

Case Based Reasoning Diagnosis Penyakit Difteri dengan Algoritma K-Nearest Neighbor

Case Based Reasoning Diagnosis of Diphtheria with K-Nearest Neighbor Algorithm

Chavid Syukri Fatoni*¹, Friandy Dwi Noviandha²

^{1,2}Magister Teknik Informatika, Universitas AMIKOM Yogyakarta

E-mail: *¹fatoni.work@gmail.com, ²frechandhy@gmail.com

Abstrak

Akhir tahun 2017, masyarakat Indonesia ramai dengan maraknya kematian pada anak-anak dan orang dewasa akibat penyakit Difteri. Ditemukan sebanyak 12 orang meninggal dunia dari 318 kasus Difteri menurut catatan Dinas Kesehatan Jawa Timur. Padahal di tahun 2016 kasus Difteri di Jawa Timur tercatat sebanyak 4 orang meninggal dunia dari 209 kasus. Hal tersebut menjadi perhatian bagi pemerintah dan tercatat sebagai kejadian luar biasa (KLB). Kenaikan angka kasus Difteri ini disebabkan karena kurangnya kesadaran masyarakat akan pentingnya imunisasi. Semakin banyaknya kasus Difteri yang terjadi dan minimnya pengetahuan masyarakat tentang Difteri, maka dibutuhkan suatu sistem pakar yang mampu membantu masyarakat maupun pemerintah dalam mendiagnosis penyakit Difteri. Penelitian mengenai Difteri ini menggunakan metode algoritma K-Nearest Neighbour (K-NN) dimana dilakukan perhitungan similaritas pada kasus lama dengan kasus baru. Penelitian penyakit Difteri ini disempurnakan dengan menggunakan penalaran berbasis kasus atau Cased Based Reasoning (CBR) agar hasil diagnosis lebih akurat. Output dari penelitian ini yaitu berupa hasil diagnosa penyakit Difteri berdasarkan gejala-gejala yang dialami dengan hasil akurasi pengujiannya sebesar 95,17%.

Kata Kunci — Difteri, Sistem Pakar, CBR, K-NN.

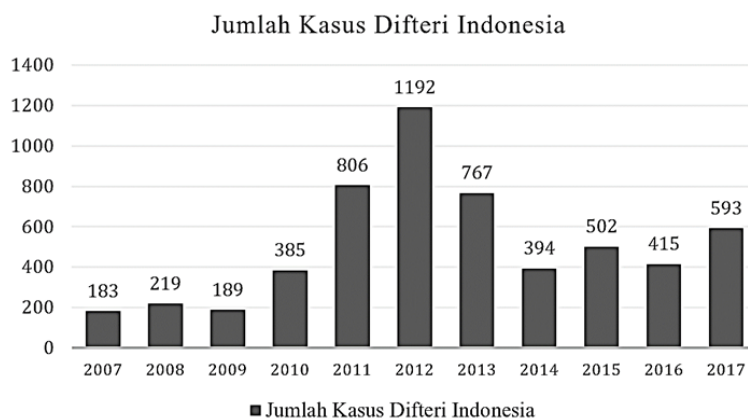
Abstract

End of 2017, the people of Indonesia enlivened so many of deaths in children and adults due to Diphtheria. Found 12 people died from 318 cases of Diphtheria according to East Java Health Office records. Whereas in the year 2016 Diphtheria cases in East Java recorded and reported as many as 4 people died from 209 cases. It's of particular concern to government and is noted as an extraordinary event (KLB). The increase in the number of Diphtheria cases is due to a lack of public awareness of the importance of immunization. Increasing number of Diphtheria cases and the lack of public knowledge about Diphtheria, it needs an expert system capable of assisting the public and the government in diagnosing Diphtheria. This research on Diphtheria uses the K-Nearest Neighbors (K-NN) algorithm method in which a similarity case study in the old case with new cases is used. The research of Diphtheria disease is enhanced by using case-based reasoning or Cased Based Reasoning (CBR) to make the diagnosis more accurate. The output of this research is the result of diagnosis of Diphtheria disease based on the symptoms experienced by the result of the accuracy of the test is 95,17%.

Keywords — Difteri, Expert System, CBR, K-NN.

1. PENDAHULUAN

Menurut Kementerian Kesehatan, penyakit Difteri merupakan penyakit menular mematikan yang menyerang saluran pernafasan bagian atas (tonsil, faring dan hidung) dan kadang pada selaput lendir dan kulit. Penyakit ini disebabkan oleh bakteri yaitu *Corynebacterium diphteriae*. Semua golongan umur bisa tertular, namun anak usia kurang dari 5 tahun dan orang tua diatas 60 tahun sangat beresiko tertular penyakit Difteri. Sangat perlu diwaspadai bahwa kecenderungan jumlah kasus penyakit ini meningkat sejak 2007 (183 kasus) dan puncaknya pada 2012 (1.192 kasus). Setelah itu menurun tapi angkanya masih ratusan kasus. Informasi dari Kementerian Kesehatan tentang jumlah kasus Difteri di Indonesia ditunjukkan pada Gambar 1 [1].



