

# KLASIFIKASI POLA PENJURUSAN BIDANG IT MENGUNAKAN ALGORITMA C4.5 STUDI KASUS SMKN MANONJAYA

Evi Dewi Sri Mulyani<sup>1</sup>, Irma Agustina<sup>2</sup>, Rubi Ismanto<sup>3</sup>, Sania Januar Susanto<sup>4</sup>, Septian Setia Adhi<sup>5</sup>,  
Tia Sri Rismayanti<sup>6</sup>

1, 2, 3, 4, 5, 6, Program Studi Teknik Informatika, STMIK Tasikmalaya

<sup>1</sup>eviajadech@gmail.com, <sup>2</sup>irmaagustina19@gmail.com, <sup>3</sup>ismanto.ruby@gmail.com

<sup>4</sup>esaniajanuars@gmail.com, <sup>5</sup>septiansetiaadhi@gmail.com, <sup>6</sup>tiasrirismayanti@gmail.com

---

## Abstrak

Penentuan jurusan di SMK merupakan hal yang sangat penting, karena siswa akan dibekali pembelajaran sesuai dengan jurusan yang telah dipilih sehingga memiliki keterampilan sebagai bekal memasuki dunia kerja khususnya dibidang IT. Namun prosesnya masih sering mengalami masalah, karena pihak sekolah masih sulit dalam menentukan kebijakan untuk penentuan jurusan calon siswanya, diperlukan banyak factor penilaian dan semua data diolah dalam waktu yang terbatas. Maka dari itu pihak sekolah perlu mengetahui pola penjurusan siswa berdasarkan nilai tes *software*, nilai tes *hardware*, dan test jaringan perbandingan nilai dari dua jurusan tersebut dan jumlah intervalnya. Nantinya pola tersebut memudahkan pihak sekolah dalam menentukan jurusan mana yang tepat untuk siswa baru tersebut. Algoritma yang digunakan merupakan algoritma C4.5 yang merupakan salah satu dari bagian teknik klasifikasi dengan kemampuan model algoritma C4.5 yang dapat melakukan proses learning dan klasifikasi yang sederhana, cepat dan secara umum mempunyai tingkat akurasi yang tinggi. Dengan menggunakan dataset 215 record data siswa baru, hasil dari pengklasifikasian data pemilihan jurusan SMK bidang IT untuk jurusan TKJ dan RPL menghasilkan akurasi sebesar 96,73%. Dengan tingkat akurasi yang baik, pola klasifikas ipenentuan jurusan ini dapat digunakan pihak SMKN Manonjaya dalam menentukan kebijakan penentuan jurusan siswa baru khususnya bidang IT.

**Kata Kunci:** Data mining, Klasifikasi penjurusan, Algoritma C4.5

---

## Abstract

*Determination of majors in vocational schools is very important, because students will be provided with learning in accordance with the majors that have been selected so therefore, they have skills as provisions to enter the workforce, especially in IT field. However, in the process there are still often problems, because the school feels difficult in determining policies to decide the direction of prospective students. Many factors are required for assessment and all the data must be processed in a limited time. Therefore, the school needs to know the pattern of student majors based on software test scores, hardware test scores, and network test comparison of the values of the two majors and the number of intervals. Later, the pattern will help the school to determine which course is suitable for the new student. The algorithm used was the C4.5 algorithm which is one of the classification techniques with the ability of the C4.5 algorithm model that can carry out learning and classification processes that are simple, fast and generally have a high degree of accuracy. By using a dataset of 215 new student data records, the results of classifying the data selection of vocational majors in IT fields for TKJ and RPL majors produced an accuracy of 96.73%. With that level of accuracy, the pattern of classification of the majors can be used by Manonjaya Vocational School in determining the policy for deciding the majors of new students, especially in the IT field.*

**Key Words:** Data mining, majors classification, C4.5 algorithm

---

## 1. Pendahuluan

Pendidikan merupakan salah satu sasaran pokok pemerintah dalam rangka meningkatkan kesejahteraan

rakyat. Pendidikan juga sebagai wadah pembinaan tenaga kerja untuk dapat menambah lapangan pekerjaan, serta untuk memperoleh status tertentu dalam

