

ABSTRAK

Tingkat persaingan perguruan tinggi mengalami peningkatan yang sangat pesat setiap tahun. Penerimaan mahasiswa baru merupakan aktivitas penting bagi perguruan tinggi untuk menjaring mahasiswa baru. Dalam proses mendapatkan mahasiswa baru setiap perguruan tinggi memiliki strategi promosi yang berbeda. Penentuan strategi promosi yang tepat dapat menekan biaya promosi mencapai sasaran promosi yang tepat. Penelitian ini bertujuan untuk mengidentifikasi pola karakteristik profil mahasiswa dan calon mahasiswa sebagai rekomendasi strategi promosi. Pengolahan Data Mining menggunakan metode *Cross Industry Standard Process for Data Mining* (CRISP-DM) dengan model klasterisasi dan klasifikasi yang memanfaatkan algoritma *K-Means Clustering* dan *Decision Tree*. Implementasi pengolahan data menggunakan *RapidMiner Studio 9.10.1*. Atribut yang digunakan dalam model klasterisasi adalah Nim, KotaAsal, Asal Sekolah, Prodi, Jurusan Sekolah, Jenis Kelamin, Jalur Seleksi, Tahunpendaftaran. Pengolahan data menggunakan algoritma *K-Means Clustering* menghasilkan 3 buah *cluster*, *cluster_1* berjumlah 91 *items* (4,56%) dominasi mahasiswa paling banyak dari program studi Sistem Informasi. *Cluster_2* berjumlah 1683 *items* (84,48%) dominasi mahasiswa paling banyak dari Kota Surabaya dan Kab. Sidoarjo, *Cluster_3* berjumlah 218 *items* (10,96%) dominasi mahasiswa paling banyak dari daerah Kawasan Indonesia Timur dengan Prodi Informastika. Nilai *Davies_Bouldin Index* (DBI) sebesar 0,477. Atribut yang digunakan dalam model klasifikasi adalah No.Peserta, Status Payment Biaya UP3, Kab./Kota Asal, Kab./Kota Asal sekolah, Jenis Sekolah, Kategori Sekolah, Sumber Informasi, Penghasilan orang tua, Pilihan Prodi. Data set yang digunakan sebanyak 3.760 *record* dengan *mode test percentage split* yaitu 70% sebanyak 2.637 sebagai data *training* dan 30% sebanyak 1.113 sebagai data *testing*. Pengolahan data menggunakan *Decision Tree* menghasilkan nilai *accuracy* pengujian sebesar 98,86%, nilai *precision* sebesar 93,33%, nilai *recall* sebesar 32,56% dan AUC sebesar 0.992. Atribut yang paling berpengaruh paling tinggi berdasarkan *Decision Tree* menghasilkan sebuah pohon keputusan untuk memudahkan pengambilan keputusan strategi yang tepat.

Kata Kunci: Data mining, Strategi Promosi, *CRISP-DM*, *K-Means*, *Decision Tree*

ABSTRACT

The level of competition in tertiary institutions has increased very rapidly every year. Acceptance of new students is an important activity for universities to recruit new students. In the process of getting new students, each university has a different promotion strategy. Determining the right promotion strategy can reduce promotion costs and help achieve the right promotion goals. This study aims to identify the profile characteristics of students and prospective students as recommendations for promotion strategies. Data Mining processing uses the Cross Industry Standard Process for Data Mining (CRISP-DM) method with a clustering and classification model that utilizes the K-Means Clustering and Decision Tree algorithms. Implementation of data processing using RapidMiner Studio 9.10.1. The attributes used in the clustering model are Student Name, City of Origin, School of Origin, Study Program, School Major, Gender, Selection Path, and year of Registration. Data processing using the K-Means Clustering algorithm produces 3 clusters, with cluster_1 totaling 91 items (4.56%). The majority of students are from the Information Systems study program. Cluster_2 has 1,683 items (84.48%) dominated by students from the cities of Surabaya and Sidoarjo, and Cluster_3 has 218 items (10.96%) dominated by students from the Eastern Indonesia region with the Informatics Study Program. The Davies-Bouldin Index (DBI) value is 0.477. The attributes used in the classification model are Participant Number, UP3 Fee Payment Status, District/City of Origin, District/City of School Origin, Type of School, School Category, Source of Information, Parents' Income, and Choice of Study Program. The data set used was 3,760 records with a percentage split test mode, namely 70% of 2,637 as training data and 30% of 1,113 as data testing. Data processing using the Decision Tree resulted in an accuracy test value of 98.86%, a precision value of 93.33%, a recall value of 32.56%, and an AUC of 0.992. The most influential attribute of the highest level based on the Decision Tree produces a decision tree to facilitate making the right strategic decisions.

Keywords: *Data mining, Promotion Strategy, CRISP-DM, K-Means, Decision Tree*