Gambar 1. Grafik Jumlah Kasus Difteri Indonesia

Mengingat Difteri merupakan penyakit lama yang mewabah kembali, sistem pakar yang dapat mendiagnosis penyakit Difteri sangat diperlukan dengan pendekatan Cased Based Reasoning (CBR). Metode CBR ini bekerja dengan cara menyelesaikan masalah sejenis yang pernah terjadi di masa lampau kemudian mengadopsi informasi serta solusi yang pernah digunakan untuk menyelesaikan masalah yang baru. K-Nearest Neighbour (K-NN) bekerja dengan cara mengklasifikasikan objek berdasarkan data pembelajaran yang jarak paling dekat dengan objek tersebut. Solusi terhadap objek atau masalah yang digunakan pada algoritma ini diambil dengan ketentuan kemiripan antara 0,8 (80%) hingga 1 (100%).

Penelitian tentang penyakit Difteri menggunakan penggabungan Cased Based Reasoning dan K-Nearest Neighbour tersebut, diharapkan bisa membantu dalam mengenali serta mendeteksi secara dini penyakit Difteri berdasarkan gejala-gejala yang dialami serta memberikan pengetahuan tentang bagaimana menangani penyakit menular tersebut, selain itu juga dapat membantu pembaca untuk penggabungan Cased Based Reasoning dan K-Nearest Neighbour menjadi referensi untuk penelitian selanjutnya.

Penelitian yang pertama menjelaskan, berdasarkan hasil pengujian sistem CBR terhadap pakar untuk 15 kasus yang diuji sistem mampu mendiagnosis dengan tepat sesuai dengan hasil diagnosis pakar sebenarnya sebesar 93.3%, namun dari 15 kasus data yang diuji belum dipaparkan untuk jumlah kasus dalam data training yang digunakan untuk pengujian similarity dari 15 kasus [2].

Penelitian yang kedua sistem dibangun dengan menerapkan pengetahuan-pengetahuan lama (Case Based Reasoning) sebagai pedoman menyelesaikan masalah diagnosa penyakit pencernaan manusia, namun belum dipaparkan untuk pengujian data dengan berapa jumlah data uji serta data training yang digunakan [3].

Penelitian yang ketiga sistem pakar ini mampu mengidentifikasi penyakit Demam dengan menginputkan 15 gejala dan kemudian dilakukan proses perhitungan menggunakan metode K-Nearest Neighbor – Certainty. Dalam penelitian tersebut K-Nearest Neighbor – Certainty Factor diperlukan data latih yang besar sehingga hasil lebih efektif dan tangguh terhadap training data

yang noisy, namun apabila menerapkan konsep Case Based Reasoning maka data training akan selalu bertambah dan diperbarui dan disimpan dalam knowledge base sehingga K-Nearest Neighbor akan lebih optimal [4].

Penelitian yang keempat yaitu menyelidiki efisiensi algoritma K-Nearest Neighbor untuk Text Categorization. Dapat disimpulkan model yang digunakan untuk mewakili dan memilih dokumen yang relevan didasarkan pada kata-kata. Namun, penggunaan kata-kata untuk mewakili isi dokumen masih menimbulkan dua masalah, yakni ambiguitas semantik kata-kata menyiratkan bahwa dokumentasi pelatihan yang tidak relevan, berbagi kata-kata yang sama dengan yang ingin kita klasifikasikan nanti dipilih bisa meningkatkan noise, berbeda dengan kasus diagnosa penyakit yang sudah jelas atribut yang digunakan merupakan hasil analisa pakar dan dari data rekam medis [5].

Penelitian yang kelima menerapkan 60 data rekam medis tentang penyakit usus yang diambil dari RSUD dr. Soetrasno Rembang. Dengan skenario 40 kasus sumber dan 20 kasus sasaran, dapat diketahui bahwa keakuratan diagnosis sistem sebesar 95%. Namun, dalam penelitian tersebut belum menerapkan jumlah data yang lebih banyak, serta atribut yang digunakan semua merupakan gejala-gejala yang berhubungan belum masuk dalam data karakteristik penyakit, dari 60 data tersebut diklasifikasikan menjadi 5 penyakit dasar pada usus belum masuk pada komplikasi penyakit [6].

Penelitian yang keenam membahas tentang jumlah mekanik yang terbatas dan minimnya pengetahuan pengguna menyebabkan berbagai kesulitan dalam perawatannya khususnya dalam mengatasi kerusakan mesin. Pembuatan aplikasi sistem cerdas melalui metode case-based reasoning dapat memberikan kemudahan melakukan diagnosis awal secara mandiri, namun dalam penelitian tersebut belum menjelaskan jumlah data training yang digunakan untuk membuat perbandingan similarity untuk menyelesaikan kasus kerusakan mesin, serta belum dijelaskan berapa tingkat akurasi yang diperoleh dalam penelitian tersebut [7].

Penelitian yang ketujuh membahas tentang sistem pakar diagnosa Penyakit Pulmonary Tuberculosis menggunakan algoritma Certainty Factor. Konsep yang digunakan dalam penelitian tersebut yaitu dengan mengumpulkan gejala-gejala yang menimbulkan penyakit tersebut, kemudian merumuskan suatu aturan-aturan atau rule yang disimpan dalam knowledge base, namun dalam penelitian tersebut belum dipadukan dengan metode penalaran berbasis kasus seperti case-based reasoning yang dapat membantu memperkaya knowledge base untuk memberikan keputusan yang lebih akurat [8].

2. METODE PENELITIAN

Penelitian memiliki beberapa tahapan mulai dari studi literatur sampai metode eksperimen dan melalui proses pengujian similarity. Metode perancangan sistem yang digunakan yaitu Reuse-Based memiliki 6 tahapan dalam proses perancangan sistem antara lain pengumpulan kebutuhan spesifikasi sistem, analisa kebutuhan komponen, memodifikasi spesifikasi sistem, menggabungkan desain sistem dengan Reuse-Based, pengembangan dari penggabungan sistem, serta melakukan proses validasi dari sistem. Pendekatan konsep tersebut memungkinkan untuk pengambilan komponen yang dapat digunakan ulang dan tergantung seberapa besar komponen yang dapat digunakan ulang serta di integrasikan dengan konsep komponen-komponen dalam perangkat lunak, dipaparkan pada Gambar 2 [9].



Gambar 2. Metode Perancangan Reuse-Based

2.1. Sistem Pakar (Expert System)

Sistem pakar merupakan cabang dari Artificial Intelligence (AI) yang cukup tua karena sistem ini mulai dikembangkan pada pertengahan 1960. Sistem pakar yang muncul pertama kali adalah General Purpose Problem Solver (BPS) yang dikembangkan oleh Newel dan Simon. Sistem Pakar adalah sistem yang berusaha mengadopsi pengetahuan para pakar dalam menyelesaikan permasalahan berbasis sistem komputer. Menurut Turban dan Aronson “sistem pakar merupakan sistem yang menggunakan pengetahuan manusia yang dimasukkan kedalam komputer untuk memecahkan permasalahan yang biasanya diselesaikan oleh pakar” [10]. Dibawah ini merupakan komponen yang dibutuhkan dalam membangun sistem pakar, antara lain adalah:

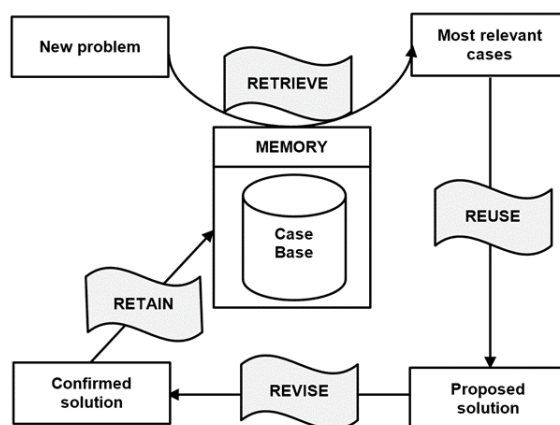
1. Memori Kerja (Working Memory)
2. Antar Muka Pengguna (User Interface)
3. Mekanisme Inferensi (Inference Machine)
4. Basis Pengetahuan (Knowledge Base)

2.2. Case Based Reasoning

Case Based Reasoning (CBR) adalah metode untuk menyelesaikan masalah dengan mengingat kejadian-kejadian yang sama/sejenis (similarity) yang pernah terjadi di masa lalu kemudian menggunakan pengetahuan atau informasi tersebut untuk menyelesaikan masalah yang baru, atau dengan kata lain menyelesaikan masalah dengan mengadopsi solusi-solusi yang pernah digunakan di masa lalu [11]. Terdapat 4 tahapan proses dalam sistem penalaran komputer berbasis kasus yaitu [12]:

1. Retrive, mendapatkan kasus-kasus yang mirip.
2. Reuse, menggunakan kembali kasus-kasus yang ada dan dicoba untuk menyelesaikan suatu masalah sekarang.
3. Revise, merubah dan mengadopsi solusi yang ditawarkan jika perlu.
4. Retain, memakai solusi baru sebagai bagian dari kasus baru, kemudian kasus baru diperbaharui ke dalam basis kasus.

Dari 4 tahapan proses dalam sistem penalaran komputer berbasis kasus diatas dapat dipaparkan pada Gambar 3 [13].



Gambar 3. Tahapan Proses Case Based Reasoning

2.3. *K-Nearest Neighbor (K-NN)*

K-Nearest Neighbor adalah sebuah algoritma untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Algoritma K-Nearest Neighbor mengelompokkan data baru yang belum diketahui class-nya dengan memilih data sejumlah k yang letaknya paling dekat dari data baru. Class paling banyak dari data terdekat sejumlah k akan dipilih sebagai class yang diprediksi untuk data baru. Pada umumnya nilai k menggunakan jumlah ganjil agar tidak terdapat jarak yang sama dalam proses klasifikasi. Jauh atau dekatnya tetangga dihitung menggunakan jarak Euclidean [4].

Kasus khusus di mana klasifikasi diprediksi berdasarkan data pembelajaran yang paling dekat (dengan kata lain, $k = 1$) disebut algoritma K-Nearest Neighbor. Rumus untuk menghitung bobot kemiripan (similarity) dengan K-Nearest Neighbor [14] ditunjukkan pada persamaan (1).

$$\text{Similarity}(T, S) = \frac{\sum_{i=1}^n f(T_i, S_i) \times W_i}{W_i} \quad (1)$$

Keterangan:

T : Kasus baru

S : Kasus yang ada dalam penyimpanan

n : Jumlah atribut dalam setiap kasus

i : Atribut individu antara 1 s.d. n

f : Fungsi *similarity* atribut i antara kasus T dan kasus S

w : Bobot yang diberikan antara atribut ke- i

3. HASIL DAN PEMBAHASAN

3.1. *Analisa Desain Sistem*

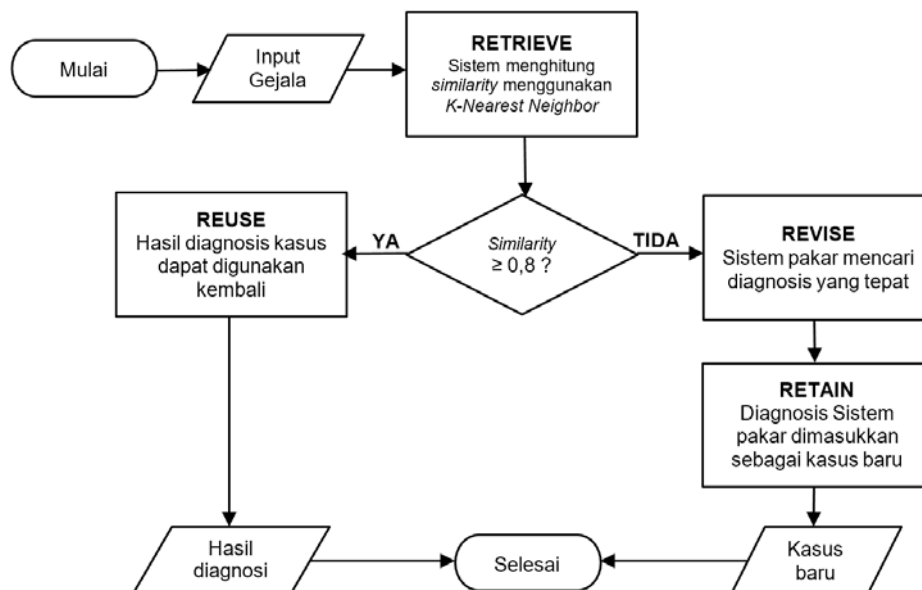
Proses analisa desain sistem sesuai dengan langkah-langkah dalam alur pembuatan sistem pakar, berikut alur proses analisa desain sistem dipaparkan pada Gambar 4.



Gambar 4. Alur Proses Analisa Desain Sistem

Proses analisa desain sistem diawali dengan pengumpulan data-data terkait dengan penyakit Difteri. Data yang dikumpulkan bisa menggunakan data dari berbagai sumber valid yang membahas tentang penyakit Difteri. Demi melengkapi data yang dikumpulkan, perlu juga mendapatkan informasi tentang penyakit Difteri dari narasumber yang mengetahui tentang penyakit tersebut. Narasumber yang cocok untuk melakukan wawancara yaitu seorang Dokter Spesialis. Apabila data dan informasi telah lengkap dan mencukupi, maka masuk ke tahap analisis penyakit Difteri menggunakan metode-metode yang telah ditetapkan hingga didapatkan sistem yang bisa mendiagnosis berdasarkan gejala yang dialami penderita serta penanganan yang tepat, selanjutnya langkah terakhir yaitu pengujian sistem pada kasus penyakit Difteri.

Sebelum masuk ke tahap proses analisa desain sistem, maka tahap selanjutnya yaitu membuat diagram alur sistem itu sendiri. Diagram alur sistem ini dibuat untuk memudahkan dalam memberikan gambaran proses inti dalam sistem analisa penyakit Difteri. Berikut diagram alur sistem dipaparkan pada Gambar 5.



Gambar 5. Diagram Alur Sistem (flowchart)

Pada diagram alur sistem, digambarkan bahwa langkah awal yang harus dilakukan yaitu menginput gejala yang dialami penderita ke dalam sistem. Kemudian dari gejala yang diinputkan, sistem akan melakukan proses Retrieve (menghitung kesamaan gejala menggunakan algoritma K-NN). Apabila similarity gejala hasilnya $\geq 0,8$ maka sistem akan melakukan proses Reuse data kasus lama untuk digunakan kembali sehingga langsung didapatkan hasil diagnosanya. Sedangkan apabila similarity gejalanya hasilnya $\leq 0,8$ maka sistem akan melakukan proses Revise dan Retain untuk menjadikan gejala penderita sebagai kasus baru. Kasus baru tersebut akan menambah kecerdasan sistem karena memperkaya sumber kasus pada sistem.

3.2. Pengumpulan Data

Proses awal analisa desain sistem yaitu dengan pengumpulan data. Data yang digunakan bisa dari beberapa sumber. Sumber data penelitian ini merupakan data rekam medis dari 148 pasien di tahun 2012, data-data pasien yang menderita penyakit Difteri di RSUD Dr. Soetomo Surabaya. Data yang dapat digunakan pada sistem ini adalah data gejala yang dialami pasien RSUD Dr. Soetomo Surabaya dan hasil diagnosa dari dokter yang bertanggung jawab. Berikut adalah data-data yang sudah dirangkum dan dipaparkan pada Tabel 1 dan Tabel 2 [15].

Tabel 1. Data Karakteristik 148 Penderita Difteri

Karakteristik	Jumlah Kasus <i>n</i> (%)
Usia (tahun) mean \pm SD (range)	5,11 \pm 0,24
Jenis Kelamin (Laki-laki : Perempuan)	1,14:1
Lama Sakit Sebelum MRS (hari)	
a. Mean	4,96 \pm 2,41
b. 0 – < 5	77 (52)
c. \geq 5	71 (48)
Tempat Tinggal	
a. Surabaya	84 (56,8)
b. Luar Surabaya	64 (43,2)
Derajat Difteri	
a. Difteri ringan	2 (1,4%)
b. Difteri sedang	62 (41,9%)
c. Difteri berat	84 (56,7%)

Tabel 1. Lanjutan.

