

PENERAPAN ALGORITMA METODE NAÏVE BAYES UNTUK PENENTUAN PENERIMAAN BANTUAN PROGRAM INDONESIA PINTAR (PIP) (Studi Kasus SMP PGRI 1 CILACAP)

Imansyah Priyanto¹, Elsa Mayorita Dewanti², Tundo³, Muhammad Nurdin⁴,
Roy Kasiono⁵

^{1,2} Teknik Informatika, Universitas Nahdlatul Ulama Al Ghazali Cilacap

³ Teknik Informatika, Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika (STIKOM CKI)

⁴ BPPTL – Kementerian Perhubungan

⁵ Ketatalaksanaan Angkutan Laut dan Pelabuhan, Sekolah Tinggi Ilmu Pelayaran (STIP)

email : imansyah811@gmail.com, elsadewanti80@admin.smp.belajar.id, asna8mujahid@gmail.com,
muhnuraldidika@gmail.com, kasionoroy@gmail.com

Abstrak

Program Indonesia Pintar (PIP) melalui Kartu Indonesia Pintar (KIP) adalah program pemerintah yang ditawarkan dalam wujud pembiayaan pendidikan langsung kepada para siswa (6-21 tahun). KIP merupakan bagian penyempurnaan dari program Bantuan Siswa Miskin (BSM) sejak akhir 2014. Sasaran PIP di SMP PGRI 1 Cilacap masih kurang tepat sarasanya, dikarenakan kurangnya kriteria penerima KKS. Oleh karena itu penulis menambahkan kriteria penerima KKS dalam penelitian. Penelitian ini dibuat berdasarkan data yang telah ada sebelumnya yaitu dengan 100 data training dan 9 data uji menggunakan metode data mining *Naïve Bayes* dan dengan 6 atribut yaitu pekerjaan orang tua, jumlah tanggungan, penghasilan orang tua, penerima KIP, penerima KPS, penerima KKS. Hasil pengujian akurasi yang didapatkan yaitu 88,89% dan perhitungan Recall 85,71%.

Kata kunci: Data Mining, Naïve Bayes, Program Indonesia Pintar

Abstract

The Smart Indonesia Program (PIP) through the Smart Indonesia Card (KIP) is a government program offered in the form of direct education financing to students (6-21 years). KIP is an improvement part of the Poor Student Assistance (BSM) program since the end of 2014. The target of PIP at SMP PGRI 1 Cilacap is still not well targeted, due to the lack of criteria for KKS recipients. Therefore, the author added criteria for KKS recipients in the research. This research was created based on previously existing data, namely 100 training data and 9 test data using the Naïve Bayes data mining method and with 6 attributes, namely parents' occupation, number of dependents, parents' income, KIP recipients, KPS recipients, KKS recipients. The accuracy test results obtained were 88.89% and the Recall calculation was 85.71%.

Keywords: Data Mining, Naïve Bayes, Smart Indonesia Program

1 PENDAHULUAN

Sesuai dengan Undang-Undang yang ada di Indonesia tentang pentingnya sebuah pendidikan bagi warga Negara yang tertera pada pasal 31 Undang-Undang Dasar 1945 yang berisikan bahwa



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

<http://journal.stmikjayakarta.ac.id/index.php/JMIJayakarta>

DOI: <https://doi.org/10.52362/jmijayakarta.v4i2.1355>

setiap warga Negara berhak memperoleh dan mengikuti pendidikan dasar dan pemerintah wajib untuk membiayai. Pendidikan memiliki peranan sangat penting dalam meningkatkan sumber daya manusia. Pendidikan akan memajukan dan mengembangkan sikap, keterampilan, serta kecerdasan intelektual untuk memberikan manusia yang terampil, cerdas, dan berakhlak mulia. Namun pendidikan seringkali tidak berjalan dengan baik karena beberapa faktor penyebab putus sekolah, salah satu yang seringkali terjadi adalah faktor ekonomi yaitu dalam pendapatan pekerjaan orang tua.

