

Analisis dan Implementasi Sistem Pendiagnosis Penyakit *Tuberculosis* Menggunakan Metode *Case-Based Reasoning*

Bimmo Satryo Wicaksono

Fakultas Informatika
Universitas Telkom
Jl. Telekomunikasi, Terusan Buah Batu
Bandung, Indonesia
bimmo.satryo@gmail.com

Ade Romadhony

Fakultas Informatika
Universitas Telkom
Jl. Telekomunikasi, Terusan Buah Batu
Bandung, Indonesia
aderomadhony@telkomuniversity.ac.id

Mahmud Dwi Sulistiyo

Fakultas Informatika
Universitas Telkom
Jl. Telekomunikasi, Terusan Buah Batu
Bandung, Indonesia
mahmuddwis@telkomuniversity.ac.id

Abstract—*Tuberculosis* atau yang dikenal dengan TB, adalah suatu penyakit yang disebabkan oleh sebuah virus yang bernama *Mycobacterium tuberculosis*. Penyakit ini termasuk mematikan karena menyerang organ paru-paru. Di Indonesia sendiri, pada tahun 2012, data dari Kementerian Kesehatan Indonesia menyebutkan bahwa 1 orang meninggal setiap 8 menit karena penyakit TB ini. Hal ini jelas menunjukkan betapa bahayanya jenis penyakit ini. Berangkat dari hal ini, penulis memiliki ide untuk mengembangkan suatu aplikasi sistem berbasis pengetahuan yang bertujuan untuk membantu dokter dalam mendiagnosis penyakit TB. Aplikasi atau sistem yang dikembangkan ini menggunakan metode *Case-Based Reasoning (CBR)*. Data yang digunakan adalah data rekam medis pasien pengidap penyakit TB dengan kasus yang baru atau pasien yang baru mengalami penyakit TB sebanyak satu kali. Data tersebut akan diolah untuk menjadi acuan saat pasien memasukkan gejala yang dialaminya. Perhitungan *similarity* antara kasus lama dan baru pada sistem menggunakan probabilitas Bayes. Berdasarkan percobaan yang telah dilakukan, sistem berbasis pengetahuan yang dibangun menggunakan metode CBR ini mampu mendiagnosis penyakit TB dengan baik. Rata-rata akurasi sistem dalam mendiagnosis TB dalam berbagai kondisi bernilai sekitar 85% dan tingkat akurasi tertinggi sistem mencapai 90%. Nilai tersebut dinilai cukup memuaskan karena telah memenuhi kualifikasi minimum yang ditentukan oleh WHO, serta mampu mencapai rata-rata kemampuan dokter di Indonesia dalam mendiagnosis TB.

Keywords—*Tuberculosis, Case-Based Reasoning, Similarity, Probabilitas Bayes*

I. PENDAHULUAN

Penyakit adalah suatu keadaan abnormal dari tubuh atau pikiran yang menyebabkan ketidaknyamanan, disfungsi atau kesukaran terhadap orang yang dipengaruhinya. Salah satu penyakit yang sedang banyak menjadi bahan perbincangan adalah *Tuberculosis* atau biasa dikenal dengan sebutan TB. Persoalan yang berkembang adalah laju pertumbuhan pasien lebih cepat dibandingkan dengan jumlah dokter yang tersedia. Hal ini jelas menjadi masalah besar karena pada dasarnya setiap manusia memiliki hak untuk mendapatkan pelayanan yang baik untuk penyakit yang dideritanya.

Terkait dengan masalah penanganan penyakit TB tersebut, maka tugas seorang dokter akan sangat terbantu apabila terdapat sebuah sistem yang dapat membantu dokter dalam mendiagnosis penyakit TB. Tujuan dari adanya sistem tersebut jelas bukan untuk menggantikan peran seorang dokter, tetapi lebih kepada memberikan rekomendasi atau kemungkinan hasil diagnosis berdasarkan gejala yang dialami oleh pasien. Selain itu, sistem tersebut juga dibangun untuk membantu dokter dalam mengurangi risiko terjadinya human error akibat begitu banyaknya pasien yang harus ditangani oleh seorang dokter dalam satu waktu.

Sistem berbasis pengetahuan merupakan upaya untuk menyimpan pengetahuan yang dimiliki oleh seorang pakar ke dalam komputer, sehingga suatu saat basis pengetahuan tersebut dapat digunakan kembali untuk menyelesaikan permasalahan yang dihadapi sesuai dengan basis pengetahuan yang telah disimpan [1]. Dalam pembangunan sistem berbasis pengetahuan untuk mendiagnosis TB ini, dapat digunakan beberapa metode, seperti *Rule-Based Reasoning (RBR)* dan *Case-Based Reasoning (CBR)*.

