

Implementasi *Naïve Bayes* Pada Siswa Penerima Bantuan

^{1*}Dito Dwi Prasetyo, ²Rony Heri Irawan, ³Umi Mahdiyah

^{1,2,3} Teknik Informatika, Universitas Nusantara PGRI Kediri

E-mail: ^{1*}ditoprasetyov@gmail.com, ²rony@unpkediri.ac.id, ³umimahdiyah@gmail.com

Penulis Korespondens : Dito Dwi Prasetyo

Abstrak— Penyaluran bantuan sosial pendidikan untuk siswa tingkat sekolah dasar sederajat di Desa Paron oleh perangkat desa selaku pengambil keputusan masih dilakukan secara manual dan kriteria tidak diketahui, sehingga menimbulkan risiko bantuan menjadi tidak tepat sasaran. Penelitian ini bertujuan menerapkan algoritma *Naïve Bayes* untuk mengonstruksi model keputusan yang objektif dan transparan bagi perangkat desa dalam menentukan kelayakan penerima bantuan. Metode penelitian yang digunakan meliputi pendekatan *Knowledge Discovery in Databases* (KDD) dengan kriteria kemiskinan seperti jumlah saudara yang sekolah, pekerjaan ayah, pendidikan akhir ayah, pekerjaan ibu, pendidikan akhir ibu, penerima Program Indonesia Pintar (PIP), status orang tua dan kehadiran siswa. Hasil pengujian menunjukkan bahwa model mampu mengenali karakteristik kemiskinan lokal secara stabil dengan akurasi 77,78 persen. Implementasi algoritma ini berhasil mengklasifikasikan kelayakan penerima bantuan secara objektif sehingga dapat membantu perangkat desa dalam menyaring siswa dengan tepat sasaran.

Kata Kunci— Algoritma *Naïve Bayes*, Bantuan Sosial, Kelayakan, Siswa

Abstract— *The distribution of educational social assistance for elementary school-equivalent students in Paron Village by village officials as decision-makers is still conducted manually and with unknown criteria, creating a risk of mistargeted aid. This study aims to apply the Naïve Bayes algorithm to construct an objective and transparent decision model for village officials in determining the eligibility of assistance recipients. The research method incorporates the Knowledge Discovery in Databases (KDD) approach based on poverty criteria such as the number of siblings attending school, father's occupation, father's education level, mother's occupation, mother's education level, Program Indonesia Pintar (PIP) recipient status, parental status, and student attendance. The test results show that the model can stably recognize local poverty characteristics with an accuracy of 77.78%. The implementation of this algorithm successfully classifies recipient eligibility objectively, thereby assisting village officials in screening targeted students accurately.*

Keywords— *Eligibility, Naïve Bayes Algorithm, Social Assistance, Students*

This is an open access article under the CC BY-SA License.



I. PENDAHULUAN

Bantuan sosial pada bidang pendidikan digunakan pemerintah untuk membantu siswa yang berasal dari golongan prasejahtera agar tetap dapat menempuh pendidikan dasar [1]. Namun, realitanya menunjukkan bahwa pengelolaan penyaluran bantuan sering kali menghadapi kendala teknis maupun administratif pada tingkat otoritas lokal [2]. Berdasarkan kenyataan di Desa Paron, Kabupaten Kediri proses menentukan kelayakan siswa penerima bantuan untuk tingkat Sekolah

Dasar (SD) dan Madrasah Ibtidaiyah (MI) masih dilakukan secara manual dan kriteria untuk menentukannya tidak diketahui. Kondisi tersebut menunjukkan bahwa kesenjangan antara regulasi penyaluran bantuan dengan praktik di lapangan menyebabkan risiko bantuan menjadi tidak tepat pada sasaran. Oleh karena itu, diperlukannya sebuah algoritma untuk menghasilkan keputusan yang objektif dan transparan.

