

KNOWLEDGE DISCOVERY DALAM PREDIKSI NILAI SISWA BERDASARKAN BULAN KELAHIRAN SISWA DI SMP NEGERI 2 TUNGKAL JAYA MENGGUNAKAN NAIVE BAYES

Muhammad Haidar Afif Mufid, Shopi Attika Putri, Sindy Gustiani,
Rintan Wirnanti, Nico Sabar Manahan, Ken Ditha Tania, Ahmad Rifai
Sistem Informasi, Universitas Sriwijaya
Jalan Lintas Palembang-Prabumulih KM 32 Indralaya Ogan Ilir, Indonesia
afifhaidar35@gmail.com

ABSTRAK

Pendidikan memiliki peran yang sangat penting dalam membentuk kualitas akademik siswa, di mana prestasi belajar dipengaruhi oleh berbagai faktor, baik yang bersifat internal maupun eksternal. Penelitian ini fokus pada hubungan antara bulan kelahiran dengan pencapaian nilai siswa di SMP Negeri 2 Tungkal Jaya, yang selama ini masih menggunakan metode tradisional tanpa mempertimbangkan faktor non-tradisional seperti bulan kelahiran. Tujuan dari penelitian ini adalah untuk membangun model prediksi nilai siswa menggunakan metode Naive Bayes, yang dapat membantu sekolah dalam mengidentifikasi siswa yang berisiko mengalami kesulitan akademik berdasarkan bulan kelahirannya, sekaligus memberikan rekomendasi untuk strategi pembelajaran yang inklusif. Metode yang digunakan dalam penelitian ini adalah *Knowledge Discovery in Databases (KDD)*, yang melalui beberapa tahapan, dimulai dari pengumpulan data dari 123 siswa, kemudian melakukan preprocessing dengan mengklasifikasikan nilai ke dalam kategori A, B, dan C, dilanjutkan dengan transformasi, dan terakhir melakukan data mining menggunakan *Naive Bayes* di *RapidMiner*. Hasil evaluasi menunjukkan bahwa akurasi tertinggi dicapai pada nilai IPS, yaitu sebesar 72,75%, sementara akurasi terendah terdapat pada Matematika dengan 40,30%. Rata-rata akurasi prediksi untuk Bahasa Indonesia, Bahasa Inggris, dan rapor berada di atas 70%.

Kata Kunci : *Naive Bayes, Bulan Kelahiran, SMP Negeri 2 Tungkal Jaya, Data Mining, Prestasi Akademik*

1. PENDAHULUAN

Pendidikan memiliki peran yang sangat penting dalam membentuk individu yang berkualitas dan mampu bersaing di era perkembangan teknologi yang pesat saat ini. Selain sebagai sarana untuk mentransfer pengetahuan secara efektif kepada siswa, sistem pendidikan juga berfungsi sebagai alat untuk mengembangkan keterampilan kritis, dorongan kreativitas, dan kemampuan adaptasi para siswa [1].

Prestasi akademik siswa menjadi salah satu tolak ukur utama dalam menilai kesuksesan proses belajar-mengajar. Tidak hanya dipengaruhi oleh faktor internal seperti motivasi dan kecerdasan, kemampuan akademik juga dipengaruhi oleh faktor eksternal seperti lingkungan pembelajaran, pola pengajaran yang digunakan, dan juga interaksi sosial dengan teman sebaya.

Prediksi nilai siswa menjadi hal yang penting dalam dunia pendidikan, terutama di sekolah-sekolah dengan contohnya SMP Negeri 2 Tungkal Jaya. Terdapat permasalahan di sekolah ini, dimana mereka kesulitan menghadapi kendala pada evaluasi prestasi akademik siswa akibat ketergantungannya pada metode konvensional, seperti intuisi guru dan pendekatan subjektif, yang dinilai kurang efektif dalam mengidentifikasi akar masalah variasi pencapaian siswa. Padahal, faktor non-tradisional seperti bulan kelahiran yang belum pernah diuji dalam penelitian sebelumnya berpotensi memberikan perspektif baru untuk memetakan risiko kesulitan

belajar. Beberapa penelitian sebelumnya telah mencoba berbagai pendekatan statistik dan machine learning untuk memprediksi nilai siswa seperti regresi linier, decision tree, dan neural network. Namun sayangnya tidak ada penelitian yang memasukkan faktor-faktor non-tradisional seperti bulan kelahiran ke dalam model prediksi tersebut.

