

SPK Pemilihan Calon Pendoror Darah Potensial dengan Algoritma C4.5 dan Fuzzy Tahani

Mahmud Yunus, Harry Soekotjo Dahlan, dan Purnomo Budi Santoso

Abstract— Penelitian ini bertujuan untuk membangun sebuah aplikasi Sistem Penunjang Keputusan (SPK) yang dapat menghasilkan daftar calon pendonor darah potensial berdasarkan hasil penambangan data pada *database* transaksi pelayanan donor darah dengan menggunakan algoritma *data mining* C4.5 dan *database* Fuzzy Tahani. Algoritma *data mining* C4.5 digunakan untuk melakukan eksplorasi data guna menemukan hubungan tersembunyi antara sejumlah variabel input dengan variabel target dalam bentuk pohon keputusan (*decision tree*). Variabel input meliputi atribut (1) golongan darah; (2) jenis kelamin; (3) usia; (4) berat badan; (5) tekanan darah; (6) kadar hemoglobin (HB); dan (7) kadar hematocrit (HCT), sedangkan variabel output/targetnya adalah atribut Donor Lagi yang berisikan nilai ‘Ya’ atau ‘Tidak’. Metode *database fuzzy* model Tahani digunakan untuk melakukan perangkangan data hasil penambangan dengan melakukan *query fuzzy* yang melibatkan variabel fuzzy (1) usia; (2) tekanan darah; (3) berat badan dan (4) kedekatan tempat tinggal.

Data pembelajaran untuk membentuk aturan pohon keputusan adalah data transaksi donor darah yang terjadi pada tahun 2010 hingga 2011. Pengukuran keakuratan prediksi terhadap data calon pendonor darah potensial yang dihasilkan, dilakukan dengan membandingkannya dengan data transaksi donor darah yang terjadi pada periode tahun 2012. Akurasi hasil prediksinya adalah (1) 90,56% untuk semua golongan darah; (2) 88,64% untuk golongan darah A; (3) 93,33% untuk golongan darah B; (4) 84,62% untuk golongan darah AB dan (5) akurasi prediksi sebesar 91,03% untuk golongan darah O. Rata-rata tingkat akurasi prediksi pemilihan calon pendonor darah potensial dengan algoritma *data mining* C4.5 adalah sebesar 89,64%.

Kata kunci: pohon keputusan, *data mining*, algoritma C4.5, dan *database fuzzy* Tahani.

I. PENDAHULUAN

SALAH satu masalah utama yang dihadapi dalam sistem pelayanan kesehatan di negara berkembang, adalah kurangnya persediaan darah aman yang memadai untuk keperluan transfusi darah bagi yang

mebutuhkannya. Aman dan memadainya persediaan darah dan produknya, bergantung pada dukungan dan komitmen pemerintah terhadap pengembangan pelayanan transfusi darah nasional, melalui organisasi yang dikelola pemerintah atau organisasi nirlaba yang ditunjuk seperti Palang Merah Indonesia (PMI) atau organisasi donor darah sukarela [11].

Persediaan darah aman dapat berubah sewaktu-waktu sementara jumlah permintaan darah dari pasien terus meningkat. Unit Transfusi Darah Cabang (UTDC) PMI Kota Malang seringkali mengalami kondisi kritis, yaitu minimnya persediaan darah aman yang dibutuhkan masyarakat. Metode yang digunakan pada kondisi kritis tersebut adalah menghubungi beberapa orang secara acak atau tidak terpolo untuk melakukan donor darah demi terpenuhinya bahan baku darah untuk diproses menjadi darah aman yang siap didistribusikan ke masyarakat yang membutuhkannya. Metode tersebut dirasa kurang efektif karena data calon pendonor yang terkumpul seringkali tidak memenuhi persyaratan donor darah.

UTDC PMI Kota Malang telah melakukan pelayanan transfusi darah lebih dari 5 tahun dan telah terbentuk *database* transaksi pelayanan transfusi darah dalam jumlah yang besar. *Database* dalam jumlah yang besar dapat dijadikan bahan analisa untuk menemukan pengetahuan, informasi, pola-pola atau kecenderungan tertentu yang sangat penting dan biasanya tidak disadari, yang dikenal dengan istilah *Data Mining* (penambangan data). *Data mining* merupakan analisis dari peninjauan kumpulan data untuk menemukan hubungan yang tidak diduga dan meringkas data dengan cara yang berbeda dengan sebelumnya, yang dapat dipahami dan bermanfaat bagi pemilik data [9]. Penambangan data pada *database* transaksi pelayanan transfusi darah dengan menggunakan algoritma tertentu, dapat dijadikan metode pemilihan calon pendonor darah potensial, yaitu orang yang memiliki kemampuan yang memungkinkan (memenuhi syarat) untuk mendonorkan darahnya secara sukarela dengan alasan kemanusiaan.

Pemilihan calon pendonor darah potensial dapat dikategorikan sebagai tindakan pengambilan keputusan. Salah satu metode pengambilan keputusan yang sistematis adalah dengan menyusun sebuah pohon keputusan (*decision tree*). Pohon keputusan merupakan sebuah struktur yang dapat digunakan untuk membagi kumpulan data yang besar menjadi himpunan-himpunan *record* yang lebih kecil dengan menerapkan serangkaian aturan keputusan, dengan masing-masing rangkaian

Mahmud Yunus adalah Dosen Teknik Informatika, STMIK PPKIA Pradnya Paramita Malang-Indonesia dan mahasiswa Program Magister Teknik Elektro, Fakultas Teknik, Universitas Brawijaya (UB) Malang-Indonesia; email: myoenoes@gmail.com.

Harry Soekotjo Dachlan adalah Dosen Teknik Elektro, Fakultas Teknik Universitas Brawijaya (UB) Malang-Indonesia; email: harysd@ub.ac.id.

Purnomo Budi Santoso adalah Dosen Teknik Industri, Fakultas Teknik Universitas Brawijaya (UB) Malang-Indonesia; email: budiakademika@yahoo.com.

pembagian, anggota himpunan hasil menjadi mirip satu dengan yang lain [1].