Dalam menyikapi permasalahan tersebut, pemerintah meluncurkan program Indonesia Pintar. Pada tahun 2015 Kebijakan Program Indonesia Pintar (PIP) didirikan lewat Permendikbud No. 12 Tahun 2015 tanggal 12 Mei 2015. Program Indonesia Pintar adalah program pemerintah yang ditawarkan dalam wujud pembiayaan pendidikan langsung kepada para siswa (6-21 tahun). Program ini ditunjukkan kepada individu dari keluarga kurang mampu secara ekonomi yang ditandai dengan menggunakan Kartu Indonesia Pintar (KIP). KIP merupakan bagian penyempurnaan dari Program Bantuan Siswa Miskin (BSM) sejak akhir 2014. SMP PGRI 1 Cilacap bertempat di Jalan Rama No.22 Trajumas, Gumilir, Kecamatan Cilacap Utara, Kabupaten Cilacap, Jawa Tengah, SMP ini merupakan salah satu SMP yang berada dikawasan Gumilir yang mendapatkan dana Program Indonesia Pintar (PIP). Sasaran PIP di SMP PGRI 1 Cilacap masih kurang tepat saranya.

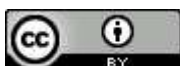
Menurut penelitian terdahulu oleh Tundo dan Shofwatul Uyun (2018) berjudul “Penerapan Decision Tree J48 dan Reptree Dalam Menentukan Prediksi Produksi Minyak Kelapa Sawit Menggunakan Metode *Fuzzy Tsukamoto*” kajian tersebut menggunakan *metode Fuzzy Tsukamoto* dengan kombinasi *decision tree* J48 dan *REPTree*, akurasi yang didapat adalah 95,2381% untuk *decision tree* J48 dan 90,4762% untuk *decision tree REPTree* [1], kemudian kajian oleh Okta Rin dan Suzi Oktavia Kunang (2021) berjudul “Implementasi Data Mining Menggunakan Metode *Naive Bayes* Untuk Penentuan Penerima Bantuan Program Indonesia Pintar (PIP) (Studi Kasus : Sd Negeri 9 Air Kumbang)” Berdasarkan uraian di atas, Implementasi data mining dengan prosedur *Naive Bayes* guna memprediksi kelayakan penerima PIP bersumber pada penambahan atribut dataset “Jumlah tanggungan” sangat berguna untuk memastikan kelayakan penerima PIP. Bersumber pada informasi yang diperoleh, penentuan penerima PIP memakai prosedur algoritma *Naive Bayes* memberikan informasi prognostik yang lebih akurat tentang siswa yang berhak menerima dukungan PIP ketimbang dengan penentuan yang dicoba oleh sekolah. Dengan demikian, metode *Naive Bayes* berhasil melakukan prediksi dengan presisi 90,00%, presisi 98,57%, *recall* 87,67%, dan AUC 0,868 menggunakan 101 data [2].

Sedangkan kajian oleh Putri Ayu Ani dan Andri (2021) yang berjudul “Penerapan Algoritma *Naive Bayes* untuk Klasifikasi Mahasiswa Penerima KIP (Studi Kasus : Universitas Bina Darma) Berdasarkan hasil penelitian dan pengujian tersebut dilakukan menggunakan pengklasifikasian algoritma *naive bayes* dengan teknik data *mining* CRISP-DM maka hasil akhir pengolahan data dengan menggunakan bantuan tools *Rapidminer* dari 342 data mahasiswa penerima KIP dengan nilai akurasi keseluruhan 49.15% dan *Class recall* Diterima 34.90%, *Class recall* tidak diterima 60.10%, *Class precision* diterima 40.31% dan *Class precision* tidak diterima 54.46% [3].

Berdasarkan hal-hal yang telah dijelaskan di atas, maka penelitian tertarik untuk melakukan penelitian tentang penentuan penerimaan Bantuan Program Indonesia Pintar (PIP) menggunakan Algoritma metode *Naive Bayes* dengan data yang diperoleh dari SMP PGRI 1 CILACAP menggunakan 6 atribut yaitu Pekerjaan Orang Tua, Jumlah Tanggungan, Penghasilan Orang Tua, Penerima KIP, Penerima KPS dan Penerima KKS. Dimana pada penelitian sebelumnya menggunakan 5 atribut untuk menentukan kelayakan penerima bantuan PIP, Berdasarkan penelitian saat ini, menggunakan 6 atribut dengan tambahan atribut penerima KKS (Kartu Keluarga Sehat) dengan maksud dan tujuan dalam pemberian kelayakan penerimaan bantuan PIP tepat sasaran.