Penelitian ini memanfaatkan metode Naive Bayes untuk memprediksikan nilai siswa berdasarkan bulan kelahiran mereka. Pemilihan Naive Bayes didasari oleh kinerja yang andal dalam mengelola data dan keandalannya dalam memberikan prediksi yang tepat meskipun dalam penggunaannya menggunakan dataset yang terbatas [2].

Selain itu Metode Naive Bayes lebih mudah digunakan karena memiliki alur perhitungan yang tidak panjang. Berdasarkan penjabaran di atas, penelitian bertujuan untuk mengevaluasi model prediksi nilai siswa yang berguna bagi sekolah dan guru dalam mengidentifikasi murid yang kemungkinan akan menghadapi kesulitan dalam hal akademik berdasarkan bulan kelahirannya. Penelitian ini juga dapat memberikan kontribusi dalam memahami faktor-faktor non akademik yang mempengaruhi prestasi akademis siswa sekaligus memberikan rekomendasi kepada pihak sekolah agar dapat merancang strategis pembelajaran yang lebih inklusif dan efisien.

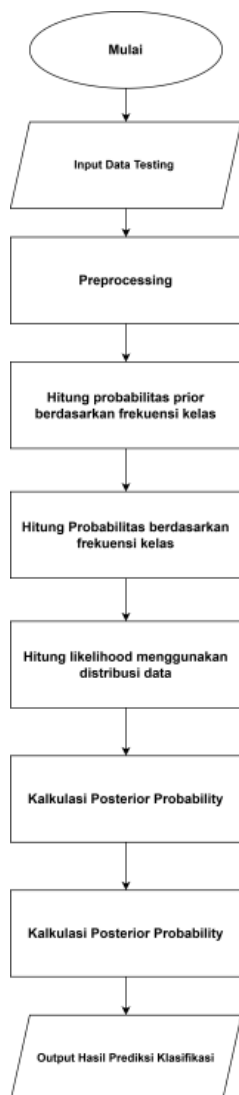
2. TINJAUAN PUSTAKA

2.1. Knowledge Discovery

Knowledge Discovery adalah proses sistematis untuk menemukan pola baru yang bermanfaat dalam data. Hal hal seperti memahami konteks, memilih data, membersihkan, mengurangi, melakukan analisis dengan algoritma pengolahan data, yang kemudian diinterpretasikan hasilnya. Tujuannya adalah untuk menemukan pola tersembunyi demi menghasilkan pemahaman baru [3][4].

2.2. Naive Bayes

Naive Bayes adalah salah satu metode pembelajaran mesin yang mampu memprediksi peluang suatu kejadian di masa depan berdasarkan data atau pengalaman sebelumnya. Selain itu, *Naive Bayes* dianggap efektif dan memiliki akurasi yang tinggi dalam klasifikasi dibandingkan dengan metode lain [5][6]. Pada tahap ini, pemodelan *Naive Bayes Classifier* digunakan untuk mengklasifikasikan siswa dan memperkirakan nilai mereka berdasarkan bulan kelahiran mereka. Prosesnya adalah sebagai berikut:



Gambar 1. Alur Proses *Naive Bayes*

Teorema Bayes yang menjadi dasar dari *Naive Bayes* adalah seperti berikut:

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

Keterangan:

P(C|X) : Probabilitas kelas C diberikan data X (posterior probability).

P(X|C) : Probabilitas mendapatkan data X jika kelasnya adalah C (likelihood).

P(C) : Probabilitas awal kelas C sebelum melihat data X (prior probability).

