

## KOMPARASI ALGORITMA C4.5 DAN NAÏVE BAYES PADA KLASIFIKASI REKOMENDASI PENERIMA BANTUAN PIP

Rahmat Hidayat<sup>1</sup>, Miftahul Huda<sup>2</sup>, Nanang Pradita<sup>3</sup>, Febri Laksana<sup>4</sup>

<sup>1,2,3,4</sup>Universitas Putra Bangsa

[ayat150190@gmail.com](mailto:ayat150190@gmail.com)<sup>1</sup>, [febrilaksana10@gmail.com](mailto:febrilaksana10@gmail.com)<sup>4</sup>

*Received 2 November 2023; revised 24 November 2023; accepted 15 December 2023.*

### ABSTRAK

Program Indonesia Pintar (PIP) merupakan program bantuan pemerintah yang ditujukan untuk peserta didik yang berasal dari keluarga miskin atau rentan miskin untuk membiayai pendidikan. Siswa penerima PIP terdiri dari dua kategori yaitu, siswa yang memiliki Kartu Indonesia Pintar (KIP) / Kartu Keluarga Sejahtera (KIS) dan siswa yang direkomendasikan oleh pihak sekolah untuk mendapatkan beasiswa. Cara konvensional yang dilakukan sekolah dalam memberikan rekomendasi beasiswa membutuhkan waktu yang lama serta rentan akan ketidaksesuaian rekomendasi sehingga dapat mengakibatkan penyaluran bantuan PIP yang tidak tepat sasaran. Berdasarkan permasalahan tersebut penulis tertarik melakukan penelitian untuk menyelesaikan masalah yang ada dengan menggunakan teknik data mining. Penelitian ini membandingkan algoritma C4.5 dan algoritma Naive Bayes yang meliputi akurasi, presisi dan recall. Pengujian dilakukan pada 139 data siswa SDN 03 Karanganyar. Hasil penelitian ini menunjukkan kedua algoritma dapat diterapkan dengan baik, berdasarkan hasil penghitungan manual dan software rapidminer. Hasil dari pengujian ini dilakukan lagi komparasi dari kedua algoritma berdasarkan akurasi, presisi dan recall dimana algoritma C4.5 lebih baik pada pengujian akurasi sebesar 92%, presisi sebesar 94% dan recall sebesar 85% lebih tinggi dibandingkan algoritma naïve bayes. Sehingga Rule yang dihasilkan oleh algoritma C4.5 ini dapat digunakan untuk membangun suatu sistem pendukung pengambil keputusan pada SDN 03 Karanganyar.

**Kata kunci:** data mining, klasifikasi, C4.5, naïve bayes, PIP

### ABSTRACT

The Smart Indonesia Program (PIP) is a government assistance program aimed at students who come from poor or vulnerable families to finance education. Students receiving PIP consist of two categories, namely, students who have an Indonesia Smart Card (KIP) / Prosperous Family Card (KIS) and students who are recommended by the school to get a scholarship. The conventional way carried out by schools in providing scholarship recommendations takes a long time and is prone to inappropriate recommendations so that it can result in the distribution of PIP assistance that is not on target. Based on these problems, the authors are interested in conducting research to solve existing

problems using data mining techniques. This research compares the C4.5 algorithm and the Naive Bayes algorithm which includes accuracy, precision and recall. Testing was carried out on 139 SDN 03 Karanganyar student data. The results of this study show that both algorithms can be applied well, based on the results of manual calculations and rapidminer software. The results of this test are carried out again comparing the two algorithms based on accuracy, precision and recall where the C4.5 algorithm is better at testing accuracy by 92%, precision by 94% and recall by 85% higher than the naïve bayes algorithm. So that the Rule generated by the C4.5 algorithm can be used to build a decision-making support system for SDN 03 Karanganyar.

**Keywords:** data mining, classification, C.45, naïve bayes, PIP

## **PENDAHULUAN**

Pendidikan merupakan hal yang wajib dimiliki oleh setiap warga negara Indonesia. Mencerdaskan kehidupan bangsa merupakan hal yang diamanatkan dalam pembukaan UUD 1945 juga berdasarkan Pasal 31 ayat 1 UUD 1945 yang menyatakan bahwa “Setiap warga negara berhak mendapat pendidikan”. Dan haruslah menjadi prioritas dalam hal pendidikan dengan sekurang-kurangnya 20 persen dari APBN dan APBD untuk pendidikan menurut UUD 1945 Pasal 31 ayat 4 (Nugraha & UUD, 1945). Meskipun fakta yang didapat saat ini, masih sulitnya bagi pemerintah dalam menyeimbangkan kewajiban pemerataan pendidikan diseluruh Indonesia hingga dapat meningkatkan mutu pendidikan (Sianipar & Maulia, 2023). Hingga saat ini angka putus sekolah jenjang SD,SMP,SMA,dan SMK Negeri dan Swasta pada tahun 2021 saja di Indonesia mencapai 83,7 ribu (Desa et al., 2023). Dan angka putus sekolah tertinggi terjadi tahun 2017 mencapai hingga 187.828 (Octavianti, 2023).

Ekonomi merupakan salah satu faktor penyebab ketidak merataannya pendidikan yang terjadi di Indonesia yang tentunya mempengaruhi rendahnya angka partisipasi masyarakat dalam pendidikan (Nurokhmah, 2021). Mahalnya biaya pendidikan saat ini mengakibatkan ketidak mampuan masyarakat dengan ekonomi rendah untuk mengakses pendidikan sehingga terjadinya kesenjangan (Rohaeni & Saryono, 2018). Salah satu bentuk upaya yang dilakukan pemerintah dalam membuat jaminan akses pendidikan dengan memberikan bantuan biaya pendidikan yang ditujukan kepada anak dengan keluarga ekonomi rendah adalah dengan adanya Program Indonesi Pintar (PIP). Dengan tujuan yaitu meminimalisir

atau memperkecil kemungkinan putus sekolah akibat kesulitan biaya pendidikan sehingga dapat melanjutkan pada jenjang pendidikan yang lebih tinggi (Dahlia & Jamaludin, 2019) dan setidaknya meningkatkan partisipasi sekolah anak berusia (6-21 tahun) untuk mendapatkan pendidikan wajar 12 tahun.

Program Indonesia Pintar (PIP) merupakan bantuan yang dikeluarkan oleh Kemendikbud berupa uang tunai, perluasan akses, atau kesempatan belajar yang diberikan kepada peserta didik dari keluarga miskin dengan tujuan untuk membantu mendapatkan pendidikan. Kemendikbud akan melakukan pencocokan data siswa yang ada di DTKS yang dikelola oleh dinas sosial dengan dapodik untuk menetapkan kelayakan penerima PIP. Sekolah mengelola menginputkan data siswa pada dapodik sesuai dengan kriteria penerima PIP (Arminy et al., 2023). Kriteria yang menjadi pertimbangan rekomendasi meliputi beberapa aspek yaitu kondisi rumah, prestasi, nilai rata-rata raport, penghasilan orang tua, dan tanggungan. Hal ini mengapa penerima PIP tidak hanya siswa yang memiliki KIP namun juga siswa yang mendapatkan rekomendasi dari sekolah berdasarkan data yang terdapat pada dapodik berdasarkan kriteria penerima PIP.