2 METODE PENELITIAN

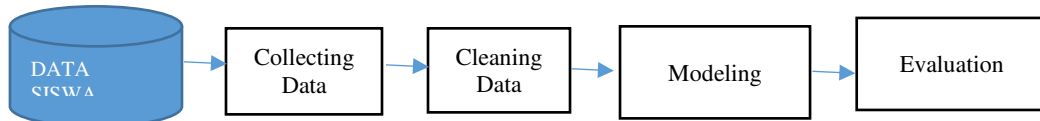
Pada penelitian ini menggunakan metode pemodelan algoritma dengan pendekatan *Knowledge Discovery in Database* (KDD) [4]. Tujuan dari metode pemodelan ini yaitu untuk menggali informasi



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).
<http://journal.stmikjayakarta.ac.id/index.php/JMIJayakarta>

DOI: <https://doi.org/10.52362/jmijayakarta.v4i2.1355>

serta pengetahuan yang belum diketahui sebelumnya dari *database* [5]. *Database* ini berisikan informasi yang disimpan dalam tabel yang saling berelasi. Adapun tahapan KDD bisa dilihat pada Gambar 1.



Gambar 1. Tahapan KDD

3.1. *Collecting Data*

Data yang telah dikumpulkan adalah data siswa SMP PGRI 1 CILACAP penerima PIP dengan atribut NPSN, Pekerjaan Orang Tua, Jumlah Tanggungan, Penghasilan Orang Tua, Penerima KIP, Penerima KPS, dan Penerima KKS. Data yang akan digunakan pada penelitian ini dikelompokkan menjadi dari 3 data yaitu :

1. Data Latih

Data training adalah data yang pada penelitian ini digunakan untuk proses pembuatan pemodelan prediksi penerima PIP SMP PGRI 1 Cilacap. Data Training (dataset) pada penelitian ini di ambil 100 data siswa.

2. Data Uji

Data Testing adalah data yang digunakan untuk melakukan uji pada pemodelan data yang terbentuk pada proses pemodelan. Data Uji pada penelitian ini di ambil 9 data siswa.

3. Data Prediksi

Data prediksi adalah data yang digunakan untuk melakukan prediksi penerima PIP SMP PGRI 1 Cilacap.

3.2. *Cleaning Data*

Data yang telah terkumpul selanjutnya dibersihkan dan beberapa diantaranya dihapus [6], proses ini meliputi data yang tidak sesuai, data yang berisikan salah satu atribut kosong, data anomali, dan membersihkan atribut yang tidak digunakan pada saat pemodelan.

3.3. *Modeling*

Setelah tahap pembersihan dan transformasi data selesai, selanjutnya dilakukan tahap pemodelan, tahap ini akan menentukan hasil klasifikasi dan prediksi [7],[8]. Pada penelitian ini menggunakan pemodelan algoritma naive bayes.

3.4 *Evaluation*

Hasil dari ujicoba pemodelan disajikan dalam bentuk *confusion matrix* atau matriks kesalahan. *Confusion matrix* ini menyajikan informasi yang aktual terhadap pemodelan yang telah dilakukan, *confusion matrix* ini juga memberikan informasi berupa hasil akurasi [9].

3 HASIL DAN PEMBAHASAN

3.1 *Collecting Data*

Pada tahap ini data yang dikumpulkan adalah *By Name By NIPD* dari penerima PIP dengan atribut Nama, NIPD, Jenis Kelamin, NISN, Tempat Lahir, Tanggal Lahir, NIK, Agama, Alamat, Jenis Tinggal, Alat Transportasi, Telepon / Hp, Penerima KPS, Data Orang Tua berupa Nama, Pekerjaan, Penghasilan, Nomer Seri Ijazah, Penerima KIP, Penerima KKS, Layak PIP (usulan dari sekolah), Sekolah Asal, Anak ke-berapa, Jumlah Saudara Kandung, Jarak Rumah ke Sekolah data tersebut adalah data awal yang diambil dari dapodik SMP PGRI 1 Cilacap. Berikut data yang di dapatkan dari SMP PGRI 1 Cilacap dilampirkan pada Tabel 1.