Untuk mengatasi permasalahan tersebut, berbagai penelitian terdahulu membuktikan keandalan algoritma yang bisa menyelesaikannya yaitu *Naïve Bayes*. [1] melakukan studi komparatif antara algoritma *Fuzzy C-Means* dan *Naïve Bayes* untuk menentukan Keluarga Penerima Manfaat (KPM), yang dimana *Naïve Bayes* menunjukkan performa lebih unggul dengan akurasi mencapai 74%. Lalu pada penelitian berikutnya, [2] juga menerapkan algoritma *Naïve Bayes Classifier* untuk menyaring penerima program bantuan sosial Keluarga Harapan (PKH) pada tingkat desa dan berhasil memperoleh hasil pengelompokan yang cukup dengan akurasi 80%. Penggunaan algoritma tersebut juga dikembangkan oleh [3] dalam klasifikasi kelayakan penerima beasiswa internal dengan mencapai akurasi 75,00%. Pada penelitian selanjutnya, [4] menganalisis penentuan penerima PKH dengan membandingkan algoritma *Naïve Bayes* dan *K-Nearest Neighbor* (KNN), dan menyimpulkan bahwa *Naïve Bayes* menghasilkan nilai akurasi yang tinggi. Algoritma juga diperkuat oleh [5] yang menerapkan algoritma *Naïve Bayes* dengan pendekatan *Knowledge Discovery in Database* (KDD) untuk memprediksi penerima beasiswa pendidikan dengan akurasi 77,11%.

Secara teoretis, algoritma *Naïve Bayes* merupakan metode klasifikasi berbasis probabilitas yang bersumber pada Teorema *Bayes* dengan asumsi independensi kuat antar atributnya [3]. Algoritma ini juga memiliki keunggulan pada kesederhanaan struktur komputasi serta efisiensi waktu dalam pemrosesan, namun tetap mampu mempertahankan tingkat akurasinya meskipun dihadapkan pada kumpulan data yang memiliki banyak kriteria [3]. Karakteristik hal ini menjadikan *Naïve Bayes* relevan untuk mentransformasikan data kriteria kemiskinan yang kompleks menjadi sebuah keputusan klasifikasi logis, transparan, serta mudah dalam memahaminya.

Meskipun beberapa penelitian sebelumnya mengeksplorasi penentuan kelayakan bantuan menggunakan algoritma C4.5 yang mampu memetakan sebuah keputusan berbasis atribut seperti pekerjaan dan tanggungan orang tua dengan akurasi berkisar 72,20% hingga 80,27% penerapan C4.5 tersebut juga belum dispesifikasikan pada jenjang sekolah dasar sederajat yang memiliki variasi kriteria lokal beragam [2]. Oleh karena penelitian ini mengimplementasikan algoritma *Naïve Bayes* pada tingkat desa, khususnya dalam menangani data siswa SD dan MI yang kurang mampu. Keterbatasan inilah yang mendasari pentingnya penelitian ini dilakukan guna mengisi kekosongan hal tersebut.

Berdasarkan permasalahan tersebut, solusi yang ditawarkan dalam penelitian ini adalah mengimplementasikan *naïve bayes* dengan judul "Implementasi *Naïve Bayes* Pada Siswa Penerima Bantuan". Tujuan dari penelitian ini adalah memperbanyak penerapan metode tersebut dalam lingkup bantuan sosial pendidikan pada tingkat desa. Secara spesifik, penelitian ini juga mengonstruksi sebuah model keputusan yang objektif, cepat, dan akurat yang berguna untuk membantu perangkat desa dalam menyaring serta menentukan kelayakan siswa penerima bantuan secara tepat sasaran. Metode penelitian yang digunakan meliputi empat tahapan utama

pengumpulan data siswa SD dan MI di Desa Paron, menentukan parameter kriteria kelayakan lokal, pemodelan menggunakan *Naïve Bayes*, serta pengujian performa sistem menggunakan confusion matrix untuk mengetahui tingkat keakuratan data yang digunakan.

II. METODE

Masalah yang dibahas selanjutnya adalah menguraikan materi, instrumen, dan metode komputasi yang diterapkan untuk menyelesaikan permasalahan klasifikasi kelayakan siswa penerima bantuan di Desa Paron. Pendekatan ini dirancang secara terstruktur guna memastikan seluruh tahapan pengolahan data dapat dihasilkan secara valid oleh penelitian di masa depan [6].

