

Data Mining Dengan Algoritma Naive Bayes Untuk Klasifikasi Minat Bakat Anak Pada Pusat Pengembangan Anak (PPA) Di Kota Medan

Data Mining Using the Naive Bayes Algorithm for Classifying Children's Interests and Talents at the Child Development Center (PPA) in Medan City

Mikhael Aditty Manalu^{a,1}, Ria Eka Sari^{b,2}

^{a,b}Sistem Informasi, Universitas Potensi Utama, Medan 20241, Sumatera Utara, Indonesia
Email: mikaeladitty21@gmail.com¹, ladiespure@gmail.com²

ABSTRAK

Pusat Pengembangan Anak (PPA) memiliki peran penting dalam membantu anak-anak mengembangkan potensi, minat, dan bakat mereka secara optimal. Namun, proses identifikasi minat dan bakat anak di PPA sering menghadapi tantangan, seperti volume data evaluasi yang besar, subjektivitas dalam penilaian, dan kesulitan menentukan program pengembangan yang sesuai. Untuk mengatasi permasalahan tersebut, penelitian ini mengembangkan sistem berbasis data mining dengan menggunakan algoritma Naive Bayes untuk klasifikasi minat dan bakat anak. Algoritma Naive Bayes dipilih karena kemampuannya dalam menganalisis data dan menghasilkan klasifikasi yang akurat berdasarkan probabilitas. Sistem yang dirancang menggunakan data hasil evaluasi anak, seperti observasi aktivitas, dan penilaian guru, untuk mengelompokkan anak ke dalam kategori minat dan bakat, seperti musik, olahraga, akademik, teknologi dan keterampilan. Hasil implementasi sistem ini menunjukkan bahwa algoritma Naive Bayes mampu mengklasifikasikan minat dan bakat anak dengan tingkat akurasi yang tinggi. Sistem ini tidak hanya membantu mengurangi subjektivitas penilaian, tetapi juga mendukung PPA dalam menentukan program pengembangan yang lebih terarah dan sesuai dengan kebutuhan individu anak. Dengan demikian, sistem ini diharapkan dapat menjadi solusi efektif untuk meningkatkan kualitas layanan di PPA.

Kata Kunci : Data Mining, Klasifikasi, Minat Bakat, Pusat Pengembangan Anak (PPA), Naïve Bayes.

ABSTRACT

Child Development Centers (PPA) have an important role in helping children develop their potential, interests, and talents optimally. However, the process of identifying children's interests and talents at PPA often faces challenges, such as large volumes of evaluation data, subjectivity in assessment, and difficulty in determining appropriate development programs. To overcome these problems, this study developed a data mining-based system using the Naive Bayes algorithm for classifying children's interests and talents. The Naive Bayes algorithm was chosen because of its ability to analyze data and produce accurate classifications based on probability. The designed system uses child evaluation data, such as activity observations, and teacher assessments, to group children into interest and talent categories, such as music, sports, academics, technology and skills. The results of the implementation of this system show that the Naive Bayes algorithm is able to classify children's interests and talents with a high level of accuracy. This system not only helps reduce the subjectivity of assessment, but also supports PPA in determining development programs that are more focused and in accordance with the needs of individual children. Thus, this system is expected to be an effective solution to improve the quality of services at PPA.

Keywords : Data Mining, Classification, Talent Interest, Child Development Center (PPA), Naïve Bayes

Info Artikel :

Disubmit:

Direview:

Diterima :

1. PENDAHULUAN

Kemajuan teknologi informasi pada saat ini telah mampu membantu manusia dalam mengambil Keputusan [1]. Perkembangan teknologi di era milenial sangatlah pesat mampu memberikan efek yang cukup besar, apa lagi dalam dunia bisnis [2]. Informasi adalah salah satu kata kunci salah satu kunci pada zaman ini. Untuk mendapatkan dan menghasilkan informasi, komputer dan teknologinya adalah salah satu alat bantu yang paling tepat [3]. Perkembangan teknologi informasi memunculkan berbagai jenis kegiatan yang berbasis perangkat digital dengan didukung oleh teknologi informasi [4].

Pusat Pengembangan Anak (PPA) merupakan lembaga yang memiliki peran penting dalam memfasilitasi tumbuh kembang anak secara menyeluruh, baik dalam aspek intelektual, emosional, sosial, maupun keterampilan praktis. Seluruh anak yang terdaftar di PPA diarahkan untuk mengikuti berbagai aktivitas dari semua bidang secara bergiliran, tanpa pembatasan awal. Setelah itu, setiap mentor memberikan penilaian berdasarkan atribut perkembangan anak, yang terdiri dari Kognitif, Motorik, Sosial-Emosional, Komunikasi, Kreativitas, dan Kepribadian. Penilaian dilakukan berdasarkan pengamatan langsung terhadap keterlibatan anak dalam setiap kegiatan, kemudian direkap sebagai data evaluasi perkembangan anak.

Dalam implementasinya, proses identifikasi ini masih dilakukan secara manual, yang menimbulkan beberapa kendala. Pertama, jumlah data penilaian dari berbagai bidang dan atribut sangat besar, sehingga mempersulit proses analisis. Kedua, keterbatasan waktu dan tenaga kerja menyebabkan pengolahan data tidak dapat dilakukan secara mendalam dan menyeluruh. Ketiga, hasil penilaian sering kali berbeda-beda antar mentor, karena masih bergantung pada cara pandang masing-masing, sehingga hasil akhirnya menjadi kurang konsisten dan tidak obyektif.

