

Penerapan Data Mining Dalam Menentukan Persediaan Obat Dengan Algoritma C4.5

R. Mahdalena Simanjorang¹, Agustina Simangunsong²

¹Manajemen Informatika, STMIK Pelita Nusantara, Medan, Indonesia

²Teknik Informatika, STMIK Pelita Nusantara, Medan, Indonesia

Email: ¹relimamahdalenasimanjorang@yahoo.co.id, ²agustinasimangunsong93@gmail.com

Email Penulis Korespondensi: emailpenuliskorespondensi@email.com

Article History:

Received Jun 12th, 202x

Revised Aug 20th, 202x

Accepted Aug 26th, 202x

Abstrak

Peranan Data Mining saat ini sangat dibutuhkan dalam berbagai bidang. Tanpa data mining maka akan terjadi banyak penumpukan data seperti contohnya di klinik Romana. Adapun permasalahan yang terjadi adalah saat ini dalam menentukan persediaan obat sering kali terjadi kendala. Persediaan dibutuhkan karena tingkat permintaan yang tidak terorganisir dengan baik dan sering terjadinya stockout (kehabisan persediaan) sehingga mengakibatkan lambatnya pelayanan kepada pasien. Tidak tersedianya stok obat disebabkan kemungkinan karena stock habis dan belum melakukan permintaan ke Gudang. Maka diperlukan sebuah algoritma yang dapat membantu dalam menentukan pola persediaan obat di klinik tersebut yaitu dengan algoritma C4.5. Dalam penelitian ini data yang digunakan yaitu dilihat dari pola data transaksi setiap hari. Data yang digunakan nantinya adalah data transaksi 3 bulan sebelumnya. Tujuan diadakan nya penelitian ini adalah agar dapat membantu mempermudah dalam menentukan pola persediaan stok obat yang dapat memudahkan pelayanan terhadap pasien dalam meresep obat dan mampu mengatasi terjadinya penimbunan stok yang berlebih dengan menggunakan teknik data mining dengan algoritma C4.5. Dengan adanya algoritma yang digunakan maka data yang didapatkan dapat diolah sehingga menjadi sebuah informasi yang berguna dan dapat membentuk sebuah pola untuk memnentukan stok obat yang diperlukan.

Kata Kunci : Algoritma C4.5, Data Mining, Stok Obat

Abstract

The role of Data Mining is currently needed in various fields. Without data mining, there will be a lot of data accumulation, for example at the Romana clinic. The problem that occurs is currently in determining drug supplies there are often obstacles. Supplies are needed because demand levels are not well organized and stockouts often occur, resulting in slow service to patients. The unavailability of drug stock is due to the possibility that the stock has run out and has not made a request to the Warehouse. So we need an algorithm that can help determine the pattern of drug supply in the clinic, namely the C4.5 algorithm. In this study the data used is seen from the pattern of transaction data every day. The data used later is transaction data from the previous 3 months. The purpose of this research is to help make it easier to determine drug stock inventory patterns that can facilitate services to patients in prescribing drugs and be able to overcome the occurrence of excess stock hoarding using data mining techniques with the C4.5 algorithm. With the algorithm used, the data obtained can be processed so that it becomes useful information and can form a pattern to determine the required drug stock.

Keyword : Data Mining, Drug Stock, C4.5 Algorithm

1. PENDAHULUAN

Perkembangan Teknologi informasi saat ini mempunyai peranan yang sangat penting dalam perkembangan ilmu pengetahuan termasuk juga aspek kehidupan, aspek perekonomian terutama dalam sistem persediaan obat. Segala hal menjadi tanpa batas akibat dari perkembangan teknologi tersebut sama halnya dengan penggunaan aplikasi database (tempat penyimpanan data). Dalam pengembangan sebuah sistem dibutuhkan aplikasi database yang berguna nantinya untuk melakukan penyimpanan banyak data data penting. Dan saat ini sudah sangat banyak media penyimpanan data dan semakin banyak data yang disimpan maka semakin besar pula ruang penyimpanan yang diperlukan Dampak dari penyimpanan data yang banyak ini maka secara tidak langsung akan terjadi penumpukan data dan tanpa disadari data yang banyak itu menyimpan suatu informasi yang bisa dimanfaatkan untuk dibuat suatu aplikasi untuk mengolah data yang ada untuk menghasilkan informasi yang berguna yang dinamakan dengan Data Mining.

Data mining adalah proses menemukan hubungan baru yang mempunyai arti, pola dan kebiasaan dengan memilah-milah sebagian besar data yang disimpan dalam media penyimpanan dengan menggunakan teknologi pengenalan pola seperti teknik statistik dan matematika. Data mining merupakan gabungan dari beberapa disiplin ilmu yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, database, dan visualisasi untuk penanganan permasalahan pengambilan informasi dari database yang besar.

Klinik Romana merupakan sebuah klinik yang berada di Jl. Besar Tanjung Anom. Klinik Romana merupakan sarana pelayanan dibidang kesehatan yang melayani masyarakat untuk melakukan pemeriksaan kehamilan, imunisasi, program keluarga berencana (KB), persalinan, rawat inap, dan pengobatan umum. Pasien yang sudah menyelesaikan proses pemeriksaan dan pengobatan di Klinik Romana akan memiliki Rekam Medis, dimana Rekam Medis berisikan catatan dan dokumen tentang identitas pasien, hasil pemeriksaan, pengobatan, tindakan dan pelayanan medis yang akan diberikan kepada pasien dan menjadi acuan bidan atau dokter apabila pasien melakukan pemeriksaan selanjutnya di klinik romana atau fasilitas kesehatan lainnya.

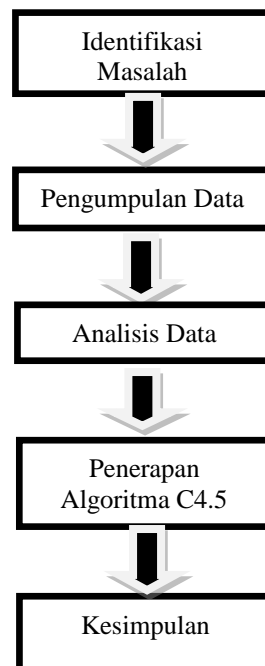
Adapun permasalahan yang terjadi adalah saat ini dalam menentukan persediaan obat sering kali terjadi kendala. Persediaan dibutuhkan karena tingkat permintaan yang tidak terorganisir dengan baik. Masalah yang sering timbul adalah terjadinya stockout (kehabisan persediaan) sehingga mengakibatkan lambatnya pelayanan kepada pasien. Tidak tersedianya stok obat disebabkan kemungkinan karena stock habis dan belum melakukan permintaan ke Gudang. Maka diperlukan sebuah algoritma yang dapat membantu dalam menentukan pola persediaan obat di klinik tersebut yaitu dengan algoritma C4.5.

Dalam penelitian ini data yang digunakan yaitu dilihat dari pola data transaksi setiap hari. Data yang digunakan nantinya adalah data transaksi 3 bulan sebelumnya. Tujuan diadakan nya penelitian ini adalah agar dapat membantu mempermudah dalam menentukan pola persediaan stok obat yang dapat memudahkan pelayanan terhadap pasien dalam meresep obat dan mampu mengatasi terjadinya penimbunan stok yang berlebih dengan menggunakan teknik data mining dengan algoritma C4.5. Dengan adanya algoritma yang digunakan maka data yang didapatkan dapat diolah sehingga menjadi sebuah informasi yang berguna dan dapat membentuk sebuah pola untuk menentukan stok obat yang diperlukan.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Tahapan-tahapan penelitian menguraikan kerangka kerja penelitian atau tahap-tahap yang dilakukan dalam melaksanakan penelitian. Tahapan penelitian ini dilakukan agar dapat menyelesaikan masalah yang akan dibahas. Hal ini dilakukan dengan tujuan untuk mengumpulkan informasi-informasi berkaitan dengan masalah yang akan diteliti. Sehingga dapat diketahui keadaan atau kedudukan masalah tersebut baik secara teoritis maupun praktis. Pengetahuan yang diperoleh dari studi pendahuluan sangat berguna untuk menyusun kerangka teoritis tentang pemecahan masalah dalam bentuk hipotesis yang akan diuji kebenarannya melalui pelaksanaan penelitian lapangan. Studi pendahuluan dapat dilakukan dengan studi kelayakan, kepustakaan dan studi lapangan. Untuk menyelesaikan penelitian ini maka digambarkan tahapan-tahapan penelitian yang dilakukan seperti pada gambar 1 berikut ini:



Gambar 1. Tahapan Penelitian

2.3 Uraian Tahapan Penelitian

Tahapan penelitian diatas akan diuraikan seperti berikut ini.

