

ABSTRAK

Tidak stabilnya rasio tingkat kelulusan mahasiswa program studi Sistem Informasi pada Universitas Sari Mutiara Indonesia menciptakan kondisi yang membuat adanya suatu penumpukan data. Teknik data mining dapat digunakan untuk memprediksi kelulusan tepat waktu mahasiswa. Penelitian menggunakan metode *K-Nearest Neighbors* yang merupakan sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data uji yang jaraknya paling dekat dengan objek tersebut. Pada penelitian ini menggunakan data mahasiswa tahun angkatan 2014 sampai dengan 2017 dengan jumlah data sebanyak 104 orang. Hasil dari perhitungan algoritma k-NN diimplementasikan dengan *jupyter notebook*. Tingkat akurasi pengujian model kelulusan mahasiswa dengan menggunakan algoritma K-Nearest Neighbor (k-NN) dipengaruhi oleh missing value indeks prestasi semester. missing value diganti dengan angka 0 maka hasil akurasi tertinggi adalah 95% dengan k=3. Jadi k terbaik adalah k=3 berdasarkan indeks prestasi semester satu sampai dengan semester enam.

Kata Kunci : *K-Nearest Neighbors (k-NN), Data Mining, Klasifikasi*



ABSTRACT

The unstable ratio of the graduation rate of students in the Information Systems study program at Sari Mutiara University Indonesia creates conditions that create a buildup of data. Data mining techniques can be used to predict student graduation on time. The research uses the K-Nearest Neighbors method which is a method for classifying objects based on test data that is closest to the object. This study uses student data from 2014 to 2017 with a total of 104 people. The results of the calculation of the k-NN algorithm are implemented with jupyter notebook. The level of accuracy of testing the student graduation model using the K-Nearest Neighbor (k-NN) algorithm is influenced by the missing value of the semester achievement index. missing value is replaced with number 0 then the highest accuracy result is 95% with k=3. So the best k is k=3 based on the achievement index for semester one to semester six.

Keywords : *K-Nearest Neighbors (k-NN), Data Mining, Classification*