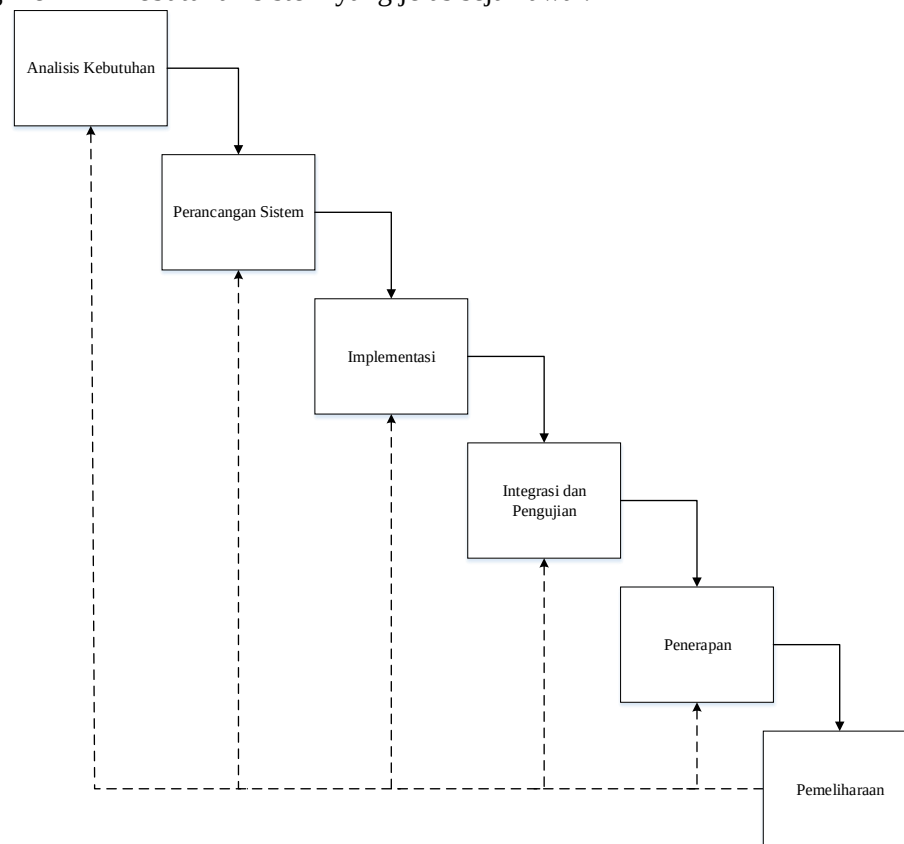
Untuk mengatasi permasalahan tersebut, dibutuhkan sebuah sistem yang mampu mengklasifikasikan minat dan bakat anak secara otomatis dan obyektif berdasarkan data penilaian yang tersedia. Salah satu solusi yang digunakan adalah data mining dengan algoritma Naive Bayes. Data mining adalah proses menganalisis set data yang besar untuk menemukan pola, korelasi, atau anomali yang sebelumnya tidak diketahui dan memiliki makna [5]. Dari data mining adalah mengubah data mentah menjadi informasi yang berguna dan dapat ditindaklanjuti. Ini adalah langkah dalam proses yang lebih besar, yaitu KDD (Knowledge Discovery in Databases), untuk menghasilkan pengetahuan atau *insight* [6]. Data mining menggabungkan berbagai disiplin ilmu, seperti statistik, machine learning (pembelajaran mesin), dan sistem basis data, untuk memprediksi tren atau perilaku di masa depan berdasarkan data historis [7]. Data mining memiliki dua tujuan utama: Deskripsi (memahami dan menjelaskan pola yang ada dalam data) dan Prediksi (menggunakan pola tersebut untuk memperkirakan hasil masa depan) [8].

Naive Bayes adalah algoritma klasifikasi yang menggunakan teorema Bayes sebagai dasarnya. Metode ini disebut "naif" (naïve) karena membuat asumsi kuat bahwa setiap fitur atau prediktor yang digunakan adalah bebas (independen) satu sama lain, meskipun pada kenyataannya sering kali tidak [9]. Algoritme NB ini bekerja dengan menghitung probabilitas suatu data masuk ke dalam suatu kategori berdasarkan fitur-fiturnya [10]. Naive Bayes terkenal karena kecepatan, kemudahan implementasinya, dan sering kali menunjukkan kinerja yang sangat efektif [11]. Penelitian "Penerapan Metode Naive Bayes Dalam Menentukan Telur Ayam Layak Konsumsi Berbasis Web". Peneliti merekomendasikan metode Naive Bayes yang akan digunakan untuk menentukan telur ayam yang layak konsumsi [12]. Penelitian "Data Mining Menentukan Balita Yang Mendapatkan Makanan Tambahan Dari Kemenkes Dengan Metode Naive Bayes (Studi Kasus: Puskesmas Pembantu Bantan)". Di mana dari 55 data balita yang melakukan kunjungan pelayanan posyandu, ada sebanyak 54 data balita yang berhasil diklasifikasikan dengan benar dan sebanyak 1 data balita tidak berhasil diklasifikasikan dengan benar [13]. Penelitian "Klasifikasi Tingkat Kalori Menu Restoran Fast Food Di Indonesia Menggunakan Metode Naive Bayes". Penggunaan algoritma Naive Bayes dapat

menghasilkan tingkat akurasi yang cukup tinggi yaitu 82,35% sehingga dapat disimpulkan bahwa hasil klasifikasi yang dilakukan oleh algoritma tersebut memiliki tingkat kecocokan dengan data asli sebesar 82.35% [14]. Penelitian “Penerapan Metode NAIVE BAYES Untuk Menganalisis Minat Belajar”. Penelitian ini menggunakan metode Naïve Bayes untuk mendapatkan hasil diagnosa minat belajar murid Mis Al-Hasanah [15].

2. METODE

Waterfall adalah salah satu model pengembangan perangkat lunak yang menggunakan pendekatan *sekuensial* atau *linear*. Dalam model ini, setiap tahap pengembangan dilakukan secara berurutan, di mana tahapan berikutnya baru dapat dimulai setelah tahapan sebelumnya selesai sepenuhnya. *Waterfall* sering dianggap sebagai pendekatan tradisional dalam rekayasa perangkat lunak, terutama untuk proyek yang memiliki kebutuhan sistem yang jelas sejak awal.



Gambar 1 Diagram Waterfall

1. *Requirement Analysis* (Analisis Kebutuhan)
Tahap ini berfokus pada pengumpulan dan analisis kebutuhan pengguna. Semua spesifikasi sistem yang akan dikembangkan didokumentasikan dengan detail untuk memastikan pemahaman yang sama antara pengembang dan pengguna.
2. *System Design* (Perancangan Sistem)
Berdasarkan hasil analisis kebutuhan, desain sistem dibuat. Desain ini mencakup arsitektur perangkat lunak, desain *database*, antarmuka pengguna, dan spesifikasi teknis lainnya.
3. *Implementation* (Implementasi)

Pada tahap ini, pengembang mulai menulis kode untuk membangun perangkat lunak berdasarkan desain yang telah disusun. Setiap modul dikembangkan secara terpisah dan diintegrasikan di akhir tahap implementasi.

