

## CASE BASED REASONING UNTUK MENDIAGNOSA PENYAKIT KEHAMILAN MENGUNAKAN *COSINE SIMILARITY*

Adelysa P. Gitasswara<sup>1</sup>, Sebastianus A. S. Mola<sup>2</sup>, E. S. Y. Pandie<sup>3</sup>  
<sup>1,2,3</sup> Jurusan Ilmu Komputer, Fakultas Sains dan Teknik, Universitas Nusa Cendana

### INTISARI

Penerapan *Case Based Reasoning* (CBR) dalam mendiagnosa penyakit kehamilan dilatarbelakangi oleh kurangnya jumlah dokter spesialis kandungan. Penggunaan CBR bertujuan untuk menyelesaikan permasalahan baru dengan cara mengadaptasi solusi yang terdapat pada kasus sebelumnya, dengan cara menghitung tingkat kemiripan (similaritas). Perhitungan nilai kemiripan menggunakan metode *cosine similarity* dengan *threshold* sebesar 100%. Sistem ini dapat mendiagnosa 6 penyakit 28 gejala yang ada. Keluaran sistem berupa penyakit yang dialami oleh pasien berdasarkan gejala-gejala yang diinputkan oleh tenaga medis non spesialis, serta solusi penanganan yang disertai presentasi kemiripan dengan kasus terdahulu untuk menunjukkan tingkat kebenaran dari kemungkinan hasil diagnosa. Berdasarkan hasil pengujian kasus didapatkan hasil sistem dapat mengambil kembali kasus lama yang tepat dan telah menggunakan rumusan metode *Cosine Similarity* dengan benar ditunjukkan dengan hasil akurasi 100%, serta dengan menggunakan 104 kasus dinilai cukup optimal untuk mendiagnosa 6 penyakit yang ditunjukkan dengan hasil rata-rata *similarity* terhadap 20 kasus yaitu sebesar 90 %.

**Kata kunci:** *Case Base Reasoning*, penyakit kehamilan, *threshold*, *Cosine Similarity*.

### ABSTRACT

*The application of case-based reasoning in diagnosing pregnancy diseases is motivated by the lack of number of obstetricians. The use of CBR aims to solve new problems by adapting the solutions contained in the previous case, by calculating the level of similarity. Calculation of similarity value using cosine similarity method, with threshold equal to 100%. This system can diagnose 6 diseases, 28 existing symptoms. System outbreaks of illness experienced by patients based on symptoms induced by non-specialist medical personnel, as well as handling solutions accompanied by a presentation of similarities with previous cases to indicate the degree of truth of possible diagnosis. Based on the results of case testing, the results obtained: the system can retrieve the exact old case and have used the cosine similarity methodology correctly, shown with 100% accuracy results, and using 104 cases is optimal enough to diagnose 6 illnesses shown with average results Similarity to 20 cases is 90%.*

**Keyword:** *Case Base Reasoning*, pregnancy diseases, *threshold cosine similarity*

## I. PENDAHULUAN

Kesehatan merupakan hal yang sangat penting bagi manusia, begitu juga dengan kesehatan pada ibu hamil karena mempengaruhi pertumbuhan janin. Survei Demografi dan Kesehatan Indonesia (SDKI) menunjukkan bahwa angka kematian ibu pada tahun 2012 adalah 359 kematian per 100.000 kelahiran hidup. Angka ini meningkat dari 220 kematian per 100.000 kelahiran hidup pada tahun 2010. Secara global sekitar 80% kematian ibu disebabkan oleh komplikasi kehamilan, terutama

disebabkan oleh perdarahan, infeksi, aborsi yang tidak aman, dan eklampsia (tekanan darah tinggi pada kehamilan) [1].

