

BAB V

SIMPULAN DAN SARAN

3.5 Kesimpulan

Berdasarkan analisis model Machine Learning pada klasifikasi diagnosa tingkat stress pada mahasiswa tingkat akhir menggunakan algoritma Random Forest, diperoleh bahwa:

- a. Model dengan oversampling SMOTE menghasilkan performa terbaik pada rasio 80% data latih, 20% data uji dengan nilai $K=5$ mendapatkan akurasi sebesar 54%, precision sebesar 59% dan recall sebesar 60%.
- b. Perbandingan akurasi yang didapatkan antara metode *Random Forest* tanpa oversampling SMOTE dan *Random Forest* dengan oversampling SMOTE memiliki perbedaan yang cukup signifikan. Hasil akurasi sebesar 71% menggunakan metode *Random Forest* dengan oversampling SMOTE mampu ditingkatkan pada data yang *imbalanced* yang sebelumnya tanpa menggunakan SMOTE memiliki akurasi sebesar 54%. Sehingga dapat disimpulkan bahwasannya metode tersebut berhasil untuk mempengaruhi dan meningkatkan evaluasi klasifikasi studi kasus klasifikasi diagnosa mahasiswa tingkat akhir.
- c. Oversampling dengan metode SMOTE berhasil menyelesaikan tantangan ketidakseimbangan data pada kelas minoritas. Berdasarkan nilai recall, pendekatan oversampling SMOTE menunjukkan performa yang lebih unggul dalam mengevaluasi seberapa baik model memprediksi data secara akurat. Selain itu, pendekatan oversampling SMOTE tidak memerlukan volume besar data latih, karena metode ini menghasilkan data latih tambahan secara sintetis, secara efektif meningkatkan jumlah sampel dalam data latih.

5.2 Saran

Berdasarkan hasil analisis penelitian yang telah dilakukan, beberapa rekomendasi untuk penelitian mendatang dapat disarankan sebagai berikut:

- a. Memerlukan pembaruan serta penambahan data agar model yang dibuat dapat mempelajari berbagai jenis data atau karakteristik terbaru terkait klasifikasi diagnosis tingkat stres mahasiswa tingkat akhir.
- b. Melakukan eksperimen menerapkan algoritma lain dari *Ensemble Learning* guna membandingkan evaluasi yang dihasilkan dan menilai apakah penggunaan metode oversampling SMOTE juga berdampak terhadap evaluasi algoritma tersebut.

- c. Melakukan uji coba menerapkan metode *Resampling* data lainnya untuk menentukan metode yang paling optimal dalam melakukan klasifikasi diagnosis tingkat stres pada mahasiswa tingkat akhir.