

INTISARI

LOCAL TRIANGULAR KERNEL-BASED CLUSTERING (LTKC) UNTUK CASE INDEXING PADA CASE-BASED REASONING

Oleh

DAMAR RIYADI
15/388463/PPA/04902

Penelitian ini bertujuan untuk meningkatkan kinerja *Case-Based Reasoning* dengan memanfaatkan analisis kluster. Analisis kluster digunakan sebagai metode *indexing* untuk mempercepat proses *case retrieval* pada CBR. Metode *clustering* menggunakan *Local Triangular Kernel-based Clustering* (LTKC). Metode *cosine coefficient* digunakan untuk proses pencarian kluster yang sesuai atau relevan sedangkan nilai similaritas antar kasus dihitung dengan tiga metode yaitu *Manhattan distance similarity*, *Euclidean distance similarity*, dan *Minkowski distance similarity*. Hasil dari ketiga metode tersebut akan dibandingkan untuk mengetahui metode mana yang menghasilkan hasil terbaik. Penelitian ini menggunakan tiga data uji, yaitu data penyakit gizi buruk, penyakit jantung, dan penyakit *thyroid*. Pengujian dilakukan dengan dua skenario. Skenario pertama menggunakan CBR tanpa *indexing*, skenario kedua menggunakan CBR dengan *indexing* LTKC. Akurasi CBR dengan *indexing* LTKC pada kasus penyakit gizi buruk serta penyakit jantung dibandingkan dengan akurasi CBR dengan *indexing* SOM dan DBSCAN pada penelitian sebelumnya. Hasil pengujian menunjukkan bahwa CBR dengan *indexing* LTKC memiliki akurasi dan waktu proses yang lebih baik dari CBR tanpa *indexing*. Nilai akurasi terbaik pada *threshold* 0.9 data penyakit gizi buruk, didapat dengan metode *Euclidean distance* yang menghasilkan akurasi 100% dengan rata-rata waktu *retrieve* 0.0722 detik. Akurasi terbaik pada *threshold* 0.9 data penyakit jantung, didapat dengan metode *Minkowski distance* yang menghasilkan akurasi 95% dengan rata-rata waktu *retrieve* 0.1785 detik. Nilai akurasi terbaik pada *threshold* 0.9 data penyakit *thyroid*, didapat dengan metode *Minkowski distance* yang menghasilkan akurasi 92,52% dengan rata-rata waktu *retrieve* 0.3045 detik. Hasil perbandingan akurasi CBR dengan *indexing* SOM, DBSCAN, dan LTKC untuk diagnosis penyakit gizi buruk serta penyakit jantung menghasilkan akurasi yang hampir sama.

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

ABSTRACT

LOCAL TRIANGULAR KERNEL-BASED CLUSTERING (LTKC) FOR CASE INDEXING ON CASE-BASED REASONING

By

DAMAR RIYADI
15/388463/PPA/04902

This study aims to improve the performance of Case-Based Reasoning by utilizing cluster analysis. Cluster analysis is used as an indexing method to speed up case retrieval process in CBR. The clustering method uses Local Triangular Kernel-based Clustering (LTKC). The cosine coefficient method is used for finding the appropriate or relevant cluster while similarity value between cases is calculated using three methods: Manhattan distance similarity, Euclidean distance similarity, and Minkowski distance similarity. Results of those three methods will be compared to find which method gives the best result. This study uses three test data: malnutrition disease data, heart disease data, and thyroid disease data. Testing is performed with two scenarios. The first scenario uses CBR without indexing and the second scenario uses CBR with LTKC-indexing. The accuracy of CBR with LTKC-indexing in malnutrition disease and heart disease will be compared with CBR with SOM-indexing and DBSCAN-indexing in the previous study. Test results showed that CBR with LTKC-indexing has better accuracy and processing time than CBR without indexing. The best accuracy on threshold 0.9 of malnutrition disease data, obtained using the Euclidean distance which produces 100% accuracy and 0.0722 seconds average retrieval time. The best accuracy on threshold 0.9 of heart disease data, obtained using the Minkowski distance which produces 95% accuracy and 0.1785 seconds average retrieval time. The best accuracy on threshold 0.9 of thyroid disease data, obtained using the Minkowski distance which produces 92.52% accuracy and 0.3045 average retrieval time. The accuracy comparison of CBR with SOM-indexing, DBSCAN-indexing, and LTKC-indexing for malnutrition diseases and heart disease resulted that all of them have almost equal accuracy.

Keywords: case-based reasoning, *indexing*, *clustering*, LTKC, *nearest neighbor retrieval*