

# Penerapan Data Mining Untuk Memprediksi Potensi Pendoror Darah Menjadi Pendoror Tetap Menggunakan Metode Decision Tree C.45

Ketut Jaya Atmaja<sup>1</sup>, Ida Bagus Gede Anandita<sup>2</sup>, Ni Kadek Ceryna Dewi<sup>3</sup>

<sup>1</sup> Ilmu Komputer (PascaSarjana, Universitas Pendidikan Ganesha)  
Singaraja, Bali, Indonesia  
*ketutjayaatmaja@stiki-indonesia.ac.id*

<sup>2</sup> Ilmu Komputer (PascaSarjana, Universitas Pendidikan Ganesha)  
Singaraja, Bali, Indonesia  
*ibganandita@stiki-indonesia.ac.id*

<sup>3</sup> Ilmu Komputer (PascaSarjana, Universitas Pendidikan Ganesha)  
Singaraja, Bali, Indonesia  
*dewichery@gmail.com*

## Abstract

Persediaan darah pada suatu rumah sakit terkadang tidak menentu, sehingga sangat dibutuhkan persediaan darah yang cukup pada bank darah, sehingga jika sewaktu – waktu dibutuhkan, darah sudah tersedia. Namun pada kenyataannya persediaan darah sering kali bersifat tidak tetap. Dengan *data mining* diharapkan mampu memprediksi data pendonor yang ada pada PMI untuk memprediksi pendonor yang berpotensi menjadi pendonor tetap. Dalam proses ini metode yang dipakai dalam melakukan data mining adalah algoritma C4.5. Dari hasil analisis yang dilakukan, dapat diketahui bahwa data yang digunakan adalah data random sebanyak 600 data, dimana data training berjumlah 500 data, dan data set berjumlah 100 data. Dari pohon *decision tree* yang didapatkan dari proses *data mining* menggunakan algoritma C4.5 yang dilakukan dapat disimpulkan bahwa pegawai swasta dengan umur diatas 26 tahun paling banyak menjadi pendonor.

Keywords : *data mining*, algoritma C4.5,

## 1. Pendahuluan

Donor darah adalah proses pengambilan darah dari seseorang secara sukarela atau pengganti untuk disimpan di bank darah sebagai stok darah untuk kemudian digunakan untuk transfusi darah. Penyumbangan darah biasa dilakukan rutin di Unit Donor Darah (UDD) PMI Pusat maupun Unit Donor Darah di daerah setiap beberapa waktu, namun ada pula penggalangan penyumbangan darah yang diadakan di pusat-pusat keramaian. Palang Merah Indonesia (PMI) adalah sebuah organisasi perhimpunan nasional di Indonesia yang bergerak dalam bidang sosial kemanusiaan. Palang Merah Indonesia dalam pelaksanaannya juga tidak melakukan pembedaan tetapi mengutamakan korban yang paling membutuhkan pertolongan segera untuk keselamatan jiwanya.

Persediaan darah sering kali bersifat tidak tetap, itu terjadi karena jumlah pendonor yang naik turun.

Sehingga akan menjadi sebuah masalah ketika jumlah permintaan darah meningkat sedangkan persediaan darah menipis. PMI Provinsi Bali selalu melakukan sosialisasi untuk menumbuhkan kesadaran masyarakat untuk sadar melakukan donor darah secara rutin dengan menyebarkan informasi secara menyeluruh ke semua elemen masyarakat dari berbagai usia, profesi dan wilayah. Metode tersebut dirasa kurang efektif karena masing-masing elemen masyarakat yang pernah melakukan donor darah memiliki karakteristik berbeda untuk menerima informasi tertentu.

Melalui data pendonor yang ada pada PMI dapat dilakukan proses mining untuk menentukan pola prediksi pendonor yang berpotensi menjadi pendonor tetap. Sehingga dapat difokuskan dimana penyebaran informasi harus dilakukan agar lebih efisien dalam menambah jumlah pendonor darah. Metode yang dipakai dalam melakukan data mining adalah metode klasifikasi menggunakan algoritma C4.5. Adapun

atribut data yang dipakai adalah jenis kelamin, pekerjaan, golongan darah, jumlah donor, dan umur. Banyak data yang dipakai adalah 600 data, dimana data training berjumlah 500 data, dan data set berjumlah 100 data. Klasifikasi dilakukan berdasarkan pekerjaan.

## 2. Landasan Teori

### 2.1. Data Mining

Data mining adalah serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basisdata. Informasi yang dihasilkan diperoleh dengan cara mengekstraksi dan mengenali pola yang penting atau menarik dari data yang terdapat dalam basisdata.

Datamining terutama digunakan untuk mencari pengetahuan yang terdapat dalam basis data yang besar sehingga sering disebut Knowledge Discovery in Databases (KDD). Proses pencarian pengetahuan ini menggunakan berbagai teknik-teknik pembelajaran komputer (machine learning) untuk menganalisis dan mengekstraksikannya. Proses pencarian bersifat iteratif dan interaktif untuk menemukan pola atau model yang sah, baru, bermanfaat, dan dimengerti. Dalam penerapannya datamining memerlukan berbagai perangkat lunak analisis data untuk menemukan pola dan relasi data agar dapat digunakan untuk membuat prediksi dengan akurat.

### 2.2. Clustering

Klasifikasi adalah proses menemukan model (fungsi) yang menjelaskan dan membedakan kelas-kelas atau konsep, dengan tujuan agar model yang diperoleh dapat digunakan untuk memprediksikan kelas atau objek yang memiliki label kelas tidak diketahui. Model yang turunan didasarkan pada analisis dari training data (yaitu objek data yang memiliki label kelas yang diketahui). Model yang diturunkan dapat direpresentasikan dalam berbagai bentuk seperti aturan IF-THEN klasifikasi, pohon keputusan, formula matematika atau jaringan syaraf tiruan. Dalam banyak kasus, pengguna ingin memprediksikan nilai-nilai data yang tidak tersedia atau hilang (bukan label dari kelas). Dalam kasus ini biasanya nilai data yang akan diprediksi merupakan data numeric. Kasus ini seringkali dirujuk sebagai prediksi. Di samping itu, prediksi lebih menekankan pada identifikasi trend dari distribusi berdasarkan pada data yang tersedia.

Pohon Keputusan (Decision Tree) merupakan metode klasifikasi dan prediksi yang sangat kuat dan

terkenal. Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan aturan. Aturan dapat dengan mudah dipahami dengan bahasa alami. Aturan ini juga dapat diekspresikan dalam bentuk bahasa basis data seperti SQL untuk mencari record pada kategori tertentu. Pohon keputusan juga berguna untuk mengeksplorasi data, menemukan hubungan tersembunyi antara sejumlah calon variabel input dengan sebuah variabel target. Karena pohon keputusan memadukan antara eksplorasi data dan pemodelan, pohon keputusan ini sangat bagus sebagai langkah awal dalam proses pemodelan bahkan ketika dijadikan sebagai model akhir dari beberapa teknik lain.

Salah satu algoritma induksi pohon keputusan yaitu ID3 (Iterative Dichotomiser 3). ID3 dikembangkan oleh J. Ross Quinlan. Dalam prosedur algoritma ID3, input berupa sampel training, label training dan atribut. Algoritma Decision Tree C4.5 merupakan pengembangan dari ID3. Sedangkan pada perangkat lunak open source WEKA mempunyai versi sendiri dari C4.5 yang dikenal sebagai J48.