Berbagai masalah yang menyebabkan kematian dapat diatasi dengan konsultasi atau pemeriksaan secara rutin dengan mengunjungi Rumah Sakit, Klinik atau tempat praktek. Masalah yang sering muncul yaitu susahny bertemu dokter spesialis kandungan atau biasa disebut dokter spesialis obgin (Obstetri dan Ginekologi) untuk mendapatkan informasi atau berkonsultasi tentang kehamilan, karena terbatasnya jam kerja dan juga kurangnya spesialis kandungan. Dalam menyelesaikan masalah ini, digunakan CBR untuk membangun sebuah sistem dengan pengambilan keputusan kasus yang baru menggunakan kasus-kasus yang sudah terjadi sebelumnya. Dengan menggunakan CBR, kasus tidak terikat pada aturan tertentu dan kasus-kasus yang mirip pun tetap dapat dipecahkan karena sistem berbasis pada kasus. Penelitian terdahulu menggunakan metode *Cosine Similarity* dalam sistem deteksi dini diagnosa ISPA (Infeksi Saluran Pernapasan Akut) [2] dan pemanfaatan metode *vector space model* dan metode *cosine similarity* pada fitur deteksi hama dan penyakit tanaman padi [3]. Rekayasa perangkat lunak konsultasi penyakit kehamilan berbasis kasus (*case based reasoning*) di puskesmas gunung talang [4].

## II. MATERI DAN METODE

### 2.1 Data Penelitian

Data kasus yang digunakan adalah data kasus didapat dari rekam medis pasien yang berobat ke RS Ibu dan Anak Kota Kupang dari bulan Oktober tahun 2014 sampai Desember tahun 2015 berjumlah 124 data dan 104 di simpan pada basis kasus dan 20 di gunakan sebagai data uji.

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

*Case-Based-Reasoning* adalah suatu pendekatan untuk menyelesaikan suatu permasalahan (*Problem Solving*) berdasarkan solusi dari permasalahan sebelumnya [5]. *Case Based Reasoning* ini merupakan suatu paradigma pemecahan masalah yang banyak mendapat pengakuan yang pada dasarnya berbeda dari pendekatan *AI* utama lainnya. Suatu masalah baru dipecahkan dengan menemukan kasus yang serupa di masa lampau, dan menggunakannya kembali pada situasi masalah yang baru. Pada dasarnya, dalam memecahkan masalah CBR digambarkan melalui empat putaran proses yang disebut dengan *Four Re's* (4 R), yaitu:

1. *Retrieve the most similar case* (mencari kembali kasus lama yang paling mirip dengan kasus baru).
2. *Reuse the case to attempt to solve the problem* (menggunakan kembali solusi kasus lama untuk menyelesaikan kasus baru).
3. *Revise the proposed solution if necessary* (jika diperlukan, lakukan adaptasi dan revisi atas solusi lama yang diusulkan agar sesuai dengan situasi sekarang).
4. *Retain the new solution as a part of a new case* (menyimpan solusi baru menjadi kasus baru ke dalam basis kasus untuk digunakan pada penyelesaian masalah baru dimasa yang akan datang).

### 2.3 Metode Cosine Similarity

*Cosine Similarity* menghitung *similarity* antara dua objek,  $x$  dan  $y$  yang dinyatakan dalam dua vektor. Fungsi *similarity* tersebut hanya berlaku untuk data matriks yang atributnya berjenis biner (0 atau 1).

Koefisien *Cosine* merupakan pengukuran *similarity* yang paling sering digunakan untuk keperluan pengelompokan data. *Cosine similarity* ini sebanding dengan sudut antara dua vektor data dan tidak terpengaruh oleh panjang data [6]. Rumus *Cosine Similarity* adalah sebagai berikut:

$$\text{Similarity}(x,y) = \frac{x \cdot y}{\|x\| \|y\|} = \frac{\sum_{i=1}^n x_i \cdot y_i}{\sqrt{\sum_{i=1}^n (x_i)^2} \cdot \sqrt{\sum_{i=1}^n (y_i)^2}} \dots\dots\dots(2.1)$$

Keterangan:

$x \cdot y$  = *vector dot product* dari  $x$  dan  $y$ , dihitung dengan  $\sum_{i=1}^n x_i \cdot y_i$

