

Analisis Keberhasilan Pembelajaran Daring pada Masa Pandemi Covid-19 menggunakan Algoritma C4.5 dan Naive Bayes

Irma Agustika Sihombing¹, Dedy Hartama², Iin Parlina³, Indra Gunawan⁴, Ika Okta Kirana⁵

^{1,2,4,5}STIKOM Tunas Bangsa Pematangsiantar, Jl. Jend. Sudirman Blok A No.1-3 Pematangsiantar, Indonesia

³AMIK Tunas Bangsa Pematangsiantar, Jl. Jend. Sudirman Blok A No.1-3 Pematangsiantar, Indonesia

irmasihombing006@gmail.com

Abstrak. Strategi melawan pandemi dengan pembatasan sosial memaksa semua institusi pendidikan menerapkan pembelajaran daring. Pembelajaran daring yang awalnya sebagai strategi menjadi kontroversi karena singkatnya proses adaptasi. Perubahan mendadak dari pembelajaran tatap muka ke pembelajaran daring pada skala besar menyebabkan respon di masyarakat. Tujuan dari penelitian ini adalah untuk menganalisis Algoritma C4.5 dan Naive Bayes pada masa pandemi COVID-19 di SMK Swasta Persiapan Pematangsiantar. Data yang diperoleh dari penelitian ini diperoleh dari Sekolah SMK Swasta Persiapan Pematangsiantar. Berdasarkan hasil pengujian Didalam mata pelajaran Adaptif terdapat Performace atau tingkat Akurasi untuk Algoritma C4.5 sebanyak 98% dan Algoritma Naive Bayes sebanyak 99%. Sehingga Algoritma Naive Bayes lebih baik digunakan untuk memperoleh hasil dari Analisis Keberhasilan Pembelajaran Daring di masa Pandemi COVID-19.

Kata Kunci : Hasil Pembelajaran, Algoritma C4.5, Naive Bayes, Pandemi.

Abstract. Strategies to fight the pandemic with social policies are forcing all educational institutions to apply bold learning. Bold learning which started as a strategy became controversial because of the shortness of the adaptation process. The change from face-to-face learning to online learning on a large scale caused a response in society. The purpose of this study was to analyze the C4.5 and Naive Bayes Algorithm during the COVID-19 pandemic at the Pematangsiantar Preparatory Private Vocational School. The data obtained from this study were obtained from the Pematangsiantar Preparatory Private Vocational School. Based on testing in Adaptive subjects, there is a performance or accuracy level for the C4.5 Algorithm as much as 98% and the Naive Bayes Algorithm as much as 99%. So that the Naive Bayes Algorithm is better used to obtain results from the Analysis of Online Learning Success during the COVID-19 Pandemic.

Keyword : Learning Outcomes, C4.5 Algorithm, Naive Bayes, Pandemic

PENDAHULUAN

Data mining merupakan bidang ilmu yang digunakan untuk menangani masalah untuk pengambilan informasi dari database yang besar dengan menggabungkan teknik dari statistik, pembelajaran mesin, visualisasi data, pengenalan pola, dan database. Tujuan dari data mining adalah untuk mengekstrak informasi dengan metode cerdas dari kumpulan data kemudian mengubah informasi menjadi struktur yang dapat dipahami untuk penggunaan lebih lanjut[1]. Output dalam data mining dapat dipergunakan sebagai alternatif dalam pengambilan keputusan atau untuk memperbaiki pengambilan keputusan di masa yang akan datang[2],[3]. Pendidikan merupakan usaha sadar dan terencana untuk mewujudkan suasana belajar dan proses pembelajaran peserta didik secara aktif mengembangkan potensi dirinya untuk memiliki kekuatan spiritual keagamaan, pengendalian diri, kepribadian, kecerdasan, akhlak mulia serta keterampilan yang diperlukan dirinya[4]. Tujuan pendidikan untuk mengembangkan potensi peserta didik, agar menjadi manusia yang beriman dan bertaqwa kepada Tuhan Yang Maha Esa, berakhlak mulia, sehat, berilmu, cakap, kreatif, mandiri dan menjadi warga Negara yang demokratis serta bertanggung jawab[5].