P(X) : Probabilitas keseluruhan data X dalam dataset (evidence).

2.3. Data Mining

Menurut Suyanto *data mining* merupakan sekumpulan disiplin ilmu komputer yang mendefinisikan proses dengan penemuan pola-pola baru dari kumpulan *big data* meliputi metode-metode yang mempunyai irisan dari *artificial intelligence*, *machine learning*, dan statistic [7].

Berdasarkan penelitian tersebut, data mining merupakan proses pencarian pola-pola yang dapat menghasilkan suatu informasi berupa hasil prediksi.

2.4. Prediksi

Prediksi merupakan suatu proses yang digunakan untuk memproyeksikan kemungkinan hasil terhadap suatu peristiwa yang akan terjadi di masa depan berdasarkan informasi masa lalu dan masa kini. Tujuan utama dari implementasi prediksi ialah untuk meminimalisir terjadinya kesalahan. Prediksi dilakukan sebagai usaha untuk memperkirakan salah satu kemungkinan yang akan terjadi selanjutnya [8].

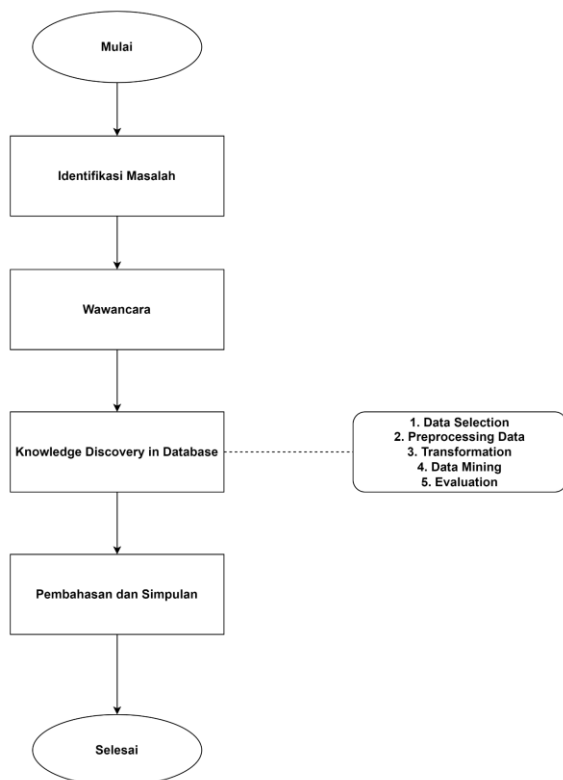
Prediksi dalam penelitian ini menggunakan teknik klasifikasi dimana klasifikasi adalah proses untuk membagi data yang ada, berdasarkan kelas kelas melalui pelatihan [9].

2.5. Rapid Miner

Rapid Miner adalah perangkat lunak pemrograman visual yang dapat digunakan untuk membuat model *data mining*, *text mining*, dan analisis prediksi tanpa harus menulis kode pemrograman. Perangkat lunak ini memiliki berbagai operator yang dapat dikombinasikan untuk membuat model yang mendukung pengambilan keputusan. Berbagai dataset, model, dan algoritma telah digunakan dalam berbagai penelitian dengan program ini [10].

3. METODE PENELITIAN

Metodologi penelitian yang digunakan dalam melakukan prediksi terhadap nilai siswa di SMP Negeri 2 Tungal Jaya dapat dilihat melalui gambar 1.



