

# Rancang Bangun Sistem Pakar untuk Deteksi Dini Katarak Menggunakan Algoritma C4.5

Ivana Herliana W. Jayawardanu, Seng Hansun  
 Program Studi Teknik Informatika, Universitas Multimedia Nusantara, Tangerang, Indonesia  
 theresia.ivana08@gmail.com, hansun@umn.ac.id

Diterima 4 November 2015  
 Disetujui 30 Desember 2015

**Abstract**-In 2010, 51% of 39 million blindness are caused by cataract. In 2013, there are 1.8% of 1.027.763 Indonesian people who suffered from cataract. Half of them are not treated yet due to their ignorance on the cataract disease. Therefore, in this research, we tried to build a system that can detect early cataract disease as the ophthalmologist would do. The system will use C4.5 algorithm that receives 150 training data set as an input, resulting in a set of rules which can be used as decision factors. To test the system, k-fold cross validation technique is been used with k equals to 10. From the analysis result, the accuracy of the system is 93.2% to detect cataract disease and 80.5% to detect the type of cataract disease one might suffered.

**Index terms**-C4.5 algorithm, cataract, k-fold cross validation, machine learning

## I. PENDAHULUAN

Banyak faktor yang dapat menyebabkan kebutaan mata. Menurut *World Health Organization* [1, 2] penyebab 51% dari 39 juta kebutaan yang terjadi pada tahun 2010 di seluruh dunia adalah penyakit mata katarak. Berdasarkan Riset Kesehatan Dasar yang dilakukan pada 1.027.763 orang, tingkat kebutaan di Indonesia pada tahun 2013 mencapai 0,4% [3]. Dalam Riskesdas juga disebutkan dari populasi tersebut, penderita katarak di Indonesia pada tahun 2013 sebesar 1,8%. Separuh dari penderita katarak tersebut belum menjalani pengobatan katarak dikarenakan faktor ketidaktahuan penderita mengenai penyakit mata yang dideritanya [3].

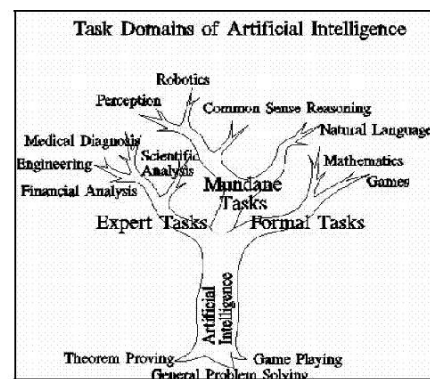
Hal yang dilakukan untuk meminimalisir ketidaktahuan tersebut adalah dengan membangun sebuah sistem yang dapat mendeteksi katarak secara dini tanpa harus mengunjungi klinik

mata. Sistem ini dibangun berbasiskan *website*, dengan mengimplementasikan metode *learning decision tree* dengan algoritma C4.5. Algoritma C4.5 dipilih karena dinyatakan sebagai algoritma *learning* yang memiliki tingkat keakuratan paling tinggi pada penelitian yang berjudul “Analisis Komparasi Algoritma Klasifikasi Data Mining untuk Prediksi Mahasiswa Non Aktif” [4].

## II. LANDASAN TEORI

### A. Artificial Intelligence

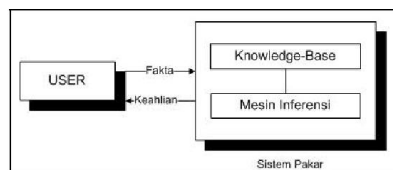
Istilah AI pertama kali dikemukakan pada tahun 1956 di konferensi Darthmouth, oleh John McCarthy [5]. Terdapat empat (4) teknik pemecahan masalah yang dapat dikembangkan secara terpisah atau dipadukan secara bersama, yaitu: *searching*, *reasoning*, *planning*, dan *learning*. Penelitian AI berkembang menjadi tiga (3) *domain* seperti pada Gambar 1.



Gambar 1. *Tasks Domains of Artificial Intelligence*[6]

## B. Sistem Pakar

Sistem pakar merupakan cabang dari AI yang memiliki *knowledge base*. Sistem pakar dapat bekerja sesuai dengan pengetahuan pakar yang dimasukkan ke dalamnya [7]. Pengetahuan tersebut diproses menggunakan algoritma yang ada untuk menghasilkan *rules*. *Rules* tersebut dijadikan dasar dalam menentukan hasil akhir dari setiap data yang diuji.