### 2.3. WEKA

Weka (Waikato Environment for Knowledge Analysis) adalah aplikasi data mining opensource berbasis Java. Aplikasi ini dikembangkan pertama kali oleh Universitas Waikato di Selandia Baru. Weka terdiri dari koleksi algoritma machine learning yang dapat digunakan untuk melakukan generalisasi / formulasi dari sekumpulan data sampling. Algoritma ini bisa diterapkan secara langsung ke dalam dataset atau bisa juga dipanggil dari kode java kita sendiri. Weka memiliki tools untuk data pre-processing, classification, regression, clustering, association rules, dan visualization. Weka mengorganisasi kelas-kelas ke dalam paket-paket dan setiap kelas di paket dapat mereferensi kelas lain di paket lain. Paket classifiers berisi implementasi dari hampir semua algoritma untuk klasifikasi dan prediksi. Kelas yang paling penting di paket ini adalah Classifier, yang mendeklarasikan struktur umum dari skema klasifikasi dan prediksi. Kelas ini memiliki dua metoda, yaitu build Classifier dan classify Instance, yang harus diimplementasikan oleh kelas-kelas yang menginduk ke kelas ini. Semua kelas yang mengimplementasikan algoritma klasifikasi menginduk ke kelas Classifier, termasuk kelas J48. J48, yang menangani himpunan data dalam format ARFF, tidak mengandung kode untuk mengkonstruksi pohon keputusan. Kelas ini mereferensi kelas-kelas lain, kebanyakan di paket weka. Classifiers j48, yang mengerjakan semua proses konstruksi pohon. Weka adalah software open

## Ketut Jaya Atmaja, Penerapan Data Mining Untuk Memprediksi Potensi Pendonor Darah menjadi Pendonor Tetap menggunakan Metode Decision Tree C4.5

source yang diterbitkan dibawah lisensi GNU General Public License.

Proses pengklasifikasian dilakukan dengan menggunakan algoritma C.45 atau dikenal dengan J48 pada WEKA.

### 3. Analisis Dan Pembahasan

#### 3.1. Data Premier

Pada tabel berikut adalah merupakan data testing Pendonor PMI yang terdiri dari 100 data testing dan 500 data training yang akan di olah dan diklasifikasi di WEKA seperti terlihat pada Tabel 1.

Tabel 1. Tabel Data Pendonor

Nama	Jenis Kelamin	Pekerjaan	Gol.Darah	Jumlah Donor	Umur
K A Heryana Putra	Laki-Laki	Peg. Swasta	B	3	36
K Alit Adnyana	Laki-Laki	Peg. Swasta	O	2	45
K Audina Permana Putri	Wanita	Pelajar	A	1	50
K Ayu S Sucandrawati	Laki-Laki	Wiraswasta	O	1	52
K Darma	Laki-Laki	Peg. Swasta	B	6	66
K Dwi Indrawinata	Laki-Laki	Wiraswasta	O	1	25
K Dwi Wahyu	Laki-Laki	Peg. Negeri	O	5	29
K Mardika Saputra P	Laki-Laki	Polri	B	2	22
K Pancawan	Laki-Laki	Peg. Swasta	O	11	38
K Patah	Laki-Laki	Peg. Swasta	B	10	31
K Sehfrudin	Laki-Laki	Peg. Negeri	B	66	51
K Sudi R	Laki-Laki	Tni	O	7	43
K Sudiara	Laki-Laki	Wiraswasta	O	24	43
K Sudiarta	Laki-Laki	Peg. Swasta	B	5	45
K. Sudana Wibawa	Laki-Laki	Wiraswasta	B	4	50
Ka Morni	Laki-Laki	Peg. Swasta	O	1	36
Kad Sugiarta Jaya	Laki-Laki	Peg. Swasta	O	1	28
Kadafi	Laki-Laki	Wiraswasta	O	2	42
Kade Agus S N	Laki-Laki	Peg. Swasta	O	8	31
Kade Agus Sudha Naryana	Laki-Laki	Peg. Swasta	Ab	8	31
Kade Ciptadi Sujana	Laki-Laki	Pelajar	O	1	22
Kade Dewi Wahyuningsih	Laki-Laki	Peg. Swasta	O	4	37
Kade Dwi Anjastari	Wanita	Peg. Swasta	O	1	28
Kade Dwi Marlinda	Laki-Laki	Mahasiswa	O	2	22
Kade Marthen Pramana	Laki-Laki	Peg. Swasta	O	7	32
L Dian Mardines	Laki-Laki	Peg. Swasta	B	1	38
L Eky Fajaria Dian C	Laki-Laki	Peg. Swasta	B	2	22
L Gede Dewi Kusumasari	Wanita	Peg. Swasta	B	2	27
L Kd Surdianggriani,Sh	Wanita	Peg. Swasta	B	1	41
L P Ani Susanti	Wanita	Peg. Swasta	O	2	36
L Pt Suyeni	Wanita	Peg. Swasta	A	1	33