A. Tahapan dan Prosedur Penelitian

Pelaksanaan penelitian ini menggunakan pendekatan *Knowledge Discovery in Databases* (KDD) yang disesuaikan juga untuk analisis data pada tingkat lokal. Eksperimen terbagi ke dalam lima tahapan sebagai :

1. Pengumpulan Data : Sebelum pengumpulan data akan dilakukan obeservasi pada lingkungan Sekolah Dasar (SD), Madrasah Ibtidaiyah (MI) dan Kantor Desa Paron.
2. *Data Selection* : Memilih data siswa SD dan MI yang tercatat pada tahun 2023.
3. *Data Preprocessing* : Melakukan pembersihan data (*data cleaning*) untuk menangani nilai yang kosong (*missing values*) serta menghilangkan redudansi data agar dataset siap diolah, yang merujuk pada standarisasi pembersihan data teks dan kategorikal oleh [6].
4. *Data Transformation* : Mengubah format data mentah menjadi bentuk data kategorikal yang sesuai dengan kebutuhan komputasi probabilitas.
5. *Data Mining (Modeling)* : Menerapkan algoritma *Naïve Bayes Classifier* untuk membentuk pola keputusan berbasis probabilitas kelas.
6. Evaluation : Mengukur performa akurasi klasifikasi *Naïve Bayes* untuk menguji tingkat keandalan model.

Eksperimen komputasi dan pengujian model ini seluruhnya dieksekusi menggunakan perangkat lunak Visual Studio Code berbasis bahasa pemrograman Python dengan pustaka *Scikit-Learn*, serta memanfaatkan efisiensi lingkungan kerja terintegrasi seperti yang divalidasi oleh [7].

B. Subjek Penelitian

Subjek utama dalam penelitian ini adalah seluruh data sekunder siswa SD, MI, dan data bantuan pendidikan desa dari di Desa Paron. Pendekatan observasi langsung terhadap data desa ini mengadopsi validasi metodologi yang dilakukan oleh [8] dalam mengklasifikasikan kelayakan penerima bantuan sosial lokal.

C. Variabel Penelitian dan Desain Data

Variabel yang diukur dalam penelitian ini dikelompokkan menjadi dua kategori utama, yaitu atribut prediktor dan target. Atribut prediktor ditentukan berdasarkan indikator kemiskinan lokal di Desa Paron, sedangkan atribut target terdiri dari dua kelas keputusan, yaitu "Layak" dan "Tidak Layak". Struktur pengelompokan variabel dan tipe data yang digunakan sebagai berikut :

Tabel 1. Atribut Prediktor dan Target

Jenis Atribut	Nama Fitur	Keterangan Nilai Atribut
Atribut Prediktor	Jumlah Saudara yang Sekolah	Banyak, Sedang, Tidak Ada
	Pekerjaan Ayah	Nelayan, Petani, Peternak, PNS/TNI/POLRI, Karyawan Swasta, Pedagang Kecil, Pedagang Besar, Wiraswasta, Wirausaha, Buruh, Pensiunan, Tenaga Kerja Indonesia, Karyawan BUMN, Sudah Meninggal, Tidak Bekerja
	Pekerjaan Ibu	Nelayan, Petani, Peternak, PNS/TNI/POLRI, Karyawan Swasta, Pedagang Kecil, Pedagang Besar, Wiraswasta, Wirausaha, Buruh, Pensiunan, Tenaga Kerja Indonesia, Karyawan BUMN, Sudah Meninggal, Ibu Rumah Tangga
	Pendidikan Akhir Ayah	Rendah, Sedang, Tinggi
	Pendidikan Akhir Ibu	Rendah, Sedang, Tinggi
	Penerima Program Indonesia Pintar (PIP)	Ya, Tidak
Atribut Target	Status Orang Tua	Lengkap, Yatim, Piatu, Yatim Piatu
	Kehadiran Siswa	Tinggi, Sedang, Rendah
	Status	Layak, Tidak Layak

Seperti yang dijabarkan pada Tabel 1, semua fitur pada atribut prediktor memiliki tipe data kategorikal. Terdapat perbedaan keterangan nilai atribut pada nama fitur pekerjaan orang tua, di mana nilai atribut ‘Tidak Bekerja’ hanya berlaku untuk pekerjaan ayah, sedangkan nilai atribut ‘Ibu Rumah Tangga’ khusus digunakan untuk pekerjaan ibu. Selanjutnya, setiap data siswa akan dianalisis berdasarkan kombinasi dari delapan atribut prediktor seperti jumlah saudara yang sekolah, pekerjaan ayah, pendidikan akhir ayah, pekerjaan ibu, pendidikan akhir ibu, penerima Program Indonesia Pintar (PIP), status orang tua dan kehadiran siswa untuk memprediksi probabilitas pada atribut target layak atau tidak layak.