Gambar 2. Alur Proses Penelitian

3.1. Pengumpulan Data

Dataset yang digunakan pada penelitian ini merujuk pada data siswa-siswi di SMP Negeri 2 Tungkal Jaya dengan total sebanyak 123 orang. Variabel data yang digunakan untuk memprediksi nilai siswa berdasarkan bulan kelahiran ini meliputi nama siswa, nilai rata-rata ujian, dan bulan kelahiran yang diperoleh dari arsip akademik melalui wawancara secara langsung terhadap Kepala Sekolah SMP Negeri 2 Tungkal Jaya.

a. Identifikasi Masalah

Dalam proses pembelajaran, sekolah sering menghadapi variasi dalam pencapaian akademik siswa. Namun, belum banyak penelitian yang membahas pengaruh bulan kelahiran terhadap pencapaian akademik siswa. Tujuan dari penelitian ini adalah untuk mengetahui apakah ada hubungan antara bulan kelahiran dan prediksi nilai akademik siswa.

b. Wawancara

Dataset yang digunakan pada penelitian ini merujuk pada data siswa-siswi Kelas IX di SMP Negeri 2 Tungkal Jaya dengan total sebanyak 123 orang. Variabel data yang digunakan untuk memprediksi nilai siswa berdasarkan bulan kelahiran ini meliputi nama siswa, nilai rata-rata ujian, dan bulan kelahiran yang diperoleh dari arsip akademik melalui wawancara secara langsung terhadap Kepala Sekolah SMP Negeri 2 Tungkal Jaya.

3.2. Knowledge Discovery in Databases

3.2.1. Data Selection

Pada tahapan ini dilakukan ekstrak data agar data yang didapat sesuai dengan tema penelitian [11]. Dalam penelitian ini, data yang dipilih mencakup informasi siswa mulai dari nama, bulan kelahiran, dan nilai rata-rata ujian yang diperoleh dari arsip akademik sekolah.

3.2.2. Processing Data

Data Preprocessing mempunyai tujuan untuk mengubah data mentah menjadi data yang mempunyai nilai sehingga layak untuk diolah pada tahapan selanjutnya [12].

3.2.3. Transformation

Tahapan ini terjadi proses transformasi data ke dalam format yang sesuai untuk analisis dan siap untuk dilakukan penambangan data [13]. Transformation data mencakup proses normalisasi, klasifikasi, dan pembuatan atribut baru agar sesuai dengan model prediksi.

3.2.4. Data Mining

Pada tahap ini, pemodelan *Naive Bayes Classifier* digunakan untuk mengklasifikasikan siswa dan memperkirakan nilai mereka berdasarkan bulan kelahiran mereka dan mendapatkan informasi yang bermanfaat [14].

3.2.5. Evaluation

Dalam *Knowledge Discovery in Databases* (KDD) tahap evaluasi bertujuan untuk mengevaluasi kualitas dan efektivitas model yang telah dibangun pada tahap *Data Mining* [15]. Pada tahap ini, hasil klasifikasi atau prediksi model (dalam hal ini *Naive Bayes Classifier*) akan diperhitungkan hasil menggunakan *confusion matrix* dengan persamaan sebagai berikut:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Keterangan:

TP (*True Positive*) = Data positif yang diprediksi benar.

TN (*True Negative*) = Data negatif yang diprediksi benar.

FP (*False Positive*) = Data negatif yang salah diprediksi sebagai positif.

FN (*False Negative*) = Data positif yang salah diprediksi sebagai negatif.

$$precision = \frac{TP}{TP + FP}$$

Keterangan:

TP (*True Positive*) = Data positif yang diprediksi benar.

FP (*False Positive*) = Data negatif yang salah diprediksi sebagai positif.

$$recall = \frac{TP}{TP + FN}$$

Keterangan:

TP (True Positive) → Data positif yang diprediksi benar.

FN (False Negative) → Data positif yang salah diprediksi sebagai negatif

Dengan evaluasi ini klasifikasi yang dilakukan dapat menggambarkan kinerja model *naïve bayes*. [16]

4. HASIL DAN PEMBAHASAN

4.1. Data

Pada bagian ini didapatkan data dengan cara wawancara langsung dengan Kepala Sekolah SMP Negeri 2 Tungkal Jaya. Setelah berdiskusi, dilakukan identifikasi untuk atribut data yang akan digunakan, data yang didapat bisa dilihat sesuai dengan tabel 1.