Karakteristik	Jumlah Kasus <i>n</i> (%)
Status Imunisasi DPT	
a. <i>Unvaccinated</i>	23 (15,5)
b. <i>Vaccinated</i>	125 (84,5)
Status Gizi	
a. Buruk	6 (4,1)
b. Kurang atau Baik	142 (95,9)

Data karakteristik dari 148 penderita Difteri menunjukkan data jenis kelamin terdiri dari 79 (53,4%) laki-laki dan 69 (46,6%) perempuan. Terdapat 91 (61,5%) penderita berusia kurang dari 5 tahun, 47 (31,8%) berusia 5–10 tahun, dan 10 (6,7%) berusia > 10 tahun, dengan rata-rata usia 5,11 tahun. Sebagian besar, 77 (52%) penderita datang ke rumah sakit sebelum hari kelima sakit.

Tabel 2. Data Gejala Klinis Penderita Difteri

Gejala	Jumlah Kasus <i>n</i> (148)
Letak Membran, <i>n</i> (%)	
a. Tonsil unilateral	19 (12,8)
b. Tonsil bilateral	125 (84,4)
c. Faring	6 (4,0)
d. Laring	3 (2,0)
e. Mata	1 (0,06)
Gambaran Klinis, <i>n</i> (%)	
a. Demam	147 (99,3)
b. Nyeri tenggorokan	92 (62,2)
c. Parau	29 (19,6)
d. Stridor	59 (39,9)
e. <i>Bullneck</i>	44 (29,7)
Kultur Positif, <i>n</i> (%)	
a. <i>Corynebacterium diphtheria</i>	33 (22,3)
Komplikasi, <i>n</i> (%)	
a. Miokarditis	17 (11,5)
b. Obstruksi Jalan Napas	6 (4,0)
c. <i>Acute Kidney Injury</i> (AKI)	6 (4,0)
Meninggal, <i>n</i> (%)	11 (7,4%)

Sebagian besar (125/148, 84,4%) penderita datang dengan lokasi pseudomembran di tonsil bilateral dan sisanya tonsil unilateral, faring, laring, dan konjungtiva. Hampir seluruh penderita datang dengan klinis panas, tetapi hanya 29 (19,6%) dengan klinis suara parau. Dari seluruh penderita yang dilakukan kultur *Corynebacterium diphtheriae* didapatkan hasil positif pada 34 pasien (22,9%). Komplikasi tersering yang didapatkan adalah Miokarditis (17/148, 11,5%). Enam (4%) penderita menjalani trakeostomi. Sebelas penderita (7,4%) meninggal dunia, dengan penyebab kematian Miokarditis (9/81,8%).

3.3. Wawancara Ahli

Setelah melakukan wawancara dengan Dokter Spesialis penyakit Difteri yang merupakan Dokter di RSUD Dr. Soetomo Surabaya. Dari hasil wawancara dan proses analisa, hasilnya adalah pembuatan tabel data analisa atribut serta analisa nilai bobot atribut pada penyakit Difteri yang ditunjukkan pada Tabel 3.

Tabel 3. Data Analisa Atribut dan Nilai Bobot

No	Atribut	A	B	C	Bobot
1.	< 5	√	√	√	1
2.	5 – 10	√		√	0,5
3.	> 10		√		0,25
4.	Laki-laki	√	√		1
5.	Perempuan	√			0,5
6.	0 – < 5	√		√	1
7.	≥ 5		√		0,5
8.	Surabaya	√	√	√	1
9.	Luar Surabaya	√			0,5
10.	Difteri ringan		√	√	0,5
11.	Difteri sedang	√	√		0,75
12.	Difteri berat	√		√	1
13.	<i>Unvaccinated</i>	√		√	1
14.	<i>Vaccinated</i>	√	√		0,5
16.	Buruk	√	√	√	1
17.	Kurang atau Baik	√	√		0,75
18.	Tonsil unilateral	√		√	0,75
19.	Tonsil bilateral	√	√	√	1
20.	Faring	√	√		0,75
21.	Laring	√		√	0,5
22.	Mata	√		√	0,25
23.	Demam	√	√	√	1
24.	Nyeri teggorokan	√	√		0,75
25.	Parau		√		0,25
26.	Stridor	√		√	0,75
27.	<i>Bullneck</i>	√	√		0,5
28.	<i>Corynebacterium diphtheria</i>	√		√	1
29.	N/A		√		0,25

Keterangan:

A : Miokarditis

B : Obstruksi Jalan Napas

C : *Acute Kidney Injury* (AKI)

Data analisa dan nilai bobot yang diberikan berdasarkan pengambilan data karakteristik dan data gejala klinis penderita Difteri dengan nilai bobot antara 0 sampai 1. Selain itu untuk pemberian nilai bobot itu sendiri berpedoman pada analisa data 148 pasien penderita dan hasil pertimbangan dari Dokter Spesialis di RSUD Dr. Soetomo Surabaya, data tersebut akan digunakan untuk dataset yang akan diolah pada perhitungan similarity.

3.4. Metode Perhitungan

Menggunakan metode *Case Base Reasoning* maka akan digunakan 4 tahapan proses dalam sistem penalaran komputer berbasis kasus yaitu:

3.4.1. Proses Retrieve

Retrieve proses akan mencari *similarity* antara kasus baru dengan kasus yang lama. Berikut contoh kasus 1 dipaparkan pada Tabel 4.