Sekolah mendaftarkan siswa pada dapodik dalam melakukan penyeleksian penerima PIP masih menggunakan cara konvensional sehingga membutuhkan waktu yang lama. Serta kemungkinan terjadi ketidak sesuaian akan rekomendasi dengan kriteria penerima PIP yang dapat mengakibatkan penyaluran program PIP yang tidak tepat sasaran (Lidia Lusiana, 2017), bahwasanya data yang ditemukan oleh BPK pada tahun anggaran 2018-2020 Kemendikbud, pelaksanaan program PIP belum memadai, hal ini dikarenakan data yang digunakan sebagai sumber pengusulan calon penerima tidak handal. Hal ini mengakibatkan penyaluran bantuan untuk PIP belum tepat sasaran dan masih banyak anak yang seharusnya mendapatkan bantuan justru tidak menerima (Edrial et al., 2022).

Permasalahan serupa juga masih terjadi pada SDN 03 Karanganyar penentuan rekomendasi yang dilakukan pada jumlah siswa yang cukup banyak menjadi beban pekerjaan yang cukup memberatkan kepada sekolah. Dengan jumlah siswa sebanyak 139 dengan kondisi dan keadaan yang berbeda tentunya menentukan rekomendasi menjadi sulit. Namun dengan data tersebut maka dapat dimanfaatkan suatu teknik yaitu data mining. Dimana Teknik data mining itu

sendiri yaitu proses menemukan pengetahuan atau informasi yang tersembunyi dari suatu data (Parapat & Sinaga, 2018). Berdasarkan permasalahan yang terjadi teknik klasifikasi adalah yang paling tepat dalam melakukan penyeleksian penerima bantuan PIP.

Pemanfaatan teknik data mining kerap digunakan dengan tujuan clusterisasi, prediksi, klasifikasi (Primajaya et al., 2020). Sebagai dasar proses analisis data, klasifikasi merupakan teknik yang dapat digunakan untuk mendapatkan informasi baru yang berguna dari suatu data dengan terbentuknya suatu rule atau aturan berdasarkan pengidentifikasian yang dilakukan terhadap obyek ke dalam kelas, kelompok, atau kategori yang didasari prosedur, karakteristik dan definisi (Jollyta et al., 2019). Data siswa pada SDN 03 Karanganyar tentunya memiliki karakteristik sehingga dapat dilakukan pengidentifikasian terhadap kelompok atau kelas berdasarkan kedekatan atribut dari data tersebut. Algoritma pada metode klasifikasi yang populer saat ini yaitu algoritma C.45 dan naive bayes (Kurniawan, 2018).

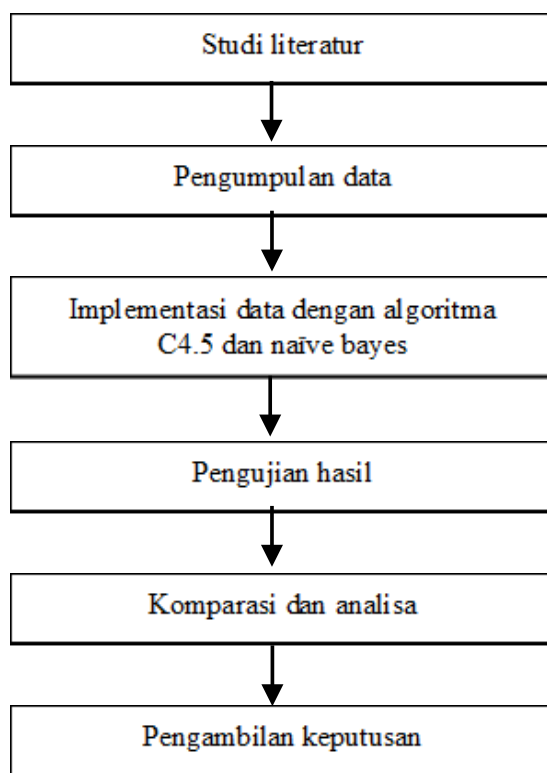
Metode klasifikasi pada data mining dapat diimplementasikan pada berbagai data, dengan algoritma C4.5 terdapat pohon keputusan serta dapat mengolah data numerik dan diskrit sehingga menghasilkan rule atau aturan yang cepat dan mudah diinterpretasikan (Haqmanullah Pambudi et al., 2018) seperti halnya penelitian yang dilakukan oleh (Astuti et al., 2023) yang mendapatkan nilai akurasi tinggi hingga 100% pada pengklasifikasian dan pengelompokan data penjualan kuota internet menggunakan algoritma C4.5 pada tahun 2021-2022. Sedangkan algoritma naïve bayes memberikan cara yang praktis dengan kombinasi data dan mudah diaplikasikan (Wicaksono et al., 2022). Pada penelitian terbaru yang dilakukan oleh (Alfitri & Rusda, 2023) menggunakan algoritma naïve bayes klasifikasi yang dihasilkan memiliki akurasi cukup tinggi sebesar 89% dalam melakukan seleksi terhadap data penerima bantuan pangan non tunai.

Sehingga penelitian ini bertujuan untuk membandingkan penggunaan algoritma C4.5 dan naive bayes berdasarkan akurasi, presisi dan recall dalam melakukan klasifikasi data rekomendasi penerima bantuan PIP. Perbandingan yang dilakukan pada kedua algoritma tersebut berguna untuk mengetahui algoritma

terbaik yang dapat digunakan dalam melakukan klasifikasi rekomendasi data siswa penerima bantuan PIP serta sebagai aspek pendukung dalam pengambilan keputusan.

## **METODE PENELITIAN**

Penelitian ini dilakukan pada 139 data siswa SDN 03 Karanganyar yang akan dilakukan pengolahan data menggunakan metode data mining dengan teknik klasifikasi menerapkan algoritma c4.5 dan naïve bayes. Berikut kerangka kerja dalam penelitian yang dilakukan.



Gambar 1. Kerangka Kerja Penelitian

Langkah kerja yang dilakukan dapat dijelaskan sebagai berikut :

### **2.1. Studi literatur**

Untuk mencapai tujuan maka pada tahap ini dilakukan dengan pencarian literatur yang dapat digunakan mengenai data mining, klasifikasi, algoritma C4.5, algoritma naïve bayes yang diambil dari berbagai bahan bacaan yang

dapat mendukung penelitian. Pencarian literatur berdasarkan pada penelitian - penelitian terdahulu hingga perkembangannya serta penelitian yang terbaru.

## 2.2. Pengumpulan data

Data yang digunakan menggunakan data siswa yang terdapat pada SDN 03 Karanganyar. Data yang didaftarkan kedalam dapodik sebagai rekomendasi penerima bantuan PIP berupa data prestasi, nilai rata-rata raport, penghasilan orang tua, tanggungan, dan kondisi rumah. Pihak sekolah berperan penting dalam memasukan data siswa rekomendasi penerima bantuan PIP kedalam dapodik (Armyny et al., 2023). Seluruh data siswa tersebut digunakan pada penelitian ini.

## 2.3. Implentasi data dengan algoritma C4.5 dan naïve bayes

Pengimplementasian digunakan untuk mendapatkan hasil perhitungan yang berupa rekomendasi penerima bantuan PIP. Pada penelitian ini penghitungan juga dilakukan secara manual.

### 2.3.1. Algoritma C4.5

Algoritma C4.5 merupakan pengembangan dari algoritma ID3 yang ditemukan oleh Ross Quinlan, Algoritma C4.5 digunakan untuk mengklasifikasikan data yang memiliki atribut berupa numerik ataupun kategorial. Algoritma C4.5 adalah satu algoritma untuk klasifikasi penggolongan data. Dalam algoritma C4.5 sebelum membangun pohon keputusan hal yang paling perlu dilakukan adalah menentukan atribut sebagai akar (Parapat & Sinaga, 2018) . Dimana aturan model atau rule dan pohon keputusan yang dibentuk dalam algoritma C4.5 bersifat mudah diinterpretasikan serta diubah (Anggraini et al., 2018).

Pada algoritma C4.5 ada beberapa tahapan dalam membuat pohon keputusan yaitu :

1. Mempersiapkan data training. Data training biasanya diambil dari data histori yang pernah terjadi sebelumnya atau disebut data masa lalu dan sudah dikelompokkan dalam kelas-kelas tertentu.
2. Menghitung akar dari pohon. Akar akan diambil dari attribute yang akan terpilih, dengan cara menghitung nilai gain dari masing-masing attribute, nilai gain yang paling tinggi yang akan menjadi

akar pertama. Sebelum menghitung nilai gain dari atribut, hitung dahulu nilai entropy. Untuk menghitung nilai entropy digunakan rumus.

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

Keterangan:

S : H<sub>i</sub>

n : Jumlah Partisi S

p<sub>i</sub> : Proporsi S<sub>i</sub> terhadap S

3. Menghitung nilai gain menggunakan persamaan berikut.

$$\text{Gain}(S,A) = \text{entropy}(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} \text{Entropy}(S_i)$$

Keterangan:

S: Himpunan Kasus

A: Fitur

n: Jumlah partisi atribut A

|S<sub>i</sub>|: Proporsi S<sub>i</sub> Terhadap S

|S|: Jumlah Kasus S

4. Ulangi Langkah ke 2 dan Langkah ke 3 hingga semua record partisi.
5. Proses partisi pohon keputusan akan berhenti saat:
  - a. Semua record dalam simpul N mendapatkan kelas yang sama.
  - b. Tidak ada atribut dalam record yang dipartisi lagi
  - c. Tidak ada record di dalam cabang yang sama

### 2.3.2. Algoritma naïve bayes

Algoritma naïve bayes merupakan teknik prediksi berbasis probabilitas sederhana yang berdasar pada penerapan aturan bayes dengan asumsi ketidaktergantungan yang kuat atau bersifat bebas (independence) (Subarkah et al., 2020). Algoritma ini juga termasuk kedalam algoritma klasifikasi sama halnya dengan C4.5. Mudah digunakan untuk data machine learning merupakan kelebihan dari algoritma naïve bayes, lalu algoritma naïve bayes hanya

membutuhkan satu kali scan data training, dan digunakan untuk penanganan nilai atribut yang hilang dan data kontinu. Algoritma naïve bayes mempunyai keunggulan lainnya yaitu karena mudah dibangun, sehingga tidak membutuhkan skema estimasi dan parameter iterative yang rumit dimana dapat langsung diterapkan pada data yang bersifat besar atau banyak (Wicaksono et al., 2022). Selain itu, algoritma naïve bayes mudah untuk ditafsirkan serta memiliki kinerja yang konsisten baik diskala yang berdimensi tinggi (Basuki, 2023). Serta metode ini dapat menentukan parameter estimasi klasifikasi hanya berdasarkan data training yang relative kecil (Anam et al., 2021). Untuk mengklasifikasikan data dengan menggunakan algoritma naïve bayes maka dapat menggunakan persamaan sebagai berikut :

$$P(R|S) = \frac{P(R)P(S|R)}{P(S)}$$

Keterangan:

R : Data yang dicari kelasnya

S : Hipotesis pada data kelas khusus

P(R|S) : Nilai probabilitas berdasarkan hipotesis R yang berdasarkan kondisi S

P(R) : Nilai probabilitas sesuai hipotesis R

P(S|R) : Nilai probabilitas S yang didasarkan dengan hipotesis R

P(S) : Nilai probabilitas S

#### 1. Pengujian hasil

Pengujian akan hasil pengimplementasian data yang dilakukan pada kedua algoritma dilakukan dengan menggunakan software rapid miner. Mekanisme pengujian dilakukan dengan menggunakan variabel yang ada berupa data dari siswa SDN 03 Karanganyar yaitu prestasi, keaktifan, penghasilan orang tua, jumlah tanggungan dan kondisi rumah siswa. Pengujian dilakukan pada algoritma C4.5 dan naïve bayes dengan standar pengukuran yang sama sehingga dapat diketahui algoritma yang lebih baik.

Hasil pengujian disajikan dalam confusion matrix atau biasa disebut juga dengan confusion table yang merupakan metode yang dapat digunakan untuk menghitung ketepatan atau akurasi dalam data mining yang kemudian dapat juga untuk mencari nilai presisi, serta *recall* dari hasil perhitungan. Table ini memiliki fungsi untuk melakukan perekapan antara hasil yang didapat sehingga ditemukan kecocokan serta digunakan untuk mengukur kinerja dari metode klasifikasi (Normawati & Prayogi, 2021). Pengukuran kinerja dalam confusion matrix dapat diinterpretasikan sebagai berikut.

Tabel 1. Confusion matrix

| Kelas   | Terklasifikasi Positif                                                 | Terklasifikasi Negatif                                       |
|---------|------------------------------------------------------------------------|--------------------------------------------------------------|
| Positif | TP (True Positif) data yang terdeteksi dengan benar                    | FN (False Negatif) data positif yang terdeteksi salah        |
| Negatif | FP (False Positif) data negative namun terdeteksi sebagai data positif | TN (True Negatif) data negative yang terdeteksi dengan benar |

Untuk dapat mengetahui kinerja dari setiap algoritma yang diujikan maka tahapan selanjutnya adalah menghitung nilai akurasi, presisi dan *recall* dari hasil pengujian tersebut (Damuri et al., 2021).

Akurasi merupakan penggambaran seberapa akurat suatu algoritma dapat mengklasifikasikan data dengan benar. Berarti disini jumlah prediksi yang tepat dibagi dengan jumlah seluruh data dirumuskan dalam persamaan berikut.

$$\text{Akurasi} = \frac{TP + TN}{TP + TN + FP + FN}$$

Presisi merupakan penggambaran akan jumlah data dengan kategori positif yang dapat diklasifikasikan secara benar dibagi dengan total data yang diklasifikasikan positif dengan persamaan sebagai berikut.

$$\text{Presisi} = \frac{FP}{FP + TP} \times 100 \%$$

*Recall* menunjukkan besaran persentasi data kategori positif yang terklasifikasi dengan benar sesuai dengan persamaan berikut.

$$Recall = \frac{TP}{FN + TP} \times 100 \%$$

## 2. Komparasi dan Analisa

Komparasi dilakukan pada hasil perhitungan serta pengujian berdasarkan hasil dari penghitungan yang dilakukan dengan menggunakan software rapid miner besaran akurasi, presisi, dan *recall* menjadi tolak ukur dari kinerja algoritma dalam mengklasifikasikan data. Komparasi yang dilakukan diantar algoritma C4.5 dan algoritma naïve bayes dapat membantu dalam pengambilan kesimpulan. Analisa yang dilakukan untuk dapat mengetahui informasi yang didapat dari komparasi hasil pengujian tersebut sehingga dapat ditentukan algoritma dengan kinerja terbaik atau efisien.

## 3. Pengambilan kesimpulan

Akhir dari penelitian ini adalah pengambilan kesimpulan berdasarkan perhitungan, komparasi, dan pengujian yang dilakukan pada hasil berdasarkan data yang dimiliki dengan menerapkan algoritma c4.5 dan algoritma naïve bayes. Pengambilan kesimpulan sebagai hasil dari penelitian yang dilakukan ini sebagai benang merah tercapainya tujuan dilakukan penelitian ini.

# HASIL PENELITIAN DAN PEMBAHASAN

Berikut merupakan hasil dan pembahasan dari penelitian ini :

## 3.1 Implementasi algoritma c4.5

Penerapan algoritma C4.5 pada data siswa SDN 03 Karanganyar akan membentuk rule dan suatu pohon keputusan (decision tree) (Halim et al., 2023). Penghitungan entropy dilakukan pada total seluruh data dan setiap atribut pada data. Pengitungan gain dilakukan juga pada seluruh atribut pada data. Nilai gain atribut terbesar di tetapkan sebagai node akar dalam iterasi pertama atau menjadi node 1. Entropy total pada penelitian ini dicari dengan

minus rekomendasi per jumlah data dikali  $\log_2$  dari rekomendasi per jumlah data ditambah minus tidak rekomendasi per jumlah data dikali  $\log_2$  dari tidak rekomendasi per jumlah data. Apabila mencari nilai entropy per atribut jumlah data diganti dengan jumlah dari rekomendasi dan tidak rekomendasi pada nilai atribut tersebut. Penghitungan nilai gain dicari dengan entropi total dikurangi jumlah nilai atribut per jumlah data dikali dengan entropy atribut tersebut dikurangi dengan hasil perhitungan dengan cara yang sama pada setiap atribut. Hasil perhitungan disajikan dalam table berikut

Tabel 2. Node 1

| Atribut        | Jumlah | Rekomendasi | Tidak | Entropy | Gain  |
|----------------|--------|-------------|-------|---------|-------|
|                | 139    | 51          | 87    | 1.271   |       |
| Nilai          |        |             |       |         |       |
| $\geq 90$      | 44     | 26          | 18    | 1.624   |       |
| $\geq 80$      | 83     | 29          | 59    | 1.096   |       |
| $\geq 70$      | 12     | 1           | 11    | 0.441   |       |
|                |        |             |       |         | 0.064 |
| Prestasi       |        |             |       |         |       |
| Banyak         | 48     | 19          | 29    | 1.315   |       |
| Cukup          | 50     | 14          | 36    | 1.055   |       |
| Kurang         | 41     | 18          | 23    | 1.397   |       |
|                |        |             |       |         | 0.025 |
| Penghasilan    |        |             |       |         |       |
| $\geq 4$ juta  | 8      | 0           | 8     | 0.000   |       |
| $\geq 3$ .juta | 12     | 3           | 9     | 0.978   |       |
| $\geq 2$ .juta | 27     | 8           | 19    | 1.096   |       |
| $< 2$ juta     | 92     | 40          | 52    | 1.390   |       |
|                |        |             |       |         | 0.054 |
| Tanggungan     |        |             |       |         |       |
| $>5$           | 30     | 11          | 19    | 1.255   |       |
| $<4$           | 109    | 40          | 69    | 1.256   |       |
|                |        |             |       |         | 0.015 |
| Kondisi rumah  |        |             |       |         |       |
| Kurang         | 41     | 34          | 7     | 1.747   |       |
| Layak          | 98     | 17          | 81    | 0.758   |       |
|                |        |             |       |         | 0.221 |

Dari perhitungan nilai gain pada setiap atribut tersebut dilakukan perbandingan dan dapat disimpulkan bahwa nilai gain dari kondisi rumah

merupakan nilai gain tertinggi maka kondisi rumah menjadi node akar (Abidin, 2023). Atribut kondisi rumah memiliki 2 nilai yaitu kurang dan layak, untuk menentukan node selanjutnya maka perlu melakukan perhitungan pada nilai kurang dan layak, karena klasifikasi dari nilai keduanya belum mutlak rekomendasi atau tidak. Bisa diartikan bahwa walaupun kondisi rumah menjadi prioritas penentuan penerima bantuan namun tidak mutlak bahwa siswa dengan kondisi rumah yang layak tidak dapat menerima bantuan karena pertimbangan juga berdasarkan beberapa atribut pendukung lainnya. Pada node selanjutnya nilai jumlah data dan entropi total diganti dengan jumlah dan entropi atribut dengan gain terbesar pada perhitungan sebelumnya. Atribut kondisi rumah tidak dilakukan perhitungan kembali karena telah menjadi node akar begitu pula seterusnya setiap iterasi nilai gain tertinggi akan diambil dan menyisakan atribut yang lain, dilakukan penghitungan hingga tersisa atribut terakhir.

Tabel 3. Node 2

| Atribut     | Jumlah | Rekomendasi | Tidak | Entropy | Gain  |
|-------------|--------|-------------|-------|---------|-------|
|             | 41     | 34          | 7     | 1.747   |       |
| Nilai       |        |             |       |         |       |
| ≥ 90        | 13     | 13          | 0     | 0       |       |
| ≥ 80        | 24     | 20          | 4     | 1.745   |       |
| ≥ 70        | 4      | 1           | 3     | 0.977   | 1.417 |
| Prestasi    |        |             |       |         |       |
| Banyak      | 12     | 10          | 2     | 1.745   |       |
| Cukup       | 12     | 10          | 2     | 1.745   |       |
| Kurang      | 17     | 14          | 3     | 1.748   | 1.231 |
| Penghasilan |        |             |       |         |       |
| ≥ 4 juta    | 0      | 0           | 0     | 0.000   |       |
| ≥ 3.juta    | 0      | 0           | 0     | 0.978   |       |
| ≥ 2.juta    | 1      | 1           | 0     | 1.096   |       |
| < 2 juta    | 33     | 26          | 7     | 1.390   | 0.054 |
| Tanggungan  |        |             |       |         |       |
| >5          | 30     | 11          | 19    | 1.255   |       |
| <4          | 109    | 40          | 69    | 1.256   | 0.015 |

Atribut nilai menjadi gain dengan nilai tertinggi, dalam atribut nilai dilakukan perbandingan antara jumlah entropy dan diambil nilai atribut dengan jumlah entropy tertinggi.

Tabel 4. Node 3

| Atribut     | Jumlah | Rekomen | Tidak | Entropy | Gain  |
|-------------|--------|---------|-------|---------|-------|
|             | 24     | 20      | 4     | 1.745   |       |
| Prestasi    |        |         |       |         |       |
| Banyak      | 12     | 10      | 2     | 1.745   |       |
| Cukup       | 12     | 10      | 2     | 1.745   |       |
| Kurang      | 17     | 14      | 3     | 1.748   |       |
|             |        |         |       |         | 1.230 |
| Penghasilan |        |         |       |         |       |
| ≥ 4 juta    | 0      | 0       | 0     | 0       |       |
| ≥ 3.juta    | 0      | 0       | 0     | 0       |       |
| ≥ 2.juta    | 1      | 1       | 0     | 0       |       |
| < 2 juta    | 33     | 26      | 7     | 1.748   |       |
|             |        |         |       |         | 1.242 |
| Tanggungan  |        |         |       |         |       |
| >5          | 8      | 8       | 0     | 0       |       |
| <4          | 32     | 25      | 7     | 1.751   |       |
|             |        |         |       |         | 189.1 |

Atribut tanggungan orang tua memiliki nilai gain tertinggi pada iterasi atau node 3, sehingga tersisa dua atribut yaitu penghasilan dan prestasi.

