

PREDIKSI INFEKSI SALURAN PERNAFASAN AKUT (ISPA) DENGAN MENGGUNAKAN METODE RANTAI MARKOV PADA KLINIK CIHIDEUNG

Acihmah Sidauruk¹⁾, Hendri Kurniawan²⁾, Adhitya Ronnie Efendi³⁾Edi Winarko⁴⁾

¹⁾ Magister Teknik Informatika STMIK AMIKOM Yogyakarta
^{2,3,4)} Universitas Gajah Mada

Jl Ring road Utara, Condongcatur, Sleman, Yogyakarta 55281

Email : chi_gals@yahoo.com¹⁾, hendr.kurniawan.p@mail.ugm.ac.id²⁾, adhityaronnie@ugm.ac.id³⁾
ewinarko@ugm.ac.id⁴⁾

Abstrak

Pertumbuhan data pada dunia teknologi sangatlah cepat. Begitu banyaknya data yang tersimpan akan menjadi sia-sia jika tidak dimanfaatkan. Hal ini terjadi pada klinik Ciheudung, Garut. Data pasien yang tersimpan selama lebih dari satu dekade belum diolah. Padahal banyak sekali informasi yang dapat digali dari data tersebut. Salah satunya adalah informasi tentang banyaknya pasien yang terkena penyakit ISPA.

Berdasarkan hal tersebut untuk mengetahui informasi yang lebih banyak pada data yang jumlahnya sangat besar digunakan teknik data mining. Sehingga dilakukan penelitian tentang prediksi penyakit Infeksi Saluran Pernafasan Akut (ISPA). Metode yang digunakan adalah metode rantai markov. Metode rantai markov adalah suatu metode statistik yang dapat mengetahui kejadian masa depan berdasarkan pada masa sekarang yang bersifat probabilistik dan kontinyu. Hasil dari penelitian ini adalah dapat diketahuinya peluang seseorang terkena penyakit ISPA sehingga tindakan dan pelayanan pada pasien yang terserang penyakit tersebut dapat diantisipasi sebelumnya. Selain itu juga bisa digunakan untuk mengambil arah kebijakan tertentu

Kata kunci: Data mining, rantai markov, ISPA, prediksi

1. Pendahuluan

Perkembangan teknologi informasi serta pemanfaatannya saat ini telah menyentuh di berbagai bidang kehidupan. Komputer tidak hanya dirancang untuk kebutuhan komputasi sesaat tapi telah dirancang dan dikembangkan untuk tujuan kemudahan bagi manusia mulai dari bidang pendidikan sampai kesehatan. Pada bidang kesehatan, komputer sangat berperan dalam fungsi penyedia informasi. Salah satu bentuk penyedia informasi adalah dengan cara mengolah data dalam jumlah yang sangat besar untuk diketahui informasi yang tersembunyi di dalamnya yang bisa dimanfaatkan untuk kebutuhan di masa mendatang.

Pemanfaatan teknologi informasi di bidang kesehatan bahkan bisa lebih sekedar dari menyimpan data pasien pada suatu rumah sakit saja. Data yang terkumpul sangat banyak dapat digunakan untuk mengambil sebuah informasi penting misalnya untuk menemukan kaitan

antar penyakit yang diderita pasien dengan teknik data mining.

Hal tersebut kemudian yang dilakukan pada klinik Ciheudung. Masih tingginya tingkat penderita penyakit infeksi saluran pernafasan akut di desa Ciheudung merupakan permasalahan yang tidak kunjung selesai. Hal tersebut dapat terjadi karena salah satunya adalah kurang dininya tindakan antisipasi padahal penyakit tersebut dapat dideteksi dengan mudah pada gejala yang muncul seperti batuk, sesak nafas dan nyeri dada[2].

Untuk itu perlu dilakukan tindakan pencegahan. Langkah awalnya adalah dengan mencari tahu peluang seseorang bisa terjangkit penyakit tersebut. Dengan mengetahui seberapa besar peluang seorang terjangkit penyakit tersebut maka pilihan tindakan pencegahan dapat lebih tepat untuk diambil. Untuk itu digunakanlah teknik data mining dengan metode rantai markov untuk memprediksi seberapa besar peluang seseorang mengidap penyakit ISPA.

Data mining adalah teknik pengambilan informasi dari data yang berjumlah banyak dengan menggunakan pola-pola khusus. Metode rantai markov adalah salah satu metode dari data mining yang dapat digunakan untuk memprediksi kejadian masa datang berdasarkan kejadian-kejadian di masa lampau yang saling terkait. Diharapkan dengan mengetahui peluang terjangkitnya penyakit pada masa datang maka dapat segera diambil tindakan pencegahan yang tepat.

