

BAB 5

SIMPULAN DAN SARAN

5.1 Simpulan

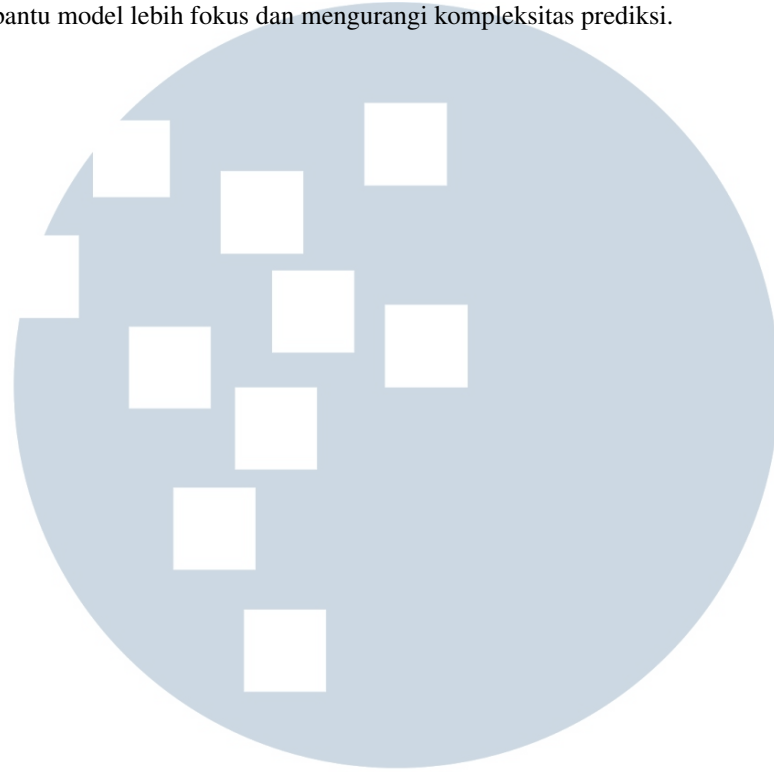
Diabetes merupakan kondisi kronis dengan kadar gula darah tinggi yang dapat menyebabkan komplikasi serius. Faktor risiko meliputi pola makan tidak sehat, kurang aktivitas fisik, faktor genetik, dan gaya hidup buruk. Berdasarkan data WHO 2022, terdapat 537 juta orang dewasa dengan diabetes di dunia, dengan prevalensi yang terus meningkat, terutama di negara berkembang seperti Indonesia. Penelitian ini bertujuan membandingkan algoritma Random Forest dan XGBoost dalam mendeteksi diabetes menggunakan data medical check-up dari Rumah Sakit Pusat Pertamina, dengan riwayat penyakit terdahulu sebagai label diabetes. Teknik evaluasi dilakukan menggunakan confusion matrix, metrik akurasi, F1-Score, serta waktu pembuatan model. Random Forest dan XGBoost dipilih karena kemampuannya dalam menganalisis data dengan banyak fitur dan menangani klasifikasi yang kompleks, termasuk data kesehatan yang tidak seimbang. Random Forest menggunakan pendekatan bagging dengan membangun banyak pohon keputusan secara paralel, sedangkan XGBoost menggunakan boosting dengan membangun pohon secara bertahap. Hasil eksperimen menunjukkan bahwa Random Forest menghasilkan akurasi pengujian 90% dengan F1-Score 0.87 dan waktu pembuatan model 0.39 detik, sementara XGBoost mencapai akurasi 88% dengan F1-Score 0.77 dan waktu pembuatan model 0.51 detik. Random Forest unggul dalam mendeteksi diabetes dengan distribusi data lebih seimbang, sementara XGBoost lebih selektif pada fitur dengan kontribusi signifikan. Penelitian ini menunjukkan bahwa distribusi data yang lebih terfokus dan seimbang dapat meningkatkan performa model dalam mendeteksi diabetes, yang diharapkan dapat mendukung upaya deteksi dini dan pengambilan keputusan klinis yang lebih efektif di masa depan.

5.2 Saran

Dalam penelitian ini, perbandingan antara algoritma Random Forest dan XGBoost dalam memprediksi diabetes menggunakan data Medical Check-up Rumah Sakit Pusat Pertamina telah dilakukan. Berdasarkan hasil percobaan yang telah dilakukan, beberapa saran dapat diberikan untuk penelitian yang serupa, sebagai berikut:

- Pastikan pemilihan fitur dan lakukan tuning hyperparameter untuk meningkatkan kinerja model.
- Gunakan teknik pengolahan data yang lebih baik dan atasi ketidakseimbangan kelas dengan metode seperti SMOTE atau undersampling/oversampling.
- Evaluasi model menggunakan berbagai metrik, termasuk F1-Score, recall, dan precision, untuk mendapatkan hasil yang lebih komprehensif.
- Lakukan penelitian lebih lanjut untuk menguji model pada dataset lain dari rumah sakit atau daerah berbeda untuk melihat generalisasi model.

- Pertimbangkan untuk menggunakan pembagian rentang usia yang lebih spesifik, seperti pada Percobaan 11, untuk meningkatkan akurasi deteksi diabetes. Pembatasan rentang usia dapat membantu model lebih fokus dan mengurangi kompleksitas prediksi.



UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA