

Perbandingan Implementasi Ct Pro Dan Algoritma C45 Menentukan Pola Penyakit Pasien Stroke (Studi Kasus Rs. Siti Fatimah Palembang)

Abi Yansa¹⁾, Diana²⁾

^{*1,2)} Jl. Jenderal Ahmad Yani, Kecamatan Seberang Ulu I, Palembang, Sumatera Selatan

^{*1,2)} Teknik Informatika, Universitas Bina Darma

e-mail: ^{*1}abiyansa98@gmail.com¹, diana@binadarma.ac.id²

Abstrak

Perkembangan data *mining* dewasa sudah merambah ke setiap segmen, hal ini berpengaruh pada bidang kesehatan. Dengan perkembangan tersebut dapat membantu proses baik dalam pelayanan kesehatan dan pencegahan penyakit pada pasien. Penyakit *stroke* merupakan penyakit yang saat ini menjadi penyebab kematian dengan presentase 10,9% di Indonesia. Hal itu terjadi karena sulitnya pada proses prediksi gejala yang dialami. Dari permasalahan tersebut, dibutuhkan teknologi dengan pemanfaatan data *mining* untuk memklusterisasi serta klasifikasi dari hasil rekam medis pasien guna menanggulangi masalah pada penyakit *stroke*. Penelitian ini menggunakan teknik clustering dengan membandingkan antara algoritma CT-Pro dan algoritma C45. Hasil pada penelitian ini menunjukkan bahwa algoritma C45 menghasilkan akurasi serta hasil yang lebih baik dibandingkan dengan algoritma CT-Pro.

Kata Kunci— Data Mining, CT Pro, C45

Abstract

The development of adult data mining has penetrated into every segment, this has an effect on the health sector. With these developments can help the process both in health care and disease prevention in patients. Stroke is a disease that is currently the cause of death with a percentage of 10.9% in Indonesia. This happens because of the difficulty in predicting the symptoms experienced. From these problems, technology is needed with the use of data mining to cluster and classify the results of patient medical records in order to overcome the problem of stroke. This study uses a clustering technique by comparing the CT-Pro algorithm and the C45 algorithm. The results of this study indicate that the C45 algorithm produces better accuracy and results than the CT-Pro algorithm.

Keywords— Data Mining, CT Pro, C45

1. PENDAHULUAN

Dalam perkembangan teknologi masa kini, banyaknya data menjadi sebuah permasalahan sekaligus kesempatan bagi sebuah instansi. Data menjadi permasalahan apabila tidak dapat disimpan, dikelola, maupun diproses dengan baik. Data yang selalu bermunculan setiap waktu akan terus menumpuk dan bila tidak didokumentasikan dengan baik, maka data tersebut akan menjadi tidak berguna untuk perusahaan. Sedangkan data menjadi sebuah kesempatan apabila dapat disimpan, dikelola dan diproses menjadi lebih berarti untuk instansi tersebut. Dengan adanya data, maka dapat ditemukan sebuah trend maupun struktur yang nantinya dapat dipergunakan untuk mendapatkan informasi di masa mendatang.

Saat ini tumpukan data pasien yang ada di rumah sakit pada umumnya hanya dimanfaatkan pada pemenuhan laporan grafik pasien, belum dilakukan pengolahan data lebih lanjut sehingga masih kurangnya informasi yang bisa berguna bagi rumah sakit dan tim medis

terkait data rekam medis pasien. Tumpukan data yang ada belum menyajikan informasi pola khusus terutama terkait ancaman penyakit berdasarkan usia di kelompok masyarakat.

Salah satu aspek yang terdapat dalam laporan pemeriksaan pasien adalah rekam medis. Dimana rekam medis merupakan informasi identitas dan riwayat berobat oleh pasien di balai pengobatan rumah sakit dan puskesmas. Rekam medis adalah berkas yang berisi catatan dan dokumen antara lain identitas pasien, hasil pemeriksaan, pengobatan yang telah diberikan, serta tindakan dan pelayanan lain yang telah diberikan kepada pasien[1]. Catatan tersebut merupakan tulisan-tulisan yang dibuat oleh dokter atau dokter gigi mengenai tindakan-tindakan yang dilakukan kepada pasien dalam rangka pelayanan kesehatan. *Data mining* merupakan sebuah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan dan *machine learning* yang mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terakrit dari berbagai *database* yang besar[2].