D. Model Statistika Algoritma *Naïve Bayes*

Model analisis statistika yang digunakan adalah algoritma *Naïve Bayes Classifier*. Algoritma ini memprediksi peluang di masa depan berdasarkan pengalaman data latih sebelumnya dengan asumsi setiap atribut prediktor bersifat independen satu dengan lainnya [8]. Persamaan dasar Teorema *Bayes* berpusat pada pencarian probabilitas posterior, yang ditulis secara terpusat pada Persamaan (1) :

$$P(H | X) = \frac{P(X | H) \cdot P(H)}{P(X)} \quad (1)$$

Di mana $P(H|X)$ merupakan probabilitas posterior dari hipotesis kelas H jika diberikan bukti atau atribut prediktor X . Karena asumsi independensi yang kuat, nilai probabilitas perkalian atribut $P(X|H)$ dihitung menggunakan rumus produktivitas bersyarat yang ditunjukkan pada Persamaan (2) :

$$P(H | X) = \prod_{i=1}^n P(x_i | H) \quad (2)$$

Berdasarkan nilai perhitungan dari Persamaan (1) dan Persamaan (2), kelas target dengan nilai probabilitas tertinggi (*Maximum A Posteriori*) akan dipilih oleh sistem sebagai hasil keputusan akhir klasifikasi layak atau tidak layak pada siswa.

E. Analisis Data dan Evaluasi Performa Model

Untuk menjamin hasil eksperimen terhindar dari bias akibat ketidakseimbangan distribusi data, skenario pembagian data dikondisikan menggunakan teknik pengujian *Stratified K-Fold Cross-Validation* [9]. Teknik ini membagi dataset menjadi K , secara proporsional memastikan data latih (*training data*) dan data uji (*testing data*) representasikan kondisi riil populasi secara seimbang [9]. Evaluasi performa akhir dari model klasifikasi *Naïve Bayes* diukur secara matematis menggunakan *Confusion Matrix*. Melalui matriks ini, nilai antara hasil prediksi sistem dan data aktual lapangan dapat dipetakan ke dalam empat parameter utama yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) [10]. Metrik pengujian yang diekstrak meliputi tingkat Akurasi, Presisi, *Recall* dan *F1-Score* untuk menentukan tingkat validitas dari model [10].

III. HASIL DAN PEMBAHASAN

Pembahasan ini merupakan temuan utama dari implementasi algoritma *Naïve Bayes Classifier* dalam menentukan kelayakan siswa penerima bantuan di Desa Paron. Analisis dilakukan secara kuantitatif melalui pengujian variasi rasio data dan secara kualitatif melalui interpretasi teoretis serta pembuktian kalkulasi manual.

A. Analisis Performa Model dan Pengujian Kuantitatif

Eksperimen klasifikasi dilakukan dengan memancangkan skenario pembagian data (*data splitting*) yang bervariasi pada dataset memiliki 50 data siswa. Berdasarkan pengujian yang telah dilakukan performa akurasi tertinggi secara signifikan ditemukan pada skenario rasio 82:18 dan 82:17. Ringkasan performa model pada beberapa rasio representatif disajikan pada Tabel 2.

Tabel 2. Hasil Performa Model

Rasio Data (Train : Test)	TP	FN	FP	TN	Akurasi	Presisi	Recall	F1-Score
81: 19	2	3	2	3	50%	50%	40%	44,44%
82 : 18	3	1	1	4	77,78%	75%	75%	75%

83 : 17	3	1	1	4	77,78%	75%	75%	75%
86 : 16	2	2	2	2	50%	50%	50%	50%
85 : 15	2	2	2	2	50%	50%	50%	50%
60 : 40	5	5	3	7	60%	62,5%	50%	55,56%
70 : 30	3	4	5	3	40%	37,5%	42,86%	40%
80 : 20	2	3	2	3	50%	50%	40%	44,44%
90 : 10	1	1	1	2	60%	50%	50%	50%

Data pada Tabel 2 menunjukkan bahwa penentuan *data train* dan *test* 82:18 dan 83:17 menghasilkan *confusion matrix* yang cukup. Dengan akurasi 77,78% hal ini membuktikan bahwa model mampu mengenali karakteristik kemiskinan lokal secara stabil jika diberikan proporsi data latih sebesar 82%. Temuan ini memperkuat hipotesis awal dan sejalan dengan penelitian [5] yang menguji prediksi beasiswa pendidikan dengan capaian akurasi 77,11%. Hasil ini juga memiliki kedekatan performa dengan studi [1] yang mencatatkan akurasi 74% dalam penentuan keluarga prasejahtera. Sebaliknya, ketika model diuji dengan metode *5-Fold Cross-Validation*, yang memiliki hasil akurasi 60%, presisi 51,5%, *recall* 56% dan *f1-score* 52,72%. Hal ini mengindikasikan adanya variasi data yang cukup tajam *fold*, yang menuntut representasi seimbang pada data latih yang berguna untuk menghindari fenomena salah klasifikasi.

