

IMPLEMENTASI CASE BASED REASONING UNTUK MENDIAGNOSIS PENYAKIT TUBERKULOSIS MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBOR

Emanuel Tes Atok¹, Derwin R. Sina², Dony M. Sihotang³

^{1,2,3} Jurusan Ilmu Komputer, Fakultas Sains dan Teknik, Universitas Nusa Cendana

Email : derwinilkom@gmail.com¹, derwinilkom@gmail.com², dmsihotang99@gmail.com³

ABSTRAK

Penalaran Berbasis Kasus menghasilkan solusi berdasarkan kemiripan terhadap kasus-kasus yang pernah terjadi sebelumnya. Solusi kasus baru dihasilkan dari pencocokan kemiripan dengan kasus lama. Pada penelitian ini penulis menerapkan CBR untuk mendiagnosa penyakit tuberkulosa. Sumber pengetahuan sistem diperoleh dengan mengumpulkan berkas rekam medis pasien tuberkulosis pada tahun 2014-2016. Perhitungan nilai kemiripan menggunakan algoritma *K-Nearest Neighbor* dengan nilai batas kewajaran 80%. Sistem ini dapat mendiagnosa 3 jenis penyakit tuberkulosis berdasarkan 25 gejala yang ada. Luaran sistem berupa jenis penyakit tuberkulosis berdasarkan gejala yang dialami pasien, solusi pengobatan dan presentasi kemiripan antara kasus baru dan kasus lama. Berdasarkan hasil pengujian dengan 51 kasus TB didapatkan hasil: (a) pengujian dengan 3 skenario pengujian kasus baru didapatkan keakuratan sistem masing-masing untuk skenario pertama akurasi yang diperoleh dengan 31 data latih (60% dari 51 kasus) dan 20 data uji (40% dari 51 kasus) akurasinya sebesar 63%, skenario kedua akurasi yang diperoleh dengan 35 data latih (70% dari 51 kasus) dan 16 data uji (30% dari 51 kasus) akurasinya sebesar 69.2% dan skenario ketiga akurasi yg diperoleh dengan 41 data latih (80% dari 51 kasus) dan 10 data uji (20% dari 51 kasus) akurasinya sebesar 90%, (b) hasil pengujian terhadap kasus lama dalam basis kasus didapatkan keakuratan sistem sebesar 100%.

Kata Kunci : Penalaran Berbasis Kasus, Algoritma *K-Nearest Neighbor* (K-NN), Penyakit Tuberkulosis.

ABSTRACT

Case-Based Reasoning produces a solution based on similarities to previous cases. New case solutions result from the placement of similarities with old cases. In this research the authors applied CBR to diagnose tuberculosis. System knowledge sources are obtained by collecting medical records of tuberculosis patients in 2014-2016. Calculation of similarity values using the K-Nearest Neighbor algorithm with a threshold value of 80%. This system can diagnose 3 types of tuberculosis based on 25 symptoms. The system output consists of the type of tuberculosis based on the symptoms experienced by the patient, treatment solutions and presentation of similarities between new cases and old cases. Based on the results of testing with 51 cases the results: (a) testing with 3 new case scenarios obtained the accuracy of each system for data scenarios obtained by 31 training data (60% of 51 cases) and 20 test data (40% of 51 cases) accuracy is 63%, the second scenario accuracy obtained with 35 training data (70% of 51 cases) and 16 test data (30% of 51 cases) accuracy is 69.2% and the third scenario accuracy obtained with 41 training data (80% of 51 cases) and 10 test data (20% of 51 cases) accuracy is 90%. (b) The results of testing of the old cases in the case base obtained 100% accuracy of the system.

Keywords : Case-Based Reasoning, K-Nearest Neighbor (K-NN) Algorithm, Tuberculosis.

1 PENDAHULUAN

Kesehatan adalah keadaan sejahtera dari badan, jiwa dan sosial yang memungkinkan setiap orang hidup produktif secara sosial dan ekonomis. Terdapat beberapa faktor yang mempengaruhi tingkat kesehatan seseorang seperti gaya hidup, pola makan dan lingkungan. Penyakit tuberkulosis (TB) adalah suatu penyakit infeksi yang disebabkan oleh bakteri *Mikobakterium tuberkulosa*. Berdasarkan data dari dinas kesehatan provinsi Nusa Tenggara Timur, penyakit TB tercatat sebanyak 4789 kasus pada tahun 2015. Khusus untuk kabupaten Belu jumlah kasus TB yang tercatat pada tahun 2014 sebanyak 439 kasus sedangkan pada tahun 2015 mencapai 560 kasus, dimana angka ini menunjukkan peningkatan dari tahun sebelumnya.