Data mining bekerja dengan menemukan pola dan informasi dari data terpilih menggunakan metode tertentu[3]. *Data mining* biasa disebut *knowledge discovery in database* (KDD), merupakan proses yang meliputi pengumpulan, pemakaian data historis guna menemukan keteraturan serta pemakaian data historis untuk menemukan keteraturan dalam set data berukuran besar[4]. Keluaran pada *data mining* dapat dipakai untuk memperbaiki pengambilan keputusan dimasa mendatang[5].

Tujuan penelitian ini adalah analis dalam implementasikan perbandingan algoritma CT-PRO dan algoritma C45 dalam menentukan pola penyakit *stroke* pada pasien RS. Siti Fatimah Palembang. Dengan adanya penelitian ini diharapkan dapat mempermudah dalam menentukan pola penyakit *stroke* pada pasien RS. Siti Fatimah Palembang serta dapat memperkuat sistem pengelolaan data pasien. Dimana store ini sendiri adalah gangguan pada fungsi otak baik lokal maupun menyeluruh yang disebabkan oleh pasokan daran ke otak yang terganggu dan terjadi dengan cepat dan berlangsung lebih dari 24 jam atau berakhir dengan kematian[6].

Randi Estian Pambudi et al. melakukan penelitian terkait dengan Klasifikasi Penyakit Stroke Menggunakan Algoritma *Decision Tree* C.45[7]. Berdasarkan penelitian ini hasil yang di dapat adalah algoritma *decision tree* C4.5, menghasilkan tingkat akurasi diagnosis penyakit *stroke* yang sangat baik dan mendapat akurasi sebesar 99.07%.

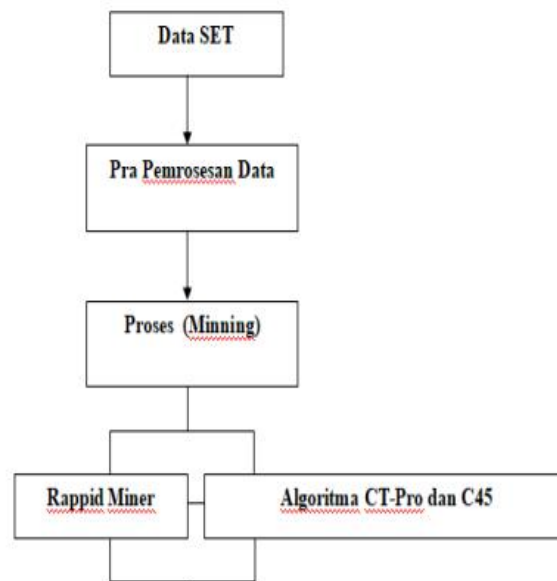
Firman Akbar et al. melakukan penelitian terkait dengan Implementasi Algoritma *Decision Tree* C45 dan *Support Vector Regression* untuk Prediksi Penyakit *Stroke*. Penelitian ini mendapatkan hasil dimana algoritma *Decission Tree* mendapatkan prediksi yang sangat bagus dalam prediksi serta nilai error yang cenderung lebih kecil pada data pasien penyakit *stroke* dan mendapat nilai RMSE 0.235, dan *Support Vector Regression* mendapat nilai dari pengukuran metrik RMSE sebesar 0.3999.

Dari penelitian diatas mendapatkan hasil dimana algoritma C45 mendapatkan hasil yang lebih baik dibandingkan algoritma yang lain. Oleh karena itu, penelitian ini dilakukan untuk mengkomparasi antara algoritma CT-Pro dan C45 pada studi kasus penyakit *stroke*. Implementasi dilakukan menggunakan *Rapid Miner*. Diharapkan dapat menemukan solusi terkait prediksi pola pada penyakit *stroke*.

2. METODE PENELITIAN

2.1 Desain Penelitian

Pada penelitian ini penulis menggunakan metode Proses pada penelitian ini yaitu mengimplementasikan Algoritma CT-PRO dan Algoritma C45 untuk menentukan pola penyakit. Karena permasalahan ini menentukan proses hasil perbandingan antara dua algoritma dengan data mining, selanjutnya akan dilakukan penerapan atau implementasi yang dapat menentukan hasil yang akan dapat digunakan pada mengimplementasikan perbandingan Algoritma CT-PRO dan Algoritma C45 pada pola penyakit menggunakan data *mining*. Untuk lebih jelasnya desain penelitian yang peneliti gunakan dapat dilihat pada Gambar 1.