Pada penelitian kali ini, metode yang digunakan adalah CBR. Metode tersebut memiliki kesamaan dengan RBR dalam hal konsep penalarannya. Namun, CBR dipilih karena memiliki keunggulan, yaitu lebih fleksibel dibandingkan RBR yang sangat terbatas pada aturan yang dibangun.

CBR adalah sistem berbasis pengetahuan yang menyelesaikan masalah dengan melakukan penalaran berdasarkan pengetahuan paling relevan yang telah dimiliki. Selanjutnya, sistem akan melakukan proses adaptasi terhadap pengetahuan tersebut untuk menyesuaikan dengan permasalahan baru. Hal tersebut membuat CBR dapat belajar dan beradaptasi terhadap kasus-kasus yang baru.

Pada penerapan metode CBR di sini, dari pasien akan diinputkan gejala apa saja yang dialaminya. Setelah itu, akan keluar hasil diagnosisnya. Dengan mekanisme kerja aplikasi tersebut, diharapkan kesimpulan yang didapat bisa sangat presisi mengenai TB jenis apa yang diderita pasien tersebut. Selain itu nantinya akan ada rekomendasi untuk pasien tersebut dan obat apa yang bisa dikonsumsi.

II. METODE

A. Case-Based Reasoning

Case-Based Reasoning (CBR) adalah suatu cabang ilmu dari Artificial Intelligence (AI) yang mampu bekerja dengan pemecahan masalah secara experienced-based [15] atau berdasarkan pengalaman (data historis). CBR memiliki paradigma yang berbeda dengan cabang ilmu AI lain, khususnya untuk problem-solving [16]. CBR merefleksikan cara kerja penyelesaian masalah manusia yang menggunakan pengetahuan yang ia miliki dalam pemecahan masalah terdahulu yang kemudian digunakan sebagai titik awal proses penyelesaian permasalahan baru yang mirip dengan permasalahan yang ia selesaikan sebelumnya [6].

Ide dasar dari CBR (Case-Based Reasoning) adalah meniru kemampuan manusia, yaitu menyelesaikan masalah baru menggunakan jawaban atau pengalaman dari masalah lama [17]. Cara kerja CBR adalah dengan membandingkan kasus baru dengan kasus lama, jika kasus baru tersebut mempunyai kemiripan dengan kasus lama maka CBR akan memberikan jawaban kasus lama untuk kasus baru tersebut. Jika tidak ada yang cocok maka CBR akan melakukan adaptasi, dengan cara memasukkan kasus baru tersebut ke dalam database penyimpanan kasus (case base), sehingga secara tidak langsung pengetahuan CBR akan bertambah.

Kelebihan utama dari CBR dibandingkan dengan Rule-Based Reasoning (RBR) adalah dalam hal akuisisi pengetahuan, di mana pada sistem CBR dapat menghilangkan kebutuhan untuk ekstrak model atau kumpulan dari aturan-aturan, seperti yang diperlukan dalam model atau sistem yang berbasis aturan. Akuisisi pengetahuan pada CBR terdapat pada kumpulan pengalaman atau kasus-kasus sebelumnya. Selain itu, dengan CBR penalaran tetap dapat dilakukan jika ada data yang tidak lengkap atau tidak tepat.

Ketika proses retrieval dilakukan, ada kemungkinan antara kasus baru dengan kasus lama pada basis kasus tidak mirip. Namun, dari ukuran kemiripan tersebut tetap dapat dilakukan penalaran dan melakukan evaluasi terhadap ketidaklengkapan atau ketidaktepatan data yang diberikan [3]. Ketidaklengkapan itu dapat diatasi dengan melakukan penambahan pengetahuan yang diambil dari dokter. Selain itu, sistem CBR bisa belajar otomatis yang sangat penting sebagai domain medis berkembang dengan waktu. Sistem berbasis aturan tidak bisa belajar secara otomatis, aturan baru biasanya dimasukkan secara manual.

Secara umum metode ini terdiri dari 4 langkah [7], yaitu:

1) Retrieve

Pada proses ini adalah proses pengenalan masalah, yaitu dengan mengidentifikasi masalah yang baru.

2) Reuse

Dalam proses Reuse, sistem akan melakukan pencarian masalah terdahulu pada database melalui identifikasi masalah baru. Kemudian menggunakan kembali informasi permasalahan terdahulu tersebut yang memiliki kesamaan untuk menyelesaikan permasalahan yang baru.