2. Pembahasan

Rantai markov adalah metode statistic yang bersifat probabilistic dan periodic serta berbasis state. Pada penelitian ini state yang diamati adalah state pasien terkena ISPA, batuk, demam, nyeri dada, sesak napas, penyakit selain yang telah disebutkan dan state pasien sehat. Sedangkan periode yang diamati adalah perpindahan state pada tiap bulannya.

Data yang diperoleh dari klinik cihideung belumlah terstandarisasi. Sehingga perlu diolah sebelum digunakan untuk prediksi. Pengolahan datanya sesuai dengan tahapan yang ada data mining

Data mining adalah sebuah proses pencarian secara otomatis informasi yang berguna dalam tempat penyimpanan data berukuran besar. Istilah lain yang sering digunakan diantaranya *knowledge discovery (mining) in databases (KDD)*, *knowledge extraction*, *data/pattern analysis*, *data archeology*, *data dredging*, *information harvesting* dan *business intelligence*. Teknik data mining digunakan untuk memeriksa basis data berukuran besar sebagai cara untuk menemukan pola yang baru dan berguna.

Tidak semua pekerjaan pencarian informasi dinyatakan sebagai data mining. Sebagai contoh pencarian record individual menggunakan *database management system (DBMS)* atau pencarian halaman web tertentu melalui query ke semua *search engine* adalah pekerjaan pencarian informasi yang erat kaitannya dengan *information retrieval*. Teknik-teknik data mining dapat digunakan untuk meningkatkan kemampuan *information retrieval system*. Data mining adalah bagian integral dari *knowledge discovery in databases (KDD)*[4].

Data mining memiliki tahapan-tahapan yang harus dipenuhi yaitu[3]:

- Data cleaning:** Proses untuk membersihkan data dari noise data dan data yang tidak konsisten.
- Data integration:** Proses pengkombinasian atau pengintergrasian beberapa sumber data.
- Data selection:** Proses mengambil data yang relevan dari basis data untuk dianalisis.
- Data transformation:** Pada tahap ini mentransformasikan data summary ataupun operasi agregasi ke dalam form yang sesuai untuk penambahan dengan melakukan operasi agregasi, misal pada data transformasi dan konsolidasi yang dilakukan sebelum proses seleksi data, terutama pada data warehousing.
- Data mining:** Merupakan proses yang esensial dimana metode digunakan mengekstrak pola data yang tersembunyi.
- Pattern evaluation:** Untuk mengidentifikasi pola sehingga merepresentasikan pengetahuan berdasarkan nilai-nilai yang menarik.
- Knowledge representation:** Dalam proses ini teknik representasi dan visualisasi data digunakan untuk mempresentasikan pengetahuan yang didapat kepada pengguna.

Pada kenyataannya data yang didapat banyak yang tidak lengkap (*missing value*), mengandung kesalahan (*error/noisy*) dan tidak konsisten dalam pengisiannya selain itu juga terdapat duplikasi data. Untuk itu perlu dilakukan proses yang disebut *preprocessing data* agar hasil yang didapatkan memiliki akurasi yang baik. Proses ini tersusun oleh beberapa tahapan yaitu: *data selection*, *data cleaning* dan *data integration*.

A. Data Selection

Data yang didapat adalah data pasien klinik Ciheudung pada tahun 2002 hingga 2011 yang

berjumlah puluhan ribu. Data tersebut berisi informasi tentang nomor rekamedis (NRM), waktu dan tanggal kunjungan, diagnosis pasien, terapi termasuk peresepan dan biaya yang dikeluarkan pasien.

Karena data yang didapat sangat banyak sehingga perlu dilakukan pengurangan data agar penelitian dapat dilakukan secara efisien dari segi waktu dan sumber daya. Untuk itu dilakukan pengambilan sampel.

Dalam penelitian ini diambil sampel data bulan Januari hingga Desember tahun 2006 karena pada tahun tersebut paling relevan dengan topik penelitian yaitu data yang paling banyak mencatat terjadinya kasus ISPA.

B. Data Cleaning

Pada penelitian ini prosesnya adalah sebagai berikut:

1. Menghilangkan atribut

Data awal terdiri dari satu tabel dengan bermacam-macam atribut meliputi NRM, tanggal periksa, diagnosis, terapi ke-1 hingga terapi ke-10, dosis ke-1 hingga dosis ke-10, obat dan biaya pemeriksaan.