Gambar 1. *Flowchart* penelitian

2.2 Pengumpulan Data

Apabila dilihat dari sumber datanya maka pengumpulan data dapat menggunakan sumber primer dan sumber sekunder. Sumber primer adalah sumber data yang langsung memberikan data kepada pengumpul data, dan sumber sekunder merupakan sumber yang tidak langsung memberikan data kepada pengumpul data misalnya lewat orang lain atau lewat dokumen[8]. Pada penelitian ini, data yang diperoleh adalah data sekunder karena diperoleh dari data pasien rumah sakit.” 51 “Pada penelitian ini prediksi dilakukan berdasarkan data-data yang sudah terjadi, maksudnya adalah data yang penulis gunakan berupa data rekam medis pasien pada tahun 2020. Jadi data yang akan diolah telah memiliki variabel tujuan yaitu pola penyakit pasien setiap tahunnya. Hal ini dimaksudkan agar dapat diketahui nilai akurasi hasil prediksi berdasarkan penerapan dari dua metode data mining yang digunakan. Penelitian ini sejalan dengan penelitian yang dilakukan oleh Pandi Ahmad, dimana pada penelitian tersebut menggunakan data rekam medis pasien pada tahun 2021 sebagai data kajian atau data set[9].

2.3 Algoritma CT-PRO

Algoritma CT-PRO merupakan salah satu algoritma pengembangan dari FP-Growth[10]. Perbedaannya terdapat pada langkah kedua dimana FP-Growth membuat FP-Tree sedangkan CT-PRO juga menggunakan Algoritma CT-PRO memiliki 3 tahap yakni :

1. Mencari data nasabah yang di input
2. Membuat struktur data CFP-Tree
3. Melakukan *mining*

2.4 Algoritma C45

Aloritma C4.5 diperkenalkan oleh Quinlan (1996) sebagai versi perbaikan dari ID3[11]. Dalam ID3 induksi decision tree hanya bisa dilakukan pada fitur bertipe kategorikal (nominal atau ordinal), sedangkan tipe numeric (interval atau rasio) tidak dapat digunakan. Perbaikan yang membedakan algoritma C4.5 dan ID3 adalah dapat menangani fitur dengan tipe numeric, melakukan pemotongan (*pruning*) decision tree, dan penurunan (*deriving*) rule set. Algoritma C4.5 juga menggunakan kriteria gain dalam menentukan fitur yang menjadi pemecah node pada pohon yang diinduksi. Ada beberapa tahapan dalam membangun sebuah pohon keputusan

dengan Algoritma C4.5 yaitu menyiapkan data training. Data training biasanya diambil dari data histori yang pernah terjadi sebelumnya dan sudah dikelompokkan ke dalam kelas kelas tertentu dan menentukan akar dari pohon.

Akar akan diambil dari atribut yang terpilih, dengan cara menghitung nilai gain dari masing-masing atribut, nilai gain yang paling tinggi yang akan menjadi akar pertama. Sebelum menghitung nilai gain dari atribut, hitung dahulu nilai entropy. Untuk menghitung nilai entropy digunakan rumus:

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

Keterangan:

S : himpunan kasus 12

A : atribut

n : jumlah partisi S

pi : proporsi dari Si terhadap S

3. Kemudian hitung nilai gain yang menggunakan rumus:

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad (2)$$

Keterangan:

S = himpunan kasus

A = fitur

n = jumlah partisi atribut

$A | S_i |$ = proporsi Si terhadap S

$| S |$ = jumlah kasus dalam S

4. Ulangi langkah ke-2 hingga semua record terpartisi.

5. Proses partisi pohon keputusan akan berhenti saat:

- a. Semua record dalam simpul N mendapat kelas yang sama.
- b. Tidak ada atribut di dalam record yang dipartisi lagi.
- c. Tidak ada record di dalam cabang yang kosong.

3. HASIL DAN PEMBAHASAN

3.1 Dataset

Jumlah data awal yang diperoleh dari pengumpulan data yaitu 309 data alat kesehatan, namun tidak semua data dapat digunakan dan tidak semua atribut digunakan karena harus melalui beberapa tahap pengolahan data awal (pra pemrosesan data). Sebuah data dikatakan tidak bersih jika mengandung data yang berupa nilai kosong atau derau dan pencilan inkonsistensi. Data yang masih mentah harus dilakukan proses pembersihan dengan cara mengisi nilai-nilai yang kosong menghaluskan data yang berderau dan membuang pencilan atau memperbaiki inkonsistensi. Sehingga setelah dilakukan pra pemrosesan, data yang digunakan pada penelitian ini hanya 309 data.