Pohon keputusan merupakan salah satu metode klasifikasi dengan mengubah data menjadi pohon keputusan dan aturan-aturan keputusan (*rule*), yang dapat digunakan untuk memprediksi atau mengklarifikasi suatu kejadian. Algoritma yang dapat dipakai dalam pembentukan pohon keputusan antara lain ID3, CART, CHAID, C4.5 dan C5.0. Pohon keputusan dengan algoritma *data mining* C4.5 dapat digunakan sebagai metode untuk memprediksi kemungkinan pengunduran diri calon mahasiswa baru di STMIK AMIKOM Yogyakarta [8]. Algoritma C5.0 memberikan rata-rata tingkat keakuratan sebesar 87,72%, CART 87,27% dan CHAID 87,15% dalam memprediksi status resiko kredit Bank X dengan menggunakan tiga algoritma *decision tree* [13].

Penelitian mengenai penerapan algoritma klasifikasi C4.5 untuk penentuan kelayakan pemberian kredit koperasi, menyimpulkan bahwa algoritma C4.5 mampu menganalisa kredit yang bermasalah dan yang tidak bermasalah sebesar 90% [5]. Penelitian mengenai prediksi status keaktifan studi mahasiswa dengan algoritma C5.0 dan *K-Nearest Neighbor* menyimpulkan bahwa keberhasilan klasifikasi yang diperoleh dari algoritma C5.0 mencapai lebih dari 90% [4]. Bujlow, Riaz dan Pedersen dalam penelitiannya yang menggunakan metode Algoritma C5.0 telah mampu mengklasifikasi dan membedakan 7 aplikasi berbeda (Skype, FTP, Torrent, Web, Web Radio, Game dan SSH) diantara 76,632 - 1,622,710 kasus yang terjadi dalam lalu lintas jaringan Internet dengan akurasi rata-rata 99,3% hingga 99,9% [2].

Penerapan algoritma *data mining* C4.5 pada Sistem Penunjang Keputusan dapat menjadi alternatif dalam menghasilkan informasi calon pendonor darah potensial yang dibutuhkan oleh UTDC PMI Kota Malang. Algoritma *data mining* C4.5 diperlukan untuk menghasilkan aturan-aturan dalam bentuk pohon keputusan dengan mengolah (1) Atribut calon pendonor seperti; (a) kesesuaian golongan darah dengan kebutuhan; (b) jenis kelamin; (c) usia; (d) berat badan; (e) tekanan darah; (f) kadar Hemoglobin (HB); dan (e) kadar Hematocrit (HCT); (2) Nilai Atribut; (3) *Entropy*, serta (4) *Gain*. Daftar aturan yang terbentuk dalam pohon keputusan digunakan untuk mencari atau menemukan calon pendonor darah yang potensial.

Daftar calon pendonor darah potensial yang dihasilkan dari proses penambangan data dengan menggunakan algoritma *data mining* C4.5, selanjutnya perlu dilakukan perangkingan sebagai usaha penentuan skala prioritas. Metode *database* Fuzzy model Tahani dapat digunakan untuk melakukan perangkingan data tersebut dengan melakukan *query* fuzzy yang melibatkan beberapa variabel fuzzy berikut (1) usia; (2) tekanan darah; (3) berat badan dan (4) kedekatan tempat tinggal.

Penelitian ini dilakukan untuk membangun sebuah Sistem Penunjang Keputusan (SPK) yang dapat melakukan pemilihan calon pendonor darah potensial

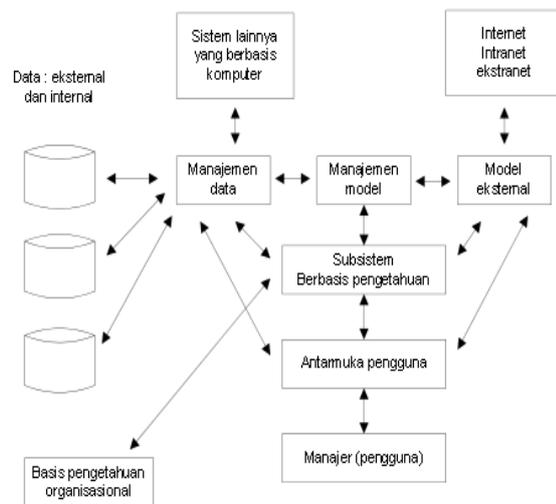
dengan menggunakan algoritma *data mining* C4.5 dan *query database* fuzzy model Tahani. Luaran (*output*) dari SPK yang dibangun, berupa daftar calon pendonor darah potensial yang dapat digunakan oleh UTDC PMI Kota Malang, sebagai bahan pertimbangan dalam menghubungi atau meminta masyarakat untuk menjadi pendonor darah sukarela utamanya di saat kondisi kritis.

II. LANDASAN TEORI

A. Sistem Penunjang Keputusan

Sistem Penunjang Keputusan (SPK) atau *Decision Support System* (DSS), merupakan sebuah sistem yang digunakan sebagai alat bantu bagi para pengambil keputusan untuk memperluas kemampuan manusia dengan tidak menggantikan perannya. SPK dikembangkan dengan tujuan untuk mendukung pengambil keputusan dengan memilih alternatif keputusan yang merupakan hasil pengolahan informasi-informasi yang diperoleh atau tersedia dengan menggunakan model-model pengambil keputusan serta untuk menyelesaikan masalah-masalah bersifat terstruktur, semi terstruktur dan tidak terstruktur.

SPK ditujukan untuk keputusan-keputusan yang memerlukan penilaian atau pada keputusan-keputusan yang sama sekali tidak dapat didukung oleh algoritma [12]. SPK dapat didefinisikan sebagai sistem berbasis komputer yang terdiri dari 3 (tiga) komponen yang saling berinteraksi : sistem bahasa (mekanisme untuk memberikan komunikasi antara pengguna dan komponen SPK lain), sistem pengetahuan (repositori pengetahuan domain masalah yang ada pada SPK baik sebagai data atau sebagai prosedur), dan sistem pemrosesan masalah (hubungan antara dua komponen lainnya, terdiri dari satu atau lebih kapabilitas manipulasi masalah umum yang diperlukan untuk pengambilan keputusan). Konsep-konsep yang diberikan oleh definisi tersebut sangat penting untuk memahami hubungan antara SPK dan pengetahuan.



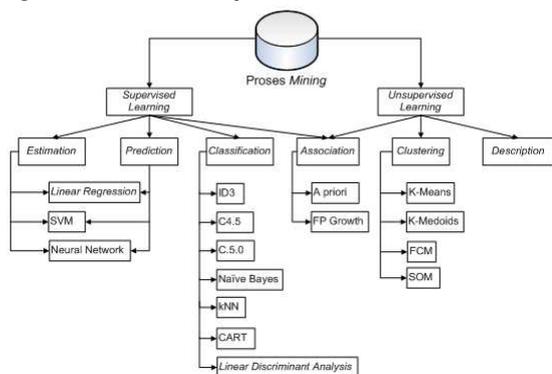
Gambar 1. Skematik Sistem Penunjang Keputusan [12]

Sistem Penunjang Keputusan (SPK) mencakup 4 komponen utama (subsistem), yaitu manajemen data, manajemen model, antar muka pengguna dan sub sistem manajemen berbasis pengetahuan [12].