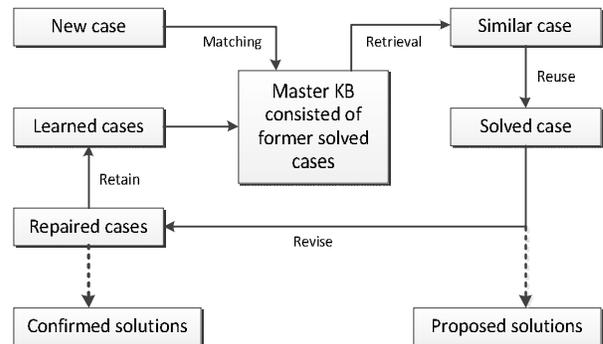
3) Retain

Pada proses terakhir ini, sistem akan menyimpan permasalahan baru ke dalam knowledge-base untuk menyelesaikan permasalahan yang akan datang.

4) Revise

Pada proses ini, informasi akan dievaluasi kembali untuk mengatasi masalah yang terjadi pada permasalahan baru. Kemudian sistem akan mengeluarkan solusi masalah baru.

Adapun skema proses yang terjadi pada metode CBR ini dapat diilustrasikan pada Gambar 1.



Gambar 1. Skema Case-Based Reasoning [7]

B. Similarity Value

Similarity value merupakan derajat tingkat kesamaan yang dalam hal ini adalah kesamaan antara input dari user dan kasus yang ada dalam database [18]. Similarity value paling rendah bernilai 0 dan paling tinggi bernilai 1. Proses penghitungan similarity value adalah dengan membandingkan jumlah gejala yang sama pada input user dengan solusi yang ada pada database kasus.

Pada kasus ini, similarity value dapat dihitung menggunakan rumus berikut.

$$SV = \frac{\text{Total gejala yang sama}}{\text{Total gejala}}$$

C. Probabilitas Bayes

Probabilitas Bayes adalah suatu cara untuk memecahkan masalah ketidakpastian dengan rumus sebagai berikut [20].

$$P(H_i|E) = \frac{P(E|H_i) * P(H_i)}{\sum_{k=i}^n P(E|H_i) * P(H_i)}$$

di mana:

$P(H_i|E)$ = probabilitas hipotesis H_i benar jika diberikan evidence E

$P(E|H_i)$ = probabilitas munculnya evidence E jika diketahui hipotesis H_i benar

$P(H_i)$ = probabilitas hipotesis tanpa memandang evidence sebelumnya

n = jumlah hipotesis yang mungkin

Berikut contoh penerapan Probabilitas Bayes untuk suatu kasus sederhana. Misalnya, jika diketahui:

A = Terjadinya Solusi A

B = Munculnya Penyakit B

$P(A) = 0,25$

$P(B) = 0,5$

$P(A|B) = 0,15$

maka,

$$P(B|A) = \frac{P(A|B) * P(B)}{P(A)}$$

$$P(B|A) = \frac{0,15 * 0,5}{0,25}$$

$$P(B|A) = 0,3$$

III. PERANCANGAN SISTEM

A. Deskripsi Umum Sistem

Dalam pembangunannya, sistem ini akan dibangun dengan bahasa pemrograman Java berbasis Object Oriented, dan akan menggunakan database Proses pengumpulan solusi pada metode CBR dilakukan dengan membandingkan input dari user dengan database kasus yang ada pada sistem. Solusi akan dipilih jika solusi tersebut memenuhi minimum SV yang ada.

Kemudian dari kumpulan solusi tersebut akan diproses lagi dengan metode bayes untuk mendapatkan solusi yang memiliki peluang paling besar. Setelah didapatkan solusi dengan peluang yang paling besar, maka sistem akan menampilkan solusi tersebut.

B. Perancangan Data

Data yang terkumpul adalah data selama 1 tahun, yaitu dari bulan Juni 2012 sampai dengan bulan Juni 2013. Sumber data berasal dari RSUD Kabupaten Sumedang. Total semua data yang diperoleh berjumlah 120 record atau baris data.

Koleksi data tersebut merupakan data rekam medis pasien khusus penderita TB yang baru pertama kali menderita penyakit tersebut (kasus baru). Data ini akan dijadikan sebagai dataset kasus dan data uji. Data tersebut dijadikan 2 skenario, dimana skenario A menggunakan 80 baris data untuk dataset kasus dan 40 baris data untuk pengujian. Sedangkan skenario B menggunakan 90 baris data untuk dataset kasus dan 30 baris data untuk pengujian.

Setiap baris data terdiri dari 34 atribut penting. 32 atribut pertama menunjukkan ada tidaknya gejala untuk 32 macam gejala, 1 atribut berikutnya menunjukkan jumlah kemunculan gejala dalam 1 baris data, dan atribut terakhir menunjukkan kelas untuk satu baris data tersebut, yaitu jenis penyakit TB.

