PREDIKSI GAGAL GINJAL KRONIK MENGGUNAKAN *IMPROVED* C4.5

SKRIPSI

OLEH
MOCH. NOOR AFFAN ANSHORI
NIM. H72214016

PROGRAM STUDI MATEMATIKA
JURUSAN SAINS
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA
SURABAYA
2018

PERNYATAAN KEASLIAN

Yang bertandatangan dibawah ini:

Nama : Moch. Noor Affan Anshori

NIM : H72214016

Program Studi : Matematika

Angkatan : 2014

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi saya yang berjudul: **Prediksi Gagal Ginjal Kronik Menggunakan *Improved* C4.5**. Apabila suatu saat nanti terbukti saya melakukan tindakan plagiat, maka saya akan menerima sanksi yang telah ditetapkan. Demikian pernyataan keaslian ini saya buat dengan sebenar-benarnya.

Surabaya, 3 Agustus 2018

METERAI TEMPEL

03DDDAFF253070167

6000

ENAM RIBU RUPIAH

Moch. Noor Affan Anshori

NIM.H72214016

LEMBAR PENGESAHAN

PREDIKSI GAGAL GINJAL KRONIK MENGGUNAKAN *IMPROVED* C4.5

Disusun oleh
Moch. Noor Affan Anshori
NIM.H72214016

Telah dipertahankan di depan Dewan Penguji
Pada tanggal 25 Juli 2018
Dan dinyatakan telah memenuhi syarat
untuk memperoleh gelar
Sarjana Matematika (S.Mat)

Dewan Penguji

Penguji I

Nurissaidah Ulinnuha, M.Kom
NIP.199011022014032004

Penguji II

Dian C. Rini Novitasari, M.Kom
NIP.198511242014032001

Penguji III

Ahmad Hanif Asyhar, M.Si
NIP.198601232014031001

Penguji IV

Putroue Keumala Intan, M.Si
NIP.198805282018012001

Mengesahkan
Dekan Fakultas Sains dan Teknologi
UIN Sunan Ampel Surabaya

Dr. Eni Purwati, M.Ag
NIP.196512211990022001

KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA
PERPUSTAKAAN
Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax.031-8413300
E-Mail: perpus@uinsby.ac.id

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama : MOCH. NOOR AFFAN ANSHORI
NIM : H72214016
Fakultas/Jurusan : SAINS DAN TEKNOLOGI / MATEMATIKA
E-mail address : affananshori2@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :
☑ Sekripsi ☐ Tesis ☐ Desertasi ☐ Lain-lain (.....)
yang berjudul :
PREDIKSI GAGAL GINJAL KRONIK MENGGUNAKAN IMPROVED C4.5

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Ekslusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara ***fulltext*** untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya, 02 Agustus 2018

Penulis

(Moch. Noor Affan A.)

PREDIKSI GAGAL GINJAL KRONIK MENGGUNAKAN
IMPROVED C4.5

ABSTRAK

Saat ini banyak sekali penyakit yang berbahaya untuk kesehatan manusia. Salah satu penyakit yang berbahaya untuk kesehatan manusia adalah penyakit gagal ginjal kronik. Ginjal memiliki peran yang sangat penting untuk menjaga stabilitas tubuh. Saat ginjal tidak dapat menjalankan fungsinya dengan baik maka akan terjadi gagal ginjal kronik. Gagal ginjal kronik memiliki resiko yang sangat serius untuk kesehatan. Tujuan penelitian ini yaitu untuk mengetahui seseorang menderita penyakit gagal ginjal atau tidak. Dalam penelitian ini dilakukan beberapa tahapan yaitu sebagai berikut: pre-proses data, pembuatan model prediksi dengan menggunakan metode *improved c4.5*, pengujian model dengan menggunakan data uji. Dari penelitian ini diperoleh hasil sebagai berikut: nilai akurasi yang diperoleh sebesar 99,21%, sensitivitas sebesar 98,70%, dan spesifikasi sebesar 100%.

Kata Kunci: Penyakit gagal ginjal kronik, *Improved C4.5*

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

CHRONIC KIDNEY DISEASE PREDICTION USING IMPROVED C4.5

ABSTRACT

Nowadays, many diseases are dangerous to human health. One of the diseases that are harmful to human health is chronic kidney disease. Kidney have a very important role to maintain body stability. When the kidney is unable to function properly, chronic kidney desease will occur. Kidneys have important role to maintain the stability of the body. The purpose of this study is to know someone suffering from chronic kidney disease or not. In this research, there are several steps to do that is: pre-process data, making prediction model using improved c4.5 method, prediction model testing using test data. The results of this research are : accuracy value obtained for 99.21%, sensitivity of 98.70%, and specifications of 100%.

Keywords: Chronic Kidney Disease, Improved C4.5.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

DAFTAR ISI

LEMBAR PEGESAHAN ..... **ii**
PERNYATAAN KEASLIAN ..... **iii**
KATA PENGANTAR ..... **iv**
DAFTAR ISI ..... **ix**
DAFTAR TABEL ..... **xi**
DAFTAR GAMBAR ..... **xii**
BAB I PENDAHULUAN **1**
A. Latar Belakang 1
B. Rumusan Masalah 5
C. Tujuan Penelitian 5
D. Batasan Masalah 5
E. Manfaat Penelitian 5
F. Sistematika Penulisan 6
BAB II TINJAUAN PUSTAKA **8**
A. Penyakit Gagal Ginjal Kronik 8
B. *Data Mining* 8
C. Klasifikasi 10
D. Decision Tree 13
E. Algoritma C4.5 16
F. *Improved* C4.5 18
G. Akurasi, Sensitivitas, dan Spesifitas 19
BAB III METODE PENELITIAN **21**
A. Jenis Penelitian 21
B. Sumber Data 21
C. Analisis Data 26
D. Rancangan Sistem 28
1. Desain Operasional 28
2. Metode Penyelesaian 29
BAB IV HASIL DAN PEMBAHASAN **31**
A. Persiapan Data 31
B. Pembuatan Model Klasifikasi 34
C. Pengujian Model Klasifikasi 69
D. Analisis Hasil 78
BAB V PENUTUP **81**

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

A. Kesimpulan 81
B. Saran..... 82
DAFTAR PUSTAKA 83

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

DAFTAR TABEL

Tabel 3.1 Informasi Atribut Data 27
Tabel 4.1 Sampel Data Pasien Gagal Ginjal Kronik 31
Tabel 4.2 Sampel Data Pasien Gagal Ginjal Kronik yang Kosong 32
Tabel 4.3 Perhitungan *Improved C4.5* Node 1 52
Tabel 4.4 Perhitungan *Improved C4.5* Node 2 53
Tabel 4.5 Perhitungan *Improved C4.5* Node 3 55
Tabel 4.6 Perhitungan *Improved C4.5* Node 4 58
Tabel 4.7 Perhitungan *Improved C4.5* Node 5 60
Tabel 4.8 Perhitungan *Improved C4.5* Node 6 63
Tabel 4.9 Perhitungan Tiap Node dengan *Improved C4.5* 70
Tabel 4.10 Perhitungan Tiap Node dengan *Improved C4.5* 76
Tabel 4.9 Data testing dan sampel pengujian 78

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BAB I

PENDAHULUAN

A. Latar Belakang

Kesehatan memiliki peranan yang sangat penting dalam setiap aktivitas manusia. Setiap manusia sangat menginginkan tubuh yang sehat, mulai dari bayi sampai berusia tua. Dalam perkembangan hidup manusia, manusia selalu berusaha mencari pola untuk hidup sehat atau menemukan obat dari suatu penyakit. Begitu juga dalam Islam, manusia selalu disuruh untuk menjaga diri untuk tidak menjatuhkan diri dalam kebinasaan (perilaku yang tidak baik) yang menjadikan tubuh menjadi sakit-sakitan, seperti dalam Alquran surat Al-Baqarah ayat 195 :

وَلَا تُلْقُوا۟ بِأَيْدِيكُمْ إِلَى ٱلتَّهْلُكَةِ وَأَحْسِنُوٓا۟ ۚ إِنَّ ٱللَّهَ يُحِبُّ ٱلْمُحْسِنِينَ ١٩٥

“ dan janganlah kamu menjatuhkan dirimu sendiri ke dalam kebinasaan, dan berbuat baiklah, karena sesungguhnya Allah menyukai orang-orang yang berbuat baik”.

Menjaga kesehatan agar terhindar dari berbagai penyakit sangat penting untuk dilakukan. Saat ini, banyak sekali penyakit yang berbahaya untuk kesehatan manusia. Salah satu penyakit yang berbahaya untuk kesehatan manusia yaitu penyakit gagal ginjal. Ginjal memiliki peran yang sangat penting dalam menjaga stabilitas tubuh dengan mengatur cairan tubuh, keseimbangan elektrolit, pengeluaran hasil metabolisme (Lucida, Trisnawati, & Suardi, 2011).

Penyakit gagal ginjal merupakan penyakit yang memiliki dampak pada masalah medik, ekonomi dan sosial yang begitu besar bagi pasien dan keluarga di negara-negara maju maupun berkembang (Hidayati,

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Kushadiwijaya, & Suhardi, 2008). Terjadinya peningkatan insidensi, prevalensi, dan tingkat morbiditasnya menunjukkan penyakit gagal ginjal merupakan masalah kesehatan yang serius di dunia. Setiap tahunnya penderita penyakit gagal ginjal kronik meningkat sebesar 20-25% berdasarkan data di *United States Renal Data System* (USRDS Annual Data Report, 2006). Penderita penyakit gagal ginjal kronik di Indonesia diperkirakan meningkat setiap tahunnya, menurut survey WHO di Indonesia sebesar 46 persen dari tahun 1955 -2025 akan terjadi peningkatan jumlah penderita penyakit gagal ginjal kronik (Melti, Arthur, & Manoppo , 2014).

Masalah kesehatan yang berkaitan dengan penyakit gagal ginjal kronis terjadi ketika ginjal tak mampu lagi melakukan fungsinya secara maksimal. Apabila laju filtrasi glomeruler (LFG) kurang dari 60 ml/menit atau 1,73 m^2 selama tiga bulan atau lebih akan mengalami gagal ginjal kronik (Hidayati, Kushadiwijaya, & Suhardi, 2008). Gagal ginjal kronik umumnya disebabkan oleh penyakit ginjal intrinsik difus dan menahun. Sekitar 60 persen penyakit ginjal kronik disebabkan oleh Glomerulonefritis, hipertensi esensial, dan pielonefritis (Sukandar, 2006). Banyak faktor yang menyebabkan gagal ginjal diantaranya usia, jenis kelamin, dan riwayat penyakit seperti diabetes, hipertensi, maupun penyakit gangguan metabolik lain yang dapat menyebabkan penurunan fungsi ginjal (Pranandari & Supadmi, 2015).

Penderita penyakit gagal ginjal memiliki resiko mengalami komplikasi penyakit diantaranya hipertensi, anemia, penyakit tulang, gagal jantung, penurunan eksresi (Smeltzer & Bare, 2001). Gejala yang ditimbulkan

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

dari gagal ginjal kronik tidak muncul secara langsung melainkan muncul secara bertahap, gejala awal yang tidak jelas menjadikan penurunan fungsi ginjal tersebut tidak dirasakan, pada tahap yang sudah parah dan sulit diobati baru diketahui (Rostanti, Bawotong, & Onibala, 2016). Selain gejala yang ditimbulkan, masalah lainnya bagi penderita gagal ginjal adalah biaya perawatan yang mahal dengan hasil yang tidak bisa dipastikan kesembuhan totalnya (Santoso, 2016).

Gejala gagal ginjal perlu diketahui sedini mungkin agar tidak berlanjut pada tahap yang parah dan sulit diobati, penderita bisa mendapat bantuan untuk merubah pola hidupnya untuk lebih menjaga kesehatannya (Sumilati & Soleha, 2015). Seperti sabda Nabi dari Abu Hurairah radhiallahu anhu, bahwa Rasulluah SAW bersabda:

مَا أَنْزَلَ اللَّهُ مِنْ دَاءٍ إِلَّا أَنْزَلَ لَهُ شِفَاءً

“Tidaklah Allah menurunkan sebuah penyakit melainkan menurunkan pola obatnya ”.

Upaya untuk mengetahui gejala gagal ginjal sedini mungkin diperlukan adanya sistem pendukung keputusan yang dapat melakukan pengambilan keputusan penyakit ginjal kronik (Amalia H. , 2018). Data yang besar dapat dimanfaatkan untuk menggali pola atau ilmu pengetahuan dengan menggunakan metode *data mining* (Amalia & Evicienna, 2017).

Teknik *data mining* seperti klasifikasi, *clustering*, dan lain-lain digunakan untuk menemukan pola dan hubungan data diantara seperangkat faktor yang besar dan untuk membangun model prediksi yang dapat diandalkan berdasarkan data masukkan yang diberikan. Proses penambangan data ini

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

dilakukan untuk menghasilkan pengetahuan yang berharga yang selanjutnya akan diimplementasikan ke dalam aplikasi sistem penunjang keputusan (Amalia & Evicienna, 2017). Di bidang kesehatan, *data mining* telah digunakan untuk diagnosis dan prognosis penyakit serta untuk memprediksi hasil prosedur medis (Silva, Cortez, Santos, Gomes, & Neves, 2008).

Pada penelitian sebelumnya menggunakan metode *data mining* dengan menggunakan dataset penyakit ginjal telah dilakukan oleh beberapa peneliti yaitu menggunakan metode algoritma C4.5 mendapatkan akurasi sebesar 91.50% (Rianto & Iswari, 2017). Penelitian tersebut termasuk dalam teknik *data mining* klasifikasi. Penelitian lainnya tentang *data mining* klasifikasi menggunakan *Improved C4.5* tentang analisis penjualan disebutkan memiliki akurasi yang lebih baik, efesiensinya meningkat pesat, pohon keputusannya lebih terstruktur sehingga informasi yang digali menjadi lebih baik lagi (Cao & Xu, 2009). Untuk itu peneliti akan melakukan pengolahan dataset penyakit ginjal kronis dengan teknik *data mining* klasifikasi menggunakan metode *Improved* C4.5 yang diharapkan lebih baik dari algoritma C4.5.

Berdasarkan kondisi tersebut, maka peneliti malakukan penelitian dengan judul "**Prediksi Penyakit Gagal Ginjal Kronik dengan Menggunakan *Improved* C4.5**". Harapan dari sistem ini adalah dapat mengklasifikasikan pasien menderita penyakit ginjal kronik atau tidak menderita penyakit gagal ginjal. Dan hasil dari sistem ini dapat digunakan untuk deteksi dini penyakit gagal ginjal kronik.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

B. Rumusan Masalah

Berdasarkan latar belakang yang sudah diuraikan tersebut diperoleh rumusan masalah sebagai berikut:

1. Bagaimana membangun model prediksi Penyakit gagal ginjal menggunakan *Improved* C4.5?
2. Bagaimana akurasi model prediksi dengan menggunakan *Improved* C4.5?

C. Tujuan Penelitian

Adapun tujuan dari penelitian ini yaitu:

1. Untuk mengetahui penyakit gagal ginjal dengan menggunakan Improved C4.5.
2. Untuk mengetahui tingkat kesalahan menggunakan *Improved* C4.5.

D. Batasan Masalah

Batasan dari penelitian ini adalah :

1. Dataset yang digunakan didapatkan dari situs penyedia data yaitu UCI repository
2. Kelas yang digunakan ada dua kelas, yaitu kelas gagal ginjal kronik dan tidak gagal ginjal kronik

E. Manfaat Penelitian

Adapun manfaat dari penulisan skripsi ini yaitu :

1. Manfaat Teoritis

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Hasil dari penelitian ini secara teoritis diharapkan dapat memberikan sumbangan pemikiran dalam memperkaya wawasan konsep praktek bidang kesehatan dalam diagnosis penyakit gagal ginjal.

2. Manfaat Praktis

Hasil dari penelitian ini secara praktis diharapkan dapat menjadi acuan bagi tenaga kesehatan untuk mendeteksi penyakit gagal ginjal sehingga dalam memberikan penanganan dapat meminimalisir resiko yang diakibatkan penyakit gagal ginjal.

F. Sistematika Penulisan

Penulisan ini disusun dengan sistematika yang terdiri dari beberapa bab atau bagian bab yaitu sebagai berikut :

Bab I Pendahuluan, berisi tentang Latar Belakang Masalah, Rumusan Masalah, Tujuan, Batasan Masalah, Manfaat Penelitian, dan Sistematika Penulisan.

Bab II Dasar Teori, berisi tentang Teori-teori, Landasan Teori, mengenai penyakit ginjal kronik, *data mining*, klasifikasi, *decision tree*, Algoritma ID3, Algoritma C4.5, dan *Improved C4.5*.

Bab III Metode Penelitian, berisi tentang jenis penelitian, waktu penelitian, pengumpulan data, analisis data, dan pengujian serta evaluasi.

Bab IV Hasil dan Pembahasan, berisi tentang hasil klasifikasi menggunakan *Improved C4.5* multikelas serta analisis dari hasil-hasil yang didapat.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Bab V Kesimpulan dan Saran, berisi tentang kesimpulan yang di dapat dari penelitian prediksi penyakit gagal ginjal kronik serta perbaikan untuk penelitian selanjutnya.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BAB II

TINJAUAN PUSTAKA

A. Penyakit Gagal Ginjal Kronik

Ginjal merupakan organ yang sangat penting untuk tubuh. Ginjal berfungsi untuk menyaring darah dalam tubuh. Ketika ginjal menyaring darah, urine akan tercipta dari zat sisa dan kelebihan cairan dalam tubuh. Ginjal juga berfungsi untuk menyaring dan membuang limbah seperti racun, garam berlebih, dan urea (limbah yang mengandung nitrogen yang diproduksi saat proses metaboisme tubuh), Mengendalikan cairan dalam tubuh, mengatur tekanan darah dan tingkat garam dalam darah, mengatur sel darah merah, serta menjaga konsentrasi mineral, seperti natrium, kalium, dan fosfor dalam darah.

Gagal ginjal terjadi ketika ginjal tidak lagi dapat melakukan fungsinya dengan benar. Gagal ginjal sangat berpengaruh terhadap kesehatan tubuh. Gagal ginjal kronik merupakan kerusakan ginjal progresif yang ditandai dengan uremia (urea dan limbah nitrogen lainnya yang beredar di dalam darah) dan berakibat fatal serta memiliki komplikasi penyakit lainnya jika tidak dilakukan dialisis atau transplantasi ginjal. Gagal ginjal kronik terjadi apabila laju filtrasi glomeruler (LFG) selama tiga bulan atau lebih kurang dari 60 ml/menit per 1,73 m^2 selama tiga bulan atau lebih (Hidayati, Kushadiwijaya, & Suhardi, 2008).

B. Data Mining

Dalam *database* terdapat informasi dan pengetahuan yang tersembunyi untuk menemukannya digunakan suatu metode yang disebut *data mining*.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac c.id

Data mining merupakan proses semi otomatik untuk mengektrasi dan mengidentifikasi informasi potensial dan berguna yang dapat dimanfaatkan menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning (Octhario & Karnila, 2013).

Pekerjaan yang berkaitan dengan *data mining* dapat dibagi menjadi empat kelompok: model prediksi, analisis cluster, analisis asosiasi, dan deteksi anomali. Model prediksi merupakan pembuatan model yang dapat melakukan pemetaan dari setiap himpunan variabel ke setiap targetnya, selanjutnya model tersebut digunakan untuk memberikan nilai target pada himpunan baru yang didapat. Ada dua jenis model prediksi, yaitu klasifikasi yang digunakan pada variabel target diskrit dan regresi digunakan pada variabel target kontinu (Prasetyo, 2014).

Melakukan deteksi terhadap jenis penyakit pasien berdasarkan nilai-nilai parameter penyakit yang diderita adalah contoh *data mining* yang menggunakan jenis klasifikasi. Target yang diharapkan adalah diskrit ini termasuk dalam jenis klasifikasi, hanya beberapa jenis kemungkinan nilai target yang didapatkan *time series* (Prasetyo, 2014).