1. Identifikasi Masalah
Pada tahap ini hal yang dilakukan yaitu tahap mengidentifikasi masalah serta mempelajari permasalahan yang akan diselesaikan.
2. Pengumpulan Data
Pada tahap ini, peneliti mengumpulkan data yang nantinya akan dijadikan acuan untuk menyelesaikan permasalahan yang ada. Teknik pengumpulan data dilakukan sebagai berikut:
 - a. Observasi
Observasi adalah kegiatan melakukan pengamatan di klinik romana terkait dengan penentuan persediaan obat.
 - b. Studi Dokumentasi
Studi dokumentasi adalah pengumpulan data sekunder yang dibutuhkan untuk penelitian yang terdiri dari tahap analisa sampai pada proses penentuan persediaan obat.
3. Analisis Data
Analisis data merupakan kegiatan yang dilakukan untuk mengubah data hasil dari penelitian menjadi informasi yang nantinya bisa digunakan dalam mengambil kesimpulan. Analisis ini dilakukan untuk menjelaskan sebab-sebab dalam fakta-fakta sosial yang terukut, menunjukkan hubungan variabel serta melakukan analisa dari data yang sudah didapatkan.
4. Penerapan Algoritma C4.5
Pada tahapan ini metode yang sudah ditentukan nantinya akan diimplementasikan dengan menerapkan semua langkah-langkah metode Algoritma C4.5 yaitu mencari nilai nilai enropy dari setiap kriteria, mencari nilai gain dari setiap atribut, pembentukan atribut sebagai akar berdasarkan gain tertinggi, pembentukan cabang berdasarkan masing masing nilai, ulangi proses untuk masing masing cabang Sehingga hasil akhir nantinya akan dijadikan acuan untuk mengambil sebuah kesimpulan tentang penentuan persediaan obat pada klinik tersebut.
5. Kesimpulan
Pada tahap ini, peneliti membuat sebuah kesimpulan terhadap hasil penelitian yang dilakukan berdasarkan penerapan algoritma C4.5. Kesimpulan yang dibuat merupakan hasil perhitungan yang didapatkan dengan menerapkan metodologi penelitian.

2.4 Data Mining

Data Mining merupakan Penambangan data atau data mining berada di antara cabang ilmu komputer dan statistik, dengan memanfaatkan kemajuan dari kedua disiplin ilmu tersebut untuk mengekstraksi informasi dari besar dari database. Data mining merupakan serangkaian proses untuk menggali nilai tambah berupapengetahuan yang selama ini belum diketahui dari suatu kumpulan data. Data mining adalah proses yang menggunakan teknik statistik,

matematika, kecerdasan buatan, dan machine learning untuk mengekstrasi dan mengidentifikasi informasi yang bermanfaat serta pengetahuan yang terkait dari berbagai database besar. [5]

Data mining adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran computer (machine learning) untuk menganalisis dan mengekstrasi pengetahuan (knowledge) secara otomatis. Definisi lain diantaranya adalah pembelajaran berbasis induksi (induction-based learning) adalah proses pembentukan definisi-definisi konsep umum yang dilakukan dengan cara mengobservasi contoh-contoh spesifik dari konsep-konsep yang akan dipelajari. Knowledge Discovery in Databases (KDD) adalah penerapan metode saintifik pada data mining. Dalam konteks ini data mining merupakan satu langkah dari proses KDD. [4]

Beberapa teknik dan sifat data mining adalah sebagai berikut:

- a. Classification [Predictive]
- b. Clustering [Descriptive]
- c. Association Rule Discovery [Descriptive]
- d. Regression [Predictive]
- e. DeviationDetection [Predictive]

2.5 Pohon Keputusan

Decision Tree memakai kinerja dalam bentuk pohon (tree). Dalam pohon keputusan, masing-masing node yang dimiliki menjelaskan atribut yang ada, cabangnya menafsirkan nilai pada atribut yang ada dan leaf mempresentasikan kelas. Pada pohon keputusan ada yang dikenal sebagai root. Root adalah node yang berada paling atas dari pohon keputusan. Decision Tree adalah metode pengklasifikasian data yang paling dikenal untuk digunakan pada data mining. Pohon keputusan menggunakan pembangunan yang relatif cepa [8]. Output dari jenis model yang dibuat mudah untuk dimengerti. Pohon keputusan terbagi 3 jenis node, yaitu sebagai berikut :

- a. Root Node, merupakan node yang berada paling atas. Pada node ini tidak memiliki input dan dapat juga tidak memiliki output atau memiliki output yang lebih dari satu.
- b. Internal Node, merupakan node yang bercabang. Pada node ini hanya memiliki 1 buah input dan memiliki output yaitu minimal 2.
- c. Leaf node atau terminal node, merupakan node yang diakhir. Pada node ini hanya memiliki 1 input dan tidak memiliki output.