1. Subsistem manajemen data, yang merupakan suatu *database* yang berisi data yang relevan untuk situasi dan dikelola oleh perangkat lunak yang disebut *Database Management System (DBMS)*.
2. Subsistem manajemen model, yang merupakan paket perangkat lunak yang memasukkan model keuangan, statistik, ilmu manajemen, atau model kuantitatif lainnya yang memberikan kapabilitas analitik dan manajemen perangkat yang tepat.
3. Subsistem antarmuka pengguna. Pengguna berkomunikasi dengan dan memerintahkan SPK melalui subsistem ini. Pengguna adalah bagian yang dipertimbangkan dari sistem.
4. Subsistem manajemen pengetahuan. Subsistem ini dapat mendukung semua subsistem lain atau bertindak sebagai suatu komponen independen. Ia memberikan kecerdasan untuk memperbesar pengetahuan pengambil keputusan

B. Penambangan Data (Data Mining)

Data mining merupakan analisis dari peninjauan kumpulan data untuk menemukan hubungan yang tidak diduga dan meringkas data dengan cara yang berbeda dengan sebelumnya, yang dapat dipahami dan bermanfaat bagi pemilik data [9]. *Data mining* merupakan teknologi yang digunakan dalam disiplin ilmu yang berbeda-beda untuk mencari hubungan yang signifikan antara variabel dalam set data yang besar [3]. Metode *data mining* dapat digunakan untuk melakukan klasifikasi data dengan mengubah data menjadi pohon keputusan (*decision tree*) dan aturan-aturannya (*rule*), yang dapat digunakan untuk memprediksi atau mengklarifikasi suatu kejadian.



Gambar 2. Pengelompokan Teknik *Data Mining*

Terdapat dua karakteristik metode pelatihan dalam *data mining*. Secara garis besar metode pelatihan yang digunakan dalam teknik-teknik *data mining* dibedakan ke dalam dua pendekatan[10], yaitu:

1. *Unsupervised learning*, metode ini diterapkan tanpa adanya latihan (*training*) dan tanpa ada guru (*teacher*). Guru di sini adalah label dari data.
2. *Supervised learning*, yaitu metode belajar dengan adanya latihan dan pelatih. Dalam pendekatan ini, untuk menemukan fungsi keputusan, fungsi pemisah atau fungsi regresi, digunakan beberapa contoh data yang mempunyai output atau label selama proses training.

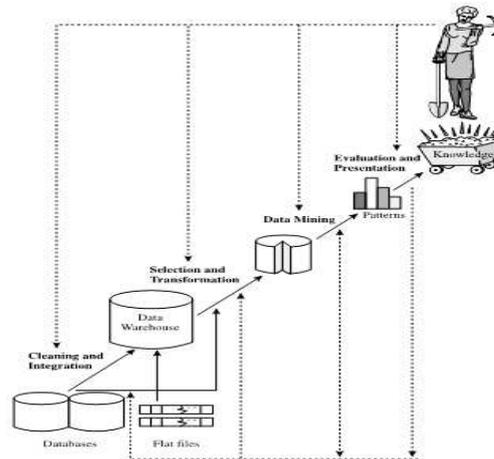
Berdasarkan tugas yang bisa dilakukan, *data mining*

beberapa teknik yaitu: deskripsi, estimasi, prediksi, klasifikasi, *clustering*, dan asosiasi [10]. Beberapa metode yang dapat digunakan berdasarkan pengelompokan *data mining* seperti pada Gambar 2. *Data mining* sebagai suatu rangkaian proses dapat dibagi menjadi beberapa tahapan yang bersifat interaktif, pemakai terlibat langsung atau dengan perantaraan *knowledge base*. Tahap-tahap *data mining* secara urut terlihat pada Gambar 3.

C. Algoritma C4.5

Algoritma yang dapat digunakan untuk memprediksi atau mengklarifikasi suatu kejadian dengan pembentukan pohon keputusan antara lain C4.5, yang merupakan salah satu algoritma induksi pohon keputusan yang dikembangkan oleh J. Ross Quinlan. Secara umum alur proses algoritma C4.5 untuk membangun pohon keputusan dalam *data mining* adalah [7]:

1. Pilih atribut sebagai simpul akar.
2. Buat cabang untuk tiap-tiap nilai.
3. Bagi kasus dalam cabang
4. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.



Gambar 3. Tahapan Dalam *Data Mining* [6]

Pemilihan atribut sebagai simpul, baik simpul akar (*root*) atau simpul internal didasarkan pada nilai *Gain* tertinggi dari atribut-atribut yang ada. Penghitungan nilai *Gain* digunakan rumus seperti dalam Persamaan 1.

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

- S : Himpunan kasus
- A : Atribut
- n : Jumlah partisi atribut A
- |S_i| : Jumlah kasus pada partisi ke-i
- |S| : Jumlah kasus dalam S

untuk menghitung nilai *Entropy* dapat dilihat pada Persamaan 2.

$$Entropy(S) = \sum_{i=1}^n - pi * \log_2 pi \dots (2)$$

- n : Jumlah partisi S
- pi : Proporsi dari S_i terhadap S

D. Database Fuzzy Tahani

Database fuzzy Tahani merupakan salah satu metode fuzzy yang menggunakan sistem basis data

sebagai pangkalan datanya. Jika suatu sistem diharapkan dapat menghasilkan informasi yang bersifat linguistik atau *ambiguous*, maka kita dapat memanfaatkan model basis data (*database*) fuzzy. Model *database* fuzzy merupakan sistem *database* standar (klasik) yang dikembangkan dengan memasukkan data hasil fuzzyfikasi (*fuzziness*) terhadap data tegas (*crisp*) yang ada dalam *database*. Informasi dari *database* fuzzy dapat diperoleh dengan menggunakan bahasa *query* standar yang dikenal dengan istilah SQL (*Structure Query Language*).

