



## Analisis Rekam Medis untuk Menentukan Pola Kelompok Penyakit Menggunakan Algoritma C4.5

Rian Rafiska<sup>a</sup>, Sarjon Defit<sup>b</sup>, Gunadi Widi Nureahyo<sup>c</sup>

<sup>a</sup>Pasca Sarjana, Magister Komputer, Universitas Putra Indonesia, rianrafiska@gmail.com

<sup>b</sup>Pasca Sarjana, Magister Komputer, Universitas Putra Indonesia, sarjonde@yahoo.co.uk

<sup>c</sup>Pasca Sarjana, Magister Komputer, Universitas Putra Indonesia, gunadiwidi@yahoo.co.id

### Abstract

*The Medical Record contains records and documents of patient identity, examination results, treatment, actions and services provided to the patient. Medical records are very important for patient care because with complete data can provide information in determining diagnostic and clinical decisions. The completeness of the medical record determines the quality of the services provided. Regarding the pattern of the tendency of disease suffered by a group of people still not excavated to be used as a reference when doing panyuluhan or prevention of disease. Finding a common pattern of disease groups in the community based on the International Classification of Diseases (ICD)-X. In this study used the classification method with algorithm C4.5 with the amount of data as much as 709 sourced from the Medical Record of General Hospital General Hospital (RSUD) Major General H.A Thalib Kerinci. Determination of the next analysis is to apply the grouping into several attributes, namely group of regions, age groups, disease groups and groups of sex. Further data is processed and done by using Rapid Miner software. The results of the calculation is a pattern that can be used to analyze patterns of disease tendency experienced by the community.*

*Keywords: Medical Record, Data Mining, Algorithm C4.5*

### Abstrak

Rekam Medis berisi catatan dan dokumen identitas pasien, hasil pemeriksaan, pengobatan, tindakan serta pelayanan yang diberikan kepada pasien. Rekam medis sangat penting untuk pelayanan pasien karena dengan data yang lengkap dapat memberikan informasi dalam menentukan keputusan diagnosis dan klinis. Kelengkapan dari rekam medis menentukan mutu dari pelayanan yang diberikan. Mengenai pola dari kecenderungan penyakit yang di derita oleh sekelompok masyarakat masih belum digali untuk dijadikan acuan apabila melakukan panyuluhan atau pencegahan penyakit. Menemukan pola kelompok penyakit yang sering terjadi di masyarakat berdasarkan kode penyakit dalam International Classification of Diseases (ICD)-10. Pada penelitian ini digunakan metode klasifikasi dengan algoritma C4.5 dengan jumlah data sebanyak 709 yang bersumber dari bagian Rekam Medis Rumah Sakit Umum Daerah (RSUD) Mayjen H.A Thalib Kerinci. Penentuan analisa selanjutnya adalah melakukan pengelompokan menjadi beberapa attribute yaitu kelompok wilayah, kelompok umur, kelompok penyakit dan kelompok jenis kelamin. Selanjutnya data diolah dan dilakukan pengujian menggunakan software Rapid Miner. Hasil dari perhitungan adalah pola yang dapat digunakan untuk menganalisa pola kecenderungan penyakit yang dialami oleh masyarakat.

*Kata kunci: Rekam Medis, Data Mining, Algoritma C4.5*

© 2018 Jurnal RESTI

### 1. Pendahuluan

Kemajuan teknologi informasi telah menyebabkan banyak orang dapat memperoleh data dengan mudah bahkan cenderung berlebihan. Data tersebut semakin lama semakin banyak dan terakumulasi, akibatnya pemanfaatan data yang terakumulasi tersebut menjadi tidak optimal. Banyaknya data yang dimiliki oleh sebuah organisasi bisa menyebabkan kesulitan dalam pengklasifikasian data tersebut untuk kepentingan organisasi. Kegiatan pengklasifikasian yang dilakukan

oleh manusia masih memiliki keterbatasan, terutama pada kemampuan manusia dalam menampung jumlah data yang ingin diklasifikasikan. Selain itu bisa juga terjadi kesalahan dalam pengklasifikasian yang dilakukan. Salah satu cara mengatasi masalah ini adalah dengan menggunakan Data Mining dengan teknik klasifikasi. Data Mining dapat membantu sebuah organisasi yang memiliki data melimpah untuk memberikan informasi yang dapat mendukung pengambilan keputusan.