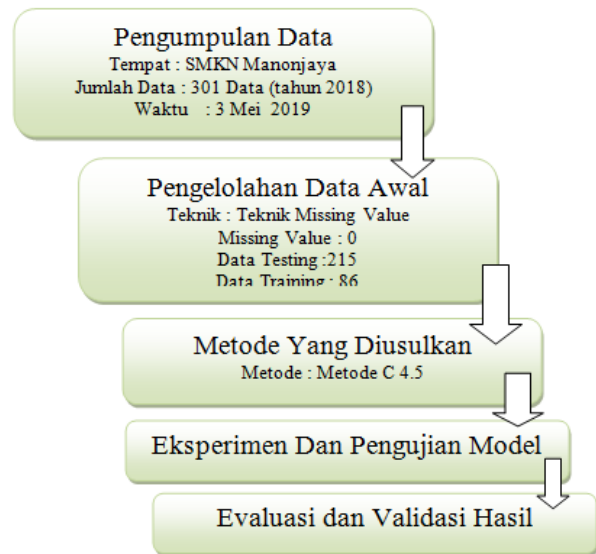
masyarakat. Paradigma dan sistem pendidikan harus disesuaikan dengan tuntutan zaman. Adanya perubahan diharapkan dapat menuju pendidikan masa depan yang lebih baik. Pendidikan sekarang harus berorientasi pada dunia kerja sehingga penekanannya tidak semata-mata aspek kognitif akan tetapi juga aspek afektif dan psikomotorik, salah satu metode sistem pendidikan yang cukup berkembang saat ini adalah sekolah kejuruan [1].

Sekolah kejuruan merupakan bagian dari sistem pendidikan yang fokusnya mempersiapkan siswanya agar lebih andal dalam bidang atau kejuruan yang digelutinya guna mempersiapkan diri terjun di dunia industri di masa mendatang [2].

Pola penjurusan tidak cukup hanya berpatokan kepada pilihan pertama dari calon siswa, tetapi pihak sekolah harus dengan jeli melihat potensi dan kemampuan yang dimiliki oleh para calon siswa ataupun siswa baru yang sudah lulus dalam tahap seleksi. Karena tidak sedikit siswa baru yang menetapkan pilihan pertama jurusannya atas dasar ikut-ikutan atau coba-coba saja, maka di sinilah peran penting sekolah untuk menempatkan para siswa baru kedalam jurusan sesuai dengan minat dan potensi yang dimiliki [3].

Salah satu dari sekolah yang menggunakan sistem kejuruan adalah SMKN Manonjaya, memiliki 3 bidang kejuruan yakni Teknik Otomotif (TO), Teknik Informatika (TI) dan Administrasi. Untuk bidang TI sendiri SMKN Manonjaya memiliki dua jurusan yakni Teknik Komputer Jaringan dan Rekayasa Perangkat Lunak. Teknik Informatika sendiri merupakan bidang kejuruan yang paling diminati oleh calon siswa SMKN Manonjaya, terbukti dalam satu angkatan terakhir bidang TI memiliki jumlah kelas paling banyak yakni enam kelas dengan TKJ sebanyak 4 kelas dan RPL 2 kelas, sedangkan bidang TO pada urutan kedua dengan jumlah 4 kelas, yakni TSM 2 kelas dan TKR 2 kelas, selanjutnya bidang administrasi sebanyak dua kelas. Melihat begitu banyak peminat bidang Teknik Informatika, maka pihak sekolah harus bisa manajemen pola penjurusan di bidang tersebut, hal tersebut sangat penting untuk mengoptimalkan potensi yang dimiliki siswa serta untuk mengembangkan bibit-bibit unggul untuk dijadikan delegasi oleh pihak sekolah dalam partisipasi kegiatan-kegiatan perlombaan, seperti contoh Lomba Kompetensi Siswa (LKS).