Kemunculan gejala ditandai dengan nilai 1, sedangkan nilai 0 menunjukkan bahwa gejala tidak muncul. Artinya, untuk penelitian kali ini, setiap gejala memiliki bobot yang sama. Untuk kelas pada penelitian kali ini, ada 6 jenis penyakit TB

yang akan dilibatkan, yaitu TB Paru, TB Meningitis, TB Milier, TB Abdomin, TB Perikarditis, dan TB Kulit.

TABEL 1. PERANCANGAN DAN ATRIBUT DATA

No	Gejala 1-32 (0/1)	nGejalaMuncul	Kelas TB
1	1001 ... 01	5	2
2	0010 ... 10	3	3
3	0100 ... 10	4	3
...

C. Data Gejala

Seperti yang telah dijelaskan sebelumnya, data gejala yang digunakan pada penelitian ini bersumber dari RSUD Kabupaten Sumedang. Ada 32 gejala yang menjadi acuan untuk sistem yang dibangun. Tabel 2 menunjukkan daftar 32 gejala tersebut beserta ID-nya masing-masing sebagai identitas gejala.

TABEL 2. DAFTAR GEJALA YANG DIGUNAKAN

ID Gejala	Nama Gejala
G1	Batuk berdarah
G2	Badan panas
G3	Berkeringat saat malam
G4	Mual
G5	Batuk berdarah
G6	Berat badan turun
G7	Badan lemas
G8	Badan gatal
G9	Batuk lebih dari 3 minggu
G10	Nafsu makan berkurang
G11	Nyeri ulu hati
G12	Nyeri punggung
G13	Muntah
G14	Penurunan kesadaran
G15	Sesak nafas
G16	Perut kembung
G17	Pusing
G18	Sakit Perut
G19	Nyeri lutut
G20	Sakit dada
G21	Mudah cape
G22	Sakit kepala
G23	Mencret
G24	Demam
G25	Sakit Buang Air Besar
G26	Feses berwarna hitam
G27	Sakit pinggang
G28	Nyeri menelan
G29	Muka anemis
G30	Timbul lesi di badan
G31	Kulit lecet dan basah
G32	Nyeri sendi

D. Kebutuhan Sistem

Sistem dibangun di atas platform Java menggunakan tools Netbeans, disertai DBMS Heidi SQL, XAMPP. Adapun berikut penjelasan kebutuhan sistem dari sisi masukan, keluaran, beserta cara mengukur performansi sistem.

1) Input Sistem

Input terhadap sistem berupa data keluhan atau gejala penyakit yang dialami oleh pasien. Pada penelitian ini, minimal ada 1 gejala yang harus diinputkan.

2) Output Sistem

Output yang dihasilkan oleh sistem adalah diagnosis penyakit yang diderita oleh pasien beserta probabilitasnya dideritanya penyakit tersebut.

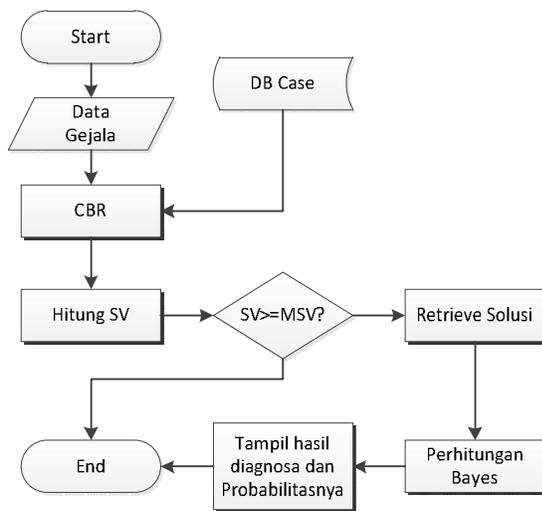
3) Performansi Sistem

Performansi sistem dihitung berdasarkan jumlah diagnosis benar yang dilakukan oleh sistem dibandingkan jumlah seluruh diagnosis yang harus dilakukan.

$$\text{Akurasi} = \frac{\text{jumlah diagnosis benar}}{\text{jumlah diagnosis benar} + \text{salah}} * 100\%$$

E. Cara Kerja Sistem

Gambar berikut menunjukkan cara kerja sistem mulai dari input gejala yang tampak pada pasien, hingga output sistem berupa jenis penyakit hasil diagnosis sistem.



Gambar 2. Flowchart sistem berbasis pengetahuan pendiagnosis TB

Untuk lebih jelasnya, berikut penjelasan lebih detail dari skema proses di Gambar 2.

1) Case Based Reasoning

Pada tahap pertama, dilakukan proses Case-Based Reasoning berdasarkan input gejala user. Pada proses ini sistem akan mengambil input gejala dari user, kemudian akan melakukan pengecekan pada database kasus. Apabila data pada database kasus memiliki Minimum Similarity Value (MSV)

dengan data input dari user, maka solusi akan dipilih dan akan dilanjutkan ke proses retrieve solusi.