B. Perhitungan Manual Algoritma *Naïve Bayes*

Untuk membuktikan validitas dan transparansi keputusan komputasi, simulasi perhitungan manual berdasarkan parameter probabilitas dari populasi dataset (50 baris data). Berikut hasil dari perhitungan peluang beryarat (*likelihood*) :

1) Menghitung Probabilitas Prior Kelas $P(H)$

$$P(\text{Layak}) : \frac{25}{50} = 0,50$$

$$P(\text{Tidak Layak}) : \frac{25}{50} = 0,50$$

2) Menghitung Probabilitas Bersyarat Atribut $P(X | H)$

- a. Jumlah Saudara yang Sekolah : Layak (Banyak=0,16, Sedang=0,56, Tidak Ada=0,28) dan Tidak Layak (Banyak=0,12, Sedang=0,60, Tidak Ada=0,28)
- b. Pekerjaan Ayah : Layak (Nelayan=0, Petani=0,16, Peternak=0, PNS/TNI/POLRI=0, Karyawan Swasta=0,28, Pedagang Kecil=0,12, Pedagang Besar=0, Wiraswasta=0,08, Wirausaha=0, Buruh=0,36, Pensiunan=0, Tenaga Kerja Indonesia=0, Karyawan BUMN=0, Sudah Meninggal=0, Tidak Bekerja=0) dan Tidak Layak (Nelayan=0,04, Petani=0,04, Peternak=0, PNS/TNI/POLRI=0,12, Karyawan Swasta=0,44, Pedagang Kecil=0,04, Pedagang Besar=0, Wiraswasta=0,12, Wirausaha=0, Buruh=0,20, Pensiunan=0, Tenaga Kerja Indonesia=0, Karyawan BUMN=0, Sudah Meninggal=0, Tidak Bekerja=0)
- c. Pendidikan Akhir Ayah : Layak (Rendah=0,36, Sedang=0,64, Tinggi=0) dan Tidak Layak (Rendah=0,32, Sedang=0,64, Tinggi=0,04)
- d. Pekerjaan Ibu : Layak (Nelayan=0, Petani=0, Peternak=0, PNS/TNI/POLRI=0, Karyawan Swasta=0,12, Pedagang Kecil=0,08, Pedagang Besar=0, Wiraswasta=0,08, Wirausaha=0, Buruh=0, Pensiunan=0, Tenaga Kerja Indonesia=0, Karyawan BUMN=0, Sudah Meninggal=0,04, Ibu Rumah Tangga=0,68) dan Tidak Layak (Nelayan=0, Petani=0, Peternak=0, PNS/TNI/POLRI=0,08, Karyawan Swasta=0,28,

- Pedagang Kecil=0,08, Pedagang Besar=0, Wiraswasta=0, Wirausaha=0, Buruh=0,08, Pensiunan=0, Tenaga Kerja Indonesia=0, Karyawan BUMN=0, Sudah Meninggal=0, Ibu Rumah Tangga=0,48)
- e. Pendidikan Akhir Ibu : Layak (Rendah=0,40, Sedang=0,52, Tinggi=0,08) dan Tidak Layak (Rendah=0,36, Sedang=0,56, Tinggi=0,08)
 - f. Penerima Program Indonesia Pintar (PIP) : Layak (Ya=0,04, Tidak=0,96) dan Tidak Layak (Ya=0,08, Tidak=0,92)
 - g. Status Orang Tua : Layak (Lengkap=0,96, Yatim=0, Piatu=0, Yatim Piatu=0,04) dan Tidak Layak (Lengkap=1, Yatim=0, Piatu=0, Yatim Piatu=0)
 - h. Kehadiran Siswa : Layak (Tinggi=0,88, Sedang=0,12, Rendah=0) dan Tidak Layak (Tinggi=0,96, Sedang=0,04, Rendah=0)