Tabel 1. Data Awal Siswa SMP PGRI 1 Cilacap

Nama	NIPD	JK	Layak PIP (usulan dari sekolah),	Sekolah Asal	Anak ke-berapa	Jumlah Saudara Kandung	Jarak Rumah ke Sekolah(K
------	------	----	-------	-------	----------------------------------	--------------	----------------	------------------------	--------------------------



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

<http://journal.stmikjayakarta.ac.id/index.php/JMIJayakarta>

DOI: <https://doi.org/10.52362/jmijayakarta.v4i2.1355>

									M)
Abib A	8631	L	Layak	SDN K 3 Kesugihan	4	3	1
Achmad W A	8731	L	Tidak	SDN G 4 Cilut	1	1	1
Adam G	8950	L	Layak	SDN Tritih 3	3	1	1
Ade D	8632	L	Layak	SDN K 3 Karangkan dri	3	1	1
:	:	:	:	:	:	:	:	:	:
Zian P	9194	L	Tidak	MI Nurul Huda	1	1	1
Zulfa N	9122	P	Layak	SDN Mertasinga 2	2	1	1

4.2 Cleaning Data

Setelah data dikumpulkan proses selanjutnya dilakukan pembersihan data, tahap ini menghapus data yang tidak digunakan sebagai data latih (dataset) seperti Jenis Kelamin, NISN, Tempat Lahir, Tanggal Lahir, NIK, Agama, Alamat, Jenis Tinggal, Alat Transportasi, Telepon / Hp, Nomer Seri Ijazah, Sekolah Asal, serta Jarak Rumah ke Sekolah. sehingga atribut yang digunakan untuk proses transformasi data yaitu Nama / NIPD, Pekerjaan Orang Tua, Jumlah Tanggungan, Penghasilan Orang Tua, Penerima KIP, Penerima KPS, dan Penerima KKS, Layak PIP (usulan dari sekolah). Berikut data yang sudah dilakukan proses Cleaning Data yang terlihat pada Tabel 2.

Tabel 2. Data Tabel setelah dilakukan Cleaning Data

No.	Nama Siswa	Pekerjaan Orang Tua	Jmlh Tanggungan	Penghasilan Orang Tua	Penerima KIP	Penerima KPS	Penerima KKS	PIP
1.	Silsi Irda P/8658	Buruh	2	Rp. 1000.000,00	Ya	Tidak	Tidak	Layak
2.	Candra R/8804	Nelayan	3	Rp. 500.000,00	Tidak	Tidak	Tidak	Layak
3.	Dinda K/8991	Buruh	3	Rp. 1000.000,00	Tidak	Tidak	Tidak	Tidak
4.	Muhammad Faiz S/8936	Nelayan	2	Rp. 500.000,00	Ya	Ya	Ya	Layak
5.	Wandika/8630	Buruh	2	Rp. 2000.000,00	Tidak	Tidak	Tidak	Tidak
6.	Rafif Ibnu/8817	Nelayan	1	Rp. 500.000,00	Ya	Ya	Ya	Layak
:	:	:	:	:	:	:	:	:
97.	Syifa Dina/8759	Buruh	1	Rp. 2.000.000	Tidak	Tidak	Tidak	Tidak
98.	Tian Fiki/8850	Nelayan	2	Rp. 2.000.000	Tidak	Tidak	Tidak	Tidak
99.	Uswatun Nur/8760	Buruh	2	Rp. 500.000	Tidak	Ya	Tidak	Layak
100.	Vania Belva/8626	Buruh	1	Rp. 2.000.000	Tidak	Tidak	Tidak	Tidak

4.3 Modeling Data

Modeling menggunakan metode NBC tahap pemodelan ini akan menghasilkan nilai akurasi dari data yang diolah. pada penelitian ini menggunakan pemodelan Algoritma Naive Bayes adapun Persamaan teorema Naive bayes pada rumus (1) [10],[11] :

$$P(A|B) = (P(B|A) * P(A))/P(B)$$

Artinya Peluang kejadian A sebagai B ditentukan dari peluang B saat A, peluang A, dan peluang B.