Gambar 2. Konsep dasar sistem pakar [8]

Konsep dasar sistem pakar dapat digambarkan seperti pada Gambar 2, sehingga dapat dijelaskan sebagai berikut:

### 1. User

Terdapat tiga (3) komponen *user* yang terlibat di dalam *environment system* [9], yaitu:

- Pakar, yaitu seorang ahli yang dipercayai pendapatnya karena memiliki pengetahuan khusus dan spesifik.
- Perekayasa sistem (*Knowledge Engineer*), yaitu seorang *engineer* yang menganalisis, merancang, dan membangun sistem.
- Pengguna (*User*), yaitu orang awam yang menggunakan sistem.

### 2. Knowledge Base

*Knowledge base* berisi pengetahuan yang dapat digunakan dalam mendapatkan hasil dari sistem. Terdapat dua (2) pendekatan yang dapat digunakan dalam membangun sistem, yaitu: *rule-based reasoning* dan *case-based reasoning*.

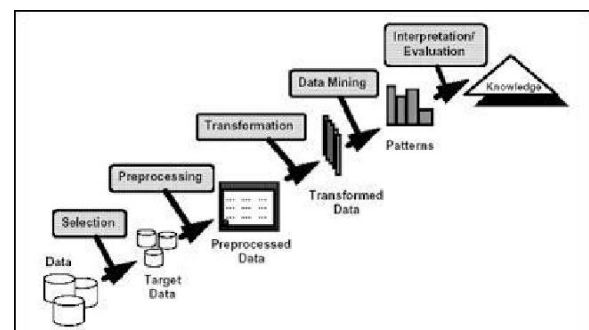
### 3. Mesin Inferensi

Mesin inferensi mengandung mekanisme pola pikir dan penalaran yang digunakan

oleh pakar dalam menyelesaikan suatu permasalahan.

## C. Data Mining

*Data mining* adalah proses mempekerjakan satu atau lebih teknik *machine learning* untuk menganalisis dan mengekstraksi *knowledge* secara otomatis [10]. Penggunaan *data mining* memiliki tujuan untuk mengetahui pola *universal* dari data-data yang ada. Dalam menghasilkan suatu *knowledge* dari pola yang ada, diperlukan penerapan metode *scientific* yang disebut dengan *Knowledge Discovery in Database* (KDD).



Gambar 3. Tahapan proses KDD [10]

Dorian Pyle pada bukunya *Data Preparation for Data Mining*, mengestimasi bahwa tahapan persiapan data pada KDD membutuhkan 60% dari keseluruhan waktu pemrosesan data [11]. Tahap pada proses KDD adalah sebagai berikut [10]:

#### 1. Selection

Pertama, lakukan seleksi data dari *multiple data source* dengan membersihkan data yang memiliki *noise* ataupun *missing value*.

#### 2. Preprocessing

Kedua, lakukan penggabungan terhadap seluruh *data source* yang telah terkumpul sebelum tahapan dari proses selanjutnya dilanjutkan.

#### 3. Transformation

Ketiga, lakukan transformasi data ke dalam bentuk yang lebih sesuai untuk dilakukan

*data mining*.

#### 4. Data Mining

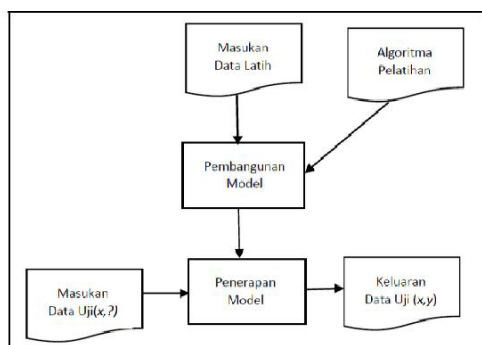
Keempat, terapkan metode *data mining* pada sistem, sehingga dapat menghasilkan pola dari data yang terkumpul.

#### 5. Interpretation/Evaluation

Kelima, lakukan interpretasi dan evaluasi dari pola yang didapatkan, sehingga dapat diidentifikasi apakah pola tersebut sudah dapat mewakili *knowledge* yang ingin dicapai.

#### D. Klasifikasi

Klasifikasi adalah salah satu teknik *data mining* yang mengelompokkan data menjadi berbagai kelas data dengan tujuan memprediksi kelas untuk data yang tidak diketahui kelasnya. Dalam klasifikasi terdapat dua (2) tahapan yang harus dilalui. Tahap pertama adalah induksi, yaitu: pembangunan model sebagai *prototype* untuk disimpan. Tahap kedua adalah deduksi, yaitu: penggunaan model yang dihasilkan untuk melakukan klasifikasi, pengenalan, dan prediksi pada suatu objek data lain agar diketahui termasuk kelas manakah objek data tersebut tergolong di dalam memori [12].



Gambar 4. Tahapan *Data Mining* [12]

Dalam menggunakan teknik klasifikasi diperlukan suatu algoritma pelatihan untuk mendapatkan suatu model yang paling memenuhi hubungan antara himpunan nilai dari setiap *variable* dengan hasil. Salah satu metode *learning* yang dapat digunakan adalah *decision tree*.