2) Perhitungan SV

Proses ini adalah proses untuk melakukan perhitungan Similarity Value (SV). Proses ini dilakukan dengan cara mencocokkan karakter antara inputan yang dimasukkan oleh user dengan database kasus yang ada. Caranya adalah mengubah terlebih dahulu input gejala menjadi rentetan karakter 0 dan 1. Sebagai contoh, gejala G yang dipilih adalah G1, G7, dan G17. Maka, inputan tersebut akan diubah menjadi:

10000010000000000100000000000000

di mana 1 artinya gejala tersebut dipilih, sedangkan 0 berarti gejala tersebut tidak dipilih. Di sini terlihat bahwa setiap gejala memiliki pembobotan yang sama. Artinya, tidak ada gejala yang dianggap memiliki prioritas lebih tinggi dibandingkan gejala yang lain.

Selanjutnya, akan dilakukan pencocokan karakter dengan database kasus. Misal, salah satu penyakit A memiliki gejala:

10000010000000000001000000000000

maka setiap karakter tersebut akan dicocokkan. Jika karakter sama, maka akan diberi nilai 1, sedangkan jika tidak, akan diberikan nilai 0.

Dari gejala penyakit A yang dicocokkan dengan gejala G, perhitungan SV-nya adalah $30/32=0,9375$. Apabila MSV yang diset adalah 0,92, maka penyakit A tersebut akan lolos dari pengecekan nilai MSV terhadap gejala G untuk kemudian diproses lebih lanjut menggunakan Probabilitas Bayes.

Setelah proses kedua dilakukan, maka akan didapat solusi yang SV-nya lolos (lebih besar atau sama) terhadap nilai MSV. Biasanya, semakin kecil MSV maka akan semakin banyak solusi yang ter-retrieve. Tetapi, penyakit yang sama dianggap menjadi satu penyakit saja, dengan jumlah gejala yang diakumulasi.

Misalnya, ter-retrieve penyakit A sebanyak 2 kali, maka akan dianggap 1 penyakit A saja, tetapi gejalanya ditambahkan. Misal, pada penyakit A yang pertama gejalanya adalah G1,G2,G3, dan pada penyakit A yang kedua adalah G1,G2 saja, maka akan diakumulasikan menjadi G1 jumlahnya 2, G2 jumlahnya 2 dan G3 jumlahnya 1. Hal ini akan berpengaruh pada saat perhitungan Probabilitas Bayes.

3) Perhitungan Bayes

Pada tahap selanjutnya, akan dilakukan perhitungan probabilitas Bayes untuk mencari satu solusi terbaik yang kemudian akan ditampilkan kepada pasien. Berikut adalah contoh kasus sederhana yang dapat menunjukkan perhitungan untuk mendapatkan probabilitas Bayes tersebut.

Misalnyadari pasien P diinputkan gejala G1 dan G3. Setelah melewati perhitungan SV,diperoleh 2 penyakit, yaitu TB Paru dan TB Milier, di imana pada 2 penyakit tersebut masing-masing memiliki 10 gejala dengan rincian, pada penyakit TB Paru muncul G1 sebanyak 5, G2 sebanyak 2, dan G3 sebanyak 3. Sedangkan untuk TB Milier, muncul G1 sebanyak 3 dan G3 sebanyak 7.

TABEL 3. PERHITUNGAN KEMUNCULAN GEJALA

Gejala/Penyakit	TB Paru	TB Milier	Jumlah
G1	5/10	3/10	8/20
G2	2/10	0/10	2/20
G3	3/10	7/10	10/20
Jumlah	10/20	10/20	1

Dengan menggunakan rumus Probabilitas Bayes, maka akan diperoleh tabel sebagai berikut.

TABEL 4. PERHITUNGAN PROBABILITAS BAYES

Penyakit/Gejala	G1	G3
TB Paru	0,625	0,3
TB Milier	0,375	0,7

Karena gejala yang diinputkan dari pasien P adalah G1 dan G3, maka yang akan dihitung hanya G1 dan G3. Dari hasil tersebut, masing-masing probabilitas gejala G_i untuk setiap penyakit akan dikalikan. Berdasarkan cara perhitungan tersebut, akan didapatkan output sistem, yaitu probabilitas TB Milier sebesar 0,2625 dan probabilitas TB Paru sebesar 0,1875.

4) *Output sistem: Diagnosis Penyakit*

Hasil diagnosis yang ditampilkan adalah jenis penyakit dan probabilitas akan terjadinya penyakit tersebut. Jenis penyakit yang dipilih ialah yang memiliki probabilitas tertinggi dari hasil perhitungan sebelumnya. Misalnya, jika menggunakan perhitungan pada kasus di atas, maka output sistem adalah TB Milier dengan probabilitas sebesar 0,2625.

