

ABSTRAK

Elbert Utama (03081200053)

PERBANDINGAN EFEKTIVITAS SUPPORT VECTOR MACHINE DAN RANDOM FOREST DALAM IDENTIFIKASI DIABETES DAN PENYAKIT JANTUNG

(x + 83 halaman: 50 Gambar, 3 Tabel, 9 Lampiran)

Penelitian ini bertujuan untuk membandingkan efektivitas algoritma Support Vector Machine (SVM) dan Random Forest (RF) dalam identifikasi penyakit diabetes dan jantung. Penyakit diabetes dan jantung merupakan dua penyakit kronis yang prevalensinya semakin meningkat dan memerlukan metode diagnosis yang cepat dan akurat untuk pengelolaan yang lebih baik. Dalam penelitian ini menggunakan *dataset* yang diperoleh dari *Kaggle*, yang terdiri dari data pasien dengan atribut-atribut relevan yang digunakan untuk diagnosis penyakit. Algoritma SVM dan RF diterapkan untuk membangun model prediksi, dan kinerja masing-masing model dievaluasi berdasarkan metrik akurasi, presisi, recall, dan F1-score. Hasil penelitian menunjukkan bahwa model Stack dan Random Forest (RF) memiliki tingkat keakuratan yang tinggi dan sama, yaitu 89,4% dalam mendeteksi penyakit diabetes. Di sisi lain, model Stack dan Support Vector Machine (SVM) memiliki tingkat keakuratan yang tinggi dan sama, yaitu 81,8% dalam mendeteksi penyakit jantung. Kedua hasil ini menunjukkan bahwa baik model Stack-RF maupun Stack-SVM efektif dalam mendeteksi penyakit diabetes dan jantung, meskipun dengan tingkat akurasi yang berbeda. Penelitian ini memberikan kontribusi dalam bidang *data mining* dan *machine learning* untuk aplikasi medis, khususnya dalam pengembangan alat diagnostik berbasis *machine learning*. Dengan demikian, hasil dari penelitian ini diharapkan dapat membantu dalam pengembangan sistem pendukung keputusan klinis yang lebih akurat dan efisien untuk deteksi dini penyakit diabetes dan jantung.

Kata Kunci: Support Vector Machine, Random Forest, prediksi penyakit, diabetes, penyakit jantung, machine learning, data mining.

Referensi: 30 (2001-2023)

ABSTRACT

Elbert Hutama (03081200053)

COMPARISON OF THE EFFECTIVENESS OF SUPPORT VECTOR MACHINE AND RANDOM FOREST IN IDENTIFYING DIABETES AND HEART DISEASES

(x + 83 pages: 50 Figures, 3 Tables, 9 Appendices)

This study aims to compare the effectiveness of Support Vector Machine (SVM) and Random Forest (RF) algorithms in identifying diabetes and heart diseases. Diabetes and heart diseases are two chronic conditions with increasing prevalence, requiring rapid and accurate diagnostic methods for better management. In this research utilized a dataset obtained from Kaggle, comprising patient data with relevant attributes used for disease diagnosis. SVM and RF algorithms were employed to build predictive models, and the performance of each model was evaluated based on accuracy, precision, recall, and F1-score metrics. The results indicate that both the Stack-RF and RF models exhibit a high level of accuracy, at 89.4%, in detecting diabetes. Conversely, both the Stack-SVM and SVM models demonstrate a similar high accuracy rate of 81.8% in detecting heart disease. These findings suggest that both Stack-RF and Stack-SVM models are effective in detecting diabetes and heart diseases, albeit with differing accuracy levels. This study contributes to the field of data mining and machine learning for medical applications, particularly in the development of machine learning-based diagnostic tools. Thus, the findings of this research are expected to aid in the development of more accurate and efficient clinical decision support systems for early detection of diabetes and heart diseases.

Keywords: Support Vector Machine, Random Forest, disease prediction, diabetes, heart disease, machine learning, data mining.

References: 30 (2001-2023)