2.6 Algoritma C4.5

Algoritma C4.5 merupakan salah satu algoritma yang digunakan untuk membentuk pohon keputusan. Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan aturan. Aturan dapat dengan mudah dipahami dengan bahasa alami. Dan mereka juga dapat diekspresikan dalam bentuk bahasa basis data seperti Structured Query Language untuk mencari record pada kategori tertentu. Algoritma Pohon Keputusan C4.5 atau Classification version 4.5 adalah pengembangan dari algoritma ID3. Oleh karena pengembangan tersebut, algoritma C4.5 mempunyai prinsip dasar kerja yang sama dengan algoritma ID3. Secara umum, proses algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut. 1. Pilih atribut sebagai 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 [4]. Secara khusus, algoritma C4.5 Decision Tree menggunakan kriteria split yang telah dimodifikasi yang dinamakan Gain Ratio dalam proses pemilihan split atribut. Split atribut merupakan proses utama dalam pembentukan pohon keputusan (Decision Tree) di C4.5 [7]

Tahapan dari algoritma C4.5 adalah sebagai berikut. a. Menghitung nilai Entropy, b. Menghitung nilai Gain Ratio untuk masing-masing atribut, c. Atribut yang memiliki Gain Ratio tertinggi dipilih menjadi akar (root) dan atribut yang memiliki nilai Gain Ratio lebih rendah dari akar (root) dipilih menjadi cabang (branches), d. Menghitung lagi nilai Gain Ratio tiap-tiap atribut dengan tidak mengikutsertakan atribut yang terpilih menjadi akar (root) di tahap sebelumnya, e. Atribut yang memiliki Gain Ratio tertinggi dipilih menjadi cabang (branches), f. Mengulangi langkah ke-4 dan ke-5 sampai dengan dihasilkan nilai Gain = 0 untuk semua atribut yang tersisa. Untuk menghitung nilai Entropy dapat dihitung dengan persamaan :

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

Dimana :

S : Himpunan Kasus

N : Himpunan Kasus

P : Jumlah partisi N

Pi: Proporsi dari Ni terhadap N

Sementara itu nilai information gain (Gain) dapat dihitung menggunakan persamaan

$$Gain(S,A) = Entropy(S) - \sum_{i=1}^n \binom{S_i}{S} x Entropy(S_i) \dots \dots \dots (2)$$

dimana:

S = himpunan kasus

A = atribut

n = jumlah partisi atribut A

|Si| = jumlah kasus pada partisi ke-i

Selanjutnya nilai Split Info dapat dihitung dengan persamaan :

$$\text{SplitInfo}(S,A) = \text{Entropy}(S) - \sum_{i=1}^n \left(\frac{S_i}{S}\right) \log_2 \frac{S_i}{S} \text{Entropy}(S_i) \dots \dots \dots (3)$$

dimana:

S = himpunan kasus

A = atribut.

Si = jumlah sampel untuk atribut i

Maka nilai Gain Ratio yang menentukan sebuah atribut dapat dijadikan akar maupun cabang suatu pohon keputusan dapat dihitung dengan persamaan:

$$\text{GainRatio}(S,A) = \frac{\text{Gain}(S,A)}{\text{SplitInfo}(S,A)} \dots \dots \dots (4)$$

Setelah selesai lalu ulangi langkah ke 2 dan ke 3 sampai semua atribut dihitung dan memiliki informasi dan yang akan dijadikan sebagai node (akar).

3. HASIL DAN PEMBAHASAN

3.1 Hasil

Pada tahapan ini peneliti melakukan seleksi data transaksi mulai dari periode Januari 2022 sampai Maret 2022.

3.1.1 Pengumpulan Data

Pengumpulan data dilakukan untuk memenuhi kebutuhan terkait data yang akan digunakan dalam penelitian ini.

