



### Penerapan Metode Klasifikasi Dengan Algoritma *Decision Tree* C4.5

#### Untuk Mendiagnosa Awal Penyakit Ginjal Kronis

**Karina Imelda<sup>1</sup>**

<sup>1</sup>Teknik Informatika, Fakultas Teknik, Universitas Pelita Bangsa

<sup>1</sup>karina.imelda@pelitabangsa.ac.id

#### *Abstract*

*Patients with kidney disease find it difficult to know because they have to go through a series of laboratory tests with a considerable amount of time at the hospital. The complexity of the detecting process can be made easier using technology with data processing or Data Mining. Data Mining is the process of mining or discovering new information that aims to overcome certain conditions by looking for certain patterns and rules of a large amount of data. To diagnose early patients with chronic kidney disease with Data Mining using the classification method with the Decision Tree C4.5 algorithm. Decision Tree or meaning a decision tree is a prediction model with an hierarchical structure that has the concept of converting data into rules and decision trees, data in decision trees are expressed in tables with attributes and records that state parameters as tree formation criteria. The study used Chronic Kidney Disease data as a dataset and applied the classification method with the Decision Tree C4.5 algorithm. This study uses RapidMiner 9.0.3 data mining tools. The results obtained from this study show an accuracy of 89.05%.*

**Keywords:** *Data Mining, Chronic Kidney Disease, Decision Tree, Algorithm C4.5.*

#### **Abastrak**

Penderita penyakit ginjal sulit untuk diketahui karena harus melewati serangkaian proses pemeriksaan laboratorium dengan waktu yang cukup lama dirumah sakit. Rumitnya proses mendeteksi bisa dipermudah menggunakan teknologi dengan pengolahan data atau *Data Mining*. *Data Mining* merupakan proses penambangan atau penemuan informasi baru yang bertujuan dapat mengatasi kondisi tertentu dengan mencari pola dan aturan tertentu dari sejumlah data dalam jumlah besar. Untuk mendiagnosa awal penderita penyakit ginjal kronis dengan *Data Mining* menggunakan metode klasifikasi dengan algoritma *Decision Tree* C4.5. *Decision Tree* atau yang berarti pohon keputusan merupakan suatu model prediksi dengan struktur berhirarki yang memiliki konsep untuk mengubah data menjadi

aturanaturan dan pohon keputusan, data dalam pohon keputusan dinyatakan ke dalam bentuk tabel dengan atribut dan record yang menyatakan parameter sebagai kriteria pembentukan pohon. Penelitian menggunakan data *Chronic Kidney Disease* sebagai dataset dan menerapkan metode klasifikasi dengan algoritma *Decision Tree* C4.5. Penelitian ini menggunakan *data mining tools RapidMiner* 9.0.3. Hasil yang didapat dari penelitian ini menunjukkan akurasi sebesar 89,05%.

**Kata kunci:** *Data Mining, Penyakit Ginjal Kronis, Decision Tree, Algoritma C4.5.*

#### **1. Pendahuluan**

Kesehatan merupakan kebutuhan pokok bagi setiap orang. Perubahan kondisi tertentu pada tubuh menjadi pengaruh besar bagi kesehatan dan dapat

memicu datangnya berbagai macam penyakit. Beberapa klasifikasi penyakit antara lain penyakit menular, tidak menular, dan kronis. Salah satu penyakit kronis yaitu penyakit ginjal. Ginjal adalah organ penting dalam tubuh manusia. Jika terdapat gangguan atau kerusakan pada ginjal maka dapat memberikan komplikasi pada tubuh. Penderita ginjal kronis mencapai 0,38 % di Indonesia. Indonesia menjadi salah satu negara dengan penderita penyakit ginjal kronis yang meningkat setiap tahunnya.[1]