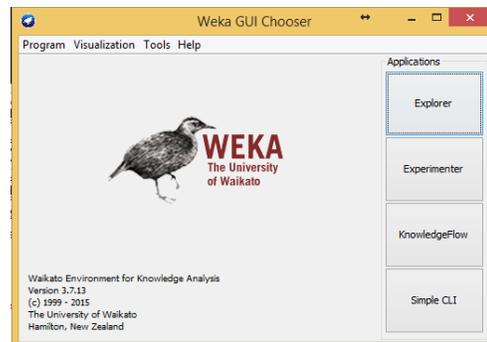
L Putu Dea Sasmita P	Wanita	Mahasiswa	B	1	21
L Putu Sri Sumartini	Laki-Laki	Peg. Swasta	O	1	51
L Putu Tresna Dewi	Laki-Laki	Peg. Swasta	O	2	22
L Rahadi Sulaksana W	Laki-Laki	Peg. Swasta	O	1	20
L Ria Dewi M	Laki-Laki	Peg. Swasta	O	2	28
L Siddharta	Laki-Laki	Wiraswasta	O	27	51
L Surya Kusuma	Laki-Laki	Tni	O	5	30
L Wisnu Aditya	Laki-Laki	Peg. Swasta	O	2	39
L Y Wawan Rukmana	Laki-Laki	Tni	O	3	48
Laasar Adu	Laki-Laki	Peg. Swasta	Ab	5	36
Lady Yugisworo	Laki-Laki	Peg. Swasta	O	1	34
Laela Nur Khoiyinah	Wanita	Wiraswasta	O	1	21
Laela Permatasari	Wanita	Peg. Swasta	A	12	28
Laila	Wanita	Peg. Swasta	B	4	87
Laila Maqiyah	Laki-Laki	Mahasiswa	O	1	24
Lailatul Qadriyah	Wanita	Peg. Swasta	O	8	22
Laini/Amei	Laki-Laki	Peg. Swasta	O	4	36
Laksana Eka S P	Laki-Laki	Peg. Swasta	O	23	32
Laksana Frizal Farli	Laki-Laki	Peg. Swasta	B	1	28
M D Budhi Darma	Laki-Laki	Peg. Swasta	B	5	38
M Darmawi	Laki-Laki	Peg. Swasta	A	1	29
M Dhofirul Usman	Laki-Laki	Wiraswasta	B	4	44
M Didik H	Laki-Laki	Peg. Swasta	Ab	2	32
M Dwipa Yudha	Laki-Laki	Peg. Swasta	O	3	34
M E Bayu Aji P	Laki-Laki	Peg. Swasta	O	3	43
M Edward Ndolu	Laki-Laki	Tni	B	3	38
M Eko Saputra	Laki-Laki	Peg. Swasta	A	1	27
M Eko Saputra	Laki-Laki	Peg. Swasta	A	11	27
M Fachlur Rozy	Laki-Laki	Mahasiswa	B	4	28
M Fahmi	Laki-Laki	Peg. Swasta	B	1	27
M Faisal Afandi	Laki-Laki	Peg. Swasta	O	5	22
M Fakhri Zaki	Laki-Laki	Mahasiswa	B	2	20
M Farchan	Laki-Laki	Peg. Swasta	Ab	2	23
M Fatikh	Laki-Laki	Peg. Swasta	Ab	1	27
M Fauzi	Laki-Laki	Wiraswasta	A	1	41
M Firman	Laki-Laki	Peg. Swasta	O	2	38
M Firsada Putra	Laki-Laki	Peg. Swasta	B	7	26
M Ghozali	Laki-Laki	Peg. Swasta	O	17	33
M Ginting	Laki-Laki	Peg. Swasta	B	2	55
M Hanafi	Laki-Laki	Peg. Swasta	B	57	50
M Haris Iskandar	Laki-Laki	Peg. Swasta	B	2	34

Ketut Jaya Atmaja, Penerapan Data Mining Untuk Memprediksi Potensi Pendonor Darah menjadi Pendonor Tetap menggunakan Metode Decision Tree C4.5

