

BAB V

SIMPULAN DAN SARAN

5.1 Simpulan

Penelitian ini menghasilkan temuan utama bahwa model *Decision Tree* memberikan performa yang lebih baik dibanding *Artificial Neural Network* (ANN) pada tugas klasifikasi diabetes berbasis data klinis tabular. Pada data pengujian, *Decision Tree* mencapai akurasi 0,8292, sedangkan ANN mencapai akurasi 0,7458. Perbedaan ini memperlihatkan bahwa, pada karakter data yang digunakan, pola pemisahan berbasis aturan yang dibentuk *Decision Tree* lebih sesuai untuk membedakan kelas “diabetes” dan “tidak diabetes” dibanding pola pembelajaran ANN. Hasil evaluasi per kelas juga menguatkan kesimpulan tersebut. *Decision Tree* menunjukkan kemampuan deteksi kasus diabetes yang lebih baik, terlihat dari *recall* kelas diabetes sebesar 0,78, sedangkan ANN memiliki *recall* kelas diabetes sebesar 0,60. Artinya, *Decision Tree* lebih sedikit melewatkan pasien yang sebenarnya diabetes. Temuan ini diperkuat oleh *confusion matrix*: pada data pengujian, *Decision Tree* menghasilkan 19 kasus diabetes yang terklasifikasi sebagai tidak diabetes, sementara ANN menghasilkan 35 kasus dengan kesalahan yang sama. Dalam konteks skrining, kesalahan ini merupakan yang paling berisiko karena dapat membuat pasien diabetes luput terdeteksi dan terlambat mendapatkan pemeriksaan lanjutan. Dengan demikian, hasil penelitian menegaskan bahwa penilaian kinerja tidak cukup berhenti pada akurasi, tetapi perlu melihat *precision*, *recall*, *F1-score*, dan *confusion matrix* agar konsekuensi salah klasifikasi dapat dipahami secara utuh.

Pemilihan dataset Pima dalam penelitian ini relevan karena dataset tersebut merupakan salah satu dataset publik yang paling sering digunakan sebagai rujukan untuk penelitian prediksi diabetes berbasis *machine learning*. Karakteristiknya berupa data klinis tabular dengan delapan variabel prediktor yang umum dikaitkan dengan risiko diabetes, seperti kadar glukosa, tekanan darah, ketebalan lipatan kulit, insulin, BMI, fungsi riwayat keluarga diabetes (*Diabetes Pedigree Function*), usia, serta riwayat jumlah kehamilan. Struktur targetnya berupa keluaran biner yang

sesuai untuk tugas klasifikasi, sehingga memudahkan pengukuran performa model melalui metrik evaluasi standar. Selain itu, dataset ini cukup representatif untuk tujuan pembuktian konsep karena memuat variasi nilai klinis yang luas, sekaligus memiliki tantangan kualitas data berupa nilai nol yang tidak realistis pada beberapa variabel fisiologis. Tantangan tersebut membuat penelitian ini tidak hanya melakukan pemodelan, tetapi juga menunjukkan pentingnya prapemrosesan, khususnya penanganan nilai tidak valid melalui imputasi median agar data lebih layak digunakan dalam pelatihan model.

Pemilihan ANN dan *Decision Tree* pada penelitian ini memiliki alasan akademik dan praktis yang jelas. ANN dipilih karena secara teori mampu mempelajari hubungan non-linear antarvariabel klinis, sehingga potensial menangkap pola risiko diabetes yang tidak selalu berbentuk hubungan linear sederhana. ANN juga banyak digunakan dalam studi klasifikasi medis karena dapat memodelkan interaksi kompleks antarfitur numerik. Sementara itu, *Decision Tree* dipilih karena sifatnya yang transparan dan mudah diinterpretasikan, yaitu menghasilkan aturan keputusan berbentuk percabangan yang dapat ditelusuri. Dalam konteks penelitian yang diarahkan untuk skrining atau dukungan keputusan awal, interpretabilitas menjadi penting karena hasil model perlu dijelaskan secara masuk akal, bukan sekadar menghasilkan angka prediksi. Oleh karena itu, kedua metode ini diposisikan sebagai perbandingan antara model yang kuat dalam menangkap pola kompleks dan model yang kuat dalam keterjelasan aturan. Berdasarkan hasil penelitian, pendekatan berbasis aturan pada *Decision Tree* terbukti lebih sesuai pada data yang digunakan, sekaligus menghasilkan kesalahan kritis yang lebih sedikit dalam mendeteksi pasien diabetes dibanding ANN.

Secara keseluruhan, penelitian ini menyimpulkan bahwa untuk klasifikasi diabetes pada data klinis tabular dengan karakter seperti dataset Pima, *Decision Tree* merupakan model yang lebih unggul dalam penelitian ini, baik dari sisi akurasi maupun kemampuan menangkap kasus diabetes, serta memiliki kelebihan tambahan berupa interpretabilitas yang lebih mudah dipahami. ANN tetap relevan sebagai pendekatan pembelajaran non-linear, namun pada hasil penelitian ini performanya lebih rendah dan cenderung melewatkan lebih banyak kasus diabetes

pada data pengujian, sehingga perlu kehati-hatian apabila dijadikan model utama untuk konteks skrining tanpa penguatan evaluasi lanjutan.

5.2 Saran

Berdasarkan temuan penelitian, model *Artificial Neural Network* disarankan untuk dijadikan rujukan utama ketika tujuan analisis adalah memperoleh performa prediksi yang lebih tinggi pada data medis terstruktur seperti *Pima Indians Diabetes Dataset*. Namun demikian, model ini tetap perlu diposisikan sebagai alat bantu pengambilan keputusan bagi tenaga kesehatan, bukan sebagai pengganti diagnosis klinis, sehingga interpretasi hasil tetap harus dikaitkan dengan pertimbangan medis dan standar praktik kedokteran yang berlaku.

Di sisi lain, algoritma *Decision Tree* masih sangat layak dimanfaatkan ketika pemangku kepentingan, seperti tenaga medis, pengelola rumah sakit, atau pihak manajemen, memerlukan penjelasan yang sederhana mengenai pola dan aturan keputusan yang dihasilkan model. Visualisasi pohon keputusan dapat membantu proses edukasi, audit model, maupun diskusi awal mengenai faktor-faktor yang paling berpengaruh terhadap status diabetes. Mengingat dataset yang digunakan berasal dari populasi Pima yang spesifik, penerapan model pada konteks klinis yang berbeda, termasuk di Indonesia, sebaiknya diawali dengan pengujian ulang menggunakan data lokal agar pola yang dipelajari benar-benar sesuai dengan karakteristik pasien di wilayah tersebut.

Untuk pengembangan penelitian selanjutnya, disarankan agar peneliti menambahkan algoritma lain seperti *Random Forest*, *Gradient Boosting*, atau *XGBoost* sehingga perbandingan kinerja menjadi lebih luas dan komprehensif. Eksperimen lanjutan juga dapat memanfaatkan teknik optimasi *hyperparameter* yang lebih sistematis, pendekatan penanganan ketidakseimbangan kelas, serta penggunaan metrik tambahan seperti *ROC-AUC* dan evaluasi kalibrasi probabilitas agar kualitas prediksi tidak hanya diukur dari akurasi semata.