*Data mining* adalah suatu istilah yang digunakan untuk menemukan pengetahuan yang tersembunyi di dalam *database* [4], yang berguna untuk memecahkan masalah dengan menganalisis data yang telah ada dalam *database* sehingga menghasilkan keputusan [5]. *Data mining* telah terbukti memegang peran penting dilini masa insdustri perbankn keuangan dalam skala besar seperti bank ataupun skala kecil seperti koperasi [6], begitu pula untuk mengatasi permasalahan penentuan jurusan dengan penerapan algoritma *Decision Tree* C4.5



**Gambar 1. Urutan Metode Penelitian**

akan dilakukan untuk menentukan pola jurusan yang akan diambil oleh siswa sesuai dengan akademik, kemampuan nilai tes, dan minat siswa. Pada proses pohon keputusan yaitu mengubah bentuk data berupa pohon keputusan menjadi model pohon, mengubah model pohon menjadi *rule* dan menyederhanakan *rule* dengan algoritma C 4.5 diharapkan calon siswa yang akan melanjutkan ke SMK dapat melewati hambatan pemilihan jurusan sesuai dengan minat dengan mudah [7].

Sesuai dengan kasus di atas maka penulis bermaksud untuk membuat penelitian yang berjudul “*Klasifikasi Pola Penjurusan Bidang Teknik Informatika Menggunakan Algoritma C4.5*”, studi kasus di SMKN Manonjaya.

## 2. Metode Penelitian

Metode penelitian yang dilakukan adalah metode penelitian eksperimen, dengan tahapan penelitian dalam skema penelitian seperti Gambar 1.

Algoritma C4.5 adalah algoritma yang digunakan untuk membentuk *decision tree*. *Decision tree* dapat diartikan suatu cara untuk memprediksi atau mengklasifikasi yang sangat kuat [8]. Algoritma C4.5 dan pohon keputusan merupakan dua model yang tak terpisahkan, karena untuk membangun sebuah pohon keputusan, dibutuhkan algoritma C4.5 [9]. Pohon keputusan tersebut kemudian diinterpretasikan kedalam bentuk aturan- aturan keputusan yang mudah dipahami dan digunakan sebagai acuan untuk memprediksi [10].

Entropi dan Informasi *Gain* sebuah obyek yang diklasifikasikan dalam pohon harus dites nilai *entropy*-nya. *Entropy* dapat dikatakan sebagai kebutuhan bit untuk menyatakan suatu kelas. Semakin kecil nilai

|     |          |                     |    |    |     |   |   |   |     |
|-----|----------|---------------------|----|----|-----|---|---|---|-----|
| 1   | 18104716 | PUPUT SUKMA         | 75 | 70 | TKJ | F | F | T | TKJ |
| 2   | 18104718 | PUPUT SUKMA         | 65 | 70 | TKJ | F | F | F | TKJ |
| 3   | 18104720 | PUTRI AVNIGA        | 90 | 80 | RPL | T | T | T | RPL |
| 4   | 18104719 | RAFLI SUHAEDIN      | 80 | 75 | RPL | T | F | T | RPL |
| 5   | 18104721 | RAIHAN HIDAYAT      | 80 | 85 | TKJ | T | T | F | TKJ |
| 6   | 18104722 | RALDY ACHMAD CHUS   | 80 | 85 | TKJ | T | T | F | TKJ |
| 7   | 18104723 | RANGGA FAJAR GUMEL  | 50 | 55 | TKJ | F | F | F | TKJ |
| 8   | 18104724 | RANGGA GUMELAR      | 50 | 55 | TKJ | F | F | F | TKJ |
| 9   | 18104725 | RANI APRILIANI      | 70 | 80 | TKJ | F | T | F | TKJ |
| 10  | 18104726 | RANI ANGGRAENI      | 85 | 75 | TKJ | T | F | T | RPL |
| 291 | 18104792 | TASYA SRI MULYANI   | 80 | 60 | RPL | T | F | T | RPL |
| 292 | 18104793 | TATANG RAMDANI      | 85 | 75 | RPL | T | F | T | RPL |
| 293 | 18104794 | TAUFIK HIDAYAT      | 65 | 75 | TKJ | F | F | F | TKJ |
| 294 | 18104795 | TEGUH ADI WIGUNA    | 65 | 75 | TKJ | F | F | F | TKJ |
| 295 | 18104796 | TIA LISMULYANI SARI | 65 | 70 | TKJ | F | F | F | TKJ |
| 296 | 18104797 | TIA RISMA           | 60 | 65 | TKJ | F | F | F | TKJ |
| 297 | 18104798 | TRIANI SRI REZEKI   | 80 | 70 | RPL | T | F | T | RPL |
| 298 | 18104799 | TSANI WULAN         | 70 | 80 | TKJ | F | T | F | TKJ |
| 299 | 18104801 | ULI AULIA           | 90 | 80 | RPL | T | T | T | RPL |
| 300 | 18104800 | UPI MUPLIHAIH       | 75 | 70 | RPL | F | F | T | RPL |
| 301 | 18104802 | USEP RENDI          | 70 | 75 | TKJ | F | F | F | TKJ |