Dari sekian banyak atribut maka dipilihlah atribut NRM, tanggal dan diagnosis. Karena atribut-atribut tersebutlah yang relevan dengan penelitian. Selain dari atribut-atribut tersebut maka dihapus. Proses pemilihan dan penghapusan data dilakukan secara manual dengan menggunakan worksheet

2. Penambahan kolom id periksa

Agar tiap record memiliki id yang unik maka ditambahkan sebuah kolom "id_periksa" yang digunakan sebagai primary key dalam pengolahan data pada DBMS.

3. Mengisi nilai yang hilang

Tahap selanjutnya adalah mengisi field yang kosong karena pada kenyataannya banyak terdapat field kosong atau nilai yang hilang dan penggunaan bahasa yang tidak sesuai dengan kaidah bahasa Indonesia yang baik dan benar atau tidak sesuai dengan standart penulisan diagnosis kedokteran

Untuk mengisi nilai yang hilang dan untuk menjadikan baku penulisan digunakan algoritma yang telah diimplementasikan dalam bahasa pemrograman. Untuk pengisian nilai kosong (*missing value*) dengan pengasumsian kasus terbanyak (modus).

Secara umum langkah-langkah yang digunakan untuk mengisi nilai yang kosong adalah teks yang berada pada kolom diagnosis dipisahkan per kata kemudian di masukkan dalam satu array. Setiap nilai pada indeks array kemudian di cek apakah memenuhi kondisi (state yang sesuai dengan penelitian) atau tidak. Jika TRUE maka field tersebut akan digantikan dengan state yang sesuai, dalam hal ini state yang diteliti adalah ISPA, batuk, demam, sesak napas, nyeri dada dan jika tidak ada yang cocok maka field akan diisi dengan asumsi yang lain yaitu "bukan ispa". Nilai yang telah diisi kemudian akan disimpan ke dalam database.

Karena pada metode markov rujukannya adalah data *by time*. Sehingga dalam satu satuan waktu tidak boleh

ada dua data pasien yang sama maka tahap selanjutnya adalah menghilangkan data redundan yang berada pada bulan sama. Untuk menghilangkan data yang redundan dalam satu bulan maka digunakan query. Query tersebut dapat dilihat pada Gambar 1.

Karena pada metode markov rujukannya adalah data *by time*. Sehingga dalam satu satuan waktu tidak boleh ada dua data pasien yang sama maka tahap selanjutnya adalah menghilangkan data redundan yang berada pada bulan sama. Untuk menghilangkan data yang redundan dalam satu bulan maka digunakan query. Query tersebut dapat dilihat pada Gambar 1.

```

"CREATE TABLE datatunggal AS SELECT distinct
NRM,extract(month from tanggal) as bulan,diagnosisfix2
FROM `ib_periksa`
ORDER BY nrm"
    
```

Gambar 1. Kode program untuk menghilangkan data redundan

Permasalahannya adalah dalam kasus ini tidak bisa diketahui pada sekumpulan data redundan tersebut data mana yang dipilih untuk dijadikan tunggal.

C. Data Integration

Tahap selanjutnya adalah proses integrasi data. Data diubah ke dari format *.xls ke*.csv yang kemudian akan diolah dalam DBMS dalam bentuk *.sql.

Masalah yang muncul ketika preprocessing data adalah DBMS ataupun resource PC tidak mampu sekaligus mengimport ribuan data dan tidak mampu sekaligus memroses data dari bentuk data inkonsisten menjadi konsisten. Sehingga pengerjaannya harus dilakukan secara bertahap. Selain itu masih terdapat beberapa data yang salah dalam update-nya.

Solusi dari hal tersebut bisa saja menggunakan PC atau DBMS yang lebih *capable* namun ada cara lain yaitu dengan optimalisasi algoritma pemrogramannya.

DATA MINING DENGAN RANTAI MARKOV

Rantai markov adalah [2] Dari sudut pandang teori probabilitas, proses stokastik paling tepat didefinisikan sebagai keluarga variable random $\{X(t), t \in T\}$, yaitu untuk setiap $t \in T$, $X(t)$ adalah variable random. Indeks t menyatakan waktu dan $X(t)$ menyatakan state (kedudukan) dari proses pada waktu t.

Proses stokastik parameter diskrit $\{X(t), t = 0, 1, 2, \dots\}$ atau proses stokastik parameter continue $\{X(t), t > 0\}$ dapat dikatakan sebagai proses markov. Proses markov mempunyai peranan sangat penting dalam aplikasi teori stokastik.

