

Optimasi Algoritma Prediksi Menggunakan Particle Swarm Optimization Untuk Memprediksi Diabetes

Adriansyah Putra

Abstrak

Penyakit diabetes adalah masalah kesehatan global yang serius, dengan prevalensi yang terus meningkat. Indonesia menempati urutan ke-5 dunia, bahkan menjadi negara urutan pertama di Asia Tenggara yang terdampak penyakit ini. Berbagai upaya pencegahan dan penanganan telah banyak dilakukan, disamping penelitian-penelitian mengenai *data mining* untuk memahami dan manajemen penyakit ini. Menggunakan *dataset public* BIT Mesra, penelitian ini berfokus pada kinerja tiga algoritma klasifikasi: k-Nearest Neighbor, Random Forest, Deep Learning dalam memprediksi diabetes mellitus setelah dioptimasi menggunakan Particle Swarm Optimization. Data yang bersih dan lengkap menjadi landasan penting sebelum menerapkan metode optimasi, khususnya dalam menangani *noisy data* dan *missing values*. Hasil penelitian menunjukkan bahwa ketiga algoritma mencapai tingkat akurasi yang baik. Bahkan setelah dioptimasi dengan PSO, k-Nearest Neighbor dan Random Forest mencapai performa sempurna dengan akurasi mencapai 100%. Sementara Deep Learning menunjukkan peningkatan akurasi menjadi 98.95%. Tetapi meski akurasi dari k-Nearest Neighbor dan Random Forest ini lebih baik dari Deep Learning, hasil uji statistik menunjukkan bahwa ketiganya adalah setara dan sama baiknya dalam penggunaan. Ini memberikan pemahaman komprehensif tentang potensi penggunaan algoritma klasifikasi dan optimasi PSO dalam memprediksi diabetes, dengan penekanan pada keseimbangan antara penanganan data yang baik dan optimalisasi algoritma.

Kata Kunci: Particle Swarm Optimization, k-Nearest Neighbor, Random Forest, Deep Learning, PSO, Diabetes Mellitus, *data mining*

Optimization of Prediction Algorithms Using Particle Swarm Optimization to Predict Diabetes

Adriansyah Putra

Abstract

Diabetes is a serious global health issue, with its prevalence steadily increasing. Indonesia ranks fifth worldwide, even emerging as the first country in Southeast Asia affected by this disease. Various prevention and management efforts have been undertaken, alongside research on data mining to understand and manage this condition. Utilizing the BIT Mesra public dataset, this study focuses on the performance of three classification algorithms: k-Nearest Neighbor, Random Forest, and Deep Learning in predicting diabetes mellitus after optimization using Particle Swarm Optimization. Clean and complete data serve as a crucial foundation before implementing optimization methods, particularly in handling noisy data and missing values. The research findings indicate that all three algorithms achieve a high level of accuracy. Even after optimization with PSO, k-Nearest Neighbor and Random Forest attain perfect performance with accuracy reaching 100%. Meanwhile, Deep Learning shows an accuracy improvement to 98.95%. However, despite the superior accuracy of k-Nearest Neighbor and Random Forest compared to Deep Learning, statistical tests reveal that all three are equivalent and equally effective in practice. This provides a comprehensive understanding of the potential use of classification algorithms and PSO optimization in predicting diabetes, with an emphasis on balancing good data handling and algorithm optimization.

Keywords: Particle Swarm Optimization, k-Nearest Neighbor, Random Forest, Deep Learning, PSO, Diabetes Mellitus, data mining