

ABSTRAK

Steven Martin (03081190039)

IMPLEMENTASI ALGORITMA C4.5 DALAM PREDIKSI PENYAKIT DIABETES MELLITUS PADA MANUSIA

(xv + 86; 52; 10; 22)

Penyakit diabetes mellitus menjadi salah satu penyakit yang sering dihadapi manusia, dengan jumlah pasien yang terus meningkat di Indonesia. Deteksi dini penyakit ini sangat penting untuk mencegah komplikasi serius. Penelitian ini mengimplementasikan algoritma C4.5 dalam prediksi penyakit diabetes mellitus pada manusia dengan menggunakan data mining. Data dianalisis dari Kaggle Pima Indians Diabetes Database dengan pembagian data *training* dan data *testing* yakni pengujian dengan proposisi data 60% data *training* dan 40% data *testing*, pengujian dengan proposisi data 75% data *training* dan 25% data *testing*, dan pengujian dengan proposisi data 90% data *training* dan 10% data *testing*. Hasil penelitian menunjukkan bahwa algoritma C4.5 memberikan tingkat akurasi, presisi, recall, dan tingkat kesalahan klasifikasi yang baik. Penggunaan algoritma C4.5 dengan data *training* 90% dan data *testing* 10% menghasilkan kinerja terbaik dengan tingkat akurasi sekitar 79,22%. Hal ini membuktikan bahwa algoritma C4.5 dapat digunakan efektif dalam prediksi penyakit diabetes mellitus pada manusia, dengan potensi untuk meningkatkan deteksi dini dan pengelolaan penyakit ini.

Kata Kunci: Data *Mining*, Diabetes Mellitus, Prediksi, Algoritma C4.5

Referensi: 14

ABSTRACT

Steven Martin (03081190039)

IMPLEMENTATION OF THE C4.5 ALGORITHM IN PREDICTION OF DIABETES MELLITUS DISEASE IN HUMAN

(xv + 86; 52; 10; 22)

Diabetes mellitus is one of the diseases that humans often face, with the number of patients continuing to increase in Indonesia. Early detection of this disease is very important to prevent serious complications. This research implements the C4.5 algorithm in predicting diabetes mellitus in humans using data mining. Data was analyzed from the Kaggle Pima Indians Diabetes Database by dividing training data and testing data, namely testing with a data proposition of 60% training data and 40% testing data, testing with a data proposition of 75% training data and 25% testing data, and testing with a data proposition of 90 % training data and 10% testing data. The research results show that the C4.5 algorithm provides good levels of accuracy, precision, recall and classification error rates. Using the C4.5 algorithm with 90% training data and 10% testing data produces the best performance with an accuracy level of around 79.22%. This proves that the C4.5 algorithm can be used effectively in the prediction of diabetes mellitus in humans, with the potential to improve early detection and management of this disease.

Keywords: Data Mining, Diabetes Mellitus, Prediction, C4.5 Algorithm

References: 14