

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan rumusan masalah dan tujuan penelitian yang telah ditetapkan, maka dapat ditarik beberapa kesimpulan sebagai berikut:

1. Penelitian ini berhasil menerapkan algoritma Light Gradient Boosting Machine (LightGBM) untuk membangun model prediksi risiko stroke menggunakan data kesehatan individu. Proses pengembangan model mencakup tahapan *preprocessing*, mulai dari penanganan *missing values* pada fitur BMI hingga penanganan ketidakseimbangan data (*class imbalance*) menggunakan parameter *scale_pos_weight*.
2. Model LightGBM menunjukkan performa yang baik dalam mengklasifikasikan kelas mayoritas dengan tingkat presisi yang sangat tinggi yaitu 98%. Model juga memiliki kemampuan deteksi yang kuat terhadap pasien berisiko stroke, yang dibuktikan dengan nilai recall yang tinggi yaitu 70%, serta rendahnya angka *false negative* pada hasil pengujian.
3. Melalui metode SHapley Additive exPlanations (SHAP), penelitian ini berhasil mengidentifikasi bahwa variabel usia (*age*), indeks massa tubuh (*bmi*), dan rata-rata kadar glukosa (*avg_glucose_level*) merupakan fitur-fitur yang paling signifikan dalam memengaruhi hasil prediksi risiko stroke. Penggunaan SHAP memberikan transparansi pada model yang bersifat *black-box* sehingga mempermudah interpretasi klinis bagi tenaga medis.

5.2 Saran

Untuk pengembangan penelitian di masa yang akan datang, khususnya pada data stroke terdapat beberapa saran yang dapat peneliti berikan sebagai berikut:

1. Penelitian ini berfokus pada algoritma LightGBM secara spesifik, penelitian selanjutnya disarankan untuk menerapkan teknik optimasi hyperparameter yang lebih luas, seperti penggunaan *Grid Search* atau *Random Search*, untuk meningkatkan nilai presisi pada kelas minoritas. Peneliti selanjutnya juga dapat menggunakan dataset yang lebih besar atau data kesehatan lokal

untuk mendapatkan model yang lebih relevan dengan karakteristik masyarakat di wilayah tertentu.

2. Penelitian ini hanya menggunakan analisis SHAP secara global saja sehingga untuk peneliti selanjutnya dapat mengkombinasikan antara analisis global dan lokal sehingga mampu meningkatkan transparansi model dan membantu interpretasi keputusan klinis berbasis data.
3. Penelitian ini juga hanya sampai tahapan pembuatan prototipe saja, sehingga untuk peneliti selanjutnya diharapkan dapat lebih mengembangkan penelitian ini ke tahapan menjadi suatu sistem yang dapat digunakan di dalam dunia medis khususnya dalam prediksi risiko penyakit stroke.

