

ABSTRAK

Fariska Desi. 2023. Performa Algoritma C4.5 dan Naive Bayes Untuk Klasifikasi Penerima Bantuan Langsung Tunai Dana Desa (Studi Kasus Desa Asemdayong Kecamatan Pemalang Kabupaten Pemalang). Skripsi. Program Studi Statistika. Universitas Muhammadiyah Semarang. Pembimbing: Dr. Rochdi Warsono, M.Si, II. Tiani Wahyu Utami, S.Si., M.Si.

Setiap pemimpin baik pemimpin daerah maupun pusat menjadikan permasalahan penduduk miskin sebagai tujuan utama yang harus dituntaskan. Program jaring pengaman sosial melalui bantuan sosial langsung tunai, merupakan salahsatu kebijakan pemerintah berupa pemberian bantuan stimulan uang tunai yang bertujuan untuk membantu masyarakat miskin agar mampu mempertahankan kehidupannya dan tidak jatuh miskin yang lebih dalam. Tujuan utama dari *data mining* adalah untuk mengekstrak informasi yang berguna dari data mentah yang sangat besar dan mengubahnya menjadi bentuk yang dapat dimengerti untuk penggunaan yang efektif dan efisien. Pedoman yang jelas diperlukan untuk mengambil keputusan yang tepat sasaran, sehingga memudahkan tim seleksi untuk mengevaluasi penerima bansos dana desa. Dengan penerapan *data mining* memanfaatkan algoritma C4.5 dan *Naïve Bayes* diharapkan dapat membantu dalam menemukan data dalam database. Berdasarkan perbandingan hasil pengujian melalui berbagai skenario terhadap algoritma C4.5 dan *Naïve Bayes* tersebut, diperoleh nilai akurasi, *recall*, *f1-score* dan *AUC (Area Under Curve)* algoritma *Naïve Bayes* memiliki nilai lebih besar yaitu 97,23%, 100,00%, 97,40% dan 99,87%, dibandingkan algoritma C4.5 yaitu 96,11%, 95,52%, 96,24% dan 98,30%. namun dari *precision* algoritma C4.5 lebih unggul yaitu 97% dibanding algoritma *Naïve Bayes* 94,96%. Dari hasil keseluruhan pengujian model dapat disimpulkan bahwa performa algoritma Naive Bayes lebih baik dibandingkan algoritma C.45

Kata Kunci: Kemiskinan, BLT DD, Data Mining, *Naïve Bayes*, C4.5

ABSTRACT

Fariska Desi. 2023. *C4.5* and Naive Bayes Algorithm for *Classification* of Village Fund Direct Cash Assistance Recipients (Case Study of Asemdayong Village, Pemalang District, Pemalang Regency). Thesis. Statistics Study Program. Muhammadiyah University of Semarang. Supervisor: Dr. Rochdi Warsono, M.Si, II. Tiani Wahyu Utami, S.Si., M.Si.

Every leader, both regional and central leaders, makes the problem of the poor as the main goal that must be resolved. the social safety net program through direct cash social assistance, is one of the government policies in the form of providing cash stimulant assistance which aims to help the poor so they are able to maintain their lives and not fall deeper into poverty. The main goal of data mining is to extract useful information from huge raw data and convert it into an understandable form for effective and efficient use. Clear guidelines are needed to make decisions that are right on target, making it easier for the selection team to evaluate village fund social assistance recipients. With the application of data mining utilizing the C4.5 Algorithm and Naïve Bayes, it is hoped that it can assist in finding data in the database. Based on the comparison of test results through various scenarios on the C4.5 and *Naïve Bayes* Algorithms, the values for Accuracy, *Recall*, f1-score and AUC (Area Under Curve) of the *Naïve Bayes* Algorithm have a greater value of 97.23%, 100.00% , 97.40% and 99.87%, compared to the C4.5 algorithm, namely 96.11%, 95.52%, 96.24% and 98.30%. but in terms of *Precision* the C4.5 algorithm is superior, namely 97% compared to the *Naïve Bayes* algorithm of 94.96%. From the results of the overall model testing, it can be concluded that the performance of the Naive Bayes Algorithm is better than the C.45 algorithm.

Keywords : Poverty, BLT DD, Data Mining, Naïve Bayes, C4.5