Tabel 4. Data Analisa Atribut dan Nilai Bobot

No	Nama	Atribut	Bobot
1.	Pasien Baru	< 5	1
		Laki-laki	1
		0 – < 5	1
		Surabaya	1
		Difteri sedang	0,75
		<i>Unvaccinated</i>	1
		Kurang atau Baik	0,75
		Tonsil unilateral	0,75
		Laring	0,5
		Demam	1
		<i>Bullneck</i>	0,5
		N/A	0,25

Kasus 1 akan dicari nilai kedekatannya dengan kasus lama mengambil 2 contoh kasus yang paling mendekati dengan kasus baru. Contoh kasus lama yang pertama dipaparkan pada Tabel 5.

Tabel 4. Data Analisa Atribut dan Nilai Bobot

No	Pasien	Atribut	Bobot		Nilai Atribut <i>Similarity</i>		Penyakit
1.	Pasien Lama 1	< 5	a ₁	1	b ₁	1	Miokarditis
		Perempuan	a ₂	0,5	b ₂	0	
		0 – < 5	a ₃	1	b ₃	1	
		Surabaya	a ₄	1	b ₄	1	
		Difteri ringan	a ₅	0,5	b ₅	0	
		<i>Vaccinated</i>	a ₆	0,5	b ₆	0	
		Kurang atau Baik	a ₇	0,75	b ₇	1	
		Tonsil bilateral	a ₈	1	b ₈	0	
		Faring	a ₉	0,75	b ₉	0	
		Mata	a ₁₀	0,25	b ₁₀	0	
		Demam	a ₁₁	1	b ₁₁	1	
		Nyeri teggorokan	a ₁₂	0,75	b ₁₂	0	
		<i>Bullneck</i>	a ₁₃	0,5	b ₁₃	1	
		<i>Corynebacterium diphtheria</i>	a ₁₄	1	b ₁₄	0	

Data-data yang telah dipaparkan pada Tabel 5 merupakan data kasus lama yang terdiri dari data pasien, atribut, bobot dan nilai kedekatan atribut *similarity* pada masing-masing gejala yang akan di proses pada perhitungan.

$$\begin{aligned}
 \text{Similarity} &= \frac{\begin{bmatrix} (a_1 \times b_1) + (a_2 \times b_2) + \\ (a_3 \times b_3) + (a_4 \times b_4) + \\ (a_5 \times b_5) + (a_6 \times b_6) + \\ (a_7 \times b_7) + (a_8 \times b_8) + \\ (a_9 \times b_9) + (a_{10} \times b_{10}) + \\ (a_{11} \times b_{11}) + (a_{12} \times b_{12}) + \\ (a_{13} \times b_{13}) + (a_{14} \times b_{14}) \end{bmatrix}}{\begin{matrix} b_1 + b_2 + b_3 + \\ b_4 + b_5 + b_6 + \\ b_7 + b_8 + b_9 + \\ b_{10} + b_{11} + b_{12} + \\ b_{13} + b_{14} \end{matrix}} & \quad \text{Similarity} = \frac{\begin{bmatrix} (1 \times 1) + (0,5 \times 0) + \\ (1 \times 1) + (1 \times 1) + \\ (0,5 \times 0) + (0,5 \times 0) + \\ (0,75 \times 1) + (1 \times 0) + \\ (0,75 \times 0) + (0,25 \times 0) + \\ (1 \times 1) + (0,75 \times 0) + \\ (0,5 \times 1) + (1 \times 0) \end{bmatrix}}{\begin{matrix} 1 + 0 + 1 + \\ 1 + 0 + 0 + \\ 1 + 0 + 0 + \\ 1 + 0 + 0 + \\ 0 + 1 + 0 + \\ 1 + 0 \end{matrix}} \\
 \text{Similarity} &= \frac{\begin{bmatrix} 1 + 0 + 1 + \\ 1 + 0 + 0 + \\ 0,75 + 0 + 0 + \\ 0 + 1 + 0 + \\ 0,5 + 0 \end{bmatrix}}{\begin{matrix} 1 + 0 + 1 + \\ 1 + 0 + 0 + \\ 1 + 0 + 0 + \\ 0 + 1 + 0 + \\ 1 + 0 \end{matrix}} & \quad \text{Similarity} = \frac{[5,25]}{6} \\
 & & \quad \text{Similarity} = 0,875
 \end{aligned}$$

Contoh kasus lama yang kedua dipaparkan pada Tabel 6.

Tabel 6. Data Analisa Atribut dan Nilai Bobot

No	Pasien	Atribut	Bobot		Nilai Atribut Similarity		Penyakit
1.	Pasien Lama 2	5 – 10	a ₁	0,5	b ₁	0	Obstruksi Jalan Napas
		Perempuan	a ₂	0,5	b ₂	1	
		≥ 5	a ₃	0,5	b ₃	0	
		Surabaya	a ₄	1	b ₄	1	
		Difteri ringan	a ₅	0,5	b ₅	0	
		Vaccinated	a ₆	0,5	b ₆	0	
		Kurang atau Baik	a ₇	0,75	b ₇	1	
		Faring	a ₈	0,75	b ₈	0	
		Laring	a ₉	0,5	b ₉	1	
		Mata	a ₁₀	0,25	b ₁₀	0	
		Nyeri teggorokan	a ₁₁	0,75	b ₁₁	0	
		Parau	a ₁₂	0,25	b ₁₂	0	
		Stridor	a ₁₃	0,75	b ₁₃	0	
		N/A	a ₁₄	0,25	b ₁₄	1	

