

ABSTRAK

Data Mining adalah proses pencarian pola-pola dan kecenderungan yang menarik dari dalam basis data berukuran besar. Sebuah *outlier* didefinisikan sebagai sebuah titik data pada suatu data set dimana sangat berbeda dibandingkan dengan titik data pada data set pada umumnya dengan suatu ukuran tertentu. *Outlier* ini walaupun mempunyai kelakuan yang abnormal, seringkali mengandung informasi yang sangat berguna. Permasalahan deteksi *outlier* ini mempunyai peran yang sangat penting pada aplikasi deteksi kecurangan, analisis kekuatan jaringan dan deteksi intrusi.

Pencarian *outlier* biasanya dengan konsep keterdekatan berdasarkan hubungannya dengan sisa data yang ada. Pada data berdimensi tinggi, kepadatan data akan semakin berkurang, akibatnya dugaan akan keterdekatan antar data menjadi gagal.

Pada tugas akhir ini akan dilakukan perbandingan metode dalam pencarian suatu *outlier* dalam data berdimensi tinggi. Metode yang akan dibandingkan yaitu: *Clustering-based*, *Density-based*, dan *Distance-based*. Dimana masing-masing metode telah mendukung data berdimensi tinggi.

Kata Kunci : data mining, *outlier*, deteksi *outlier*, metode deteksi *outlier*.