Data Uji yang dipilih agar cocok untuk langkah penambahan data serta penggunaan data diskrit dan continue untuk melakukan pengujian data. Setelah melakukan pembersihan data sebagai dataset (Data Latih) sebanyak 100 data siswa, serta Data Uji sebanyak 9 data siswa terlihat pada tabel 3.



DOI: <https://doi.org/10.52362/jmijayakarta.v4i2.1355>

Tabel 3. Data Tabel Uji sebanyak 9 data siswa

Nama Siswa / NIPD	Pekerjaan Orang Tua	Jmlh Tanggungan	Penghasilan Orang Tua	Penerima KIP (Kartu Indonesia Pintar)	Penerima KPS (Keluarga Pra Sejahtera)	Penerima KKS (Kartu Keluarga Sejahtera)	Layak PIP Atau Tidak ?
Ida Biora / 9245	Buruh	1	1.000.000	Ya	Tidak	Tidak	?
Juanita Alicya Putri / 8901	Petani	2	2.000.000	Tidak	Tidak	Tidak	?
Rangga Khaerul Dzakirin / 8876	Karyawan Swasta	1	2.000.000	Tidak	Tidak	Tidak	?
Maolana Nurhidayat / 9213	Buruh	2	500.000	Ya	Ya	Tidak	?
Putri Laura Dani / 8842	Nelayan	1	500.000	Ya	Ya	Tidak	?
Retno Gupito/ 8653	Pedagang Kecil	1	500.000	Ya	Ya	Ya	?
Riski Agung Sahputra / 8655	Buruh	3	500.000	Ya	Tidak	Ya	?
Zulfa Nur Ramadhani / 9122	Buruh	1	1.000.000	Tidak	Tidak	Tidak	?
Isbat Suyaki / 8682	Pedagang Kecil	2	500.000	Tidak	Ya	Tidak	?

4.3.1 Langkah 1 : Menghitung Probabilitas kemunculan setiap nilai untuk atribut Pekerjaan Orang Tua(C1) pada tabel 4.

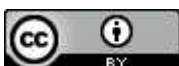
Tabel 4. Hasil Perhitungan atribut Pekerjaan Orang Tua (C1)

Pekerjaan Orang Tua	Probabilitas (C1)		Jumlah Kejadian "Dipilih"	
	Layak	Tidak	Layak	Tidak
Buruh	0,6	0,46	30	23
Nelayan	0,26	0,12	13	6
Petani	0,04	0,06	2	3
Pedagang Kecil	0,06	0,08	3	4
Wiraswasta	0,02	0,1	1	5
Karyawan Swasta	0,02	0,14	1	7
PNS/TNI/Polri	0	0,04	0	2

4.3.2 Langkah 2 : Menghitung Probabilitas kemunculan setiap nilai untuk atribut Jumlah Tanggungan(C2) pada tabel 5.

Tabel 5. Hasil Perhitungan atribut Jumlah Tanggungan (C2)

Jumlah Tanggungan (C2)	Layak	Tidak
	2	3
	3	2
	2	1
	1	1
	3	1
	1	1
	1	1



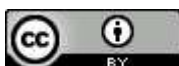
DOI: <https://doi.org/10.52362/jmijayakarta.v4i2.1355>

	2	1
	3	1
	3	3
	2	1
	2	3
	1	1
	5	1
	1	2
	1	4
	1	1
	2	1
	2	1
	2	2
	1	3
	1	1
	1	1
	1	3
	1	3
	1	2
	4	2
	6	1
	1	1
	1	1
	1	1
	1	1
	1	1
	1	1
	1	1
	2	1
	1	1
	1	1
	1	1
	2	1
	3	1
	1	3
	1	1
	1	1
	1	1
	1	1
	1	2
	1	3
	1	1
	3	2
	1	1
Mean (μ)	2,2	1,6
DS (Deviasi Standar)	0,83666	0,894427