#### E. Decision Tree

*Decision tree* adalah sebuah struktur data yang terdiri dari kumpulan *node* yang saling terhubung melalui *edge*. Simpul pada sebuah pohon dibedakan menjadi: *root*, *branch*, dan *leaf*. *Decision tree* dapat merepresentasikan secara sederhana teknik klasifikasi untuk sejumlah kelas berhingga, dimana simpul cabang maupun simpul akar ditandai dengan nama atribut, sementara setiap rusuknya diberi label nilai dari atribut tersebut (*record*) [10]. Dalam mendapatkan solusi, tahapan yang harus dilakukan adalah mengelompokkan *set-data* ke dalam tabel, lalu mengubah tabel tersebut menjadi bentuk pohon dengan menggunakan algoritma yang ditentukan, kemudian mengubah model pohon menjadi *rules* [13].

#### F. Algoritma C4.5

Algoritma C4.5 merupakan pengembangan dari algoritma ID3 yang dilakukan Quinlan [14]. Algoritma ini dapat menyelesaikan masalah secara sistematis dengan membentuk suatu *decision tree* dengan langkah-langkah sebagai berikut:

1. Pilih atribut sebagai akar,
2. Buat cabang untuk masing-masing *record* dari atribut,
3. Membagi kasus ke dalam cabang,
4. Ulangi proses untuk masing-masing cabang sampai semua kasus pada cabang menghasilkan suatu keputusan yang sesuai.

Dalam memilih sebuah atribut menjadi akar, dilakukan perhitungan nilai dari atribut yang ada. Nilai *gain* yang paling tinggi dijadikan *root* pada pohon keputusan. Untuk menghitung nilai *gain* digunakan rumus:

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} \times Entropy(S_i) \quad \dots \text{ Rumus 2.1}$$

dengan

- : himpunan kasus
- : atribut
- : jumlah partisi atribut A
- : jumlah kasus pada partisi ke-i

Sedangkan untuk perhitungan nilai *entropy* dapat dilakukan dengan rumus:

$$Entropy(S) = -\sum_{i=1}^n p_i \times \log_2 p_i \quad \dots \text{ Rumus 2.2}$$

dengan:

- S : himpunan kasus
- n : jumlah partisi S
- pi: proporsi dari Si terhadap S

Setelah perhitungan selesai, maka dihasilkan pohon keputusan yang dapat diubah menjadi *rules* menggunakan kaidah IF-THEN-ELSE.

#### G. Katarak

Katarak adalah Bahasa Yunani, *Katarrahakies* yang berarti air terjun. Katarak merupakan suatu kelainan pada mata dimana lensa mata terlihat keruh [15]. Kekeuhan lensa dapat terjadi pada saat perkembangan serat lensa masih berlangsung atau sesudah serat lensa berhenti perkembangannya [16]. Faktor yang dapat menyebabkan seseorang mengalami katarak adalah [17]:

1. Usia
2. Cedera Mata
3. Penyakit Sistemik
4. Radiasi Sinar Ultraviolet
5. Campak (Infeksi Virus rubella)
6. Riwayat Keturunan
7. Riwayat Obat Steroid

Gejala yang dapat dialami oleh penderita katarak, antara lain:

1. Pandangan mata kabur
2. Perubahan dalam persepsi warna
3. Penglihatan ganda pada satu mata (*Halo*)
4. Penglihatan di malam hari lebih berkurang
5. Penglihatan menguning
6. Seperti ada titik gelap di depan mata
7. Susah untuk melihat di tempat yang terang karena silau
8. Pandangan kurang tajam

Jenis katarak berdasarkan penyebab katarak adalah [16]:

#### 1. Katarak Primer

Katarak yang diklasifikasikan berdasarkan usia manusia.

- a. Katarak Kongenital  
Katarak yang terlihat pada usia di bawah 1 tahun.
- b. Katarak Juvenil  
Katarak yang terlihat pada usia di atas 1 tahun dan di bawah 40 tahun.
- c. Katarak Presenilis  
Katarak yang terjadi pada penderita di bawah 30 – 40 tahun.
- d. Katarak Senilis  
Katarak yang terjadi pada penderita berusia lebih dari 40 tahun.

#### 2. Katarak Sekunder

Katarak sekunder dapat terjadi apabila penderita telah menjalani bedah lensa dan mengalami kekeruhan lensa kembali.

