

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil analisis yang telah dilakukan, dapat disimpulkan bahwa kinerja algoritma *Random Forest* dan *XGBoost* untuk analisis perilaku konsumen *e-commerce* memberikan performa yang sangat baik. Kedua algoritma tersebut mencapai nilai *accuracy* di atas 98%, dengan algoritma *Random Forest* sedikit lebih unggul (98.43%) dibandingkan algoritma *XGBoost* (98.40%). Begitu pula pada nilai *recall* algoritma *Random Forest* yang dapat mengungguli algoritma *XGBoost* (98.40% vs 98.00%). Sebaliknya, algoritma *XGBoost* lebih unggul daripada algoritma *Random Forest* pada nilai *precision* (97.23% vs 96.48%) dan nilai F1 (97.61% vs 97.43%). Adapun nilai AUC yang sangat tinggi (0.9992 untuk algoritma *Random Forest* dan 0.9993 untuk algoritma *XGBoost*) mengindikasikan bahwa kedua algoritma tersebut memiliki kemampuan untuk membedakan kelas dengan hampir sempurna.

Dengan demikian, dalam menangani *dataset* besar yang tidak seimbang dan kompleks untuk analisis perilaku konsumen *e-commerce*, algoritma *XGBoost* menunjukkan sedikit keunggulan daripada algoritma *Random Forest*. Walaupun perbedaannya kecil, algoritma *XGBoost* mampu memperoleh nilai *precision* yang lebih tinggi, nilai F1 yang lebih baik, dan nilai AUC yang sedikit lebih tinggi. Selain itu, algoritma *XGBoost* memiliki kemampuan yang lebih baik dalam menangani ketidakseimbangan kelas. Hal ini dibuktikan dengan hasil prediksi yang sedikit lebih seimbang (253 loyal dan 583 tidak loyal) dibandingkan dengan hasil prediksi algoritma *Random Forest* (256 loyal dan 580 tidak loyal). Oleh sebab itu, perlu diketahui bahwa kedua algoritma ini mampu berkinerja dengan sangat baik dan kompatibel untuk menangani tugas analisis perilaku konsumen *e-commerce*.

5.2 Saran

Adapun saran yang dapat diberikan sebagai upaya perbaikan maupun pengembangan penelitian di masa depan, yaitu dengan memanfaatkan *model tuning* untuk meneliti lebih lanjut pengaruhnya terhadap hasil prediksi pada kedua model. Selain itu, dapat menambahkan teknik *ensemble learning* lain untuk membantu menentukan algoritma mana yang berkinerja lebih optimal untuk analisis perilaku konsumen *e-commerce*. Ataupun dapat mencoba menganalisis perilaku konsumen *e-commerce* dengan menggunakan *dataset* berukuran kecil atau bahkan yang lebih besar dari ukuran *dataset* ini.