3.1.2 Data Selection

Pada tabel dibawah ini merupakan data penjualan obat dengan sejumlah transaksi diambil dari data penjualan selama 3 bulan yakni bulan Januari – Maret 2023 dan data sampel yang digunakan sebanyak 30 sampel data.

3.1.3 Data Preprocessing / Data Cleaning

Data Cleaning diterapkan untuk menambahkann isi atribut yang hilang atau kosong dan merubah data yang tidak konsisten. Dan pada penelitian ini setelah dilakukan data cleaning maka tidak ada ditemukan data yang tidak konsisten sehingga bisa dilanjutkan ke tahap berikutnya.

3.1.4 Data Transformation

Dalam proses ini, data ditransferkan kedalam bentuk yang sesuai untuk proses data mining.

3.1.5 Data Reduction

Reduksi data dilakukan dengan menghilangkan atribut yang tidak diperlukan sehingga ukuran dari database menjadi kecil dan hanya menyertakan atribut yang diperlukan dalam proses data mining, karena akan lebih efisien terhadap data yang lebih kecil. Masalah klasifikasi berakhir dengan dihasilkan sebuah pengetahuan yang dipresentasikan dalam bentuk diagram yang biasa disebut pohon keputusan(decision tree). Data berikut ini dipergunakan untuk sampel data. Data selengkapnya tampak pada tabel dibawah ini:

Tabel 5.1
Data Sampel Yang Digunakan

No	Nama Obat	Jenis	Harga	Bulan	Keterangan
1	Acyclovir Cream	Obat Generik	Rp 28.000	Januari	Tidak Laku
2	Acyclovir Tablet 200mg	Obat Generik	Rp 6.500	Januari	Laku
3	Acyclovir Tablet 400mg	Obat Generik	Rp 13.000	Januari	Laku
4	Albendazol Tablet 400 mg	Obat Generik	Rp 7.000	Januari	Laku
5	Allupurinol Tablet 100mg	Obat Generik	Rp 10.000	Januari	Tidak Laku
6	Ambroxol Tablet 30 mg	Obat Generik	Rp 7.500	Januari	Tidak Laku
7	Ambroxol Syrup	Obat Generik	Rp 15.000	Januari	Laku
8	Aminophillin Injeksi 24mg/ml – 10 ml	Obat Generik	Rp 5.000	Januari	Laku
9	Aminophillin Tablet 30mg	Obat Generik	Rp 15.000	Januari	Tidak Laku
10	Amitriptillin Tablet Salut 25mg	Obat Generik	Rp 17.000	Januari	Laku
11	Amlodipin 5mg Tablet	Obat Generik	Rp 35.000	Februari	Laku
12	Amlodipin 10 mg Tablet	Obat Generik	Rp 44.900	Februari	Laku
13	Amoksisilin 125mg perSach	Obat Generik	Rp 7.000	Februari	Tidak Laku
14	Amoksisilin Kaplet 500mg	Obat Generik	Rp 5.000	Februari	Tidak Laku
15	Amoksisilin Kapsul 250mg	Obat Generik	Rp 5.000	Februari	Laku
16	Amoksisilin Sirup Kering 125mg/5ml	Obat Generik	Rp 7.000	Februari	Laku
17	Ampicillin Serbuk Injeksi	Obat Generik	Rp 47.000	Maret	Tidak Laku
18	Enervon-C	Multivitamin	Rp 34.000	Maret	Tidak Laku
19	Renovit Multivitamin & Mineral	Multivitamin	Rp 65.000	Maret	Laku

20	Wellness Multivitamin	Multivitamin	Rp 423.000	Maret	Tidak Laku
21	Youvit Gummy Multivitamin	Multivitamin	Rp 32.000	Maret	Laku
22	Redoxon Fortimun	Multivitamin	Rp 60.000	Maret	Tidak Laku
23	Betadine Feminine	Obat Bebas	Rp 33.000	Maret	Tidak Laku
24	Betadine Moutwash	Obat Bebas	Rp 32.000	Maret	Laku
25	Betadine Wash Natural	Obat Bebas	Rp 38.000	Maret	Laku
26	Bbetadine Natural Defense	Obat Bebas	Rp 20.000	Maret	Laku
27	Thephylline	Obat Bebas	Rp 100.000	Maret	Laku
28	Kalpanax	Obat Bebas	Rp 13.000	Maret	Laku
29	Rivanol Kompres	Obat Bebas	Rp 9.000	Maret	Laku
30	Anusol Supositoria	Obat Bebas	Rp 29.000	Maret	Laku