Menurut Gonunescu dalam *data mining* ada tiga tahapan utama yaitu :

1. Eksplorasi atau pemrosesan awal data

   Eksplorasi atau pemrosesan awal data terdiri dari penanganan data yang salah, normalisasi data, pembersihan data, transformasi data, reduksi dimensi, pemilihan subset fitur, dan sebagainya.

2. Membangun model dan melakukan validasi terhadapnya

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Membangun model dan melakukan validasi terhadapnya berarti melakukan analisis untuk mendapatkan suatu model terbaik. Dalam langkah ini digunakan metode-metode seperti klasifikasi, regresi, analisis cluster, deteksi anomali, analisis asosiasi, analisis pola skuensial, dan sebagainya.

3. Penerapan

   Penerapan berarti menerapkan model yang telah didapatkan pada data yang baru untuk mendapatkan perkiraan/prediksi masalah yang diinvestigasi.

C. Klasifikasi

Klasifikasi dapat didefinisikan secara detail sebagai suatu pekerjaan yang melakukan pelatihan/pembelajaran terhadap fungsi target f yang memetakan setiap vektor x ke dalam satu dari sejumlah label kelas y yang tersedia. Pelatihan tersebut dilakukan untuk menghasilkan suatu model yang kemudian disimpan sebagai memori (Prasetyo, 2014). Klasifikasi adalah proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk daapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui (Aradea, Ariyan, & Yuliana , 2011). Dalam klasifikasi, terdapat target variabel kategori. Untuk mengklasifikasikan suatu obyek dalam decision tree tidak menggunakan vektor jarak. Seringkali data observasi mempunyai atribut-atribut yang bernilai nominal. Misalkan obyeknya adalah sekumpulan hewan yang bisa dibedakan berdasarkan atribut bentuk, warna, ukuran dan jumlah kaki. Bentuk, warna, ukuran dan jumlah

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

kaki adalah besaran nominal. Dengan nilai-nilai atribut ini, kemudian dibuat decision tree untuk menentukan suatu obyek termasuk jenis hewan apa jika nilai tiap-tiap atribut diberikan (Mardi, 2015).

Dalam proses pengklasifikasian biasa terdapat dua proses yang harus dilakukan, yaitu:

1. Proses *Training*

   Pada proses ini akan digunakan data *training* set atau data sampel yang telah diketahui label-label atau atribut dari data sampel tersebut untuk membangun model.

1. Proses *Testing*

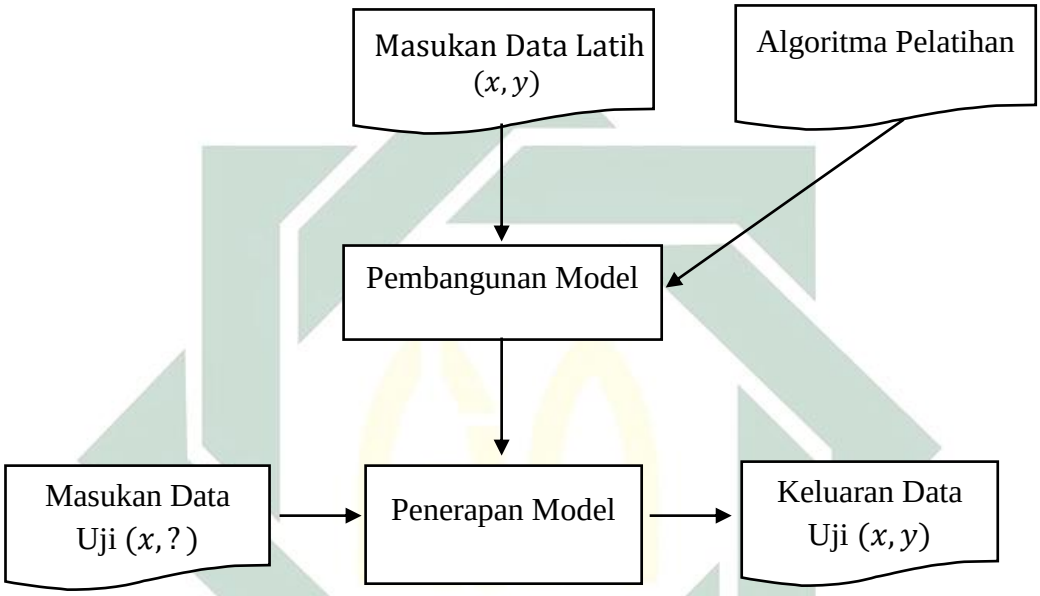
   Pada proses *testing* ini dilakukan suatu uji untuk mengetahui keakuratan model yang telah dibuat pada proses training sehingga dibangun data yang disebut dengan data testing untuk mengklasifikasikan label-labelnya.

Model dalam klasifikasi mempunyai arti yang sama dengan *black box*, dimana ada suatu model yang dapat menerima masukan kemudian melakukan analisis dengan menggunakan model tersebut untuk memberikan jawaban sebagai keluaran dari hasil pemikirannya.

Pada Gambar 2.1 diberikan sejumlah data latih (x, y) untuk digunakan sebagai data pembangun model, selanjutnya menggunakan model tersebut untuk melakukan prediksi kelas dari data uji $(x, ?)$ (sehingga data uji $(x, ?)$ diketahui kelas y yang seharusnya. Kerangka kerja yang ditunjukan pada Gambar 2.1 meliputi dua langkah proses yaitu induksi dan deduksi. Induksi merupakan suatu proses untuk membuat model klasifikasi dari data latih yang

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

diberikan, disebut juga proses pelatihan, sedangkan deduksi merupakan suatu langkah yang digunakan untuk menerapkan model tersebut pada data uji sehingga data uji dapat diketahui kelas yang sesungguhnya atau disebut juga proses prediksi.



Gambar 2.1 Proses Klasifikasi

Data baru yang belum diketahui label kelasnya dapat diprediksi menggunakan model yang dibangun pada saat pelatihan. Dalam pembangunan model selama proses pelatihan tersebut diperlukan adanya suatu algoritma yang digunakan untuk membangun model yang disebut sebagai algoritma pelatihan (Prasetyo, 2014)

Pada proses klasifikasi data yang diimputkan adalah data *record* atau data sampel. Pada setiap *record* dikenal sebagai instance atau contoh yang ditentukan oleh sebuah tuple (x, y). Dimana x adalah himpunan atribut dan y adalah atribut menyatakan sebagai label kelas.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Komponen-komponen utama dari proses klasifiksi antara lain:

1. Kelas, merupakan label dari hasil klasifikasi yang variabel tidak bebas. sebagai contoh adalah kelas loyalitas pelanggan, kelas badai, kelas gagal ginjal, dan lain-lain.
2. Prediktor, merupakan variabel bebas suatu model berdasarkan dari karakteristik atribut data yang diklasifikasi, misalnya merokok, minum-minuman beralkohol, dan sebagainya.
3. Set data pelatihan, merupakan sekumpulan data lengkap yang berisi kelas dan prediktor untuk dilatih sehingga mendapatkan suatu model agar dapat mengelompokkan data ke dalam kelas yang tepat.
4. Set data uji, berisi data-data baru yang digunakan untuk menguji suatu model dengan mengelompokkan data berdasarkan suatu model tertentu guna megetahui akurasi dari model yang telah dibuat.

D. Decision Tree

Decision tree adalah pohon keputusan yang digunakan sebagai prosedur penalaran untuk mendapatkan suatu jawaban dari masalah yang dimasukkan. Pohon yang dibentuk tidak selalu berupa pohon biner. Jika fitur dalam data set menggunakan dua macam nilai kategorikal maka pohon yang dihasilkan berupa pohon biner. Jika berisi lebih dari dua macam nilai kategorikal atau menggunakan tipe numerik maka bentuk pohon keputusan yang dihasilkan biasanya tidak berupa pohon biner (Prasetyo, 2014).

Decision tree merupakan sebuah diagram alir yang memiliki bentuk seperti struktur pohon yang setiap *internal node* menyatakan pengujian

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

terhadap suatu atribut, setiap cabang menyatakan output dari pengujian tersebut dan *leaf node* menyatakan kelas-kelas atau distribusi kelas. *Node* yang paling awal disebut sebagai *node* akar. Sebuah *node* akar akan memiliki sebuah *root node*. Sebuah *root node* memiliki beberapa *edge* keluar tetapi tidak memiliki *edge* masuk, *Internal node* akan memiliki satu *edge* masuk dan beberapa edge keluar, sedangkan *leaf node* hanya akan memiliki satu *edge* masuk dan beberapa *edge* keluar, sedangkan *leaf node* hanyakan memiliki satu *edge* masuk tanpa memiliki *edge* keluar (Basuki & Syarif, 2003).

Untuk mengklasifikasikan suatu sampel data yang belum diketahui kelasnya dapat digunakan metode *decision tree* untuk diketahui masuk ke dalam kelas-kelas yang sudah ada. Jalur pengujian data yang pertama melalui *root node* dan terakhir melalui *leaf node* yang akan menyimpulkan prediksi kelas bagi data tersebut. Atribut data yang digunakan harus berupa data kategorik, bila data kontinu maka atribut harus didiskreditasi terlebih dahulu (Basuki & Syarif, 2003).

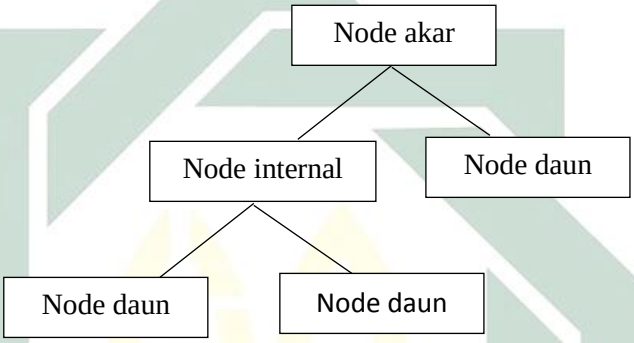
Kefleksibelan membuat metode ini atraktif, karena dapat memberikan keuntungan berupa visualisasi saran (dalam suatu bentuk pohon keputusan) yang menjadikan prosedur prediksinya mudah diamati. *Decision tree* sering digunakan untuk menyelesaikan kasus penentuan keputusan seperti di bidang kedokteran (diagnosis penyakit pasien), ilmu komputer (struktur data), psikologi (teori pengambilan keputusan), dan sebagainya (Prasetyo, 2014).

Decision tree merupakan teknik model prediksi yang dapat digunakan untuk klasifikasi dan prediksi tugas. Teknik yang digunakan dalam *decision*

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

tree adalah teknik membagi dan menaklukkan untuk membagi ruang pencarian masalah menjadi himpunan masalah (Dunham, 2003). Proses pada *decision tree* yaitu mengubah data yang berbentuk tabel menjadi sebuah model *tree* yang dapat digunakan untuk mengambil sebuah kesimpulan. Model *tree* akan menghasilkan *rule* dan disederhanakan (Basuki & Syarif, 2003).

Konsep dasar *decision tree* pada Gambar 2.2 dibawah ini



Gambar 2.2 *Decision tree*

Karakteristik dari *decision tree* seperti pada gambar 1 dibentuk sejumlah elemen sebagai berikut (Tan, Steinbach, & Kumar , 2006):

a. Node akar, tidak mempunyai lengan masukan dan mempunyai nol atau lebih lengan keluaran.
b. Node internal, setiap node yang bukan daun yang mempunyai tepat satu lengan masukan dan dua atau lebih lengan keluaran. Node ini menyatakan pengujian yang didasarkan pada nilai fitur.
c. Lengan, setiap cabang menyatakan nilai hasil pengujian di node bukan daun.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

d. Node daun, node yang mempunyai tepat satu lengan masukan dan tidak mempunyai lengan keluaran. Node ini menyatakan label kelas (keputusan).

E. Algoritma C4.5

Algoritma C4.5 diperkenalkan oleh Quinlan (1996) sebagai versi perbaikan dari ID3. Algoritma C4.5 merupakan algoritma yang digunakan untuk permasalahan klasifikasi di dalam sebuah mesin pembelajaran dan himpunan data. dengan nilai data yang berbeda-beda, dimana kejadian diuraikan oleh koleksi kategori dan mempunyai salah satu dari satu set kelas ekslusif (Prasetyo, 2014).

Dalam ID3, induksi *decision tree* hanya bisa dilakukan pada fitur bertipe kategorikal, sedangkan data bertipe numerik tidak dapat digunakan. Perbaikan yang membedakan antara ID3 dan C4.5 adalah dapat menanganani data bertipe numerik, melakukan pemotongan *decision tree* dan penurunan rule set (Prasetyo, 2014). Algoritma C4.5 juga menggunakan kriteria Gain dalam menentukan fitur yang menjadi pemecah node pada pohon yang diinduksi.

Ada 3 kelompok penting dalam dalam syarat pengujian node:

1. Fitur Biner

   Dalam fitur ini, nilai yang ada hanya dua jenis sehingga disebut fitur biner. Dalam pengujian ketika menjadi node, fitur ini hanya mempunyai dua pilihan cabang.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

2. Fitur Kategorikal

   Untuk fitur yang nilainya bertipe kategorikal (nominal atau ordinal) bisa mempunyai beberapa nilai berbeda. Fitur ini mempunyai cabang sesuai dengan berapa jumlah kategori yang ditentukan.

3. Fitur Numerik

   Untuk fitur bertipe numerik, syarat pengujian dalam node dinyatakan dengan pengujian perbandingan (A< v) atau (A> v) dengan hasil biner, atau untuk multi dengan hasil berupa jangkauan nilai dalam bentuk ($v_i < A > v_{i+1}$) untuk i = 1, 2, ..., k.

Kriteria yang paling sering digunakan untuk memilih fitur sebagai pemecah dalam algoritma C4.5 adalah rasio gain, dengan persamaan berikut:

$$RasioGain\ (s.j) = \frac{Gain\ (s.j)}{SplitInfo\ (s.j)} \tag{2.1}$$

Persamaan diatas menyatakan nilai rasio gain pada fitur ke-j. Untuk perhitungan $Gain\ (s.j)$ dan $SplitInfo\ (s.j)$ didapat dari persamaan berikut

$$Gain\ (s.j) = E(s) - \sum_{i=1}^{n} p(v_i|s) \times E(s_i) \tag{2.2}$$

$p(v_i|s)$ adalah proporsi nilai v muncul pada kelas dalam node. $E(s_i)$ adalah entropi komposisi nilai v dari kelas ke$-j$ dalam data ke$-i$ node tersebut. n adalah jumlah nilai berbeda dalam node.

$$SplitInfo\ (s.j) = -\sum_{i=1}^{k} p(v_i|s) \log_2 p(v_i|s) \tag{2.3}$$

k menyatakan jumlah pemecahan.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

$$E(s) = -\sum_{i=1}^{m} p(\omega_i|s)\log_2 p(\omega_i|s) \tag{2.4}$$

$p(\omega_i|s)$ adalah proporsi kelas ke$-i$ dalam semua data latih yang diproses di node s. $p(\omega_i|s)$ didapatkan dari jumlah semua baris data dengan label kelas i dibagi jumlah semua baris semua data. m adalah jumlah nilai berbeda dalam data.

F. *Improved* C4.5

Improved C4.5 merupakan perbaikan dari algoritma C4.5 untuk mendapatkan efesiensi perhitungan sehingga mempercepat dalam membuat pohon keputusan. Dalam *improved* C4.5 tidak hanya mempercepat pohon keputusan saja, tapi juga mendapatkan pohon keputusan yang lebih terstruktur, sehingga informasi yang lebih baik dapat digali. *Improved* C4.5 juga lebih ringkas dibandingkan dengan algoritma C4.5 biasa. Dalam *improved* C4.5 perhitungan Rasio Gain dipermudah dengan menggunakan Persamaan 2.3.

$$RasioGain\ (S.A) = \frac{\left(\frac{p_n}{N} - \left\{\left(\frac{s_{11}s_{12}}{s_1}\right) + \left(\frac{s_{11}s_{12}}{s_2}\right)\right\}\right)}{\left(\frac{s_1 s_2}{N}\right)} \tag{2.3}$$

p = jumlah data positif dari semua data

n = jumlah data negatif dari semua data

N = jumlah semua data

S_1 = jumlah data positif di A

S_2 = jumlah data negatif di A

S_{11} = jumlah data positif di A dan atribut bernilai positif

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

S_{12} = jumlah data positif di A dan atribut bernilai negatif

S_{21} = jumlah data negatif di A dan atribut bernilai positif

S_{21} = jumlah data negatif di A dan atribut bernilai negatif

Dalam rumus perhitungan gain tersebut hanya menggunakan penambahan, pengurangan, perkalian, pembagian, dan tidak ada perhitungan logaritma sehingga memudahkan dalam penghitungan dan efesiensi waktu menjadi lebih baik (Cao & Xu, 2009).

G. Akurasi, Sensitivitas, dan Spesifitas

Suatu sistem tidak dapat melakukan klasifikasi semua data set dengan semua benar, selalu terdapat kesalahan dalam proses klasifikasi. Oleh karena itu, pengukuran kinerja dari sistem klasifikasi penyakit gagal ginjal menggunakan akurasi, sensitivitas, dan spesifisitas. Akurasi menunjukkan kedekatan hasil pengukuran dengan nilai sesungguhnya. Semakin besar nilai akurasi yang didapatkan semakin baik hasil yang diperoleh. Sensitivitas didefinisikan sebagai kemampuan alat tes untuk mendeteksi mereka yang terkena penyakit. Spesifitas didefinisikan sebagai kemampuan alat tes untuk mendeteksi mereka yang tidak terkena penyakit. Prediktor yang baik memberikan nilai sensitivitas sebesar 100% (misalnya semua orang yang sebenarnya sakit diidentifikasi sebagai sakit) dan nilai spesifitas sebesar 100% (misalnya semua orang yang sebenarnya tidak sakit diidentifikasi sebagai tidak sakit).

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

$$Akurasi = \frac{jumlah\ data\ yang\ diklasifikasi\ benar}{jumlah\ klasifikasi\ yang\ dilakukan} \quad (2.4)$$

$$Sensitivitas = \frac{TP}{TP + FN} \quad (2.5)$$

$$Spesifisitas = \frac{TN}{FP + TN} \quad (2.6)$$

Keterangan :

TP = Orang sakit yang didiagnosis secara benar sebagai sakit

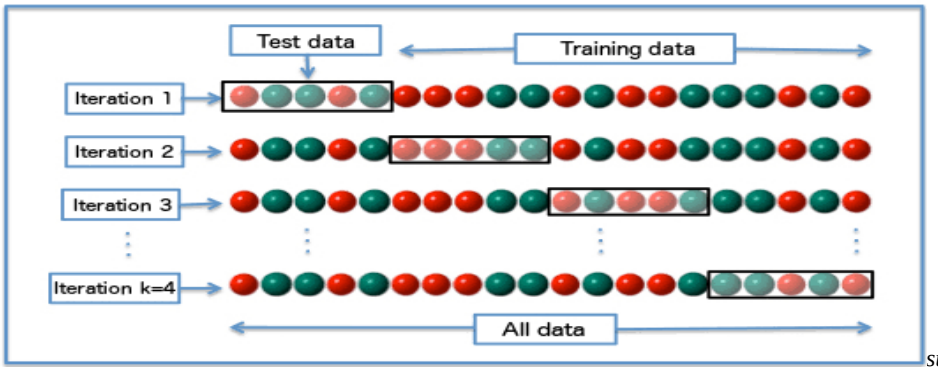
FP = Orang sehat yang didiagnosis secara salah sebagai sakit

TN = Orang sehat yang didiagnosis secara benar sebagai sehat

FN = Orang sakit yang didiagnosis secara salah sebagai sehat

H. K-Fold Cross Validation

K-fold cross validation merupakan sebuah metode yang digunakan untuk menguji tingkat akurasi pada pohon keputusan. Metode ini digunakan untuk menghindari *overlapping* pada data testing. Tahapan-tahapan pada k-fold cross validation adalah membagi data menjadi k subset yang berukuran sama. Selanjutnya gunakan setiap subset untuk data testing dan sisanya untuk data training.



sumber: en.wikipedia.org/wiki/cross-validation_(statistics)

Gambar 2.2 K-fold Croos Validation

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BAB III

METODE PENELITIAN

A. Jenis Penelitian

Jenis penelitian yang digunakan dalam penelitian ini merupakan penelitian kuantitatif. Jenis penelitian ini merupakan jenis penelitian yang spesifikasinya bersifat sistematis, terencana, dan terstruktur dengan jelas mulai dari awal hingga pembuatan desain penelitiannya. Jenis penelitian kuantitatif dapat diartikan sebagai metode penelitian yang berladaskan pada filsafat positivisme, digunakan untuk dalam meneliti suatu populasi atau sampel tertentu.