Masalah kesehatan seperti TB tentunya hanya dapat diatasi dengan bantuan petugas medis. Terkait dengan penanganan kasus TB ini hanya bisa dilakukan oleh dokter agar ketepatan mendiagnosa penyakit TB lebih akurat. Permasalahan yang sering muncul yaitu susahny menemukan tenaga medis dalam hal ini dokter, terlebih di puskesmas peskemas pembantu. Permasalahan ini terjadi karena jam kerja dari dokter yang terbatas serta mahalny biaya jika harus melakukan pengobatan di dokter praktek. Berdasarkan persoalan tersebut, munculah sebuah konsep penalaran berbasis kasus (*case based reasoning*). Sistem CBR (*case based reasoning*) ini menghasilkan kesimpulan berdasarkan kemiripan terhadap kasus-kasus yang pernah terjadi sebelumnya.

Penelitian terdahulu menggunakan algoritma *k-nearest neighbor* untuk pencocokan kasus yang mirip pada diagnosis penyakit leukemia [1]. Kemudian pada kelainan neuropsikiatri[2]. Penelitian yang dilakukan oleh Nurdiansa tahun 2013 [3] serta Moh. Zainuddindkk [4].

2 MATERI DAN METODE

2.1 Data Penelitian

Data rekam medis yang dipakai adalah data penyakit tuberkulosis tahun 2014 sampai dengan 2016 berjumlah 51 kasus yang terdapat di Rumah Sakit Umum Daerah Mgr. Gabriel Manek Atambua.

2.2 Penalaran Berbasis Kasus/*Case Based Reasoning* (CBR)

Penalaran Berbasis Kasus/*Case-Based Reasoning* adalah suatu pendekatan untuk menyelesaikan suatu permasalahan (*problem solving*) berdasarkan solusi dari permasalahan sebelumnya[5]. CBR atau penalaran berbasis kasus merupakan salah satu teknik pemecahan masalah berdasarkan pengalaman pada masa lalu.

2.2.1. Retrieve dalam CBR

Retrieve adalah pengambilan kembali (*retrieval*) kasus-kasus sebelumnya untuk menyelesaikan masalah baru. Proses *retrieve* akan melakukan dua langkah pemrosesan, yaitu pengenalan masalah dan penilaian kesamaan masalah pada basis kasus.

2.2.2. Reuse dalam CBR

Proses *reuse* dalam siklus CBR adalah tanggung jawab dalam memberikan solusi dari sebuah kasus baru berdasarkan penyelesaian-penyelesaian kasus yang diambil kembali.

2.2.3. Revise dalam CBR

Tahap ini mengubah dan mengadopsi solusi yang ditawarkan jika perlu, dua tugas utama dari tahapan ini adalah evaluasi solusi dan memperbaiki kesalahan.

2.2.4. Retain dalam CBR

Retain merupakan tahap terakhir dalam siklus CBR yang menghasilkan penyelesaian masalah terbaru yang digabung dalam sistem pengetahuan. telah diterjemahkan menjadi berbagai pendekatan untuk merekam hasil dari penyelesaian masalah sebagai sebuah kasus baru dan dapat ditambahkan dalam basis kasus.

2.3 Indexing

Indeks adalah struktur data yang mengatur *record* data pada disk untuk mengoptimalkan beberapa jenis operasi pengambilan (*retrieval*) tertentu. Indeks memungkinkan untuk secara efektif mengambil semua *record* yang memenuhi syarat pencarian pada *fieldsearch key* pada indeks [6].

2.4 Algoritma *K-Nearest Neighbor*

Algoritma *nearest neighbor* merupakan pendekatan untuk mencari kasus dengan kedekatan antara kasus baru dengan kasus lama, yaitu berdasarkan pada kecocokan bobot sejumlah fitur yang ada [7]. Metode ini mencari jarak terhadap tujuan dari data yang telah disimpan sebelumnya. Setelah didapatkan jaraknya kemudian dicari jarak terdekat.