3) Simulasi Pengujian Data Siswa Baru (X)

Misalkan data 1 siswa baru (X) ingin diuji kelayakan dengan karakteristik seperti contoh berikut, tapi sebelum melakukan pengujian $P(\text{Layak})$ dan $P(\text{Tidak Layak})$ memiliki nilai sebesar 0,50 :

- a. Jumlah Saudara yang Sekolah = Sedang (Layak=0,50 dan Tidak Layak=0,60)
- b. Pekerjaan Ayah = Buruh (Layak=0,56 dan Tidak Layak=0,20)
- c. Pendidikan Akhir Ayah = Rendah (Layak=0,36 dan Tidak Layak=0,32)
- d. Pekerjaan Ibu = Ibu Rumah Tangga (Layak=0,36 dan Tidak Layak=0,48)
- e. Pendidikan Akhir Ibu = Sedang (Layak=0,68 dan Tidak Layak=0,56)
- f. Penerima PIP = Tidak (Layak=0,52 dan Tidak Layak=0,92)
- g. Status Orang Tua = Lengkap (Layak=0,96 dan Tidak Layak=1)
- h. Kehadiran = Tinggi (Layak=0,88 dan Tidak Layak=0,96)

Selanjutnya perhitungan kelas Layak :

$$\begin{aligned} \text{Skor (Layak)} &= 0,50 \times 0,56 \times 0,36 \times 0,36 \times 0,68 \times 0,52 \times 0,96 \times 0,96 \times 0,88 \\ &= 0,0104064 \end{aligned}$$

Lalu masukkan perhitungan kelas Tidak Layak :

$$\begin{aligned} \text{Skor (Tidak Layak)} &= 0,52 \times 0,60 \times 0,20 \times 0,32 \times 0,48 \times 0,56 \times 0,92 \times 1 \times 0,96 \\ &= 0,0045582 \end{aligned}$$

4) Keputusan Klasifikasi

Langkah terakhir yaitu mengambil nilai tertinggi dari kedua hasil perhitungan tersebut yang memiliki skor layak 0,0104064 dan skor tidak layak 0,0045582, dari hasil skor tersebut maka sistem mengambil keputusan pada 1 data siswa baru (X) dengan diklasifikasikan pada kelas layak menerima bantuan.

C. Pembahasan Teoretis dan Implikasi Praktis

Secara teoretis, lonjakan akurasi pada rasio 82:18 (77,78%) menggambarkan bahwa algoritma Naïve Bayes sangat bergantung pada kecukupan kuantitas sampel pembelajaran. Pada rasio yang lebih rendah (seperti 70:30), pemangkasan data latih mengakibatkan hilangnya informasi kriteria sensitif seperti status pekerjaan ayah/ibu atau penerima PIP, sehingga meningkatnya *False Positive*.

Hasil dari penelitian ini sangat krusial bagi perangkat di Kantor Desa Paron. Melalui akurasi model yang menyentuh angka 77,78%, proses verifikasi manual cenderung subjektif dapat digantikan oleh sistem dengan cukup objektif. Hal ini sesuai dengan pemikiran [2] menegaskan bahwa standardisasi digitalisasi kriteria kemiskinan di tingkat desa terbukti mampu menekan konflik sosial dari distribusi dana jaminan pendidikan yang tidak merata.

D. Batasan Penelitian dan Rekomendasi Masa Depan

Meskipun model *Naïve Bayes* menghasilkan akurasi yang memuaskan pada skenario 82:18 dan 83:17, batasan utama dari penelitian ini terletak pada ukuran volume dataset yang sangat sedikit 50 baris data. Keterbatasan kuantitas data ini rentan memunculkan masalah *zero probability* apabila di masa mendatang terdapat entri kriteria baru yang belum pernah dipelajari oleh sistem.

Penambahan data secara berkala menjadi sangat penting, untuk penelitian di masa depan, disarankan untuk mengintegrasikan sistem klasifikasi *Naïve Bayes* ini dengan teknik *Laplace Smoothing* secara permanen yang berguna untuk menangani nilai probabilitas nol, serta menggunakan website atau program agar yang dapat dioperasikan secara langsung oleh perangkat desa.