Diberikan proses stokastik $\{X(t), t \in T\}$ dengan ruang parameter T dan ruang state S. $X(t)$ disebut sebagai proses markov jika untuk setiap n dan untuk $t_1 < t_2 < t_3 < \dots < t_n$ dengan $t, t_r \in T(r = 1, 2, 3, \dots, n)$ maka berlaku [5]:

$$P_r\{X(t) \in A | X(t_1) = x_1, \dots, X(t_n) = x_n\} = P\{X(t) \in A | X(t_n) = x_n\}$$

Untuk ACS (1) Persamaan tersebut berarti bahwa probabilitas bersyarat dari $X(t) \in A$ untuk nilai-nilai yang diberikan oleh $X(t_1), X(t_2), X(t_3), \dots, X(t_n)$ hanya bergantung pada $X(t_n)$.

Rantai markov adalah sebuah sekuen variable random yang berkorespondensi antar statenya dalam sebuah sistem yang berarti bahwa setiap state bergantung pada state sebelumnya.

Di klinik Cihedueng, Garut pada bulan Maret-April 2006 memiliki jumlah pasien 1299. Pada bulan Maret pasien yang menderita ISPA sebanyak 144 pasien, yang menderita batuk sebanyak 60 pasien, yang menderita nyeri dada sebanyak 2 pasien, yang menderita demam sebanyak 81 pasien, yang menderita sesak napas sebanyak 8 pasien, yang menderita sakit selain itu ada 499 pasien dan yang dalam kondisi sehat ada 505 pasien.

Pada bulan April pasien yang berkunjung mengalami penyakit yang berbeda. Hal tersebut terlihat pada tabel 1.

Tabel 1. DATA KUNJUNGAN PASIEN BULAN APRIL

Kondisi Pasien	Jumlah Pasien
ISPA	53
Batuk	71
Nyeri dada	1
Demam	75
Sesak napas	7
Penyakit selain ISPA	430
Sehat	662

Dilihat dari bulan Maret hingga April tampak perubahan jumlah pasien pada penyakit tertentu. Hal tersebut dapat dilihat padatabel 3.2.

Tabel 2. DATA PERUBAHAN STATE BULAN MARET-APRIL

STATE	ISPA	Batuk	Nyeri Dada	Demam	Sesak	Bukan Ispa	Sehat
ISPA	2	2	0	0	0	4	136
Batuk	2	0	0	0	0	4	54
Nyeri Dada	0	0	0	0	0	0	2
Demam	1	1	0	1	0	6	72
Sesak	0	0	0	0	0	2	6
Bukan Ispa	6	9	0	5	2	85	392
Sehat	42	59	1	69	5	329	0

Tabel 2. menjelaskan bahwa jumlah pasien ISPA pada bulan Maret yang pada bulan April tetap pada kondisi sakit ISPA berjumlah 2 pasien, yang berpindah dari kondisi sakit ISPA ke sakit batuk ada 2 pasien, yang berpindah dari kondisi sakit ISPA ke sakit demam dan sesak napas tidak ada, yang berpindah dari kondisi sakit ISPA ke sakit bukan ISPA dan gejalanya ada 4 pasien dan yang pada bulan April berpindah dari kondisi ISPA ke sehat ada 136 pasien.

Jumlah dalam 1 baris menunjukkan jumlah state awal sedangkan jumlah dalam 1 kolom menunjukkan jumlah state bulan ke-2.

Sehingga dari tabel 2 dapat dibuat matrik transisinya. Matrik transisi tersebut dapat dilihat pada persamaan (2).

$$\begin{pmatrix} 0.014 & 0.014 & 0 & 0 & 0 & 0.028 & 0.94 \\ 0.033 & 0 & 0 & 0 & 0 & 0.067 & 0.9 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.99 \\ 0.012 & 0.012 & 0 & 0.012 & 0 & 0.07 & 0.89 \\ 0 & 0 & 0 & 0 & 0 & 0.25 & 0.75 \\ 0.01 & 0.02 & 0 & 0.01 & 0.01 & 0.17 & 0.79 \\ 0.08 & 0.12 & 0.01 & 0.14 & 0.01 & 0.65 & 0 \end{pmatrix}$$

Angka-angka tersebut diperoleh dari jumlah perpindahan state pada tiap penyakit dibagi dengan jumlah state awal. Sebagai contoh pada matrik kolom 2 baris 1 dengan nilai 0.014 yaitu perpindahan state dari ISPA ke batuk diperlihatkan pada persamaan (3).

