

INTISARI

METODE INDEXING PADA CASE BASED REASONING (CBR) MENGGUNAKAN DENSITY BASED SPATIAL CLUSTERING APPLICATION WITH NOISE (DBSCAN)

Oleh

HERDIESEL SANTOSO

14/371854/PPA/04605

Penelitian ini bertujuan meningkatkan kinerja *Case-Based Reasoning* dengan memanfaatkan *clustering*. Analisis *cluster* digunakan sebagai metode *indexing* untuk mempercepat pencarian pada proses *case retrieval*. Metode *clustering* menggunakan *Density Based Spatial Clustering Application with Noise* (DBSCAN) dibandingkan dengan *Self-organizing maps* (SOM). Metode *cosine similarity* digunakan untuk proses pencarian *cluster* yang relevan dan perhitungan similaritas dengan membandingkan tiga metode yaitu *nearest neighbor similarity*, *euclidean distance similarity* dan *minkowski distance similarity*. Penelitian ini menggunakan dua data uji, yaitu data penyakit gizi buruk dan data penyakit jantung. Pengujian dilakukan dengan tiga skenario. Skenario pertama menggunakan CBR *non-indexing*, skenario kedua menggunakan CBR dengan *indexing* SOM, dan skenario ketiga menggunakan CBR dengan *indexing* DBSCAN. Hasil pengujian menunjukkan bahwa CBR dengan *indexing* menggunakan *clustering* memiliki akurasi dan waktu proses yang lebih baik dari pada CBR *non-indexing*. Akurasi terbaik pada $threshold \geq 90$ data penyakit gizi buruk, didapat dengan menggunakan metode *minkowski distance* yang diimplementasikan pada CBR dengan *indexing* DBSCAN menghasilkan akurasi 100% dengan waktu *retrieve* rata-rata 0.02305 detik. Nilai akurasi terbaik pada $threshold \geq 90$ data penyakit jantung, didapat dengan menggunakan metode *minkowski distance* yang diimplementasikan pada CBR dengan *indexing* DBSCAN menghasilkan akurasi 95% dengan waktu *retrieve* rata-rata 0.05327 detik.

Kata kunci: penalaran berbasis kasus, *indexing*, *clustering*, SOM, DBSCAN, *nearest neighbor retrieval*

ABSTRACT

INDEXING METHODE ON CASE BASED REASONING (CBR) USING DENSITY BASED SPATIAL CLUSTERING APPLICATION WITH NOISE (DBSCAN)

By

HERDIESEL SANTOSO

14/371854/PPA/04605

This study aims to improve the performance of Case-Based Reasoning by utilizing clustering. Cluster analysis is used as a method of indexing to speed up the search in case retrieval process. Clustering method using Density Based Spatial Clustering Application with Noise (DBSCAN) compared with Self-organizing maps (SOM).. Cosine similarity method is used for the process of finding relevant cluster and the calculation of similarity by comparing three methods: nearest neighbor similarity, euclidean distance similarity and Minkowski distance similarity. This study used two test data, ie data malnutrition disease and heart disease data. Testing is done with three scenarios. The first scenario uses CBR non-indexing, the second scenario using CBR with indexing SOM, and the third scenario using CBR with indexing DBSCAN. The test results showed that, CBR with indexing using clustering has better accuracy and processing time than the CBR non-indexing. Best accuracy on the threshold of ≥ 90 malnutrition disease data, obtained using the Minkowski distance is implemented at CBR with indexing DBSCAN produces an accuracy of 100% with an average time of retrieving 0.02305 seconds. Best accuracy in the threshold value ≥ 90 heart disease data, obtained using the Minkowski distance is implemented at CBR with indexing DBSCAN produce 95% accuracy with an average time of retrieving 0.05327 seconds.

Keywords : Case-based reasoning, indexing, clustering, SOM, DBSCAN, nearest neighbor retrieval