Gambar 2. Dataset

|    |          |                     |    |    |     |   |   |   |     |
|----|----------|---------------------|----|----|-----|---|---|---|-----|
| 1  | 18104716 | PUPUT SUKMA         | 75 | 70 | TKJ | F | F | T | TKJ |
| 2  | 18104718 | PUTRI AVNIGA        | 65 | 70 | TKJ | F | F | F | TKJ |
| 3  | 18104720 | RAFLI RAVENA        | 90 | 80 | RPL | T | T | T | RPL |
| 4  | 18104719 | RAFLI SUHAEDIN      | 80 | 75 | RPL | T | F | T | RPL |
| 5  | 18104721 | RAIHAN HIDAYAT      | 80 | 85 | TKJ | T | T | F | TKJ |
| 6  | 18104722 | RALDY ACHMAD CHUS   | 80 | 85 | TKJ | T | T | F | TKJ |
| 7  | 18104723 | RANGGA FAJAR GUMEL  | 50 | 55 | TKJ | F | F | F | TKJ |
| 8  | 18104724 | RANGGA GUMELAR      | 50 | 55 | TKJ | F | F | F | TKJ |
| 9  | 18104725 | RANI APRILIANI      | 70 | 80 | TKJ | F | T | F | TKJ |
| 10 | 18104726 | RANI ANGGRAENI      | 85 | 75 | TKJ | T | F | T | RPL |
| 76 | 18104792 | TASYA SRI MULYANI   | 80 | 60 | RPL | T | F | T | RPL |
| 77 | 18104793 | TATANG RAMDANI      | 85 | 75 | RPL | T | F | T | RPL |
| 78 | 18104794 | TAUFIK HIDAYAT      | 65 | 75 | TKJ | F | F | F | TKJ |
| 79 | 18104795 | TEGUH ADI WIGUNA    | 65 | 75 | TKJ | F | F | F | TKJ |
| 80 | 18104796 | TIA LISMULYANI SARI | 65 | 70 | TKJ | F | F | F | TKJ |
| 81 | 18104797 | TIA RISMA           | 60 | 65 | TKJ | F | F | F | TKJ |
| 82 | 18104798 | TRIANI SRI REZEKI   | 80 | 70 | RPL | T | F | T | RPL |
| 83 | 18104799 | TSANI WULAN         | 70 | 80 | TKJ | F | T | F | TKJ |
| 84 | 18104801 | ULI AULIA           | 90 | 80 | RPL | T | T | T | RPL |
| 85 | 18104800 | UPI MUPLIHAIH       | 75 | 70 | RPL | F | F | T | RPL |
| 86 | 18104802 | USEP RENDI          | 70 | 75 | TKJ | F | F | F | TKJ |