Pada awalnya, penyakit ginjal tidak menunjukkan gejala yang khas sehingga penyakit ini sering terlambat diketahui. Tanda dan gejala yang timbul karena penyakit Ginjal sangat umum dan dapat ditemukan pada penyakit lain seperti tekanan darah tinggi, perubahan frekuensi buang air kecil dalam sehari, adanya darah dalam urin, mual dan muntah serta bengkak, terutama pada kaki dan pergelangan kaki. Saat ini para penderita penyakit ginjal sulit untuk mengetahui penyakit yang dideritanya, karena untuk mengetahuinya harus melewati seangkaian proses pemeriksaan laboratorium di rumah sakit yang rumit dan memerlukan waktu yang cukup lama hingga hasil pemeriksaannya keluar.[2]

Deteksi dini untuk penyakit ginjal yang dirasa cukup sulit dan apabila pasien telah sampai pada gagal ginjal maka selanjutnya adalah pengobatan yang menjadi sangat sulit dilakukan. Fungsi ginjal pada penderita gagal ginjal tidak mampu untuk membuang sisa-sisa zat dari metabolisme tubuh dan dalam menyeimbangkan jumlah cairan yang terdapat didalam tubuh juga sulit.[3]

Rumitnya proses dalam mendeteksi suatu penyakit bisa kita permudah dengan peran serta teknologi didalamnya, salah satu caranya ialah dengan mengolah suatu data menggunakan metode agar dapat menghasilkan sebuah keputusan atau yang disebut *Data Mining*. Data mining merupakan sebuah proses otomatisasi untuk mencari data dalam sebuah database dimana data tersebut berskala besar yang mampu menghasilkan informasi baru yang dapat dimanfaatkan.[4]

Pemilihan Algoritma yang tepat sangat mempengaruhi pada hasil dari proses *Data Mining*, maka dari itu penulis memilih metode klasifikasi dan algoritma *decision tree* C4.5 dalam penelitian. Banyak teknik klasifikasi di Data Mining salah satunya adalah algoritma *decision tree* C4.5 yang menyajikan klasifikasi data kedalam bentuk pohon

keputusan. Klasifikasi adalah proses menempatkan objek atau konsep tertentu kedalam satu set kategori, berdasarkan sifat objek atau konsep yang bersangkutan.[5] Kelebihan pada algoritma C4.5 memiliki tingkat akurasi yang baik dengan membuat pohon keputusan yang efisien dalam menangani atribut tipe diskrit dan tipe diskritnumerik, mudah untuk diinterpretasikan. Kelemahan Algoritma C4.5 terdapat di skalabilitas yaitu data training hanya dapat digunakan dan disimpan secara keseluruhan pada waktu yang bersamaan di memori.[6]

Dari pembahasan diatas maka penulis mengusulkan menggunakan *Data Mining* untuk “Penerapan Metode Klasifikasi Dengan Algoritma *Decision Tree* C4.5 Untuk Mendiagnosa Awal Penyakit Ginjal Kronis”, penghitungan tingkat akurasi dilakukan dengan menggunakan aplikasi *RapidMiner*.

**2. Metode Penelitian**

Penelitian kali ini merupakan *dataset* yang didapat dari UCI *Machine Learning Repository* pada bagian *Chronic Kidney Disease Dataset* yang terdiri dari 25 atribut didalamnya dapat dilihat dalam tabel 1.

Tabel 1. Atribut *Dataset*

No	Atribut	No	Atribut
1	<i>Age (age)</i>	14	<i>Potassium (pot)</i>
2	<i>Blood Pressure (bp)</i>	15	<i>Hemoglobin (hemo)</i>
3	<i>Specific Gravity (sg)</i>	16	<i>Packed Cell Volume (pcv)</i>
4	<i>Albumin (al)</i>	17	<i>White Blood Cell Count (wbcc)</i>
5	<i>Sugar (su)</i>	18	<i>Red Blood Cell Count (rbcc)</i>
6	<i>Red Blood Cells (rbc)</i>	19	<i>Hypertension (htn)</i>
7	<i>Pus Cell (pc)</i>	20	<i>Diabetes Mellitus (dm)</i>
8	<i>Pus Cell clumps (pcc)</i>	21	<i>Coronary Artery Disease (cad)</i>
9	<i>Bacteria (ba)</i>	22	<i>Appetite (appet)</i>
10	<i>Blood Glucose (bgr)</i>	23	<i>Pedal Edema (pe)</i>
11	<i>Blood Urea (bu)</i>	24	<i>Anemia (ane)</i>
12	<i>Serum Creatinine (sc)</i>	25	<i>Class (kelas)</i>
13	<i>Sodium (sod)</i>		