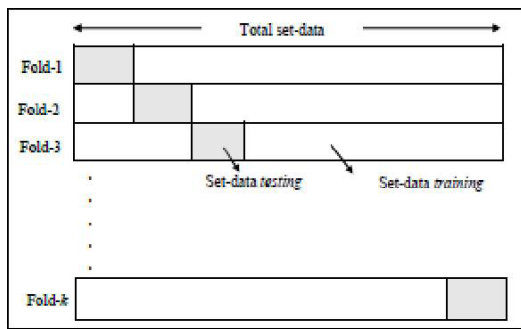
#### 3. Katarak Komplikata

Katarak Komplikata adalah katarak yang dapat terjadi akibat adanya komplikasi dari penyakit lain, penggunaan obat steroid dalam jangka panjang, dan adanya cedera mata yang pernah dialami oleh penderita.

Katarak stadium awal, dapat diatasi dengan menggunakan kacamata. Apabila penderita merasa kataraknya sudah sangat mengganggu, maka diperlukan tindakan operasi [18].

#### H. *K-fold Cross Validation*

*K-fold cross validation* adalah teknik umum yang digunakan dalam mengestimasi performa dari teknik klasifikasi pada jumlah *set-data* yang kecil. Cara untuk melakukan teknik ini adalah dengan menentukan nilai *k* yang digunakan sebagai ukuran pembagi *set-data* yang diuji. Penggunaan nilai *k* terbaik untuk menguji validitas data *training* adalah sebesar 10 [19].

Gambar 5. *K-fold cross validation* [20]

### III. METODOLOGI PENELITIAN DAN PERANCANGAN SISTEM

#### A. Metodologi Penelitian

Berdasarkan prinsip *software process* menurut Pressman [21], maka langkah-langkah yang dilakukan dalam pembangunan sistem ini adalah sebagai berikut:

##### 1. Studi Literatur

Studi literatur dilakukan untuk memperdalam wawasan dan pengetahuan mengenai topik penelitian yang dilakukan secara *valid*, serta untuk mendapatkan gambaran mengenai topik yang diambil.

##### 2. Pengumpulan Data

Pengumpulan data dilakukan dengan membuat dan memberikan *form* diagnosa kepada dokter spesialis mata untuk diisi.

##### 3. Analisis Model Sistem

Pada tahap ini, dilakukan penentuan atribut *decision tree* yang digunakan dalam sistem. Selain itu, juga dilakukan analisis model sistem secara terstruktur.

##### 4. Perancangan Sistem

Pada tahap ini, dilakukan perancangan *user interface*, perancangan *database*, dan perancangan alur sistem.

##### 5. Pembuatan program

Program yang dibuat adalah berupa *website* dengan Bahasa pemrograman PHP dan *database* MYSQL.

#### 6. Pengujian

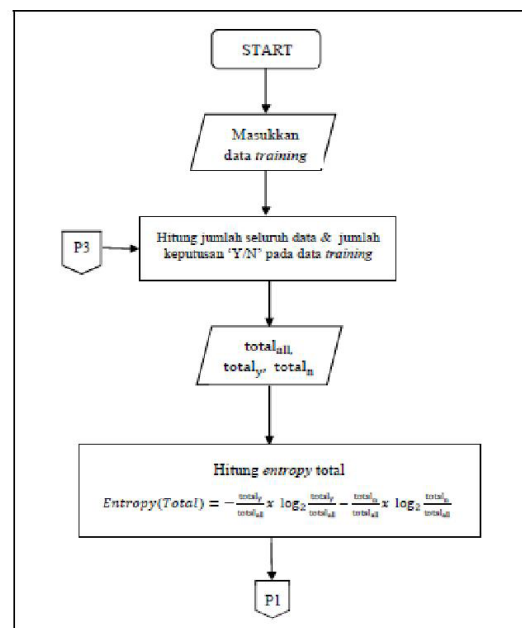
Pengujian adalah proses pelatihan sistem yang telah dibuat secara spesifik, dalam menemukan *error* sebelum digunakan oleh pengguna.

#### 7. Penarikan Simpulan

Tahapan terakhir adalah menarik simpulan dari hasil pengujian yang telah dilakukan.