Gambar 4. Data Training

| No  | Nm   | Nama             | Tes Software | Tes Hardware & Jaringan | Pilihan Jurusan | Software >= 80 | Jaringan >= 80 | Software >= Jaringan | Decision |
|-----|------|------------------|--------------|-------------------------|-----------------|----------------|----------------|----------------------|----------|
| 1   | H000 | ASH AF           | 65           | 70                      | TKJ             | T              | T              | F                    | TKJ      |
| 2   | H000 | ARTAR RAHD       | 80           | 80                      | TKJ             | T              | T              | F                    | TKJ      |
| 3   | H000 | ANDRIAN          | 70           | 80                      | TKJ             | F              | T              | F                    | TKJ      |
| 4   | H000 | ANAN NURHAYAT    | 65           | 80                      | TKJ             | F              | T              | F                    | TKJ      |
| 5   | H000 | ALIA SANGA RAHAR | 80           | 80                      | TKJ             | T              | T              | F                    | TKJ      |
| 200 | H000 | IRPKA LAFITE     | 80           | 80                      | TKJ             | T              | T              | F                    | TKJ      |
| 210 | H000 | SABRIL SYAHAN    | 70           | 60                      | RPL             | F              | F              | T                    | RPL      |
| 220 | H000 | HENDRA OKLEHAR   | 70           | 60                      | RPL             | F              | F              | T                    | RPL      |
| 230 | H000 | RETMARENS        | 85           | 70                      | RPL             | T              | F              | T                    | RPL      |
| 240 | H000 | HENDRA OKLEHAR   | 70           | 60                      | RPL             | F              | F              | T                    | RPL      |
| 250 | H000 | SUNYALIANI       | 85           | 70                      | RPL             | T              | F              | T                    | RPL      |

Gambar 3. Data Testing

entropy, maka akan semakin entropy digunakan dalam mengekstrak suatu kelas. Entropy digunakan untuk mengukur ketidak aslian sistem informasi atau disebut dengan processing system.

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

Keterangan :

S : Himpunan kasus

A : Fitur

n : Jumlah partisi S

pi : Proporsi dari Si terhadap S

Information gain adalah salah satu attribute selection measure yang digunakan untuk memilih test attribute tiap node pada tree. Atribut dengan informasi gain tertinggi dipilih sebagai test atribut dari suatu node. Gain (S,A) merupakan perolehan informasi dari atribut A relative terhadap output data S. Perolehan informasi didapatdari output data atau variable dependent S yang dikelompokkan berdasarkan atribut A, dinotasikan dengan gain (S,A).

### 3. Hasil dan Pembahasan

Gambar 2 menunjukkan Dataset awal dari SMK Negeri Manonjaya yang didapat pada tanggal 3 Mei 2019 dan dataset ini belum diolah dengan teknik missing value.

Setelah data diolah selanjutnya data dibagi menjadi dua yaitu data testing dan data training, adapun data yang akan diuji merupakan data testing yang berjumlah 215 record dari total keseluruhan 301 record data.

Data testing merupakan data yang diambil dari data keseluruhan dengan tujuan untuk diuji dengan metode

tertentu, adapun pengujian yang kami lakukan adalah dengan menggunakan algoritma C4.5, seperti ditunjukkan Gambar 3, sedangkan Data Training adalah data latihan untuk memprediksi kelas data baru yang belum pernah ada (Gambar 4).

### 3.1 Eksperimen dan Pengujian Hasil

Data testing yang disiapkan selanjutnya dikelompokkan berdasarkan atribut-atribut yang ada sehingga terlihat pada Tabel 1 berikut.

