*, vol. 1, no. 1, pp. 82–87, 2021.
- Amik tunas bangsa and S. Utara, "Teknik Data Mining Dalam Prediksi Jumlah Siswa Baru Dengan Algoritma Naive Bayes," *Kesatria: Jurnal Penerapan Sistem Informasi (Komputer & Manajemen)*, vol. 2, no. 2, pp. 108–117, 2021.

D. Alita, I. Sari, and A. Rahman Isnain, "Penerapan Naïve Bayes Classifier Untuk Pendukung Keputusan Penerima Beasiswa," *Jurnal Data Mining dan Sistem Informasi (JDMSI)*, vol. 2, no. 1, pp. 17–23, 2021.