#### B. Analisis Perancangan

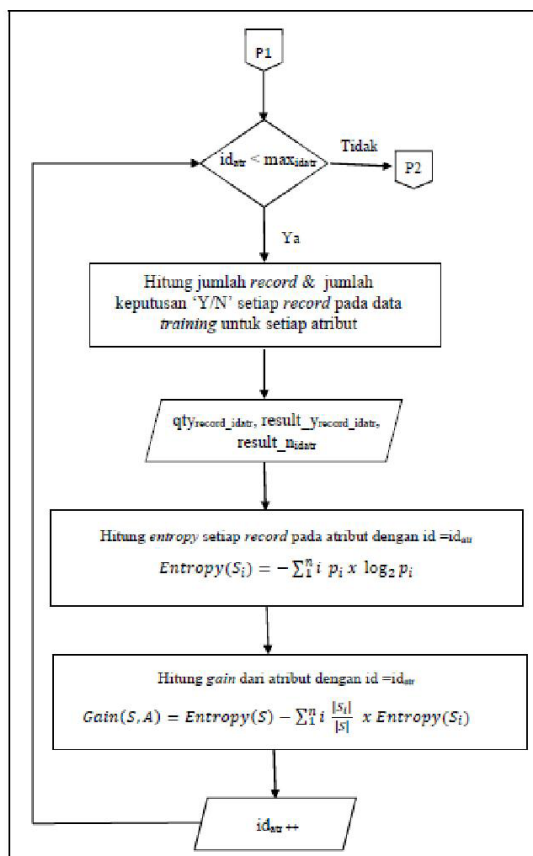
Perancangan sistem pakar deteksi dini penyakit katarak ini menggunakan algoritma C4.5. *Flowchart* algoritma C4.5 digambarkan pada tiga (3) gambar 6.

Gambar 6. *Flowchart* inialisasi awal data training

Alur proses pada algoritma C4.5 pada sistem ini dimulai dengan memasukkan *data training*, kemudian dihitung jumlah data dan jumlah keputusan pada *data training* tersebut. Kemudian, hasil dari perhitungan dijadikan *variable* untuk menghitung *entropy* total,

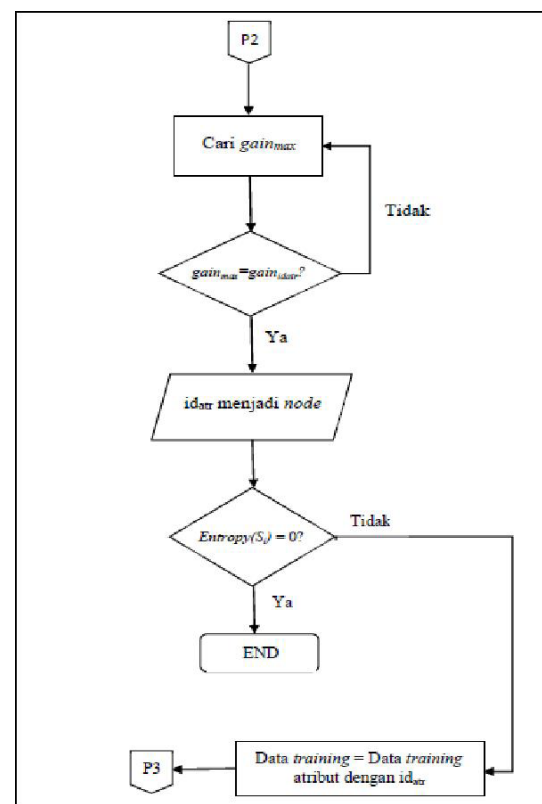
kemudian dilanjutkan ke P1.

Pada P1, pertama-tama dilakukan perhitungan jumlah *record* dan jumlah keputusan setiap *record* dari masing-masing atribut. Setelah itu, *output* yang didapatkan dijadikan *variable* untuk menghitung *entropy* dari masing-masing *record*. Setelah seluruh atribut telah diketahui jumlah *record*, jumlah keputusan setiap *record*, dan *entropy record*, maka dilanjutkan dengan menghitung nilai *gain* dari atribut dengan mengurangkan *entropy (total)* yang sudah didapat dengan probabilitas masing-masing hasil keputusan dikali *entropy* dari masing-masing *record* atribut tersebut. Jika seluruh atribut yang ada di dalam sistem telah dihitung nilai *gain*-nya, maka akan diteruskan ke P2.



Gambar 7. Flowchart proses mendapatkan variabel yang dibutuhkan setiap *record* pada atribut

Pada P2, dilakukan pencarian *maxgain*. Jika *gain* dari suatu atribut tidak sama dengan *maxgain*, maka pencarian akan diulangi dengan mencari nilai *gain* dari atribut dengan *id* berikutnya sampai ditemukan bahwa *gain* dengan *id* atribut = *idattr* merupakan *gain* yang memiliki nilai *gain* tertinggi (*maxgain*). Atribut terpilih tersebut dijadikan sebagai *node* dari pohon keputusan, sementara *record* dari atribut tersebut sebagai *leaf* dari pohon keputusan. Kemudian, dilakukan pengecekan pada masing-masing *record* yang menjadi *leaf*, apakah *entropy* yang dimiliki *record* tersebut adalah 0? Jika 'Ya', maka *leaf* tersebut langsung menghasilkan sebuah keputusan. Jika 'Tidak', maka masuk ke dalam proses *me-generate* data *training* baru sesuai dengan kondisi dari *node-leaf* tersebut, kemudian dilanjutkan ke P3. Pada P3, dilakukan perhitungan jumlah data dan jumlah keputusan yang memenuhi kondisi *node-leaf* tersebut. Lalu, dilakukan kembali perhitungan *entropy* dan *gain* yang dilanjutkan dengan penentuan *node* berikutnya.