Berdasarkan tujuan pendidikan tersebut maka kualitas pembelajaran di sekolah atau lembaga pendidikan perlu ditingkatkan karena mempengaruhi tingkat keberhasilan pembelajaran seorang siswa. Wabah dari Pandemi Corona Virus Disease 19 atau disebut COVID-19 menimbulkan situasi yang menuntut setiap lapisan masyarakat untuk tetap menjaga kesehatan dan melakukan aktivitas dari



rumah. Pembatasan sosial berdampak pada ruang publik, termasuk kantor, sekolah atau kampus. Didalam pendidikan strategi ini melarang proses belajar tatap muka dan memaksa semua institusi pendidikan menghentikan proses belajar dan menggantinya dengan pembelajaran daring.

Pembelajaran secara konvensional merupakan proses pembelajaran yang dilakukan dengan menggabungkan satu atau lebih metode pembelajaran dan guru mempunyai peran penting dalam pendekatan ini. Pembelajaran daring dapat didefinisikan sebagai pembelajaran berbasis teknologi dimana bahan belajar dikirim secara elektronik kepeserta didik jarak jauh menggunakan jaringan komputer. Kementerian Pendidikan dan Kebudayaan Republik Indonesia mempertegas kebijakan pembelajaran daring dalam masa pandemi dengan mengeluarkan kebijakan belajar dari rumah. Untuk itu mewajibkan siswa menggunakan smartphone, Gadget, komputer dan aplikasi sebagai pengganti tatap muka. Hal yang mempengaruhi tingkat keberhasilan pembelajaran daring yaitu teknologi, karakteristik belajar dan karakteristik siswa, akan tetapi banyak kendala yang terjadi di Indonesia seperti infrastruktur dan minimnya access point.

Dari permasalahan diatas, penulis melakukan penelitian yang terkait tentang analisis keberhasilan pembelajaran daring dengan menggunakan data mining yaitu Algoritma C4.5 dan Naïve Bayes[6]. Algoritma C4.5 merupakan algoritma yang digunakan untuk mengeksplorasi data yaitu untuk menemukan hubungan variable input atau attribute kriteria dengan variable target atau decision attribute (atribut keputusan)[7], sedangkan Algoritma Naïve Bayes yaitu salah satu algoritma klasifikasi berdasarkan teorema Bayesin pada statistika[8],[9],[10]. Algoritma Naïve Bayes dapat digunakan untuk memprediksi probabilitas keanggotaan suatu kelas. Penulis menggunakan Data Nilai Siswa SMK Swasta Persiapan Pematangsiantar tahun ajaran 2020/2021 sebanyak 100 data yang akan diimplementasikan dengan algoritma C4.5 dan Naïve Bayes, Sehingga dihasilkan model aturan yang diperoleh untuk klasifikasi keberhasilan siswa dalam mempelajari setiap mata pelajaran yang dikirimkan lewat pembelajaran daring oleh guru. Penelitian ini Menggunakan perbandingan antara algoritma C4.5 dan Naive Bayes untuk mengetahui bagaimana tingkat keberhasilan seorang siswa dalam proses pembelajaran daring.

