

INTISARI

CASE-BASED REASONING DENGAN SELEKSI FITUR *RELIEF-F* DAN SIMILARITAS MAHALANOBIS DISTANCE UNTUK DIAGNOSIS PENYAKIT JANTUNG

Oleh

**HUSNUL KHATIMAH
15/388478/PPA/04917**

CBR merupakan salah satu teknik penalaran yang menyelesaikan masalah dengan cara menggunakan kembali kasus lama untuk menemukan solusi atas masalah tersebut. Terdapat dua tantangan yang harus dihadapi CBR dalam menyelesaikan masalah. Pertama, pemilihan fitur yang mampu mewakili suatu kasus dan kedua, pengambilan kasus (*retrieve*) serupa dari *case-base*. Algoritma seleksi fitur yang dapat digunakan salah satunya adalah *Relief-F*. *Relief-F* memanfaatkan teknik bobot untuk mengukur signifikansi fitur dan fitur yang memiliki nilai bobot di atas *threshold* yang digunakan akan dipilih untuk mewakili suatu *instance*. Proses *retrieve* dapat menggunakan metode *mahalanobis distance* yang mengukur jarak dengan memperhitungkan korelasi antara atribut menggunakan invers matriks varians-kovarians.

Penelitian ini melakukan diagnosis penyakit jantung dengan CBR yang menggunakan seleksi fitur *Relief-F* untuk mendapatkan fitur yang dapat mewakili kasus dalam *case-base* serta menggunakan metode *mahalanobis distance* untuk proses *retrieve*.

Hasil penelitian menunjukkan bahwa seleksi fitur menggunakan *Relief-F* dengan nilai *threshold* 0,0005 menghasilkan fitur yang relevan sebanyak 47 fitur dari 55 jumlah fitur gejala dan resiko, serta hasil tersebut tidak mempengaruhi hasil diagnosis awal dari basis kasus. Sedangkan untuk proses diagnosis penyakit jantung menggunakan metode *retrieve mahalanobis distance*, sistem menunjukkan unjuk kerja dengan nilai rata-rata sensitivitas sebesar 72,48%, nilai rata-rata spesifisitas sebesar 92,89% dan nilai rata-rata akurasi 73,01%.

Kata kunci : CBR diagnosis penyakit jantung, seleksi fitur, *Relief-F*, *mahalanobis distance*

ABSTRACT

CASE-BASED REASONING WITH RELIEF-F FEATURE SELECTION AND MAHALANOBIS DISTANCE SIMILARITY FOR HEART DISEASE DIAGNOSIS

By

HUSNUL KHATIMAH
15/388478/PPA/04917

CBR is one of the reasoning techniques that solves the problem by reusing the old case to find a solution to the problem. There are two challenges to be faced by CBR in solving the problem. First, the selected features have to represent a case and secondly, the retrieval process of the most similar case from the case-base. Feature selection algorithms that can be used one of which is Relief-F. Relief-F utilizes weighting techniques to measure the significance of features and features that have a weight value above the threshold used will be selected to represent an instance. The retrieving process can use the mahalanobis distance method by calculating the correlation between attributes using the inverse variance-covariance matrix.

This study diagnosed heart disease with CBR using the Relief-F feature selection to obtain features that could represent cases in the case-base and the mahalanobis distance method used for the retrieve process.

The results showed that feature selection using Relief-F with a threshold value of 0,0005 produce 47 relevant features from 55 symptom and risk features, and the results did not affect the initial diagnosis outcome of the case-base. Then, the process of heart disease diagnosis using retrieve mahalanobis distance method, the system shows the performance with average sensitivity is 72,48%, the average specificity is 92,89% and the average accuracy is 73,01%.

Keywords : CBR heart disease diagnosis, feature selection, Relief-F, mahalanobis distance