

BAB V PENUTUP

A. Kesimpulan

Penelitian ini bertujuan untuk membandingkan performa dua algoritma klasifikasi, yaitu *Support Vector Machine* (SVM) dan *Random Forest* (RF), dalam memprediksi risiko penyakit diabetes berdasarkan *dataset Pima Indian Diabetes Database*. Selain itu, penelitian ini juga bertujuan untuk mengimplementasikan kedua algoritma tersebut agar dapat dimanfaatkan dalam membantu pencegahan dini terhadap risiko diabetes.

Berdasarkan hasil yang diperoleh melalui tahapan *preprocessing*, pembagian data (70:30), penyeimbangan data menggunakan teknik SMOTE, optimasi model menggunakan *GridSearchCV*, serta evaluasi menggunakan metrik akurasi, *precision*, *recall*, *F1-score*, dan ROC AUC, maka dapat disimpulkan:

1. Perbandingan Kinerja Algoritma

Algoritma *Random Forest* (RF) memberikan performa yang lebih baik dibandingkan *Support Vector Machine* (SVM), baik pada data tanpa SMOTE maupun setelah diterapkan SMOTE. Hal ini ditunjukkan dengan nilai evaluasi yang lebih stabil, khususnya pada kelas positif (diabetes). *Random Forest* (RF) dengan SMOTE memiliki nilai *F1-score* dan ROC AUC yang lebih tinggi, sehingga dinilai sebagai model paling optimal dalam memprediksi diabetes pada penelitian ini. Berikut merupakan hasil evaluasi model:

- a. SVM tanpa SMOTE menghasilkan akurasi 73,59%, *precision* 76%, *recall* 49%, *F1-score* 57%, dan AUC 0,8318.
- b. RF tanpa SMOTE menghasilkan akurasi 75,75%, *precision* 77%, *recall* 51%, *F1-score* 59%, dan AUC 0,8322.
- c. SVM dengan SMOTE meningkat menjadi akurasi 77,48%, *precision* 67%, *recall* 72%, *F1-score* 69%, dan AUC 0,8408.

d. RF dengan SMOTE memberikan performa terbaik dengan akurasi 78,35%, *precision* 84%, *recall* 72%, *F1-score* 70%, dan AUC 0,8429.

2. Implementasi Algoritma untuk Prediksi Diabetes

Kedua algoritma, *Support Vector Machine* (SVM) dan *Random Forest* (RF), telah berhasil diimplementasikan dan diuji secara sistematis dalam proses prediksi diabetes. Model terbaik yang dihasilkan (*Random Forest* dengan SMOTE) mampu mengidentifikasi pasien berisiko dengan tingkat akurasi yang cukup tinggi. Hal ini membuktikan bahwa *Machine Learning* dapat dimanfaatkan sebagai alat bantu dalam upaya deteksi dini dan pencegahan risiko diabetes, khususnya dalam sistem pendukung keputusan di bidang kesehatan. Namun, perlu diingat bahwa hasil prediksi ini tidak dapat menggantikan diagnosis medis, sehingga tetap diperlukan pemeriksaan lebih lanjut oleh tenaga medis profesional.

Secara umum, hasil penelitian ini menunjukkan bahwa pemilihan algoritma yang tepat, disertai pengolahan data yang cermat seperti *balancing* data dan tuning parameter, berpengaruh signifikan terhadap keberhasilan model prediksi. Dengan demikian, penelitian ini dapat menjadi dasar untuk pengembangan sistem klasifikasi berbasis kecerdasan buatan dalam bidang kesehatan.

B. Saran

Berdasarkan hasil penelitian dan kesimpulan yang telah diperoleh, penulis menyampaikan saran sebagai berikut: untuk menggunakan dataset yang lebih besar dan beragam di masa depan, agar model yang dikembangkan dapat lebih generalizable dan akurat saat diterapkan pada populasi yang berbeda dari data latih awal.

1. Penelitian ini masih memiliki ruang pengembangan, terutama dalam hal eksplorasi algoritma klasifikasi lainnya seperti *XGBoost*, KNN, atau *deep learning* untuk mengetahui apakah ada algoritma lain yang

dapat memberikan performa lebih baik. Selain itu, penambahan atribut atau fitur medis yang lebih kompleks dapat meningkatkan akurasi model. Disarankan juga untuk menguji model pada data real-time atau data lokal yang lebih relevan dengan populasi Indonesia agar hasilnya lebih aplikatif secara lokal.

2. Hasil dari penelitian ini dapat dijadikan acuan awal dalam mengembangkan sistem pendukung keputusan (*decision support system*) untuk deteksi dini diabetes. Model yang telah diuji dapat diimplementasikan dalam sistem informasi kesehatan masyarakat guna membantu skrining awal pada masyarakat berisiko tinggi, terutama di daerah yang memiliki keterbatasan tenaga medis.
3. Dengan adanya sistem prediksi berbasis *Machine Learning* seperti yang diteliti ini, masyarakat dapat memperoleh manfaat dalam bentuk deteksi dini risiko diabetes. Masyarakat diharapkan lebih sadar akan pentingnya pengecekan kesehatan secara rutin dan memanfaatkan teknologi sebagai bagian dari upaya pencegahan penyakit kronis.
4. Penelitian ini dapat menjadi rujukan bagi mahasiswa dan dosen dalam pengembangan sistem prediktif berbasis data kesehatan. Universitas dapat mendukung penelitian serupa melalui kolaborasi dengan instansi kesehatan, serta memfasilitasi pengembangan sistem berbasis machine learning yang dapat diimplementasikan untuk pengabdian kepada masyarakat.