3.2 Pengolahan Data

Data yang digunakan dalam penelitian ini adalah sumber data rekam medis pasien. Data diperoleh dari arsip yang dimiliki oleh pihak rumah sakit sebanyak 309 data. Data ini berisi mengenai jenis kelamin pasien, usia pasien, kategori penyakit, bulan masuk pasien, status

pasien dan disposisi. Tabel berisi data yang 53 akan diolah di *rapidminer*. Data sampel yang diuji dalam penelitian ini ialah 309 data. Format seperti berikut :

Tabel 1 Contoh Data Sampel

No.	Umur	Jenis Kelamin	Disposisi	Status	Tensi Darah	Gula Darah	Stroke
1	36	Laki-laki	Rawat Jalan	Umum	Tinggi	110	Ringan
2	51	Laki-laki	Rawat Jalan	BPJS	Rendah	115	Berat
3	19	Perempuan	Rawat Jalan	BPJS	Tinggi	120	Berat
4	24	Perempuan	Rawat Inap	BPJS	Rendah	125	Ringan
5	32	Perempuan	Rawat Jalan	Umum	Tinggi	135	Ringan
6	27	Laki-laki	Rawat Jalan	Umum	Rendah	120	Ringan
7	56	Perempuan	Rawat Jalan	Umum	Rendah	110	Berat
8	20	Laki-laki	Rawat Jalan	BPJS	Rendah	110	Berat
9	31	Perempuan	Rawat Inap	BPJS	Tinggi	115	Ringan
10	22	Perempuan	Rujuk	Umum	Rendah	120	Ringan
11	3	Laki-laki	Rawat Jalan	BPJS	Tinggi	125	Ringan
12	2	Laki-laki	Rawat Jalan	BPJS	Rendah	135	Berat
13	45	Laki-laki	Rawat Inap	Umum	Tinggi	120	Berat
14	23	Perempuan	Rawat Jalan	Umum	Rendah	110	Ringan

Sampel diatas akan didistribusikan menjadi tabulasi menjadi binomial agar dapat diolah berdasarkan metode C45. Pada tabel berikut ini akan menampilkan sampel tabulasi yang digunakan :

Tabel 2 Sampel Tabulasi

No.	Umur	Jenis Kelamin	Disposisi	Status	Tensi Darah	Gula Darah	Stroke
1	36	Laki-laki	Rawat Jalan	Umum	Tinggi	110	Ringan
2	51	Laki-laki	Rawat Jalan	BPJS	Rendah	115	Berat
3	19	Perempuan	Rawat Jalan	BPJS	Tinggi	120	Berat
4	24	Perempuan	Rawat Inap	BPJS	Rendah	125	Ringan
5	32	Perempuan	Rawat Jalan	Umum	Tinggi	135	Ringan
6	27	Laki-laki	Rawat Jalan	Umum	Rendah	120	Ringan
7	56	Perempuan	Rawat Jalan	Umum	Rendah	110	Berat
8	20	Laki-laki	Rawat Jalan	BPJS	Rendah	110	Berat
9	31	Perempuan	Rawat Inap	BPJS	Tinggi	115	Ringan
10	22	Perempuan	Rujuk	Umum	Rendah	120	Ringan

11	3	Laki-laki	Rawat Jalan	BPJS	Tinggi	125	Ringan
12	2	Laki-laki	Rawat Jalan	BPJS	Rendah	135	Berat
13	45	Laki-laki	Rawat Inap	Umum	Tinggi	120	Berat
14	23	Perempuan	Rawat Jalan	Umum	Rendah	110	Ringan

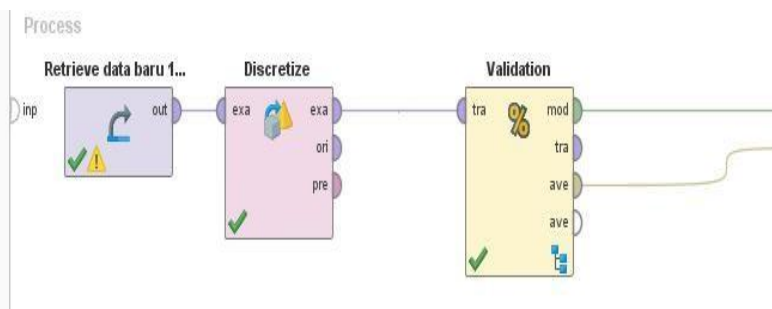
Sampel diatas akan didistribusikan menjadi tabulasi bilangan biner agar dapat diolah berdasarkan metode FP-Grwoth. Tabulasi FP-Grwoth menampilkan bilangan 1 dan 0. Dimana bilangan 1 itu dapat diartikan dengan ya atau positif sedangkan 0 itu tidak atau negatif. Pada tabel berikut ini akan menampilkan sampel tabulasi yang digunakan :