IV. PENGUJIAN DAN ANALISIS

Tujuan dari pengujian ini adalah untuk mengukur sejauh mana sistem berbasis pengetahuan yang menerapkan CBR ini dalam mendiagnosis penyakit TB. Menurut data dari WHO, kemampuan dokter yang baik dalam mendiagnosis TB adalah sekurang-kurangnya 85%. Di Indonesia sendiri, tingkat keakuratan dokter-dokternya dalam mendiagnosis TB sudah mencapai 90%. Dengan demikian, diharapkan sistem berbasis pengetahuan yang dibangun pada penelitian ini mampu menyamainya, yaitu mencapai angka 90%, atau minimal dapat memenuhi kualifikasi minimum sebagaimana yang telah ditetapkan oleh WHO, yakni sebesar 85%.

Total data yang digunakan pada pengujian kali ini adalah 120 baris data. Pengujian yang dilakukan di sini dibagi menjadi 2 skenario uji. Skenario yang pertama (A) adalah dengan menggunakan 80 data kasus dan 40 data uji. Sedangkan skenario yang kedua (B) adalah dengan menggunakan 90 data kasus dan 30 data uji. Kedua skenario tersebut bertujuan untuk menganalisis pengaruh porsi data kasus dan uji terhadap performansi sistem.

Untuk setiap skenario, dilakukan observasi terhadap variabel Minimum Similarity Value (MSV). Sebagaimana dijelaskan sebelumnya, nilai MSV dapat menentukan seberapa banyak rule penyakit yang ter-retrieve. Besar kecilnya jumlah tersebut dapat mempengaruhi performansi sistem. Observasi dilakukan untuk mendapatkan nilai MSV yang optimal terhadap akurasi sistem dalam mendiagnosis penyakit TB.

Dari pengamatan yang dilakukan sebelumnya terhadap data, diketahui bahwa pada satu penyakit minimal memiliki 1 gejala dan maksimal 7 gejala. Dengan demikian, observasi terhadap nilai MSV dilakukan dari (32-7)/32 sampai (32-1)/32 atau pada rentang nilai 0,78 sampai 0,96. Selama observasi, akan dipilih beberapa nilai MSV yang menghasilkan tingkat akurasi sistem terbaik.

A. Implementasi Sistem

Metode CBR yang digunakan pada kasus pendiagnosis penyakit TB ini diimplementasikan menggunakan bahasa Java. Melalui aplikasi yang dibuat, pengguna dapat memasukkan gejala-gejala penyakit yang muncul pada pasien. Setting nilai parameter untuk MSV juga dapat dilakukan pada aplikasi tersebut sebelum melakukan diagnosis.



Gambar 3. Flowchart sistem berbasis pengetahuan pendiagnosis TB

Setelah diagnosis dilakukan, diperoleh hasil diagnosis yang menunjukkan jenis penyakit beserta probabilitas masuknya ke jenis penyakit tersebut. Keluaran proses penalaran yang dilakukan oleh CBR ini tentunya tidak dapat menggantikan peran seorang dokter dan hanya menjadi bahan pertimbangan bagi dokter dalam melakukan diagnosis.

B. Hasil Observasi MSV

Dari hasil observasi MSV yang dilakukan, diperoleh 4 nilai MSV yang terbaik untuk masing-masing skenario sebagaimana ditunjukkan pada Tabel 5 dan Tabel 6.

TABEL 5. HASIL OBSERVASI MSV PADA SKENARIO A

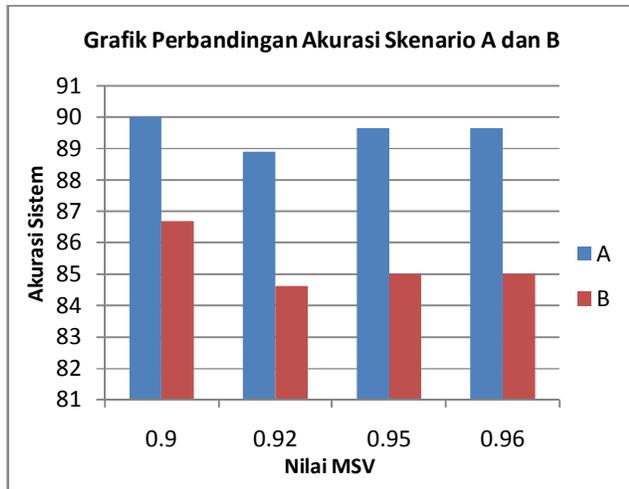
MSV	Benar	Salah	SV<MSV	Akurasi
0.90	36	4	0	0,9
0.96	26	3	11	0,896552
0.92	32	4	4	0,888889
0.95	26	3	11	0,896552

TABEL 6. HASIL OBSERVASI MSV PADA SKENARIO B

MSV	Benar	Salah	SV<MSV	Akurasi
0.90	26	4	0	0,8667

0.96	17	3	10	0,85
0.92	22	4	4	0,846154
0.95	17	3	10	0,85

Dari kedua tabel di atas, untuk masing-masing skenario diberikan nilai MSV yang sama yaitu 0,90; 0,92; 0,95; dan 0,96. Keempat nilai MSV tersebut menghasilkan tingkat akurasi yang bervariasi.