Rumus 2 :

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$$



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).
<http://journal.stmikjayakarta.ac.id/index.php/JMIJayakarta>

DOI: <https://doi.org/10.52362/jmijayakarta.v4i2.1355>

$\mu = \bar{x} = \text{Mean}$

Diketahui : $e = 2,7183$

$s = \sigma = \text{Standar Deviasi}$

4.3.3 Langkah 3 : Menghitung Probabilitas kemunculan setiap nilai untuk atribut Penghasilan Orang Tua(C3) pada tabel 6.

Tabel 6. Hasil Perhitungan atribut Penghasilan Orang Tua (C3)

	Layak	Tidak
	1.000.000	1.000.000
	500.000	2.000.000
	500.000	1.000.000
	500.000	1.000.000
	1.000.000	1.000.000
	500.000	2.000.000
	1.000.000	1.000.000
	1.000.000	1.000.000
	500.000	1.000.000
	500.000	2.000.000
	500.000	2.000.000
	1.000.000	2.000.000
	1.000.000	2.000.000
	1.000.000	1.000.000
	1.000.000	4.000.000
	1.000.000	3.000.000
	2.000.000	2.000.000
	500.000	2.000.000
	2.000.000	2.000.000
	2.000.000	4.000.000
	1.000.000	4.000.000
	2.000.000	2.000.000
	2.000.000	2.000.000
	1.000.000	2.000.000
	500.000	2.000.000
	2.000.000	1.000.000
	1.000.000	2.000.000
	2.000.000	2.000.000
	500.000	2.000.000
	2.000.000	1.000.000
	1.000.000	3.000.000
	1.000.000	2.000.000
	1.000.000	2.000.000
	1.000.000	4.000.000
	2.000.000	4.000.000
	500.000	2.000.000
	1.000.000	3.000.000
	500.000	4.000.000
	1.000.000	1.000.000
	500.000	2.000.000
Penghasilan Orang Tua (C3)	500.000	4.000.000



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).
<http://journal.stmikjayakarta.ac.id/index.php/JMIJayakarta>

DOI: <https://doi.org/10.52362/jmijayakarta.v4i2.1355>

	500.000	2.000.000
	500.000	1.000.000
	500.000	1.000.000
	500.000	2.000.000
	500.000	4.000.000
	500.000	2.000.000
	1.000.000	2.000.000
	1.000.000	2.000.000
	500.000	2.000.000
Mean (μ)	700000	1200000
DS (Deviasi Standar)	273861,2788	447213,5955

- 4.3.4 Langkah 4 : Menghitung Probabilitas kemunculan setiap nilai untuk atribut Penerima KIP(C4) pada tabel 7.

Tabel 7. Hasil Perhitungan atribut Penerima KIP (C4)

Penerima KIP	Probabilitas (C4)		Jumlah Kejadian "Dipilih"	
	Layak	Tidak	Layak	Tidak
Ya	0,72	0	36	0
Tidak	0,28	1	14	50

- 4.3.5 Langkah 5 : Menghitung Probabilitas kemunculan setiap nilai untuk atribut Penerima KPS(C5) pada tabel 8.

Tabel 8. Hasil Perhitungan atribut Penerima KPS (C5)

Penerima KPS	Probabilitas (C5)		Jumlah Kejadian "Dipilih"	
	Layak	Tidak	Layak	Tidak
Ya	0,52	0	26	0
Tidak	0,48	1	24	50

- 4.3.6 Langkah 6 : Menghitung Probabilitas kemunculan setiap nilai untuk atribut Penerima KKS(C6) pada tabel 9.

Tabel 9. Hasil Perhitungan atribut Penerima KKS (C6)

Penerima KKS	Probabilitas (C6)		Jumlah Kejadian "Dipilih"	
	Layak	Tidak	Layak	Tidak
Ya	0,36	0	18	0
Tidak	0,64	1	32	50

- 4.3.7 Langkah 7 : Menghitung Probabilitas kemunculan setiap nilai untuk atribut Dipilih sebagai Layak PIP(C7) pada tabel 10.