Tabel 1. Hasil Perhitungan Node 1 pada Data Testing

| Node 1               | Jumlah | TKJ | RPL | Entropy     | Gain        |
|----------------------|--------|-----|-----|-------------|-------------|
| 1                    | 2      | 3   | 4   | 5           | 6           |
| Total                | 215    | 143 | 72  | 0,919838422 |             |
| Pilihan pertama      |        |     |     |             | 0,738690874 |
| Pil TKJ              | 138    | 137 | 1   | 0,061927389 |             |
| Pil RPL              | 77     | 6   | 71  | 0,394814846 |             |
| Software >= 80       |        |     |     |             | 0,228381403 |
| TRUE                 | 85     | 29  | 56  | 0,92594006  |             |
| FALSE                | 130    | 114 | 16  | 0,538141185 |             |
| Jaringan >= 80       |        |     |     |             | 0,056020043 |
| TRUE                 | 85     | 70  | 15  | 0,672294817 |             |
| FALSE                | 130    | 73  | 57  | 0,989045324 |             |
| Software >= Jaringan |        |     |     |             | 0,628588157 |
| TRUE                 | 89     | 17  | 72  | 0,703582101 |             |
| FALSE                | 126    | 126 | 0   | 0           |             |

Dari Tabel 1 di atas, dapat diketahui bahwa atribut yang memiliki gain tertinggi adalah atribut pilihan jurusan pertama. Hasil tersebut didapatkan sesuai dengan rumus yang dicantumkan sebelumnya yaitu  $0,919838422 - ((138/215 * 0,061927389) + (77/215 * 0,394814846)) = 0,738690874$ . Dengan demikian, atribut pilihan pertama dapat menjadi node akar. Ada dua nilai dari atribut pilihan pertama yaitu Pil TKJ dan Pil RPL.

Dari Tabel 2 di atas, dapat diketahui bahwa atribut yang memiliki gain tertinggi selanjutnya adalah atribut Software >= Jaringan. Hasil tersebut didapatkan sesuai dengan rumus yang dicantumkan sebelumnya yaitu  $0,061927389 - ((18/138 * 0,309543429) + (120/318 * 0,92594006)) = 0,628588157$ .

**Tabel 2. Hasil Perhitungan Node 1.1 pada Data Testing**

| Node 1.1             | Jumlah | TKJ | RPL | Entropy     | Gain        |
|----------------------|--------|-----|-----|-------------|-------------|
| 1                    | 2      | 3   | 4   | 5           | 6           |
| Total                | 138    | 137 | 1   | 0,061927389 |             |
| Software >= 80       |        |     |     |             | 0,016826119 |
| TRUE                 | 28     | 27  | 1   | 0,222284831 |             |
| FALSE                | 110    | 110 | 0   | 0           |             |
| Jaringan >= 80       |        |     |     |             | 0,006547996 |
| TRUE                 | 64     | 64  | 0   | 0           |             |
| FALSE                | 74     | 73  | 1   | 0,103275085 |             |
| Software >= Jaringan |        |     |     |             | 0,021552159 |
| TRUE                 | 18     | 17  | 1   | 0,309543429 |             |
| FALSE                | 120    | 120 | 0   | 0           |             |

0)) = 0,021552159. Dengan demikian, atribut Software ≥ Jaringan dapat menjadi node 1.1., dan ada dua nilai dari atribut tersebut yaitu TRUE dan FALSE.

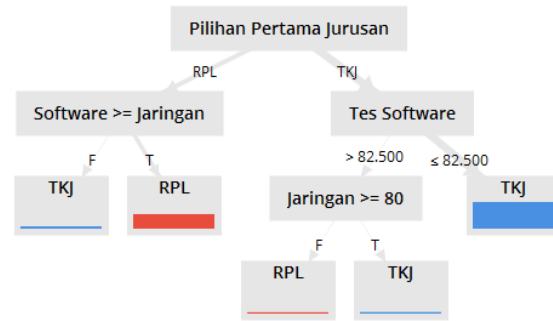
**Tabel 3. Hasil Perhitungan Node 1.2 Pada Data Testing**