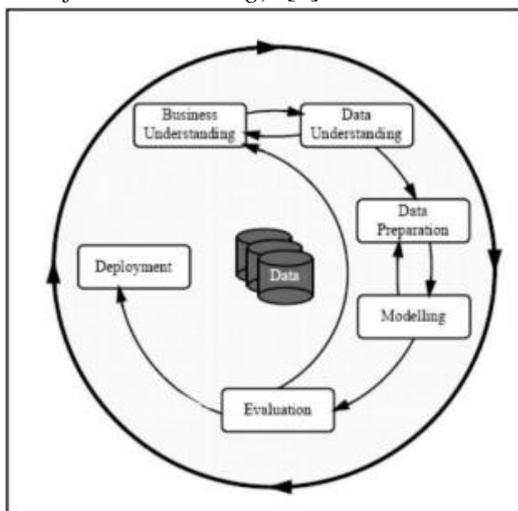
Pendekatan yang digunakan dalam penelitian ini ialah pendekatan kuantitatif yang berarti pendekatan ini menggunakan skala numerik yang

berbasis pola alur kumpulan teori hingga menghasilkan konsep yang kemudian merumuskan hipotesis, melakukan uji hipotesis, lalu menarik kesimpulan yang banyak menggunakan angka diawali dari pengumpulan data, pengertian pada data tersebut hingga menampilkan hasil. Jenis penelitian yang digunakan merupakan penelitian eksperimental dimana penelitian ini bersifat uji coba, juga mempengaruhi hal-hal yang berkaitan dengan seluruh variabel maupun atribut yang melibatkan evaluasi dan pengembangan. Metode pengumpulan data yang digunakan dalam penelitian ini adalah metode pengumpulan data sekunder. Data yang digunakan bersumber dari data publik yaitu UCI *Machine Learning Repository* pada bagian *Chronic Kidney Disease* yang diperoleh dari *GitHub* sebanyak 337 data.[7]

Tabel 2. *Dataset* Ginjal Kronis

age	bp	sg	al	su	...
48	80	1.02	1	0	...
7	50	1.02	4	0	...
62	80	1.01	2	3	...
48	70	1.005	4	0	...
51	80	1.01	2	0	...
60	90	1.015	3	0	...
68	70	1.01	0	0	...
52	100	1.015	3	0	...
53	90	1.02	2	0	...
50	60	1.02	2	4	...
...	...	...	...	...	...

Teknik analisis data pada penelitian ini menggunakan metodologi pengembangan data mining yaitu CRISP-DM (*Cross Industry Standard Process for Data Mining*). [8]



Gambar 1. Siklus CRISP-DM

- Business understanding* digunakan untuk memahami permasalahan pada diagnosa awal penyakit ginjal kronis yang menggunakan data mining agar dapat menentukan pola dan tujuan. Hasil diagnosa digunakan sebagai acuan agar penderita penyakit ginjal bisa ditangani sejak gejala awal sehingga dapat meminimalisir tingkat kematian.
- Data understanding* bertujuan untuk mengetahui data yang akan di gunakan pada penelitian. Data yang digunakan pada penelitian ini merupakan data public yaitu UCI *Machine Learning Repository* pada bagian *Chronic Kidney Diseases*.
- Data preparation* digunakan untuk memanipulasi data mentah untuk dianalisa tingkat keberhasilannya. Tujuan lain dari data *preparation* yaitu meningkatkan kualitas data. Pada tahap ini peneliti menggunakan tahap KDD yaitu data *preprocessing* dengan tahap data *cleaning* dan data *reduction*. [9]

1) *Data Reduction* adalah teknik yang digunakan untuk mengurangi representasi data yang digunakan dengan tetap mempertahankan keaslian data. Setelah dilakukan reduksi maka atribut terpilih yang digunakan dalam penelitian ini terdapat dalam tabel dibawah ini.