Rumus untuk menghitung bobot kemiripan (*similarity*) antara kasus baru dan kasus lama yang ada dalam basis kasus ditunjukkan dalam persamaan 2.1 yg diperkenalkan oleh Oktaviani sebagai berikut:

$$\text{Similarity}(\text{problem}, \text{case}) = \frac{s_1 * w_1 + s_2 * w_2 + \dots + s_n * w_n}{w_1 + w_2 + \dots + w_n} \dots \dots \dots (2.1)$$

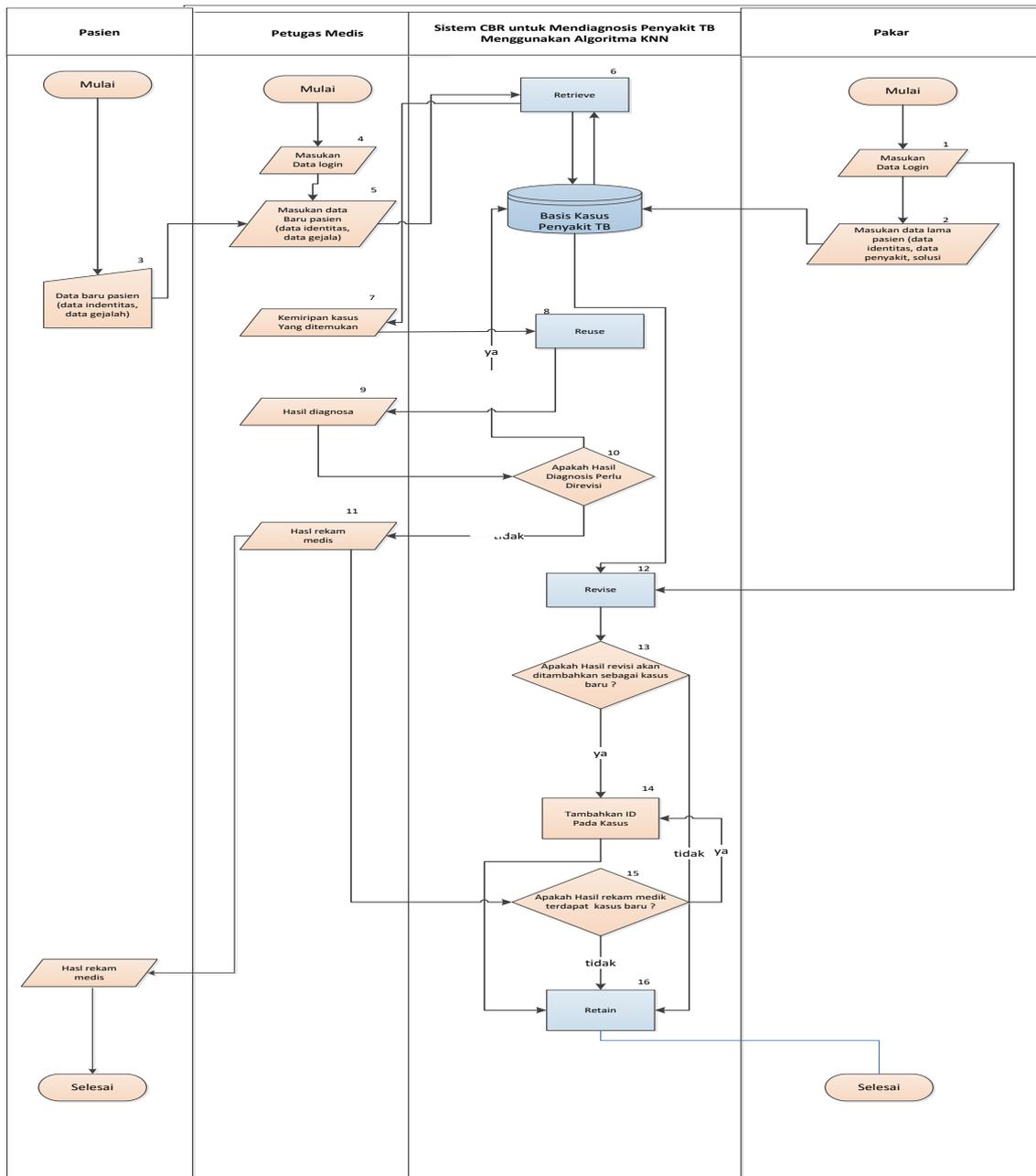
Keterangan:

s = *similarity*(nilai kemiripan). Jika gejala pada kasus baru terdapat juga pada kasus lama maka s diberi nilai 1, bila tidak maka s diberi nilai 0.