$$\frac{2}{144} = 0.014$$

Sehingga dari matrik tersebut dapat dihitung kemungkinan seorang pasien akan mengalami suatu penyakit tertentu pada masa yang akan datang. Jika diasumsikan pasien A mengalami penyakit ISPA pada bulan Maret (state awal) maka peluangnya menderita sakit ISPA adalah 1 sehingga dapat ditulis untuk matrik probabilitasnya pada kondisi awal

$$(1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0)$$

Maka dengan mengalikan matrik probabilitas tersebut dengan matrik transisi pada persamaan (2) dapat dicari peluang pada bulan April, yaitu:

$$(1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0) \times$$

$$\begin{pmatrix} 0.014 & 0.014 & 0 & 0 & 0 & 0.028 & 0.94 \\ 0.033 & 0 & 0 & 0 & 0 & 0.067 & 0.9 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.99 \\ 0.012 & 0.012 & 0 & 0.012 & 0 & 0.07 & 0.89 \\ 0 & 0 & 0 & 0 & 0 & 0.25 & 0.75 \\ 0.01 & 0.02 & 0 & 0.01 & 0.01 & 0.17 & 0.79 \\ 0.08 & 0.12 & 0.01 & 0.14 & 0.01 & 0.65 & 0 \end{pmatrix}$$

$$= (0.014 \ 0.014 \ 0 \ 0 \ 0 \ 0.028 \ 0.94)$$

Dan untuk bulan Mei adalah:

$$(0.014 \ 0.014 \ 0 \ 0 \ 0 \ 0.028 \ 0.94) \times$$

$$\begin{pmatrix} 0.14 & 0.14 & 0 & 0 & 0 & 0.28 & 0.94 \\ 0.33 & 0 & 0 & 0 & 0 & 0.67 & 0.9 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.99 \\ 0.12 & 0.12 & 0 & 0.12 & 0 & 0.07 & 0.89 \\ 0 & 0 & 0 & 0 & 0 & 0.25 & 0.75 \\ 0.01 & 0.02 & 0 & 0.01 & 0.01 & 0.17 & 0.79 \\ 0.08 & 0.12 & 0.01 & 0.14 & 0.01 & 0.65 & 0 \end{pmatrix} =$$

$$(0.076 \ 0.113 \ 0.0094 \ 0.132 \ 0.009 \ 0.617 \ 0.479)$$

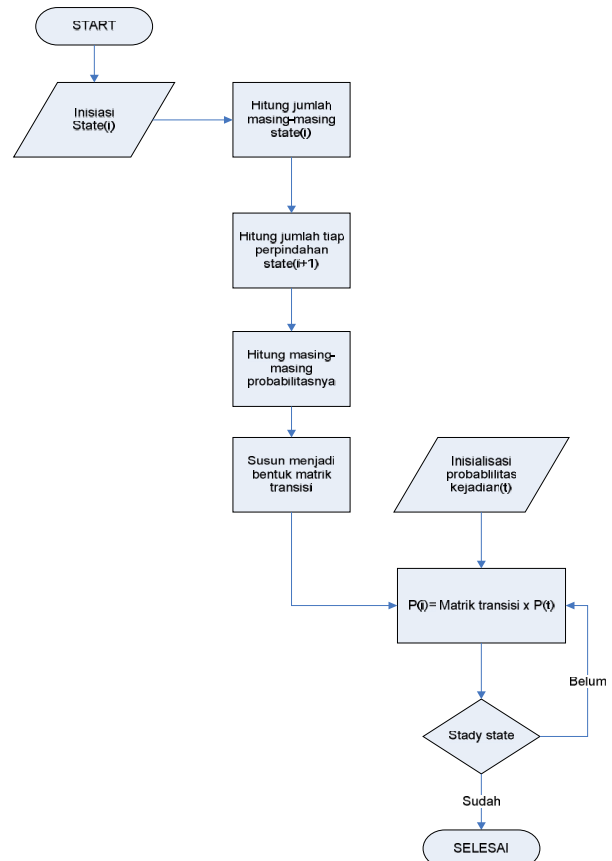
Kesimpulannya adalah pasien yang pada bulan Maret mengalami penyakit ISPA maka peluang pasien tersebut sakit pada bulan Mei adalah ISPA sebanyak 7,6%, batuk sebanyak 11,3%, nyeri dada sebanyak 0,9%, demam

sebanyak 13,2%, sesak napas sebanyak 0,9%, penyakit selain itu sebanyak 61,7% dan peluang sehatnya sebanyak 47,9%.