Tabel 3. Atribut yang digunakan

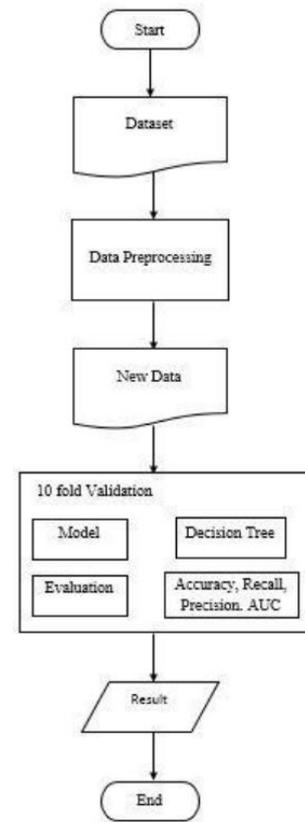
No	Atribut	Satuan
1	Age	Tahun
2	Ane	Ya/Tidak
3	Dm	Ya/Tidak
4	Htn	Ya/Tidak
5	Class	Ckd/not ckd

*Data Reduction* dilakukan menggunakan operator *select attribute* pada *Rapidminer*.

- Pada *dataset Chronic Kidney Disease* dilakukan tahapan data *transformation*, yaitu dengan mengklasifikasi bentuk data atribut *age*.
- Data cleaning* adalah data yang terdapat isian-isian yang tidak sempurna seperti *missing* data, data tidak valid, kesalahan pengetikan,

atribut tidak relevan yang seharusnya dibuang karena akan mengurangi kualitas atau akurasi dari hasil.

- d. Tahap *modeling* menentukan teknik data *mining* dengan menggunakan metode dan algoritma untuk mengolah data yang siap digunakan. Metode yang dipakai dalam penelitian ini adalah metode klasifikasi dengan algoritma *decision tree* C4.5 menggunakan tools *RapidMiner*. Dilakukan validasi data untuk mengetahui tingkat keberhasilan dalam pengujian data menggunakan *10 fold cross validation*. Dalam *10 fold cross validation* dilakukan pengujian sebanyak 10 kali pada 10 iterasi.
- e. *Evaluation* pada penelitian yang dilakukan menggunakan *confusion table* yang akan menghasilkan *accuracy*, *precision*, dan *recall*, serta menggunakan ROC yang akan menghasilkan AUC untuk tahap evaluasi.
- f. *Deployment* atau pembentukan model dilakukan dengan analisa dan pengukuran pada tahap sebelumnya, pada tahap ini diterapkan model atau rule yang paling akurat dan selanjutnya dapat digunakan untuk mengevaluasi data baru.



Gambar 2. Model yang diusulkan

Alur penelitian ini dimulai dari pengambilan *dataset Chronic Kidney Disease* dari *GitHub*. Lalu dilakukan proses *preprocessing* dengan melewati tiga tahap yang pertama untuk memilih atau menyeleksi atribut yang akan digunakan dengan melakukan data *reduction*, tahap kedua melakukan proses data *transformation* dimana data dengan atribut *age* diklasifikasikan, kemudian melakukan tahap ketiga yaitu menghilangkan data *missing value* pada tahap *cleaning*. Setelah melewati proses *preprocessing* maka dihasilkan data yang sudah matang dan siap diolah. Data tersebut kemudian di proses validasi data menggunakan metode *10 fold cross validation* agar mengetahui model algoritma yang telah dibuat berapa besar tingkat keberhasilannya.[10] Dalam proses *10 fold cross validation* data dibagi menjadi 2 yaitu data *training* dan data *testing*. Data *training* digunakan untuk membuat model atau pola dengan algoritma *decision tree* untuk selanjutnya dilakukan proses evaluasi menggunakan data *testing* yang akan menghasilkan *accuracy*, *recall*, *precision*, dan AUC. Tahapan penelitian dapat dilihat pada gambar 2.