Pada penelitian ini menggunakan pendekatan deskriptif. Penelitian deskriptif yaitu penelitian yang dilakukan untuk mengetahui nilai variabel mandiri. Penelitian kuantitatif deskriptif merupakan data yang diperoleh dari sampel populasi penelitian dianalisis sesuai dengan metode statistik yang digunakan. Penelitian kuntitatif deskriptif dalam penelitian ini dimaksudkan untuk mendapatkan gambaran dan keterangan-keteangan mengenai data yang digunakan dalam penelitian ini.

B. Sumber Data

Pada penelitian ini, dataset yang digunakan didapatkan dari situs penyedia data yaitu UCI repository. Data yang digunakan merupakan data sekunder sebanyak 400 record yang terdiri dari 24 atribut dan satu label. Data sekunder yaitu data yang diperoleh bukan dari pengumpulan data secara langsung oleh peneliti.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Atribut-atribut data yang digunakan yaitu:

1. Umur, yaitu satuan waktu yang mengukur waktu keberadaan suatu benda atau makhluk, baik yang hidup maupun yang mati. Rentang yang digunakan yaitu 0 – 30 tahun adalah muda, 31- 45 tahun adalah dewasa, 46 tahun dan seterusnya adalah lanjut usia.
2. Tekanan darah, yakni tekanan yang terjadi pada pembuluh darah arteri ketika darah dipompa oleh jantung untuk dialirkan ke seluruh anggota tubuh. Tekanan darah normal berada di rentang 60 – 80 mm/Hg.
3. *Specific gravity* urin, yang menunjukkan proporsi relatif dari komponen-komponen yang larut pada keseluruhan volume spesimen urin. Pengetahuan akan *specific gravity* dibutuhkan dalam menginterprestasi hasil-hasil dari tes pada analisis urin.
4. *Albumin*, yakni jenis protein monomer yang larut dalam air atau garam dan mengalami koagulasi saat terpapar panas. Fungsinya memelihara tekanan osmotik yang ditimbulkan oleh *albumin* akan memelihara fungsi ginjal.
5. Gula darah, yakni tingkat glukosa dalam darah. Jika level gula darah terlalu rendah akan menimbulkan hipokligemia, gejalanya adalah perasaan lelah, kehilangan kesadaran. Jika level gula darah teralu tinggi akan menimbulkan hiperkligemia. Dalam jangka panjang dapat menimbulkan penyakit diabetes, ginjal, kerusakan mata, dan saraf.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

6. Status sel darah merah. Sel darah merah adalah sel darah yang berfungsi membawa oksigen ke jaringan-jaringan tubuh melalui darah. Kekurangan sel darah merah dapat memicu keberadaan anemia, gejalanya adalah mudah lelah dan imun semakin melemah. Kelebihan sel darah merah dapat mengakibatkan penggumpalan darah, kerusakan ginjal, limpa, dan mata. Status sel darah merah terbagi dua yakni normal atau tidak normal.
7. Status *pus cell*, yakni kumpulan cairan, sel darah putih, mikroorganisme dan bahan selular yang menunjukan adanya luka yang terinfeksi atau abses.
8. *Pus cell clumps* adalah sel darah putih yang bergerombol dalam pemeriksaan urin, biasanya salah satu pertanda adanya infeksi.
9. Status kehadiran bakteri. Bakteri yang masuk melalui uretra dan melakukan perjalanan ke kandung kemih mengakibatkan infeksi kandung kemih. Biasanya tubuh dapat menghilangkan bakteri dengan menyiram bakteri agar keluar saat buang air kecil.
10. Kadar glukosa dalam darah. Glukosa yang tinggi dapat menunjukan diabetes, tetapi banyak penyakit dan kondisi lain juga dapat disebabkan glukosa dalam darah. Kadar glukosa dalam darah yang normal berada di rentang 70 – 130 mgs / dl.
11. *Blood urea nitrogen.* Kadar nitrogen dalam darah merupakan banyaknya kandungan nitrogen dalam darah. Jika terjadi masalah pada kemampuan kerja ginjal maka jumlah konsentrasinya dalam

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

darah juga akan terganggu. *Blood urea nitrogen* yang normal berada pada rentang 15 – 40 mgs/dl.

12. Kadar *serum creatinine.* Kadar kreatinin merupakan limbah lain dari proses pengolahan pembntukan otot. Ketika terjadi masalah dengan fungsi ginjal, maka kadar kreatinin dalam darah akan meningkat. Kadar kreatinin yang normal berada pada rentang 0,4 – 1,4 mgs / dl.
13. *Sodium blood test.* Sodium tes ini digunakan untuk mendeteksi konsentrasi abnormal natrium, termasuk sodium rendah dan sodium tinggi. Hal ini sering digunakan sebagai bagian dari *electrolyte* atau panel metabolik dasar untuk pemeriksaan kesehatan rutin. Kadar sodium yang normal berada pada rentang 135 – 145 mEq/ l.
14. *Potassium test. Potassium* tes digunakan untuk mendeteksi konsentrasi abnormal *potassium.* Hal ini sering digunakan sebagai bagian dari *electrolyte* atau panel metabolic dasar untuk pemeriksaan kesehatan rutin. Kadar potassium yang normal berada pada rentang 3,5 – 5,0 mEq/l.
15. Hemoglobin. Kadar hemoglobin yaitu metaloprotein (protein yang mengandung zat besi) di dalam sel darah merah yang berfungsi sebagai pengangkut oksigen dari paru-paru ke seluruh tubuh. Kadar hemoglobin normal berada pada rentang 10 – 18 gms.
16. *Packed cell volume,* yakni persentase volume eritrosit dalam darah yang dimampatkan dengan cara diputar pada kecepatan tertentu dan

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

dalam waktu tertentu. Tujuan dilakukan uji ini adalah untuk mengetahui konsentrasi eritrosit dalam darah. Keadaan normal *packed cell volume* adalah 36 – 50 %.

17. Jumlah sel darah putih, yakni menunjukkan peningkatan dan penurunan yang terjadi pada sel darah putih. Jumlah sel darah putih normal adalah 4500 – 10000 *cell/cum*
18. Jumlah sel darah merah, yakni menunjukkan peningkatan dan penurunan yang terjadi pada sel darah merah. Jumlah sel darah merah normal adalah 3,9 – 5,7 *millions/ cmm*.
19. Riwayat hipertensi. Hipertensi adalah kondisi kronis dimana tekanan darah pada dinding arteri meningkat. Hipertensi dapat menyebabkan kerusakan serius pada jantung serta organ lain seperti ginjal, otak, dan mata.
20. Riwayat diabetes mellitus. Diabetes mellitus merupakan penyakit yang ditandai dengan kadar gula darah yang tinggi yang disebabkan oleh gangguan pada sekresi insulin atau gangguan kerja insulin. Kadar glukosa darah tinggi membuat proses penyaringan pada ginjal menjadi lebih berat. Hal ini dapat menyebabkan sejumlah masalah kesehatan akut dan kronis.
21. Riwayat jantung koroner. Penyakit jantung koroner merupakan penyakit di mana zat lilin yang disebut plak menumpuk di dalam arteri koroner. Arteri ini memasok darah yang kaya oksigen ke otot jantung.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

22. Tingkat nafsu makan. Nafsu makan merupakan keiinginan untuk makan makanan, merasa lapar.
23. Riwayat *pedal edema. Pedal edema* merupakan penyakit dimana terdapat koleksi abnormal cairan di ruang antara sel-sel.
24. Riwayat anemia, Anemia merupakan salah satu kondisi yang terjadi ketika jumlah hemoglobin dalam darah seseorang turun di bawah batas normal.
25. *Class* adalah variabel yang menunjukkan pasien terjangkit penyakit gagal ginjal atau tidak.

C. Analisis Data

Analisis data disebut juga pengelolaan data dan penafsiran data. Tujuan analisis data adalah meyederhanakan data dalam bentuk yang lebih mudah dibaca dan diinterprestasikan. Analisis data dimaksudkan untuk memahami apa yang ada pada data dengan mengelompokkannya, meringkasnya, serta menemukan pola umum yang timbul dari data tersebut. Penjelasan atribut-atribut data dapat dilihat pada Tabel 3.1.

Data yang sudah dihimpun disimpan dalam bentuk tabel ke dalam format csv yang ada pada microsoft excel. Data yang diperoleh dikelompokkan untuk memudahkan dalam proses masukkan data. Selanjutnya data yang telah dikelompokkan dinormalisasi data sehingga tidak ada yang kosong. setelah semua data normal data dibagi menjadi data training dan data testing. Data training digunakan untuk proses pembentukan model prediksi, data training

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

diambil dari 66% dari semua data. Data testing digunakan untuk menguji model prediksi tersebut sudah valid atau belum.

Tabel 3.1 Informasi Atribut Data

No	Nama atribut	Jenis data	Nilai
1	Umur (age)	Numerik	Umur dalam th
2	Tekanan Darah (bp)	Numerik	Bp dalam mm/Hg
3	*Specfic Gravity (*Sg)	Kategorikal	1005,1010,1015,1020,1025
4	*Albumin* (al)	Kategorikal	0,1,2,3,4,5
5	Gula Darah (su)	Kategorikal	0,1,2,3,4,5
6	Status sel darah merah (rbc)	Kategorikal	Normal, tidak normal
7	Status *Pus cell* (pc)	Kategorikal	Normal, tidak normal
8	*Pus cell clumps* (pcc)	Kategorikal	Ada, tidak ada
9	Status kehadiran Bakteri (ba)	Kategorikal	Ada, tidak ada
10	Kadar Glukosa dalam darah (bgr)	Numerik	Bgr dalam mgs/dl
11	*Blood Urea Nitrogen* (bu)	Numerik	Bu dalam mgs/dl
12	Kadar serum kreatinin (sc)	Numerik	Sc dalam mgs/dl
13	Hasil *Sodium blood test* (sod)	Numerik	Sod dalam mEq/l
14	Hasil *Potassium test* (pot)	Numerik	Pot dalam mEq/l
15	Kadar Hemoglobin (hemo)	Numerik	Hemo dalam gms
16	*Packed cell volume* (pcv)	Numerik	Pcv dalam %
17	Jumlah sel darah putih (wc)	Numerik	Wc dalam cell/cum
18	Jumlah sel darah merah (rc)	Numerik	Rc dalam milions/cmm
19	Riwayat Hipertensi (htn)	Kategorikal	Ya, tidak
20	Riwayat diabete mellitus (dm)	Kategorikal	Ya, tidak
21	Riwayat Penyakit Jantung Koroner (cad)	Kategorikal	Ya,tidak
22	Tingkat nafsu makan (appet)	Kategorikal	Baik, kekurangan
23	Riwayat *Pedal edema* (pe)	Kategorikal	Ya,tidak
24	Riwayat Anemia (ane)	Kategorikal	Ya, tidak
25	Class	Kategorikal	Ckd, notckd

Data training diproses menggunakan metode *Improved C4.5* untuk mendapatkan model prediksinya. Setelah didapatkan model prediksinya dilakukan pengujian terhadap metode yang digunakan untuk mengetahui kevalidan dari model tersebut.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

D. Rancangan Sistem

1. Desain Operasional

a. Membuat Laporan Penelitian, yaitu menyusun pembuatan proposal penelitian yang ditunjukkan kepada program yang menyelenggarakan pendidikan pada berbagai jenjang tadi.

b. Mencari literatur-literatur penelitian, yaitu proses pengumpulan peelitian-penelitian sebelumnya sebagai refrensi yang digunakan dalam penyusunan laporan penelitian.

c. Mencari data penelitian, yaitu mencari data yang digunakan dalam penelitian, pencarian data dilakukan berdasarkan variabel-variabel yang digunakan dalam penelitian.

d. Membuat program, yaitu membuat model dari dari data yang diperoleh sesuai dengan data dan metode yang digunakan dengan menggunakan aplikasi pengembang Matlab.

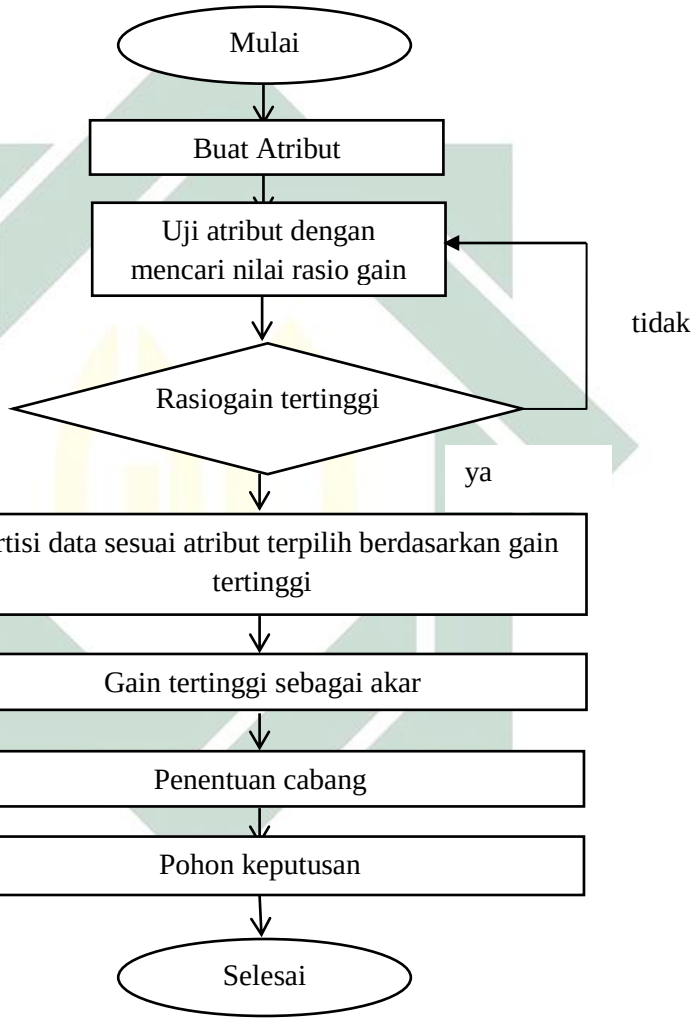
e. Penyelesaian program, yaitu tahap selanjutnya dari pembuatan program, proses ini dilakukan untuk menyelesaikan sebuah program yang sudah valid.

f. Simulasi pogram, yaitu memperagakan hasil dari progam berdasarkan data yang diperoleh untuk menguji apakah program yang digunakan dapat bekerja sesuai dengan apa yang diprogramkan.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

g. Membuat laporan penelitian, yaitu penyusunan laporan dari penelitian sesuai dengan hasil yang diperoleh dari proses penelitan.

2. Metode Penyelesaian



Gambar 3.1 Flowchart Improved C4.5

Tahapan-tahapan klasifikasi penderita ginjal kronik seperti yang ditunjukkan Gambar 3.1 dengan menggunakan *Improved* C4.5 yang

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

pertama menentukan atribut yaitu menentukan atribut yang digunakan dalam penelitian ini untuk dilakukan perhitungan. Selanjutnya melakukan uji atribut dengan mencari nilai rasio gain dengan menggunakan rumus yang ada. Dengan melakukan perhitungan yang ada akan didapatkan nilai rasiogain tertinggi. Mempartisi data sesuai dengan atribut terpilih berdasarkan nilai rasiogain tertinggi. Nilai rasiogain tertinggi menjadi akar. Selanjutnya menghitung rasio gain kembali untuk mendapatkan cabang-cabang dari pohon keputusan. Menentukan cabang-cabang berdasarkan nilai rasiogain pada perhitungan selanjutnya. Setelah semua data yang dicari nilai rasio gainnya terpartisi membuat pohon keputusan berdasarkan nilai rasio gain tertinggi dari setiap perhitungan node.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BAB IV

HASIL DAN PEMBAHASAN

Pada bab ini berisi uraian hasil penelitian yang telah dilakukan, analisis hasil penelitian dan pembahasan hasil penelitian.

A. Persiapan Data

Tahapan awal yang dilakukan dalam penelitian ini adalah melakukan persiapan data yang didapatkan dari situs penyedia data yaitu UCI Repository. Data yang didapatkan merupakan data sekunder sebanyak 400 data yang terdiri dari 24 atribut dan satu label. Atribut-atribut yang digunakan yaitu Umur, Tekanan Darah, *Specific Gravity, Albumin,* Gula Darah, Status Sel Darah Merah, Status *Pus Cell, Pus Cell Clumps,* Status Kehadiran Bakteri, Kadar Glukosa dalam Darah, *Blood Urea Nitrogen,* Kadar *Serum Creatinin,* Hasil *Sodium Blood Test,* Hasil *Pottasium test,* Kadar Hemoglobin, *Packed cell volume,* Jumlah Sel Darah Putih, Jumlah Sel Darah Merah, Riwayat Hipertensi, Riwayat Diabetes Mellitus, Riwayat Penyakit Jantung Koroner, Tingkat Nafsu Makan, Riwayat *Pedal Edema,* Riwayat Anemia dan satu label yaitu *Class.*

Data yang didapatkan tersebut disimpan ke dalam bentuk tabel untuk mempermudah dalam proses masukan data. Data yang didapatkan terdapat beberapa atribut yang tidak memiliki nilai atau kosong sehingga diperlukan normalisasi data untuk mendapatkan data yang lebih baik. Terdapat banyak metode yang dapat digunakan untuk melakukan normalisasi data diantaranya penyisipan data, pengubahan data, dan penghapusan data. Setelah semua data

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

normal selanjutnya data dibagi menjadi data training dan data testing. Data penderita penyakit gagal ginjal sebanyak 231 data dan tidak penderita gagal ginjal sebanyak 150 data.

Data yang sudah disimpan ke bentuk tabel tersebut terdapat beberapa atribut yang tidak memiliki nilai atau kosong. Untuk itu maka dilakukan proses normalisasi data untuk menghilangkan data yang bermasalah.

Pada proses normalisasi ini dilakukan penghapusan data dan penyisipan data yang tidak memiliki nilai atau kosong. Penghapusan data dilakukan terhadap data yang memiliki banyak atribut yang tidak memiliki nilai atau kosong. Penyisipan data diilakukan terhadap data yang memiliki satu atau dua atribut yang masih kosong saja. Penyisipan data dilakukan dengan mencari rata-rata dari atribut yang tidak memiliki nilai atau kosong. Dari proses normalisasi data didapatkan data sebesar 381 data yang akan digunakan dalam proses klasifikasi.