Data-data yang telah dipaparkan pada Tabel 6 merupakan data kasus lama yang terdiri dari data pasien, atribut, bobot dan nilai kedekatan atribut *similarity* pada masing-masing gejala yang akan di proses pada perhitungan.

$$\begin{aligned}
 \text{Similarity} &= \frac{\begin{bmatrix} (a_1 \times b_1) + (a_2 \times b_2) + \\ (a_3 \times b_3) + (a_4 \times b_4) + \\ (a_5 \times b_5) + (a_6 \times b_6) + \\ (a_7 \times b_7) + (a_8 \times b_8) + \\ (a_9 \times b_9) + (a_{10} \times b_{10}) + \\ (a_{11} \times b_{11}) + (a_{12} \times b_{12}) + \\ (a_{13} \times b_{13}) + (a_{14} \times b_{14}) \end{bmatrix}}{\begin{matrix} b_1 + b_2 + b_3 + \\ b_4 + b_5 + b_6 + \\ b_7 + b_8 + b_9 + \\ b_{10} + b_{11} + b_{12} + \\ b_{13} + b_{14} \end{matrix}} & \quad \text{Similarity} = \frac{\begin{bmatrix} (0,5 \times 0) + (0,5 \times 1) + \\ (0,5 \times 0) + (1 \times 1) + \\ (0,5 \times 0) + (0,5 \times 0) + \\ (0,75 \times 1) + (0,75 \times 0) + \\ (0,5 \times 1) + (0,25 \times 0) + \\ (0,75 \times 0) + (0,25 \times 0) + \\ (0,75 \times 0) + (0,25 \times 1) \end{bmatrix}}{\begin{matrix} 0 + 1 + 0 + \\ 1 + 0 + 0 + \\ 1 + 0 + 1 + \\ 0 + 0 + 0 + \\ 0 + 1 \end{matrix}} \\
 \text{Similarity} &= \frac{\begin{bmatrix} 0 + 0,5 + 0 + \\ 1 + 0 + 0 + \\ 0,75 + 0 + 0,5 + \\ 0 + 0 + 0 + \\ 0 + 0,25 \end{bmatrix}}{\begin{matrix} 0 + 1 + 0 + \\ 1 + 0 + 0 + \\ 1 + 0 + 1 + \\ 0 + 0 + 0 + \\ 0 + 1 \end{matrix}} & \quad \text{Similarity} = \frac{[3]}{5} \\
 & & \quad \text{Similarity} = 0,6
 \end{aligned}$$

3.4.2. Proses Reuse

Reuse proses akan menggunakan diagnosis kasus sebelumnya (kasus lama) untuk menghitung *similarity* pada masing-masing gejala. Jika nilai *similarity* minimum berkisar antara nilai 0,8 sampai 1, maka diagnosis kasus lama akan digunakan kembali untuk kasus baru. Nilai *similarity* pada contoh kasus 1 dengan kasus lama yang pertama mendapatkan nilai kemiripan 0,875, sehingga diagnosis dari kasus lama yang pertama yaitu Miokarditis, dapat digunakan lagi sebagai *knowledge base* pada kasus pasien baru.

3.4.3. Proses Revise

Revise proses akan menyimpan kasus yang memiliki nilai *similarity* di bawah 0,8. Maka kasus baru yang memiliki *similarity* rendah tidak akan dibuang akan tetapi disimpan sebagai *knowledge base* masa depan.

3.4.4. Proses Retain

Retain proses akan mengumpulkan kasus yang telah dianalisa oleh pakar dan dapat didiagnosis pada tahap revisi serta disimpan ke database dan menjadi sumber baru untuk *knowledge base* masa depan.

3.5. Pengujian

Pengujian terhadap 148 data rekam medis pasien penderita penyakit Difteri di RSUD Dr. Soetomo Surabaya, dengan 148 data training dan 44 data menggunakan Cross Validation diperoleh hasil dari detail perhitungan manual dengan akurasi 95%.

Perhitungan Manual :

1. Data *Training* 148 data dari rekam medis.
2. Nilai *alpha* : 70% berarti 148 data *training* dan 30% atau 44 data *testing*.
3. Langkah selanjutnya pengujian sebanyak “K” kali. Nilai K dipilih secara *random*, K=10.
4. Kemudian pengujian sebanyak 10 kali dengan data *training* yang berbeda-beda di 148 data tersebut.

Pengujian data *training* dengan catatan akurasi pada masing-masing pengujian sebanyak 10 kali dipaparkan pada Tabel 7.

Tabel 7. Hasil Pengujian Data Kasus Baru

No	Jumlah Pengujian Data <i>Training</i> – Data <i>Testing</i>	Jumlah Data Sesuai	Jumlah Data Tidak Sesuai	Akurasi (%)
1	148 - 44	42	2	95.45%
2	148 - 44	41	3	92,7%
3	148 - 44	42	2	95.45%
4	148 - 44	43	1	97,7%
5	148 - 44	42	2	95.45%
6	148 - 44	40	4	90,9%
7	148 - 44	42	2	95.45%
8	148 - 44	43	1	97,7%
9	148 - 44	42	2	95.45%
10	148 - 44	42	2	95.45%
Rata-rata				95.17%

Hasil pengujian data kasus baru menunjukkan hasil uji data rekam medis dengan jumlah pengujian data 44 data kasus baru terhadap 148 data *training*, diperoleh tingkat akurasinya sebesar 95,17%.