Tabel 10. Hasil Perhitungan atribut Dipilih sebagai Layak PIP (C7)

	Layak	Tidak
P(Layak / Tidak)	0,5	0,5
Total Data	50	50



DOI: <https://doi.org/10.52362/jmijayakarta.v4i2.1355>

4.3.8 Langkah 8 : Menghitung Class Prediksi Layak PIP atau Tidak dari 9 data Uji pada tabel 11.

Tabel 11. Hasil Perhitungan Class Prediksi Layak PIP atau Tidak

Class Prediksi	Layak	Tidak
Layak	87985,45334	0
Tidak	788,349662	24000
Tidak	4582,575695	56000
Layak	32941,75373	0
Layak	14274,75995	0
Layak	1852,973647	0
Layak	17104,37213	0
Tidak	11825,24493	184000
Layak	1281,068201	0

4.3.9 Langkah 9 : Membandingkan Hasil Class Prediksi dengan data yang sudah ada dari instansi SMP PGRI 1 Cilacap, disajikan dalam bentuk tabel 12.

Tabel 12. Hasil Perbandingan Class Prediksi dengan data Instansi

Layak PIP Atau Tidak ? (Instansi)	Class Prediksi
Layak	Layak
Tidak	Tidak
Tidak	Tidak
Layak	Layak
Layak	Layak
Layak	Layak
Layak	Layak
Layak	Tidak
Layak	Layak

4.3.10 Contoh perhitungan Class Prediksi untuk menentukan Layak atau Tidak Sebagai Penerima PIP yaitu dengan perhitungan $C1 \cdot C2 \cdot C3 \cdot C4 \cdot C5 \cdot C6$ dengan diketahui sesuai tabel 13.

Tabel 13. Menghitung Layak atau Tidak untuk Class Prediksi

Nama Siswa / NIPD	Pekerjaan Orang Tua	Jmlh Tanggungan	Penghasilan Orang Tua	Penerima KIP	Penerima KPS	Penerima KKS	Layak PIP Atau Tidak ?
IDA BIORA / 9245	Buruh	1	1.000.000	Ya	Tidak	Tidak	?

L. Class Prediksi Layak : $0,6 \times 0,83666 \times 273861,2788 \times 0,72 \times 0,48 \times 0,64 = 87985,45334$

L. Class Prediksi Tidak : $0,46 \times 0,894427 \times 447213,5955 \times 0 \times 1 \times 1 = 0$

Jadi, dapat diperoleh nilai probabilitas Class Prediksi dapat dihitung dengan melakukan normalisasi terhadap likelihood tersebut sehingga jumlah nilai yang diperoleh = 1. Perbandingan sampel perhitungan untuk menentukan Layak atau Tidak pada Class Prediksi yaitu Layak.



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

<http://journal.stmikjayakarta.ac.id/index.php/JMIJayakarta>

3.4 Evaluation

Untuk mencari nilai akurasi dan *recall* pada pemodelan yang telah dilakukan terhadap data uji, langkah selanjutnya adalah melakukan pengujian terhadap data uji yang telah dipisahkan sebelumnya, nilai akurasi dan nilai *recall* disajikan dalam bentuk tabel 14.

Confusion Matrix.

$$\begin{aligned} \text{Rumus 3 : Akurasi} &= \frac{TP+TN}{TP+TN+FP+FN} \\ &= \frac{6 + 2}{6 + 2 + 0 + 1} \\ &= \frac{8}{9} = 0,8889 \times 100\% = \mathbf{88,89\%} \end{aligned}$$

Tabel 14. *Confusion Matrix*

Akurasi	88,89%		<i>confusion matrix</i>
	<i>class label</i>		
Prediksi	Layak	Tidak	
Layak	6	1	
Tidak	0	2	

$$\begin{aligned} \text{Rumus 4 : Perhitungan Recall} &= \frac{TP}{TP+FN} \\ &= \frac{6}{6 + 1} = \frac{6}{7} = 0,8571 \times 100\% = \mathbf{85,71\%} \end{aligned}$$

TP = True Positive
TN = True Negative
FP = False Positive
FN = False Negative

4 KESIMPULAN

Berdasarkan penelitian yang dilakukan, dapat ditarik kesimpulan bahwa, penerapan Data Mining menggunakan Metode Naive Bayes untuk memprediksi kelayakan penerima bantuan PIP berdasarkan dataset dengan menambahkan atribut Penerima KKS sangat membantu dalam menentukan kelayakan penerima bantuan PIP.