Tabel 5. Node 4

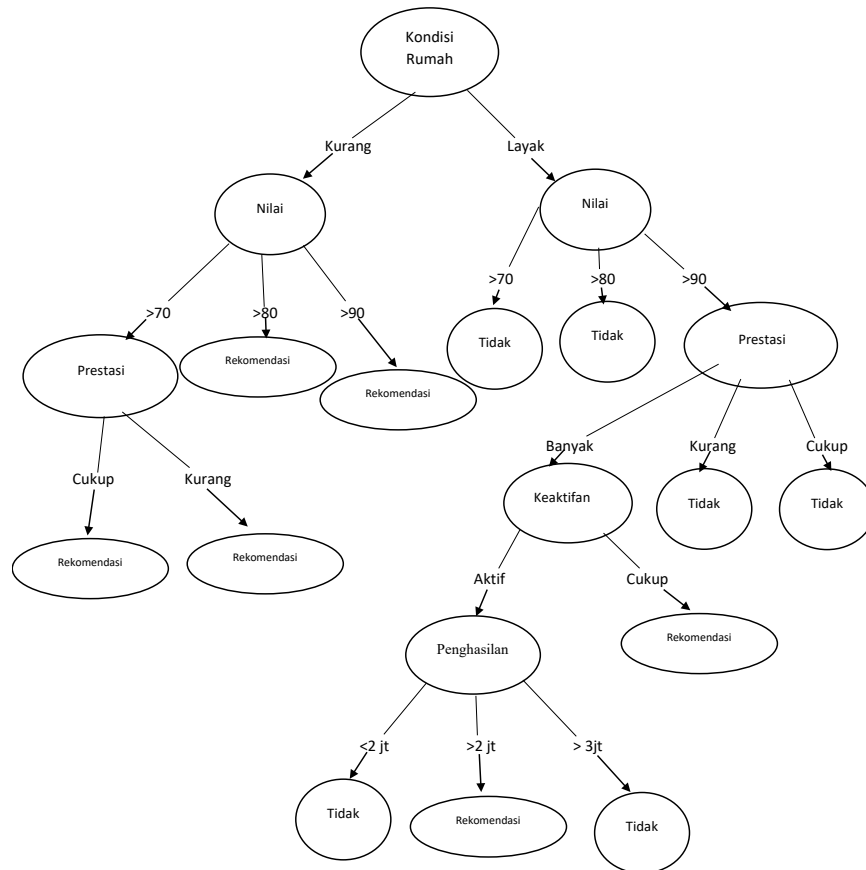
| Atribut     | Jumlah | Rekomendasi | Tidak | Entropy | Gain  |
|-------------|--------|-------------|-------|---------|-------|
|             | 32     | 25          | 7     | 1.751   |       |
| Prestasi    |        |             |       |         |       |
| Banyak      | 6      | 4           | 2     | 1.698   |       |
| Cukup       | 5      | 4           | 1     | 1.752   |       |
| Kurang      | 6      | 5           | 1     | 1.745   |       |
|             |        |             |       |         | 1.572 |
| Penghasilan |        |             |       |         |       |
| ≥ 4 juta    | 0      | 0           | 0     | 0       |       |
| ≥ 3.juta    | 0      | 0           | 0     | 0       |       |
| ≥ 2.juta    | 1      | 1           | 0     | 0       |       |
| < 2 juta    | 16     | 12          | 4     | 1.745   |       |
|             |        |             |       |         | 1.550 |

Tersisa atribut terakhir yaitu penghasilan orang tua, sebagai node terakhir dalam perhitungan menggunakan algoritma c4.5 pada penelitian ini.

Tabel 6. Node 5

| Atribut     | Jumlah | Rekomendasi | Tidak | Entropy | Gain  |
|-------------|--------|-------------|-------|---------|-------|
|             | 5      | 4           | 1     | 1.752   |       |
| Penghasilan |        |             |       |         |       |
| ≥ 4 juta    | 0      | 0           | 0     | 0       |       |
| ≥ 3.juta    | 0      | 0           | 0     | 0       |       |
| ≥ 2.juta    | 1      | 1           | 0     | 0       |       |
| < 2 juta    | 16     | 12          | 4     | 1.745   |       |
|             |        |             |       |         | 1.551 |

Dari penghitungan yang dilakukan akan terbentuk pohon keputusan atau decision tree berdasarkan pohon keputusan tersebut dapat digunakan dalam pengambilan keputusan berdasarkan dari rule yang ada dengan menafsirkan solusi untuk masalah yang terjadi (Adinda, 2022). Decision tree berfungsi untuk mengeksplorasi data, menemukan hubungan tersembunyi antar variable dalam data. Pohon keputusan pada penelitian ini tentunya mengenai rekomendasi berdasarkan tingkat prioritas dari penghitungan entropi dan gain pada data siswa SDN 03 Karanganyar. Sama halnya dalam pengambilan suatu keputusan akan di pengaruhi oleh banyak factor menjadikan pertimbangan sebelum melakukan pengambilan keputusan.



Gambar 2. Pohon Keputusan

### 3.2 Implementasi algoritma naïve naves

Merupakan algoritma yang mengklasifikasikan probabilistik sederhana serta menghitung sekumpulan probalilistik data dengan menjumlahkan frekuensi atau atribut dari data yang dimiliki. Metode yang terdapat pada algoritma naïve bayes perlunya diketahui bahwa dalam proses klasifikasi memerlukan petunjuk untuk menentukan kelas atau kategori untuk data yang dianalisis maka dari itu dibutuhkan data testing dan juga data training. Algoritma naïve bayes ini banyak digunakan pada klasifikasi teks yang memiliki permasalahan dengan berbagai kelas, hal ini karena algoritma ini relative membutuhkan waktu yang sedikit untuk mengelola data. Dimulai dengan menghitung nilai probabilitas total atau probabilitas prior pada data training yaitu pada penelitian ini dengan atribut rekomendasi atau tidak rekomendasi dibagi degan jumlah data. Penghitungan probabilitas menjadi

tumpuan dalam algoritma (Yudhana et al., 2022) ini maka dilakukan pada semua atribut yang dimiliki.

Tabel 7. Probabilitas prior

| $(P(C_i))$ | Rekomendasi | Tidak |
|------------|-------------|-------|
|            | 0.366       | 0.633 |

Dalam naïve bayes menggunakan metode yang sama dalam memprediksi probabilitas kelas yang berbeda berdasarkan berbagai atribut. Pada penelitian ini dalam penentuan probabilitas atribut dilakukan pada tiap nilai pada atribut dibagi dengan jumlah atribut yang termasuk kedalam rekomendasi atau tidak direkomendasi.