Tabel 1. Data Bulan Lahir dan Nilai Rata-Rata Siswa Kelas IX SMPN 2 Tungkal Jaya Angkatan 2024

No	Bulan Lahir	MTK	B. Ing	B. Indo	IPA	IPS	Rapor
1	Mar	72,00	76,31	77,22	76,52	78,59	79,31
2	Jun	80,04	80,70	80,66	78,50	83,97	83,55
3	Mar	80,50	78,65	76,11	79,91	82,40	80,28
4	Mar	73,79	77,50	74,34	77,34	79,84	78,77
5	Okt	76,17	78,17	79,29	78,59	78,54	79,31
6	Feb	80,13	81,25	86,29	81,42	85,21	85,76
7	Mei	78,67	82,92	89,84	81,92	82,67	86,84
8	Feb	75,75	76,34	77,09	74,63	77,63	79,74
9	Jul	81,96	80,54	78,13	80,04	85,71	82,95
10	Des	72,52	75,84	75,41	77,48	77,95	79,30
...	...	...	...	...	...	...	...

4.2. Data Selection

Komponen yang paling penting untuk analisis dan klasifikasi dipilih pada tahap pemilihan data. Beberapa atribut, seperti rata-rata Bahasa Inggris, rata-rata Bahasa Indonesia, rata-rata IPA, dan rata-rata rapor adalah atribut yang akan dilakukan prediksi sebagai label.

Berdasarkan relevansi tujuan penelitian adalah untuk mengklasifikasikan siswa berdasarkan nilai rata-rata dan bulan lahir mereka. Hasil dari proses seleksi data ini disajikan dengan merepresentasikan hanya nilai rata-rata Matematika dalam bentuk tabel 2.

Tabel 2. Data Selection

No	Bulan Lahir	Rata-rata MTK
1	Maret	72,00
2	Juni	80,04
3	Maret	80,50
4	Maret	73,79
5	Oktober	76,17
6	Februari	80,13
7	Mei	78,67
8	Februari	75,75
9	Juli	81,96
10	Desember	72,52
...	...	...

4.3. Preprocessing Data

Selanjutnya dilakukan pra-pemrosesan untuk mendapatkan representasi dari data nilai agar dapat dilakukan klasifikasi. Data ini selanjutnya diubah dengan ketentuan nilai 86–100 berpredikat A, nilai 75–85 berpredikat B, nilai 56–75 berpredikat C, dan nilai 41–55 berpredikat E seperti pada tabel 3.

Tabel 3. Data Processing

No	Bulan Lahir	Rata-rata MTK
1	Maret	A
2	Juni	B
3	Maret	B
4	Maret	C
5	Oktober	C
6	Februari	C
7	Mei	C
8	Februari	B
9	Juli	B
10	Desember	C

4.4. Transformation

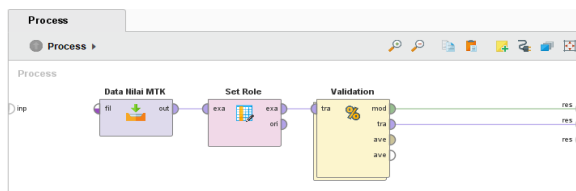
Pada tahap ini, setelah data sudah siap maka langkah selanjutnya dilakukan proses *input* data pada *RapidMiner* yang di mana Nama Siswa dijadikan sebagai id dan Bulan sebagai label seperti pada Gambar 2.

Row No.	Nama Siswa	Rata-rata M...	Bulan Lahir (...)
1	Ade Ikhsan S...	B	Juni
2	Afrada Engga...	B	April
3	Ardi Rahmad...	C	September
4	Arifka Amelia	C	Mei
5	Aulia Ririn Ari...	C	Agustus
6	Berliana Selvi...	B	Maret
7	Carira Shevina	B	Februari
8	Choirunisa U...	B	Januari
9	Difa Hadira	C	Desember
10	Fabio Arthur ...	A	Februari