M Harun Ar Rosyid	Laki-Laki	Peg. Swasta	O	5	26
M Hidayat	Laki-Laki	Peg. Swasta	O	27	42
M Hidayat	Laki-Laki	Tni	O	2	48
Nanang Fathoni	Laki-Laki	Peg. Swasta	A	7	44
Nanang Hadiyanto	Laki-Laki	Peg. Swasta	O	2	54
Nanang Hartoyo	Laki-Laki	Peg. Swasta	A	8	27
Nanang Herimawan	Laki-Laki	Peg. Swasta	A	3	43
Nanang Hermawan	Laki-Laki	Peg. Swasta	A	8	43
Nanang Kasim	Laki-Laki	Peg. Swasta	A	1	35
Nanang Kristanto	Laki-Laki	Peg. Swasta	A	2	30
Nanang Nr	Laki-Laki	Peg. Swasta	A	12	51
Nanang Setiawan	Laki-Laki	Bumn	B	11	32
Nanang Setyadi	Laki-Laki	Peg. Swasta	O	15	33
Nanang Siswoko	Laki-Laki	Peg. Swasta	O	5	43
Nanang Supriyadi	Laki-Laki	Wiraswasta	B	1	28
Nanang Temo Winadi	Laki-Laki	Peg. Swasta	A	1	22
Nanang Tri Suseno	Laki-Laki	Peg. Swasta	Ab	2	47
Nanang Uhsin N	Laki-Laki	Tni	B	1	34
Nanang Winadi	Laki-Laki	Peg. Swasta	O	33	50
Nancik Jasmiali	Laki-Laki	Tni	A	6	52
Nancy Katarina	Laki-Laki	Peg. Swasta	O	3	30
Nanda Dipinto	Laki-Laki	Mahasiswa	A	4	23
Nanda Ibu	Laki-Laki	Peg. Swasta	A	1	42
Nanda Kh	Laki-Laki	Peg. Swasta	A	1	22
Nanda Khrimaya Sofyan	Laki-Laki	Mahasiswa	A	1	25
Nanda Noor Gustiriani D	Wanita	Pelajar	B	2	21
Nanda Putri Chintia	Wanita	Mahasiswa	B	5	20
Nanda Sendiawan	Laki-Laki	Wiraswasta	O	1	62

### 3.2. Learning Data Di WEKA

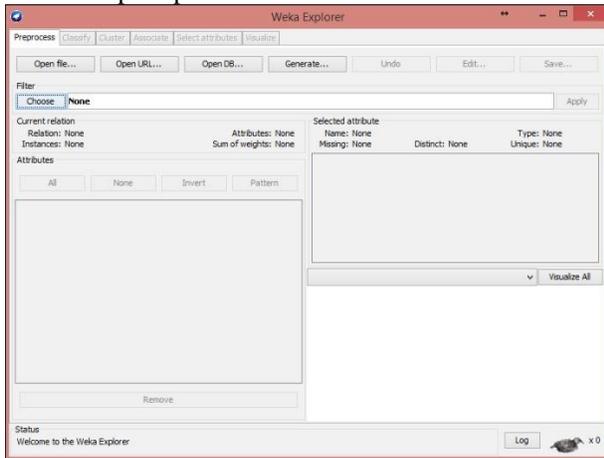
Pada Gambar 1, merupakan tampilan awal pada aplikasi WEKA yang terdapat menu bar Program, Visualization, Tools dan Help. Dan terdapat juga empat menu aplikasi yaitu Explorer, Experimenter, KnowledgeFlow, Simple CLI.



Gambar 1. Tampilan Awal WEKA

a. Langkah I

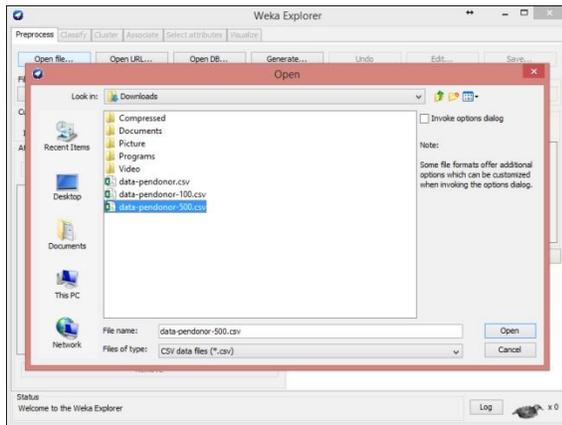
Pada tampilan awal WEKA pilih menu Applications Explorer sehingga muncul window baru seperti pada Gambar 2 berikut :