Metode fuzzy *database* model Tahani menggunakan teori himpunan fuzzy untuk mendapatkan informasi query-nya. *Database* fuzzy model Tahani tersusun atas tahapan, yaitu (1) tahap menggambarkan fungsi keanggotaan dalam kurva yang menunjukkan pemetaan titik-titik input data ke dalam nilai (derajat) keanggotaannya yang memiliki interval nilai antara 0 sampai 1. Beberapa fungsi keanggotaan yang dapat digunakan adalah (a) representasi kurva Linier; (b) representasi kurva Segitiga dan (c) representasi kurva Trapesium; (2) tahap fuzzyfikasi yaitu perubahan nilai tegas ke nilai fuzzy; (3) tahap *fuzzyfikasi query* yaitu sebuah *query* konvensional (*non-fuzzy*) DBMS yang menerapkan sistem dasar logika fuzzy *query* (*fuzzy logic based querting system*).

III. METODE PENELITIAN

A. Data Penelitian

Data penelitian yang digunakan adalah data identitas pendonor darah, transaksi donor darah pada tahun 2010 hingga tahun 2012, dan semua data yang berhubungan dengan pelayanan transfusi darah. Data primer tersebut selanjutnya dilakukan proses normalisasi data untuk mendapatkan data skunder berupa tabel data kasus transaksi donor darah yang siap untuk dilakukan proses penambangan data. Selanjutnya data skunder pada penelitian dibagi menjadi dua kelompok data;

1. Data *training* dan data *testing*. Menggunakan data transaksi donor darah yang terjadi pada tahun 2010 hingga tahun 2011. Data pembelajaran digunakan untuk pembentukan pohon keputusan berdasarkan algoritma penambangan data C4.5. Data pengujian digunakan untuk memperoleh daftar calon pendonor darah yang dianggap potensial berdasarkan sekumpulan aturan yang ada dalam struktur pohon keputusan yang telah terbentuk.
2. Data pembandingan hasil prediksi SPK. Menggunakan data transaksi donor darah yang terjadi pada tahun 2012. Data pembandingan ini digunakan untuk mengukur ketepatan prediksi calon pendonor darah potensial yang dihasilkan oleh sistem. Pengukuran ketepatan prediksi dilakukan dengan cara membandingkan daftar calon pendonor darah potensial yang dihasilkan oleh sistem dengan fakta yang terjadi pada data transaksi donor darah tahun 2012. Jika calon pendonor potensial yang dihasilkan oleh sistem memang melakukan donor darah pada tahun 2012, maka hasil prediksi sistem tepat, jika sebaliknya maka hasil prediksi sistem tidak tepat.

B. Variabel Penelitian

Variabel penelitian meliputi variabel *input* dan variabel *output* atau target. Variabel input dinyatakan oleh beberapa atribut (1) berat badan; (2) jenis kelamin; (3) usia; (4) tekanan darah; (5) kadar hemoglobin (HB); dan (6) kadar hematocrit (HCT). Penyederhanaan nilai atribut usia, berat badan, tekanan darah, kadar HB dan kadar HCT, perlu dilakukan karena jangkauan nilainya yang lebar, sehingga dapat mengakibatkan proses pengenalan pola data dan pembentukan pohon keputusan menjadi lama. Penyederhanaan atau pengelompokan nilai atribut terlihat pada Tabel 1, 2, 3, 4 dan Tabel 5.

Variabel targetnya adalah atribut 'Donor Lagi' yang berisikan nilai 'Ya' atau 'Tidak'. Nilai atribut 'Donor Lagi' diperoleh dari proses penentuan nilai ambang batas (*threshold*) pada atribut 'Jumlah Donor'. Misalkan nilai ambang batas yang ditetapkan adalah 2, maka nilai atribut 'Donor Lagi' adalah 'Ya', jika nilai 'Jumlah Donor' ≥ 2 , sebaliknya jika nilai 'Jumlah Donor' < 2 , maka nilainya adalah 'Tidak'

TABEL I
KELOMPOK USIA

Kelompok	Rentang Usia
0	< 17 Tahun
1	17 s/d 29 Tahun
2	30 s/d 44 Tahun
3	45 s/d 59 Tahun
4	≥ 60 Tahun

TABEL II
KELOMPOK BERAT BADAN

Kelompok	Rentang Berat Badan
0	< 47 Kg.
1	47 s/d 69 Kg.
2	70 s/d 89 Kg.
3	≥ 90 Kg.

TABEL III
KELOMPOK TEKANAN DARAH

Kelompok	Rentang Tekanan Darah
0	< 80
1	80 s/d 109
2	110 s/d 139
3	≥ 140

TABEL IV
KELOMPOK KADAR HB

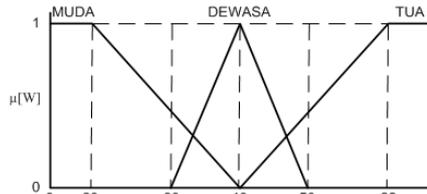
Kelompok	Rentang HB
0	< 9
1	9 s/d 13
2	14 s/d 17
3	≥ 18

TABEL V
KELOMPOK KADAR HCT

Kelompok	Rentang HCT
0	< 26%
1	26% s/d 32%
2	33% s/d 39%
3	$\geq 40\%$

Proses perangkan data calon pendonor darah potensial terpilih, dilakukan dengan menggunakan metode *query database* fuzzy model Tahani. *Query database* ini menggunakan beberapa variabel fuzzy sebagai berikut;

1. Variabel Usia yang dibagi dalam himpunan keanggotaan Muda, Dewasa dan Tua.



Gambar 4. Grafik Fungsi Keanggotaan Usia

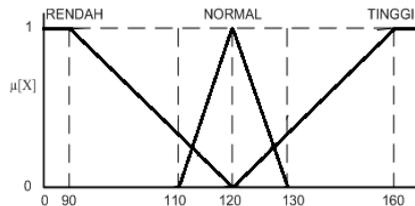
Fungsi linear untuk masing-masing himpunan Usia adalah sebagai berikut;

$$\mu_{\text{MUDA}}[W] = \begin{cases} 1; & W \leq 20 \\ \frac{40 - W}{40 - 20}; & 20 \leq W \leq 40 \\ 0; & W \geq 40 \end{cases}$$

$$\mu_{\text{DEWASA}}[W] = \begin{cases} 0; & W \leq 30 \\ \text{atau} \\ \frac{W - 30}{40 - 30}; & 30 \leq W \leq 40 \\ \frac{50 - W}{50 - 40}; & 40 \leq W \leq 50 \\ 0; & W \geq 50 \end{cases}$$

$$\mu_{\text{TUA}}[W] = \begin{cases} 0; & W \leq 40 \\ \frac{W - 40}{60 - 40}; & 40 \leq W \leq 60 \\ 1; & W \geq 60 \end{cases}$$