Tabel 3 Sampel Tabulasi FP-Grwoth

Dewasa	Tua	Tensi Rendah	Tensi Tinggi	Gula Darah Rendah	Gula Darah Tinggi	BPJS	Umum	Rawat Jalan	Rawat Inap	Rujukan	Stroke Ringan
1	0	0	0	0	0	1	0	0	0	1	0
0	0	0	0	0	0	0	0	1	0	0	1
0	0	0	0	0	0	1	0	0	0	0	1
1	0	0	0	0	1	0	0	0	0	0	1
1	0	0	0	1	0	0	0	0	0	1	0
1	0	0	0	1	0	0	0	0	0	1	0
1	0	1	0	0	0	0	0	0	0	1	0
1	0	0	0	0	0	1	0	0	0	0	1
1	0	0	0	0	1	0	0	0	0	0	1
1	0	0	0	1	0	0	0	0	0	1	0

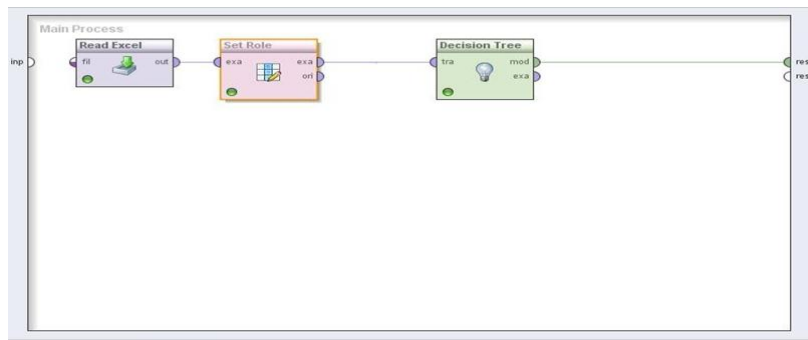
3.3 Implementasi Metode C4.5 dengan RapidMiner

Penelitian ini di lakukan untuk mengetahui klasifikasi pola nasabah. Penggunaan parameter berpengaruh pada hasil akurasi dan model yang akan dihasilkan oleh algoritma C4.5 dan CT-Pro. Untuk menentukan akurasi dari algoritma, penelitian ini menggunakan *Cross Validation*. Desain model C4.5 yang terdapat pada *RapidMiner* yang ditunjukkan pada gambar 2



Gambar 2. Cross Validation

Pada objek *Discretize attribute* yang dipilih yaitu *subnet* yang memiliki nilai *continuous* saja seperti label lancar atau macet. Pada objek atau operator *Cross Validation* terdapat sub proses yang terdiri dari operator *Decision Tree* sebagai algoritma yang digunakan pada proses data *mining Classification* yang akan dilakukan, operator *Apply Model* dan operator *Performance* sebagai operator untuk menghasilkan pengolahan data mining berupa *Decision Tree*. Adapun model sub proses dapat di lihat pada Gambar 3.

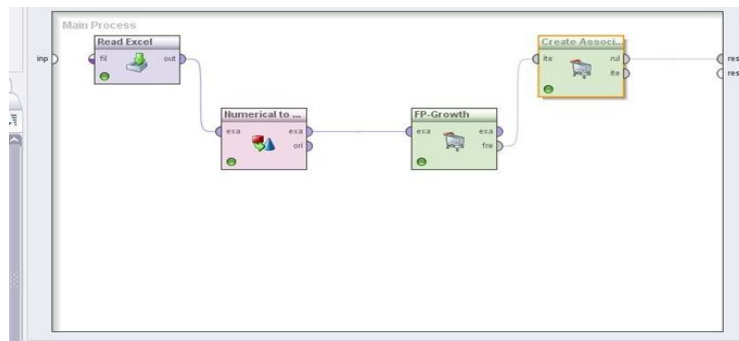


Gambar 3. Rapidminer Model Sub Proses Clasification Decision Tree

Salah satu tujuan penelitian ini adalah untuk mengetahui nilai akurasi dari algoritma Decision tree yang digunakan untuk mengklasifikasi Label. Didalam kolom training terdapat algoritma klasifikasi yang diterapkan yaitu decision tree, sedangkan didalam kolom set role terdapat pengauran lbel yang digunakan untuk menjalankan model model decision tree dan Performance untuk mengukur performa dari model decision Tree Tersebut.