Salah satu penerapan ilmu Data Mining, yaitu pada permasalahan penumpukan data rekam medis di Rumah Sakit. Rekam medis merupakan berkas yang berisikan rangkaian catatan dan dokumen medis yang berisi tentang identitas pasien, pemeriksaan, pengobatan, tindakan serta pelayanan yang diberikan kepada pasien[1]. Rekam medis harus dibuat secara tertulis, lengkap, dan jelas atau secara elektronik. Penyelenggaraan rekam medis dengan menggunakan teknologi informasi elektronik diatur oleh peraturan tersendiri. Informasi dalam rekam medis dijaga kerahasiannya oleh dokter, tenaga kesehatan dan petugas pengelola serta pimpinan sarana pelayanan kesehatan. Data rekam medis terus terakumulasi setiap hari seiring dengan aktivitas rumah sakit[2].

Tumpukan data yang ada baik di dinas kesehatan maupun di rumah sakit saat ini hanya sebatas memberikan grafik atau statistik jumlah pasien yang berobat dengan penyakit yang dideritanya beserta laporan kepulauan pasien tersebut. Laporan dari data inilah yang saat ini dijadikan oleh dinas kesehatan untuk melakukan kebijakan-kebijakan apabila akan memberikan penyuluhan kepada masyarakat. Mengenai pola dari kecenderungan penyakit yang diderita oleh sekelompok masyarakat masih belum digali untuk dijadikan acuan apabila melakukan penyuluhan atau pencegahan penyakit.

Berdasarkan latar belakang di atas adapun yang menjadi pokok permasalahan dalam penelitian ini adalah bagaimana mengimplementasikan algoritma C4.5 untuk menemukan pola penyakit yang sering terjadi di masyarakat.

## 2. Tinjauan Pustaka

Knowledge discovery in database (KDD) merupakan proses untuk menemukan informasi yang berguna dalam database. Seluruh proses KDD biasanya terdiri dari langkah-langkah, yaitu memahami bidang aplikasi, membuat data target yang ditetapkan dari data mentah yang tersimpan dalam database, pembersihan data dan preprocessing data [3]

### 2.1 Data Mining

Data Mining adalah proses menemukan atau penggalian pola-pola baru dari kumpulan data besar yang melibatkan metode dari statistik dan kecerdasan buatan [4]. Data Mining adalah kegiatan yang meliputi pengumpulan, pemakaian data historis yang menentukan keteraturan, pola, dan hubungan dalam set data berukuran besar [5]. Dalam Data Mining terdapat beberapa teknik yang bisa digunakan, antara lain : deskripsi, estimasi, prediksi, klasifikasi, clustering, dan asosiasi [6].

### 2.2 Klasifikasi

Klasifikasi adalah teknik Data Mining yang memetakan data kedalam kelompok-kelompok yang telah ditetapkan atau kelas [7]. Menurut [8], klasifikasi adalah proses penemuan model atau fungsi yang menggambarkan dan membedakan kelas data atau konsep yang bertujuan agar bisa digunakan untuk memprediksi kelas dari objek yang label kelas nya tidak diketahui.

### 2.3 Algoritma C4.5

Algoritma ini merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal. Algoritma C4.5 adalah pengembangan dari algoritma iterative dichotomizer 3 (ID3). Algoritma C4.5 mempunyai prinsip dasar kerja yang sama dengan algoritma ID3, hanya saja algoritma C4.5 menggunakan pendekatan induksi dimana, algoritma C4.5 membagi-bagi data berdasarkan kriteria yang dipilih untuk membuat pohon keputusan[9].

Terdapat 4 langkah dalam menentukan pohon keputusan menggunakan algoritma C4.5 [10] adalah:

1. Memilih atribut sebagai akar (root).
2. Membuat cabang untuk tiap-tiap nilai.
3. Membagi kasus dalam cabang
4. Mengulangi proses dalam setiap cabang, sampai semua kasus dalam cabang memiliki kelas yang sama.