Tabel 8. Probabilitas atribut

| Atribut               | Rekomendasi | Tidak    |
|-----------------------|-------------|----------|
| Nilai rata-rata       |             |          |
| 75                    | 0.020       | 0.125    |
| 80                    | 0.314       | 0.420    |
| 85                    | 0.157       | 0.250    |
| 90                    | 0.431       | 0.170    |
| 95                    | 0.078       | 0.034    |
| Prestasi              |             |          |
| Kurang                | 0.353       | 0.261    |
| Cukup                 | 0.275       | 0.409    |
| Banyak                | 0.373       | 0.330    |
| Keaktifan             |             |          |
| Aktif                 | 0.490196    | 0.568182 |
| Cukup                 | 0.509804    | 0.431818 |
| Penghasilan orang tua |             |          |
| 800000                | 0.157       | 0.068    |
| 850000                | 0.078       | 0.011    |
| 900000                | 0.235       | 0.080    |
| 950000                | 0           | 0.023    |
| 1000000               | 0.196       | 0.148    |
| 1100000               | 0.039       | 0.023    |
| 1200000               | 0.039       | 0.080    |
| 1300000               | 0           | 0.045    |
| 1400000               | 0.020       | 0.023    |
| 1500000               | 0.020       | 0.091    |
| 2000000               | 0.059       | 0.114    |

|               |       |       |
|---------------|-------|-------|
| 2500000       | 0.098 | 0.102 |
| 3000000       | 0.059 | 0.034 |
| 3500000       | 0     | 0.068 |
| 4000000       | 0     | 0.023 |
| 4200000       | 0     | 0.034 |
| 4300000       | 0     | 0.011 |
| 4500000       | 0     | 0.023 |
| Tanggungan    |       |       |
| 2             | 0.019 | 0.011 |
| 3             | 0.313 | 0.295 |
| 4             | 0.450 | 0.477 |
| 5             | 0.156 | 0.204 |
| 7             | 0.058 | 0.011 |
| Kondisi rumah |       |       |
| Kurang        | 0.666 | 0.079 |
| Layak         | 0.333 | 0.920 |

Penghitungan dengan algoritma naïve bayes dilanjutkan dengan menghitung probabilitas data testing berdasarkan hasil dari probabilitas tiap atribut data training. Dimana dalam penelitian ini setiap data dilakukan penghitungan kepada kemungkinan untuk direkomendasi atau tidak direkomendasi, apabila salah satu dari kedua kelas tersebut lebih tinggi nilainya maka masuk kedalam kelas dengan nilai tertinggi tersebut.

Penghitungan data testing dilakukan untuk kelas prediksi rekomendasi sebagai berikut :

$$P(\text{Nama siswa Rekomendasi}) = P(\text{nilai rata-rata raport Rekomendasi}) \times P(\text{prestasi rekomendasi}) \times P(\text{Keaktifan rekomendasi}) \times P(\text{Penghasilan rekomendasi}) \times P(\text{Tanggungan rekomendasi}) \times P(\text{Kondisi Rumah rekomendasi})$$

Sedangkan untuk perhitungan pada data testing untuk kelas prediksi tidak direkomendasi sebagai berikut.

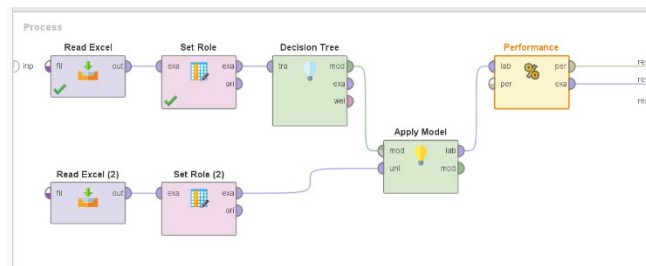
$$P(\text{Nama siswa Tidak direkomendasi}) = P(\text{nilai rata-rata raport Tidak direkomendasi}) \times P(\text{prestasi Tidak direkomendasi}) \times P(\text{Keaktifan Tidak direkomendasi}) \times P(\text{Penghasilan Tidak direkomendasi}) \times P(\text{Tanggungan Tidak direkomendasi}) \times P(\text{Kondisi Rumah Tidak direkomendasi})$$

Hasil dari penghitungan tersebut merupakan prediksi yang didapat dalam penerapan algoritma naïve bayes. Tentunya bisa ditemui ketidaksaan akan

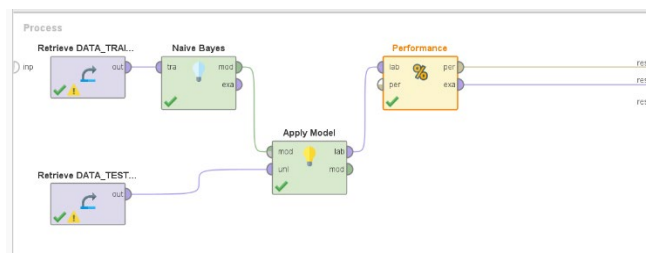
kelas yang dihasilkan hal ini dikarenakan nilai akurasi yang dimiliki oleh algoritma naïve bayes itu sendiri. Algoritma naïve bayes tergolong kedalam algoritma yang mudah dan cepat dalam memprediksi data.

### 3.3 Pengujian hasil

Dalam melakukan pengujian tentunya harus dilakukan dengan standar yang sama sehingga dapat diketahui yang terbaik diantara algoritma dari perbandingan tersebut. Pengujian yang dilakukan pada penelitian ini dengan melakukan penghitungan secara manual dengan microsoft excel. Dan menggunakan software rapid miner yang menerapkan algoritma C4.5 dan algoritma naïve bayes dengan menggunakan fitur-fitur yang dimiliki oleh software rapid miner sesuai kebutuhan untuk menentukan hasil prediksi dari algoritma C4.5 dan naïve bayes. Berdasarkan dari penghitungan yang dilakukan dengan cara manual dan menggunakan software rapid miner menunjukan hasil yang sama.



Gambar 3. Penghitungan riped miner algoritma c4.5



Gambar 4. Penghitungan riped miner algoritma naïve bayes

Berikut merupakan hasil dari penghitungan menggunakan rapid miner dengan algoritma C4.5 dan naïve bayes yang dengan data training dan data testing yang sama.

Tabel 9. Confusion table hasil pengujian pada rapid miner dengan algoritma C4.5

| Predicted   | Class       |       |
|-------------|-------------|-------|
|             | Rekomendasi | Tidak |
| Rekomendasi | 48          | 8     |
| Tidak       | 3           | 80    |

Tabel 10. Confusion table hasil pengujian pada rapid miner dengan algoritma naïve bayes

| Predicted   | Class       |       |
|-------------|-------------|-------|
|             | Rekomendasi | Tidak |
| Rekomendasi | 35          | 9     |
| Tidak       | 16          | 79    |

Untuk dapat mengetahui kinerja dari setiap algoritma yang diujikan maka tahapan selanjutnya adalah menghitung nilai akurasi, presisi dan *recall* dari hasil penghitungan tersebut. Berikut merupakan hasil pengujian terhadap data dengan nilai akurasi, presisi dan recall.

Tabel 11. Tabel pengujian

| Pengujian      | Algoritma |             |
|----------------|-----------|-------------|
|                | C4.5      | Naïve bayes |
| <i>Akurasi</i> | 92 %      | 82 %        |
| <i>Presisi</i> | 94 %      | 68 %        |
| <i>Recall</i>  | 85 %      | 79 %        |

Dari pengujian yang dilakukan pada data yang menerapkan algoritma C4.5 dan naïve bayes berdasarkan akurasi dapat dilihat bahwa akurasi yang didapatkan oleh algoritma C4.5 sebesar 92% lebih tinggi dari pada akurasi yang didapatkan dari algoritma naïve bayes sebesar 82%. Akurasi didapatkan dari jumlah prediksi yang tepat berdasarkan data testing dan prediksi dari data training dibagi dengan jumlah seluruh data. Hal ini dapat disimpulkan bahwa keakuratan penghitungan data dengan menggunakan algoritma C4.5 lebih baik karena dapat mengklasifikasikan data secara benar dengan lebih banyak (Rahman et al., 2022).