4. *Integration and Testing* (Integrasi dan Pengujian)
Setelah implementasi selesai, semua modul yang dikembangkan diintegrasikan menjadi satu sistem utuh. Sistem tersebut kemudian diuji untuk memastikan bahwa fungsionalitasnya sesuai dengan kebutuhan yang telah ditentukan.
5. *Deployment* (Penerapan)
Perangkat lunak yang telah diuji kemudian diimplementasikan di lingkungan pengguna. Pada tahap ini, perangkat lunak mulai digunakan oleh pengguna dalam situasi nyata.
6. *Maintenance* (Pemeliharaan)
Setelah perangkat lunak digunakan, tahap pemeliharaan dilakukan untuk memperbaiki *bug*, menyesuaikan sistem dengan kebutuhan baru, atau meningkatkan kinerja perangkat lunak.

3. HASIL DAN PEMBAHASAN

3.1 Metode Naive Bayes

Untuk menyelesaikan metode *Naïve Bayes* dapat dilakukan dengan persamaan-persamaan sebagai berikut[16]:

1. Baca data training.
2. Hitung jumlah dan probabilitas prior.

$$P(C_i) = \frac{S_i}{s} \dots\dots\dots (1)$$

3. Hitung probabilitas training.

$$P(K|D) = \frac{P(D|K) \times P(K)}{P(D)} \dots\dots\dots (2)$$

Di mana D adalah data *class* yang belum diketahui. Selanjutnya K adalah hipotesis suatu *class* yang spesifik. $P(K|D)$ adalah Probabilitas hipotesis K berdasarkan kondisi D. $P(K)$ yaitu probabilitas sesuai dengan hipotesis D. $P(K|D)$ merupakan probabilitas D kondisi yang berdasarkan pada hipotesis D. Sedangkan $P(D)$ adalah probabilitas K.

4. Hitung nilai probabilitas data testing.

$$P(K|a_1, \dots, a_n) = P(D)P(a_1, \dots, a_n|D) \dots\dots\dots (3)$$

5. Menghitung probabilitas akhir

$$P(K|D) = P(D|K) \times P(K|D) \dots\dots\dots (4)$$

6. Menghitung nilai akhir.

$$C = \arg \max P(D|K) \dots\dots\dots (5)$$

3.2 Contoh Perhitungan Metode Naive Bayes

Adapun langkah-langkah penyelesaian metode Naive Bayes dapat dijelaskan sebagai berikut :

1. Data Training/Sampel

Dalam penelitian ini menggunakan data nilai 150 sampel yang dikonversikan kedalam bentuk nilai yang berisikan data anak pada Pusat Pengembangan Anak (PPA), dimana tabel tersebut sudah diberi nilai untuk menentukan jenis *class* yang dapat dilihat sebagai berikut :

Tabel 1 Data Latih (Training/Sampel)

No	Nama Anak	Kognitif	Motorik	Sosial-Emosional	Komunikasi	Kreativitas	Kepribadian	Kelas
1	Eko Lubis	4	5	4	5	4	2	Akademik

2	Ruth Tio	2	5	2	3	2	5	Musik
3	Tio Ombun Manurung	4	2	4	4	5	4	Keterampilan
4	Rika Hutabarat	3	3	3	2	4	3	Musik
5	Monarika Gultom	3	4	3	5	4	4	Keterampilan
6	Willy Siahaan	4	3	5	3	3	3	Olahraga
7	Sadrakh mendrofa	3	3	2	4	5	2	Musik
8	Putri Sinaga	4	3	5	4	4	2	Teknologi
9	Riki opung sungguh	2	2	4	4	5	5	Teknologi
10	Yogirang laia	2	2	4	2	3	3	Musik
...	...	...	...	...	...	...	...	...
145	Renny Manullang	3	3	5	4	5	3	Keterampilan
146	Sartika Simamora	3	3	3	4	4	2	Musik
147	Sabrina Dachi	5	3	5	2	3	5	Keterampilan
148	Rina Wati halawa	4	4	5	5	4	5	Akademik
149	Wati Rina Sigalingging	4	4	4	4	5	2	Keterampilan
150	Saprida Sihombing	2	5	4	4	2	5	Teknologi

Pada tabel di atas, telah melakukan konversi yang melibatkan transformasi data kategorikal menjadi format yang dapat diinterpretasikan oleh algoritma *Naive Bayes*. Adapun ketentuan konversinya dapat dijelaskan sebagai berikut :

Tabel 2 Atribut

No.	Atribut	Keterangan	Nilai
1	Kognitif	Sangat Baik	5
		Baik	4
		Cukup	3
		Kurang	2
		Sangat Kurang	1
2	Motorik	Sangat Baik	5
		Baik	4
		Cukup	3
		Kurang	2
		Sangat Kurang	1
3	Sosial-Emosional	Sangat Baik	5
		Baik	4
		Cukup	3
		Kurang	2
		Sangat Kurang	1
4	Komunikasi	Sangat Baik	5
		Baik	4
		Cukup	3
		Kurang	2
		Sangat Kurang	1
5	Kreativitas	Sangat Baik	5
		Baik	4
		Cukup	3
		Kurang	2
		Sangat Kurang	1
6	Kepribadian	Sangat Baik	5
		Baik	4
		Cukup	3
		Kurang	2
		Sangat Kurang	1

2. Hitung Jumlah Dan Probabilitas Prior.

Hitung jumlah setiap kelas yang digunakan, adapun jumlah setiap kelas dapat dilihat pada contoh perhitungan dibawah ini.

a. Musik

$$P_1 = \frac{64}{150} = 0,427$$

b. Olahraga

$$P_2 = \frac{28}{150} = 0,1867$$

c. Teknologi

$$P_3 = \frac{16}{150} = 0,113$$

d. Keterampilan

$$P_4 = \frac{16}{150} = 0,107$$

e. Akademik

$$P_5 = \frac{25}{150} = 0,167$$

Sehingga untuk hasil keseluruhanya dapat dilihat pada tabel berikut.