w = *weight* (bobot yang diberikan untuk gejala)

2.5. Tahap Perancangan Sistem CBR

Flowchart sistem CBR untuk mendiagnosis penyakit Tuberkulosis dimulai dari pasien memberikan data diri dan data gejala kepada petugas medis. Petugas medis melakukan login dan memasukkan data baru kasus pasien (data identitas, data gejala) ke sistem CBR untuk diolah. Selanjutnya sistem akan memberikan hasil diagnosis yang tepat berdasarkan kesamaan dengan kasus lama pada basis kasus. Jika informasi mengenai hasil diagnosis tidak direvisi berarti kasus tersebut memiliki nilai kemiripan lebih besar atau sama dengan nilai *threshold* yaitu 80% sedangkan jika lebih kecil dari 80% maka kasus akan direvisi oleh pakar.



Gambar 1 Flowchart sistem CBR diagnosis Penyakit Tuberkulosis

3 HASIL DAN PEMBAHASAN

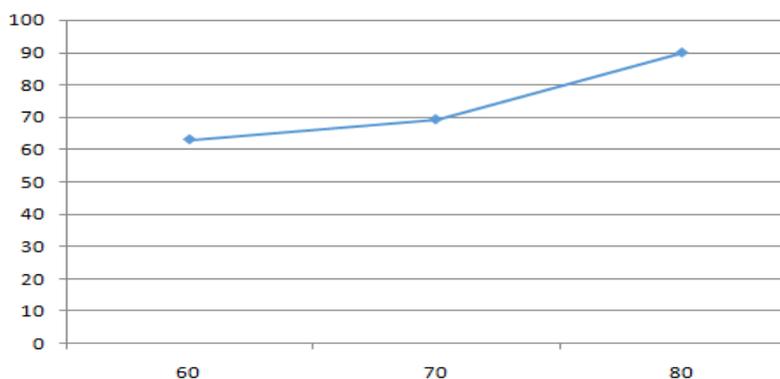
3.1 Pembahasan

Langkah awal dalam pembahasan yaitu pengujian sistem yang bertujuan untuk mengukur keakuratan sistem dalam mendiagnosa penyakit TB. Dalam pengujian sistem ini diterapkan pengujian *confusion matrix* dengan dua skenario pengujian yaitu pengujian sistem dengan menggunakan kasus yang tersimpan dalam basis kasus dan yang kedua pengujian sistem dengan menggunakan kasus baru yang belum pernah tersimpan dalam basis kasus.

3.1.1. Pengujian Sistem

Pengujian sistem CBR ini diberikan 3 cara pengujian dan setiap cara akan dibandingkan dengan semua kasus yang terdapat dalam basis kasus. Pengujian cara pertama dilakukan dengan memasukan kasus baru yang mempunyai gejala yang sama persis dengan salah satu kasus dalam *case base*(basis kasus) melalui empat tahapan sistem *CBR*. Nilai akurasi yang didapat sebesar 100%.

Pengujian cara kedua dilakukan dengan cara menguji sistem CBR dengan kasus baru dan nilai akurasi sistemnya dihitung menggunakan *confusion matrix*. Pengujian ini menerapkan 3 skenario pengujian yaitu skenario pertama 60% (31 kasus TB) data latih dan 40% (20 kasus TB)+10 kasus non TB data uji, skenario kedua 70% (35 kasus TB) data latih dan 30% (16 kasus TB) +10 kasus non TB data uji, dan skenario ketiga 80% (41 kasus TB) data latih dan 20% (10 kasus TB) +10 kasus non TB data uji. Secara berurutan nilai akurasi yang didapat dari ketiga skenario pengujian ini yaitu 63%, 69.2% dan 90%. Dari pengujian cara kedua ini dapat disimpulkan bahwa, makin banyak data latih dalam basis kasus maka baik pula nilai akurasi yang diperoleh. Grafik pengaruh data latih terhadap nilai akurasi dapat dilihat pada gambar 2.