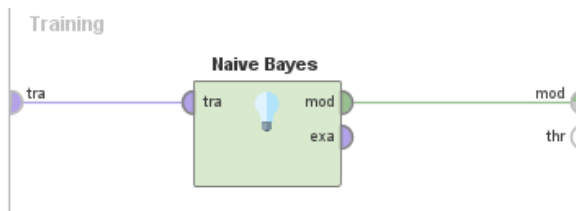
Gambar 3. Hasil Transformasi Data

4.5. Data Mining

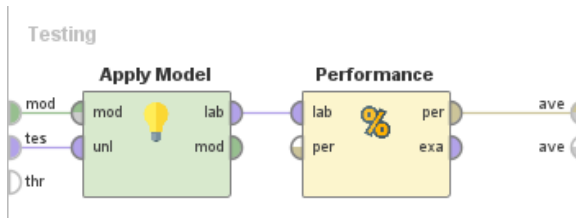
Proses *naive bayes* dalam memproses data dimulai dengan input dataset yang telah melalui *preprocessing*. Selanjutnya, menghitung probabilitas prior berdasarkan frekuensi kemunculan *data training* dan distribusi bulan kelahiran untuk setiap kategori nilai. Kemudian, menggunakan teorema Bayes untuk melakukan kalkulasi dan memprediksi kategori nilai siswa berdasarkan bulan lahir. Dan memilih kategori nilai dengan probabilitas tertinggi sebagai hasil prediksi. Gambar dibawah ini menunjukkan implementasi *RapidMiner* dengan menggunakan algoritma *Naive Bayes* untuk memprediksi nilai siswa SMPN 2 Tungal Jaya di 5 mata pelajaran dan rapor.



Gambar 3. Proses



Gambar 4. Data Training



Gambar 5. Data Testing

Berdasarkan gambar 3, 4, dan 5 bahwa telah dilakukan pemrosesan data. Proses diawali dengan pengolahan data menggunakan operator *Set Role* untuk menentukan atribut target dan fitur, lalu dilakukan validasi model dengan metode *cross-validation* untuk membagi data menjadi *training* dan *testing*. Model kemudian dilatih menggunakan *Naive Bayes*, yang mempelajari pola hubungan antar fitur dan menghasilkan model klasifikasi. Setelah itu, model diuji dengan data baru menggunakan *Apply Model*, dan performanya dievaluasi melalui operator *Performance*, yang mengukur metrik seperti akurasi, presisi, *recall*, dan F1-score untuk menilai keakuratan prediksi.

4.6. Evaluation

Setelah dilakukan penerapan metode *Naive Bayes* untuk prediksi menggunakan *RapidMiner*, berikut adalah hasil akurasi:

accuracy: 60.00%

	true A	true B	true C	class precision
pred. A	0	0	0	0.00%
pred. B	3	14	7	58.33%
pred. C	0	0	1	100.00%
class recall	0.00%	100.00%	12.50%	

Gambar 6. Akurasi Nilai Matematika

accuracy: 83.33%

	true A	true B	true C	class precision
pred. A	0	0	0	0.00%
pred. B	3	20	1	83.33%
pred. C	0	0	0	0.00%
class recall	0.00%	100.00%	0.00%	

Gambar 7. Akurasi Nilai IPA

accuracy: 64.00%

	true B	true A	true C	class precision
pred. B	16	5	1	72.73%
pred. A	3	0	0	0.00%
pred. C	0	0	0	0.00%
class recall	84.21%	0.00%	0.00%	

Gambar 8. Akurasi Nilai IPS

accuracy: 70.83%

	true B	true A	true C	class precision
pred. B	17	3	4	70.83%
pred. A	0	0	0	0.00%
pred. C	0	0	0	0.00%
class recall	100.00%	0.00%	0.00%	

Gambar 9. Akurasi Nilai Bahasa Indonesia

accuracy: 76.00%

	true B	true A	true C	class precision
pred. B	19	4	2	76.00%
pred. A	0	0	0	0.00%
pred. C	0	0	0	0.00%
class recall	100.00%	0.00%	0.00%	

