

ABSTRAK

Dapodik digunakan untuk menjaring semua data terkait data kelembagaan dan kurikulum sekolah, data siswa, data guru dan karyawan, serta data sarana prasarana setiap sekolah di seluruh Indonesia. Sarana adalah alat langsung untuk mencapai tujuan pendidikan, lalu prasarana adalah alat tidak langsung untuk mencapai tujuan. Penelitian ini menggunakan Data Dapodik SD dan SMP. Karna pada bidang program dan pelaporan mengalami permasalahan dalam pemberian bantuan sarana dan prasarana pada jenjang SD, SMP sehingga hasil dari klasifikasi dapat ditemukan sekolah yang mengalami sarana dan prasarana yang kurang merata. Penelitian ini dilakukan pengklasifikasian sarana dan prasarana berdasarkan data Dapodik SD, SMP Karawang dengan menggunakan dua algoritma pengklasifikasian. Teknik klasifikasi yang digunakan yaitu *Naïve Bayes* dan *C.45*. klasifikasi yang akan dilakukan menggunakan 3 kelompok yaitu banyak, sedang, sedikit pada Data Dapodik SD, SMP tahun 2020/2021. Data yang digunakan 100 dibagi menjadi data training 70% dan data testing 30% lanjut perhitungan algoritma *Naïve Bayes* dengan nilai akurasi 63% sedang kan algoritma *C.45* dengan nilai akurasi 73%.

Kata Kunci : *C.45*, Dapodik, *Data Mining*, Klasifikasi, Sarana dan Prasarana, *Naive Bayes*.



ABSTRACT

The utilizing of Dapodik is to collect all data related to institutional and school curriculum data, student data, teacher and employee data, as well as data on facilities and infrastructure for each school throughout Indonesia. Facilities are direct tools to achieve educational goals, then infrastructure means indirect tools to achieve goals. This research uses data from Dapodik SD and SMP. Because in the field of programming and reporting, there are problems in providing facilities and infrastructure assistance at the elementary and junior high school levels, which are not yet appropriate, this is a problem of uneven distribution of facilities and infrastructure. The solution to the problem of uneven distribution of facilities and infrastructure for elementary and junior high schools is by using the classification of facilities and infrastructure based on data from Dapodik Elementary School, Junior High School so that the results of the classification can be found in schools that experience uneven facilities and infrastructure. This research was conducted to classify facilities and infrastructure based on data from Dapodik SD, SMP Karawang using two classification algorithms. The classification technique used is Naïve Bayes and C.45. The classification that will be carried out uses 3 groups, namely many, moderate, and few in the Dapodik SD and SMP data in 2020/2021. The data used is 100 divided into 70% training data and 30% testing further calculating the Naïve Bayes algorithm and C.45 using 30% data. The results of the calculation of the Naïve Bayes algorithm with an accuracy value of 63% while the C.45 algorithm with an accuracy value of 73%.

Keywords: C.45, Dapodik, Data Mining, Classification, Facilities and Infrastructure, Naive Bayes.