$\|x\|$  = panjang vektor  $x$  dihitung dengan  $\sqrt{\sum_{i=1}^n (x_i)^2}$

$\|y\|$  = panjang vektor  $y$  dihitung dengan  $\sqrt{\sum_{i=1}^n (y_i)^2}$

$x_i = (x_1, x_2, x_3, \dots, x_n)$  = gejala – gejala pada kasus baru

$y_i = (y_1, y_2, y_3, \dots, y_n)$  = gejala – gejala pada kasus lama

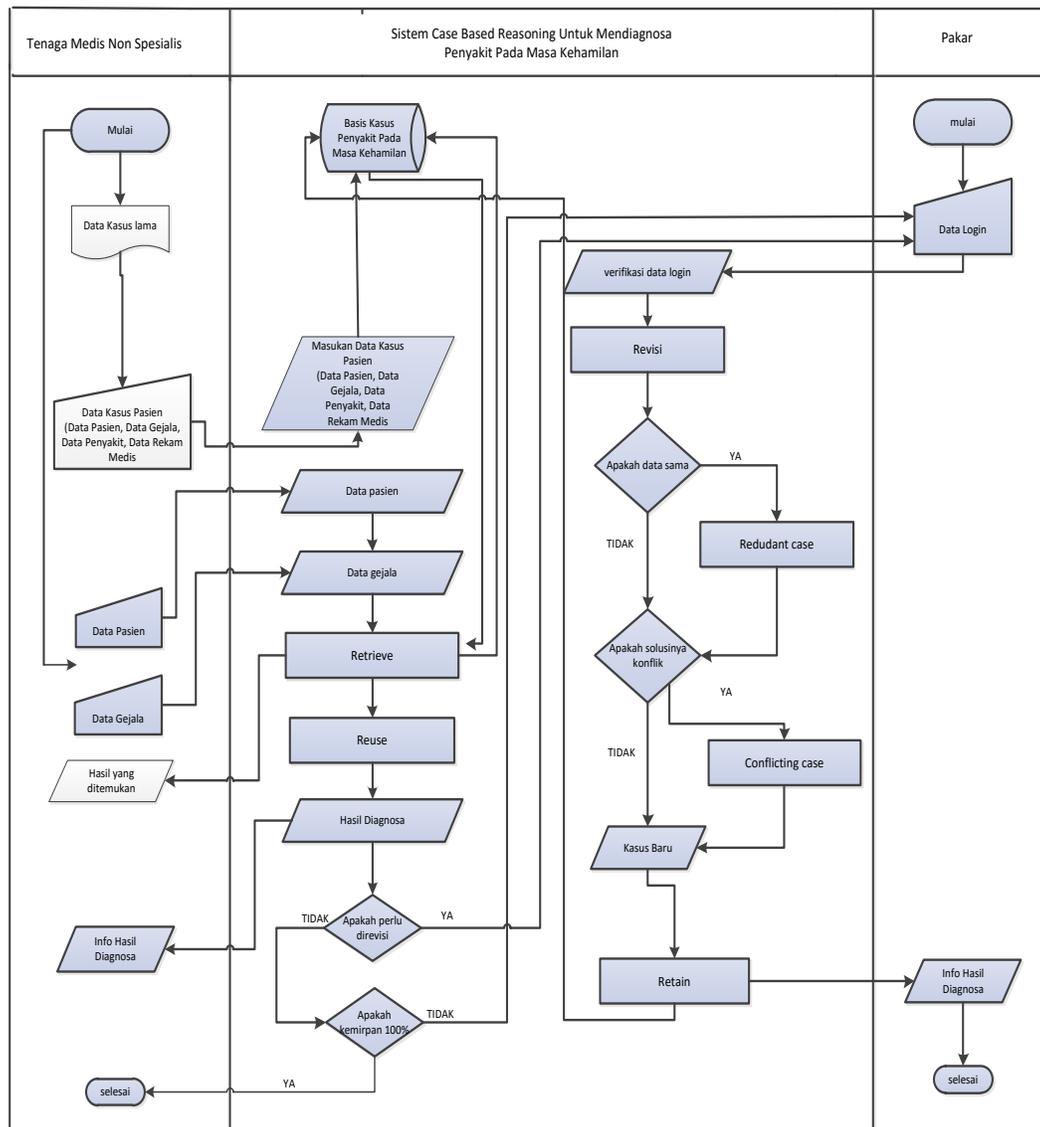
$n$  = jumlah gejala

### 2.4 Indexing

Indeks adalah struktur data yang mengatur *record* data pada *disk* untuk mengoptimalkan beberapa jenis operasi pengambilan (*retrieval*) tertentu. *Indeks* merupakan struktur data tersendiri yang tidak bergantung kepada struktur tabel. Indeks memungkinkan pengguna untuk secara efektif mengambil semua *record* yang memenuhi syarat pencarian pada *field search key* dari indeks. Hal ini membuat pencarian data akan lebih cepat dan tidak banyak menghabiskan *resource* komputer.