Melihat dari peluang kejadian yang ada maka dapat dikatakan bahwa pasien yang pada bulan pertama mengalami penyakit ISPA pada bulan ketiga dapat menderita sakit selain ISPA.

Hal yang paling utama pada proses training data adalah membuat modelnya dengan cara membentuk matrik transisi. Langkah-langkah yang dilakukan dalam membentuk matrik tersebut adalah pertama kali tentukan dulu periode yang akan dijadikan dasar perhitungan. Pada kasus ini maka bulan mana yang akan dipilih. Kemudian hitung masing-masing state awal. Lalu hitung jumlah perpindahan state. Dengan syarat bahwa data yang dihitung adalah data yang sama antara bulan pertama dengan bulan kedua.

Langkah selanjutnya adalah menghitung masing-masing probabilitas perpindahan state dengan cara jumlah state yang pindah dibagi total state awal. Algoritmanya secara keseluruhan dapat dilihat pada Gambar 2.



Gambar 2. Algoritma menghitung probabilitas state yang akan datang

Prosesnya diawali ketika memilih data mana yang akan dijadikan dasar training. Dalam hal ini maka yang dimaksud adalah nama bulan. Selanjutnya proses training dilakukan dengan cara membandingkan data pasien yang ada pada bulan i dengan bulan i+1. Setiap perubahan state

akan dicatat. Kode programnya dapat dilihat pada Gambar 3.

```

1. $query=$this->db->query("SELECT NRM,bulan,diagnosisfix2
2. FROM datatunanggalix
3. WHERE bulan = $bulan or bulan = $bulan2");
4. for($j=0;$j<$query->num_rows();$j++){
5. if($query->num_rows() > 0) {
6. $row = $query->row($j);
7. $row2 = $query->next_row(); //return $query->result();
8. }if ($row->NRM == $row2->NRM){
9. $j=$j+1;
10. $pos=$row->diagnosisfix2.$row2->diagnosisfix2;
11. if (isset($a[$pos])){
12. $a[$pos]=$a[$pos]+1;}
13. }elseif ($bulan == $row->bulan){ //sakit->sehat
14. $pos=$row->diagnosisfix2."sehat";
15. $a[$pos]=$a[$pos]+1;
16. }elseif ($bulan2 == $row->bulan){ //sehat->sakit
17. $pos="sehat".$row->diagnosisfix2;
18. $a[$pos]=$a[$pos]+1;
19. }
    
```

Gambar 3. Inisiasi perubahan state

Langkah awalnya adalah memilih periode yang akan ditraining. Kemudian menggunakan query sql data tersebut dipilih. Fungsi query->num_rows() adalah fungsi untuk mendapatkan banyaknya row dari table yang didapat. Selama perulangan (for) berlangsung maka data akan di pilih satu persatu kemudian dibandingkan dengan bulan selanjutnya.

Setiap perubahan state akan dicatat melalui variable \$a[\$pos]. ketika perulangan selesai (telah mencapai maksimum row) maka selanjutnya adalah dihitung probabilitas tiap perubahannya dengan cara jumlah state yang berubah dibagi jumlah state awal. Kode programnya dapat dilihat pada Gambar 4.

Karena data yang didapat adalah data pemeriksaan pasien, tentunya tidak setiap bulan (periode) pasien selalu periksa. Sehingga pada proses ini dilakukan pengasumsian ketika data pasien tidak muncul maka pasien dianggap sedang dalam status "sehat". Hal ini dilakukan agar tingkat error bisa diminimalkan.

```

1. $jumlahIspe=$jumlahDemam=$jumlahBatuk=$jumlahNy
2. eri=$jumlahSesak=$jumlah
3. hBukanIspe=$jumlahSehat=0.01;
4. for($x=0;$x<=6;$x++){
5. $jumlahIspe=$jumlahIspe+$state[0][$x];
6. $jumlahBatuk=$jumlahBatuk+$state[1][$x];
7. $jumlahNyeri=$jumlahNyeri+$state[2][$x];
8. $jumlahDemam=$jumlahDemam+$state[3][$x];
9. $jumlahSesak=$jumlahSesak+$state[4][$x];
10. $jumlahBukanIspe=$jumlahBukanIspe+$state[5][$x];
11. $jumlahSehat=$jumlahSehat+$state[6][$x];
    
```