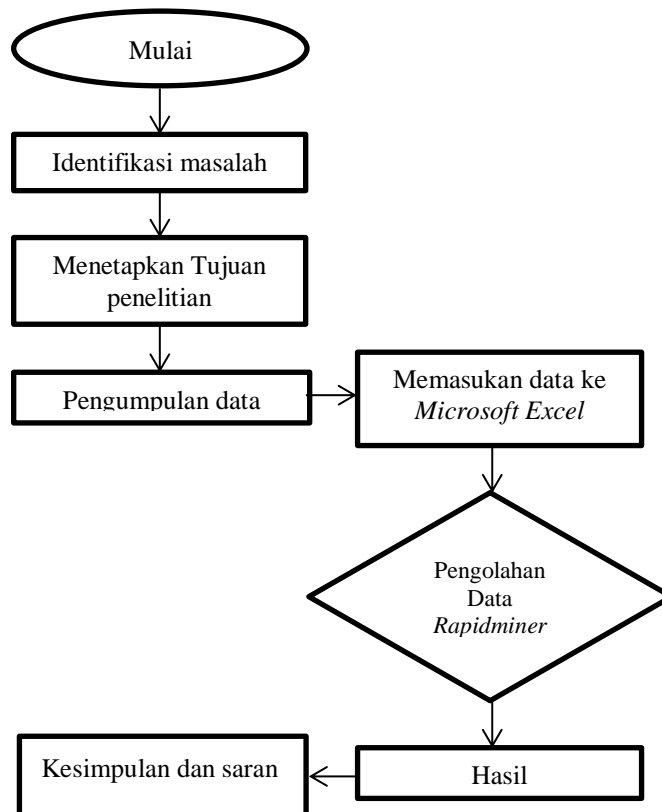
METODOLOGI PENELITIAN

Dalam penelitian ini akan dilakukan analisis keberhasilan pembelajaran daring pada masa pandemi covid-19 di SMK Swasta Persiapan Pematangsiantar menggunakan Algoritma C4.5 dan Naïve Bayes dimana kedua metode tersebut akan dilakukan perbandingan yang akan menghasilkan hasil yang lebih optimal, serta memperbaiki tingkat keberhasilan siswa dalam belajar pada masa pandemi covid-19. Data diperoleh berasal dari data nilai siswa tahun ajaran 2020/2021 di SMK Swasta Persiapan Pematangsiantar. Rancangan penelitian yang dilakukan untuk Analisis keberhasilan pembelajaran daring dengan menggunakan Algoritma C4.5 dan Naïve Bayes dimulai dari studi kepustakaan yang merupakan langkah untuk mendapatkan ataupun mengumpulkan informasi yang terkait terhadap topik, metode atau masalah yang terjadi pada penelitian ini. Maka pada tahap ini pengumpulan informasi di peroleh dari berbagai jenis jurnal yang terkait dengan topik penelitian ini, selanjutnya dilakukan pengumpulan data dan mengidentifikasi permasalahan yaitu bagaimana keberhasilan pembelajaran daring pada masa pandemi, selanjutnya mengolah data dengan Algoritma C4.5 dan Algoritma Naïve Bayes, penyelesaian permasalahan dan membuat kesimpulan.

Pada pengolahan data atribut yang akan dijadikan Prediktor untuk menghasilkan atribut target, Yaitu (1) Mata pelajaran Agama, Pkn, Bahasa Indonesia, Penjaskes, Senibudaya Variabel Normatif. (2) Adaptif diambil dari Bahasa Inggris, Matematika, Ipa, Kimia, Fisika, Ips. (3) Produktif diambil dari Mata pelajaran Mendiagnosis Permasalahan Pengoperasian, Melakukan perbaikan Setting Ulang, Melakukan perbaikan Peripheral, Melakukan perbaikan PC, dan Instalasi Software. (4) Muatan lokal diambil dari Conversation yang merupakan mata pelajaran bahasa tambahan. Data akan



diolah sehingga menghasilkan model aturan yang berisikan faktor yang paling mempengaruhi keberhasilan siswa dalam pembelajaran daring. Dalam pengolahan data transformasi menggunakan perangkat lunak RapidMiner versi 5.3 dengan diagram aktifitas kerja seperti pada gambar 1 berikut:



Gambar 1. Diagram Aktifitas

Gambar diatas menjelaskan bagaimana mengidentifikasi masalah dan menentukan tujuan penelitian yang dilakukan, mengumpulkan data dan kemudian memasukan data kedalam Microsoft excel , kemudian dilakukan analisis Algoritma C4.5 dan Naïve Bayes kedalam RapidMiner untuk mendapatkan hasil dan tingkat ke akurasiannya.