Tabel 3 Probabilitas Kelas

Kelas	Jumlah	Probabilitas
Musik	64	0,427
Olahraga	28	0,187
Teknologi	17	0,113
Keterampilan	16	0,107
Akademik	25	0,167

3. Hitung Probabilitas Training.

Hitung jumlah setiap probabilitas pada data training yang digunakan, adapun jumlah setiap probabilitas dapat dilihat pada perhitungan dibawah ini.

a. Kognitif

1) Musik

$$P_{1.1} = \frac{0}{64} = 0,000$$

$$P_{1.2} = \frac{23}{64} = 0,359$$

$$P_{1.3} = \frac{22}{64} = 0,344$$

$$P_{1.4} = \frac{11}{64} = 0,172$$

$$P_{1.5} = \frac{8}{64} = 0,125$$

2) Olahraga

$$P_{2.1} = \frac{0}{28} = 0,000$$

$$P_{2.2} = \frac{7}{28} = 0,250$$

$$P_{2.3} = \frac{6}{28} = 0,214$$

$$P_{2.4} = \frac{13}{28} = 0,464$$

$$P_{2.5} = \frac{2}{28} = 0,071$$

3) Teknologi

$$P_{3.1} = \frac{0}{17} = 0,000$$

$$P_{3.2} = \frac{2}{17} = 0,118$$

$$P_{3.3} = \frac{4}{17} = 0,235$$

$$P_{3.4} = \frac{8}{17} = 0,471$$

$$P_{3.5} = \frac{3}{17} = 0,176$$

4) Keterampilan

$$P_{4.1} = \frac{0}{16} = 0,000$$

$$P_{4.2} = \frac{0}{16} = 0,000$$

$$P_{4.3} = \frac{7}{16} = 0,438$$

$$P_{4.4} = \frac{5}{16} = 0,313$$

$$P_{4.5} = \frac{4}{16} = 0,250$$

5) Akademik

$$P_{5.1} = \frac{0}{25} = 0,000$$

$$P_{5.2} = \frac{3}{25} = 0,120$$

$$P_{5.3} = \frac{5}{25} = 0,200$$

$$P_{5.4} = \frac{7}{25} = 0,280$$

$$P_{5.5} = \frac{10}{25} = 0,400$$

Sehingga untuk hasil keseluruhannya dapat dilihat pada tabel berikut:

Tabel 4 Probabilitas Kognitif

Kognitif	Kelas					Probabilitas				
	1	2	3	4	5	1	2	3	4	5
1	0	0	0	0	0	0,000	0,000	0,000	0,000	0,000
2	23	7	2	0	3	0,359	0,250	0,118	0,000	0,120
3	22	6	4	7	5	0,344	0,214	0,235	0,438	0,200
4	11	13	8	5	7	0,172	0,464	0,471	0,313	0,280
5	8	2	3	4	10	0,125	0,071	0,176	0,250	0,400

b. Motorik

1) Musik

$$P_{1.1} = \frac{0}{64} = 0,000$$

$$P_{1.2} = \frac{21}{64} = 0,328$$

$$P_{1.3} = \frac{21}{64} = 0,328$$

$$P_{1.4} = \frac{13}{64} = 0,203$$

$$P_{1.5} = \frac{9}{64} = 0,141$$

2) Olahraga

$$P_{2.1} = \frac{0}{28} = 0,000$$

$$P_{2.2} = \frac{4}{28} = 0,143$$

$$P_{2.3} = \frac{9}{28} = 0,321$$

$$P_{2.4} = \frac{10}{28} = 0,357$$

$$P_{2.5} = \frac{5}{28} = 0,179$$

3) Teknologi

$$P_{3.1} = \frac{0}{17} = 0,000$$

$$P_{3.2} = \frac{3}{17} = 0,176$$

$$P_{3.3} = \frac{6}{17} = 0,353$$

$$P_{3.4} = \frac{3}{17} = 0,176$$

$$P_{3.5} = \frac{5}{17} = 0,294$$

4) Keterampilan

$$P_{4.1} = \frac{0}{16} = 0,000$$

$$P_{4.2} = \frac{2}{16} = 0,125$$

$$P_{4.3} = \frac{5}{16} = 0,313$$

$$P_{4.4} = \frac{4}{16} = 0,250$$

$$P_{4.5} = \frac{5}{16} = 0,313$$

5) Akademik

$$P_{5.1} = \frac{0}{25} = 0,000$$

$$P_{5.2} = \frac{3}{25} = 0,120$$

$$P_{5.3} = \frac{2}{25} = 0,080$$

$$P_{5.4} = \frac{7}{25} = 0,280$$

$$P_{5.5} = \frac{13}{25} = 0,520$$

Sehingga untuk hasil keseluruhannya dapat dilihat pada tabel berikut:

Tabel 5 Probabilitas Motorik

Motorik	Kelas					Probabilitas				
	1	2	3	4	5	1	2	3	4	5
1	0	0	0	0	0	0,000	0,000	0,000	0,000	0,000
2	21	4	3	2	3	0,328	0,143	0,176	0,125	0,120
3	21	9	6	5	2	0,328	0,321	0,353	0,313	0,080
4	13	10	3	4	7	0,203	0,357	0,176	0,250	0,280
5	9	5	5	5	13	0,141	0,179	0,294	0,313	0,520

c. Sosial-Emosional

1) Musik

$$P_{1.1} = \frac{0}{64} = 0,000$$

$$P_{1.2} = \frac{24}{64} = 0,375$$

$$P_{1.3} = \frac{18}{64} = 0,281$$

$$P_{1.4} = \frac{12}{64} = 0,188$$

$$P_{1.5} = \frac{10}{64} = 0,156$$

2) Olahraga

$$P_{2.1} = \frac{0}{28} = 0,000$$

$$P_{2.2} = \frac{5}{28} = 0,179$$

$$P_{2.3} = \frac{2}{28} = 0,071$$

$$P_{2.4} = \frac{8}{28} = 0,286$$

$$P_{2.5} = \frac{13}{28} = 0,464$$

3) Teknologi

$$P_{3.1} = \frac{0}{17} = 0,000$$

$$P_{3.2} = \frac{5}{17} = 0,294$$

$$P_{3.3} = \frac{4}{17} = 0,235$$

$$P_{3.4} = \frac{6}{17} = 0,353$$

$$P_{3.5} = \frac{2}{17} = 0,118$$

4) Keterampilan

$$P_{4.1} = \frac{0}{16} = 0,000$$

$$P_{4.2} = \frac{1}{16} = 0,063$$

$$P_{4.3} = \frac{4}{16} = 0,250$$

$$P_{4.4} = \frac{7}{16} = 0,438$$

$$P_{4.5} = \frac{4}{16} = 0,250$$

5) Akademik

$$P_{5.1} = \frac{0}{25} = 0,000$$

$$P_{5.2} = \frac{2}{25} = 0,080$$

$$P_{5.3} = \frac{4}{25} = 0,160$$

$$P_{5.4} = \frac{11}{25} = 0,440$$

$$P_{5.5} = \frac{8}{25} = 0,320$$

Sehingga untuk hasil keseluruhannya dapat dilihat pada tabel berikut:

Tabel 6 Probabilitas Sosial-Emosional

Sosial- Emosional	Kelas					Probabilitas				
	1	2	3	4	5	1	2	3	4	5
1	0	0	0	0	0	0,000	0,000	0,000	0,000	0,000
2	24	5	5	1	2	0,375	0,179	0,294	0,063	0,080
3	18	2	4	4	4	0,281	0,071	0,235	0,250	0,160
4	12	8	6	7	11	0,188	0,286	0,353	0,438	0,440
5	10	13	2	4	8	0,156	0,464	0,118	0,250	0,320

d. Komunikasi

1) Musik

$$P_{1.1} = \frac{0}{64} = 0,000$$

$$P_{1.2} = \frac{20}{64} = 0,313$$

$$P_{1.3} = \frac{18}{64} = 0,281$$

$$P_{1.4} = \frac{14}{64} = 0,219$$

$$P_{1.5} = \frac{12}{64} = 0,188$$

2) Olahraga

$$P_{2.1} = \frac{0}{28} = 0,000$$

$$P_{2.2} = \frac{6}{28} = 0,214$$

$$P_{2.3} = \frac{8}{28} = 0,286$$

$$P_{2.4} = \frac{8}{28} = 0,286$$

$$P_{2.5} = \frac{6}{28} = 0,214$$

3) Teknologi

$$P_{3.1} = \frac{0}{17} = 0,000$$

$$P_{3.2} = \frac{1}{17} = 0,059$$

$$P_{3.3} = \frac{2}{17} = 0,118$$

$$P_{3.4} = \frac{4}{17} = 0,235$$

$$P_{3.5} = \frac{10}{17} = 0,588$$

4) Keterampilan

$$P_{4.1} = \frac{0}{16} = 0,000$$

$$P_{4.2} = \frac{3}{16} = 0,188$$

$$P_{4.3} = \frac{3}{16} = 0,188$$

$$P_{4.4} = \frac{6}{16} = 0,375$$

$$P_{4.5} = \frac{4}{16} = 0,250$$

5) Akademik

$$P_{5.1} = \frac{0}{25} = 0,000$$

$$P_{5.2} = \frac{1}{25} = 0,040$$

$$P_{5.3} = \frac{3}{25} = 0,120$$

$$P_{5.4} = \frac{7}{25} = 0,280$$

$$P_{5.5} = \frac{14}{25} = 0,560$$

Sehingga untuk hasil keseluruhannya dapat dilihat pada tabel berikut:

Tabel 7 Probabilitas Komunikasi

Komunikasi	Kelas					Probabilitas				
	1	2	3	4	5	1	2	3	4	5
1	0	0	0	0	0	0,000	0,000	0,000	0,000	0,000
2	20	6	1	3	1	0,313	0,214	0,059	0,188	0,040
3	18	8	2	3	3	0,281	0,286	0,118	0,188	0,120
4	14	8	4	6	7	0,219	0,286	0,235	0,375	0,280
5	12	6	10	4	14	0,188	0,214	0,588	0,250	0,560

e. Kreativitas

1) Musik

$$P_{1.1} = \frac{0}{64} = 0,000$$

$$P_{1.2} = \frac{28}{64} = 0,438$$

$$P_{1.3} = \frac{18}{64} = 0,281$$

$$P_{1.4} = \frac{11}{64} = 0,172$$

$$P_{1.5} = \frac{7}{64} = 0,109$$

2) Olahraga

$$P_{2.1} = \frac{0}{28} = 0,000$$

$$P_{2.2} = \frac{8}{28} = 0,286$$

$$P_{2.3} = \frac{9}{28} = 0,321$$

$$P_{2.4} = \frac{7}{28} = 0,250$$

$$P_{2.5} = \frac{4}{28} = 0,143$$

3) Teknologi

$$P_{3.1} = \frac{0}{17} = 0,000$$

$$P_{3.2} = \frac{4}{17} = 0,235$$

$$P_{3.3} = \frac{4}{17} = 0,235$$

$$P_{3.4} = \frac{7}{17} = 0,412$$

$$P_{3.5} = \frac{2}{17} = 0,118$$

4) Keterampilan

$$P_{4.1} = \frac{0}{16} = 0,000$$

$$P_{4.2} = \frac{2}{16} = 0,125$$

$$P_{4.3} = \frac{2}{16} = 0,125$$

$$P_{4.4} = \frac{4}{16} = 0,250$$

$$P_{4.5} = \frac{8}{16} = 0,500$$

5) Akademik

$$P_{5.1} = \frac{0}{25} = 0,000$$

$$P_{5.2} = \frac{1}{25} = 0,040$$

$$P_{5.3} = \frac{4}{25} = 0,160$$

$$P_{5.4} = \frac{9}{25} = 0,360$$

$$P_{5.5} = \frac{25}{25} = 0,440$$

Sehingga untuk hasil keseluruhannya dapat dilihat pada tabel berikut:

Tabel 8 Probabilitas Kreativitas

Kreativitas	Kelas					Probabilitas				
	1	2	3	4	5	1	2	3	4	5
1	0	0	0	0	0	0,000	0,000	0,000	0,000	0,000
2	28	8	4	2	1	0,438	0,286	0,235	0,125	0,040
3	18	9	4	2	4	0,281	0,321	0,235	0,125	0,160
4	11	7	7	4	9	0,172	0,250	0,412	0,250	0,360
5	7	4	2	8	11	0,109	0,143	0,118	0,500	0,440

f. Kepribadian

1) Musik

$$P_{1.1} = \frac{0}{64} = 0,000$$

$$P_{1.2} = \frac{29}{64} = 0,453$$

$$P_{1.3} = \frac{20}{64} = 0,313$$

$$P_{1.4} = \frac{10}{64} = 0,156$$

$$P_{1.5} = \frac{5}{64} = 0,078$$

2) Olahraga

$$P_{2.1} = \frac{0}{28} = 0,000$$

$$P_{2.2} = \frac{8}{28} = 0,286$$

$$P_{2.3} = \frac{8}{28} = 0,286$$

$$P_{2.4} = \frac{8}{28} = 0,286$$

$$P_{2.5} = \frac{4}{28} = 0,143$$

3) Teknologi

$$P_{3.1} = \frac{0}{17} = 0,000$$

$$P_{3.2} = \frac{4}{17} = 0,235$$

$$P_{3.3} = \frac{3}{17} = 0,176$$

$$P_{3.4} = \frac{5}{17} = 0,294$$

$$P_{3.5} = \frac{5}{17} = 0,294$$

4) Keterampilan

$$P_{4.1} = \frac{0}{16} = 0,000$$

$$P_{4.2} = \frac{3}{16} = 0,188$$

$$P_{4.3} = \frac{2}{16} = 0,125$$

$$P_{4.4} = \frac{7}{16} = 0,438$$

$$P_{4.5} = \frac{4}{16} = 0,250$$

5) Akademik

$$P_{5.1} = \frac{0}{25} = 0,000$$

$$P_{5.2} = \frac{3}{25} = 0,120$$

$$P_{5.3} = \frac{2}{25} = 0,080$$

$$P_{5.4} = \frac{4}{25} = 0,160$$

$$P_{5.5} = \frac{16}{25} = 0,640$$

Sehingga untuk hasil keseluruhannya dapat dilihat pada tabel berikut:

Tabel 9 Probabilitas Kepribadian

Kepribadian	Kelas					Probabilitas				
	1	2	3	4	5	1	2	3	4	5
1	0	0	0	0	0	0,000	0,000	0,000	0,000	0,000
2	29	8	4	3	3	0,453	0,286	0,235	0,188	0,120
3	20	8	3	2	2	0,313	0,286	0,176	0,125	0,080
4	10	8	5	7	4	0,156	0,286	0,294	0,438	0,160
5	5	4	5	4	16	0,078	0,143	0,294	0,250	0,640

4. Hitung Nilai Probabilitas Data Testing/Uji.

Sampel dibawah ini, sampel yang digunakan untuk pengujian data, Adapun untuk datanya dapat dilihat pada tabel berikut.

Tabel 10 Data Uji (Testing)

No.	Nama Anak	A1	A2	A3	A4	A5	A6	Kelas
151	David Sihombing	2	3	2	5	4	5	?

Hitung setiap probabilitas data testing yang dapat dijelaskan pada tabel berikut :

Tabel 11 Probabilitas Data Uji (Testing)

David Sihombing	Musik	Olahraga	Teknologi	Keterampilan	Akademik
	1	2	3	4	5
Kognitif (2)	0,359	0,250	0,118	0,000	0,120
Motorik (3)	0,328	0,321	0,353	0,313	0,080
Sosial-Emosional (2)	0,375	0,179	0,294	0,063	0,080
Komunikasi (5)	0,188	0,214	0,588	0,250	0,560
Kreativitas (4)	0,172	0,250	0,412	0,250	0,360
Kepribadian (5)	0,078	0,143	0,294	0,250	0,640

Kalikan semua probabilitas pada data uji, contoh perhitungan ya dapat dijelaskan dibawah ini :

$$\begin{aligned}
 P_{\text{musik}} &= 0,359 \times 0,328 \times 0,375 \times 0,188 \times 0,172 \times 0,078 \\
 &= 0,0001113 \\
 P_{\text{olahraga}} &= 0,250 \times 0,321 \times 0,179 \times 0,214 \times 0,250 \times 0,143 \\
 &= 0,0001098 \\
 P_{\text{teknologi}} &= 0,118 \times 0,353 \times 0,294 \times 0,588 \times 0,412 \times 0,294 \\
 &= 0,0008700 \\
 P_{\text{keterampilan}} &= 0,000 \times 0,313 \times 0,063 \times 0,250 \times 0,250 \times 0,250 \\
 &= 0,0000000 \\
 P_{\text{akademik}} &= 0,120 \times 0,080 \times 0,080 \times 0,560 \times 0,360 \times 0,640 \\
 &= 0,0000991
 \end{aligned}$$

5. Menghitung Probabilitas Akhir.

Hitung semua probabilitas akhir dengan cara kalikan semua hasil probabilitas setiap kelas yang dijelaskan sebagai berikut :

$$\begin{aligned}
 P_{\text{musik}} &= 0,427 \times 0,0001113 \\
 &= 0,0000475 \\
 P_{\text{olahraga}} &= 0,187 \times 0,0001098 \\
 &= 0,0000205 \\
 P_{\text{teknologi}} &= 0,113 \times 0,0008700 \\
 &= 0,0000986 \\
 P_{\text{keterampilan}} &= 0,107 \times 0,0000000 \\
 &= 0,0000000 \\
 P_{\text{akademik}} &= 0,167 \times 0,0000991 \\
 &= 0,0000165
 \end{aligned}$$

Sehingga untuk hasil keseluruhannya dapat dilihat pada **tabel berikut:**

Tabel 12 Probabilitas Akhir

Kelas	$P(D K)$	$P(K D)$	Probabilitas
Musik	0,427	0,0001113	0,0000475
Olahraga	0,187	0,0001098	0,0000205
Teknologi	0,113	0,0008700	0,0000986
Keterampilan	0,107	0,0000000	0,0000000
Akademik	0,167	0,0000991	0,0000165

6. Menghitung Nilai Akhir.

Normalisasikan setiap nilai probabilitas, kemudian dibagi dengan nilai probabilitas itu sendiri sehingga untuk hasil yang didapat dapat dilihat pada perhitungan berikut.