Selanjutnya data tersebut dibagi menjadi data training dan data testing. Data yang digunakan untuk data training sebanyak 254 data dan data yang digunakan sebagai data testing sebanyak 127 data. Data training digunakan untuk proses pembentukan model prediksi dan data testing digunakan untuk menguji model prediksi tersebut sudah valid atau belum.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.1 Sampel Data Pasien Gagal Ginjal Kronik

age	bp	sg	al	su	rbc	pc	pcc	ba	bgr	bu	sc	Sod	pot	hemo	pcv	wc	rc	Htn	dm	cad	appet	pe	ane	class
48	100	1,015	0	2	2	2	2	2	103	79	5,3	135	6,3	6,3	19	7200	2,6	Yes	no	yes	poor	no	no	ckd
48	110	1,015	3	0	2	1	1	2	106	215	15,2	120	5,7	8,6	26	5000	2,5	Yes	no	yes	good	no	yes	ckd
54	90	1,025	1	0	1	2	2	2	150	18	1,2	140	4,2	9	26	10000	3,8	No	no	no	poor	yes	yes	ckd
59	70	1,010	1	3	2	2	2	2	424	55	1,7	138	4,5	12,6	37	10200	4,1	Yes	yes	yes	good	no	no	ckd
56	90	1,010	4	1	1	2	1	2	176	309	13,3	124	6,5	3,1	9	5400	2,1	yes	yes	no	poor	yes	yes	ckd
40	80	1,025	0	0	1	1	2	2	140	10	1,2	135	5	15	48	10400	4,5	no	no	no	good	no	no	notckd
23	80	1,025	0	0	1	1	2	2	70	36	1	150	4,6	17	52	9800	5	no	no	no	good	no	no	notckd
45	80	1,025	0	0	1	1	2	2	82	49	0,6	147	4,4	15,9	46	9100	4,7	no	no	no	good	no	no	notckd
57	80	1,025	0	0	1	1	2	2	119	17	1,2	135	4,7	15,4	42	6200	6,2	no	no	no	good	no	no	notckd
51	60	1,025	0	0	1	1	2	2	99	38	0,8	135	3,7	13	49	8300	5,2	no	no	no	good	no	no	notckd

Keterangan :

Age : umur
Bp : tekanan darah
Sg : *specific gravity*
Al : albumin
Su : gula darah
Rbc : status sel darah merah
Pc : *pus cell*
Pcc : pus cell clumps
Ba : bakteri
bgr : glukosa

bu : *blood urea nitrogen*
sc : serum kreatinin
sod : sodium blood test
pot : potassium test
hemo : hemoglobin
pcv : *packed cell volume*
wc : jumlah sel darah putih
rc : jumlah sel darah merah
htn : hipertensi
dm : diabetes mellitus

cad : penyakit jantung koroner
appet : tingkat nafsu makan
pe : *pedal edema*
ane : anemia
class : kelas
ckd : gagal ginjal kronik
notckd : tidak gagal ginjal kronik

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.2 Sampel Data Pasien Gagal Ginjal Kronik yang Kosong

age	bp	Sg	al	Su	rbc	Pc	pcc	Ba	bgr	Bu	sc	sod	Pot	hemo	pcv	wc	Rc	Htn	dm	cad	appet	pe	ane	Class
48	60	1.020	0	0	1	1	2	2	112	44	1.2	142	4.9	14.5	44	9400	6.4	No	no	no	good	no	no	notckd
24	70	1.025	0	0	1	1	2	2	140	23	0.6	140	4.7	16.3	48	5800	5.6	No	no	no	good	no	no	notckd
47	80	?	?	?	?	?	2	2	93	33	0.9	144	4.5	13.3	52	8100	5.2	No	no	no	good	no	no	notckd
55	80	1.025	0	0	1	1	2	2	130	50	1.2	147	5	15.5	41	9100	6.0	No	no	no	good	no	no	notckd
20	70	1020	0	0	1	1	2	2	123	44	1.0	135	3.8	14.6	44	5500	4.8	No	no	no	good	no	no	notckd
60	70	1020	0	0	1	1	2	2	?	?	?	?	?	16.4	43	10800	5.7	No	no	no	good	no	no	notckd
71	70	1.020	0	0	1	1	2	2	81	18	0.8	145	5.0	14.7	44	9800	6.0	No	no	no	good	no	no	notckd
39	70	1.025	0	0	1	1	2	2	124	22	0.6	137	3.8	13.4	43	?	?	No	no	no	good	no	no	notckd
56	70	1.025	0	0	1	1	2	2	70	46	1.2	135	4.9	15.9	50	11000	5.1	?	?	?	good	no	no	notckd
54	70	1.020	0	0	?	?	?	?	76	28	0.6	146	3.5	14.8	52	8400	5.9	No	No	no	good	no	no	notckd

Keterangan :

Age : umur
Bp : tekanan darah
Sg : *specific gravity*
Al : albumin
Su : gula darah
Rbc : status sel darah merah
Pc : *pus cell*
Pcc : pus cell clumps
Ba : bakteri
bgr : glukosa

bu : *blood urea nitrogen*
sc : serum kreatinin
sod : sodium blood test
pot : potassium test
hemo : hemoglobin
pcv : *packed cell volume*
wc : jumlah sel darah putih
rc : jumlah sel darah merah
htn : hipertensi
dm : diabetes mellitus

cad : penyakit jantung koroner
appet : tingkat nafsu makan
pe : *pedal edema*
ane : anemia
class : kelas
ckd : gagal ginjal kronik
notckd : tidak gagal ginjal kronik

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

B. Pembuatan Model Klasifikasi

Pembuatan model klasifikasi yang akan dilakukan oleh peneliti ini meliputi pembuatan model dengan menggunakan metode *improved* C4.5 dan pengujian model untuk mengetahui seberapa besar akurasi yang didapatkan.

1. Perhitungan *Improved* C4.5 pada *Training* Pertama

Di bawah ini merupakan perhitungan menggunakan excel dari penerapan *Improved* C4.5 untuk klasifikasi penyakit gagal ginjal dengan menggunakan 254 data *training*. Pada *Improved* C4.5 harus menentukan pohon keputusan yang kemudian akan menjadi *rule* untuk mengklasifikasikan penyakit gagal ginjal kronik.

a. Perhitungan Node 1

Perhitungan Node 1 adalah perhitungan awal yang digunakan untuk menentukan akar dari pohon keputusan yang dibuat.

Diketahui :

Jumlah kasus (N) = 254

Jumlah gagal ginjal kronik (p) = 154

Jumlah tidak gagal ginjal kronik (n) = 100

1) Perhitungan rasio gain umur

Perhitungan rasio gain umur (age) dihitung menggunakan Persamaan 2.3 sebagai berikut:

Age < 30 th (s1) = 26

Age 30 – 45 th (s2) = 55

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Age > 45 th (s3) = 174

Age < 30 th ckd (s11) = 12

Age < 30 th notckd (s12) = 14

Age 30 – 45 th ckd (s21) = 15

Age 30 – 45 th notckd (s21) = 40

Age >45 th ckd (s31) = 128

Age >45 th notckd (s32) = 46

Rasio Gain (Total, Age)

$$= \frac{\frac{p_n}{N} - \left\{\left(\frac{s_{11}s_{12}}{s_1}\right) + \left(\frac{s_{21}s_{22}}{s_2}\right) + \left(\frac{s_{31}s_{32}}{s_3}\right)\right\}}{\frac{s_1 s_2 s_3}{N}}$$

$$= \frac{\frac{154\times100}{254} - \left\{\left(\frac{12\times14}{26}\right) + \left(\frac{15\times40}{55}\right) + \left(\frac{128\times46}{174}\right)\right\}}{\frac{(26\times55\times174)}{254}}$$

$$= 0{,}00961632$$

2) Perhitungan rasio gain tekanan darah

Perhitungan rasio gain tekanan darah (bp) dihitung menggunakan Persamaan 2.3 sebagai berikut:

Bp < 60 mm/Hg (s1) = 3

Bp 60 – 80 mm/Hg (s2) = 203

Bp > 80 mm/Hg (s3) = 48

Bp < 60 mm/Hg ckd (s11) = 3

Bp < 60 mm/Hg notckd (s12) = 0

Bp 60 – 80 mm/Hg ckd (s21) = 103

Bp 60 – 80 mm/Hg notckd (s21) = 100

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Bp > 80 mm/Hg ckd (s31) = 48

Bp 60 – 80 mm/Hg notckd (s32) = 0

RasioGain (Total, Tekanan darah)

$$= \frac{\frac{154 \times 100}{254} - \left\{\left(\frac{3 \times 0}{3}\right) + \left(\frac{103 \times 100}{203}\right) + \left(\frac{48 \times 0}{48}\right)\right\}}{\frac{(3 \times 203 \times 48)}{254}}$$

$$= 0{,}08594401$$

3) Perhitungan rasio gain *Specific Gravity*

Perhitungan rasio gain *specific gravity* (Sg) dihitung menggunakan Persamaan 2.3 sebagai berikut:

Sg 1005 (s1) = 11

Sg 1010 (s2) = 57

Sg 1015 (s3) = 59

Sg 1020 (s4) = 78

Sg 1025 (s5) = 49

Sg 1005 ckd (s11) = 11

Sg 1005 notckd (s12) = 0

Sg 1010 ckd (s21) = 57

Sg 1010 notckd (s21) = 0

Sg 1015 ckd (s31) = 58

Sg 1015 notckd (s32) = 1

Sg 1020 ckd (s41) = 21

Sg 1020 notckd (s42) = 57

Sg 1025 ckd (s51) = 7

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Sg 1025 notckd (s52) = 42

RasioGain (Total, Specific Gravity)

$$= \frac{\frac{154\times100}{254} - \left\{\left(\frac{11\times0}{11}\right)+\left(\frac{57\times0}{57}\right)+\left(\frac{58\times1}{59}\right)+\left(\frac{21\times57}{78}\right)+\left(\frac{7\times42}{49}\right)\right\}}{\frac{(11\times57\times59\times78\times49)}{254}}$$

$$= 0{,}00006881$$

4) Perhitungan rasio gain *albumin*

Perhitungan rasio gain *albumin* (al) dihitung menggunakan Persamaan 2.3 sebagai berikut:

Al 0 (s1) = 133

Al 1 (s2) = 29

Al 2 (s3) = 42

Al 3 (s4) = 32

Al 4 (s5) = 17

Al 5 (s6) = 1

Al 0 ckd (s11) = 33

Al 0 notckd (s12) = 100

Al 1 ckd (s21) = 29

Al 1 notckd (s21) = 0

Al 2 ckd (s31) = 42

Al 2 notckd (s32) = 0

Al 3 ckd (s41) = 32

Al 3 notckd (s42) = 0

Al 4 ckd (s51) = 17

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Al 4 notckd (s52) = 0

Al 5 ckd (s61) = 1

Al 5 notckd (s62) = 0

RasioGain (Total, albumin)

$$= \frac{\frac{154\times100}{254} - \left\{\left(\frac{33\times100}{133}\right)+\left(\frac{29\times0}{29}\right)+\left(\frac{42\times0}{42}\right)+\left(\frac{32\times0}{32}\right)+\left(\frac{17\times0}{17}\right)+\left(\frac{1\times0}{1}\right)\right\}}{\frac{(133\times29\times42\times32\times17\times1)}{254}}$$

$$= 0{,}000103237$$

5) Perhitungan rasio gain gula darah

Perhitungan rasio gain gula darah (su) dihitung menggunakan Persamaan 2.3 sebagai berikut:

Su 0 (s1) = 187

Su 1 (s2) = 15

Su 2 (s3) = 29

Su 3 (s4) = 9

Su 4 (s5) = 13

Su 5 (s6) = 1

Su 0 ckd (s11) = 92

Su 0 notckd (s12) = 95

Su 1 ckd (s21) = 15

Su 1 notckd (s21) = 0

Su 2 ckd (s31) = 24

Su 2 notckd (s32) = 5

Su 3 ckd (s41) = 9

Su 3 notckd (s42) = 0

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Su 4 ckd (s51) = 13

Su 4 notckd (s52) = 0

Su 5 ckd (s61) = 1

Su 5 notckd (s62) = 0

RasioGain (Total, Gula darah)

$$= \frac{\frac{154\times100}{254} - \left\{\left(\frac{92\times95}{187}\right)+\left(\frac{15\times0}{15}\right)+\left(\frac{24\times0}{24}\right)+\left(\frac{9\times0}{9}\right)+\left(\frac{13\times0}{13}\right)+\left(\frac{1\times0}{1}\right)\right\}}{\frac{(187\times15\times24\times9\times13\times1)}{254}}$$

$$= 0{,}00026032$$

6) Perhitungan rasio gain status sel darah merah

Perhitungan rasio gain status sel darah merah (rbc) dihitung menggunakan Persamaan 2.3 sebagai berikut:

Rbc normal (s1) = 160

rbc abnormal (s2) = 94

rbc normal ckd (s11) = 60

rbc normal notckd (s12) = 100

rbc abnormal ckd (s21) = 94

rbc abnormal notckd (s21) = 0

RasioGain (Total, Status sel darah merah)

$$= \frac{\frac{154\times100}{254} - \left\{\left(\frac{60\times100}{160}\right)+\left(\frac{94\times0}{94}\right)\right\}}{\frac{(160\times94)}{254}}$$

$$= 0{,}39062500$$

7) Perhitungan rasio gain Status *Pus Cell*

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Perhitungan rasio gain Status *pus cell* (pc) dihitung menggunakan Persamaan 2.3 sebagai berikut:

Pc normal (s1) = 186

Pc abnormal (s2) = 68

Pc normal ckd (s11) = 86

Pc normal notckd (s12) = 100

Pc abnormal ckd (s21) = 68

Pc abnormal notckd (s21) = 0

RasioGain (Total, Status *pus cell*)

$$= \frac{\frac{154 \times 100}{254} - \left\{\left(\frac{86 \times 100}{186}\right) + \left(\frac{68 \times 0}{68}\right)\right\}}{\frac{(186 \times 68)}{254}}$$

$$= 0{,}28905076$$

8) Perhitungan rasio gain *Pus Cell Clumps*

Perhitungan rasio gain *pus cell clumps* (pcc) dihitung menggunakan Persamaan 2.3 sebagai berikut:

Pcc ada (s1) = 226

Pcc tidak ada (s2) = 28

Pcc ada ckd (s11) = 126

Pcc ada notckd (s12) = 100

Pcc tidak ada ckd (s21) = 28

Pcc tidak ada notckd (s21) = 0

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

RasioGain (Total, *Pus Cell Clumps*)

$$= \frac{\frac{154 \times 100}{254} - \left\{\left(\frac{126 \times 100}{126}\right) + \left(\frac{28 \times 0}{28}\right)\right\}}{\frac{(126 \times 28)}{254}}$$

$= 0{,}1957867$

9) Perhitungan rasio gain Status kehadiran bakteri

Perhitungan rasio gain Status kehadiran bakteri (ba) dihitung menggunakan Persamaan 2.3 sebagai berikut:

Ba ada (s1)	= 242
Ba tidak ada (s2)	= 12
Ba ada ckd (s11)	= 142
Ba ada notckd (s12)	= 100
Ba tidak ada ckd (s21)	= 12
Ba tidak ada notckd (s21)	= 0

RasioGain (Total, Status kehadiran bakteri)

$$= \frac{\frac{154 \times 100}{254} - \left\{\left(\frac{142 \times 100}{242}\right) + \left(\frac{12 \times 0}{12}\right)\right\}}{\frac{(242 \times 12)}{254}}$$

$= 0{,}17075336$

10) Perhitungan rasio gain kadar glukosa

Perhitungan rasio gain kadar glukosa (bgr) dihitung menggunakan Persamaan 2.3 sebagai berikut:

Bgr < 70 mgs / dl (s1)	= 7
Bgr 70 – 130 mgs / dl (s2)	= 150

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Bgr > 130 mgs / dl (s3) = 97

Bgr < 70 mgs / dl ckd (s11) = 7

Bgr < 70 mgs / dl notckd (s12) = 0

Bgr 70 – 130 mgs / dl ckd (s21) = 66

Bgr 70 – 130 mgs / dl notckd (s21) = 84

Bgr > 130 mgs / dl ckd (s31) = 81

Bgr >130 mgs / dl notckd (s31) = 16

RasioGain (Total, Kadar Glukosa)

$$= \frac{\frac{154 \times 100}{254} - \left\{\left(\frac{7 \times 0}{7}\right) + \left(\frac{66 \times 84}{150}\right) + \left(\frac{81 \times 16}{97}\right)\right\}}{\frac{(7 \times 150 \times 97)}{254}}$$

$$= 0{,}02570948$$

11) Perhitungan rasio gain *Blood urea nitrogen*

Perhitungan rasio gain *Blood urea nitrogen* (bun) dihitung menggunakan Persamaan 2.3 sebagai berikut:

Bun < 15 mgs / dl (s1) = 2

Bun 15 – 40 mgs / dl (s2) = 132

Bun > 40 mgs / dl (s3) = 120

Bun < 15 mgs / dl ckd (s11) = 2

Bun < 15 mgs / dl notckd (s12) = 0

Bun 15 – 40 mgs / dl ckd (s21) = 61

Bun 15 – 40 mgs / dl notckd (s21) = 71

Bun > 40 mgs / dl ckd (s31) = 93

Bun >40 mgs / dl notckd (s31) = 27

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

RasioGain (Total, *Blood urea nitrogen*)

$$= \frac{\frac{154 \times 100}{254} - \left\{\left(\frac{2 \times 0}{2}\right) + \left(\frac{61 \times 71}{132}\right) + \left(\frac{93 \times 27}{120}\right)\right\}}{\frac{(2 \times 132 \times 120)}{254}}$$

$$= 0{,}05527639$$

12) Perhitungan rasio gain Kadar serum kreatinin

Perhitungan rasio gain kadar serum kreatinin (sc) dihitung menggunakan Persamaan 2.3 sebagai berikut:

sc < 0,4 mgs / dl (s1)	= 0
sc 0,4 – 1,4 mgs / dl (s2)	= 140
sc > 1,4 mgs / dl (s3)	= 114
sc < 0,4 mgs / dl ckd (s11)	= 0
sc < 0,4 mgs / dl notckd (s12)	= 0
sc *0,4* – 1,4 mgs / dl ckd (s21)	= 113
sc 0,,4 – 1,4 mgs / dl notckd (s21)	= 27
sc > 1,4 mgs / dl ckd (s31)	= 42
sc >1,4 mgs / dl notckd (s31)	= 72

RasioGain (Total, Kadar serum kreatini)

$$= \frac{\frac{154 \times 100}{254} - \left\{\left(\frac{0 \times 0}{0}\right) + \left(\frac{113 \times 27}{140}\right) + \left(\frac{42 \times 72}{114}\right)\right\}}{\frac{(0 \times 140 \times 114)}{254}}$$

$$= 0{,}000000$$

13) Perhitungan rasio gain *Sodium blood test*

Perhitungan rasio gain *Sodium blod test* (sod) dihitung menggunakan Persamaan 2.3 sebagai berikut:

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Sod < 135 mEq / l (s1) =64

Sod 135 – 145 mEq / l (s2) = 159

Sod > 145 mEq / l (s3) = 31

Sod < 135 mEq / l ckd (s11) = 64

Sod < 135 mEq / l notckd (s12) = 0

Sod 135 – 145 mEq / l ckd (s21) = 85

Sod 135 – 145 mEq / l notckd (s21) = 74

Sod > 145 mEq / l ckd (s31) = 5

Sod > 145 mEq / l notckd (s31) = 26

RasioGain (Total, *Sodium blood test*)

$$= \frac{\frac{154 \times 100}{254} - \left\{\left(\frac{64 \times 0}{64}\right) + \left(\frac{85 \times 74}{159}\right) + \left(\frac{5 \times 26}{31}\right)\right\}}{\frac{(64 \times 159 \times 31)}{254}}$$

$$= 0{,}01358878$$

14) Perhitungan rasio gain *Potassium test*

Perhitungan rasio gain *potassium test* (pot) dihitung menggunakan Persamaan 2.3 sebagai berikut:

Pot < 3,5 mEq / l (s1) = 30

Pot 3,5 – 5,0 mEq / l (s2) = 202

Pot > 5,0 mEq / l (s3) = 22

Pot < 3,5 mEq / l ckd (s11) = 29

Pot < 3,5 mEq / l notckd (s12) = 1

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Pot 3,5 – 5,0 mEq / l ckd (s21) = 103

Pot 3,5 – 5,0 mEq / l notckd (s21) = 99

Pot > 5,0 mEq / l ckd (s31) = 22

Pot > 5,0 mEq / l notckd (s31) = 0

RasioGain (Total, *Potassium test*)

$$= \frac{\frac{154 \times 100}{254} - \left\{\left(\frac{29 \times 1}{30}\right) + \left(\frac{103 \times 99}{202}\right) + \left(\frac{22 \times 0}{22}\right)\right\}}{\frac{(30 \times 202 \times 22)}{254}}$$

$$= 0{,}01749547$$

15) Perhitungan rasio gain Kadar Hemoglobin

Perhitungan rasio gain Status hemoglobin (hemo) dihitung menggunakan Persamaan 2.3 sebagai berikut:

Hemo < 10 gms (s1) = 48

Hemo 10 – 18 gms (s2) = 206

Hemo > 18 gms (s3) = 0

Hemo < 10 gms ckd (s11) = 48

Hemo < 10 gms notckd (s12) = 0

Hemo 10 – 18 gms ckd (s21) = 106

Hemo < 10 – 18 gms notckd (s21) = 100

Hemo >18 gms ckd (s31) = 0

Hemo > 18 gms notckd (s31) = 0

RasioGain (Total, Hemoglobin)

$$= \frac{\frac{154 \times 100}{254} - \left\{\left(\frac{48 \times 0}{48}\right) + \left(\frac{106 \times 100}{206}\right) + \left(\frac{0 \times 0}{0}\right)\right\}}{\frac{(48 \times 206 \times 0)}{254}}$$

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

$= 0{,}000000$

16) Perhitungan rasio gain *Packed cell volume*

Perhitungan rasio gain *packed cell volume* (pcv) dihitung menggunakan Persamaan 2.3 sebagai berikut:

pcv < 36 % (s1)	= 98
pcv 36 - 50 % (s2)	= 139
pcv >50 % (s3)	= 17
pcv < 36 % ckd (s11)	= 98
pcv < 36 % notckd (s12)	= 0
pcv 36 - 50 % ckd (s21)	= 53
pcv 36 - 50 % notckd (s21)	= 86
pcv > 50 % ckd (s31)	= 3
pcv >50 % notckd (s31)	= 14

RasioGain (Total, *packed cell volume*)

$$= \frac{\frac{154 \times 100}{254} - \left\{\left(\frac{98 \times 0}{98}\right) + \left(\frac{53 \times 86}{139}\right) + \left(\frac{3 \times 14}{17}\right)\right\}}{\frac{(98 \times 139 \times 17)}{254}}$$

$= 0{,}02782464$

17) Perhitungan rasio gain sel darah putih

Perhitungan rasio gain sel darah putih (wc) dihitung menggunakan Persamaan 2.3 sebagai berikut:

wc < 4500 cell/cum (s1)	= 13
wc 4500 – 10000 cell/cum (s2)	= 189
wc >10000 cell/cum (s3)	= 52

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

wc < 4500 cell/cum ckd (s11) = 9

wc < 4500 cell/cum ckd (s12) = 4

wc 4500 – 10000 cell/cum ckd (s21) = 105

wc 4500 – 10000 cell/cum ckd (s21) = 84

wc >10000 cell/cum ckd (s31) = 40

wc >10000 cell/cum ckd (s31) = 12

RasioGain (Total,sel darah putih)

$$= \frac{\frac{154 \times 100}{254} - \left\{\left(\frac{9 \times 4}{13}\right) + \left(\frac{105 \times 84}{189}\right) + \left(\frac{40 \times 12}{52}\right)\right\}}{\frac{(13 \times 189 \times 52)}{254}}$$

$$= 0{,}00390303$$

18) Perhitungan rasio gain sel darah merah

Perhitungan rasio gain sel darah merah (rc) dihitung menggunakan Persamaan 2.3 sebagai berikut:

rc < 3,9 millions/cmm (s1) = 62

rc 3,9 – 5,7 millions/cmm (s2) = 161

rc >5,7 millions/cmm (s3) = 31

rc < 3,9 millions/cmm ckd (s11) = 62

rc < 3,9 millions/cmm notckd (s12) = 0

rc 3,9 – 5,7 millions/cmm ckd (s21) = 89

rc 3,9 - 5,7 millions/cmm notckd (s21) = 72

rc > 5,7 millions/cmm ckd (s31) = 28

rc >5,7 millions/cmm notckd (s31) = 3

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

RasioGain (Total,sel darah merah)

$$= \frac{\frac{154 \times 100}{254} - \left\{\left(\frac{62 \times 0}{62}\right) + \left(\frac{89 \times 72}{161}\right) + \left(\frac{28 \times 3}{31}\right)\right\}}{\frac{(62 \times 161 \times 31)}{254}}$$

$$= 0{,}01487266$$

19) Perhitungan rasio gain Hipertensi

Perhitungan rasio gain hipertensi (htn) dihitung menggunakan Persamaan 2.3 sebagai berikut:

Htn ya (s1)	= 98
Htn tidak (s2)	= 156
Htn ya ckd (s11)	= 98
Htn ya notckd (s12)	= 0
Htn tidak ckd (s21)	= 56
Htn tidak notckd (s21)	= 100

RasioGain (Total, Hipertensi)

$$= \frac{\frac{154 \times 100}{254} - \left\{\left(\frac{98 \times 0}{98}\right) + \left(\frac{56 \times 100}{156}\right)\right\}}{\frac{(98 \times 156)}{254}}$$

$$= 0{,}41091387$$

20) Perhitungan rasio gain Diabetes Mellitus

Perhitungan rasio gain diabetes mellitus (dm) dihitung menggunakan Persamaan 2.3 sebagai berikut:

dm ya (s1)	= 85
dm tidak (s2)	= 169
dm ya ckd (s11)	= 85

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

dm ya notckd (s12) = 0

dm tidak ckd (s21) = 69

dm tidak notckd (s21) = 100

RasioGain (Total, Diabetes Mellitus)

$$= \frac{\frac{154 \times 100}{254} - \left\{\left(\frac{85 \times 0}{85}\right) + \left(\frac{69 \times 100}{169}\right)\right\}}{\frac{(85 \times 169)}{254}}$$

$$= 0{,}35012780$$

21) Perhitungan rasio gain Penyakit Jantung Koroner

Perhitungan rasio gain penyakit jantung koroner (cad) dihitung menggunakan Persamaan 2.3 sebagai berikut:

cad ya (s1) = 233

cad tidak (s2) = 21

cad ya ckd (s11) = 133

cad ya ckd (s12) = 100

cad tidak ckd (s21) = 21

cad tidak tidak ckd (s21) = 0

RasioGain (Total, Penyakit Jantung Koroner)

$$= \frac{\frac{154 \times 100}{254} - \left\{\left(\frac{133 \times 100}{233}\right) + \left(\frac{21 \times 0}{21}\right)\right\}}{\frac{(233 \times 21)}{254}}$$

$$= 0{,}1841994$$

22) Perhitungan rasio gain tingkat nafsu makan

Perhitungan rasio gain tingkat nafsu makan (appet) dihitung menggunakan Persamaan 2.3 sebagai berikut:

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

appet baik (s1) = 200

appet kurang (s2) = 54

appet baik ckd (s11) = 100

appet baik notckd (s12) = 100

appet kurang ckd (s21) = 54

appet kurang notckd (s21) = 0

RasioGain (Total, tingkat nafsu makan)

$$= \frac{\frac{154 \times 100}{254} - \left\{\left(\frac{100 \times 100}{200}\right) + \left(\frac{54 \times 0}{54}\right)\right\}}{\frac{(200 \times 54)}{254}}$$

$= 0{,}250000$

23) Perhitungan rasio gain *Pedal edema*

Perhitungan rasio gain *pedal edema* (pe) dihitung menggunakan Persamaan 2.3 sebagai berikut:

pe ya (s1) = 208

pe tidak (s2) = 46

pe ya ckd (s11) = 108

pe ya notckd (s12) = 100

pe tidak ckd (s21) = 46

pe tidak notckd (s21) = 0

RasioGain (Total, *Pedal edema*)

$$= \frac{\frac{154 \times 100}{254} - \left\{\left(\frac{108 \times 100}{208}\right) + \left(\frac{46 \times 0}{46}\right)\right\}}{\frac{(208 \times 46)}{254}}$$

$= 0{,}2311391$

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

24) Perhitungan rasio gain Anemia

Perhitungan rasio gain anemia (ane) dihitung menggunakan Persamaan 2.3 sebagai berikut:

ane ya (s1) = 215

ane tidak (s2) = 39

ane ya ckd (s11) = 115

ane ya notckd (s12) = 100

ane tidak ckd (s21) = 39

ane tidak notckd (s21) = 0

RasioGain (Total, Anemia)

$$= \frac{\frac{154 \times 100}{254} - \left\{\left(\frac{115 \times 100}{215}\right) + \left(\frac{39 \times 0}{39}\right)\right\}}{\frac{(215 \times 39)}{254}}$$

$$= 0{,}216333$$

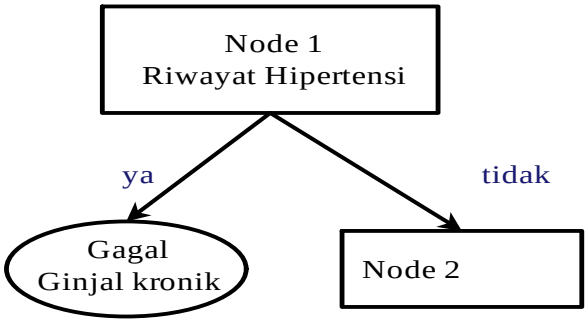
Berdasarkan perhitungan dengan *Improved C4.5* maka dapat disajikan model perhitungan node 1 seperti pada Tabel 4.3 dibawah. Tabel 4.3 merupakan hasil perhitungan pada Node 1, diketahui bahwa nilai RasioGain terbesar yaitu pada atribut riwayat hipertensi yaitu 0.41091387, sehingga atribut riwayat hipertensi menjadi node akar. Pada atribut riwayat hipertensi memiliki dua nilai atribut yaitu ya dan tidak. Nilai atribut yang pertama yaitu bernilai ya, dimana nilai tersebut sudah mengklasifikasikan data menjadi satu hasil yaitu hasil gagal ginjal kronik, sehingga tidak diperlukan perhitungan lebih lanjut, sedangkan nilai atribut yang kedua yaitu bernilai tidak belum mengklasifikasikan kasus menjadi satu keputusan sehingga perlu dilakukan perhitungan lebih lanjut.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.3 Perhitungan *Improved C4.5* Node 1

Atribut	RasioGain
Umur	0.00961632
Tekanan Darah	0.08594401
Specfic Gravity	0.00006881
Albumin	0.00010324
Gula Darah	0.00026032
Status sel darah merah	0.39062500
Status *Pus cell*	0.28905076
Pus cell clumps	0.19578667
Status kehadiran Bakteri	0.17075336
Kadar Glukosa dalam darah	0.02570948
Blood Urea Nitrogen	0.05527639
Kadar serum kreatinin	0.00000000
Hasil *Sodium blood test*	0.01358878
Hasil *Potassium test*	0.00000000
Kadar Hemoglobin	0.23564898
Packed cell volume	0.02782464
Jumlah sel darah putih	0.00390303
Jumlah sel darah merah	0.01487266
Riwayat Hipertensi	0.41091387
Riwayat diabete mellitus	0.35012780
Riwayat Penyakit Jantung Koroner	0.18419938
Tingkat nafsu makan	0.25000000
Riwayat *Pedal edema*	0.23113905
Riwayat Anemia	0.21633315

Berdasarkan Tabel 4.3 perhitungan Node 1 dapat dibentuk pohon keputusan sesuai dengan nilai rasio gain tertinggi atribut hipertensi yaitu sebesar 0.41091387, sehingga pohon keputusan seperti pada Gambar 4.2



Gambar 4.1 Hasil Node 1 *Improved C4.5*

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Gambar di atas merupakan pohon keputusan yang menjadi node akar yaitu atribut riwayat hipertensi yang memiliki dua cabang yaitu pada nilai ya mengklasifikasikan gagal ginjal kronik, sedangkan pada nilai tidak masih belum diketahui sehingga menjadi node 2.

b. Perhitungan Node 2

Berdasarkan perhitungan pada node 2 dengan *Improved C4.5* untuk menentukan atribut yang menjadi cabang kedua, maka dapat didapatkan nilai rasio gain seperti pada Tabel 4.4.

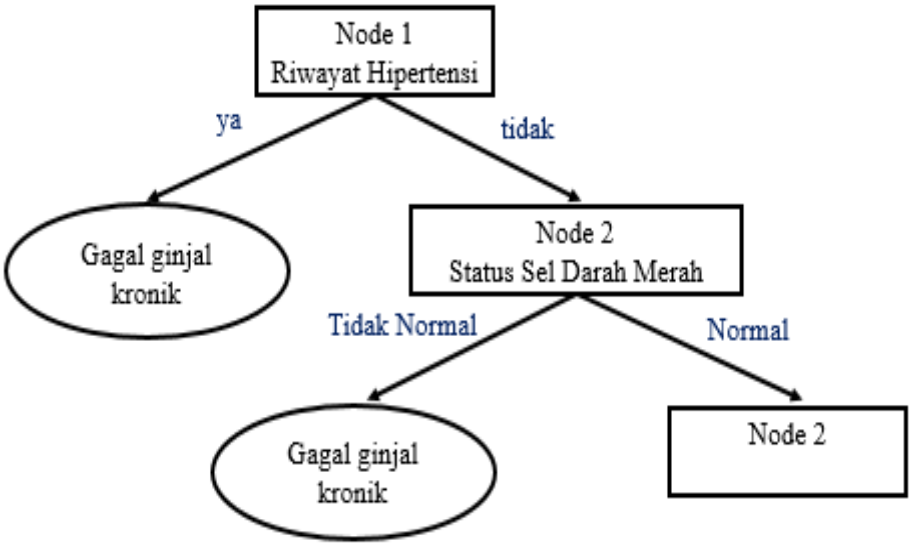
Tabel 4.4 Perhitungan *Improved C4.5* Node 2

Atribut	Rasio Gain
Umur	0.0040171
Tekanan Darah	0.2095803
Specfic Gravity	0.0014070
Albumin	0.0000000
Gula Darah	0.0000000
Status sel darah merah	0.7694675
Status *Pus cell*	0.5406574
Pus cell clumps	0.4565376
Status kehadiran Bakteri	0.4444444
Kadar Glukosa dalam darah	0.0407737
Blood Urea Nitrogen	0.0080687
Kadar serum kreatinin	0.0000000
Hasil *Sodium blood test*	0.0279543
Hasil *Potassium test*	0.0943803
Kadar Hemoglobin	0.0000000
Packed cell volume	0.0384719
Jumlah sel darah putih	0.0096754
Jumlah sel darah merah	0.0319909
Riwayat Hipertensi	0.0000000
Riwayat diabete mellitus	0.5327934
Riwayat Penyakit Jantung Koroner	0.4216563
Tingkat nafsu makan	0.4890215
Riwayat *Pedal edema*	0.4627701
Riwayat Anemia	0.4504302

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.4 merupakan hasil perhitungan pada Node 2, diketahui bahwa nilai RasioGain terbesar yaitu pada atribut status sel darah merah yaitu 0.7694675, sehingga atribut status sel darah merah menjadi node cabang kedua. Pada atribut status sel darah merah memiliki dua nilai atribut yaitu normal dan tidak normal. Nilai atribut yang pertama yaitu normal belum mengklasifikasikan kasus menjadi satu keputusan sehingga perlu dilakukan perhitungan lebih lanjut, sedangkan nilai atribut yang kedua yaitu tidak normal sudah mengklasifikasi menjadi satu nilai yaitu dengan hasil gagal ginjal kronik.

Berdasarkan Tabel 4.4 perhitungan node 2 dapat dibentuk pohon keputusan sesuai dengan nilai rasio gain tertinggi status sel darah merah yaitu sebesar 0.7694675, sehingga pohon keputusan seperti pada Gambar 4.2.



Gambar 4.2 Hasil Node 2 Improved C4.5

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Gambar 4.2 merupakan pohon keputusan yang menjadi node 2 yaitu atribut status sel darah merah yang memiliki dua cabang yaitu sesuai dengan nilai. Pada nilai tidak normal mengklasifikasikan menjadi gagal ginjal kronik, sedangkan pada nilai normal masih belum diketahui sehingga menjadi node 3.

c. Perhitungan Node 3

Berdasarkan perhitungan pada node 3 dengan *Improved C4.5* untuk menentukan atribut yang menjadi cabang ketiga, maka dapat didapatkan nilai rasio gain seperti pada Tabel 4.7.

Tabel 4.5 Perhitungan *Improved C4.5* Node 3

Atribut	RasioGain
Umur	0,0010976
Tekanan Darah	1,0521000
Specfic Gravity	0,0000000
Albumin	0,0000000
Gula Darah	0,0000000
Status sel darah merah	0,0000000
Status *Pus cell*	0,8573388
Pus cell clumps	0,8116224
Status kehadiran Bakteri	0,7971939
Kadar Glukosa dalam darah	0,0688502
Blood Urea Nitrogen	0,0016577
Kadar serum kreatinin	0,0000000
Hasil *Sodium blood test*	0,0427667
Hasil *Potassium test*	0,3492517
Kadar Hemoglobin	0,0000000
Packed cell volume	0,0691317
Jumlah sel darah putih	0,0112172
Jumlah sel darah merah	0,0404423
Riwayat Hipertensi	0,0000000
Riwayat diabete mellitus	0,8734387
Riwayat Penyakit Jantung Koroner	0,0000000
Tingkat nafsu makan	0,8264463
Riwayat *Pedal edema*	0,8116224
Riwayat Anemia	0,7971939

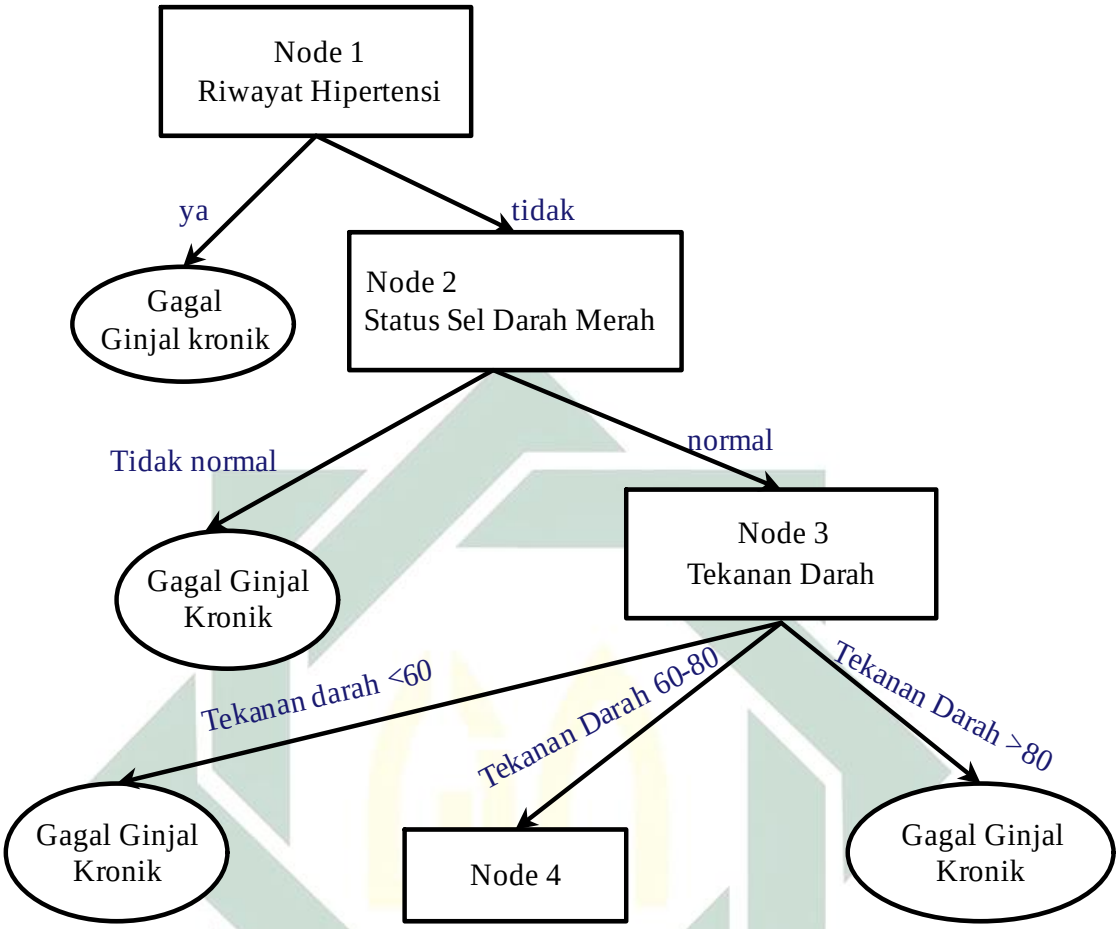
digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.7 merupakan hasil perhitungan pada node 3, diketahui bahwa nilai RasioGain terbesar yaitu pada atribut tekanan darah yaitu 1,0521000, Sehingga atribut tekanan darah menjadi node cabang ketiga. Pada atribut tekanan darah memiliki tiga nilai atribut yaitu tekanan darah < 60 mm/Hg, tekanan darah 60 – 80 mm/Hg dan tekanan darah > 80 mm/Hg. Nilai atribut yang pertama yaitu tekanan darah < 60 mm/Hg sudah mengklasifikan menjadi satu yaitu gagal ginjal kronik. Nilai atribut kedua yaitu tekanan darah 60 – 80 mm/Hg belum mengklasifikasikan kasus menjadi satu keputusan sehingga perlu dilakukan perhitungan lebih lanjut, sedangkan nilai atribut yang ketiga yaitu yaitu tekanan darah > 80 mm/Hg sudah mengklasifikan menjadi satu yaitu gagal ginjal kronik.