Untuk memilih atribut akar, didasarkan pada nilai *gain* tertinggi dari atribut-atribut yang ada. Untuk menghitung *gain* digunakan rumus seperti yang tertera dalam persamaan 1 berikut.

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

Di mana :

- S : himpunan kasus
- A : atribut
- N : jumlah partisi atribut A
- |S<sub>i</sub>| : jumlah kasus pada partisi ke-i
- |S| : jumlah kasus dalam S

Sementara itu, perhitungan nilai entropi dapat dilihat pada persamaan 2 berikut.

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

Di mana :

- S : himpunan kasus
- N : jumlah partisi S
- p<sub>i</sub> : proporsi dari S<sub>i</sub> terhadap S.

Commented [a1]: (1)

Commented [a2]: (2)

Commented [a3]:

**3. Metodologi Penelitian**

Pada penelitian ini, sebelum melakukan perhitungan menggunakan algoritma C4.5, ada beberapa tahapan yang harus dilakukan terlebih dahulu, yaitu :

1. Data Selection

Menyeleksi data rekam medis yang akan digunakan untuk perhitungan algoritma C4.5. Data yang telah diseleksi di masukkan ke dalam excel dan di simpan dalam format .xls.

2. Data Cleaning

Proses cleaning mencakup membuang duplikasi data, memeriksa data yang tidak lengkap, dan memperbaiki kesalahan pada data seperti kesalahan cetak.

3. Data Transformation

Merubah data tiap atribut menjadi beberapa kategori. Hal ini dilakukan untuk memudahkan dalam melakukan perhitungan dengan menggunakan algoritma C4.5

Setelah mendapatkan data transformation, proses selanjutnya yaitu pengolahan data menggunakan algoritma C4.5 yang langkah-langkahnya dapat dilihat pada Gambar 1.



Gambar 1. Proses pada Algoritma C4.5

Dari Gambar 1 tersebut dapat diketahui langkah-langkah dalam melakukan perhitungan algoritma C4.5, yang mana pada tahap awlanya memasukkan data rekam medis yang telah di transformasi, kemudian dari data transformasi tersebut dapat dilakukan perhitungan untuk menghitung entropy, setelah mendapatkan entropy selanjutnya melakukan perhitungan untuk menghitung gain dan menentukan gain yang tertinggi. Gain tertinggi inilah yang akan dijadikan sebagai node cabang dalam pohon keputusan. Selanjutnya lakukan lagi perhitungannya sampai semua cabang memiliki kelas yang sama.

**4. Hasil dan Pembahasan**

Data yang nantinya akan diolah mempunyai beberapa kriteria yang merupakan syarat dalam pengolahan Data Mining dengan menggunakan teknik algoritma C4.5. terdapat 4 atribut dalam penelitian ini, yaitu : (1) jenis kelamin yang terdiri dari laki-laki (LK) dan perempuan (PR); (2) umur yang terdiri dari kategori bayi dan anak (<15 tahun), muda & dewasa (15-50 tahun), tua (>50 tahun); (3) alamat yang dikelompokkan menjadi pedesaan, pegunungan, perkotaan; (4) kode ICD-X

berdasarkan kode penyakit, yaitu : A00-B99, C00-D48, D50-D89, E00-E90, F00-F99, G00-G99, H00-H59, H60-H95, I00-I99, J00-J99, K00-K93, L00-L99, M00-M99, N00-N99, O00-O99, P00-P96, Q00-Q99, R00-R99, S00-T98, V01-Y98, Z00-Z99.

Setelah melakukan klasifikasi, selanjutnya terbentuk format data akhir yang dapat di lihat pada Tabel 1.