Gambar 8. Flowchart pembentukan *node* dari pohon keputusan

#### IV. IMPLEMENTASI DAN UJI COBA

Sistem yang dibangun dalam penelitian ini berupa sebuah *website* yang dibedakan menjadi bagian *frontend* dan *backend*. Berikut penjelasan hasil implementasi yang diterapkan pada masing-masing bagian.

##### A. Frontend

Terdapat empat halaman utama pada *frontend website* ini, yaitu:

##### 1. Halaman Index

Halaman ini merupakan halaman utama yang ditampilkan saat *user* masuk pertama kali ke dalam sistem yang dibangun.



Gambar 9. Hasil implementasi halaman *index* dari *frontend website*

##### 2. Halaman Diagnose

Tampilan halaman *diagnose* pertama kali adalah seperti pada Gambar 10. *Visitor* diminta untuk menginputkan nama dan lokasinya. Setelah itu, *visitor* dapat mengklik *button* 'Mulai', sehingga *pop-up* pertanyaan dapat muncul seperti pada Gambar 11.



Gambar 10. Hasil implementasi halaman *diagnose* dari *frontend website*



Gambar 11. *Pop-up* pertanyaan pada halaman *diagnose frontend website*

##### 3. Halaman News

Pada halaman *news*, *visitor* dapat melihat berbagai berita yang terakhir di-*posting*. Tampilan halaman *news* dapat dilihat pada Gambar 12.



Gambar 12. Hasil implementasi halaman *news* dari *frontend website*

##### 4. Halaman Nearby

Tampilan halaman *nearby* pertama kali adalah seperti pada Gambar 13.



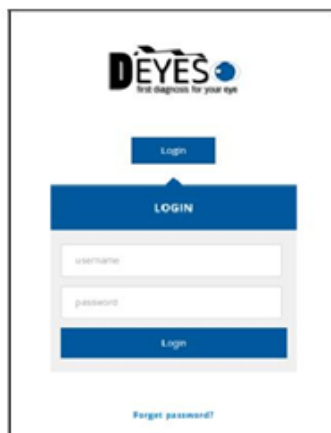
Gambar 13. Hasil implementasi halaman *nearby* dari *frontend website*



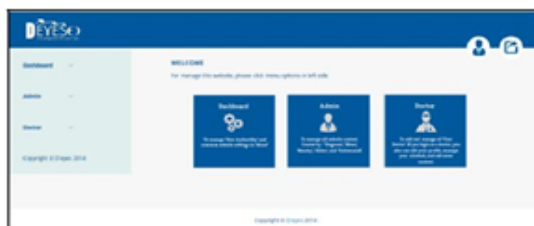
Gambar 14. Tampilan setelah memilih lokasi pada halaman *nearby*

### B. Backend

Untuk masuk ke dalam halaman *backend*, *user* harus melakukan *login* terlebih dahulu. Apabila *user* melakukan *login* dengan benar, maka langsung masuk ke halaman *home* dari *website* sesuai dengan otoritas yang *user* miliki. Terdapat dua (2) otoritas *user* pada *website* ini yaitu *admin* dan *doctor*. Apabila *user* memiliki otoritas sebagai *admin*, maka *user* dapat melihat tampilan halaman *home* seperti pada Gambar 16.



Gambar 15. Halaman *login backend website*



Gambar 16. Halaman *home* dari *backend* untuk *user* dengan otoritas *admin*

Sementara, apabila *user* memiliki otoritas sebagai *doctor*, maka *user* dapat melihat tampilan halaman *home* seperti pada Gambar 17.



Gambar 17. Halaman *home* dari *backend* untuk *user* dengan otoritas *doctor*

### C. Uji Fungsionalitas

Pada pembangunan sistem ini *set-data* yang digunakan berupa data pasien yang telah melakukan deteksi mata. Data ini didapatkan dengan membuat sebuah *form diagnosa* yang berisi pernyataan dan jawaban yang berhubungan dengan penyebab dan gejala katarak. Seluruh *form* tersebut diisi langsung oleh pakar yang terlibat.

Pakar yang terlibat pada penelitian ini adalah dokter spesialis mata. Terdapat tiga (3) dokter spesialis mata yang turut membantu pada penelitian ini, yaitu: Dr. Melissa Yulita Sp.M sebanyak 67 *form*, Dr. Cynthia Sp.M sebanyak 60 *form*, dan Dr. Budi Suryanto Sp.M sebanyak 50 *form*. Total seluruh *form* adalah 177 *form*. Setelah *form* terkumpul, selanjutnya dilakukan pengolahan berdasarkan tahapan KDD. Tahapan tersebut adalah:

#### 1. Selection

Data yang masuk ke tahap berikutnya sebanyak 150 data.