Dalam pelaksanaan penelitian diperlukan peralatan sebagai pendukung terlaksananya penelitian berupa perangkat keras (*hardware*) juga perangkat lunak (*software*). Adapun yang dibutuhkan antara lain :

- 1) Kebutuhan *Software* :
  - a. *Operating system Windows 7* versi 64 bit
  - b. *RapidMiner* versi 9.0.3
- 2) Kebutuhan *Hardware* :
  - a. Laptop *intel Pentium*
  - b. RAM 6 GB
  - c. Hardisk 300 GB
  - d. *Printer*

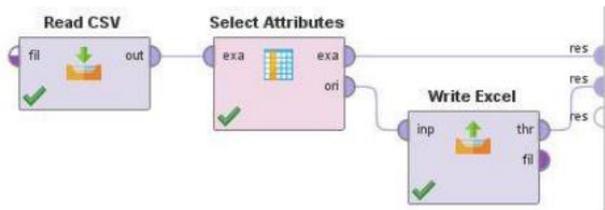
### 3. Hasil dan Pembahasan

Perhitungan *Decision Tree* C4.5 dilakukan dengan mulai menghitung menggunakan *dataset Chronic Kidney Disease* yang telah melewati tahap *preprocessing* hingga menghasilkan 5 atribut dan 337 data. Atribut yang siap dilakukan perhitungan dengan algoritma *Decision Tree* C4.5, terdapat 5 atribut dengan 1 atribut yang menjadi atribut label.

Tabel 4. Atribut Pengujian

age	htn	dm	ane	class
a/p/s/ot/otp	yes/no	yes/no	yes/no	ckd/notckd

Pada pengujian ini diawali dengan melakukan *import dataset*, proses *preprocessing* memakai fitur *select attributes*, data *transformation* dan *replace missing value*. Algoritma *Decision Tree C4.5* digunakan untuk membuat model dalam pengujian, terdapat 2 tahapan proses pengujian menggunakan tool *RapidMiner*, berikut adalah gambar dari pengujian :



Gambar 3. Proses *RapidMiner* 1

- a. Proses *Import Dataset*  
*Dataset Chronic Kidney Disease* termasuk kedalam format CSV sehingga dalam melakukan pengujian dilakukan *import dataset* dengan menggunakan fitur *Read CSV*. *Dataset* setelah di import menunjukkan 337 data.
- b. Proses *Select Atribut*  
 Setelah *dataset* dilakukan *import* tahapan berikutnya melakukan proses *preprocessing* dengan menggunakan fitur *select attributes* untuk menyeleksi atribut yang akan digunakan dalam pengujian. Dari ke 25 atribut yang tersedia penelitian ini hanya memakai 5 atribut untuk proses selanjutnya.

ExampleSet (337 examples, 1 special attribute, 4 regular attributes)

Row No.	class	age	htn	dm	ane
1	ckd	48	yes	yes	no
2	ckd	7	no	no	no
3	ckd	62	no	yes	yes
4	ckd	48	yes	no	yes
5	ckd	51	no	no	no
6	ckd	60	yes	yes	no
7	ckd	68	no	no	no
8	ckd	52	yes	yes	yes
9	ckd	53	yes	yes	yes
10	ckd	50	yes	yes	yes

Gambar 4. Hasil *Select Attributes*

- c. Proses *Data Transformation*  
 Proses data *transformation* bertujuan untuk *dataset* yang telah melewati proses seleksi disimpan ke dalam format baru, agar bisa dilakukan proses klasifikasi pada atribut *age* untuk bisa dilakukan perhitungan. Proses data *transformation*

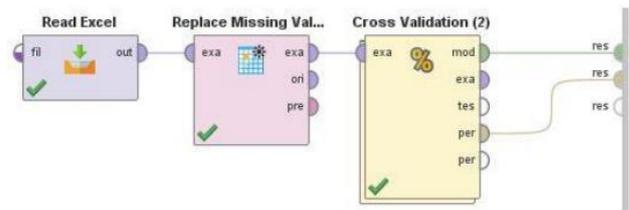
dilakukan dengan menggunakan fitur *write excel*.

