

BAB V

SIMPULAN DAN SARAN

5.1 Simpulan

Penelitian yang dilakukan merupakan penerapan *educational data mining* dalam pembuatan sistem prediksi waktu kelulusan mahasiswa. Penelitian membandingkan 3 algoritma klasifikasi yang dioptimalkan menggunakan algoritma Particle Swarm Optimization, yaitu Naïve Bayes, Support Vector Machine, dan Neural Network yang digunakan sebagai model prediksi waktu kelulusan mahasiswa. *Framework* CRISP-DM digunakan untuk mengolah data dengan 6 tahapan yaitu *business understanding, data understanding, data preparation, modeling, evaluation, dan deployment*.

Neural Network yang dioptimasi menggunakan algoritma Particle Swarm Optimization mendapatkan hasil terbaik dibandingkan Naïve Bayes maupun Support Vector Machine, dimana algoritma PSO dapat memperoleh akurasi sebanyak 8% lebih tinggi dibandingkan Neural Network dengan parameter *default*. Neural Network dengan PSO ditentukan sebagai model terbaik karena memiliki nilai akurasi, presisi, sensitivitas, spesifisitas, dan *f1-score* yang tertinggi yang telah dibahas secara lebih detail pada sub-bab hasil dan diskusi. Evaluasi lanjutan juga dilakukan dengan menggunakan *confusion matrix* dimana algoritma Neural Network memiliki total *true positive* dan *true negative* terbanyak, serta *false positive* dan *false negative* paling sedikit pada sebagian besar kelas luaran.

Model dengan algoritma Neural Network sebagai algoritma terbaik disimpan dengan ekstensi .h5 dan digunakan untuk implementasi pada *website*. Pengimplementasian diawali dengan melakukan prediksi data mahasiswa berstatus aktif sebagai data baru menggunakan model prediksi yang telah ditentukan. Hasil label dari prediksi akan disimpan pada *database*. Data yang tersimpan pada *database* ditampilkan pada *website* dengan 2 tipe pengguna, yaitu mahasiswa dan dosen pembimbing.

5.2 Saran

Penelitian lanjutan dapat dilakukan untuk meningkatkan hasil performa algoritma hingga menemukan algoritma yang lebih baik serta sesuai terhadap prediksi waktu kelulusan mahasiswa. Mengacu pada penelitian ini, berikut terdapat beberapa hal sebagai saran untuk penelitian lanjutan:

- 1) Memperbanyak jumlah data yang digunakan untuk pelatihan model sehingga model semakin dapat mempelajari hubungan pola performa mahasiswa dengan waktu kelulusannya.
- 2) Menambahkan algoritma perbandingan seperti Decision Tree, Random Forest, dan algoritma lainnya, serta melakukan perbandingan *hyperparameter tuning* lainnya, seperti grid search atau random search untuk mencari algoritma terbaik.
- 3) Memperbarui model klasifikasi dengan data mahasiswa terbaru sebagai data *training* untuk menjaga relevansi pada saat diimplementasikan di *website* akademik resmi Universitas Multimedia Nusantara