Gambar 10. Akurasi Nilai Bahasa Inggris

accuracy: 84.00%

	true B	true A	class precision
pred. B	21	4	84.00%
pred. A	0	0	0.00%
class recall	100.00%	0.00%	

Gambar 11. Akurasi Nilai Rapor Siswa

Berdasarkan gambar-gambar diatas didapatkan bahwa akurasi yang beragam dari prediksi nilai setiap mapel Siswa Kelas IX SMPN 2 Tungal Jaya dengan rincian berikut:

- Prediksi nilai Matematika mendapatkan akurasi 60% dengan recall kelas B 100%, kelas C 12.50% dan kelas A 0%.
- Prediksi nilai IPA mendapatkan akurasi 83.33%, tetapi model hanya memprediksi kelas B, sehingga kelas A dan C tidak diprediksi.
- Prediksi nilai IPS mendapatkan akurasi 70.83% yang hanya dapat mengenali kelas B saja.
- Prediksi nilai Bahasa Indonesia mendapatkan akurasi 70.83%, namun model masih hanya mengenali kelas B.
- Prediksi nilai Bahasa Inggris mendapatkan akurasi 76% dengan memprediksi semua data sebagai kelas B, menyebabkan recall kelas B 100% tetapi kelas A dan C 0%.

- f. Prediksi nilai Rapor mendapatkan akurasi 84% yang hanya mengenali kelas B.

5. KESIMPULAN DAN SARAN

Penelitian ini menunjukkan bahwa metode Naive Bayes dapat digunakan untuk memprediksi nilai siswa berdasarkan bulan kelahiran, meskipun hasilnya masih bervariasi. Beberapa mata pelajaran, seperti IPS dan Bahasa Inggris, menunjukkan akurasi di atas 70%, sementara prediksi untuk Matematika masih cukup rendah, hanya mencapai akurasi 60%. Hasil ini memberikan gambaran bahwa bulan kelahiran mungkin memiliki keterkaitan dengan performa akademik siswa, meskipun bukan satu-satunya faktor yang dapat menentukan nilai siswa. Kedepannya penelitian ini bisa dikembangkan lebih lanjut dengan menggunakan dataset yang lebih luas, mempertimbangkan faktor lain seperti kehadiran siswa dan metode pengajaran, serta menguji algoritma lain seperti *Decision Tree* atau *Random Forest*. Dengan peningkatan tersebut, model ini berpotensi dapat membantu sekolah dalam merancang strategi pembelajaran yang lebih efektif dan mendukung perkembangan akademik siswa secara lebih optimal.