3.2 Pembahasan

Tahapan yang dilakukan untuk pengolahan data menggunakan Algoritma C4.5 seperti berikut ini:

Mencari Nilai Entropy

1. Entropy Total = Entropy(S) = $\sum_{i=1}^n - p_i * \log_2 p_i$
 $= ((-11/30 * \log_2 (11/30)) + (-9/30 * \log_2 (9/30)))$
 $= -11/30 * 0,30103 (11/30) + (-9/30 * 0,30103 (9/30))$
 $= -11/30 * 0,110378 + (-9/30 * 0,090309)$
 $= (0,04047) + (0,02709)$
 $= 0,06756$
2. Entropy Jenis Obat
 - Generik = Entropy(S) = $\sum_{i=1}^n - p_i * \log_2 p_i$
 $= ((-7/17 * \log_2 (7/17)) + (-10/17 * \log_2 (10/17)))$
 $= (-7/17 * 0,30103 (7/17)) + (-10/17 * \log_2 (10/17))$
 $= -7/17 * 0,123954 + (-10/17 * 0,177076)$
 $= 0,05104 + 0,10416$
 $= 0,1552$
 - Multivitamin = Entropy(S) = $\sum_{i=1}^n - p_i * \log_2 p_i$
 $= ((-3/5 * \log_2 (3/5)) + (-2/5 * \log_2 (2/5)))$
 $= 0,15654$
 - Obat Bebas = Entropy(S) = $\sum_{i=1}^n - p_i * \log_2 p_i$
 $= ((-1/8 * \log_2 (1/8)) + (-7/8 * \log_2 (7/8)))$
 $= 0,23518$
3. Entropy Merek
 - Acyclovir = Entropy(S) = $\sum_{i=1}^n - p_i * \log_2 p_i$
 $= ((-1/3 * \log_2 (1/3)) + (-1/3 * \log_2 (1/3)))$
 $= 1,58496$
 - Albendazol = Entropy(S) = $\sum_{i=1}^n - p_i * \log_2 p_i$
 $= ((-1/1 * \log_2 (1/1))$
 $= 0$
 - Allupurinol = Entropy(S) = $\sum_{i=1}^n - p_i * \log_2 p_i$
 $= ((-1/1 * \log_2 (1/1))$
 $= 0$
 - Ambroxol = Entropy(S) = $\sum_{i=1}^n - p_i * \log_2 p_i$
 $= ((-1/2 * \log_2 (1/2)) + (-1/2 * \log_2 (1/2)))$
 $= 1$
 - Aminophilin = Entropy(S) = $\sum_{i=1}^n - p_i * \log_2 p_i$
 $= ((-1/2 * \log_2 (1/2)) + (-1/2 * \log_2 (1/2)))$
 $= 1$
 - Ampicilin = Entropy(S) = $\sum_{i=1}^n - p_i * \log_2 p_i$
 $= ((-1/1 * \log_2 (1/1))$
 $= 0$
 - Amoksisilin = Entropy(S) = $\sum_{i=1}^n - p_i * \log_2 p_i$
 $= ((-2/4 * \log_2 (2/4)) + (-2/4 * \log_2 (2/4)))$
 $= 1$
 - Ampicilin = Entropy(S) = $\sum_{i=1}^n - p_i * \log_2 p_i$
 $= (-1/1 * \log_2 (1/1))$
 $= 0$
 - Enervonce = Entropy(S) = $\sum_{i=1}^n - p_i * \log_2 p_i$

$$= (-2/5 * \log_2 (2/5)) + (-3/5 * \log_2 (3/5))$$

$$= 1,32193 + 0,73697$$

$$= 2,0589$$

- Betadine = Entropy(S) = $\sum_{i=1}^n - p_i * \log_2 p_i$
 $= (-1/4 * \log_2 (1/4)) + (-3/4 * \log_2 (3/4))$
 $= 2$

- Thephiline = Entropy(S) = $\sum_{i=1}^n - p_i * \log_2 p_i$
 $= (-0/4 * \log_2 (0/4)) + (-4/4 * \log_2 (4/4))$
 $= 2,41504$

4. Entropy Bulan

- Januari = Entropy(S) = $\sum_{i=1}^n - p_i * \log_2 p_i$
 $= (-4/10 * \log_2 (4/10)) + (-6/10 * \log_2 (6/10))$
 $= 2,0589$