| Node 1.2             | Jumlah | TKJ | RPL | Entropy     | Gain        |
|----------------------|--------|-----|-----|-------------|-------------|
| 1                    | 2      | 3   | 4   | 5           | 6           |
| Total                | 77     | 6   | 71  | 0,394814846 |             |
| Software >= 80       |        |     |     |             | 0,044964087 |
| TRUE                 | 57     | 2   | 55  | 0,219297308 |             |
| FALSE                | 20     | 4   | 16  | 0,721928095 |             |
| Jaringan >= 80       |        |     |     |             | 0,159418327 |
| TRUE                 | 21     | 6   | 15  | 0,863120569 |             |
| FALSE                | 56     | 0   | 56  | 0           |             |
| Software >= Jaringan |        |     |     |             | 0           |
| TRUE                 | 71     | 0   | 71  | 0           |             |
| FALSE                | 6      | 6   | 0   | 0           |             |

Dari Tabel 3 di atas dapat diketahui bahwa atribut yang memiliki gain tertinggi selanjutnya adalah atribut Jaringan > 80. Hasil tersebut didapatkan sesuai dengan rumus yang dicantumkan sebelumnya yaitu  $0,394814846 - ((21/77 * 0,863120569) + (56/77 * 0)) = 0,159418327$ . Dengan demikian, atribut Jaringan > 80 dapat menjadi node 1.2, dan ada dua nilai dari atribut tersebut yaitu TRUE dan FALSE. Adapun pohon keputusannya berdasarkan RapidMiner adalah sebagai berikut:

Dari pohon keputusan tersebut didapatkan hasil jumlah siswa baru yang diterima pada ke dalam jurusan TKJ adalah sebanyak 143 siswa dan RPL sebanyak 72 siswa. Adapun rule nya akan dijelaskan pada Gambar 6.

Gambar 6 adalah hasil dari *decision tree data testing*, yang menghasilkan jika jurusan pertamanya RPL dan nilai *software*nya lebih kecil dari pada jaringan ada 6 orang yang harusnya masuk TKJ dan nilai *software*nya lebih besar dari jaringan ada 71 orang yang cocok masuk RPL. Sedangkan jika pilhan pertamanya TKJ dan nilai *software*nya lebih dari 82.500 dan nilai test jaringannya



**Gambar 5. Pohon Keputusan**

**Tree**

```

Pilihan Pertama Jurusan = RPL
| Software >= Jaringan = F: TKJ {TKJ=6, RPL=0}
| Software >= Jaringan = T: RPL {TKJ=0, RPL=71}
Pilihan Pertama Jurusan = TKJ
| Tes Software > 82.500
| | Jaringan >= 80 = F: RPL {TKJ=0, RPL=1}
| | Jaringan >= 80 = T: TKJ {TKJ=2, RPL=0}
| Tes Software ≤ 82.500: TKJ {TKJ=135, RPL=0}
    
```

**Gambar 6. Rule Decision Tree**

kurang dari 80 ada 1 orang yang harusnya masuk RPL dan ada 2 orang yang nilai jaringannya lebih besar dari 80 yang pas masuk TKJ. Lalu yang test *software*nya dibawah 82.500 terdapat 135 orang.

**3.2 Evaluasi dan Validasi Hasil**

Dari pohon keputusan dan *rule* yang terbentuk dimana status RPL dan TKJ dilihat dari atribut yang ada, dan Tabel 4 berikut menampilkan rincian hasil pengujiannya.

**Tabel 4. Hasil Pengujian Dari Data Testing**

| No | Rule  | Jumlah |     |
|----|---|--------|-----|
|    |   | TKJ    | RPL |
| 1  | If pilihan pertama TKJ and Software ≥ hardware = F than TKJ | 137    | -   |
| 2  | If pilihan pertama RPL and Software ≥ hardware = T than RPL | -      | 71  |
| 3  | If pilihan pertama TKJ and Software ≥ hardware = T than RPL | -      | 1   |
| 4  | If pilihan pertama RPL and Software ≥ hardware = F than TKJ | 6      | 1   |

Dari hasil pengujian yang dilakukan didapatkan hasil jumlah siswa baru yang diterima pada ke dalam jurusan TKJ adalah sebanyak 143 siswa dan RPL sebanyak 72 siswa.