Gambar 2. Tampilan Antarmuka Application WEKA Explorer

b. Langkah II

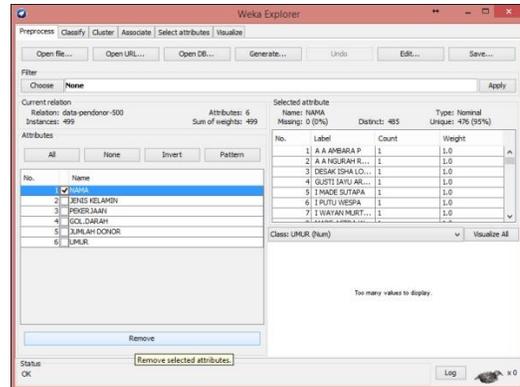
Selanjutnya klik button “Open file” dan cari tempat file data-pendonor-500.csv disimpan, yang akan digunakan sebagai data training, seperti pada Gambar 3 berikut :



Gambar 3. Tampilan Antarmuka Open File

c. Langkah III

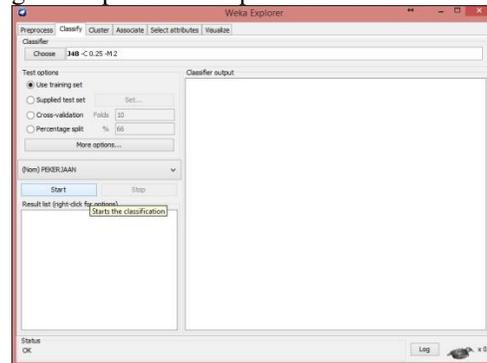
Selanjutnya akan tampil atribut – atribut yang ada di dalam tabel file data-pendonor-500.csv. Remove atau singkirkan atribut yang tidak digunakan dalam proses mining data, pada contoh disini atribut yang dihilangkan adalah atribut “nama” centang pada atribut yang akan dihilangkan dan klik button “remove” seperti pada Gambar 4 berikut



Gambar 4. Tampilan Antarmuka Pemilihan Atribut Data

d. Langkah IV

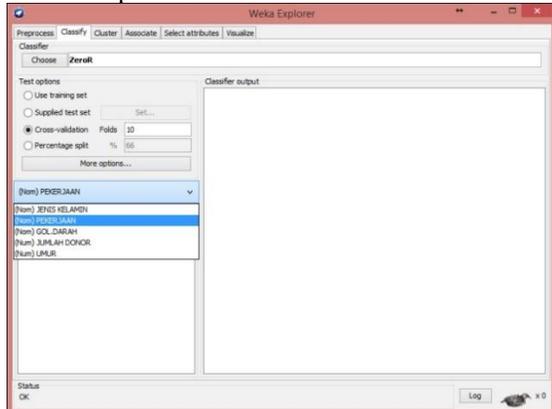
Langkah selanjutnya adalah memilih metode yang akan digunakan untuk mining data pendonor, contohnya disini menggunakan metode klasifikasi, maka di tab metode pilih tab “classify” sehingga tampilan window, pada panel Test options pilih “Use training set” seperti terlihat pada Gambar 5.