Gambar 2 Pengaruh jumlah data latih terhadap nilai akurasi

Pengujian cara ketiga Sistem CBR diuji dengan 6 kasus berbeda yang sudah tersimpan sebelumnya dalam basis kasus. Keenam kasus berbeda tersebut terdiri dari dua jenis TB paru, dua jenis TB milier, dua jenis TB meningitis. Pengujian ini bertujuan untuk mengetahui apakah sistem mampu mengambil kembali ketiga jenis TB pada kasus lama tersebut sehingga dapat memberikan hasil diagnosis dan solusi yang sesuai. Pengujian kasus ke-1, ke-2, ke-3, ke-4, ke-5 dan ke-6 pada kasus lama menghasilkan nilai kemiripan yang sama yaitu 100%. Hal ini membuktikan bahwa sistem CBR mampu diimplementasikan untuk mendiagnosa penyakit TB dengan baik.

4 KESIMPULAN DAN SARAN

4.1 Kesimpulan

Berdasarkan hasil penelitian *Case Based Reasoning* untuk mendiagnosa penyakit TB menggunakan algoritma *K-Nearest Neighbor*, maka dapat disimpulkan bahwa:

1. Sistem CBR untuk mendiagnosa penyakit TB dengan menggunakan algoritma *K-NN* diuji dengan kasus baru yang tidak pernah tersimpan dalam basis kasus. Pengujian menggunakan kasus baru dilakukan dengan tiga skenario pengujian yaitu yang pertama dengan 40% data uji TB (20 kasus) +10 kasus non TB, 60% data latih dan kedua 30% data uji TB (16 kasus) +10 kasus non TB, 70% data latih dan skenario yang ketiga dengan 20% data uji TB (10 kasus) +10 kasus non TB, 80% data

latih. Berdasarkan ketiga skenario pengujian ini didapat masing masing nilai akurasi yang berbeda yaitu untuk skenario pertama didapat nilai akurasi sebesar 0.63 (63%), akurasi skenario kedua 0.692 (69,2%) dan skenario ketiga didapat nilai akurasi sebesar 0.90 (90%). Dari ketiga skenario ini dapat diambil kesimpulan bahwa makin banyak data latih dalam basis kasus, maka semakin baik nilai akurasi yang diperoleh.

2. Sistem CBR untuk mendiagnosa penyakit TB dengan menggunakan algoritma *K-NN* diuji dengan kasus lama yg suda tersimpan sebelumnya dalam basis kasus. Pengujian dengan kasus lama ini menggunakan 6 data uji, dari keenam data uji yang digunakan, semuanya mendapatkan nilai similarity sebesar 100% atau kemiripan yang sama persis dengan kasus yang tersimpan dalam basis kasus.

4.2. Saran

Saran yang dapat diberikan dari hasil penelitian adalah sebagai berikut:

1. Penelitian selanjutnya dapat mengembangkan sistem dengan menggunakan metode lain seperti *cosine similarity* dan *block city* untuk menghasilkan penelitian-penelitian yang bermanfaat.
2. Proses pembobotan dalam penelitian ini adalah pembobotan yang dilakukan oleh satu pakar dan tentunya bersifat subjektif, sehingga untuk penelitian selanjutnya disarankan pembobotan dilakukan oleh beberapa pakar dan diambil rata-rata agar diperoleh pembobotan yang lebih baik.

DAFTAR PUSTAKA

- [1] Aribowo., Khomsah., 2012, Penalaran Berbasis Kasus untuk Deteksi Dini Penyakit Leukemia, *jurnal SemnasIF*, UPN Yokyakarta.
- [2] Khatoon, S. and Agarwal, K., 2014, Case-Based Reasoning System for Diagnosis of Neuropsychiatric Abnormality, *International Journal of Computational Engineering Research (IJCER)*, Deptt of CSE, Integral University.Lucknow.
- [3] Nurdiansyah, Y., 2013, Case-Based Reasoning Untuk Pendukung Diagnosa Gangguan Pada Anak Autis, *Informatics Journal*, Universitas Jember.
- [4] Zainuddin, M, dkk., 2016, penerapan case base reasoning untuk mendiagnosa penyakit stroke menggunakan algoritma k-nn, *jurnal CITISEE*, STMIK Bumigora, Mataram.
- [5] Aamodt A., Plaza E., 1994, Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches, *AICom - Artificial Intelligence Communications*, IOS Press, Vol. 7: 1, pp. 39-59.
- [6] Labellapansa, A., 2007, Penalaran Berbasis Kasus Untuk Mengetahui Biaya Sewa Kost di Yogyakarta, *Jurnal Ilmu Komputer*, Yogyakarta.
- [7] Kusrini dan Emha Taufiq Luthfi, 2009, *Algoritma Data Mining*, C.V Andi Offset, Yogyakarta.