ExampleSet (337 examples, 1 special attribute, 4 regular attributes)

Row No.	class	age	htn	dm	ane
1	ckd	p	yes	yes	no
2	ckd	a	no	no	no
3	ckd	p	no	yes	yes
4	ckd	p	yes	no	yes
5	ckd	p	no	no	no
6	ckd	p	yes	yes	no
7	ckd	s	no	no	no
8	ckd	p	yes	yes	yes
9	ckd	p	yes	yes	yes
10	ckd	p	yes	yes	yes

Gambar 5. *Dataset* yang sudah ditransformasi

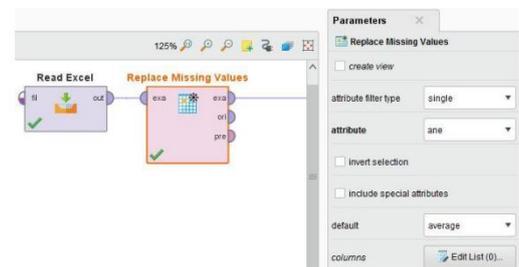
Setelah tahapan proses *RapidMiner* 1 telah semua dilakukan maka berikutnya tahapan proses *RapidMiner* 2. Berikut adalah gambar tahapannya :



Gambar 3. Proses *RapidMiner* 2

Pada tahapan proses *RapidMiner* 2 melakukan *Replace missing value* dan melakukan proses validasi.

- a) *Replace missing value*  
 Fitur ini berguna untuk mengurangi atau menghilangkan *record* yang *missing* yang terdapat didalam atribut. Berikut ini merupakan gambar *replace missing value* :



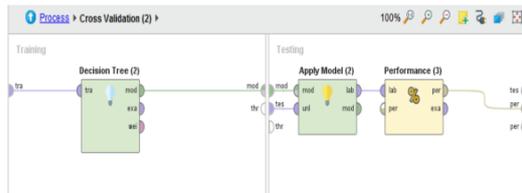
Gambar 4. *Replace Missing Value*

Pada kolom parameter, *attribute filter by type* pilih *single* karena pada proses *replace missing value* berfokus pada 1 atribut saja yaitu atribut *ane*, pada *attribute* pilih *ane*, pada *default* pilih *average* yang bertujuan untuk mengisi *record* yang *missing* dengan *record* rata-rata pada atribut. Fitur *replace missing value* digunakan dalam pengujian karena terdapat 1 record yang mengalami *missing*

dalam atribut ane. *Dataset Chronic Kidney Disease* telah melewati semua proses tahapan Preprocessing yang dibutuhkan yaitu *select attribute*, data *transformation* dan *replace missing value* hingga menghasilkan data yang siap untuk dilakukan tahap validasi.

b) Proses Validasi

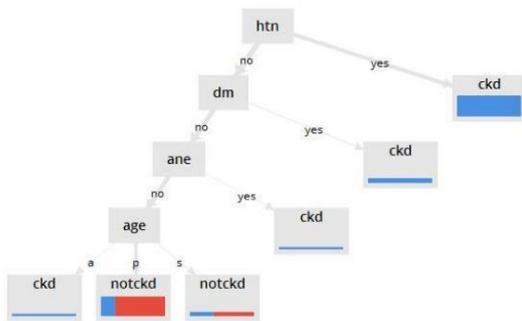
Validasi dilakukan untuk memeriksa seberapa besar tingkat keberhasilan model yang telah dibuat. Validasi pada penelitian ini menggunakan fitur *Cross Validation* dengan menginput nilai *k = 10 folds* pada parameter.



Gambar 5. Model *Cross Validation*

Pada tahap permodelan *Cross Validation* diatas menunjukkan bagian dalamnya dibagi menjadi 2 bagian yaitu data *training* dan data *testing*. Dalam bagian *training* hanya terdapat 1 fitur didalamnya yaitu menggunakan fitur *decision tree C4.5*. Sedangkan pada bagian *testing* terdapat 2 fitur didalamnya yaitu *apply model* dan *confusion table*. Untuk mengaplikasikan model menggunakan fitur *apply model*, setelah model diaplikasikan maka untuk menampilkan *confusion table* yang berisi *accuracy*, *recall*, *precision* dan *AUC* menggunakan fitur *performance*.