Berdasarkan Tabel 4.5 perhitungan node 3 dapat dibentuk pohon keputusan sesuai dengan nilai rasio gain tertinggi tekanan darah yaitu sebesar 1.0521000, sehingga pohon keputusan pada node 3 seperti pada Gambar 4.2.

Gambar 4.2 merupakan pohon keputusan yang menjadi node 3 yaitu atribut tekanan darah yang memiliki tiga cabang yaitu sesuai dengan nilai atribut. Pada rentang nilai < 60 mm/Hg mengklasifikasikan menjadi gagal ginjal kronik. Pada rentang nilai 60 – 80 mm/Hg masih belum mengklasifikan nilai atribut masih membutuhkan perhitungan lagi sehingga menjadi node 4. Pada rentang nilai > 80 mm/Hg mengklasifikasikan menjadi gagal ginjal kronik.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id



Gambar 4.3 Hasil Node 3 Improved C4.5

d. Perhitungan Node 4

Berdasarkan perhitungan pada node 4 dengan *Improved C4.5* untuk menentukan atribut yang menjadi cabang keempat, maka dapat didapatkan nilai rasio gain seperti pada Tabel 4.6.

Tabel 4.6 merupakan hasil perhitungan pada Node 4, diketahui bahwa nilai RasioGain terbesar yaitu pada atribut Riwayat diabetes mellitus yaitu 0,9245562, sehingga atribut riwayat diabetes mellitus menjadi node cabang keempat. Pada atribut riwayat diabetes mellitus memiliki dua nilai atribut yaitu ya dan tidak. Nilai atribut yang pertama yaitu ya sudah mengklasifikan

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

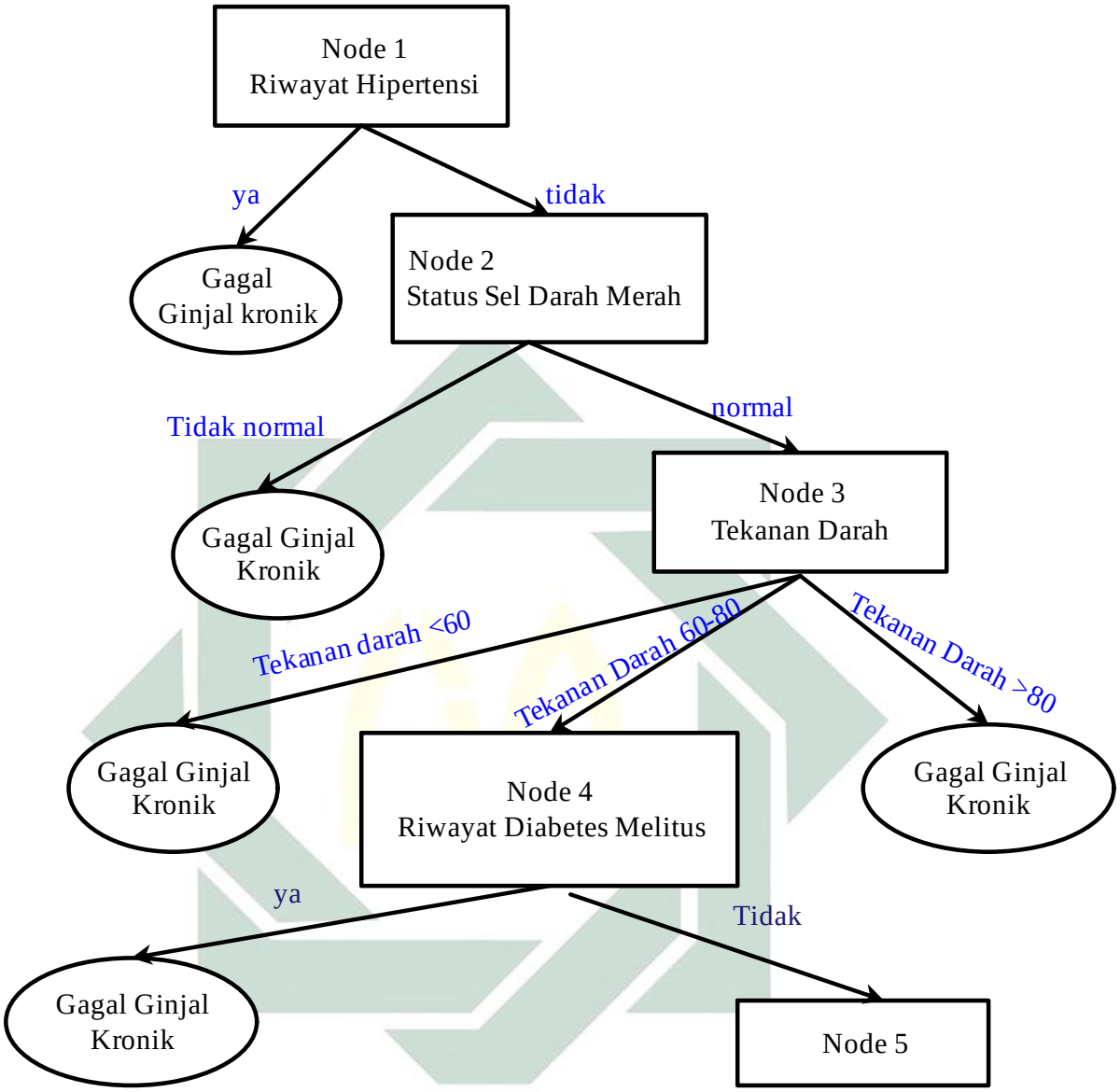
menjadi satu yaitu gagal ginjal kronik. Nilai atribut kedua yaitu tidak belum mengklasifikasikan kasus menjadi satu keputusan sehingga perlu dilakukan perhitungan lebih lanjut.

Tabel 4.6 Perhitungan *Improved C4.5* Node 4

Atribut	RasioGain
Umur	0,0002936
Tekanan Darah	0,0000000
Specfic Gravity	0,0000000
Albumin	0,0000000
Gula Darah	0,0000000
Status sel darah merah	0,0000000
Status *Pus cell*	0,8899964
Pus cell clumps	0,8573388
Status kehadiran Bakteri	0,8573388
Kadar Glukosa dalam darah	0,0925220
Blood Urea Nitrogen	0,0005546
Kadar serum kreatinin	0,0000000
Hasil *Sodium blood test*	0,0452749
Hasil *Potassium test*	0,7645282
Kadar Hemoglobin	0,0000000
Packed cell volume	0,0706749
Jumlah sel darah putih	0,0864068
Jumlah sel darah merah	0,0425658
Riwayat Hipertensi	0,0000000
Riwayat diabete mellitus	0,9245562
Riwayat Penyakit Jantung Koroner	0,0000000
Tingkat nafsu makan	0,8734387
Riwayat *Pedal edema*	0,8734387
Riwayat Anemia	0,8573388

Berdasarkan Tabel 4.5 perhitungan node 4 dapat dibentuk pohon keputusan sesuai dengan nilai rasio gain tertinggi yaitu riwayat diabetes mellitus sebesar 0.9245562, sehingga pohon keputusan pada node 4 seperti pada Gambar 4.4.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id



Gambar 4.4 Hasil Node 4 *Improved C4.5*

Gambar 4.4 merupakan pohon keputusan yang menjadi node 4 yaitu atribut Riwayat diabetes mellitus yang memiliki dua cabang yaitu ya dan tidak sesuai dengan nilai atribut. Pada nilai ya mengklasifikasikan menjadi gagal ginjal kronik. Pada nilai tidak masih belum mengklasifikan nilai atribut masih membutuhkan perhitungan lagi sehingga menjadi node 5.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

e. Perhitungan Node 5

Berdasarkan perhitungan pada node 5 dengan *Improved C4.5* untuk menentukan atribut yang menjadi cabang kelima, maka dapat didapatkan nilai rasio gain seperti pada Tabel 4.8.

Tabel 4.7 Perhitungan *Improved C4.5* Node 5

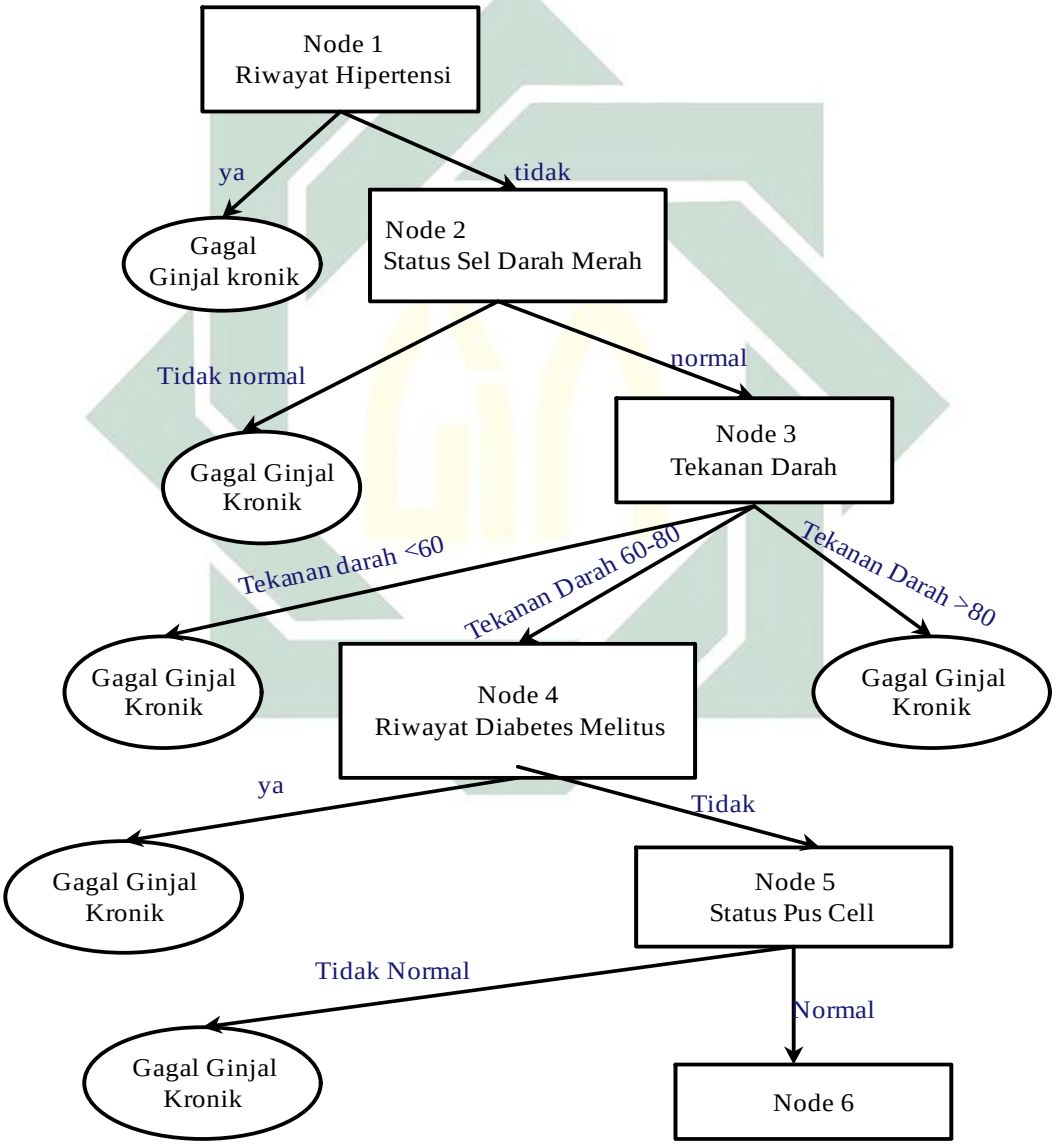
Atribut	RasioGain
Umur	0,0001551
Tekanan Darah	0,0000000
Specfic Gravity	0,0000000
Albumin	0,0000000
Gula Darah	0,0000000
Status sel darah merah	0,0000000
Status *Pus cell*	0,9611688
Pus cell clumps	0,3141986
Status kehadiran Bakteri	0,3141986
Kadar Glukosa dalam darah	0,0676846
Blood Urea Nitrogen	0,0000908
Kadar serum kreatinin	0,0000000
Hasil *Sodium blood test*	0,0487009
Hasil *Potassium test*	0,7590222
Kadar Hemoglobin	0,0000000
Packed cell volume	0,0798406
Jumlah sel darah putih	0,0006098
Jumlah sel darah merah	0,0478476
Riwayat Hipertensi	0,0000000
Riwayat diabete mellitus	0,0000000
Riwayat Penyakit Jantung Koroner	0,0000000
Tingkat nafsu makan	0,0000000
Riwayat *Pedal edema*	0,9425959
Riwayat Anemia	0,0000000

Tabel 4.8 merupakan hasil perhitungan pada Node 5, diketahui bahwa nilai Rasio Gain terbesar yaitu pada atribut Status *Pus cell* yaitu 0,9611688, sehingga atribut Status *Pus Cell* menjadi node cabang kelima. Pada atribut Status *Puss Cell* memiliki dua nilai atribut yaitu normal dan tidak normal. Nilai atribut yang pertama yaitu tidak normal sudah mengklasifikan menjadi satu yaitu gagal ginjal kronik. Nilai atribut kedua yaitu normal belum

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

mengklasifikasikan kasus menjadi satu keputusan sehingga perlu dilakukan perhitungan lebih lanjut.

Berdasarkan Tabel 4.8 perhitungan node 5 dapat dibentuk pohon keputusan sesuai dengan nilai rasio gain tertinggi yaitu status *pus cell* sebesar 0.9245562, sehingga pohon keputusan pada node 4 seperti pada Gambar 4.5.



Gambar 4.5 Hasil Node 5 *Improved C4.5*

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Gambar 4.5 node 5 *improved C4.5* merupakan pohon keputusan yang menjadi node 5 yaitu atribut Status *Pus cell* yang memiliki dua cabang yaitu normal dan tidak normal sesuai dengan nilai atribut. Pada nilai tidak normal mengklasifikasikan menjadi gagal ginjal kronik. Pada nilai normal masih belum mengklasifikan nilai atribut masih membutuhkan perhitungan lagi sehingga menjadi node 6.

f. Perhitungan Node 6

Berdasarkan perhitungan pada node 6 dengan *Improved C4.5* untuk menentukan atribut yang menjadi cabang kelima, maka dapat didapatkan nilai rasio gain seperti pada Tabel 4.8. Tabel 4.8 merupakan hasil perhitungan pada Node 6, diketahui bahwa nilai RasioGain terbesar yaitu pada atribut *Packed cell volume* yaitu 0,0814223. Sehingga atribut *Packed Cell volume* menjadi node cabang keenam. Pada atribut *Packed Cell volume* memiliki tiga nilai atribut yaitu pcv < 36%, 36%< pcv < 50%, dan pcv> 50%. Nilai atribut yang pertama yaitu pcv < 36% sudah mengklasifikan menjadi satu yaitu gagal ginjal kronik. Nilai atribut kedua yaitu 36%< pcv < 50% sudah mengklasifikasikan kasus menjadi satu keputusan yaitu tidak gagal ginjal. Nilai atribut kedua yaitu pcv > 50% sudah mengklasifikasikan kasus menjadi satu keputusan yaitu tidak gagal ginjal kronik.

Gambar 4.9 Node 6 *Improved* C4.5 merupakan hasil dari pohon keputusan, diketahui bahwa semua kasus sudah masuk dalam kelas. Sehingga akan tebentuk sejumlah aturan dalam pohon keputusan tersebut.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.8 Perhitungan *Improved C4.5* Node 6

Atribut	Rasio Gain
Umur	0,0001424
Tekanan Darah	0,0000000
Specfic Gravity	0,0000000
Albumin	0,0000000
Gula Darah	0,0000000
Status sel darah merah	0,0000000
Status *Pus cell*	0,0000000
Pus cell clumps	0,0000000
Status kehadiran Bakteri	0,0000000
Kadar Glukosa dalam darah	0,0000000
Blood Urea Nitrogen	0,0004031
Kadar serum kreatinin	0,0000000
Hasil *Sodium blood test*	0,0509538
Hasil *Potassium test*	0,0000000
Kadar Hemoglobin	0,0000000
Packed cell volume	0,0814223
Jumlah sel darah putih	0,0001803
Jumlah sel darah merah	0,0000000
Riwayat Hipertensi	0,0000000
Riwayat diabete mellitus	0,0000000
Riwayat Penyakit Jantung Koroner	0,0000000
Tingkat nafsu makan	0,0000000
Riwayat *Pedal edema*	0,0000000
Riwayat Anemia	0,0000000

Berdasarkan Tabel 4.8 perhitungan node 6 dapat dibentuk pohon keputusan sesuai dengan nilai rasio gain tertinggi yaitu status *packed cell volume* sebesar 0.9245562, sehingga pohon keputusan pada node 6 seperti pada Gambar 4.6. Perhitungan node berhenti setelah semua data training terpartisi, sehingga semua data masuk dalam kelas.