Tabel 1. Format Data Transformasi Rekam Medis

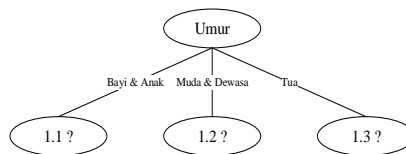
No	Jekel	Umur	Alamat	Icd-x
1	LK	Muda & Dewasa	Gunung	R00-R99
2	LK	Tua	Desa	I00-I99
3	PR	Muda & Dewasa	Desa	O00-O99
4	PR	Tua	Kota	G00-G99
5	LK	Tua	Desa	K00-K93
6	LK	Bayi & Anak	Kota	R00-R99
7	LK	Bayi & Anak	Desa	A00-B99
8	LK	Tua	Gunung	D50-D89
9	PR	Muda & Dewasa	Desa	R00-R99
10	LK	Tua	Kota	S00-T98

Langkah awal pengelohan data adalah mencari atribut akar dengan melakukan perhitungan terhadap entropy total dan menentukan gain tertinggi. Hasilnya dapat dilihat pada Tabel 2.

Tabel 2. Tabel Perhitungan Gain Tertinggi

Node 1	A00-B99	..	Z00-Z99	Ent	Gain
Total	709	130	..	11	3,1144
Jekel					1,2108
L	286	66	..	0	0
P	423	64	..	11	3,1907
Umur					1,8821
Bayi & Anak	155	70	..	0	0
Muda & Dewasa	277	26	..	9	0
Tua	277	34	..	2	3,1543
Alamat					0,4449
Desa	427	84	..	5	2,9947
Kota	89	9	..	1	0
Gunung	193	37	..	5	3,1809

Dari Tabel 2 dapat dilihat nilai entropy dan nilai gain untuk setiap atribut. Untuk mendapatkan pohon keputusan node 1, dapat menggunakan nilai gain yang tertinggi. Pada pencarian pohon keputusan nilai gain yang terbesar berada pada atribut Umur yaitu sebesar 1,8821. Dengan demikian Umur dapat dijadikan sebagai node 1. Ada 3 klasifikasi pada atribut Umur yaitu Bayi & Anak, Muda & Dewasa, dan Tua. Lihat Gambar 2.



Gambar 2. Pohon Keputusan Node 1

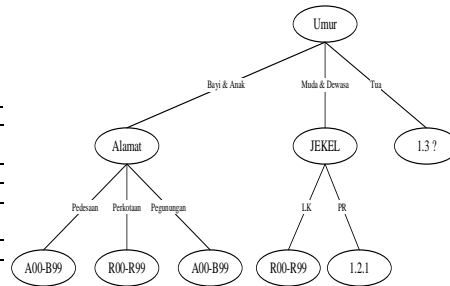
Commented [a4]: 1 space

Setelah mendapatkan atribut yang menjadi cabang node, maka selanjutnya mencari atribut mana yang akan menjadi cabang level 1.1. Pada cabang level 1.1 terdapat 3 perhitungan klasifikasi yang berbeda yaitu atribut Umur klasifikasi Bayi & Anak, Muda & Dewasa dan Tua. Adapun perhitungan selanjutnya yang akan dilakukan adalah pencarian pohon keputusan cabang level 1.1. dengan cara yang sama seperti diatas.

Tabel 3. Tabel Perhitungan Cabang Level 1.1

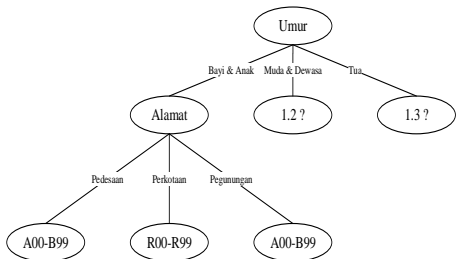
Node 1.1 (Umur_Anak)					
	A00-B99	..	Z00-Z99	Ent	Gain
Total	155	70	..	0	1,8849
Jekel					1,8849
L	75	36	..	0	0
P	80	34	..	0	0
Alamat					1,8849
Desa	90	45	..	0	0
Kota	20	4	..	0	0
Gunung	45	21	..	0	0

Dengan demikian Jekel dapat dijadikan sebagai cabang level 1.2. Ada 2 klasifikasi dari atribut Jekel yaitu LK dan PR. Untuk klasifikasi PR perlu dilakukan perhitungan kembali untuk mencari cabang level 1.2.1. Lihat Gambar 4.