Gambar 4. Grafik perbandingan akurasi skenario A dan B

Grafik pada gambar 4 menunjukkan hasil perbandingan tingkat akurasi pada skenario A dan B dengan menggunakan keempat nilai MSV yang diujikan. Berdasarkan grafik tersebut, setidaknya ada 3 hal yang dapat dianalisis, di antaranya sebagai berikut.

Pertama, dari rentang nilai MSV yang diobservasi, yaitu dari 0,78 sampai 0,96, ternyata 4 nilai MSV terbaik yang menghasilkan akurasi sistem yang optimal untuk kedua skenario adalah sama, yaitu 0,9; 0,92; 0,95; dan 0,96. Hal ini menunjukkan bahwa secara umum dengan nilai-nilai tersebut solusi-solusi yang ter-retrieve oleh sistem tidak terlalu banyak, tidak pula terlalu sedikit. Artinya, calon-calon solusi yang ditawarkan sebagai output sistem dengan nilai MSV tersebut sudah dalam kondisi konvergen, yaitu mengarah pada solusi yang sesungguhnya. Sebaliknya, untuk nilai-nilai MSV di bawahnya, solusi yang ditawarkan terlalu divergen karena terlalu banyak solusi yang lolos atau ter-retrieve, sehingga membuat sistem menjadi lebih sulit untuk memutuskan keluarannya.

Kedua, di antara 4 nilai MSV terbaik, nilai MSV 0,9 merupakan yang paling optimal untuk diterapkan pada kasus pendiagnosis TB dengan data kasus yang diberikan. Meskipun perbedaannya tidak terlalu jauh dibandingkan ketiga nilai MSV lainnya, nilai tersebut mampu memberikan hasil terbaik untuk kedua skenario yang dirancang, yaitu pada skenario A dan B. Selain itu, nilai MSV tersebut mampu membuat sistem berbasis pengetahuan pendiagnosis TB ini menyamai rata-rata kualifikasi dokter di Indonesia dalam mendiagnosis TB.

Ketiga, secara keseluruhan, akurasi rata-rata yang dihasilkan pada skenario A lebih baik dibanding pada skenario B. Pada skenario A, banyaknya data kasus lebih sedikit daripada pada skenario B. Sebaliknya, pada skenario A, data uji yang digunakan lebih banyak daripada skenario B.

Pada saat dilakukan pengujian, sistem berbasis CBR akan belajar memperbaiki pengetahuannya (knowledge-base) dengan adanya data-data baru. Semakin banyak data uji, maka semakin banyak pula kesempatan sistem untuk mencapai performansi yang lebih baik. Hal inilah yang menyebabkan akurasi sistem pada skenario A lebih baik daripada skenario B. Pada skenario B, data uji yang digunakan masih belum cukup untuk membuat sistem mencapai performansi yang optimal.

Dengan demikian, dapat dikatakan bahwa jumlah data kasus yang digunakan penting guna membangun pengetahuan awal sistem. Namun, performansi sistem dapat menjadi lebih baik melalui proses pengujian karena sistem akan belajar dan memperbaiki basis pengetahuannya dari penambahan data kasus baru yang diberikan.

V. KESIMPULAN

Penerapan metode *Case-Based Reasoning* (CBR) untuk studi kasus diagnosis TB berhasil mencapai akurasi rata-rata lebih dari 85% dengan nilai MSV 0,90. Hasil yang diperoleh dari skenario A tersebut mampu menyamai rata-rata kualifikasi dokter di Indonesia, yaitu sebesar 90%, dalam mendiagnosis penyakit TB.

Jika dibandingkan dengan metode *Rule-Based Reasoning* (RBR), CBR mampu memperbaiki pengetahuan secara fleksibel dari setiap data baru yang diterimanya pada proses pengujian. Penambahan konsep probabilitas Bayes juga mempengaruhi performansi sistem karena mampu melakukan penanganan pada ketidakpastian dan kebenaran parsial, yaitu untuk data yang belum pernah ada di database kasus.

VI. SARAN

Pada penelitian kali ini, gejala yang dikenali dianggap memiliki prioritas yang sama karena peneliti belum memiliki informasi dari pakar atau dokter (*expert judgement*). Hal tersebut dapat menjadi bahan untuk pengembangan sistem pada penelitian selanjutnya. Pengetahuan pakar untuk melengkapi pembobotan setiap gejala diharapkan mampu membuat sistem menjadi lebih cerdas dan intuitif layaknya seorang dokter.

