

ABSTRAK

Nama Mahasiswa (1) : Denis Nila Cahyani
NIM : 41820010111
Nama Mahasiswa (2) : Aisyah Tri Deanita
NIM : 41820010122
Nama Mahasiswa (3) : Muhammad Anand Rizki Andinta
NIM : 41820010069
Pembimbing TA : Dr. Bambang Jekonowo, S.Si., M.T.I.
Judul Tugas Akhir : Penerapan Data Mining Menggunakan Algoritma Naïve Bayes, C4.5, dan K-Nearest Neighbor untuk Klasifikasi Kemiskinan di DKI Jakarta

Kemiskinan masih menjadi salah satu permasalahan fundamental yang sulit dihadapi, termasuk di Provinsi DKI Jakarta sebagai Ibu kota negara Indonesia, yang juga tidak luput dari permasalahan kemiskinan. Penelitian ini menggunakan *data mining* untuk melakukan klasifikasi terhadap tingkat kemiskinan di Provinsi DKI Jakarta dengan data yang diperoleh dari Badan Pusat Statistik (BPS). Tujuan penelitian ini adalah untuk secara efektif dan akurat mengklasifikasikan data kemiskinan dan membandingkan hasil dari algoritma Naïve Bayes, C4.5N dan K-NN guna mencari hasil akurasi terbaik. Melalui penerapan algoritma tersebut, didapatkan hasil bahwa performa algoritma dapat dipengaruhi oleh pembagian data, dan peningkatan rasio data uji cenderung meningkatkan akurasi serta konsistensi hasil klasifikasi. Pada rasio pembagian data 70:30, Naïve Bayes mencapai akurasi 81%, C4.5 sebesar 76%, dan K-NN sebesar 71%. Pada rasio 80:20, Naïve Bayes menunjukkan akurasi 93%, C4.5 sebesar 79%, dan K-NN sebesar 86%. Sementara pada rasio 90:10, Naïve Bayes mencapai akurasi 100%, C4.5 sebesar 71%, dan K-NN sebesar 86%. Dengan mempertimbangkan variasi performa ketiga algoritma, dapat disimpulkan bahwa Naïve Bayes lebih unggul sebagai algoritma yang stabil dan dapat diandalkan dalam berbagai skenario pembagian dataset, menunjukkan akurasi tinggi dan kemampuan baik dalam mengidentifikasi kasus positif.

Kata kunci:

kemiskinan, klasifikasi, Naïve Bayes, C4.5, K-Nearest Neighbor

ABSTRACT

Name (1) : Denis Nila Cahyani
Student Number : 41820010111
Name (2) : Aisyah Tri Deanita
Student Number : 41820010122
Name (3) : Muhammad Anand Rizki Andinta
Student Number : 41820010069
Counsellor : Dr. Bambang Jekonowo, S.Si., M.T.I.
Title : Penerapan Data Mining Menggunakan Algoritma Naïve Bayes, C4.5, dan K-Nearest Neighbor untuk Klasifikasi Kemiskinan di DKI Jakarta

Poverty remains one of the fundamental problems that is difficult to overcome, including in DKI Jakarta Province as the capital of Indonesia, which is also not immune to poverty. This research uses data mining to classify the poverty level in DKI Jakarta Province with data obtained from the Central Bureau of Statistics (BPS). The purpose of this study is to effectively and accurately classify poverty data and compare the results of the Naïve Bayes, KNN, and C4.5 algorithms to find the best accuracy results. Through the application of such algorithms, the results were obtained that the performance of the algorithm could be influenced by the division of data, and increased ratio of test data tended to improve the accuracy as well as consistency of classification results. At the data division ratio of 70:30, Naïve Bayes achieved accuracy of 81%, C4.5 of 76%, and K-NN of 71%. At the ratio of 80:20, Naïve Bayes showed accurateness of 93%, C4.5 of 79%, and K-NN of 86%. Whereas at the ratios of 90:10, Naïve Bayes achieves accurateness of 100%, C4.5 of 71%, and K-NN of 86%. By considering the performance variations of the third algorithm, it can be concluded that Naïve Bayes is superior as a stable and reliable algorithm in a variety of data set splitting scenarios, showing high accuracy and good ability in identifying positive cases.

Keywords:

poverty, classification, Naïve Bayes, C4.5, K-Nearest Neighbor