- Februari = Entropy(S) = $\sum_{i=1}^n - p_i * \log_2 p_i$
 $= (-2/6 * \log_2 (2/6)) + (-4/6 * \log_2 (4/6))$
 $= 2,0589$

- Maret = Entropy(S) = $\sum_{i=1}^n - p_i * \log_2 p_i$
 $= (-5/14 * \log_2 (5/14)) + (-9/14 * \log_2 (9/14))$
 $= 1,48543$

5. Mencari Nilai Gain

- Gain Jenis Obat
 $= 0,06756 - ((17/30) * 0,1552) + (5/30) * 0,15654 + (8/30) * 0,23518$
 $= 0,020387 + 0,026909 + 0,062715$
 $= 0,109191$

- Gain Merk
 $= 0,06756 - ((3/30) * 1,58496) + (1/30) * 0 + (1/30) * 0 + ((2/30) * 1 + ((2/30) * 1 + ((1/30) * 0)) + ((4/30) * 1 + ((1/30) * 0 + ((5/30) * 2,0589 + ((4/30) * 2 + ((4/30) * 2,41504$
 $= 0,090936 + 0 + 0 + 0,66667 + 0,66667 + 0 + 0 + 0,133333 + 0 + 0,34315 + 0,266667 + 0,322005$
 $= 2,489431$

- Gain Bulan
 $= 0,06756 - ((10/30) * 2,0589) + (6/30) * 2,0589 + (14/30) * 1,48543$
 $= 0,61874 + 0,41178 + 0,693201$
 $= 1,723721$

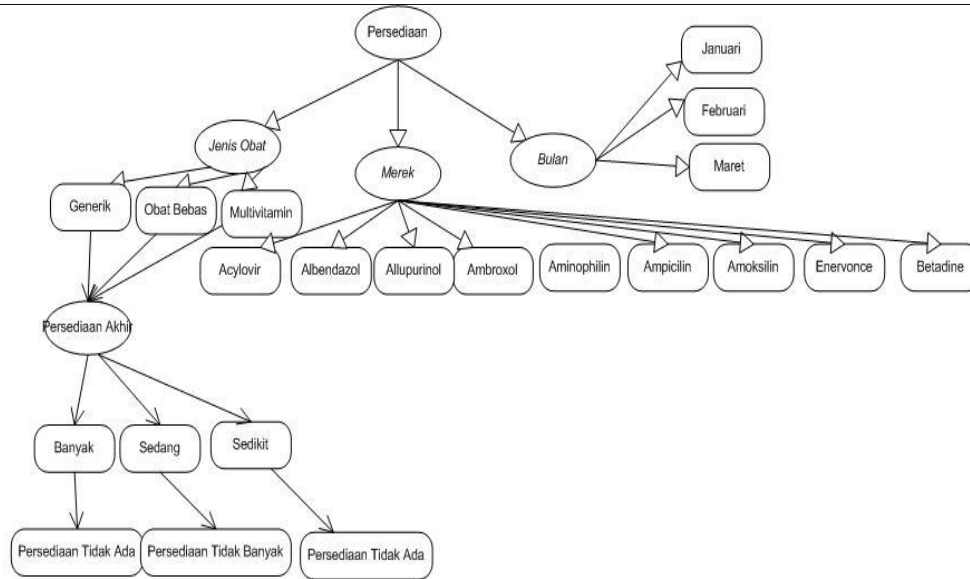
Hasil perhitungan ditunjukkan pada tabel 5.2. Perhitungan Node 1

Tabel 2

Hasil perhitungan Entropy dan Gain

Node	Jumlah	Tidak Laku	Laku	Entropy	Gain
Akar 1					
Total	30	11	19	0,06756	
Jenis Obat					0,109191
Persediaan	17	7	10	0,1552	
Jenis Obat					0,109191
	Generik	5	3	0,15654	
	Multivitamin	8	1	0,23518	
	Obat Bebas				0,109191
Merek	3	1	2	1,58496	
	Acylovir	1	0	0	
	Albendazol	1	0	0	
	Allupurinol	2	1	1	
	Aminophilin	2	1	1	
	Ampicilin	4	2	2	
	Amoksilin	5	2	3	2,0589
	Enervonce	4	1	3	2
	Betadine	4	0	4	2,41504
	Thephiline				2,489431
Bulan	10	4	6	2,0589	
	Januari	6	4	2,0589	
	Februari	14	5	1,48543	
	Maret				1,723721