Untuk selanjutnya, perlu dilakukan penelitian untuk jumlah database kasus yang lebih besar agar basis pengetahuan sistem yang dibangun di awal dapat lebih menyeluruh dalam menanggapi berbagai variasi kasus serta lebih akurat. Metode CBR bisa jadi akan lebih baik pula jika dikombinasikan dengan metode lain seperti *fuzzy systems* atau jaringan syaraf tiruan.

DAFTAR PUSTAKA

- [1] Badiru, A. and Cheung Y. *Fuzzy Expert System with Neural Network Applications*. New York, USA: John Wiley & Sons, 2002.
- [2] Hartati dan Sri Iswanti S. *Sistem Berbasis Pengetahuan dan Pengembangannya*. Yogyakarta, Indonesia: Graha Ilmu, 2008.
- [3] Pal, S.K. dan S.C.K. Shiu. *Foundation of Soft Case-Based Reasoning*. Hoboken, New Jersey: John Wiley & Sons, 2004.

- [4] Kusumadewi, Sri. Artificial Intelligence (Teknik dan Aplikasinya). Yogyakarta: Graha Ilmu, 2003.
- [5] Unknown. Website Provinsi Yogyakarta [Online]. Diakses: April 2013. Tersedia: <http://www.plazainformasi.jogjaprov.go.id/index.php/tentangkami/layanan/daftar-artikel/1168-tbc.html>
- [6] Salam, M. and M.L. Sánchez. Adaptive Case-Based Reasoning using Retention and Forgetting Strategies. 2010.
- [7] Begum, Shahina. A Case-Based Reasoning System for The Diagnosis of Individual Sensitivity to Stress in Psychophysiology. 2009.
- [8] Widiastuti, W., Destiani D., dan J. Damiri. Aplikasi Sistem Berbasis Pengetahuan Deteksi Dini pada Penyakit Tuberculosis. Garut, Indonesia, 2012.
- [9] Kemenkes. Keputusan Kementerian Kesehatan tentang Penanganan Tuberculosis. Jakarta, Indonesia, 2009.
- [10] Giarratano, J. dan G. Riley. Expert Systems: Principles and Programming, 2nd edition. USA: PWS Publishing Co., 1994.
- [11] Ocampo, E., et al. Comparing Bayesian Inference and Case-Based Reasoning as Support Techniques in the Diagnosis of Acute Bacterial Meningitis. New York, USA, 2011.
- [12] Koton, P. Using Experience in Learning and Problem Solving. Massachusetts Institute of Technology, USA: Ph.D. Thesis MIT/LCS/TR-441, 1989.
- [13] Simpson, R. Computer Model of Case-Based Reasoning in Problem Solving: An Investigation in the Domain of Dispute Mediation. Atlanta, US: Technical Report GIT-ICS-85/18, Georgia Institute of Technology, School of Information and Computer Science, 1985.
- [14] Ahmed, M., S. Begum, P. Funk, and N. Xiong. Fuzzy Rule-Based Classification to Build Initial Case Library for Case-Based Stress Diagnosis. Innsbruck, Austria: Artificial Intelligence and Applications(AIA), 2009.
- [15] Bergmann, R. and A. Klaus-Dieter. Case-Based Reasoning Introduction and Recent Developments. Germany, 2009.
- [16] Aamodt, A. and E. Plaza. Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches. Norway: Published in AI Com - Artificial Intelligence Communications, Vol. 7, No. 1, pp. 1-2, 15-16, 1994.
- [17] Watson, I. Applying Case-Based Reasoning: Techniques for Enterprise Systems. California, USA: Morgan Kaufmann Publishers, Inc., 1997.
- [18] Ceccaroni, L. Integration of A Rule-Based Expert System, A Case-Based Reasoner and An Ontological Knowledge-Base in the Wastewater Domain. Berlin, Germany: 2nd ECAI Workshop on Binding Environmental Sciences and Artificial Intelligence (BESAI' 2000), pp. 8:1-8:10, 2000.
- [19] Magne, S., B. Robert, K. Uzay, L. and Hans. Similarity Measures in Fuzzy Rule Base Simplification. Eindhoven, Netherland: Eindhoven University of Technology, 1998.
- [20] Tiwari, M., and B. Mishra. Application of Cluster Analysis in Expert System. India: University Kanpur, 2011.
- [21] Richter, M. Modeling Uncertainty and Similarity-Based Reasoning-Challenges. Germany: 8th European Workshop on Uncertainty and Fuzziness in CBR, pp. 191-199, 2006.