Gambar 4. Pohon Keputusan Node 1.2

Dari Tabel 3 dapat dilihat bahwa nilai *gain* sama besar, yaitu 1,8849. Dikarenakan nilai *gain* yang sama, maka dapat dipilih salah satu yang akan menjadi cabang level 1.1. Pada perhitungan ini peneliti memilih atribut Alamat untuk dijadikan cabang level 1.1. Ada 3 klasifikasi pada atribut Alamat yaitu Pedesaan, Perkotaan, dan Pegunungan. Lihat Gambar 3.



Gambar 3. Pohon Keputusan Node 1.1

Setelah pencarian cabang level 1.1, maka dilanjutkan pencarian cabang level 1.2. Pada cabang level 1.2 nilai entropy atribut yang digunakan adalah Umur Muda & Dewasa.

Tabel 4. Tabel Perhitungan Cabang Level 1.1

Node 1.1 (Umur_Muda)					
	A00-B99	..	Z00-Z99	Ent	Gain
Total	277	26	..	9	3,0433
Jekel					0,8928
L	80	13	..	0	0
P	197	13	..	9	3,0237
Alamat					0,8380
Desa	159	18	..	4	2,9790
Kota	40	3	..	1	3,4296
Gunung	78	5	..	4	0

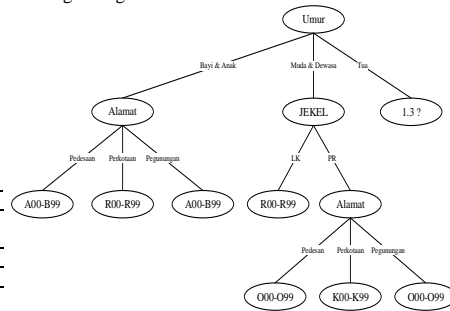
Dari hasil pada Tabel 4 dapat diketahui bahwa atribut dengan *Gain* tertinggi adalah Jekel sebesar 0,8928.

Setelah pencarian cabang level 1.2, maka dilanjutkan pencarian cabang level 1.2.1. Pada cabang level 1.2.1 nilai entropy atribut yang digunakan adalah Jekel PR.

Tabel 5. Tabel Perhitungan Cabang Level 1.1

Node 1.1 (Umur_Muda) (Jekel_PR)					
	A00-B99	..	Z00-Z99	Ent	Gain
Total	197	13	..	9	3,0237
Alamat					3,0237
Desa	112	8	..	0	0
Kota	29	1	..	1	0
Gunung	56	4	..	4	0

Dari hasil pada Tabel 5 dapat diketahui bahwa atribut Alamat memiliki nilai *gain* sebesar 3,0237. Dengan demikian Alamat menjadi cabang level 1.2.1. Ada 3 klasifikasi atribut Alamat yaitu Pedesaan, dan Pegunungan. Lihat Gambar 5.



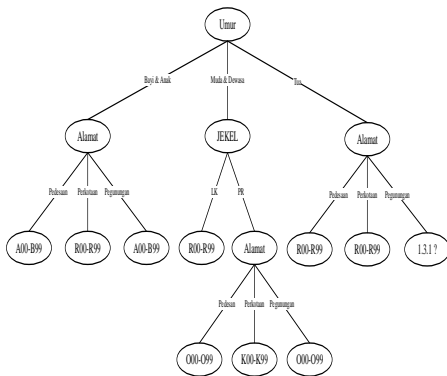
Gambar 5. Pohon Keputusan Node 1.2.1

Setelah pencarian cabang level 1.2.1, maka dilanjutkan pencarian cabang level 1.3. Pada cabang level 1.3. nilai entropy atribut yang digunakan adalah Umur Tua.

Tabel 6. Tabel Perhitungan Cabang Level 1.3

Node 1.1 (Umur_Tua)					
	A00- B99	..	Z00- Z99	Ent	Gain
Total	277	34	..	2	3,1543
Jekel	1,4825				
L	131	17	..	0	0
P	146	17	..	2	3,1717
Alamat	2,3097				
Desa	178	21	..	1	0
Kota	29	2	..	0	0
Gunung	70	11	..	1	3,3420

Dari hasil pada Tabel 6 dapat diketahui bahwa atribut dengan *Gain* tertinggi adalah Alamat sebesar 2,3097. Dengan demikian Alamat dapat dijadikan sebagai cabang level 1.3. Ada 3 klasifikasi dari atribut Alamat yaitu Pedesaan, Perkotaan dan Pegunungan. Untuk klasifikasi Pegunungan perlu dilakukan perhitungan kembali untuk mencari cabang level 1.3.1. Lihat Gambar 6.