Setelah pohon keputusan didapatkan dibentuk aturan sebagai berikut:

“ JIKA riwayat hipertensi = ya MAKA class = GAGAL GINJAL KRONIK”

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

“ JIKA riwayat hipertensi = tidak DAN Status sel darah merah = tidak normal MAKA class = GAGAL GINJAL KRONIK”

“ JIKA riwayat hipertensi = tidak DAN Status sel darah merah = normal DAN Tekanan darah < 60 mm/Hg MAKA class = GAGAL GINJAL KRONIK”

“ JIKA riwayat hipertensi = tidak DAN Status sel darah merah = normal DAN Tekanan darah > 80 mm/Hg MAKA class = GAGAL GINJAL KRONIK”

“ JIKA riwayat hipertensi = tidak DAN Status sel darah merah = normal DAN Tekanan darah 60 – 80 mm/Hg DAN riwayat diabetes mellitus = ya MAKA class = GAGAL GINJAL KRONIK”

“ JIKA riwayat hipertensi = tidak DAN Status sel darah merah = normal DAN Tekanan darah 60 – 80 mm/Hg DAN riwayat diabetes mellitus = tidak DAN status *pus cell* = tidak normal MAKA class = GAGAL GINJAL KRONIK”

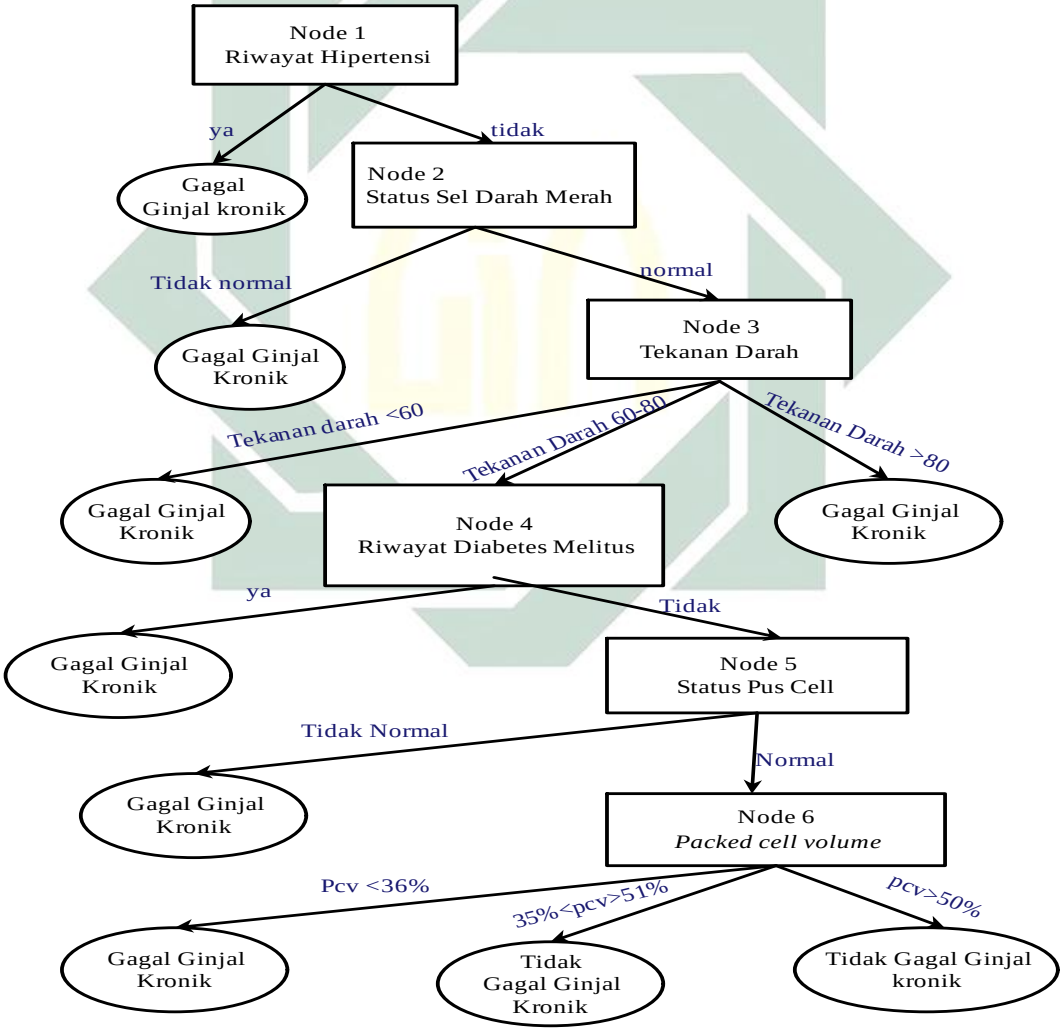
“ JIKA riwayat hipertensi = tidak DAN Status sel darah merah = normal DAN Tekanan darah 60 – 80 mm/Hg DAN riwayat diabetes mellitus = tidak DAN status *pus cell* = normal DAN *packed cell volume* < 36 % MAKA class = GAGAL GINJAL KRONIK”

“ JIKA riwayat hipertensi = tidak DAN Status sel darah merah = normal DAN Tekanan darah 60 – 80 mm/Hg DAN riwayat diabetes mellitus =

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

tidak DAN status *pus cell* = normal DAN *packed cell volume* 36 – 50% MAKA class = TIDAK GAGAL GINJAL KRONIK”

“ JIKA riwayat hipertensi = tidak DAN Status sel darah merah = normal DAN Tekanan darah 60 – 80 mm/Hg DAN riwayat diabetes mellitus = tidak DAN status *pus cell* = normal DAN *packed cell volume* >50% MAKA class = TIDAK GAGAL GINJAL KRONIK”



Gambar 4.6 Hasil Node 6 *Improved C4.5*

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

2. Perhitungan *Improved* C4.5 pada *Training* Kedua

Perhitungan *Improved C4.5* pada *training* kedua ini merupakan perhitungan menggunakan excel dari penerapan *Improved* C4.5 untuk klasifikasi penyakit gagal ginjal dengan menggunakan 254 data *training*. Pada *Improved* C4.5 harus menentukan pohon keputusan yang kemudian akan menjadi *rule* untuk mengklasifikasikan penyakit gagal ginjal kronik. Perhitungan *improved C4.5* dapat dilihat pada Tabel 4.9.

Berdasarkan Tabel 4.9 dapat diketahui nilai rasio gain pada tiap node. Perhitungan nilai rasio gain tertinggi pada node pertama adalah status sel darah merah sebesar 0.3718025, sehingga status sel darah merah sebagai node pertama. Perhitungan nilai rasio gain tertinggi pada node kedua adalah riwayat hipertensi sebesar 0.6718624, sehingga riwayat hipertensi sebagai node kedua.

Perhitungan nilai rasio gain tertinggi pada node ketiga adalah status *pus cell* sebesar 0.7971939, sehingga status *pus cell* sebagai node ketiga. Perhitungan nilai rasio gain tertinggi pada node keempat adalah riwayat diabetes mellitus sebesar 0.8573388, sehingga riwayat diabetes mellitus sebagai node keempat. Perhitungan nilai rasio gain tertinggi pada node kelima adalah status bakteri sebesar 0.8734387, sehingga status bakteri sebagai node kelima. Perhitungan nilai rasio gain tertinggi pada node keenam adalah tekanan darah sebesar 0.9425959, sehingga tekanan darah sebagai node keenam.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Perhitungan nilai rasio gain tertinggi pada node ketujuh adalah riwayat penyakit jantung sebesar 0.9611688, sehingga riwayat penyakit jantung sebagai node ketujuh. Perhitungan nilai rasio gain tertinggi pada node kedelapan adalah kadar glukosa sebesar 0.0814223, sehingga kadar glukosa sebagai node kedelapan. Perhitungan node berhenti pada node kedelapan setelah semua data terpartisi.

Setelah pohon keputusan didapatkan dibentuk aturan sebagai berikut:

" JIKA status sel darah merah = abnormal MAKA class = GAGAL GINJAL KRONIK"

" JIKA status sel darah merah = normal DAN riwayat hipertensi = ya MAKA class = GAGAL GINJAL KRONIK"

" JIKA status sel darah merah = normal DAN riwayat hipertensi = tidak DAN status *pus cell* = tidak normal MAKA class = GAGAL GINJAL KRONIK"

"JIKA status sel darah merah = normal DAN riwayat hipertensi = tidak DAN status *pus cell* = normal DAN riwayat diebetes mellitus = ya MAKA class = GAGAL GINJAL KRONIK"

"JIKA status sel darah merah = normal DAN riwayat hipertensi = tidak DAN status *pus cell* = normal DAN riwayat diebetes mellitus = tidak DAN status bakteri = ada MAKA class = GAGAL GINJAL KRONIK"

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

"JIKA status sel darah merah = normal DAN riwayat hipertensi = tidak DAN status *pus cell* = normal DAN riwayat diebetes mellitus = tidak DAN status bakteri = tidak ada DAN tekanan darah < 60 mm/Hg MAKA class = GAGAL GINJAL KRONIK"

"JIKA status sel darah merah = normal DAN riwayat hipertensi = tidak DAN status *pus cell* = normal DAN riwayat diebetes mellitus = tidak DAN status bakteri = tidak ada DAN tekanan darah > 80 mm/Hg MAKA class = GAGAL GINJAL KRONIK"

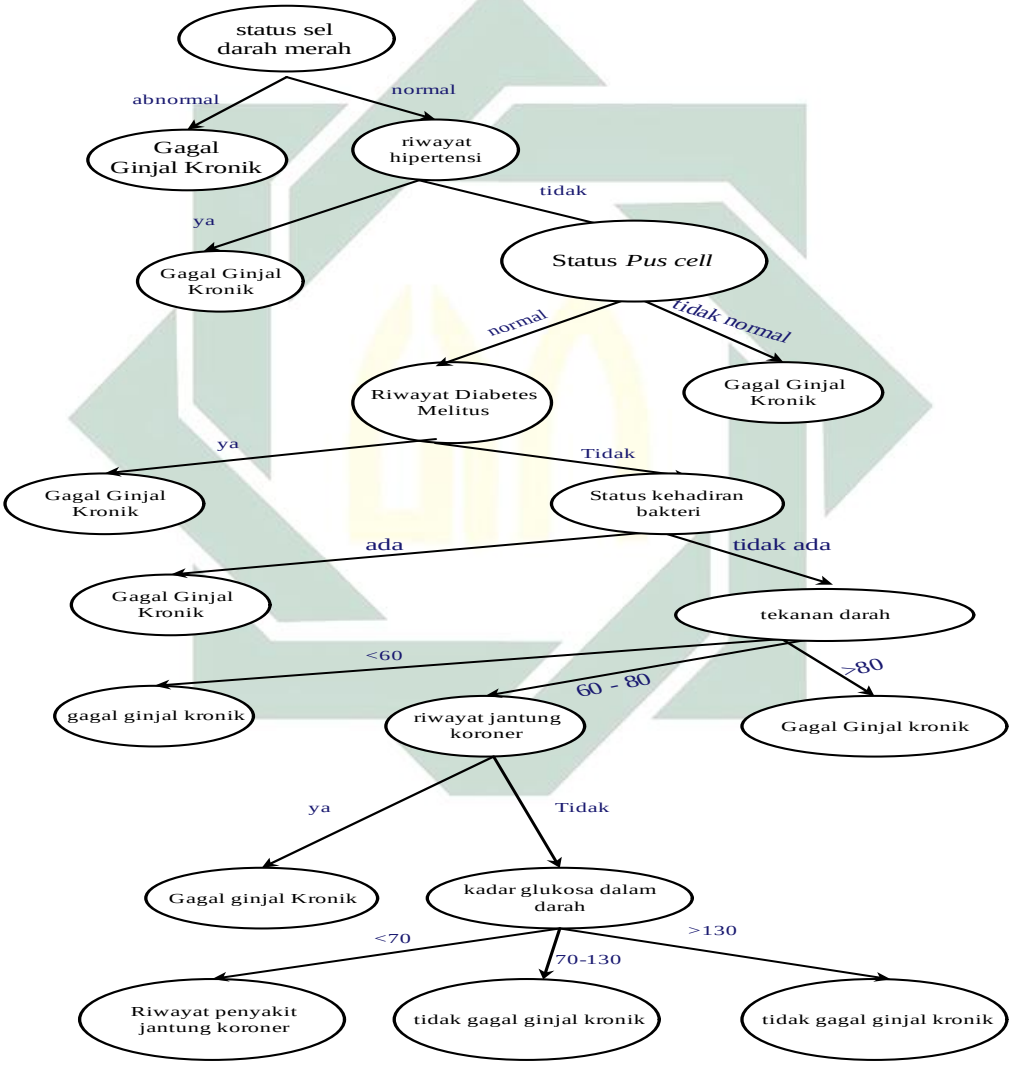
"JIKA status sel darah merah = normal DAN riwayat hipertensi = tidak DAN status *pus cell* = normal DAN riwayat diebetes mellitus = tidak DAN status bakteri = tidak ada DAN tekanan darah 60 – 80 mm/Hg DAN riwayat penyakit jantung = ya MAKA class = GAGAL GINJAL KRONIK"

"JIKA status sel darah merah = normal DAN riwayat hipertensi = tidak DAN status *pus cell* = normal DAN riwayat diebetes mellitus = tidak DAN status bakteri = tidak ada DAN tekanan darah 60 – 80 mm/Hg DAN riwayat penyakit jantung = tidak DAN kadar glukosa < 70 mgs/dl MAKA class = GAGAL GINJAL KRONIK"

"JIKA status sel darah merah = normal DAN riwayat hipertensi = tidak DAN status *pus cell* = normal DAN riwayat diebetes mellitus = tidak DAN status bakteri = tidak ada DAN tekanan darah 60 – 80 mm/Hg DAN riwayat penyakit jantung = tidak DAN kadar glukosa 70 – 130 mgs/dl MAKA class = TIDAK GAGAL GINJAL KRONIK"

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

"JIKA status sel darah merah = normal DAN riwayat hipertensi = tidak DAN status *pus cell* = normal DAN riwayat diebetes mellitus = tidak DAN status bakteri = tidak ada DAN tekanan darah 60 – 80 mm/Hg DAN riwayat penyakit jantung = tidak DAN kadar glukosa > 130 mgs/dl MAKA class = TIDAK GAGAL GINJAL KRONIK"



Gambar 4.7 Pohon Keputusan Training Kedua

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.9 Perhitungan Tiap Node dengan *Improved C4.5*

Atribut	Rasio gain Node 1	Rasio gain Node 2	Rasio gain Node 3	Rasio gain Node 4	Rasio gain Node 5	Rasio gain Node 6	Rasio gain Node 7	Rasio gain Node 8
Umur (age)	0,00411	0,00665	0,00052	0,00026	0,00061	0,00075	0,00008	0,00015
Tekanan Darah (bp)	0,04482	0,16065	0,30992	0,66750	0,78550	0,94260	0,00000	0,00000
*Specfic Gravity (*Sg)	0,00017	0,00231	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
Albumin (al)	0,00010	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
Gula Darah (su)	0,00019	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
Status sel darah merah (rbc)	0,37180	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
Status *Pus cell* (pc)	0,28597	0,52510	0,79719	0,00000	0,00000	0,00000	0,00000	0,00000
Pus cell clumps (pcc)	0,19069	0,45043	0,73051	0,81162	0,00000	0,00000	0,00000	0,00000
Status kehadiran Bakteri (ba)	0,17955	0,41091	0,71818	0,81162	0,87344	0,00000	0,00000	0,00000
Kadar Glukosa dalam darah (bgr)	0,01816	0,10700	0,05603	0,06126	0,07541	0,07629	0,08010	0,08142
Blood Urea Nitrogen (bu)	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
Kadar serum kreatinin (sc)	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
Hasil *Sodium blood test* (sod)	0,01431	0,06442	0,03994	0,04403	0,04566	0,05213	0,00000	0,00000
Hasil *Potassium test* (pot)	0,01532	0,06128	0,29157	0,44759	0,00000	0,00000	0,00000	0,00000
Kadar Hemoglobin (hemo)	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
Packed cell volume (pcv)	0,02074	0,03936	0,04964	0,05453	0,05477	0,05330	0,05825	0,00000
Jumlah sel darah putih (wc)	0,00643	0,02207	0,02838	0,01503	0,00473	0,00553	0,00048	0,00022
Jumlah sel darah merah (rc)	0,01447	0,02511	0,03733	0,04191	0,04319	0,04155	0,04390	0,00000
Riwayat Hipertensi (htn)	0,35013	0,67186	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
Riwayat diabete mellitus (dm)	0,32653	0,64000	0,75614	0,85734	0,00000	0,00000	0,00000	0,00000
Riwayat Penyakit Jantung Koroner (cad)	0,18108	0,41091	0,68301	0,81162	0,87344	0,89000	0,96117	0,00000
Tingkat nafsu makan (appet)	0,23565	0,48225	0,71818	0,82645	0,87344	0,00000	0,00000	0,00000
Riwayat *Pedal edema* (pe)	0,23114	0,48902	0,73051	0,81162	0,87344	0,89000	0,00000	0,00000
Riwayat Anemia (ane)	0,21042	0,46913	0,71818	0,81162	0,87344	0,89000	0,96117	0,00000

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

3. Perhitungan *Improved* C4.5 pada *Training* Ketiga

Perhitungan *Improved C4.5* pada *training* ketiga ini merupakan perhitungan menggunakan excel dari penerapan *Improved* C4.5 untuk klasifikasi penyakit gagal ginjal dengan menggunakan 254 data *training*. Pada *Improved* C4.5 harus menentukan pohon keputusan yang kemudian akan menjadi *rule* untuk mengklasifikasikan penyakit gagal ginjal kronik. Perhitungan *improved C4.5* dapat dilihat pada Tabel 4.10.

Berdasarkan Tabel 4.10 dapat diketahui nilai rasio gain pada tiap node. Perhitungan nilai rasio gain tertinggi pada node pertama adalah riwayat hipertensi sebesar 0.427186, sehingga riwayat hipertensi sebagai node pertama. Perhitungan nilai rasio gain tertinggi pada node kedua adalah status sel darah merah sebesar 0.6718624, sehingga status sel darah merah sebagai node kedua.

Perhitungan nilai rasio gain tertinggi pada node ketiga adalah status *pus cell* sebesar 0.7971939, sehingga status *pus cell* sebagai node ketiga. Perhitungan nilai rasio gain tertinggi pada node keempat adalah riwayat diabetes mellitus sebesar 0.841680, sehingga riwayat diabetes mellitus sebagai node keempat. Perhitungan nilai rasio gain tertinggi pada node kelima adalah status bakteri sebesar 0.857339, sehingga status bakteri sebagai node kelima. Perhitungan nilai rasio gain tertinggi pada node keenam adalah riwayat penyakit jantung sebesar 0.873439, sehingga riwayat penyakit jantung sebagai node keenam.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Perhitungan nilai rasio gain tertinggi pada node ketujuh adalah riwayat anemia sebesar 0.889996, sehingga riwayat anemia sebagai node ketujuh. Perhitungan nilai rasio gain tertinggi pada node kedelapan adalah tekanan darah sebesar 0.9702364, sehingga tekanan darah sebagai node kedelapan. Perhitungan nilai rasio gain tertinggi pada node kesembilan adalah kadar glukosa sebesar 0.0778029, sehingga kadar glukosa sebagai node kesembilan. Perhitungan node berhenti pada node kedelapan setelah semua data terpartisi.