Gambar 5. Tampilan Antarmuka Pemilihan Metode dan Test Options

e. Langkah V

Selanjutnya pilih atribut yang digunakan untuk pengklasifikasian data yang ada di bawah panel Test Options. Pada contoh ini digunakan atribut Pekerjaan seperti terlihat pada Gambar 6 berikut

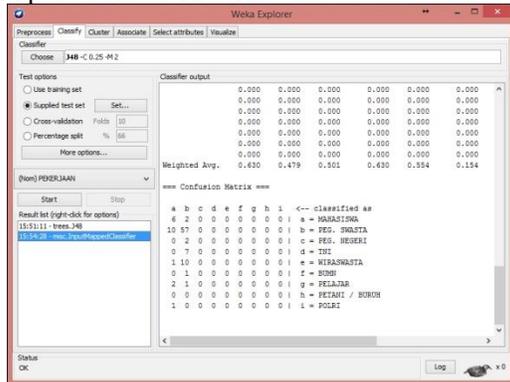


Gambar 6. Tampilan Antarmuka Pemilihan Atribut yang digunakan sebagai klasifikasi



k. Langkah XI

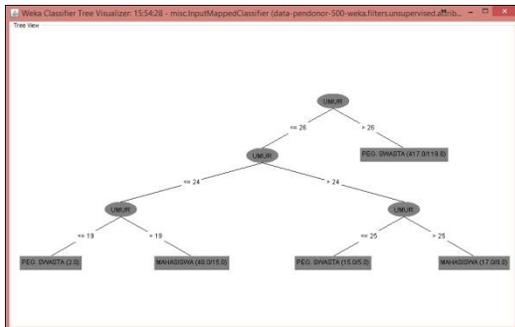
Setelah data-pondonor-100.csv di open, pilih atribut Pekerjaan yang juga digunakan pada saat pengolahan data training tadi dan klik button “Close” kemudian selanjutnya kembali klik button “Start” untuk melakukan pengolahan data testing sehingga hasilnya seperti pada Gambar 12 berikut ini.



Gambar 12. Tampilan Antarmuka Hasil Proses Data Test

l. Langkah XII

Perhatikan kembali bagian panel Result list akan muncul result atau hasil pengolahan baru bernama misc.InputMappedClassifier. Klik kanan pada result tersebut dan pilih “Visualize tree” akan tampil pohon keputusan kembali seperti pada Gambar 13 berikut ini.



Gambar 13. Hasil Pohon Keputusan Data Test

4. Kesimpulan

Dari hasil analisis yang dilakukan, dapat diketahui bahwa data yang digunakan adalah data random sebanyak 600 data, dimana data training berjumlah 500 data, dan data set berjumlah 100 data. Metode yang dipakai dalam melakukan data mining adalah metode klasifikasi menggunakan algoritma C4.5. Dari pohon decision tree yang didapatkan dari proses data mining yang dilakukan dapat disimpulkan bahwa pegawai swasta dengan umur diatas 26 tahun paling banyak menjadi pendonor.

References

- [1] Eko Prasetyo, Data Mining-Mengolah Data Menjadi Informasi Menggunakan Matlab, Yogyakarta, 2014.
- [2] Abidin, Taufik Fuadi, Pengantar WEKA, 2013
- [3] Teguh Budi Santoso, “Analisa dan Penerapan Metode C4.5 Untuk Prediksi Loyalitas Pelanggan”, Jurnal Ilmiah Fakultas Teknik LIMIT’S, Vol.10, No.1, ISSN:0216-1184.
- [4] Indri Rahmayuni, “Perbandingan Performansi Algoritma C4.5 dan CART Dalam Klasifikasi Data

**First Author** Ketut Jaya Atmaja lahir di Singaraja pada tanggal 26-11-1991. Telah menyelesaikan pendidikan S1 Manajemen Teknik Informatika di STMIK STIKOM Indonesia pada tahun 2015 dan saat ini sedang menyelesaikan studi S2 Ilmu Komputer di Universitas Pendidikan Ganesha

**Second Author** Ida Bagus Gede Anandita lahir di Dawan pada tanggal 05-03-1992. Telah menyelesaikan pendidikan S1 Manajemen Teknik Informatika di STMIK STIKOM Indonesia pada tahun 2015 dan saat ini sedang menyelesaikan studi S2 Ilmu Komputer di Universitas Pendidikan Ganesha

**Third Author** Ni Kadek Ceryna Dewi lahir di Klungkung 4 Juli 1990. Telah menyelesaikan pendidikan S1 Teknik Informatika di Sekolah Tinggi Teknologi Adisutjipto Yogyakarta Pada Tahun 2013 dan saat ini sedang menyelesaikan studi S2 Ilmu Komputer di Universitas Pendidikan Ganesha.