Gambar 4. Kode program untuk menghitung jumlah total masing-masing state

Setelah total tiap state didapatkan selanjutnya adalah menghitung probabilitas tiap statenya dengan cara membagi jumlah perpindahan tiap state dengan jumlah total state awal. Kode programnya dapat dilihat pada Gambar 5.

```

1. for($x=0;$x<=6;$x++){
2. $prob[0][$x]=$state[0][$x]/$jumlahIspe;
3. $prob[1][$x]=$state[1][$x]/$jumlahBatuk;
4. $prob[2][$x]=$state[2][$x]/$jumlahNyeri;
5. $prob[3][$x]=$state[3][$x]/$jumlahDemam;
6. $prob[4][$x]=$state[4][$x]/$jumlahSesak;
7. $prob[5][$x]=$state[5][$x]/$jumlahBukanIspe;
8. $prob[6][$x]=$state[6][$x]/$jumlahSehat;
9. }
    
```

Gambar 5. Kode program untuk menghitung probabilitas matrik transisi

Matrik transisi yang telah terbentuk kemudian bisa digunakan untuk menghitung probabilitas state pada masa yang akan datang. Dengan cara mengalikan matrik state awal dengan matrik transisi sebanyak periode yang diinginkan. Kode programnya dapat dilihat pada Gambar 6.

```

1. while($step <= $bulan){
2. if($step > 1){
3. $test=$store;}
4. for($i=0;$i<=6;$i++){
5. $temp=0;
6. for($j=0;$j<=6;$j++){
7. $kali=$test[$i]*$prob[$j][$i];
8. $temp=$temp+(round($kali,3));
9. }
10. $pred[$i]=$temp;
11. $store[$i]=$pred[$i];
12. }
13. $step++;
14. $sekal=$bulan+$step;
15. }
    
```

Gambar 6. Kode program perkalian matrik transisi

Variable \$step adalah variable yang menunjukkan berapa kali periode proses dilakukan. Sedangkan variable \$test adalah variable yang bernilai array yang menunjukkan kondisi seorang pasien. Misalkan seorang pasien pada bulan pertama menderita sakit ISPA maka isi dari \$test adalah [1 0 0 0 0 0]. Hasil perubahan dari matrik inilah yang menunjukkan peluang seorang pasien menderita suatu penyakit.

II. HASIL IMPLEMENTASI

Hasil dari perhitungan matrik transisi adalah berupa angka-angka probabilistik

STATE	ISPA	Batuk	Nyeri	Demam	Sesak	Lainnya	Sehat
ISPA	0.018	0.018	0	0	0	0.025	0.944
Batuk	0.023	0	0	0	0	0.067	0.9
Nyeri	0	0	0	0	0	0	0.999
Demam	0.012	0.012	0	0.012	0	0.074	0.889
Sesak	0	0	0	0	0	0.25	0.75
Lain-lain	0.012	0.018	0	0.01	0.004	0.17	0.786
Sehat	0.02	0.117	0.002	0.137	0.01	0.651	1

Gambar 3. Interface matrik transisi dengan keterangan statenya

Hasil perhitungan prediksi dapat dilihat pada tabel 3. Tabel 3 menunjukkan prosentase probabilitas seorang pasien terkena suatu penyakit hingga periode ke-24 dengan state awal ISPA. Tabel tersebut juga menjelaskan

bahwa probabilitas mendekati konstan mulai periode ke-15. Inilah yang disebut dengan kondisi *steady state*.

Selain itu informasi yang juga dapat dilihat adalah secara rinci dijelaskan tentang prosentase prediksi penyakit. Sebagai contoh pada periode kedua peluang seorang terkena penyakit ISPA adalah 7.90 %, terserang sakit batuk sebanyak 11.10%, terkena penyakit nyeri dada sebanyak 0.20%, berpeluang sakit demam 12.90%, berpeluang terkena sesak napas 0.90%, terkena penyakit selain ISPA dan gejalanya sebanyak 62.10% dan berpeluang sehat sebanyak 4.80%.

