

ABSTRAK

Kemiskinan merupakan salah satu masalah sosial yang sering terjadi di Indonesia. Setiap daerah di Indonesia tentu memiliki penduduk yang berstatus miskin, namun masih banyak daerah yang penduduk miskinnya terpetahkan dengan baik salah satunya di desa Kenotan, Nusa Tenggara Timur. Masalah tersebut tentunya berdampak tidak baik bagi masyarakat miskin seperti bantuan yang tidak tepat sasaran dan masih banyak lagi. Untuk mengatasi hal tersebut, perlu adanya upaya klasifikasi untuk mengelompokan penduduk yang berstatus miskin dengan tepat. Salah satu cara yakni klasifikasi dengan menggunakan data mining. Ada banyak metode data mining yang digunakan untuk melakukan klasifikasi namun diantaranya memiliki tingkat akurasi yang berbeda. Tujuan dari penelitian ini adalah untuk melakukan klasifikasi penduduk miskin menggunakan dua metode data mining C4.5 dan metode *Support Vector Machine* (SVM) serta membandingkan tingkat akurasi dari kedua metode tersebut untuk melihat metode apa yang memiliki akurasi lebih baik dari kedua metode tersebut dalam proses klasifikasi khususnya klasifikasi penduduk miskin di desa Kenotan, Nusa Tenggara Timur. Terdapat 9 atribut yang digunakan untuk klasifikasi keluarga miskin dalam penelitian ini yakni jenis kelamin, tempat tanggal lahir, umur, pekerjaan, pendapatan / bulan, pendidikan terakhir, tanggungan, NIK dan KK yang mana pada tahap preprocessing atribut yang dipakai dalam proses pengujian menjadi 6 atribut yakni jenis kelamin, umur, pekerjaan, pendapatan / bulan, pendidikan terakhir dan tanggungan. Penelitian ini menghasilkan akurasi tertinggi pada algoritma C4.5 dengan hasil akurasi sebesar X pada $k\text{-fold} = X$ dan $\text{minimal leaf sampel} = X$.

Kata kunci : C4.5, Support Vector Machine, *k-fold cross validation*, Penduduk Miskin

ABSTRACT

Poverty is one of the social problems that often occurs in Indonesia. Every region in Indonesia certainly has a poor population, but there are still many areas where the poor population is well defined, one of which is Kenotan village, East Nusa Tenggara. These problems certainly have a negative impact on poor people, such as aid that is not well targeted and many more. To overcome this, there needs to be a classification effort to group the population with poor status appropriately. One way is classification using data mining. There are many data mining methods used to carry out classification, but they have different levels of accuracy. The aim of this research is to classify the poor using two data mining methods C4.5 and the Support Vector Machine (SVM) method and compare the level of accuracy of the two methods to see which method has better accuracy than the two methods in the classification process especially the classification of poor people in Kenotan village, East Nusa Tenggara. There are 9 attributes used to classify poor families in this research, namely gender, place, date of birth, age, occupation, income/month, highest level of education, dependents, NIK and KK. At the preprocessing stage, the attributes used in the testing process become 6 attributes. Namely gender, age, occupation, income/month, highest level of education and dependents. This research produces the highest accuracy in the C4.5 algorithm with accuracy results of X at k -fold = X and minimum sample leaf = X .

Keywords : C4.5, Support Vector Machine, k -fold cross validation, *Penduduk Miskin*