HASIL DAN PEMBAHASAN

Hasil penelitian sesuai dengan kriteria dan alternatif yang digunakan pada saat pengumpulan data. Data yang telah dikumpulkan untuk diolah atau ditransformasikan ke format data Ms.Excel 2010. Data yang ditransformasikan tersebut digunakan sebagai syarat dalam pengolahan Algoritma C4.5 dan Naïve Bayes. Analisa data menggunakan tools RapidMiner sebagai implementasi dari pengolahan data. Hasil dari penelitian digunakan untuk menentukan keputusan untuk keberhasilan dari pembelajaran daring di masa pandemi Covid-19.

a. Pengolahan data dengan algoritma C4.5

Pengolahan data menggunakan Data Mining dengan Algoritma C4.5. Algoritma C4.5 dilakukan untuk memberikan model aturan dari data yang digunakan.

Hasil dari perhitungan nilai Entropy, nilai Gain, SplitInfo dan GainRatio untuk tiap atribut dapat dilihat pada Tabel 4.1 dengan keterangan SB =Sangat Baik, B =Baik, C =Cukup, K =Kurang dan SK =Sangat Kurang sebagai berikut:

$$Entropy(S) = \sum_{i=1}^n - p_i * \log_2 p_i$$

Tabel 1. Perhitungan Gain Node 1

No	Atribut		Jumlah Kasus	Iya	Tidak	Entropy	Gain	Split Info	Gain Ratio
Total			100	96	4	0,2423			
1	NORMATIF						0,2062	0,5257	0,3923
		Sangat Baik	4	4	0	0,0000			
		Baik	91	91	0	0,0000			
		Cukup	5	1	4	0,7219			
		Kurang	0	0	0	0,0000			
2	ADAPTIF						0,2062	0,4785	0,4309
		Sangat Baik	3	3	0	0,0000			
		Baik	92	92	0	0,0000			
		Cukup	5	1	4	0,7219			
		Kurang	0	0	0	0,0000			
3	PRODUKTIF						0,1663	1,0536	0,1578
		Sangat Baik	24	24	0	0,0000			
		Baik	72	71	1	0,1056			
		Cukup	1	1	0	0,0000			
		Kurang	3	0	3	0,0000			
4	MULOK						0,2423	1,0512	0,2305
		Sangat Baik	26	26	0	0,0000			
		Baik	70	70	0	0,0000			
		Cukup	4	0	4	0,0000			
		Kurang	0	0	0	0,0000			
		Sangat Kurang	0	0	0				

Dari Tabel.1 diketahui nilai GainRatio terbesar adalah atribut Adaptif dengan nilai 04309, kategori Cukup dengan nilai 0,7219 dan dijadikan simpul akar untuk mengulangi perhitungan nilai Entropy dan nilai Gain selanjutnya.

Hasil dari perhitungan nilai Entropy, nilai Gain, SplitInfo dan GainRatio untuk tiap atribut dapat dilihat pada Tabel 4.2 dengan keterangan SB =Sangat Baik, B =Baik, C =Cukup, K =Kurang dan SK =Sangat Kurang sebagai berikut Hasil perhitungan nilai Entropy dan nilai Gain untuk atribut Adaptif untuk kategori Cukup, dapat dilihat pada Tabel 2 berikut :

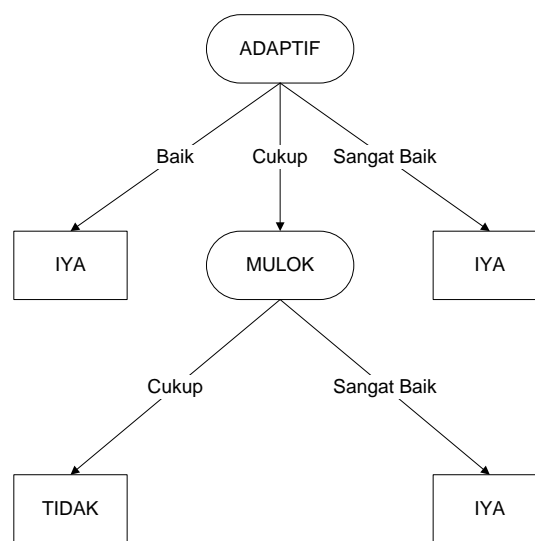
Tabel 2. Perhitungan Gain Node 1.1

No	Atribut		Jumlah Kasus	Iya	Tidak	Entropy	Gain	Split Info	Gain Ratio
ADAPTIF – CUKUP			5	1	4	0,7219			
1	NORMATIF						0,0000	0,0000	0,0000
		Sangat Baik	0	0	0	0,0000			
		Baik	0	0	0	0,0000			
		Cukup	5	1	4	0,7219			
		Kurang	0	0	0	0,0000			