3.4 Implementasi Metode CT-Pro dengan RapidMiner

Pada implementasi dengan algoritma CT-Pro, adapun tampilan dari struktur model data mining yang di gunakan terdiri dari beberapa objek atau operator antara lain read file XLSX, selected attribute dan Cross Validation yang tampil pada proses utama untuk menjalankan proses data mining. Untuk menentukan akurasi dari algoritma, penelitian ini menggunakan Assosiation Desain model algoritma CT-Pro yang terdapat pada Rapid Miner.

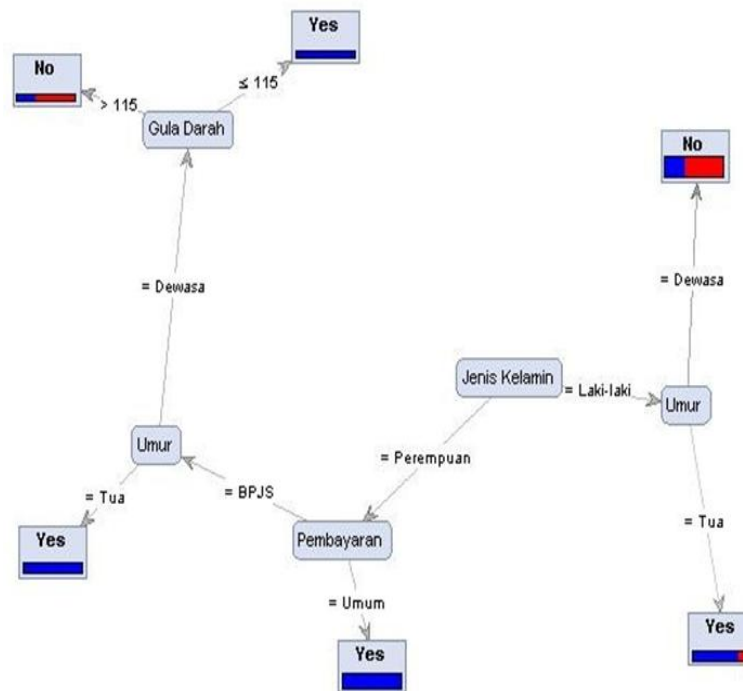


Gambar 4. Rapidminer Model Sub Proses KlasifikasiCT-Pro

Pada objek atau operator Cross Validation terdapat sub proses yang terdiri dari operator algoritma CT-Pro sebagai algoritma yang digunakan pada proses data *mining Classification* yang akan dilakukan, operator FP Grwoth dan operator *Assosiation* sebagai operator untuk menghasilkan pengolahan data *mining* berupa algoritma CT-Pro.

3.5 Hasil Implementasi Klasifikasi Menggunakan Metode C4.5

Implementasi proses pengolahan data menggunakan diskritisasi data yang digunakan untuk mengurangi jumlah nilai *atribute numeric* dengan cara membagi *range* atribut ke dalam interval. Label interval kemudian dapat digunakan untuk menggantikan nilai-nilai aktual data. Hasil pemodelan yang telah di proses oleh *Tool RapidMiner* selain menghasilkan dalam bentuk pola pemodelan ini juga dapat mengetahui tingkat akurasi, *recall*, dan *precision*. Hasil pohon keputusan dalam implementasi C4.5 yang dapat dilihat pada gambar 5 sebagai berikut :



Gambar 5. Pohon Keputusan Algoritma C45

3.6 Implementasi Klasifikasi Menggunakan Metode Algoritma CT-Pro Pada Rapid Miner

Hasil permodelan yang telah di proses oleh *Tool RapidMiner* dengan metode algoritma CT-Pro dapat diketahui nilai akurasi data tersebut seperti terlihat pada gambar 6 di bawah ini.

No.	Premises	Conclusion	Support	Confidence	LaPlace	Gain	p-s	Lift	Conviction
1	Umum	BPJS	0.538	0.913	0.968	-0.641	0.221	1.696	5.308
22	Tua	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞
28	Perempuan	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞
32	Laki-laki	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞
34	Gula Darah Tinggi	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞
67	Tensi Tinggi, Tua	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞
82	Tensi Tinggi, Perempuan	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞
92	Tensi Tinggi, Laki-laki	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞
97	Tensi Tinggi, Gula Darah Tinggi	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞
117	Tensi Rendah, Tua	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞
132	Tensi Rendah, Perempuan	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞
142	Tensi Rendah, Laki-laki	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞
147	Tensi Rendah, Gula Darah Tinggi	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞
165	Tua, Perempuan	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞
177	Tua, Laki-laki	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞
183	Tua, Gula Darah Tinggi	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞
195	Perempuan, Laki-laki	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞
201	Perempuan, Gula Darah Tinggi	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞
207	Laki-laki, Gula Darah Tinggi	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞
247	Tensi Tinggi, Tensi Rendah, Tua	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞
280	Tensi Tinggi, Tensi Rendah, Perempuan	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞
302	Tensi Tinggi, Tensi Rendah, Laki-laki	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞
313	Tensi Tinggi, Tensi Rendah, Gula Darah Tinggi	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞
352	Tensi Tinggi, Tua, Perempuan	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞
378	Tensi Tinggi, Tua, Laki-laki	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞
391	Tensi Tinggi, Tua, Gula Darah Tinggi	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞
417	Tensi Tinggi, Perempuan, Laki-laki	Dewasa	0.410	1	1	-0.410	0.242	2.438	∞