### 2.5. Tahap Perancangan Sistem CBR

Sistem CBR diagnosis penyakit kehamilan dimulai dengan tenaga medis non spesialis memasukkan data pasien dan data gejala berdasarkan pada keluhan pasien dan yang tampak dari pengamatan terhadap pasien. Selanjutnya data kasus baru akan diproses dalam sistem. Bila hasil diagnosis sama dengan 100% maka kasus baru tidak perlu direvisi.



Gambar 1. Flowchart sistem CBR diagnosis penyakit kehamilan

### III. HASIL DAN PEMBAHASAN

#### 3.1 Hasil Pengujian Sistem CBR

Sistem CBR telah diberikan pengujian dengan 2 (dua) cara uji. Pengujian pertama yaitu pengujian sistem CBR dengan kasus lama, cara ke-1 sistem diuji dengan memasukkan gejala yang sama dengan gejala pada salah satu kasus dalam database pada pengujian ini sistem mampu mengambil kembali kasus, dan cara ke-2 sistem diuji dengan memasukkan satu kasus yang telah dihitung kemiripannya secara manual didapatkan hasil yang sama dengan perhitungan sistem.

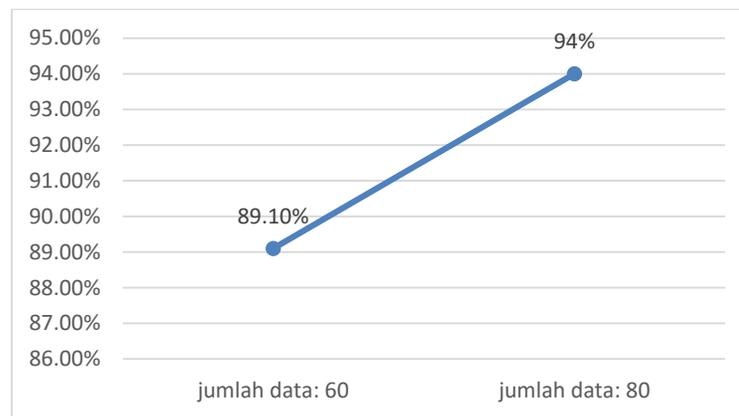
Pengujian kedua yaitu pengujian sistem CBR dengan kasus baru, cara ke-1 dilakukan dengan memasukkan data kasus baru yang belum ada pada basis pengetahuan. Pada pengujian sistem CBR

ini dimasukkan 20 data kasus baru yang belum ada pada basis pengetahuan. Dari 20 kasus terdapat 18 kasus yang memiliki hasil diagnosa yang sama dengan diagnosa yang dilakukan oleh pakar, dan terdapat 1 kasus yang memiliki hasil diagnosa yang sama tetapi dibawah *threshold* dan 1 kasus yang hasilnya tidak sama dengan hasil diagnosa yang dilakukan oleh pakar. Cara ke-2 Pengujian di lakukan dengan mengurangi jumlah basis kasus, sistem akan di uji dengan jumlah basis kasus 60, setelah itu di uji lagi dengan jumlah basis kasus 80 untuk melihat proses dari CBR dengan jumlah basis kasus yang berbeda – beda.

**Tabel 1.** Hasil pengujian kedua cara kedua

No	Jumlah Basis Kasus	Kasus Uji	Rata – Rata <i>Similarity</i>
1.	60	12	89,1%
2.	80	17	94%

Diagram pengaruh *similarity* terhadap jumlah kasus adalah sebagai berikut:



**Gambar 1.** Diagram *Similarity*

Diagram diatas menunjukkan, saat menggunakan 60 kasus didapatkan rata-rata nilai *similarity* sebesar 89,1%. Saat menggunakan 80 kasus, nilai *similarity* mengalami peningkatan sebesar 4,9% menjadi 94%.

Pengujian diatas menunjukkan semakin banyak kasus maka nilai *similarity* akan semakin tinggi karena kasus yang tersedia semakin lengkap sehingga rata-rata kasus revisi pun semakin rendah.

