

BAB V

SIMPULAN DAN SARAN

5.1 Simpulan

Klasifikasi penyakit *stroke* yang dilakukan dengan tiga algoritma berbasis pohon keputusan yaitu *Decision Tree*, *Extra tree classifier*, dan *Xgboost*. Model *Decision Tree* menunjukkan performa yang kuat dengan menunjukkan hasil akurasi, presisi, dan sensitivitas diatas 95% dengan waktu pemrosesan yang relatif lama karena waktu pemrosesannya berada di urutan kedua dibanding dengan *Extra tree classifier*, dan *Xgboost*. *Extra tree classifier* sendiri menunjukkan tingkat akurasi yang lebih rendah yaitu 89% namun waktu pemrosesannya merupakan yang tercepat. *Xgboost* menampilkan hasil yang memuaskan dengan performa terbaik di antara ketiganya dengan nilai akurasi di atas 97% namun membutuhkan waktu pemrosesan yang paling lama. Evaluasi dari hasil ketiganya dalam klasifikasi memberikan kelebihan dan kekurangan dari masing-masing algoritma sehingga dengan sesuai kebutuhan data ketiga algoritma menunjukkan performa yang baik.

Penerapan algoritma optimasi berbasis *swarm intelligence* seperti *Particle swarm optimization (PSO)* dan *Ant Colony Optimization (ACO)* sebagai metode untuk klasifikasi penyakit *stroke* berbasis pohon keputusan melibatkan proses iteratif algoritma mencari *subset* fitur yang paling relevan dari sekumpulan fitur yang ada sesuai dengan sifatnya. *ACO* meniru perilaku semut dalam menemukan jalur terbaik menuju sumber makanan dengan meninggalkan jejak feromon yang berfungsi sebagai informasi yang akan diikuti oleh semut lainnya. Sementara itu, *PSO* berusaha untuk menemukan solusi terbaik dengan memperbarui posisi partikel dalam ruang pencarian berdasarkan pengalaman terbaiknya dan pengalaman dari partikel tetangganya. Penerapan *ACO* dan *PSO* sebagai seleksi fitur diharapkan dapat menghasilkan *subset* yang lebih optimal, meningkatkan klasifikasi dan

mengurangi dimensi fitur yang tidak relevan, sehingga membantu dalam identifikasi dan diagnosis penyakit *stroke* dengan lebih efisien.

Seleksi fitur menggunakan *PSO* dan *ACO* berhasil meningkatkan akurasi, presisi, sensitivitas dan terutama waktu pemrosesan. Hasil menggunakan seleksi fitur *PSO* pada *Xgboost* berhasil mereduksi fitur sampai 17 fitur terpilih dengan hasil akurasi 97.68%, presisi 97.79%, sensitivitas 97.67% dan mampu menurunkan waktu pemrosesan menjadi 1.5429 detik yang menunjukkan performa terbaik dari sembilan model. Kenaikan performa tertinggi adalah seleksi fitur menggunakan *ACO* pada *Extra Tree Classifier* yang meningkatkan akurasi dari 88.57% hingga 93.75%, presisi 94.07%, dan sensitivitas 93.73%, serta mengurangi waktu pemrosesan hingga menjadi yang tercepat yaitu 0.2508 detik. Hasil menunjukkan bahwa seleksi fitur menggunakan algoritma *PSO* dan *ACO* mampu meningkatkan efisiensi dan efektivitas algoritma berbasis pohon keputusan dalam klasifikasi penyakit stroke.

5.2 Saran

Berdasarkan hasil yang didapatkan dari penelitian ini, terdapat beberapa saran yang dapat dipertimbangkan untuk dilakukan penelitian selanjutnya yang berkaitan dengan *feature selection* dan klasifikasi.

- 1) Penerapan teknik *pruning* pada algoritma berbasis keputusan dan melihat pengaruhnya terhadap metrik evaluasi yang diukur. Teknik *pruning* dapat dilakukan pada algoritma berbasis pohon keputusan yang fungsinya dapat mencegah *overfitting*. Melalui penerapan teknik *pruning* yang sesuai diharapkan diperoleh hasil yang lebih baik untuk model klasifikasi *dataset indicator of heart disease 2022*.
- 2) Melakukan komparasi berbagai macam teknik *resample* untuk mengetahui dampaknya terhadap akurasi karena *dataset indicator of heart disease* memiliki ketidakseimbangan pada kelas minoritas dan mayoritas pada

variabel *Hadstroke*. Membandingkan metode-metode seperti *Oversampling*, *Undersampling*, *SMOTE (Synthetic Minority Over-sampling Technique)*, dan *ADASYN (Adaptive Synthetic Sampling Approach)*.

- 3) Melakukan komparasi algoritma berbasis pohon keputusan yang digunakan pada penelitian ini dan yang tidak digunakan pada penelitian ini seperti *Random forest*, *Extra trees classifiers*, dan *Spanning Tree*. Dengan menganalisis perbedaan hasil dari masing-masing algoritma diharapkan dapat memberikan wawasan lebih dalam tentang kelebihan dan keterbatasan algoritma berbasis pohon keputusan.
- 4) Menggunakan algoritma *Swarm Intelligence* yang berbeda seperti *Bee Colony Optimization*, *Salp Swarm Algorithm* dan *Firefly Algorithm* untuk seleksi fitur untuk dikomparasikan pada algoritma klasifikasi lain untuk mendapatkan hasil kombinasi terbaik untuk akurasi, presisi, sensitivitas dan waktu pemrosesan.
- 5) Melakukan analisis mendalam terhadap fitur-fitur terpilih yang belum mendapatkan referensi dari artikel ilmiah. Langkah ini merupakan langkah penting untuk memastikan validitas fitur-fitur tersebut sebagai faktor penyebab stroke. Dengan melakukan analisis mendalam terhadap fitur-fitur terpilih diharapkan memperoleh pemahaman yang lebih akurat dan komprehensif.