

ABSTRAK

Clustering merupakan salah satu teknik *data mining* untuk mengelompokan data-data tak berlabel atau *unsupervised*, dimana data-data tersebut memiliki kemiripan karakteristik dengan data lain pada klaster yang sama dan berbeda dengan data-data pada klaster lain. Data yang memiliki perbedaan yang besar dengan data lain dapat dikatakan suatu *outlier*. Tugas akhir ini difokuskan pada analisis metode *Density-based Outlier Detection* yaitu *Local Outlier Factor* untuk mencari suatu *outlier* pada kumpulan data. *Density-based Outlier Detection* merupakan sebuah metode pendeteksian *outlier* yang menggunakan prinsip klaster kerapatan data. Salah satu cara mendeteksi *outlier* adalah dengan menggunakan metode *Local Outlier Factor (LOF)*, yaitu metoda menghitung *k-distance*, *k-distance neighborhood*, *reachability distance* dari suatu titik, dan *local reachability density* sehingga dihasilkan nilai LOF yang kemudian akan diurutkan nilainya. LOF akan menghasilkan nilai *outlier* tiap data yang menentukan apakah data tersebut anomali atau tidak. Dikatakan lokal karena tergantung bagaimana objek tersebut terpisah di antara objek-objek yang mengelilinginya. Suatu data p dikatakan bukan *local outlier* jika nilai $LOF(p)$ mendekati 1. Jika suatu data memiliki nilai LOF terbesar, maka dapat dikategorikan sebagai *outlier*. LOF dari suatu data tergantung *MinPts*, yaitu jumlah *nearest neighbors* yang digunakan untuk mendefinisikan *local neighborhood* dari suatu objek.

Kata kunci : *density-based outlier detection, local outlier factor*