accuracy: 96.73% +/- 4.50% (micro average: 96.74%)

|              | true TKJ | true RPL | class precision |
|--------------|----------|----------|-----------------|
| pred TKJ     | 137      | 1        | 99.28%          |
| pred RPL     | 5        | 71       | 92.21%          |
| class recall | 95.80%   | 98.61%   |                 |

**Gambar 7. Tabel Confusion Matrix Algoritma C4.5**

Hasil akhir dari pengujian pada data testing selanjutnya diuji tingkat akurasi menggunakan aplikasi Rapid Miner, adapun hasil tingkat akurasi klasifikasi data testing yang didapatkan adalah seniai 96,73%.

#### 4. Kesimpulan

Dengan menggunakan algoritma C4.5 dalam penentuan klasifikasi pemilihan jurusan di SMKN Manonjaya maka proses pemilihan jurusan dapat dilakukan dengan cepat, klasifikasi ini telah melalui proses pengujian dengan memiliki akurasi sebesar 96.73%.

#### Daftar Pustaka

- [1] L. N. Rani, "Klasifikasi nasabah menggunakan algoritma c4. 5 sebagai dasar pemberian kredit," *INOVTEK Polbeng-Seri Informatika*, vol. 1, no. 2, pp. 126–132, 2016.
- [2] E. B. Sambani and F. Nuraeni, "Penerapan algoritma c4. 5 untuk klasifikasi pola penjurusan di sekolah menengah kejuruan (smk) kota tasikmalaya," *CSRID (Computer Science Research and Its Development Journal)*, vol. 9, no. 3, pp. 149–157, 2017.
- [3] F. Nasari, P. Keputusan, and D. Selection, "Penerapan algoritma c4. 5 dalam pemilihan bidang peminatan program studi sistem informasi di stmik potensi utama medan," *KOMITE PROGRAM*, vol. 12, p. 30, 2014.
- [4] S. Sumpena, Y. Akbar, and N. Nirat, "Penerimaan calon siswabarur dan penentuan penjurusan dengan algoritma c 4.5 smk plus pgri 1 cibinong," *CKI ON SPOT*, vol. 11, no. 2, pp. 181–191, 2018.
- [5] R. Rismayanti, "Decision tree penentuan masa studi mahasiswa prodi teknik informatika (studi kasus: Fakultas teknik dan komputer universitas harapan medan)," *Query: Journal of Information Systems*, vol. 2, no. 1, pp. 16–24, 2018.
- [6] E. D. S. Mulyani, A. Rihadisha, N. Saputri et al., "Klasifikasi penentuan kelayakan pemberian kredit menggunakan metode naive bayes classifier (studi kasus: Koperasi simpan pinjam simpanan pameungkeut banda tasikmalaya)," *Jurnal VOI (Voice Of Informatics)*, vol. 9, no. 2, pp. 81–92, 2020.
- [7] A. Novandya, "Penerapan algoritma klasifikasi data mining c4. 5 pada dataset cuaca wilayah bekasi," *Konferensi Nasional Ilmu Sosial dan Teknologi*, vol. 1, no. 1, pp. 368–372, 2017.
- [8] H. D. Darmawan, D. Yuniarti, and Y. N. Nasution, "Klasifikasi lama masa studi mahasiswa menggunakan perbandingan metode algoritma c. 45 dan algoritma classification and regression tree," *EKSPONENSIAL*, vol. 8, no. 2, pp. 151–160, 2017.
- [9] E. Hasmin and S. Aisa, "Penerapan algoritma c4. 5 untuk penentuan penerima beasiswa mahasiswa," *CogITO Smart Journal*, vol. 5, no. 2, pp. 308–320, 2019.
- [10] E. W. Perdani and K. H. Al Fatta, "Pengujian kinerja algoritma c4. 5 untuk prediksi pemilihan minat lanjutan pada siswa smk," *JURNAL TRANSFORMASI*, vol. 13, no. 2, pp. 91–97, 2018.