

Abstrak

Diabetes merupakan salah satu penyakit penyebab kematian no. 3 terbesar di Indonesia, yang diperkirakan pada tahun 2015 telah ada sebanyak 10 juta penderita diabetes di Indonesia dengan presentasi kematian sebesar 6,7%. Tingginya angka kematian akibat penyakit ini maka dibutuhkan sebuah sistem yang dapat melakukan klasifikasi terhadap pasien kemungkinan diabetes. Penelitian ini mengimplementasikan metode Modified Balanced Random Forest (MBRF) sebagai metode klasifikasi. Algoritma MBRF adalah pengembangan dari algoritma Random Forest yang dapat menangani data dengan jumlah yang besar, metode ini kemudian divalidasi dengan *K-Fold Cross Validation* untuk mencari nilai optimal dari model MBRF. Banyaknya Tree akan mempengaruhi akurasi dari prediksi Random Forest. MBRF adalah pengembangan dari Random Forest yang mampu menangani imbalanced data yang menjadi masalah umum pada klasifikasi machine learning karena penyebaran rasio yang tidak proporsional di setiap kelas. Pengujian menunjukkan model sistem MBRF memiliki hasil akurasi sebesar 97,8%, presentasi ini meningkat dari model Random Forest yang menghasilkan akurasi 92,8%.

Kata kunci : Diabetes, *Supervised Learning*, *Random Forest*, *Modified Balanced Random Forest*, **Klasifikasi**