Gambar 6. Hasil Algoritma CTPro

3.7 Hasil Association Rules Algoritma CT-Pro

Adapun hasil *association rules*, dapat dilihat sebagai berikut :

```

AssociationRules

Association Rules
[Umum] --> [BPJS] (confidence: 0.913)
[Tensi Tinggi] --> [Tensi Rendah] (confidence: 0.913)
[BPJS] --> [Umum] (confidence: 1.000)
[Tensi Rendah] --> [Tensi Tinggi] (confidence: 1.000)
[Tua] --> [Tensi Tinggi] (confidence: 1.000)
[Perempuan] --> [Tensi Tinggi] (confidence: 1.000)
[Laki-laki] --> [Tensi Tinggi] (confidence: 1.000)
[Gula Darah Tinggi] --> [Tensi Tinggi] (confidence: 1.000)
[Dewasa] --> [Tensi Tinggi] (confidence: 1.000)
[Rujukan] --> [Rasat Jalan] (confidence: 1.000)
[Tua] --> [Tensi Rendah] (confidence: 1.000)
[Perempuan] --> [Tensi Rendah] (confidence: 1.000)
[Laki-laki] --> [Tensi Rendah] (confidence: 1.000)
[Gula Darah Tinggi] --> [Tensi Rendah] (confidence: 1.000)
[Dewasa] --> [Tensi Rendah] (confidence: 1.000)
[Tua] --> [Perempuan] (confidence: 1.000)
[Perempuan] --> [Tua] (confidence: 1.000)
[Tua] --> [Laki-laki] (confidence: 1.000)
[Laki-laki] --> [Tua] (confidence: 1.000)
[Tua] --> [Gula Darah Tinggi] (confidence: 1.000)
[Gula Darah Tinggi] --> [Tua] (confidence: 1.000)
[Tua] --> [Dewasa] (confidence: 1.000)
[Dewasa] --> [Tua] (confidence: 1.000)
[Perempuan] --> [Laki-laki] (confidence: 1.000)
[Perempuan] --> [Tensi Tinggi, Gula Darah Tinggi] (confidence: 1.000)
[Tensi Tinggi, Perempuan] --> [Gula Darah Tinggi] (confidence: 1.000)
[Gula Darah Tinggi] --> [Tensi Tinggi, Perempuan] (confidence: 1.000)
[Tensi Tinggi, Gula Darah Tinggi] --> [Perempuan] (confidence: 1.000)
[Perempuan, Gula Darah Tinggi] --> [Tensi Tinggi] (confidence: 1.000)
[Perempuan] --> [Tensi Tinggi, Dewasa] (confidence: 1.000)
[Tensi Tinggi, Perempuan] --> [Dewasa] (confidence: 1.000)
[Dewasa] --> [Tensi Tinggi, Perempuan] (confidence: 1.000)
[Tensi Tinggi, Dewasa] --> [Perempuan] (confidence: 1.000)
[Perempuan, Dewasa] --> [Tensi Tinggi] (confidence: 1.000)
[Laki-laki] --> [Tensi Tinggi, Gula Darah Tinggi] (confidence: 1.000)
[Tensi Tinggi, Laki-laki] --> [Gula Darah Tinggi] (confidence: 1.000)
[Gula Darah Tinggi] --> [Tensi Tinggi, Laki-laki] (confidence: 1.000)
[Tensi Tinggi, Gula Darah Tinggi] --> [Laki-laki] (confidence: 1.000)
[Laki-laki, Gula Darah Tinggi] --> [Tensi Tinggi] (confidence: 1.000)
[Laki-laki] --> [Tensi Tinggi, Dewasa] (confidence: 1.000)
[Tensi Tinggi, Laki-laki] --> [Dewasa] (confidence: 1.000)
[Dewasa] --> [Tensi Tinggi, Laki-laki] (confidence: 1.000)
[Tensi Tinggi, Dewasa] --> [Laki-laki] (confidence: 1.000)
[Laki-laki, Dewasa] --> [Tensi Tinggi] (confidence: 1.000)
[Gula Darah Tinggi] --> [Tensi Tinggi, Dewasa] (confidence: 1.000)
[Tensi Tinggi, Gula Darah Tinggi] --> [Dewasa] (confidence: 1.000)
[Dewasa] --> [Tensi Tinggi, Gula Darah Tinggi] (confidence: 1.000)
[Tensi Tinggi, Dewasa] --> [Gula Darah Tinggi] (confidence: 1.000)
[Gula Darah Tinggi, Dewasa] --> [Tensi Tinggi] (confidence: 1.000)
[Tua] --> [Tensi Rendah, Perempuan] (confidence: 1.000)
... 1680 other rules ...