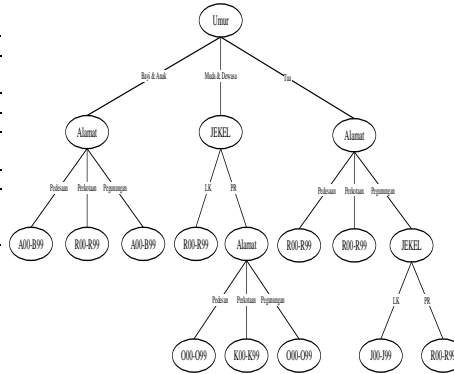
Gambar 6. Pohon Keputusan Node 1.3.1

Setelah pencarian cabang level 1.3, maka dilanjutkan pencarian cabang level 1.3.1. Pada cabang level 1.3.1 nilai entropy atribut yang digunakan adalah Alamat Pegunungan.

Tabel 7. Tabel Perhitungan Cabang Level 1.3.1

Node 1.1 (Umur_Tua) (Alamat_Pegunungan)					
	A00- B99	..	Z00- Z99	Ent	Gain
Total	70	11	..	1	3,3420
Jekel	3,3420				
L	29	5	..	0	0
P	41	6	..	1	0

Dari hasil pada Tabel 7 dapat diketahui bahwa atribut Jekel memiliki nilai *gain* sebesar 3,3420. Dengan demikian Jekel menjadi cabang level 1.3.1. Ada 2 klasifikasi atribut Jekel yaitu LK, dan PR. Lihat Gambar 7.



Gambar 7. Pohon Keputusan Node 1.3.1

Rule hasil dari prediksi berdasarkan pada pohon keputusan terakhir yang terbentuk sesuai dengan perhitungan Entropy dan *Gain*. Melalui pohon keputusan tersebut diperoleh 12 aturan (rule) dalam memprediksi jumlah mahasiswa yang mengulang mata kuliah. Adapun aturan atau rule yang terbentuk dari perhitungan *Gain* dan Entropy setiap variabel sampai menghasilkan Node 1.5 adalah sebagai berikut :

1. IF Umur = Bayi & Anak, AND Alamat = Pedesaan, THEN Kode ICD-X = A00-B99.I
2. F Umur = Bayi & Anak, AND Alamat = Perkotaan, THEN Kode ICD-X = R00-R99.
3. IF Umur = Bayi & Anak, AND Alamat = Pegunungan, THEN Kode ICD-X = A00-B99.
4. IF Umur = Muda & Dewasa, AND Jenis Kelamin = Laki-Laki, THEN Kode ICD-X = R00-R99.
5. IF Umur = Muda & Dewasa, AND Jenis Kelamin = Perempuan, AND Alamat = Pedesaan, THEN Kode ICD-X = O00-O99.
6. IF Umur = Muda & Dewasa, AND Jenis Kelamin = Perempuan, AND Alamat = Perkotaan, THEN Kode ICD-X = K00-K99.
7. IF Umur = Muda & Dewasa, AND Jenis Kelamin = Perempuan, AND Alamat = Pegunungan, THEN Kode ICD-X = O00-O99.
8. IF Umur = Tua, AND Alamat = Pedesaan, THEN Kode ICD-X = R00-R99.
9. IF Umur = Tua, AND Alamat = Perkotaan, THEN Kode ICD-X = R00-R99.
10. IF Umur = Tua, AND Alamat = Pegunungan, AND Jenis Kelamin = Laki-Laki, THEN Kode ICD-X = J00-J99.
11. IF Umur = Tua, AND Alamat = Pegunungan, AND Jenis Kelamin = Perempuan, THEN Kode ICD-X = R00-R99.

## 5. Kesimpulan

Bagian terdiri atas simpulan dan saran atas hasil penelitian.