Dari hasil perhitungan diatas maka didapatkan hasil pohon keputusan akhir yaitu



Gambar 1. Hasil Pohon Keputusan

4. KESIMPULAN

Berdasarkan pembahasan yang dilakukan terkait dengan persediaan obat, maka penulis dapat mengambil beberapa kesimpulan adalah data yang didapatkan dari klinik diolah dengan proses data mining dengan menerapkan algoritma C4.5 mulai dari penerapan data preprocessing, penentuan Entropy dan Gain hingga langkah terakhir yaitu penentuan pohon keputusan yang terbentuk berdasarkan hasil perhitungan entropy dan gain, dan dengan penerapan algoritma C4.5 dapat memberikan hasil yang tepat terkait dengan persediaan obat sehingga tidak kesulitan dalam menangani permasalahan tersebut.

UCAPAN TERIMA KASIH

1. Bapak John F. Marpaung dan Ibu Ritha Z. Tarigan, S.E.,MM. selaku pimpinan yayasan demokrat cemerlang.
2. Ibu Murni Marbun, S.Si., MM.,M.Kom, selaku Ketua STMIK Pelita Nusantara.
3. Bapak Jijon Raphita Sagala, M.Kom selaku Ketua Program Studi Teknik Informatika.
4. Bapak Amran Sitohang, M.Kom selaku Ketua Program Studi Manajemen Informatika.
5. Ibu Sulindawaty, S.Kom.,M.Kom selaku ketua LPM STMIK Pelita Nusantara.
6. Bapak Penda Sudarto Hasugian, S.Kom.,M.Kom selaku ketua LPPM STMIK Pelita Nusantara.
7. Rekan-rekan dosen STMIK Pelita Nusantara.
8. Klinik Romana Medan

DAFTAR PUSTAKA

- [1] Anggraini, S., Defit, S., & Nurcahyo, G. W. (2018). Analisis Data Mining Penjualan Ban Menggunakan Algoritma C4. 5. *Jurnal Ilmu Teknik Elektro ...*, 5(2), 0–7. <https://core.ac.uk/download/pdf/295348196.pdf>
- [2] Dewanti, F. P., Setiyowati, S., & Harjanto, S. (2022). Prediksi Persediaan Obat Untuk Proses Penjualan Menggunakan Metode Decision Tree Pada Apotek. *Jurnal Teknologi Informasi Dan Komunikasi (TIKOMSiN)*, 10(1). <https://doi.org/10.30646/tikomsin.v10i1.604>
- [3] Dewi, K. R., Mauladi, K. F., & Masrurroh. (2020). Analisa Algoritma C4.5 untuk Prediksi Penjualan Obat Pertanian di Toko Dewi Sri. *Seminar Nasional Inovasi Teknologi*, 25, 109–114.
- [4] Eska, J. (2016). Penerapan Data Mining Untuk Prekdiksi Penjualan Wallpaper Menggunakan Algoritma C4.5 STMIK Royal Ksiaran. *JURTEKSI (Jurnal Teknologi Dan Sistem Informasi)*, 2, 9–13.
- [5] Lukhayu Pritalia, G. (2018). Penerapan Algoritma C4.5 untuk Penentuan Ketersediaan Barang E-commerce. *Indonesian Journal of Information Systems*, 1(1), 47–56. <https://doi.org/10.24002/ijis.v1i1.1727>
- [6] Made, I., Ramayu, S., Susanto, F., & Mahendra, G. S. (2022). Penerapan Data Mining Dengan Algoritma C4.5 Dalam Pemesanan Obat Guna Meningkatkan Keuntungan Apotek. *Online) SENADA*, 5, 237–245. <http://senada.idbbali.ac.id>
- [7] Novianti, B., Rismawan, T., & Bahri, S. (2016). Implementasi Data Mining Dengan Algoritma C4.5 Untuk Penjurusan Siswa (Studi Kasus: Sma Negeri 1 Pontianak). *Jurnal Coding, Sistem Komputer Untan*, 04(3), 75–84.
- [8] Siahaan, S. W., Sianipar, K. D. R., R.H Zer, P. P. P. A. N. . F. I., & Hartama, D. (2020). Penerapan Algoritma C4.5 Dalam Meningkatkan Kemampuan Bahasa Inggris Pada Mahasiswa. *Petir*, 13(2), 229–239. <https://doi.org/10.33322/petir.v13i2.1029>