Berdasarkan hasil pengujian yang dilakukan dengan menghitung presisi dari hasil penghitungan prediksi penerima bantuan PIP dengan penerapan

algoritma c4.5 dan naïve bayes maka dapat dilihat bahwa nilai presisi yang dihasilkan algoritma c4.5 sebesar 94 % dan nilai presisi yang dihasilkan algoritma naïve bayes sebesar 68 %. Nilai presisi yang dihasilkan oleh algoritma c4.5 lebih besar dibandingkan yang dihasilkan oleh algoritma naïve bayes. Dalam pengujian dengan menghitung nilai presisi dapat disimpulkan bahwa algoritma c4.5 lebih baik dari pada algoritma naïve bayes hal ini dikarenakan semakin besar hasil presisi yang dihasilkan maka kemungkinan perubahan dalam setiap penghitungan akan lebih kecil (Solihat et al., 2022), sehingga hasil dari perhitungan akan berpeluang memiliki hasil yang sama atau konsisten walaupun dilakukan penghitungan beberapa kali pengulangan atau percobaan.

Pengujian yang dilakukan dengan menghitung nilai *recall* maka penghitungan data dengan menggunakan algoritma C4.5 mendapatkan nilai sebesar 85% dan hasil penghitungan *recall* dengan menggunakan algoritma naïve bayes sebesar 79%. Dari hasil tersebut maka dapat diartikan bahwa algoritma C4.5 lebih baik dibandingkan dengan algoritma naïve bayes. Hal ini dikarenakan besaran nilai *recall* menunjukkan bahwa besaran data dengan kategori positif (Sari et al., 2023) apabila dalam penelitian ini adalah data yang tergolong direkomendasi dapat diklasifikasikan dengan benar lebih banyak menggunakan algoritma C4.5.

Dari pengujian yang dilakukan dengan menghitung akurasi, presisi, dan *recall* dari hasil kedua algoritma yang gunakan yaitu C4.5 dan naïve bayes maka algoritma C4.5 lebih baik pada pengujian akurasi, presisi dan *recall*, hal ini dikarenakan nilai hasil pengujian yang didapatkan algoritma C4.5 lebih tinggi dibandingkan algoritma naïve bayes. Pengujian ini dilakukan dalam membantu pengambilan keputusan serta untuk dapat mengetahui algoritma yang lebih baik diantara algoritma C4.5 dan naïve bayes dalam menghitung prediksi rekomendasi penerima bantuan PIP.

## **SIMPULAN**

Berdasarkan hasil penghitungan dan pengujian yang telah dilakukan pada penelitian ini, maka dapat disimpulkan bahwa. Pengimplementasian dapat

dilakukan dengan baik pada data siswa SDN 03 Karanganyar untuk mengetahui rekomendasi penerima bantuan PIP dengan menggunakan algoritma C4.5 dan algoritma naïve bayes. Berdasarkan pengujian yang dilakukan pada hasil penghitungan rekomendasi penerima bantuan PIP maka algoritma C4.5 lebih baik pada pengujian akurasi sebesar 92%, presisi sebesar 94% dan *recall* sebesar 85%. Berdasarkan komparasi yang dilakukan hasil tersebut dapat dinyatakan bahwa algoritma C4.5 lebih baik dalam menghitung rekomendasi penerima bantuan PIP pada data siswa SDN 03 Karanganyar, dikarenakan memiliki hasil pengujian akurasi lebih tinggi dimana berarti dapat mengkalsifikasikan data secara benar dengan lebih banyak, dan juga dapat mengklasifikasikan data positif dalam penelitian ini data yang masuk kedalam kategori direkomendasi dengan lebih baik. Selain itu algoritma C4.5 juga cukup baik dan cepat dalam melakukan klasifikasi data serta memiliki nilai presisi yang lebih besar dibandingkan algoritma naïve bayes dimana hasil yang didapatkan cenderung lebih konsisten karena hasil yang diperoleh memiliki perbedaan yang relative kecil disetiap penghitungan walaupun dilakukan pada beberapa kali pengulangan. Sehingga rule yang didapatkan dari algoritma C4.5 ataupun naïve bayes ini dapat digunakan sebagai landasan untuk membuat sebuah system dalam membantu SDN 03 karanganyar dalam menentukan siswa yang berhak untuk diberikan rekomendasi untuk mendapatkan beasiswa.

## **DAFTAR PUSTAKA**

- Abidin, Z. (2023). Eka Nurhana 2) , Permata 3) , Faruq Ulum 4) 1,2,3 Sistem Informasi, Universitas Teknokrat Indonesia 4 Informatika, Universitas Teknokrat Indonesia 1,2,3,4 JL. ZA. Pagar Alam, 17(2), 567–582. [www.kaggle.com](http://www.kaggle.com)
- Adinda, P. R. (2022). Pohon Keputusan C4. 5 Algoritma Untuk Klasifikasi Program Bantuan Belajar. *Jurnal Portal Data*, 2(9), 1–13. <http://portaldata.org/index.php/portaldata/article/view/215%0Ahttp://portaldata.org/index.php/portaldata/article/download/215/204>
- Alfitri, M. M., & Rusda, D. (2023). *Evaluasi Performa Algoritma Naïve Bayes Dalam Mengklasifikasi Penerima Bantuan Pangan Non Tunai*. 7, 1433–1445. <https://doi.org/10.30865/mib.v7i3.6151>
- Anam, M. K., Pikir, B. N., & Firdaus, M. B. (2021). Penerapan Naïve Bayes Classifier, K-Nearest Neighbor (KNN) dan Decision Tree untuk Menganalisis Sentimen pada Interaksi Netizen dan Pemerintah. *MATRIK: Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, 21(1), 139–150.