### 5.1 Simpulan

1. Pemilihan variabel Umur, Jenis Kelamin dan Alamat, dapat menjadi kriteria penilaian terhadap pelanggan aktif dan tidak aktif dengan menggunakan Algoritma C4.5.
2. Algoritma C4.5 telah berhasil di terapkan dalam menganalisis data rekam medis di RSUD Mayjen H.A. Thalib Kerinci
3. Implementasi algoritma C4.5 dengan memanfaatkan software RapidMiner dalam menganalisis data rekam medis menghasilkan parameter-parameter keputusan berupa pohon keputusan yang baik dalam pengambilan keputusan di RSUD Mayjen H.A. Thalib Kerinci

### 5.2 Saran

1. Diharapkan kepada penelitian selanjutnya dapat menggunakan data yang lebih banyak, karena pada penelitian ini peneliti menggunakan data untuk 1 bulan saja
2. Pada penelitian ini, penulis mencoba salah satu teknik yang digunakan yaitu teori decision tree algoritma C4.5. Untuk mendapatkan hasil prediksi yang baik dapat digunakan beberapa atau penggabungan beberapa teknik klasifikasi. Oleh sebab itu, perlu dilakukan penelitian lebih lanjut untuk membandingkan hasil klasifikasi dan menentukan teori mana yang menghasilkan klasifikasi yang baik.

## Daftar Rujukan

- [1] Xinyu J., Yizheng L., Chunhui M., 2016. Medical Record Semantic Analysis Based on Weighted LDA. *International Symposium on Computational Intelligence and Design*, 9, pp. 591-596
- [2] Wenefrida T.I., 2016. Klasifikasi Data Rekam Medis Berdasarkan Kode Penyakit Internasional Menggunakan Algoritma C4.5. *Jurnal Media Elektro*, 1 (3), pp. 105-110.
- [3] Yudha A.F., Sarjon D., Yuhandri, 2017. Penerapan Algoritma C4.5 untuk Klasifikasi Data Rekam Medis berdasarkan International

Classification Diseases (ICD-10). *Jurnal Rekayasa Sistem dan Teknologi Informasi*, 1 (2), pp. 82-89.

- [4] Patel B.R., Rana K.K., 2012. A Survey on Decision Tree Algorithm for Classification. *International Journal of Engineering Development and Research*, 2 (1), pp. 1-5.
- [5] Selvia L.B.G., Zarman W., Hamidah I., 2014. Analisis dan Penerapan Algoritma C.45 Dalam Data Mining Untuk Memprediksi Masa Studi Mahasiswa Berdasarkan Data Nilai Akademik. *Prosiding Seminar Nasional Aplikasi Sains dan Teknologi*, pp. 263-272.
- [6] Ridwan M., Suyono H., Sarosa M., 2013. Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier. *Electrical Power, Electronics Control, Communication, and Informatics Seminar*, 7 (1), pp. 59-64.
- [7] Adhatrao K., Gaykar A., Dhawan A., Jha R., Honrao V., 2013. Predicting Students' Performance Using ID3 And C4.5 Classification Algorithms. *International Journal of Data Mining & Knowledge Management Process*, 3 (5), pp. 39-52.
- [8] Elmande Y., Widodo P., 2012. Pemilihan Criteria Splitting dalam Algoritma Iterative Dichotomiser 3 (ID3) untuk Penentuan Kualitas Beras Perum Bulog Drive Lampung. *Jurnal Telematika Mkom*, 4 (1), pp. 73-82.
- [9] Guterress J.A.D., Mudjihartono P., Ernawati., 2012. Analisis Efektivitas Algoritma C4.5 dalam Menentukan Peserta Pemenang Tender. *Projek Seminar Nasional Teknologi Informasi dan Komunikasi*, 3, pp. 7-12.
- [10] Luvia S.Y., Hartama D., Windarto A.P., Solikhun., 2016. Penerapan Algoritma C4.5 Untuk Klasifikasi Predikat Keberhasilan Mahasiswa Di Amik Tunas Bangsa. *Jurnal Riset Sistem Informasi dan Teknik Informatika*, 1 (1), pp. 75-59.