```

Gambar 7. Hasil *Association Rules*

4. KESIMPULAN

Pada penelitian ini dilakukan pemodelan menggunakan algoritma Apriori dan FP-Growth dengan menggunakan data yang sudah diolah menggunakan metode KDD *Knowledge Discovery in Database*. Dari hasil observasi terhadap dataset pasien stroke RSUD Siti Fatimah Palembang melalui proses menggunakan metode klasifikasi dan algoritma C45 dan CTPRO dengan tools Rapid Miner. Sehingga, diperoleh kesimpulan bahwa metode C45 dapat dikatakan lebih baik dari pada metode CT-Pro. Hal tersebut dikarenakan metode C45 memiliki pohon keputusan yang jelas dan mudah dimengerti sedangkan pada metode CT-PRO menampilkan begitu banyak asosiasi.

5. SARAN

Untuk penelitian selanjutnya perlu dilakukan perhitungan yang tepat dan baik dengan menggunakan tools selain rapid minner.

DAFTAR PUSTAKA

- [1] W. Handiwidjojo, "Rekam medis elektronik," *Jurnal Eksplorasi Karya Sistem Informasi dan Sains*, vol. 2, no. 1, 2015.
- [2] C. Anam and H. B. Santoso, "Perbandingan kinerja algoritma c4. 5 dan naive bayes untuk klasifikasi penerima beasiswa," *Energy-Jurnal Ilmiah Ilmu-Ilmu Teknik*, vol. 8, no. 1, pp. 13–19, 2018.
- [3] R. Amanda and E. S. Negara, "Analysis and implementation machine learning for youtube data classification by comparing the performance of classification algorithms," *Jurnal Online Informatika*, vol. 5, no. 1, pp. 61–72, 2020.
- [4] E. Buulolo, "Implementasi algoritma apriori pada sistem persediaan obat (studi kasus: Apotik rumah sakit estomih medan)," *Pelita Inform. Budi Dharma*, vol. 4, pp. 71–83, 2013.
- [5] B. Santosa, "Data mining teknik pemanfaatan data untuk keperluan bisnis," *Yogyakarta: Graha Ilmu*, vol. 978, no. 979, p. 756, 2007.
- [6] R. S. Rohman, R. A. Saputra, and D. A. Firmansah, "Komparasi algoritma c4. 5 berbasis pso dan ga untuk diagnosa penyakit stroke," 2020.
- [7] R. E. Pambudi, S. Sriyanto, and F. Firmansyah, "Klasifikasi penyakit stroke menggunakan algoritma decision tree c. 45," *TEKNIKA*, vol. 16, no. 2, pp. 221–226, 2022.
- [8] F. Silkyanti, "Analisis peran budaya sekolah yang religius dalam pembentukan karakter siswa," *Indonesian Values and Character Education Journal*, vol. 2, no. 1, pp. 36–42, 2019.
- [9] P. Ahmad, "Perancangan aplikasi prediksi jumlah persediaan alat kesehatan dengan metode anfis (studi kasus: Rsud deli serdang)," *Informasi dan Teknologi Ilmiah (INTI)*, vol. 4, no. 2, 2017.
- [10] R. Putri *et al.*, "Implementasi algoritma ct-pro untuk menemukan pola pada data siswa sma (studi kasus: Madrasah aliyah negeri (man) karanganon klaten)," *Panduan untuk belajar jurnal*, vol. 1, no. 1, 2017.
- [11] D. Merawati and R. Rino, "Penerapan data mining penentu minat dan bakat siswa smk dengan metode c4. 5," *ALGOR*, vol. 1, no. 1, pp. 28–37, 2019.