$$C_1 = \frac{0,0000475}{0,0000475 + 0,0000205 + 0,0000986 + 0,0000000 + 0,0000165} = 0,2594064$$

$$C_2 = \frac{0,0000205}{0,0000475+0,0000205+0,0000986+0,0000000+0,0000165} = 0,1119458$$

$$C_3 = \frac{0,0000986}{0,0000475+0,0000205+0,0000986+0,0000000+0,0000165} = 0,5384594$$

$$C_4 = \frac{0,0000000}{0,0000475+0,0000205+0,0000986+0,0000000+0,0000165} = 0,0000000$$

$$C_5 = \frac{0,0000165}{0,0000475+0,0000205+0,0000986+0,0000000+0,0000165} = 0,0901883$$

Sehingga untuk hasil keseluruhannya dapat dilihat pada tabel berikut:

Tabel 13 Probabilitas Akhir

Kelas	Probabilitas	Normalisasi
Musik	0,0000475	0,2594064
Olahraga	0,0000205	0,1119458
Teknologi	0,0000986	0,5384594
Keterampilan	0,0000000	0,0000000
Akademik	0,0000165	0,0901883

Maka hasil dari perhitungan menggunakan metode *naïve bayes* berdasarkan nilai tertinggi yaitu sebesar 0,5384594 “Teknologi”. Sehingga pada data yang diuji masuk kedalam kategori “Teknologi”, dan untuk tingkat persentase ya $0,5384594 \times 100 = 53.85\%$. Untuk hasil ya dapat dilihat pada tabel berikut.

3.3 Tampilan Hasil

Berikut ini dijelaskan tentang tampilan hasil dari perancangan sistem yang dibangun dapat dilihat sebagai berikut :

1. Tampilan Proses Analisa

Tampilan proses analisa pada sistem terlihat pada gambar berikut.

Data Uji: Lidya Tambunan

Nama Anak	Kognitif	Motorik	Sosial-Emosional	Komunikasi	Kreativitas	Kepribadian	Minat Bakat
Lidya Tambunan	4	3	3	2	2	3	7

Hasil Perhitungan Probabilitas untuk Lidya Tambunan

Subatribut	Musik		Olahraga		Teknologi		Keterampilan		AI
ID>Nama	Jumlah	Probabilitas	Jumlah	Probabilitas	Jumlah	Probabilitas	Jumlah	Probabilitas	Jumlah
4 - Baik	11	0.172	13	0.464	8	0.471	5	0.313	7
3 - Cukup	21	0.328	9	0.321	6	0.353	5	0.313	2
2 - Kurang	18	0.281	2	0.071	4	0.235	4	0.250	4
2 - Kurang	20	0.313	6	0.214	1	0.059	3	0.188	1
2 - Kurang	28	0.438	8	0.286	4	0.235	2	0.125	1
3 - Cukup	20	0.313	8	0.286	3	0.176	2	0.125	2
Hasil Final	0.0002891		0.0000348		0.0000108		0.0000076		0.
Hasil Normalisasi	0.8442774		0.1016341		0.0315879		0.0222774		0.
Kesimpulan Minat Bakat	Musik dengan nilai normalisasi 0.8442774 (84.43%)								

Kesimpulan untuk Lidya Tambunan

Hasil dari perhitungan menggunakan metode *Naive Bayes* berdasarkan nilai tertinggi yaitu sebesar 0.8442774 untuk kategori "Musik". Sehingga data uji ini masuk kedalam kategori "Musik", dengan tingkat persentase 84.43 %.

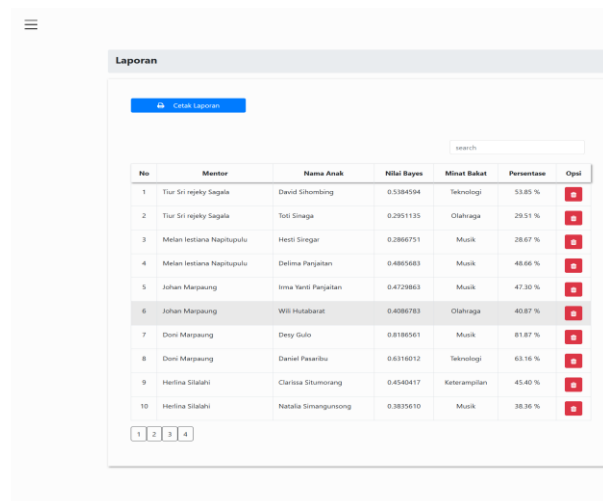
Simpan Semua Hasil Analisa

Gambar 2 Tampilan Proses Analisa

Tampilan Proses Analisa merupakan menu kunci yang menggunakan metode *Naive Bayes* untuk menganalisis minat dan bakat anak. Sistem akan melakukan analisis semua data yang diuji berdasarkan data sampel yang digunakan, kemudian sistem akan menampilkan detail hasil perhitungan, termasuk probabilitas dan kesimpulan minat bakat anak berdasarkan algoritma *Naive Bayes*. Menu ini bersifat eksklusif dan hanya dapat diakses oleh admin untuk memastikan akurasi dan keamanan proses analisis.

2. Tampilan Laporan

Tampilan laporan pada sistem terlihat pada gambar berikut.



No	Mentor	Nama Anak	Nilai Bayes	Minat Bakat	Persentase	Opsi
1	Tsur Sri regky Segala	David Sihombing	0.5384594	Teknologi	53.85 %	+
2	Tsur Sri regky Segala	Tuti Sinaga	0.2951135	Olahraga	29.51 %	+
3	Melan Iestiana Napitigulu	Hesti Siregar	0.2866751	Musik	28.67 %	+
4	Melan Iestiana Napitigulu	Defina Perjantan	0.4865683	Musik	48.66 %	+
5	Johan Marpaung	Ima Yanti Perjantan	0.4729863	Musik	47.30 %	+
6	Johan Marpaung	WIS Hutabarat	0.4086783	Olahraga	40.87 %	+
7	Doni Marpaung	Deny Gulo	0.8188561	Musik	81.87 %	+
8	Doni Marpaung	Daniel Pasaribu	0.6376012	Teknologi	63.76 %	+
9	Herlina Silalahi	Clarissa Situmorang	0.4540417	Keterampilan	45.40 %	+
10	Herlina Silalahi	Natalia Simangunsong	0.3875610	Musik	38.76 %	+

Gambar 3 Tampilan Laporan

Tampilan Laporan berfungsi untuk menampilkan hasil analisis metode *Naive Bayes* dalam bentuk data terstruktur yang mudah dibaca. Menu ini memungkinkan admin dan pimpinan untuk melihat rekapitulasi perhitungan minat dan bakat anak, termasuk detail probabilitas dan rekomendasi hasil analisis. Fitur cetak laporan tersedia untuk mendokumentasikan hasil dalam bentuk fisik atau *file*. Dengan antarmuka yang jelas dan terorganisir, menu ini membantu pihak terkait dalam mengevaluasi dan mengambil keputusan berdasarkan data yang akurat. Akses terbatas pada admin dan pimpinan menjaga kerahasiaan dan validitas informasi.