Berikut ini merupakan gambar dari hasil pemodelan *decision tree C4.5* :



Gambar 6. Pemodelan *Decision Tree C4.5*

Berdasarkan gambar diatas model menunjukkan beberapa *rule* dengan hasil (*Chronic Kidney Disease*) *ckd* dan *rule* dengan hasil *notckd* (*Not Chronic Kidney*

*Disease*). Adapun deskripsi dalam *RapidMiner* dari pemodelan diatas adalah sebagai berikut :

Tree

```

htn = no
| dm = no
| | ane = no
| | | age = a: ckd {ckd=9, notckd=0}
| | | age = p: notckd {ckd=29, notckd=101}
| | | age = s: notckd {ckd=8, notckd=14}
| | ane = yes: ckd {ckd=9, notckd=0}
| dm = yes: ckd {ckd=27, notckd=0}
htn = yes: ckd {ckd=140, notckd=0}
    
```

Gambar 7. Deskripsi Pemodelan *Decision Tree C4.5*

Adapun hasil dari pemodelan nya sebagai berikut :

1. Akan menghasilkan *rule* positif apabila :
  - a. Hipertensi = yes
  - b. Diabetes mellitus = yes
  - c. Anemia = yes
  - d. Age = a
2. Akan menghasilkan *rule* negatif apabila :
  - a. Age = p
  - b. Age = s

Berikut ini merupakan hasil *accuracy*, *recall*, *precision*, dan *AUC* dari *confusion table* dengan menggunakan algoritma *decision tree C4.5* sebagai berikut:

accuracy: 89.05% +/- 7.91% (micro average: 89.02%)			
	true ckd	true notckd	class precision
pred. ckd	185	0	100.00%
pred. notckd	37	115	75.66%
class recall	83.33%	100.00%	

Gambar 8. Hasil *Accuracy*

Berdasarkan gambar diatas menunjukkan bahwa hasil *accuracy* yaitu sebesar 89,05%. Hasil *recall* penelitian ini sebesar 100%. Pada *class ckd* menghasilkan *class recall* sebesar 83,33% sedangkan pada *class not ckd* menghasilkan *class recall* sebesar 100%.

recall: 100.00% +/- 0.00% (micro average: 100.00%) (positive class: notckd)			
	true ckd	true notckd	class precision
pred. ckd	185	0	100.00%
pred. notckd	37	115	75.66%
class recall	83.33%	100.00%	

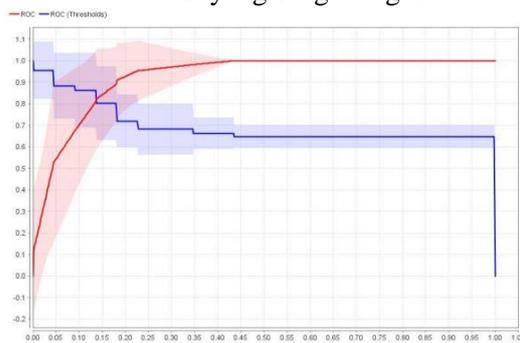
Gambar 9. Hasil *Recall*

Hasil *precision* penelitian ini sebesar 77,96%. Pada *class ckd* menghasilkan *class precision* sebesar 100% sedangkan pada *class not ckd* menghasilkan *class precision* sebesar 75,66%.

precision: 77.96% +/- 12.92% (micro average: 75.66%) (positive class: notckd)			
	true ckd	true notckd	class precision
pred. ckd	185	0	100.00%
pred. notckd	37	115	75.66%
class recall	83.33%	100.00%	

Gambar 10. *Precision*

Hasil pengujian pada algoritma *Decision Tree* menghasilkan nilai AUC sebesar 0,925. Nilai yang dihasilkan menunjukkan bahwa *excellent classification* yang berarti memiliki akurasi yang sangat bagus.



