

ABSTRAK

Deteksi outlier merupakan salah satu task penting dalam data mining, bertujuan untuk mencari data yang memiliki karakteristik berbeda dengan kebanyakan data lainnya. Dalam dataset yang besar, deteksi *outlier* dianggap merupakan permasalahan yang sangat sulit jika dibandingkan dengan dataset yang kecil. Kebanyakan metode deteksi *outlier* sering tidak cocok diaplikasikan di dunia nyata karena data di dunia nyata biasanya memiliki distribusi yang tersebar. Struktur data tersebar tidak secara eksplisit mempresentasikan perilaku data normal dan tidak ada nya label yang menunjukkan data tersebut dikatakan sebagai outlier.

Pada deteksi outlier dengan algoritma LDOF, pencarian data yang menyimpang dapat ditemukan. Metode LDOF ini berbasiskan pendekatan jarak, dengan memperhitungkan nilai LDOF yang merepresentasikan derajat suatu data lalu membandingkannya dengan nilai Lower Bound LDOF yaitu batas nilai yang digunakan untuk mengetahui apakah data termasuk dalam *outlier* atau bukan dan mengurutkannya berdasarkan nilai LDOF terbesar sampai terkecil dimana data dengan nilai LDOF terbesar merupakan *outlier* yang dicari. Pengujian dilakukan dengan beberapa skenario untuk mengetahui nilai akurasi, error rate, dan FPR. LDOF dapat mendeteksi *outlier* dengan tingkat akurasi yang baik pada data yang memiliki distribusi data yang berbeda-beda.

Kata Kunci : *Outlier, LDOF, deteksi outlier, distance based approach*