

ABSTRAK

Dengan menggunakan teknik data mining maka di perguruan tinggi dapat memanfaatkan data akademik mahasiswa yaitu Indeks Prestasi (IP) untuk memprediksi masa studi mahasiswa. diharapkan teknik data mining ini dapat digunakan oleh pihak akademik untuk mengevaluasi proses pembelajaran apakah sudah efektif. Data yang diolah dalam penelitian ini sebanyak 100 data yang bersumber dari Tumpukan data yang digunakan merupakan data akademik di Jurusan Fisika Universitas Andalas Padang Indonesia. Berdasarkan analisis terhadap Data, Klasifikasi ini menggunakan salah satu metode algoritma data mining yaitu k-Nearest Neighbor (k-NN). Algoritma k-NN bekerja berdasarkan jarak terpendek antara dua objek dengan menentukan nilai k. Nilai k merupakan parameter untuk menentukan jarak terdekat antara objek baru terhadap objek lama. Dalam teknik data mining ini terdiri dari data testing dan data training dengan inputan NIM dan nilai k. Nilai k yang terbaik, tergantung pada jumlah data yang digunakan. Ukuran nilai k yang besar untuk memprediksi masa studi mahasiswa belum tentu menjadi nilai k yang terbaik dengan tingkat keberhasilan yang tinggi begitupun juga sebaliknya. Hasil dari pengujian terhadap metode ini Pihak akademik dapat menentukan seberapa banyak mahasiswa yang akan lulus tepat pada waktunya berdasarkan hasil akurasi dengan k=5 adalah 77,19 % dan k=3 adalah 91,23 %.

Kata Kunci: Data mining, Klasifikasi, Algoritma k-Nearest Neighbor, RapiMiner.

ABSTRACT

By using data mining techniques, tertiary institutions can utilize student academic data, namely the Achievement Index (IP) to predict students' study periods. It is hoped that this data mining technique can be used by academics to evaluate whether the learning process has been effective. The data processed in this study as many as 100 data sourced from the data pile used is academic data in the Department of Physics, Andalas University, Padang Indonesia. Based on the analysis of data, this classification uses one of the methods of data mining algorithms, namely k-Nearest Neighbor (k-NN). The k-NN algorithm works based on the shortest distance between two objects by determining the value of k. The value of k is a parameter to determine the closest distance between a new object and the old object. In this data mining technique consists of data testing and training data with NIM input and k values. The best k value depends on the amount of data used. The size of a large k value to predict the period of study of students is not necessarily the best k value with a high level of success and vice versa. The results of testing this method Academic parties can determine how many students will graduate on time based on the results of accuracy with k = 5 is 77.19% and k = 3 is 91.23%.

Keywords: Data mining, Classification, k-Nearest Neighbor Algorithm, RapiMiner