### 3.2 Pembahasan

Pada sistem CBR diagnosis penyakit kehamilan diberikan nilai batas kewajaran (*threshold*) sebesar 100%. Pada pengujian sistem CBR dengan kasus baru akan dilakukan dengan memasukkan data kasus baru yang belum ada pada basis pengetahuan. Pengujian ini dilakukan untuk menguji apakah sistem mampu mendiagnosa kasus baru yang belum ada dalam database. Pada pengujian sistem CBR ini dimasukkan 20 data kasus baru yang belum ada pada basis pengetahuan. Dari 20 kasus terdapat 18 kasus yang memiliki hasil diagnosa yang sama dengan diagnosa yang dilakukan oleh pakar, dan terdapat 1 kasus yang memiliki hasil diagnosa yang sama tetapi dibawah *threshold* dan 1 kasus yang hasilnya tidak sama dengan hasil diagnosa yang dilakukan oleh pakar. Presentase keakuratan sistem CBR dalam mendiagnosa kasus dengan cara membandingkan hasil diagnosa sistem dan hasil diagnosa pakar pada 20 kasus baru adalah 90 %.

Pada pengujian dengan mengurangi jumlah basis kasus, sistem akan di uji dengan jumlah basis kasus 60, setelah itu di uji lagi dengan jumlah basis kasus 80 untuk melihat proses dari CBR dengan

jumlah basis kasus yang berbeda – beda. Pada pengujian dengan menggunakan basis kasus 60 diuji menggunakan 12 kasus baru hasilnya terdapat 6 kasus dibawah *threshold* sehingga menghasilkan rata-rata *similarity* 89,1%, setelah dinaikkan basis kasus menjadi 80 basis kasus dan di uji menggunakan 17 kasus baru hasilnya terdapat 4 kasus dibawah *threshold* sehingga menghasilkan rata-rata *similarity* 94 %.

## IV. PENUTUP

### 4.1 Kesimpulan

Berdasarkan hasil penelitian pada sistem *Case Based Reasoning* untuk mendiagnosa penyakit kehamilan menggunakan *Cosine Similarity*, maka dapat disimpulkan sebagai berikut:

1. Pengujian keakuratan/validitas sistem yaitu dengan cara membandingkan hasil diagnosa yang dihasilkan oleh sistem dan pakar menggunakan 20, diperoleh hasil 90%.
2. Pengujian dengan jumlah basis kasus 60 dan 80 di peroleh nilai rata-rata *similarity* 89,1% dan 94% walaupun nilai rata-rata *similarity* tinggi tetapi terdapat beberapa data yang perlu direvisi oleh pakar.

### 4.2 Saran

Sebagai saran dari hasil penelitian adalah diharapkan:

- a. Penelitian selanjutnya dapat mengembangkan sistem berbasis *web*.
- b. Penelitian selanjutnya dapat menerapkan algoritma lain untuk mencari nilai kemiripan dalam penelitian dengan tujuan untuk peningkatan dalam tingkat akurasi.

## DAFTAR PUSTAKA

- [1] <http://www.depkes.go.id/infodatin>.
- [2] Cahyani.D., 2013, Sistem Deteksi Dini Diagnosa ISPA (Infeksi Saluran Pernapasan Akut) Pada Anak Dengan Metode *Cosine Similarity*, *Skripsi*, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Sebelas Maret, Surakarta.
- [3] Triana, A, 2014, Pemanfaatan Metode *Vector Space Model* dan Metode *Cosine Similarity* pada Fitur Deteksi Hama dan Penyakit Tanaman Padi, Naskah Publikasi, Unversitas Sebelas Maret, Surakarta.
- [4] Eva Yulianti dan Eka Gusriani., 2013, Rekayasa Perangkat Lunak Konsultasi Penyakit Kehamilan Berbasis Kasus (*Case Based Reasoning*) di Puskesmas Gunung Talang, Jurnal teknologi dan informasi pendidikan, Universitas Negeri Padang, Vol 6, No 1.
- [5] Aamodt, A., dan Plaza, E., 1994, *Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches*. *AI Communications*, Vol. 7, 39- 59.
- [6] Karhendana, A., 2008. Pemanfaatan dokumen *Clustering* pada Agregator Berita. *Skripsi* Terpublikasi, Bandung: Institut Teknologi Bandung.