

Abstrak

Diabetes merupakan salah satu penyakit yang disebabkan karena gula darah di dalam tubuh yang tinggi atau melampaui batas normal. Penderita diabetes di Indonesia mengalami peningkatan yang cukup signifikan, Riset Kesehatan Dasar menyebutkan penderita diabetes di Indonesia yang semula dari tahun 2013 sebesar 6,9% menjadi 8,5% di tahun 2018 dengan perkiraan jumlah penderita lebih dari 16 juta orang. Oleh karena itu, sangat diperlukan suatu teknologi yang dapat mendeteksi penyakit diabetes dengan kinerja yang baik, tingkat analisis akurat, sehingga penyakit diabetes dapat ditangani lebih awal untuk mengurangi jumlah penderita, kecacatan, dan kematian. Nilai skala yang berbeda tiap atribut pada data Gula Karya Medika dapat mempersulit proses klasifikasi, untuk itu peneliti menggunakan dua metode normalisasi data yaitu *Min-max normalization*, *Z-score normalization*, dan satu tanpa metode normalisasi data dengan *Random Forest* (RF) sebagai metode klasifikasi. *Random Forest* (RF) sebagai metode klasifikasi telah teruji di beberapa penelitian sebelumnya, metode ini mampu menghasilkan kinerja yang baik dengan akurasi yang tinggi. Berdasarkan hasil penelitian, akurasi terbaik dihasilkan model 1 (*Min-max normalization*-RF) sebesar 83.19%, model 2 (*Z-score normalization*-RF) sebesar 82.64%, dan model 3 (Tanpa normalisasi data-RF) sebesar 81.63%. Dari hasil tersebut disimpulkan bahwa model 1 (*Min-max normalization*-RF) lebih baik dibandingkan dua model normalisasi data lainnya dan mampu meningkatkan performansi klasifikasi *Random Forest* sebesar 83.19%.

Kata kunci: *diabetes, klasifikasi, min-max normalization, z-score normalization, random forest.*