Gambar 11. AUC

Penelitian menggunakan metode klasifikasi algoritma *Decision Tree C4.5* pada *RapidMiner 9.0.3* dengan tahap *preprocessing* yaitu menggunakan fitur *select attributes*, *data transformation* dan *replace missing value*. Fitur *select attribute* digunakan untuk menyeleksi atribut terpilih sebagai proses penelitian. Fitur *data transformation* digunakan untuk merubah data agar dapat diklasifikasikan. Fitur *replace missing value* digunakan untuk mengganti *record* yang tidak memiliki nilai. Berikut ini hasil pengujian dapat dilihat pada tabel berikut :

Tabel 5. Hasil Pengujian

Accuracy	Recall	Precision	AUC
89,05%	100%	77,96%	0,925

Penelitian ini menunjukkan bahwa algoritma *Decision Tree C4.5* baik dalam akurasi dan menghasilkan akurasi sebesar 89,05%. Hasil ini dapat digunakan untuk membantu tenaga medis dalam mendiagnosa awal penyakit ginjal kronis.

#### 4. Kesimpulan

Berdasarkan hasil pengujian yang dilakukan, hasil akurasi yang didapat sebesar 89,05% sehingga dapat disimpulkan bahwa metode klasifikasi dengan algoritma *Decision Tree C4.5* dapat digunakan untuk mendiagnosa awal penderita penyakit ginjal kronis.

#### Referensi

[1] Andriani, A. (2023). Sistem Pendukung Keputusan Berbasis Decision Tree Alam Pemberian Beasiswa Studi Kasus : Amik “ Bsi Yogyakarta ,” 2022(Sentika).

[2] Chimwayi, K. B., Haris, N., Caytiles, R. D., & Iyengar, N. C. S. N. (2020). Risk Level Prediction of Chronic Kidney Disease Using Neuro- Fuzzy and Hierarchical Clustering Algorithm ( s ) Risk Level Prediction of ChronicKidney Disease Using Neuro-, 12(8), 23–36.

[3] Elisa, E. (2020). Analisa dan Penerapan Algoritma C4.5 Dalam Data Mining Untuk Mengidentifikasi Faktor-Faktor Penyebab Kecelakaan Kerja Kontruksi PT.Arupadhatu Adisesanti. *Jurnal Online Informatika*, 2(1), 36. <https://doi.org/10.15575/join.v2i1.71>

[4] Hadiri, A. D. (2022). Penerapan Particle Swarm Optimization Untuk Seleksi Atirbut Pada Metode Decision Tree C 4 . 5 Untuk Persetujuan Abstract : Bad credit is one of the credit risk faced by the financial and banking industry . Improved accuracy of credit ratings can be done b.

[5] Henny Leidiyana. (2023). Penerapan algoritma k-nearest neighbor untuk penentuan resiko kredit kepemilikan kendaraan bermotor, 1(1), 65–76.

[6] Idaiani, S., Andayasari, L., Widowati, L., Gitawati, R., Sihombing, M., & Tjahja, I. (2024). Faktor Risiko Penyakit Ginjal Kronik : Studi Kasus Kontrol di Empat Rumah Sakit di Jakarta Tahun 2014, 17–26.

[7] Kurniawan, M. Y., & Rosadi, M. E. (2021). Optimasi Decision Tree Menggunakan Particle Swarm Optimization Pada Data Siswa Putus Sekolah, 15–22.

[8] Luvia, Y. S., Windarto, A. P., Solikhun, S., & Hartama, D. (2021). Penerapan Algoritma C4.5 Untuk Klasifikasi Predikat Keberhasilan Mahasiswa Di Amik Tunas Bangsa. *Jurasik (Jurnal Riset Sistem Informasi Dan Teknik Informatika)*, 1(1), 75. <https://doi.org/10.30645/jurasik.v1i1.12>

[9] Sucipto, H., Kusriani, & Al fatta, H. (2021). Sistem pendukung keputusan penerimaan siswa baru dan penjurusan pada sekolah menengah kejuruan (smk), 5(2).

- [10] Yanto, R. (2019). Implementasi Data Mining Estimasi Ketersediaan Lahan Pembuangan Sampah menggunakan Algoritma Regresi Linear, 2(1), 361– 366.