Tabel 3. PERCOBAAN PREDIKSI HINGGA BULAN KE-24

Bln Ke-	ISPA	Batuk	Nyeri	Demam	Sesak	Lainnya	Sehat
1	1.40%	1.40%	0%	0%	0%	2.80%	94.40%
2	7.90%	11.10%	0.20%	12.90%	0.90%	62.10%	4.80%
3	1.80%	2%	0%	1.50%	0.20%	15.80%	78.70%
4	6.80%	9.50%	0.20%	11%	0.90%	54.20%	17.30%
5	2.60%	3.20%	0%	3%	0.40%	22.30%	68.20%
6	6.10%	8.40%	0.10%	9.50%	0.80%	48.80%	25.90%
7	3.30%	4.10%	0.10%	4.10%	0.50%	26.90%	60.80%
8	5.60%	7.70%	0.10%	8.70%	0.70%	45%	32%
9	3.70%	4.70%	0.10%	5%	0.50%	30%	55.80%
10	5.40%	7.20%	0.10%	8%	0.70%	42.40%	36.20%
11	3.90%	5.20%	0.10%	5.40%	0.60%	32.20%	52.60%
12	5.20%	6.90%	0.10%	7.60%	0.60%	40.70%	39%
13	4.10%	5.50%	0.10%	5.80%	0.60%	33.60%	50.40%
14	5%	6.70%	0.10%	7.30%	0.60%	39.50%	40.90%
15	4.30%	5.70%	0.10%	6.10%	0.60%	34.40%	48.70%
16	4.90%	6.50%	0.10%	7.10%	0.60%	38.70%	42.10%
17	4.40%	5.80%	0.10%	6.30%	0.60%	35.10%	47.60%
18	4.80%	6.40%	0.10%	7%	0.60%	38.10%	43.10%
19	4.50%	5.90%	0.10%	6.40%	0.60%	35.70%	46.90%
20	4.70%	6.30%	0.10%	6.90%	0.60%	37.80%	43.70%
21	4.50%	6%	0.10%	6.50%	0.60%	36%	46.40%
22	4.70%	6.20%	0.10%	6.80%	0.60%	37.40%	44.20%
23	4.50%	6.10%	0.10%	6.50%	0.50%	36.30%	45.90%

3. Kesimpulan

Dari seluruh proses penelitian yang telah dilakukan dapat disimpulkan, yaitu dengan menggunakan rantai markov dapat memprediksi penyakit ISPA. Hasilnya adalah berupa angka-angka probabilitistik pada masa yang akan datang seperti terlihat pada tabel 3.

Sample data yang digunakan sangat mempengaruhi hasil penelitian hanya saja periode untuk mencapai kondisi mendekati *steady state* tidak dapat dipastikan.

Daftar Pustaka

[1] [1] Ching, Wai- Ki and K. Ng., Michael, 2006, *Markov chains model, Algorithms and Application*, Hongkong.
 [2] Corwin, E.J., 2009, *Buku Saku Patofisiologi*, EGC, Jakarta
 [3] Han, J., Kamber, M., Pei, J., 2006, *Data Mining: Concepts and Techniques*, Third Edition.
 [4] Larose, D.T., 2005, *Discovering Knowledge in Data Introduction to Data Mining*, Wiley Interscience, New Jersey.

[5] Tan, W. Y., 2002, *Stochastic Models with Applications to Genetics, Cancers, Aids and Other Biomedical System*, World scientific, Singapore.

[4] A.S.R. Ansori, M. Hariadi, W. Endah, "Pemodelan Retakan Tiga Dimensi Akibat Ledakan Untuk Serious Games", in *Proc. Semnasteknomedia 2013*, pp.13-1, Januari 13, 2013.

Biodata Penulis

Acihmah Sidauruk, memperoleh gelar Sarjana Komputer (S.Kom), Jurusan Sistem Informasi STMIK AMIKOM Yogyakarta, lulus tahun 2010. Memperoleh gelar Magister Komputer (M.Kom) Program Pasca Sarjana Magister Teknik Informatika STMIK AMIKOM Yogyakarta, lulus tahun 2014, Saat ini menjadi Dosen di STMIK AMIKOM Yogyakarta.

Hendri Kurniawan P, memperoleh gelar Ahli Media Komputer (Amd) Jurusan Ilmu Komputer dan Elektronika Universitas Gajah Mada Yogyakarta, lulus tahun 2011. Memperoleh gelar Sarjana Komputer Program SiST Universitas Gajah Mada Yogyakarta, lulus tahun 2014, Saat ini menjadi Staff di Universitas Gajah Mada Yogyakarta.

Adhitya Ronnie Effendi memperoleh gelar Sarjana Fakultas MIPA Institut Teknologi Bandung, lulus tahun 1998. Memperoleh gelar Master Mathematics, Rijks Universiteit Groningen (Rug) Netherlands lulus tahun 2002, Memperoleh gelar Doktor Mathematics Universitas Gajah Mada Yogyakarta lulus tahun 2011. Saat ini menjadi Dosen di Universitas Gajah Mada.

Edi Winarko memperoleh gelar sarjana Mathematics Universitas Gajah Mada Yogyakarta, Master Computer Science, Queen's University Canada, Doktor Computer Science, Flinders University of Shouth Australia. Saat ini menjadi Dosen di Universitas Gajah Mada Yogyakarta.