4. KESIMPULAN

Adapun kesimpulan yang dapat diambil oleh penulis adalah sebagai berikut :

1. Penggunaan algoritma Naive Bayes terbukti mampu mengklasifikasikan minat dan bakat anak ke dalam kategori seperti musik, olahraga, teknologi, keterampilan, dan akademik secara lebih sistematis dan objektif.
2. Sistem ini mampu mengolah data evaluasi anak secara cepat dan akurat, sehingga mengurangi beban kerja mentor dan mempercepat pengambilan keputusan program pengembangan anak.
3. Dengan pemrosesan otomatis berbasis probabilitas, hasil klasifikasi menjadi lebih konsisten dan tidak bergantung pada penilaian personal dari masing-masing mentor.

UCAPAN TERIMA KASIH

Sebagai penutup, penulis mengucapkan terima kasih kepada Universitas Potensi Utama yang telah memberikan kesempatan dan dukungan dalam penyelesaian Karya Tulis Ilmiah (KTI) ini.

REFERENSI

- [1] Siringoringo, C. D., & Sari, R. E. (2023). Metode Waspas dalam sistem pendukung keputusan pemilihan karyawan baru berbasis web. *Jurnal Info Digit (JID)*, 1(2), 596-611.
- [2] Sari, R. E. (2023). Perancangan Sistem Peramalan Kebutuhan Sparepart Mobil Dengan Metode Trend Projection Pada Pt. Jakarta Teknologi Utama. *Jurnal Info Digit (JID)*, 1(2), 437-449.

- [3] Sari, R. E., Saleh, A., & Afandi, M. R. (2021). Pemanfaatan Website Dalam Mengembangkan Kreativitas Siswa Di Dunia Usaha. *Jurnal TUNAS*, 2(2), 43-47.
- [4] Sitanggang, A. N., & Sari, R. E. (2024). PENERAPAN TEKNOLOGI AUGMENTED REALITY PADA MEDIA PEMBELAJARAN PERANGKAT KERAS KOMPONEN JARINGAN KOMPUTER PADA SMK TELADAN MEDAN MENGGUNAKAN METODE MARKERLESS BERBASIS ANDROID. *JURNAL WIDYA*, 5(2), 1059-1073.
- [5] Harahap, A. Y. N., Sari, R. E., Gunawan, H., & Buyung, A. (2024). Evaluation of Product Sales Data Using Clustering Method and Hierarchical Divisive Clustering at PT. AYN.
- [6] Nurisyifa, A. F., & Agustin, F. (2024). Penerapan Data Mining Menggunakan Metode C. 45 Dalam Cari Pola Minat Prangko Pada PT. Pos Indonesia Berbasis Web. *Jurnal Rekayasa Sistem (JUREKSI)*, 2(2), 474-486.
- [7] Gunawan, D. N., & Harahap, A. Y. N. (2024). Penerapan Data Mining Menggunakan Metode Clustering Dan Hierarki Divisive Clustering Dalam Evaluasi Data Penjualan. *Jurnal Rekayasa Sistem (JUREKSI)*, 2(3 A), 1568-1579.
- [8] Putera, D. Y., & Rahayu, S. L. (2023). Penerapan Data Mining Dalam Memprediksi Tingkat Perceraian Pasangan Muda Di Kota Medan Menggunakan Metode NAIVE BAYES. *Jurnal Rekayasa Sistem (JUREKSI)*, 1(1), 18-29.
- [9] Nst, M. A. P., & Harahap, F. A. (2024). Mengukur Tingkat Kepuasan Mahasiswa Terhadap Pelayanan Administrasi Keuangan Mahasiswa (AKM) pada Universitas Potensi Utama Menggunakan Metode Naïve Bayes. *Jurnal Rekayasa Sistem (JUREKSI)*, 2(3 A), 1360-1375.
- [10] Gunawan, N., & Sembiring, N. S. B. (2024). Memprediksi Produk Kosmetik Terlaris Menggunakan Metode Trend Moment. *Jurnal Rekayasa Sistem (JUREKSI)*, 2(3 A), 1401-1414.
- [11] Syiddiq, M. S., & Wahyuni, L. (2024). Sistem Pendukung Keputusan Penerimaan Bantuan Siswa Miskin Menggunakan Metode Naive Bayes Berbasis Online Pada SD Swasta Budi Mulia. *Jurnal Rekayasa Sistem (JUREKSI)*, 2(1), 289-312.
- [12] Syahri, A., & Setiawan, A. (2024). Penerapan Metode Naive Bayes Dalam Menentukan Telur Ayam Layak Konsumsi Berbasis Web. *Jurnal Rekayasa Sistem (JUREKSI)*, 2(3 A), 1264-1278.
- [13] Hasanah, U. (2022). Data Mining Menentukan Balita Yang Mendapatkan Makanan Tambahan Dari Kemenkes Dengan Metode Naive Bayes (Studi Kasus: Puskesmas Pembantu Bantan).
- [14] Munir, M. S., Nugroho, D. A., & Ahsan, M. (2024). KLASIFIKASI TINGKATAN KALORI MENU RESTORAN FAST FOOD DI INDONESIA MENGGUNAKAN METODE NAIVE BAYES. *BIMASAKTI: Jurnal Riset Mahasiswa Bidang Teknologi Informasi*, 6(2), 151-161.
- [15] Zuhra, K., & Rahmad, I. F. (2023). Penerapan Metode NAIVE BAYES Untuk Menganalisis Minat Belajar. *Jurnal Info Digit (JID)*, 1(1), 295-307.
- [16] Damanik, A. R., Sumijan, S., & Nurcahyo, G. W. (2021). Prediksi tingkat kepuasan dalam pembelajaran daring menggunakan algoritma Naïve Bayes. *Jurnal Sistim Informasi Dan Teknologi*, 88-94.