

ABSTRAK

SMK PGRI 2 Karawang setiap tahun akan memberikan beasiswa pendidikan dan merekomendasikan siswa-siswinya nya kepada perusahaan berdasarkan prestasi dalam capaian Ujian Nasional namun mengalami kesulitan dalam pengelompokan siswa yang berprestasi. Setelah Ujian Nasional ditiadakan dari tahun 2020 karena covid-19 untuk itu tujuan penelitian ini yaitu mengelompokan prestasi siswa berdasarkan nilai akademik menggunakan data mining dengan menerapkan algoritma *K-Means* dan *Fuzzy C-Means* serta membandingkan hasilnya untuk mengetahui algoritma terbaik, penelitian ini bertempat di SMK PGRI 2 Karawang dan data yang digunakan merupakan nilai raport siswa dari semester 1 sampai 6. Hasil perhitungan manual dan program *python* algoritma *K-Means* sebanyak 52 siswa berprestasi, 25 siswa dengan prestasi sedang dan 28 siswa tidak berprestasi sedangkan pada algoritma *Fuzzy C-Means* sebanyak 48 siswa berprestasi, 29 siswa dengan prestasi sedang, dan 28 siswa tidak berprestasi. Pada perhitungan *rapidminer* studio hanya bisa menggunakan algoritma *K-Means* saja karena *Fuzzy C-Means* tidak tersedia dan berhasil mendapatkan 49 siswa berprestasi, 27 siswa dengan prestasi sedang dan 29 siswa tidak berprestasi serta berdasarkan nilai evaluasi dengan metode metode *Davies Bouldien Index* (DBI) dari setiap jenis perhitungan menunjukan algoritma *K-Means* merupakan algoritma terbaik karena memiliki nilai DBI yang mendekati 0 dan tidak negatif.

Kata Kunci: data mining, *davies bouldien index*, *fuzzy c-means*, *k-means*, pengelompokan, *python*

ABSTRACT

SMK PGRI 2 Karawang every year will provide educational scholarships and recommend their students to companies based on their achievements in the National Examination but having difficulty in classifying students who excel. After the National Examination was abolished from 2020 due to covid-19, the purpose of this research is to classify student achievement based on academic value using data mining by applying the *K-Means* and *Fuzzy C-Means* algorithms and comparing the results to find out the best algorithm. PGRI 2 Karawang and the data used is the value of student report cards from semesters 1 to 6. The results of manual calculations and the *python* program *K-Means* algorithm are 52 high achieving students, 25 students with moderate achievement and 28 students not achieving while in the *Fuzzy C-Means* algorithm as many as 48 students excel, 29 students with moderate achievement, and 28 students do not excel. In the *rapidminer* studio calculation, it can only use the *K-Means* algorithm because *Fuzzy C-Means* is not available and managed to get 49 high achieving students, 27 students with moderate achievement and 29 students not achieving and based on evaluation scores using the *Davies Bouldien Index* (DBI) method from each type of calculation shows the *K-Means* algorithm is the best algorithm because it has a DBI value that is close to 0 and is not negative.

Keyword: clustering, data mining, *davies bouldien index*, *fuzzy c-means*, *k-means*, *python*