2. Variabel Tekanan Darah yang dibagi dalam himpunan keanggotaan Rendah, Normal dan Tinggi.



Gambar 5. Grafik Fungsi Keanggotaan Tekanan Darah

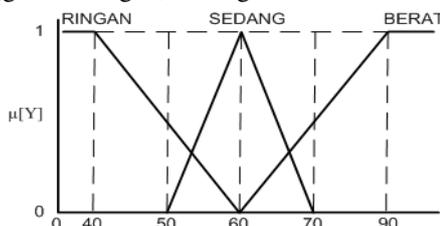
Fungsi linear untuk masing-masing himpunan Tekanan Darah adalah sebagai berikut;

$$\mu_{\text{RENDAH}}[X] = \begin{cases} 1; & X \leq 90 \\ \frac{120 - X}{120 - 90}; & 90 \leq X \leq 120 \\ 0; & X \geq 120 \end{cases}$$

$$\mu_{\text{NORMAL}}[X] = \begin{cases} 0; & X \leq 110 \\ \text{atau} \\ \frac{X - 110}{120 - 110}; & 110 \leq X \leq 120 \\ \frac{130 - X}{130 - 120}; & 120 \leq X \leq 130 \\ 0; & X \geq 130 \end{cases}$$

$$\mu_{\text{TINGGI}}[X] = \begin{cases} 0; & X \leq 120 \\ \frac{X - 120}{160 - 120}; & 120 \leq X \leq 160 \\ 1; & X \geq 160 \end{cases}$$

3. Variabel Berat Badan yang dibagi dalam himpunan keanggotaan Ringan, Sedang dan Berat.



Gambar 6. Grafik Fungsi Keanggotaan Berat Badan

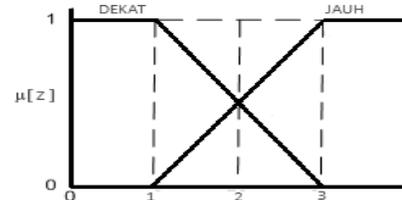
Fungsi linear untuk masing-masing himpunan Berat Badan adalah sebagai berikut;

$$\mu_{\text{RINGAN}}[Y] = \begin{cases} 1; & Y \leq 40 \\ \frac{60 - Y}{60 - 40}; & 40 \leq Y \leq 60 \\ 0; & Y \geq 60 \end{cases}$$

$$\mu_{\text{SEDANG}}[Y] = \begin{cases} 0; & Y \leq 50 \\ \text{atau} \\ \frac{Y - 50}{60 - 50}; & 50 \leq Y \leq 60 \\ \frac{70 - Y}{70 - 60}; & 60 \leq Y \leq 70 \\ 0; & Y \geq 70 \end{cases}$$

$$\mu_{\text{BERAT}}[Y] = \begin{cases} 0; & Y \leq 60 \\ \frac{Y - 60}{90 - 60}; & 60 \leq Y \leq 90 \\ 1; & Y \geq 90 \end{cases}$$

4. Kedekatan Tempat Tinggal (TT) yang dibagi dalam himpunan keanggotaan Dekat, dan Jauh.



Gambar 7. Grafik Fungsi Keanggotaan Kedekatan TT

Fungsi linear untuk masing-masing himpunan Tempat Tinggal adalah sebagai berikut;

$$\mu_{\text{DEKAT}}[Z] = \begin{cases} 1; & Z \leq 1 \\ \frac{3 - Z}{3 - 1}; & 1 \leq Z \leq 3 \\ 0; & Z \geq 3 \end{cases}$$

$$\mu_{\text{JAUH}}[Z] = \begin{cases} 0; & Z \leq 1 \\ \frac{Z - 1}{3 - 1}; & 1 \leq Z \leq 3 \\ 1; & Z \geq 3 \end{cases}$$

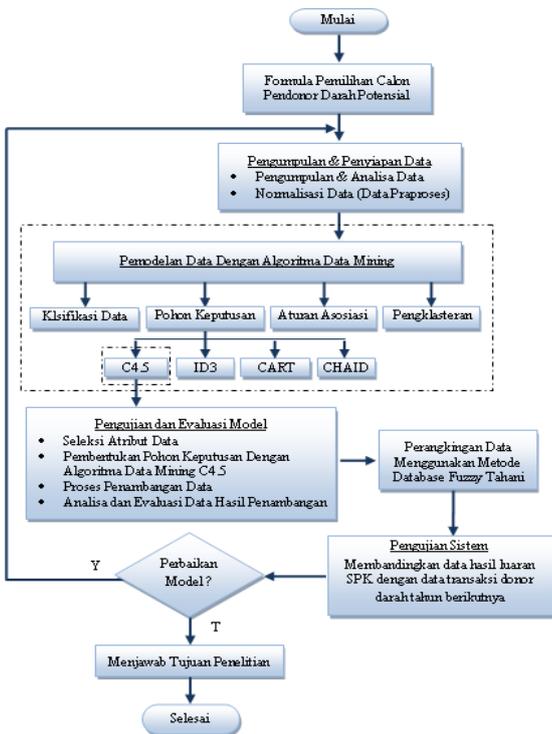
Berdasarkan keempat fungsi keanggotaan himpunan fuzzy tersebut maka proses perangkungan daftar calon pendonor darah potensial, memiliki 54 kombinasi aturan fuzzy dengan menggunakan operator dasar fuzzy maximum (AND) dan minimum (OR).