No	Atribut		Jumlah Kasus	Iya	Tidak	Entropy	Gain	Split Info	Gain Ratio
		Sangat Kurang	0	0	0	0,0000			
3	PRODUKTIF						0,7219	1,3710	0,5266
		Sangat Baik	0	0	0	0,0000			
		Baik	1	0	1	0,0000			
		Cukup	1	1	0	0,0000			
		Kurang	3	0	3	0,0000			
4		Sangat Kurang	0	0	0	0,0000			
	MULOK						0,7219	0,7219	1,0000
		Sangat Baik	1	1	0	0,0000			
		Baik	0	0	0	0,0000			
		Cukup	4	0	4	0,0000			
		Kurang	0	0	0	0,0000			
	Sangat Kurang	0	0	0					

Dari Tabel 2 diketahui nilai GainRatio terbesar adalah atribut Mulok dengan nilai 1,000. Nilai Atribut Fasilitas memiliki satu Entropy, maka proses Node 1.1 berhenti. Setelah tidak ada proses dilakukan lagi, diperoleh pohon keputusan yang dapat dilihat pada Gambar 2 berikut :



Gambar 2. Pohon Keputusan

Berdasarkan pohon keputusan pada gambar 2 aturan atau Rule model yang didapatkan adalah sebagai berikut :

1. Jika Adaptif = Baik, maka hasilnya = Iya (Iya = 92 dan Tidak = 0)
2. Jika Adaptif = Cukup dan Mulok = Cukup, maka hasilnya = Tidak (Iya = 0 dan Tidak = 4)
3. Jika Adaptif = Cukup dan Mulok = Sangat Baik, maka hasilnya = Iya (Iya = 1 dan Tidak = 0)
4. Jika Adaptif = Sangat Baik, maka hasilnya = Iya (Iya = 3 dan Tidak = 0)

b. Pengolahan Data dengan Algoritma Naïve Bayes

Setelah proses pengolahan data dengan Algoritma C4.5 selanjutnya pengolahan data dengan Algoritma Naïve Bayes berikut:

Menghitung nilai berhasil Iya dan berhasil Tidak. Dari 100 data yang digunakan, diketahui kelas Iya sebanyak 96 data, dan kelas tidak sebanyak 4 data. Perhitungan probabilitas prior kemungkinan Keberhasilan Iya yaitu :

$$P(iya) = \frac{96}{100} = 0,96$$

Sedangkan perhitungan probabilitas keberhasilan Tidak yaitu :

$$P(Tidak) = \frac{4}{100} = 0,04$$

Setelah probabilitas dari masing-masing prior telah diketahui, selanjutnya penulis menghitung masing-masing probabilitas dari setiap kriteria yang digunakan. Kriteria yang digunakan penulis yaitu Normatif, Adaptif, Produktif dan Mulok. Dalam menentukan probabilitas setiap kriteria, penulis menghitung bagian-bagian yang terdapat pada setiap kriteria, pada penelitian ini bagian-bagian yang terdapat dalam setiap kriteria menggunakan skala linkert 5 yaitu Sangat Baik (SB), Baik (B), Cukup (C), Kurang (K), dan Sangat Kurang (SK). Sehingga dalam menentukan probabilitas setiap kriteria dilakukan dengan menghitung jumlah Iya dan Tidak berhasil pada skala linkert yang digunakan. Sehingga probabilitas kriteria normatif dapat dilihat pada tabel 3 berikut :

Tabel 3. Probabilitas Kriteria Normatif

Normatif	Jumlah Kejadian Dipilih		Probabilitas	
	IYA	Tidak	IYA	TIDAK
Sangat Baik	4	0	0,0417	0
Baik	91	0	0,9479	0
Cukup	1	4	0,0104	1
Kurang	0	0	0	0
Sangat Kurang	0	0	0	0
Jumlah	96	4	1	1