Berdasarkan data yang telah diperoleh, proses penentuan penerima bantuan PIP menggunakan metode Naive Bayes menghasilkan informasi prediksi siswa yang layak mendapatkan bantuan PIP lebih akurat dibandingkan dengan penentuan yang dilakukan pihak sekolah. Dengan demikian metode Naive Bayes ini berhasil memprediksi dengan presentase *Accuracy* sebesar 88,89% dan perhitungan *Recall* 85,71% dengan menggunakan data sebanyak 100 data siswa sebagai Data uji sebanyak 9 Data Siswa.

DAFTAR PUSTAKA

- [1] T. Tundo and S. Uyun, "Penerapan Decision Tree J48 Dan Reptree Dalam Menentukan Prediksi Produksi Minyak Kelapa Sawit Menggunakan Metode Fuzzy Tsukamoto," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 3, pp. 483–492, 2020.
- [2] O. Rini and S. O. Kunang, "Implementasi Data Mining Menggunakan Metode Naive Bayes Untuk Penentuan Penerima Bantuan Program Indonesia Pintar (Pip) (Studi Kasus : Sd Negeri 9 Air Kumbang)," *Bina Darma Conf.*, pp. 714–722, 2021.
- [3] P. A. Ani and A. Andri, "Penerapan Algoritma Naive Bayes Untuk Klasifikasi Mahasiswa Penerima Kip Pada Universitas Bina Darma," *Bina Darma Conf. Comput. Sci.*, pp. 172–180, 2022.
- [4] R. Rachman and R. N. Handayani, "Klasifikasi Algoritma Naive Bayes Dalam Memprediksi Tingkat Kelancaran Pembayaran Sewa Teras UMKM," *J. Inform.*, vol. 8, no. 2, pp. 111–122, 2021.
- [5] N. A. Khairudin *et al.*, "Implementasi Sistem Informasi Pengelolaan Bantuan Warga Kampung



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).
<http://journal.stmikjayakarta.ac.id/index.php/JMIJayakarta>

DOI: <https://doi.org/10.52362/jmijayakarta.v4i2.1355>

- Pulojahe Jakarta Timur Berbasis Web,” *SELAPARANG J. Pengabd. Masy. Berkemajuan*, vol. 6, no. 1, p. 355, 2022.
- [6] A. Pebdika, R. Herdiana, and D. Solihudin, “Klasifikasi Menggunakan Metode Naive Bayes Untuk Menentukan Calon Penerima Pip,” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 452–458, 2023.
- [7] M. H. Rifqo and A. Wijaya, “Implementasi Algoritma Naive Bayes Dalam Penentuan Pemberian Kredit,” *Pseudocode*, vol. 4, no. 2, pp. 120–128, 2017.
- [8] T. Tundo and F. Mahardika, “Fuzzy Inference System Tsukamoto – Decision Tree C 4 . 5 in Predicting the Amount of Roof Tile Production in Kebumen,” *JTAM (Jurnal Teor. dan Apl. Mat.*, vol. 7, no. 2, pp. 533–544, 2023.
- [9] T. Tundo and S. 'Uyun, “Konsep Decision Tree Reptree Untuk Melakukan Optimasi Rule Dalam Fuzzy Inference System Tsukamoto,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 9, no. 3, 2022.
- [10] Heliyanti Susana, “Penerapan Model Klasifikasi Metode Naive Bayes Terhadap Penggunaan Akses Internet,” *J. Ris. Sist. Inf. dan Teknol. Inf.*, vol. 4, no. 1, pp. 1–8, 2022.
- [11] D. Ariadi and K. Fithriasari, “Indonesian News Classification Using Naive Bayesian Classification Method and Support Vector Machine With Confix Stripping Stemmer,” *J. Sains dan Seni ITS*, vol. 4, no. 2, pp. 2337–3520, 2015.