C. Konsep Solusi Masalah

Konsep solusi permasalahan pada penelitian ini sebagaimana terlihat pada Gambar 8, dapat dijelaskan langkah-langkah pemecahan masalah sebagai berikut:

1. Menentukan formulasi yang tepat untuk pemilihan calon pendonor darah potensial dengan melihat latar belakang permasalahan dan ketersediaan data
2. Pengumpulan dan evaluasi terhadap data transaksi donor darah serta melakukan proses normalisasi data yang telah terkumpul
3. Pemodelan SPK pemilihan calon pendonor darah potensial dengan menggunakan algoritma data mining C4.5 dengan tahapan (a) pemilihan atribut data; (b) pembentukan pohon keputusan dengan algoritma data mining C4.5; (c) melakukan proses penambangan data berdasarkan pohon keputusan yang telah terbentuk dan (d) melakukan analisa dan evaluasi terhadap data hasil penambangan
4. Perangkungan data hasil penambangan, dengan menggunakan metode *query database* fuzzy model Tahani yang melibatkan beberapa variabel fuzzy

5. Pengujian sistem apakah sudah dapat menjawab permasalahan penelitian, dengan membandingkan data hasil luaran SPK yang terbentuk dengan data transaksi donor darah pada tahun berikutnya

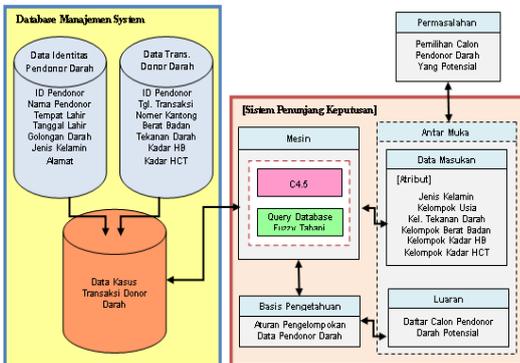


Gambar 8. Konsep Solusi Permasalahan

D. Desain Sistem

Desain sistem yang dibangun adalah sebagai berikut;

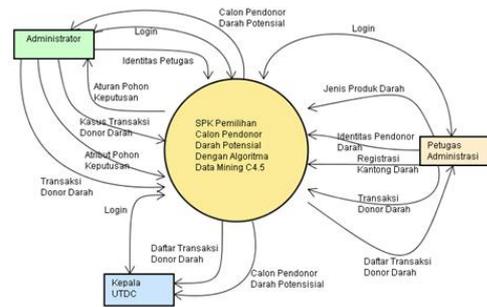
1. Arsitektur SPK. Pada Gambar 9 membagi SPK ke dalam empat lingkungan, yaitu manajemen basis data, basis pengetahuan, mesin, dan antar muka.



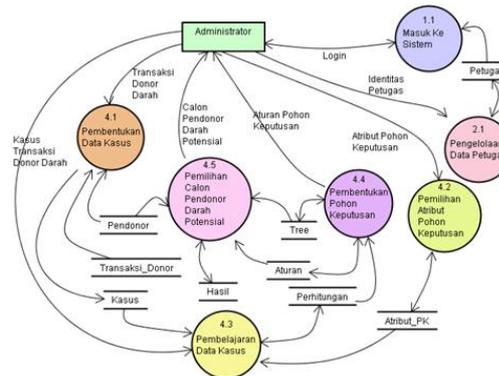
Gambar 9. Arsitektur Sistem Penunjang Keputusan

2. Pemodelan Sistem Dengan Diagram Alir Data. Diagram Alir Data pada SPK pemilihan calon pendoror darah potensial ini melibatkan 3 entitas yaitu: Administrator, Petugas Administrasi dan Kepala UTDC, seperti terlihat pada Gambar 10, 11 dan 12.

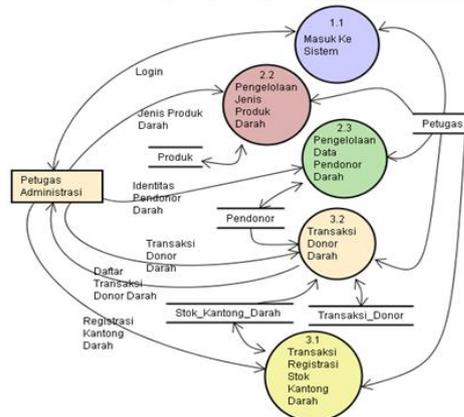
3. Pemodelan Database Dengan ERD. Entity Relationship Diagram (ERD) merupakan salah satu cara untuk pemodelan database relasional yang dibuat berdasarkan kumpulan objek-objek dasar yang dinamakan entitas dan digunakan dalam sebuah sistem, serta hubungan antara entitas. ERD dari SPK yang dibangun terlihat pada Gambar 13



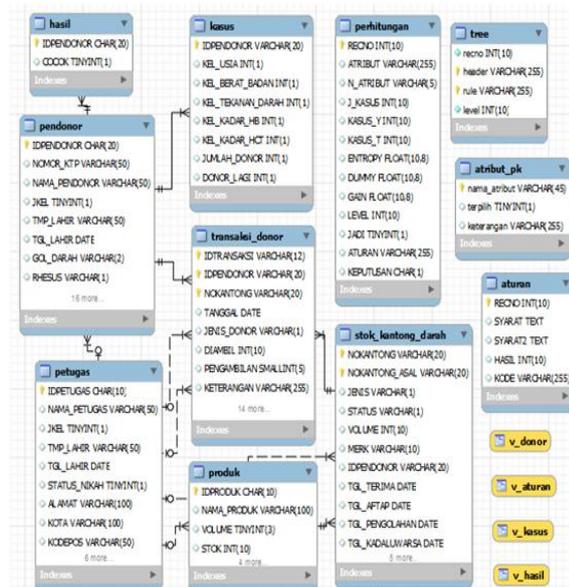
Gambar 10. Context Diagram SPK Pemilihan Calon Pendoror Darah Potensial



Gambar 11. DAD Level 1 Entitas Administrator

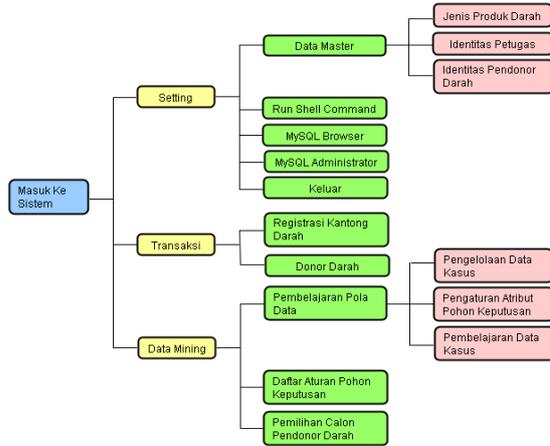


Gambar 12. DAD Level 1 Entitas Petugas Administrasi



Gambar 13. ERD SPK Pemilihan Calon Pendoror Darah Potensial

4. Desain Menu Aplikasi. Seperangkat pilihan dalam aplikasi untuk menunjukkan kemampuan dan fasilitas yang dimiliki oleh sebuah program aplikasi kepada pengguna. Desain menu SPK yang dibangun seperti pada Gambar 14.