Setelah pohon keputusan didapatkan dibentuk aturan sebagai berikut:

" JIKA riwayat hipertensi = ya MAKA class = GAGAL GINJAL KRONIK"

" JIKA riwayat hipertensi = tidak DAN status sel darah merah = abnormal MAKA class = GAGAL GINJAL KRONIK"

" JIKA riwayat hipertensi = tidak DAN status sel darah merah = normal DAN status *pus cell* = tidak normal MAKA class = GAGAL GINJAL KRONIK"

" JIKA riwayat hipertensi = tidak DAN status sel darah merah = normal DAN status *pus cell* = normal DAN riwayat diabetes mellitus = ya MAKA class = GAGAL GINJAL KRONIK"

" JIKA riwayat hipertensi = tidak DAN status sel darah merah = normal DAN status *pus cell* = normal DAN riwayat diabetes mellitus = tidak DAN status bakteri = ada MAKA class = GAGAL GINJAL KRONIK"

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

“ JIKA riwayat hipertensi = tidak DAN status sel darah merah = normal DAN status *pus cell* = normal DAN riwayat diabetes mellitus = tidak DAN status bakteri = tidak ada DAN riwayat penyakit jantung = ya MAKA class = GAGAL GINJAL KRONIK”

“ JIKA riwayat hipertensi = tidak DAN status sel darah merah = normal DAN status *pus cell* = normal DAN riwayat diabetes mellitus = tidak DAN status bakteri = tidak ada DAN riwayat penyakit jantung = tidak DAN riwayat anemia = ya MAKA class = GAGAL GINJAL KRONIK”

“ JIKA riwayat hipertensi = tidak DAN status sel darah merah = normal DAN status *pus cell* = normal DAN riwayat diabetes mellitus = tidak DAN status bakteri = tidak ada DAN riwayat penyakit jantung = tidak DAN riwayat anemia = tidak DAN tekanan darah < 60 mm/Hg MAKA class = GAGAL GINJAL KRONIK”

“ JIKA riwayat hipertensi = tidak DAN status sel darah merah = normal DAN status *pus cell* = normal DAN riwayat diabetes mellitus = tidak DAN status bakteri = tidak ada DAN riwayat penyakit jantung = tidak DAN riwayat anemia = tidak DAN tekanan darah > 80 mm/Hg MAKA class = GAGAL GINJAL KRONIK”

“ JIKA riwayat hipertensi = tidak DAN status sel darah merah = normal DAN status *pus cell* = normal DAN riwayat diabetes mellitus = tidak DAN status bakteri = tidak ada DAN riwayat penyakit jantung = tidak DAN

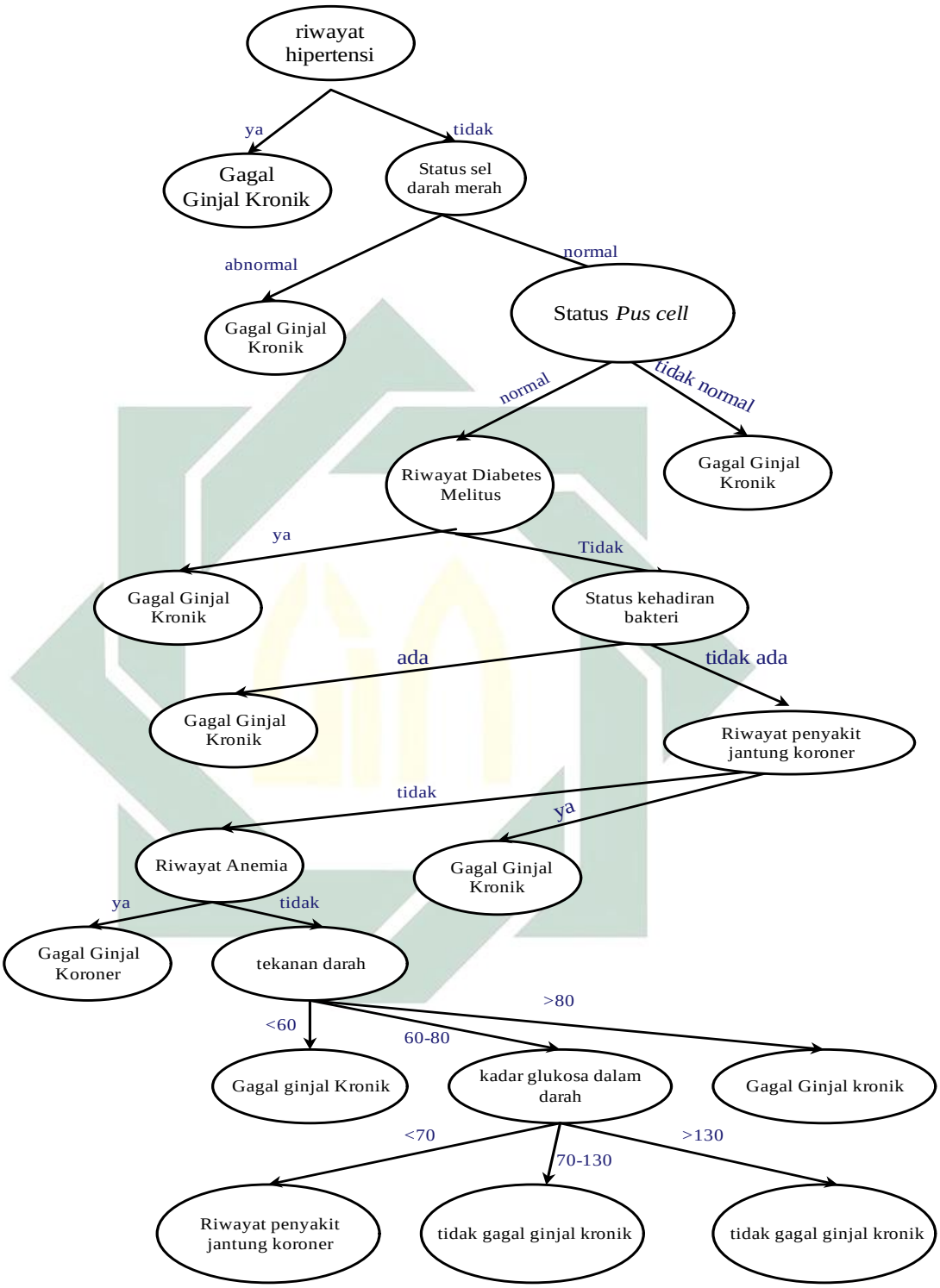
digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

riwayat anemia = tidak DAN tekanan darah 60 – 80 mm/Hg DAN kadar glukosa < 70 mgs/dl MAKA class = GAGAL GINJAL KRONIK”

“ JIKA riwayat hipertensi = tidak DAN status sel darah merah = normal DAN status *pus cell* = normal DAN riwayat diabetes mellitus = tidak DAN status bakteri = tidak ada DAN riwayat penyakit jantung = tidak DAN riwayat anemia = tidak DAN tekanan darah 60 – 80 mm/Hg DAN kadar glukosa 70 – 130 mgs/dl MAKA class = TIDAK GAGAL GINJAL KRONIK”

“ JIKA riwayat hipertensi = tidak DAN status sel darah merah = normal DAN status *pus cell* = normal DAN riwayat diabetes mellitus = tidak DAN status bakteri = tidak ada DAN riwayat penyakit jantung = tidak DAN riwayat anemia = tidak DAN tekanan darah 60 – 80 mm/Hg DAN kadar glukosa > 130 mgs/dl MAKA class = TIDAK GAGAL GINJAL KRONIK”

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id



Gambar 4.8 Pohon Keputusan Training Ketiga

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.10 Perhitungan Tiap Node dengan *Improved C4.5*

Atribut	Rasio gain Node 1	Rasio gain Node 2	Rasio gain Node 3	Rasio gain Node 4	Rasio gain Node 5	Rasio gain Node 6	Rasio gain Node 7	Rasio gain Node 8	Rasio gain Node 9
Umur (age)	0,00389	0,00131	0,00226	0,00151	0,00122	0,00145	0,00160	0,00090	0,00022
Tekanan Darah (bp)	0,03832	0,12209	0,29759	0,59333	0,62840	0,78550	0,80097	0,97024	0,00000
*Specfic Gravity (*Sg)	0,00011	0,00146	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
Albumin (al)	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
Gula Darah (su)	0,00013	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
Status sel darah merah (rbc)	0,35431	0,67186	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
Status *Pus cell* (pc)	0,28597	0,56532	1,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
Pus cell clumps (pcc)	0,19237	0,48225	0,71818	0,81162	0,00000	0,00000	0,00000	0,00000	0,00000
Status kehadiran Bakteri (ba)	0,17075	0,45043	1,00000	0,81162	0,85734	0,00000	0,00000	0,00000	0,00000
Kadar Glukosa dalam darah (bgr)	0,02263	0,03524	0,08581	0,06272	0,07467	0,07541	0,07541	0,07780	0,07780
Blood Urea Nitrogen (bu)	0,05382	0,01327	0,00413	0,00173	0,00052	0,00089	0,00027	0,00019	0,00019
Kadar serum kreatinin (sc)	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
Hasil *Sodium blood test* (sod)	0,01577	0,02998	0,03983	0,04591	0,04806	0,05177	0,05245	0,05375	0,05375
Hasil *Potassium test* (pot)	0,01719	0,25427	0,70673	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
Kadar Hemoglobin (hemo)	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
Packed cell volume (pcv)	0,02335	0,03506	0,04478	0,04667	0,05146	0,05151	0,05158	0,05170	0,05170
Jumlah sel darah putih (wc)	0,00454	0,01544	0,01995	0,01214	0,01507	0,01632	0,01773	0,01930	0,01930
Jumlah sel darah merah (rc)	0,01587	0,02954	0,03736	0,04101	0,04625	0,04621	0,04917	0,04889	0,04889
Riwayat Hipertensi (htn)	0,42719	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
Riwayat diabete mellitus (dm)	0,34602	0,49593	0,74316	0,84168	0,00000	0,00000	0,00000	0,00000	0,00000
Riwayat Penyakit Jantung Koroner (cad)	0,18904	0,43858	0,68301	0,81162	0,85734	0,87344	0,00000	0,00000	0,00000
Tingkat nafsu makan (appet)	0,23114	0,47562	0,71818	0,82645	0,85734	0,00000	0,89000	0,00000	0,00000
Riwayat *Pedal edema* (pe)	0,23114	0,48902	0,74316	0,81162	0,85734	0,87344	0,00000	0,00000	0,00000
Riwayat Anemia (ane)	0,20850	0,47562	0,71818	0,82645	0,85734	0,87344	0,96117	0,00000	0,00000

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

C. Pengujian Model Klasifikasi

Setelah mendapatkan pohon keputusan dan aturan klasifikasi selanjutnya tahap pengujian model yang dilakukan menggunakan matlab dengan tujuan untuk melihat nilai akurasi yang diperoleh. Pengujian model dilakukan dengan menggunakan data testing sebanyak 127 data untuk menguji ketepatan klasifikasi dari model yang telah dibentuk.

Pengujian dilakukan dengan mengunakan 3- *fold cross validation* untuk mengetahui tingkat akurasi yang didapatkan. Pada pengujian yang pertama atribut-atribut yang digunakan untuk melakukan pengujian tersebut adalah atribut riwayat hipertensi, status sel darah merah, tekanan darah, riwayat diabetes mellitus, status *pus cell*, dan *packed cell volume* sesuai dengan pohon keputusan dan aturan yang didapatkan *training* pertama.

Pada pengujian yang kedua atribut-atribut yang digunakan untuk melakukan pengujian tersebut adalah atribut status sel darah merah, riwayat hipertensi, status *pus cell*, riwayat diabetes mellitus, riwayat bakteri, tekanan darah, riwayat penyakit jantung dan kadar glukosa sesuai dengan pohon keputusan dan aturan yang didapatkan pada *training* yang kedua.

Pada pengujian yang ketiga atribut-atribut yang digunakan untuk melakukan pengujian tersebut adalah atribut riwayat hipertensi, status sel darah merah, status *pus cell*, riwayat diabetes mellitus, riwayat bakteri, riwayat penyakit jantung, riwayat anemia, tekanan darah, dan kadar glukosa sesuai dengan pohon keputusan dan aturan yang didapatkan pada *training* yang ketiga.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Berikut ini merupakan sampel tabel data *testing* pada pengujian pertama dan hasil setelah dilakukan perhitungan:

Tabel 4.11 Data testing dan sampel pengujian

No	Riwayat hipertensi	Status sel darah merah	Tekanan darah	Riwayat diabetes mellitus	Status *pus cell*	*Packed cell volume*	Class	Hasil uji
1	Ya	0	70	tidak	1	41	ckd	ckd
2	Ya	1	70	Ya	1	41	ckd	ckd
3	Ya	1	80	Ya	0	41	ckd	ckd
4	Tidak	1	70	tidak	0	40	ckd	notckd
5	Tidak	0	120	tidak	1	36	ckd	ckd
6	Ya	1	100	Ya	1	20	ckd	ckd
7	Ya	1	70	Ya	1	31	ckd	ckd
8	Tidak	1	60	tidak	1	42	notckd	notckd
9	Tidak	1	70	tidak	1	52	notckd	notckd
10	Tidak	1	70	tidak	1	43	notckd	notckd
11	Tidak	1	80	tidak	1	50	notckd	notckd

Setelah dilakukan pengujian data didapatkan pengujian kerja klasifikator memiliki nilai yang sama pada tiap data *testing* sebesar:

Nama	Pengujian pertama	Pengujian kedua	Pengujian ketiga
Akurasi	99,21%	99,21%	99,21%
Sensitifitas	98,70%	98,70%	98,70%
Spesifitas	100%	100%	100%

D. Analisis Hasil

Proses perhitungan untuk mendapatkan pohon keputusan dilakukan dengan melakukan perhitungan pada data training. Metode yang digunakan adalah klasifikasi dengan menggunakan *Improved C4.5.* Pengujian algoritma tersebut dilihat dari nilai akurasi, spesifikasi, dan sensitifitas.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Data yang digunakan dalam proses perhitungan sebanyak 381 data. Data tersebut dibagi menjadi 254 data untuk data training dan 127 data untuk data testing. Data training diolah menggunakan metode *Improved C4.5* sehingga didapatkan pohon keputusan. Setelah didapatkan pohon keputusan data testing digunakan untuk melakukan uji pada model yang didapatkan.

Pada perhitungan training yang pertama didapatkan pohon keputusan dengan atribut riwayat hipertensi, status sel darah merah, tekanan darah, riwayat diabetes mellitus, status *pus cell*, dan *Packed cell volume.* Pada pengujian pertama didapatkan pengujian kerja klasifikator dengan akurasi sebesar 99.21% , sensitifitas sebesar 98.70 % dan spesifikasi sebesar 100%.

Pada perhitungan training yang kedua didapatkan pohon keputusan dengan atribut status sel darah merah, riwayat hipertensi, status *pus cell*, riwayat diabetes mellitus, status kehadiran bakteri, tekanan darah, riwayat jantung koroner, dan kadar glukosa dalam darah. Pada pengujian kedua didapatkan pengujian kerja klasifikator dengan akurasi sebesar 99.21%, sensitifitas sebesar 98.70 % dan spesifikasi sebesar 100%.

Pada perhitungan training yang ketiga didapatkan pohon keputusan dengan atribut riwayat hipertensi, status sel darah merah, status *pus cell*, riwayat diabetes mellitus, status kehadiran bakteri, riwayat penyakit jantung koroner, riwayat anemia, tekanan darah, dan kadar glukosa dalam darah. Pada pengujian ketiga didapatkan pengujian kerja klasifikator dengan akurasi sebesar 99.21%, sensitifitas sebesar 98.70 % dan spesifikasi sebesar 100%.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Dari pengujian yang telah dilakukan dengan *Improved C4.5* dapat disimpulkan bahwa *Improved C4.5* memiliki tingkat akurasi yang sangat baik. Oleh karena itu untuk memprediksi penyakit gagal ginjal kronik dapat berpedoman pada *Improved C4.5*. Penentuan penyakit gagal ginjal kronik dapat berdasarkan pohon keputusan yang dapat dijadikan pedoman pada Gambar 4.10 pohon keputusan *Improved C4.5*.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BAB V

PENUTUP

A. Simpulan

Berdasarkan hasil perhitungan dan uji coba *Improved C4.5* yang telah dilakukan untuk prediksi penyakit gagal ginjal kronik, dapat disimpulkan bahwa:

1. Model prediksi yang didapatkan dari perhitungan nilai RasioGain dengan *Improved C4.5* merupakan aturan yang dibentuk dari pohon keputusan yang telah diperoleh. Pohon keputusan pada perhitungan pertama yang didapatkan seperti Gambar 4.6. Atribut-atribut pohon keputusan yang dibentuk adalah riwayat hipertensi, status sel darah merah, tekanan darah, riwayat diabetes mellitus, status *pus cell,* dan *Packed cell volume.* Atribut-atribut yang tidak terdapat dalam pohon keputusan tidak berpengaruh terhadap pengambilan keputusan. Pada pohon keputusan kedua atribut yang diperoleh adalah atribut status sel darah merah, riwayat hipertensi, status *pus cell,* riwayat diabetes mellitus, status kehadiran bakteri, tekanan darah, riwayat jantung koroner, dan kadar glukosa dalam darah. Pada pohon keputusan ketiga didapatkan pohon keputusan dengan atribut riwayat hipertensi, status sel darah merah, status *pus cell,* riwayat diabetes mellitus, status kehadiran bakteri, riwayat penyakit jantung koroner, riwayat anemia, tekanan darah, dan kadar glukosa dalam darah.
2. Berdasarkan pohon keputusan dan aturan yang yang didapatkan dapat digunakan untuk melakukan prediksi penyakit gagal ginjal kronik. Hasil

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

dari tiga kali pengujian menunjukkan nilai yang sama sebesar 99,21% , sensitifitas sebesar 98,70%, dan spesifikasi sebesar 100%.

B. Saran

Penelitian mengenai prediksi gagal ginjal kronik menggunakan *Improved C4.5* ini masih banyak kekurangan. Sehingga, peneliti menyarankan beberapa hal sebagai bahan perbaikan kedepannya, diantaranya adalah :

1. Membandingkan *Improved C4.5* dengan metode klasifikasi lain untuk mengetahui seberapa metode klasifikasi apa yang terbaik.
2. Data yang digunakan dalam penelitian ini masih banyak kekurangan yaitu banyak data yang masih kosong, sehingga harus dilakukan normalisasi dulu. Untuk kedepannya lebih baik menggunakan data yang lebih lengkap lagi.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

DAFTAR PUSTAKA

Amalia, H. (2018). Perbandingan Metode *Data Mining* SVM dan NN untuk Klasifikasi Penyakit Ginjal Kronis. *Jurnal Pilar Nusa Mandiri*, vol. 14 no. 1.

Amalia, H., & Evicienna, E. (2017). Komparasi metode *Data Mining* untuk Penentuan Proses Persalinan Ibu Melahirkan. *Jurnal Sistem Informasi*, 103-109.

Aradea, S., Ariyan, Z., & Yuliana , A. (2011). Penerapan Decission tree untuk Penentuan Pola Data Penerimaan Mahasiswa Baru.

Basuki, A., & Syarif, I. (2003). *Modul Ajar Decision Tree.* Surabaya: PENS-ITS.

Cao, R., & Xu, L. (2009). Improved C4. Algorithm for the Analysis of Sales.

Dunham, M. (2003). *Data Mining Introductory and advaced topics.* . Upper Saddle River, NJ: Pearson Education, Inc.

Hidayati, T., Kushadiwijaya, H., & Suhardi. (2008). Hubungan antara Hipertensi, Merokok, dan Minuman Suplemen Energi dan Kejadian Penyakit Ginjal Kronik. *Berita Kedokteran Masyarakat*.

Lucida, H., Trisnawati, R., & Suardi, M. (2011). Analisis Aspek Farmakokinetika Klinik Pasien Gagal Ginjal pada Irna Penyakit dalam RSUP Dr. M. Djamil Padang. *Jurnal Sains dan Tekhnologi Farmasi*, 16 (2), 144-155.

Mardi, Y. (2015). *Data Mining* : Klasifikasi menggunakan Algoritma C4.5. *2.i2* (213-219).

Melti, T., Arthur, M., & Manopp o , F. (2014). Hubungan antara Derajat Penyakit Ginjal Kronik dengan Nilai Agregasi Trombosit. *Jurnal E-Biomedik (eBM), 2*, 09-513.

Octhario, R., & Karnila, S. (2013). Sistem *Data Mining* Untuk mengetahui Tingkat Kecenderungan Memilih Menu Makanan Dengan Metode Association Rule Mining. *13*(1).

Pranandari, R., & Supadmi, W. (2015). Faktor Resiko Gagal Ginjal Kronik di Unit Hemodialisis RSUD Wates Kulon Progo. *Majalah Farmaseutik, 11*.

Prasetyo, E. (2014). *Data Mining Mengolah Data Menjadi Informasi Menggunakan Matlab*. Yogyakarta: CV. Andi Offset.

Rianto, R., & Iswari, N. M. (2017). Rancang Bangung Aplikasi Pendektesi Penyakit Ginjal Kronis dengan Menggunakan Algoritma C4.5. *Ultimatics*.

Rostanti, A., Bawotong, J., & Onibala, F. (2016). Faktor-Faktor yang Berhubungan dengan Kepatuhan Menjalani Terapi Hemodialisa pada Penyakit Ginjal Kronik di Ruangan Dahlia dan Melati RSUP Prof. Dr. R. D Kandau Manado. *Ejournal Keperawatan, 4 nomor 2*.

Santoso, R. B. (2016). Hubungan Lama Hemodialisis dengan Penurunan Nafsu Makan pada Pasien Gagal Ginjal Kronik di Unit Hemodialisa RSUD Ulin Banjarmasin. *Dinamika Kesehatan*, 139-151.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Silva, A., Cortez, P., Santos, F., Gomes, L., & Neves, J. (2008). Rating organ failure via adverse events using *data mining* in the intensive care unit. *43*(no. 4, pp. 2431-2448).

Smeltzer, S. C., & Bare, B. G. (2001). *Buku Ajar Keperawatan Medikal-Bedah Brunner & Suddarth.* Jakarta: EGC.

Sukandar, E. (2006). *Nefrologi Kliik.* Bandung: Universitas Padjajaran Press.

Sumilati, S., & Soleha, U. (2015). Hubungan Tingkat Pengetahuan Dengan Kepatuhan Diet pada Pasien Gagal Ginjal Kronik Yang dilakukan Hemodialisis Reguler di Rumah Sakit Darmo Surabaya . *Jurnal Ilmiah Kesehatan*, 131-136.

USRDS Annual Data Report. (2006). Retrieved from USRD (United States Renal Data System)

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id