

ABSTRAK

Kebutuhan untuk pengolahan data merupakan hal yang sangat penting untuk era saat ini. Dalam bidang pendidikan, *data mining* sangat berguna untuk mengembangkan sebuah metode yang dapat menemukan keunikan dari sebuah data yang berasal dari sistem pendidikan tersebut, dan menggunakan metode tersebut untuk lebih memahami siswa, sehingga dapat mengembangkan sebuah sistem yang sesuai (Barker & Yacef, 2009).

Salah satu metode dalam *data mining* untuk mengenali data yang bersifat unik di sebut deteksi *outlier*. Teknik deteksi *outlier* di kategorikan menjadi beberapa kelompok, yaitu *Statistical Approaches*, *Proximity-based Approaches*, *Clustering-based Approaches*, dan *Classification Approaches*.

Akan di buat sebuah aplikasi atau perangkat lunak untuk mendeteksi data yang unik tersebut menggunakan salah satu algoritma *data mining* yaitu algoritma *Connectivity-based Outlier Factor* dengan pendekatan *density based*. Data yang di gunakan dalam penelitian ini adalah data akademik mahasiswa Teknik Informatika Universitas Sanata Dharma angkatan 2007-2008 yang berupa hasil test penerimaan mahasiswa baru, nilai final dan nilai Indeks Prestasi Semester (IPS) dari semester satu sampai semester empat.

Pengujian akan di lakukan untuk menemukan data yang dinyatakan sebagai *outlier*. Penelitian ini bertujuan untuk membantu Kaprodi TI Universitas Sanata Dharma dalam mendeteksi kejadian langka atau yang menyimpang dari pola umum data akademik mahasiswa.

Kata kunci: *data mining*, deteksi *outlier*, *connectivity-based outlier factor*

ABSTRACT

The need for data processing is very important for the current era. In education, data mining is very useful to develop a method that can discover the uniqueness of a data derived from the education system, and using these methods to better understand the students, so that they can develop an appropriate system (Barker & Yacef, 2009).

One method in data mining to identify data that is unique is called outlier detection. Outlier detection techniques categorized into several groups, Statistical Approaches, Proximity-based Approaches, Clustering-based Approaches, and Classification Approaches.

It will be created an application or software to detect the unique data using one of the data mining algorithm is an algorithm-based Outlier Factor Connectivity with density-based approach. The data used in this study were student academic data Sanata Dharma University of Information Engineering 2007-2008 force in the form of new admissions test results, final value and the value of their GPA (IPS) from one semester to semester four.

Tests will be done to find the data that is declared as outliers. This research aims to help the chair person of Computer Science of Sanata Dharma University in detecting rare events or that deviate from the general pattern of student academic data.

Keyword : data mining, outlier detection, connectivity-based outlier factor