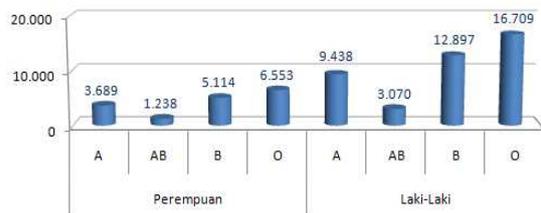


Gambar 14. Desain Menu Aplikasi SPK Pemilihan Calon Pendoron Darah Potensial

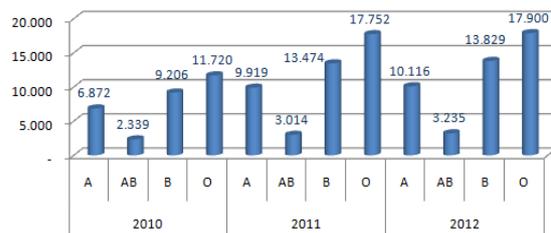
IV. HASIL DAN PEMBAHASAN

Pada tahap pengujian sistem, data yang digunakan sudah melalui proses normalisasi data dan ditransformasikan dalam bentuk kelompok data atribut yang lebih sederhana. Data yang digunakan dalam pengujian adalah data Pendoron Darah dan data transaksi donor darah yang terjadi selama tahun 2010 s/d tahun 2012, dengan jumlah data sebanyak 119.376 record. Data transaksi donor darah dibagi menjadi 2 bagian yaitu (1) data transaksi tahun 2010 s/d 2011 sebagai data pembelajaran dan pengujian sistem, serta (2) dan data transaksi tahun 2012 sebagai data pembandingan. Pembentukan data kasus transaksi donor darah dalam penelitian ini menggunakan nilai Ambang Batas ≥ 2 , yang menghasilkan 40.324 record.

Hasil pembelajaran data Kasus dengan menggunakan atribut (1) Berat Badan; (2) Jenis Kelamin; (3) Kadar Hemoglobin (HB); (4) Kadar Hematocrit (HCT); (5) Tekanan Darah dan (6) Usia, menghasilkan 1.431 record data perhitungan Entropy dan Gain dengan rata-rata waktu pembelajaran 3 menit 2,57 detik



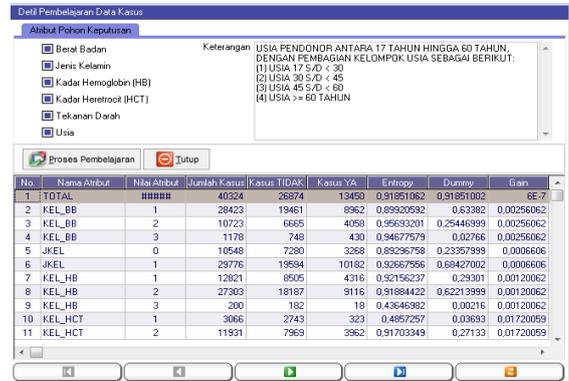
Gambar 15. Grafik Jumlah Pendoron Darah Per-Jenis Kelamin



Gambar 16. Grafik Jumlah Transaksi Donor Darah Per-Tahun

TABEL VI
JUMLAH DATA KASUS TRANSAKSI DONOR DARAH

Jenis Kelamin	Asumsi Donor Lagi	Jumlah Data
Perempuan	Ya	3.268
	Tidak	7.280
Laki-Laki	Ya	10.182
	Tidak	19.594
Total		40.324

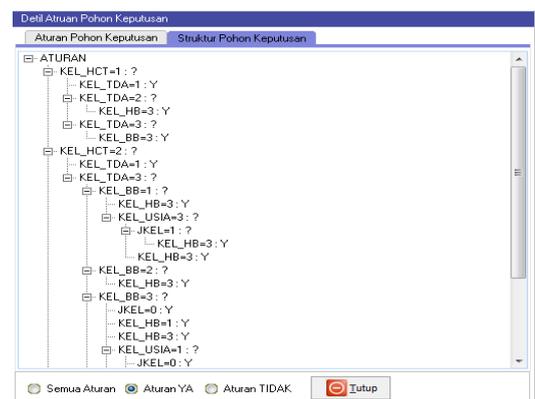


Gambar 17. Pembelajaran Data Kasus Transaksi Donor Darah

TABEL VII
ATURAN KEPUTUSAN 'YA' PADA POHON KEPUTUSAN YANG TERBENTUK

No.	Aturan Keputusan
1	If Kel_Hct=1 And Kel_TDA=1 Then Ya
2	If Kel_Hct=2 And Kel_Tda=1 Then Ya
3	If Kel_Hct=1 And Kel_Tda=2 And Kel_Hb=3 Then Ya
4	If Kel_Hct=1 And Kel_Tda=3 And Kel_Bb=3 Then Ya
5	If Kel_Hct=2 And Kel_Tda=3 And Kel_Hb=3 Then Ya
-	-
-	-
-	-
24	If Kel_Hct=3 And Kel_Usia=3 And Kel_Tda=3 And Kel_Bb=3 And Jkel=0 And Kel_Hb=2 Then Ya
25	If Kel_Hct=3 And Kel_Usia=4 And Kel_Tda=2 And Kel_Bb=3 And Jkel=0 And Kel_Hb=1 Then Ya
26	If Kel_Hct=3 And Kel_Usia=4 And Kel_Tda=3 And Jkel=1 And Kel_Hb=1 And Kel_Bb=3 Then Ya

Keterangan :
 Kel_Bb : Kel. Berat Badan Jkel : Jenis Kelamin
 Kel_Hb : Kel. Kadar Hemoglobin Kel_Hct : Kel. Hematocrit
 Kel_Tda : Kel. Tekanan Darah Kel_Usia : Kel. Usia



Gambar 18. Aturan Keputusan 'Ya' Pada Pohon Keputusan

Berdasarkan hasil pembelajaran data Kasus juga dihasilkan sebanyak 134 aturan keputusan dengan 26 aturan keputusan 'Ya' dan 108 aturan 'Tidak'. Pemilihan calon pendoron darah potensial dilakukan pada data Kasus transaksi donor darah dengan menggunakan syarat (a) aturan-aturan keputusan 'Ya'; (b) selang waktu lebih dari 60 hari setelah donor darah terakhir dan (c) calon pendoron tidak dicekal.

Berdasarkan hasil pengujian, diperoleh 180 orang calon pendonor darah potensial dengan rincian pada Tabel 8.

Gambar 19. Hasil Pemilihan Calon Pendonor Darah Potensial

Pengurutan untuk menentukan skala prioritas calon pendonor darah potensial yang harus dihubungi, dilakukan dengan menggunakan *query database* fuzzy model Tahani. Himpunan keanggotaan fuzzy yang digunakan sesuai dengan Gambar 4, 5, 6 dan Gambar 7.