IV. KESIMPULAN

Penelitian ini berhasil dalam mengimplemtasikan algoritma *Naïve Bayes* untuk bantuan siswa dengan objektif, yang berguna untuk membantu perangkat desa dalam mengklasifikasi kelayakan penerima bantuan dengan tepat sasaran, penelitian ini terbukti dapat mengenali karakteristik prasejahtera pada tingkat lokal, namun sebagai saran untuk pengembangan di masa depan, sistem ini perlu diintegrasikan dengan teknik *Laplace Smoothing* secara permanen yang berguna untuk mengatasi risiko nilai probabilitas nol karena keterbatasan dari jumlah data, serta dapat juga dikembangkan ke website agar dapat dioperasikan dengan mudah.

DAFTAR PUSTAKA

- [1] P. S. Saputra, “Perbandingan Algoritma Fuzzy C-Means Dan Algoritma Naïve Bayes Dalam Menentukan Keluarga Penerima Manfaat (Kpm) Berdasarkan Status Sosial Ekonomi (Sse) Terendah,” *JST (Jurnal Sains dan Teknol.*, vol. 10, no. 1, pp. 1–8, 2021, doi: 10.23887/jstundiksha.v10i1.23340.
- [2] A. A. Saputro, “Sistem Pendukung Keputusan Penerimaan Bantuan Sosial Program Keluarga Harapan (PKH) dengan Menggunakan Metode Naïve Bayes Classifier (Studi Kasus di Balai Desa Bendungan Kraton Pasuruan),” *J. Ilm. Edutic Pendidik. dan Inform.*, vol. 9, no. 1, pp. 40–48, 2022, doi: 10.21107/edutic.v9i1.12232.
- [3] A. Nugraha, F. Al-ghifari, A. Panigoran, W. Desmawan, and L. Lisawita, “Implementasi Algoritma Naïve Bayes Classifier Untuk Klasifikasi Penerima Beasiswa BUMN Sekota Dumai,” *J. Karya Ilm. Multidisiplin*, vol. 4, no. 2, pp. 77–85, 2024, doi: 10.31849/jurkim.v4i2.20944.
- [4] S. Nuraeni, H. Harliana, and T. Prabowo, “Analisis Akurasi Naïve Bayes Dan Knn Dalam Penentuan Penerima Pkh Di Lombok Utara,” *J. Inf. Syst. Manag.*, vol. 5, no. 2, pp. 121–126, 2024, doi: 10.24076/joism.2024v5i2.1205.
- [5] E. Simamora *et al.*, “Implementasi Algoritma Naïve Bayes Dalam,” *J. FASILKOM (teknologi Inf. dan Ilmu KOMputer)*, vol. 9, no. 4, pp. 6655–6660,

- 2025, doi: <https://doi.org/10.37859/jf.v15i1.8959>.
- [6] M. Z. Haq, C. S. Octiva, A. Ayuliana, U. W. Nuryanto, and D. Suryadi, "Algoritma Naïve Bayes untuk Mengidentifikasi Hoaks di Media Sosial," *J. Minfo Polgan*, vol. 13, no. 1, pp. 1079–1084, 2024, doi: [10.33395/jmp.v13i1.13937](https://doi.org/10.33395/jmp.v13i1.13937).
- [7] Shilpa Mehta, "Playing Smart with Numbers: Predicting Student Graduation Using the Magic of Naive Bayes," *Int. Trans. Artif. Intell.*, vol. 2, no. 1, pp. 60–75, 2023, doi: <https://doi.org/10.33050/italic.v2i1.405>.
- [8] N. Riyanah and F. Fatmawati, "Penerapan Algoritma Naive Bayes Untuk Klasifikasi Penerima Bantuan Surat Keterangan Tidak Mampu," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 2, no. 4, pp. 206–213, 2021, doi: [10.35746/jtim.v2i4.117](https://doi.org/10.35746/jtim.v2i4.117).
- [9] S. Widodo, H. Brawijaya, and S. Samudi, "Stratified K-fold cross validation optimization on machine learning for prediction," *Sinkron*, vol. 7, no. 4, pp. 2407–2414, 2022, doi: [10.33395/sinkron.v7i4.11792](https://doi.org/10.33395/sinkron.v7i4.11792).
- [10] E. Helmud, E. Helmud, F. Fitriyani, and P. Romadiana, "Classification Comparison Performance of Supervised Machine Learning Random Forest and Decision Tree Algorithms Using Confusion Matrix," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 13, no. 1, pp. 92–97, 2024, doi: [10.32736/sisfokom.v13i1.1985](https://doi.org/10.32736/sisfokom.v13i1.1985).