4. KESIMPULAN

Hasil dari penelitian ini dapat disimpulkan bahwa Case Based Reasoning akan tepat jika diterapkan untuk mengidentifikasi diagnosa penyakit Difteri, karena melihat dari persebaran penyakit Difteri yang saat ini sudah digolongkan menjadi kejadian luar biasa (KLB), dan dilihat dari data kasus penyakit Difteri pada tahun terakhir tercatat semakin meningkat yaitu 318 kasus dan 12 orang meninggal dunia. Dengan Case Based Reasoning diharapkan dapat membantu pakar dalam menangani kasus penyakit Difteri, memanfaatkan data terdahulu sebagai knowledge base sebagai data perbandingan dengan kasus baru dengan model perhitungan similarity (K-NN), dan nantinya akan selalu diperbarui dengan kasus-kasus baru yang akan muncul di masa depan. Hasil pengujian dari 148 data rekam medis pasien penderita penyakit Difteri di RSUD Dr. Soetomo Surabaya menunjukkan hasil akurasi mencapai 95.17% dari 44 pengujian data kasus baru.

5. SARAN

Penelitian Case Based Reasoning Diagnosis Penyakit Difteri dengan Algoritma K-Nearest Neighbor bisa dilakukan penambahan data atribut dari data rekam medis jika ada penambahan data pada rekam medis pasien. Seiring bertambahnya waktu maka semakin banyak pula kasus yang akan terjadi, selain itu dalam pendekatan klasifikasi dari data yang ada memungkinkan untuk penerapan algoritma lain untuk perbandingan demi mendapatkan akurasi yang lebih baik.

DAFTAR PUSTAKA

- [1] Kementerian Kesehatan Republik Indonesia, 2017, Profil Kesehatan Indonesia <http://www.depkes.go.id/folder/view/01/structure-publikasi-pusdatin-profil-kesehatan.html>, diakses tanggal 14 Desember 2017.
- [2] Zainuddin, M., Hidjah, K., Tunjung, I. W., 2016, Penerapan *Case Based Reasoning (CBR)* untuk Mendiagnosis Penyakit Stroke Menggunakan Algoritma *K-Nearest Neighbor*, *CITISEE*, Yogyakarta, 23-24 Agustus.
- [3] Ernawati, 2017, Sistem Pakar Diagnosa Penyakit Pencernaan Manusia Menggunakan Metode *Case Based Reasoning*, *Jurnal SISTEMASI*, No. 2, Vol. 6, Hal. 35-44.
- [4] Nuramilus, S.E., Regasari, M.P.R., Arwan, A., 2017, Sistem Pakar Diagnosis Penyakit Demam: DBD, Malaria dan Tifoid Menggunakan Metode *K-Nearest Neighbor – Certainty Factor*, *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, Universitas Brawijaya, No. 5, Vol. 1, Hal. 426-435.
- [5] Barigou, F., 2016, Improving K-Nearest Neighbor Efficiency for Text Categorization, *Neural Network World*, Hal. 45-65.
- [6] Vedayoko, L. G., Sugiharti, E., dan Muslim, M. A., 2017, Expert System Diagnosis of Bowel Disease Using Case Based Reasoning with Nearest Neighbor Algorithm, *Scientific Journal of Informatics*, No. 2, Vol. 4, Hal. 134-142.
- [7] Kosasi, S., 2015, Pembuatan Aplikasi Diagnosa Kerusakan Mesin Sepeda Motor Matic dengan Case-Based Reasoning, *Citec Journal*, No. 3, Vol. 2, Hal. 192-206.
- [8] Hartatik, 2016, Diagnosa Penyakit Pulmonary Tuberculosis Dan Extrapulmonary Tuberculosis Menggunakan Algoritma Certainty Factor (CF), *CSRID Journal*, No.1, Vol.8, Hal. 11-24.
- [9] Sommerville, I., 2010, *Software Engineering*, Ed 9, Addison-Wesley, Boston.
- [10] Turban, Efraim, J.E. Aronson and Ting Peng Liang, 2005, *Decision Support Systems and Intelligent Systems*, Ed 7, Prentice-Hall, New Jersey.
- [11] Suriyanti, 2013, *Aplikasi Sistem Pakar Pendeteksian Kerusakan Printer dengan Case Based Reasoning*, Pelita Informatika Budi Darma, No.3, Vol.V, Hal. 35-39.
- [12] Adriana, S.A., Indarto, dan Abdiansah, 2008, *Sistem Penalaran Komputer Berbasis Kasus (Case Based Reasoning - CBR)*, Penerbit Ardana Media, Yogyakarta.
- [13] Aamodt, A., E. Plaza., 1994, Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches, *AI Communications*, no.1, vol.7, Hal. 39-59
- [14] Kusriani, 2009, *Algoritma Data Mining*, Penerbit Andi, Yogyakarta.
- [15] Puspitasari, D., Supatmini, E., Husada, D., 2012, Gambaran Klinis Penderita Difteri Anak Di RSUD Dr. Soetomo, *Jurnal Ners*, No. 2, Vol. 7, Hal. 136–141.
-