DAFTAR PUSTAKA

- [1] E. Mulyanti, W. Retnadi, and R. Hayu, "Inovasi Pendidikan : Strategi Meningkatkan Kreativitas dan Berpikir Kritis Siswa Melalui Project Based Learning Dengan demikian , guru menggunakan metode yang efektif melalui yang berfokus pada peserta didik melalui pelaksanaan proyek . Tujuannya adalah," vol. 1, no. 3, pp. 198–208, 2024.
- [2] A. Senika, R. Rasiban, and D. Iskandar, "Implementasi Metode Naïve Bayes Dalam Penilaian Kinerja Sales Marketing Pada PT. Pachira Distrinusa," *J. Media Inform. Budidarma*, vol. 6, no. 1, p. 701, 2022, doi: 10.30865/mib.v6i1.3331.
- [3] X. Shu and Y. Ye, "Knowledge Discovery: Methods from data mining and machine learning," *Soc. Sci. Res.*, vol. 110, no. October 2022, p. 102817, 2023, doi: 10.1016/j.ssresearch.2022.102817.
- [4] N. M. Putry, "Komparasi Algoritma Knn Dan Naïve Bayes Untuk Klasifikasi Diagnosis Penyakit Diabetes Mellitus," *EVOLUSI J. Sains dan Manaj.*, vol. 10, no. 1, 2022, doi: 10.31294/evolusi.v10i1.12514.
- [5] F. Dwiramadhan, M. I. Wahyuddin, and D. Hidayatullah, "Sistem Pakar Diagnosa Penyakit Kulit Kucing Menggunakan Metode Naive Bayes Berbasis Web," *J. JTIK (Jurnal Teknol. Inf. dan Komunikasi)*, vol. 6, no. 3, pp. 429–437, 2022, doi: 10.35870/jtik.v6i3.466.
- [6] A. Z. Macfud, A. P. Kusuma, and W. D. Puspitasari, "Analisis Algoritma Naive Bayes Classifier (Nbc)," vol. 7, no. 1, pp. 87–94, 2023.
- [7] R. A. Kelvin Gautama, "Penerapan Data Mining Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode K-Nearest Neighbor," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 6, pp. 11184–11190, 2024, doi: 10.47065/bits.v3i4.1408.
- [8] S. R. Dwi Budi Srisulistiowati, Muhamad Khaerudin,*, "Sistem Informasi Prediksi Penjualan Alat Tulis Kantor Dengan Metode Fp-Growth (Studi Kasus Toko Koperasi Sekolah Bina Mulia)," *J. Sist. Inf. Univ. Suryadarma*, vol. 8, no. 2, 2021, doi: 10.35968/jsi.v8i2.739.
- [9] F. L. D. Cahyanti, F. Sarasati, W. Astuti, and E. Firasari, "Klasifikasi Data Mining Dengan Algoritma Machine Larning Untuk Prediksi Penyakit Liver," *Technol. J. Ilm.*, vol. 14, no. 2, p. 134, 2023, doi: 10.31602/tji.v14i2.10093.
- [10] M. Sholeh, E. K. Nurnawati, and U. Lestari, "Penerapan Data Mining dengan Metode Regresi Linear untuk Memprediksi Data Nilai Hasil Ujian Menggunakan *RapidMiner*," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 8, no. 1, pp. 10–21, 2023, doi: 10.14421/jiska.2023.8.1.10-21.
- [11] D. Fitrihanah, S. Dwiasnati, H. H. H, and K. A. Baihaqi, "Penerapan Metode Machine Learning untuk Prediksi Nasabah Potensial menggunakan Algoritma Klasifikasi Naïve Bayes," *Fakt. Exacta*, vol. 14, no. 2, p. 92, 2021, doi: 10.30998/faktorexacta.v14i2.9297.
- [12] F. Alghifari and D. Juardi, "Penerapan Data Mining Pada Penjualan Makanan Dan Minuman Menggunakan Metode Algoritma Naïve Bayes," *J. Ilm. Inform.*, vol. 9, no. 02, pp. 75–81, 2021, doi: 10.33884/jif.v9i02.3755.
- [13] F. H. Muhammad Rayhan Kurniawan, Oktariani Nurul Pratiwi, "Klasifikasi soal menggunakan multi-label problem transformation dengan metode random forest dan k-nearest neighbor," *JIIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 10, no. 1, pp. 367–380, 2025.
- [14] N. Wakhidah and S. N. Rochmah, "Klasifikasi kualitas mutu susu pasteurisasi menggunakan metode klasifikasi k-Nearest Neighbor," *Aiti*, vol. 21, no. 1, pp. 58–71, 2024, doi: 10.24246/aiti.v21i1.58-71.
- [15] R. S. Wijaya, A. Qur'ania, and I. Anggraeni, "Klasifikasi Penyakit Cacar Monyet Menggunakan Support Vector Machine (SVM)," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 4, pp. 1253–1260, 2024, doi: 10.57152/malcom.v4i4.1417.
- [16] I. M. A. A. D. Putra, I. M. G. Sunarya, and I. G. A. Gunadi, "Perbandingan Algoritma Naive Bayes Berbasis Feature Selection Gain Ratio dengan Naive Bayes Kovensional dalam Prediksi Komplikasi Hipertensi," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 6, no. 1, pp. 37–49, 2024, doi: 10.35746/jtim.v6i1.488.