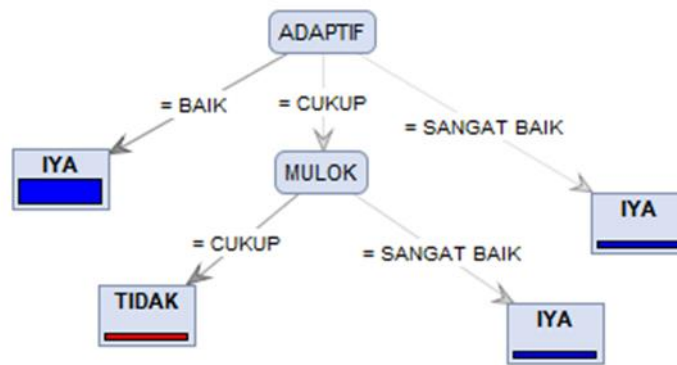
Pada tabel 3 dapat dijelaskan bahwa Probabilitas pada kriteria Normatif yaitu pada kategori iya dengan skala Sangat Baik memiliki probabilitas 0,0417, Baik memiliki probabilitas 0,9479, Cukup memiliki probabilitas 0,0104, serta Kurang dan Sangat Kurang memiliki probabilitas 0. Sehingga jumlah probabilitas puas yaitu 1. Sedangkan pada kategori tidak dengan skala Sangat Baik memiliki probabilitas 0, Baik memiliki probabilitas 0, Cukup memiliki probabilitas 1, serta Kurang dan Sangat Kurang memiliki probabilitas 0. Sehingga jumlah probabilitas tidak yaitu 1.

Selanjutnya dihitung probabilitas kemungkinan dari kriteria adaftif, kriteria produktif, dan kriteria mulok.

c. Implementasi Algoritma C4.5 dengan Tools Rapidminer

Model Algoritma C4.5 di tools RapidMiner dengan menggunakan dua operator. Kedua operator tersebut adalah operator Retrive Data yang digunakan untuk memasukkan data yang telah diimpor dari Excel Sheet dan operator Decision Tree yang digunakan untuk operator yang mengelolah data menjadi informasi dengan Algoritma C4.5 pada RatioGain. Setelah membuat desain model dari Algoritma C4.5, dapat dihasilkan gambar berupa pohon keputusan seperti berikut:





Gambar 3. Tampilan Pohon Keputusan

Pada gambar 3 dapat dilihat bahwa Node tertinggi dari pohon keputusan adalah Adaptif dengan kategori Sangat baik, Baik dan Cukup. Selanjutnya Node Mulok dengan kategori Sangat Baik dan Cukup. Setelah mendapatkan hasil pohon keputusan pada tools RapidMiner, dapat dilihat model aturan berdasarkan pohon keputusan berikut:

```
Tree
ADAPTIF = BAIK: IYA {IYA=92, TIDAK=0}
ADAPTIF = CUKUP
| MULOK = CUKUP: TIDAK {IYA=0, TIDAK=4}
| MULOK = SANGAT BAIK: IYA {IYA=1, TIDAK=0}
ADAPTIF = SANGAT BAIK: IYA {IYA=3, TIDAK=0}
```

Gambar 4. Tampilan Model Aturan

Berdasarkan gambar 4 dapat dilihat model aturan dengan keberhasilan Iya tertinggi adalah Jika Adaptif = Baik, Maka hasilnya Iya {Iya=92 dan Tidak=0}. Sedangkan model aturan dengan keberhasilan Tidak tertinggi adalah Jika Adaptif = Cukup dan Mulok = Cukup, Maka hasilnya Tidak {Iya=0 dan Tidak=4}.

- d. Implementasi Algoritma Naïve Bayes dengan Tools Rapidminer
Desain Model Algoritma Naïve Bayes menggunakan tools RapidMiner yang digunakan sebagai tampilan instalasi Algoritma Naïve Bayes. Hasil Algoritma Naïve Bayes diperoleh berdasarkan Model dari Algoritma Naïve Bayes dari tools RapidMiner sebagai berikut:

```
SimpleDistribution
Distribution model for label attribute BERHASIL

Class IYA (0.960)
4 distributions

Class TIDAK (0.040)
4 distributions
```

Gambar 5. Hasil Algoritma Naïve Bayes

Pada gambar 5 dapat dilihat bahwa hasil dari Algoritma Naïve Bayes diperoleh kelas berhasil Iya dengan nilai 0,960 dan kelas berhasil Tidak dengan nilai 0,040. Artinya kelas Iya merupakan kelas yang tertinggi pada keberhasilan pembelajaran daring di masa pandemi Covid-19.