#### 2. Preprocessing

Data yang terseleksi digabungkan menjadi sebuah *file* berekstensi *comma separated value (.csv)*.

#### 3. Transformation

Pada tahap ini dilakukan langkah-langkah sebagai berikut pada sistem:

- a. Masuk ke sistem sebagai *user admin*,



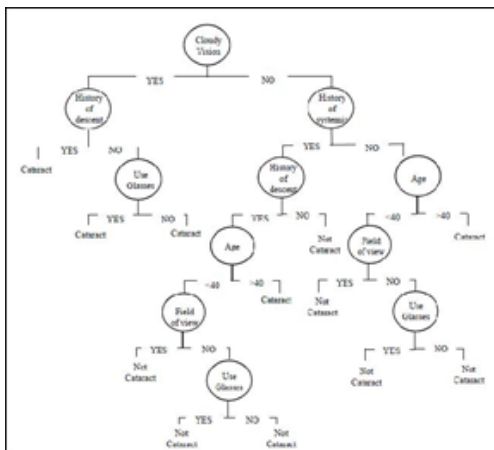
- b. Pilih *side-menu* 'Admin' > pilih *sub-menu* 'Diagnose' > pilih menu 'Training',
- c. Klik *button* 'Add New Training Set', setelah itu akan berpindah ke halaman 'New Training Set',
- d. Klik 'Browse' > pilih *file* dengan *format* .csv yang sudah dibuat sebelumnya > klik *button* 'Confirm'.

4. Data Mining

Setelah semua data *training* sudah ter-*upload*, selanjutnya dilakukan penerapan metode *decision tree*.

5. Interpretation/ Evaluation

Setelah melewati proses *mining*, maka ditemukanlah pola dari *decision tree* seperti pada Gambar 18. Dalam gambar 18, terlihat bahwa atribut pandangan mata kabur (*cloudy vision*) menjadi akar (*root*) dari pohon tersebut. Sementara beberapa atribut lainnya ada yang digunakan sebagai percabangan, ada yang tidak.



Gambar 18. Halaman *home* dari *backend* untuk *user* dengan otoritas *doctor*

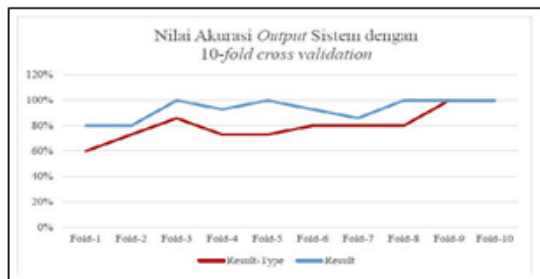
D. Analisis Hasil Uji

Jumlah *set* data yang digunakan pada sistem ini adalah 150 dan angka *k* yang digunakan pada sistem ini adalah 10. Jadi di setiap *fold* terdapat 15 *set-data* sebagai data *testing* dan 135 *set-data* sebagai data *training*.

Setelah *10-fold cross validation* telah selesai dilakukan, maka selanjutnya adalah mencari nilai rata-rata dari nilai akurasi yang telah didapat tersebut. Nilai akurasi pada penelitian ini dibagi menjadi dua jenis, yaitu nilai akurasi *result* beserta *type* dari katarak dan nilai akurasi *result* saja. Perbedaan nilai akurasi keduanya dapat dilihat pada Tabel 1.

Tabel 1. Rata-rata nilai akurasi *output* sistem dengan *10-fold cross validation*

Fold ke-	Nilai Akurasi (Result-Type)	Nilai Akurasi (Result)
1	60%	80%
2	73%	80%
3	86%	100%
4	73%	93%
5	73%	100%
6	80%	93%
7	80%	86%
8	80%	100%
9	100%	100%
10	100%	100%
<b>Rata-rata</b>	<b>80.5%</b>	<b>93.2%</b>



Gambar 19. Nilai akurasi *output* sistem pada uji *10-fold cross validation*

V. SIMPULAN DAN SARAN

A. Simpulan

Simpulan yang dapat didapat dari penelitian yang telah dilakukan adalah:

1. Sistem pakar untuk mendeteksi adanya penyakit mata katarak secara dini, sudah berhasil dibangun. Sistem ini berbasis *website*, dengan mengimplementasikan algoritma C4.5 dari metode *learning decision tree*.
2. Sistem pakar ini dibangun berdasarkan data-data yang didapatkan dari *form* diagnosa

yang telah dikumpulkan dari dokter spesialis mata.