Pengukuran keakuratan prediksi terhadap 185 orang calon pendonor darah potensial terpilih, dilakukan dengan cara membandingkannya dengan data transaksi donor darah pada tahun 2012, apakah data pendonor potensial muncul pada data transaksi di tahun 2012 atau tidak. Hasil pengujianya terlihat pada Tabel 8.

TABEL VIII
AKURASI HASIL PREDIKSI PEMILIHAN CALON PENDONOR DARAH POTENSIAL

Golongan Darah	Jumlah Data		Keakuratan Prediksi
	Hasil Prediksi	Cocok/ Sesuai	
Semua	180	163	90,56%
A	44	39	88,64%
B	45	42	93,33%
AB	13	11	84,62%
O	78	71	91,03%
Rata-Rata Keakuratan Hasil Prediksi			89,64%

V. KESIMPULAN

Berdasarkan hasil pengujian dan pembahasan dapat disimpulkan sebagai berikut;

1. Proses pembentukan pohon keputusan dengan algoritma *data mining* C4.5 dilakukan pada data kasus transaksi donor darah yang terjadi pada tahun 2010 s/d 2011, menghasilkan 134 aturan keputusan dengan rincian (1) 26 aturan keputusan ‘Ya’ dan (2) 108 aturan ‘Tidak’, dengan rata-rata waktu pembelajaran 3 menit 2,57 detik.
2. Pemilihan calon pendonor darah potensial dilakukan pada data kasus transaksi donor dengan menggunakan syarat aturan keputusan ‘Ya’, yang menghasilkan 180 orang calon pendonor darah potensial dengan rincian; (1) 44 orang golongan darah A; (2) 45 orang golongan darah B; (3) 13 orang golongan darah AB dan (4) 78 orang calon pendonor untuk golongan darah O.
3. Pengukuran keakuratan prediksi terhadap data calon pendonor darah potensial yang dihasilkan, dilakukan dengan cara membandingkannya dengan data transaksi donor darah yang terjadi pada tahun 2012.

4. Akurasi hasil prediksinya adalah (a) 90,56% untuk semua golongan darah; (b) 88,64% untuk golongan darah A; (c) 93,33% untuk golongan darah B; (d) 84,62% untuk golongan darah AB dan (e) 91,03% untuk golongan darah O. Rata-rata tingkat akurasi prediksi adalah sebesar 89,64%.

VI. SARAN

Penelitian ini masih dapat dikembangkan dengan memperhatikan beberapa rekomendasi berikut;

1. Penyederhanaan kelompok data pada atribut kelompok Usia, Kadar HB, Kadar HCT, Berat Badan dan Tekanan Darah dapat dikembangkan dengan jangkauan yang lebih lebar sehingga akan terbentuk aturan keputusan ‘Ya’ yang lebih banyak dan data calon pendonor darah potensial yang dihasilkan juga akan lebih banyak.
2. Mengintegrasikan aplikasi SPK ini dengan sistem informasi inventori stok darah yang ada di UTDC Kota Malang.
3. SPK Pemilihan Calon Pendonor Darah Potensial ini masih menampilkan struktur pohon keputusan secara horizontal, akan lebih baik jika struktur pohon keputusan ditampilkan secara vertikal.

REFERENSI

- [1] Berry, Michael J.A. dan Linoff, Gordon S. 2004. *Data Mining Techniques for Marketing, Sales, Customer Relationship Management*. Second Edition. Wiley Publishing Inc.
- [2] Bujlow Tomasz, Riaz Tahir, Pedersen Jens Myrup. 2011. *A method for classification of network traffic based on C5.0 Machine Learning Algorithm*. Section for Networking and Security, Department of Electronic Systems Aalborg University, DK-9220 Aalborg East, Denmark.
- [3] Erdogan, S.Z. dan M. Timor. 2005. A Data Mining Application In A Student Database. *Journal Of Aeronautics And Space Technologies*. Volume 2 Number 2: 53-57.
- [4] Ernawati Iin. 2008. Tesis: *Prediksi Status Keaktifan Studi Mahasiswa Dengan Algoritma C5.0 Dan K-Nearest Neighbor*. IPB. Bogor.
- [5] Firmansyah. 2011. Tesis: *Penerapan Algoritma Klasifikasi C4.5 Untuk Penentuan Kelayakan Pemberian Kredit Koperasi*. STMIK NUSA MANDIRI. Jakarta.
- [6] Han, J. dan M. Kamber. 2006. *Data Mining: Concepts and Techniques, Second Edition*. Morgan Kaufmann Publishers. San Francisco.
- [7] Kusriani dan E.M. Luthfi. 2009. *Algoritma Data Mining*. Penerbit ANDI. Yogyakarta.
- [8] Kusriani., Hartati, Sri., Wardoyo, Retantyo dan Harjoko, Agus. 2009. *Perbandingan Metode Nearest Neighbor Dan Algoritma C4.5 Untuk Menganalisa Kemungkinan Pengunduran Diri Calon Mahasiswa Di STMIK AMIKOM Yogyakarta*. Jurnal DASI Vol. 10 No.1 Maret 2009. ISSN: 1411-3201. Yogyakarta.
- [9] Larose, Daniel T. 2005. *Discovering Knowledge in Data: An Introduction to Data Mining*. John Wiley & Sons, Inc.
- [10] Santosa, B. 2007. *Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Graha Ilmu. Yogyakarta.
- [11] Soedarmono, Yuyun SM., Dr. MSc., Susanti, Dr., dan Kartabrata, Dr. 2005. *Pedoman Pelayanan Transfusi Darah: Perhitungan Biaya Unit Transfusi Darah*. UTDP PMI, Jakarta.
- [12] Turban, E., J.E. Aronson dan T.P. Liang. 2005. *Decision Support System and Intelligent Systems - 7th ed*. Pearson Education, Inc. Dwi Prabantini (penterjemah). 2005. *Sistem Pendukung Keputusan dan Sistem Cerdas*. Penerbit ANDI. Yogyakarta.
- [13] Yusuf, Yogi W. 2007. *Perbandingan Performansi Algoritma Decision Tree C5.0, CART dan CHAID: Kasus Prediksi Status Resiko Kredit di Bank X*. Seminar Nasional Aplikasi Teknologi Informasi (SNATI) Juni 2007. ISSN: 1907-5022. Yogyakarta.