- <https://doi.org/10.30812/matrik.v21i1.1092>
- Anggraini, S., Defit, S., & Nurcahyo, G. W. (2018). Analisis Data Mining Penjualan Ban Menggunakan Algoritma C4. 5. *Jurnal Ilmu Teknik Elektro ...*, 5(2), 0–7. <https://core.ac.uk/download/pdf/295348196.pdf>
- Armyny, S. B., Mutiah, N., & Sari, R. P. (2023). Penentuan Penerimaan Beasiswa PIP Menggunakan Metode Moora Pada SD Negeri 11 Sandai. 11(1), 43–50.
- Astuti, E. V., Afandi, A., & Efendi, D. M. (2023). Classification and Clustering of Internet Quota Sales Data Using C4.5 Algorithm and K-Means. *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika (JITEKI)*, 9(2), 268–283. <https://doi.org/10.26555/jiteki.v9i2.25970>
- Basuki, A. (2023). Sentiment Analysis of Customers' Review on Delivery Service Provider on Twitter Using Naive Bayes Classification. *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika (JITEKI)*, 9(2), 420–428. <https://doi.org/10.26555/jiteki.v9i2.26327>
- Dahlia, & Jamaludin. (2019). Implementasi Kebijakan Kartu Indonesia Pintar Dalam Meningkatkan Partisipasi Anak Melanjutkan Pendidikan Di Smp Negeri 6 Tolitoli Utara Kecamatan Tolitoli Utara Kabupaten Tolitoli. *Media Publikasi Ilmiah Prodi PPKn*, 5(01), 21–29.
- Damuri, A., Riyanto, U., Rusdianto, H., & Aminudin, M. (2021). Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako. *JURIKOM (Jurnal Riset Komputer)*, 8(6), 219. <https://doi.org/10.30865/jurikom.v8i6.3655>
- Desa, D. I., Seteko, T., Ogan, K., Tanjung, I. N., Village, S., & District, O. I. (2023). *JOLL 6 (1) (2023) Journal of Lifelong Learning*. 6(1).
- Edrial, Putrama, R., & Sujastiawan, A. (2022). Evaluasi Kebijakan Program Indonesia Pintar (Pip) Di Sma Negeri 1 Utan Tahun 2019-2020. *Jurnal Kapita Selekt Administrasi Publik*, 3(1), 109–116. <http://e-journalppmunsa.ac.id/index.php/ksap>
- Halim, K., Erny Herwindiati, D., & Sutrisno, T. (2023). Penerapan Metode Decision Tree Untuk Prakiraan Cuaca Kota Bekasi. *Jurnal Ilmu Komputer Dan Sistem Informasi*, 11(2), 1–5. <https://doi.org/10.24912/jiksi.v11i2.26026>
- Haqmanullah Pambudi, R., Darma Setiawan, B., & Indriati. (2018). Penerapan Algoritma C4.5 Untuk Memprediksi Nilai Kelulusan Siswa Sekolah Menengah Berdasarkan Faktor Eksternal. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2(7), 2637–2643. <http://j-ptiik.ub.ac.id>
- Jollyta, D., Gusrianty, G., & Sukrianto, D. (2019). Analysis of Slow Moving Goods Classification Technique: Random Forest and Naïve Bayes. *Khazanah Informatika : Jurnal Ilmu Komputer Dan Informatika*, 5(2), 134–139. <https://doi.org/10.23917/khif.v5i2.8263>
- Kurniawan, Y. I. (2018). Perbandingan Algoritma Naive Bayes dan C.45 dalam Klasifikasi Data Mining. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 5(4), 455. <https://doi.org/10.25126/jtiik.201854803>

- Lidia Lusiana. (2017). Implementasi Program Indonesia Pintar (PIP) Di Desa Sukomulyo Kecamatan Sepaku Kabupaten Penajam Paser Utara (Studi Kasus di SDN 011 dan SDN 013). *EJournal Administrasi Negara*, 6(1), 6991–7005. [https://ejournal.ap.fisip-unmul.ac.id/site/wp-content/uploads/2018/01/Jurnal LIDYA \(01-17-18-07-33-51\).pdf](https://ejournal.ap.fisip-unmul.ac.id/site/wp-content/uploads/2018/01/Jurnal LIDYA (01-17-18-07-33-51).pdf)
- Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 5(2), 697–711.
- Nugraha, J. T., & UUD. (1945). *No 主観的健康感を中心とした在宅高齢者における健康関連指標に関する共分散構造分析* Title. 105(3), 129–133. <https://webcache.googleusercontent.com/search?q=cache:BDsuQOHoCi4J:hhttps://media.neliti.com/media/publications/9138-ID-perlindungan-hukum-terhadap-anak-dari-konten-berbahaya-dalam-media-cetak-dan-ele.pdf+&cd=3&hl=id&ct=clnk&gl=id>
- Nurokhmah, N. (2021). Program Indonesia Pintar (PIP): Implementasi Kebijakan Kesejahteraan dalam Upaya Meningkatkan Angka Partisipasi Pendidikan. *Jurnal Paradigma: Jurnal Multidisipliner Mahasiswa Pascasarjana Indonesia*, 2(1), 37–48. <https://jurnal.ugm.ac.id/paradigma/article/view/66256>
- Octavianti, F. S. (2023). Implementasi Konvensi Hak Anak Di Kota Jakarta Utara (Studi Kasus: Pemenuhan Hak Asasi Manusia Perspektif Pendidikan Bagi Pelajar SMA). *Pedagogi: Jurnal Ilmu Pendidikan*, 23(1), 30–37. <https://doi.org/10.24036/pedagogi.v23i1.1476>
- Parapat, J. S., & Sinaga, A. S. (2018). Data Mining Algoritma C4.5 Pada Klasifikasi Kredit Koperasi Simpan Pinjam. *Jurnal Ilmu Teknik Elektro Komputer Dan Informatika (JITEKI)*, 4(2), 144–154.
- Primajaya, A., Sari, B. N., & Khusaeri, A. (2020). Prediksi Potensi Kebakaran Hutan dengan Algoritma Klasifikasi C4.5 Studi Kasus Provinsi Kalimantan Barat. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 6(2), 188. <https://doi.org/10.26418/jp.v6i2.37834>
- Rahman, F. Y., Purnomo, I. I., & Hijriana, N. (2022). Penerapan Algoritma Data Mining Untuk Klasifikasi Kualitas Air. *Technologia : Jurnal Ilmiah*, 13(3), 228. <https://doi.org/10.31602/tji.v13i3.7070>
- Rohaeni, N. E., & Saryono, O. (2018). Implementasi Kebijakan Program Indonesia Pintar (PIP) Melalui Kartu Indonesia Pintar (KIP) dalam Upaya Pemerataan Pendidikan. *Journal of Education Management and Administration Review*, 2(1), 193–204.
- Sari, L., Romadloni, A., & Listyaningrum, R. (2023). Penerapan Data Mining dalam Analisis Prediksi Kanker Paru Menggunakan Algoritma Random Forest. *Infotekmesin*, 14(1), 155–162. <https://doi.org/10.35970/infotekmesin.v14i1.1751>
- Sianipar, F. P., & Maulia, S. T. (2023). Kebijakan Pendidikan : Pemerataan dan

- Perluasan Akses ( Rehabilitasi Ruang Kelas Yang Rusak ). *Jurnal Pendidikan Sastra Inggris*, 3(1), 191–198.
- Solihat, I., Tirta, A. P., Ramdani, A. P., & Nandang Roziyanto, A. (2022). Verifikasi Metode Pengujian Kadar Nitrit dalam Air Limbah Secara Spektrofotometri UV-Visibel. *KOVALEN: Jurnal Riset Kimia*, 8(1), 53–59. <https://doi.org/10.22487/kovalen.2022.v8.i1.15756>
- Subarkah, P., Pambudi, E. P., & Hidayah, S. O. N. (2020). Perbandingan Metode Klasifikasi Data Mining untuk Nasabah Bank Telemarketing. *MATRIK : Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, 20(1), 139–148. <https://doi.org/10.30812/matrik.v20i1.826>
- Wicaksono, A. H., Supianto, A. A., Wijoyo, S. H., Krisnandi, D., Heryana, A., Brawijaya, U., & Korespondensi, P. (2022). *Klasifikasi Siswa Slow Learner Untuk Mendukung Sekolah Algoritma Naïve Bayes Classification of Slow Learner Students To Support Schools in Improving Student Understanding Using the Naïve Bayes*. 9(3), 589–596. <https://doi.org/10.25126/jtiik.202295509>
- Yudhana, A., Riadi, I., & Djou, M. R. (2022). Pengembangan Layanan Kependudukan Dan Pencatatan Sipil Menggunakan Algoritma Naïve Bayes. *JURIKOM (Jurnal Riset Komputer)*, 9(4), 1062. <https://doi.org/10.30865/jurikom.v9i4.4515>