KESIMPULAN

Faktor yang paling dominan dalam meningkatkan prestasi belajar siswa dengan menggunakan Algoritma C4.5 dan Naïve Bayes adalah mata pelajaran Adaptif. Didalam mata pelajaran Adaptif terdapat Performance atau tingkat Akurasi untuk Algoritma C4.5 sebanyak 98% dan Algoritma Naïve Bayes sebanyak 99%. Setelah dilakukan Perbandingan diantara dua algoritma tersebut, yang memiliki tingkat ke Akurasian lebih tinggi yaitu Algoritma Naïve Bayes. Optimalisasi Model aturan didapat dengan menggunakan Induction Rule adalah “ Jika Adaptif = Baik, Maka hasilnya Iya { Iya=92 dan Tidak=0}. Sedangkan model aturan dengan keberhasilan tidak tertinggi adalah jika Adaptif = Cukup dan Mulok = Cukup, Maka Hasilnya Tidak {Iya=0 dan Tidak=4} .

DAFTAR PUSTAKA

- [1] D. Firdaus, “Penggunaan Data Mining dalam Kegiatan Sistem Pembelajaran Berbantuan Komputer,” *J. Format*, vol. 6, no. 2, pp. 91–97, 2017.
- [2] Dwi Retnosari, “Sistem Aplikasi Data Mining Untuk Menampilkan,” *J. Integr. Sist. Ind. UMJ*, vol. 1, no. 2, pp. 13–20, 2014.
- [3] M. Ridwan, H. Suyono, and M. Sarosa, “Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier,” *J. EECCIS*, vol. 7, no. 1, p. pp.59-64, 2013.
- [4] Nurkholis, “PENDIDIKAN DALAM UPAYA MEMAJUKAN TEKNOLOGI Oleh: Nurkholis Doktor Ilmu Pendidikan, Alumnus Universitas Negeri Jakarta Dosen Luar Biasa Jurusan Tarbiyah STAIN Purwokerto,” vol. 1, no. 1, pp. 24–44, 2013.
- [5] Supardi U. S., “Arah Pendidikan Di Indonesia dalam Tataran Kebijakan dan Implementasi,” *J. Form. J. Ilm. Pendidik. MIPA*, vol. 2, no. 2, pp. 111–121, 2012.
- [6] K. F. Irranda, D. Hartama, and A. P. Windarto, “Analisa Klasifikasi C4.5 Terhadap Faktor Penyebab Menurunnya Prestasi Belajar Mahasiswa Pada Masa Pandemi,” *J. Media Inform. Budidarma*, vol. 5, no. 1, pp. 327–331, 2021, doi: 10.30865/mib.v5i1.2763.
- [7] P. B. N. Setio, D. R. S. Saputro, and Bowo Winarno, “Klasifikasi Dengan Pohon Keputusan Berbasis Algoritme C4.5,” *Prism. Pros. Semin. Nas. Mat.*, vol. 3, pp. 64–71, 2020.
- [8] S. Syarli and A. Muin, “Metode Naive Bayes Untuk Prediksi Kelulusan (Studi Kasus: Data Mahasiswa Baru Perguruan Tinggi),” *J. Ilm. Ilmu Komput.*, vol. 2, no. 1, pp. 22–26, 2016.
- [9] J. Sulaksono, R. H. Irawan, and I. N. Fahmi, “Penerapan Metode Naive bayes Terhadap Bantuan Sosial Keluarga PraSejahtera,” *Nusant. Eng.*, vol. 3, no. 2, pp. 52–61, 1945.
- [10] F. Y. Marianto and D. A. I. Maruddani, “3 1,2,3,” vol. 9, pp. 16–25, 2020.