3. Berdasarkan hasil uji yang telah dilakukan dengan menggunakan *10-fold cross validation* didapatkan nilai akurasi *type output* sistem sebesar 80.5%, sementara nilai akurasi dari *result output* sistem sebesar 93.2%.

#### B. Saran

Saran yang dapat disampaikan bagi penelitian selanjutnya adalah:

1. Dalam penelitian selanjutnya, dapat digunakan lebih banyak lagi *set-data* yang bervariasi pada *data-training* dalam membangun *decision tree*.
2. Dalam penelitian selanjutnya, dapat digunakan algoritma lain dari metode *decision tree* seperti algoritma C5.0. Dapat juga diterapkan metode *learning* lain seperti *Neural Network* karena metode diprediksi memiliki nilai akurasi kedua terbesar setelah metode *decision tree* [4].
3. Sistem ini dapat dikembangkan untuk dapat mendeteksi penyakit mata lainnya atau melakukan klasifikasi jenis seluruh penyakit katarak yang ada.

#### DAFTAR PUSTAKA

- [1] World Health Organization (WHO). Priority Eye Diseases [Online]. Available at: <http://www.who.int/blindness/causes/priority/en/index1.html> [Accessed 24 April 2014].
- [2] World Health Organization (WHO). Global Data On Visual Impairment [Online]. Available at: <http://www.who.int/blindness/publications/globaldata/en/> [Accessed 11 February 2015].
- [3] Badan Penelitian dan Pengembangan Kesehatan Kementerian Kesehatan RI. 2013. *Riset Kesehatan Dasar*. Jakarta.
- [4] Hastuti, Khafiizh. 2012. Analisis Komparasi Algoritma Klasifikasi Data Mining Untuk Prediksi Mahasiswa Non-aktif. *Seminar Nasional Teknologi Informasi & Komunikasi Terapan*. Semarang, 23 Juni 2012, hal.241-249.
- [5] Suyanto. 2014. *Artificial Intelligence: Searching, Reasoning, Planning dan Learning*. Bandung: Informatika Bandung.
- [6] Rich, E. dan Knight, K. 1991. *Artificial Intelligence. 2nd Edition*. New York: Mc-Graw-Hill.
- [7] Patterson, D. 1990. *Introduction to Artificial Intelligence and Expert Systems*. USA: Prentice Hall.
- [8] Arhami, Muhammad. 2006. *Konsep Dasar Sistem Pakar*. Jilid 1. Yogyakarta: Penerbit Andi.
- [9] Turban, E. 1995. *Decision Support System and Expert Systems*. 4th Edition. USA: Prentice Hall International.
- [10] Hermawati, Fajar Astuti. 2013. *Data Mining*. Yogyakarta: Penerbit Andi.
- [11] Larose, Daniel T. 2005. *Discovering Knowledge in Data: An Introduction to Data Mining*. John Willey & Sons, Inc.
- [12] Prasetyo, Eko. 2012. *Data Mining Konsep dan Aplikasi menggunakan MATLAB*. Yogyakarta: Penerbit Andi.
- [13] Basuki, A. dan Syarif, Iwan. 2003. *Modul ajar Decision Tree*. Surabaya: PENS-ITS.
- [14] Quinlan, J.R. 1992. *C4.5: Programs for machine learning*. San Mateo, California: Morgan Kaufmann Publishers.
- [15] Ilyas, Sidarta, dkk. 2002. *Ilmu Penyakit Mata: Untuk Dokter Umum dan Mahasiswa Kedokteran*. Edisi kedua. Jakarta: Sagung Seto.
- [16] Ilyas, Sidarta. 2001. *Penuntun Ilmu Penyakit Mata*. Edisi Kedua. Jakarta: Balai Penerbit FKUI.

- 
- [17] Adriansah, Muhammad Akbar. Karakteristik Penderita Katarak di Puskesmas Ciputat Tahun 2006-2010 [Skripsi]. Jakarta: UIN Syarif Hidayatullah. 2011.
- [18] Ilyas, Sidarta. 1989. *Masalah Kesehatan Mata Anda dalam Pertanyaan-pertanyaan*. Jakarta: Balai Penerbit FKUI.
- [19] Breiman, L., Spector, P. 1992. Submodel Selection and Evaluation in Regression. The X-Random Case. *International Statistical Review*, Vol. 60 No. 3, December, hal. 291-319.
- [20] Hastie, T., Tibshirani, R., dan Friedman, J. 2008. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd Edition*. Stanford, California: Springer.
- [21] Pressman, Roger S. 2010. *Software Engineering: A practitioner's Approach. 7th Edition*. New York: McGraw-Hill.