

BAB V

PENUTUP

5.1 Kesimpulan

Kesimpulan berikut ini diperoleh berdasarkan hasil penelitian yang menggunakan berbagai algoritma machine learning:

1. Algoritma machine learning yang digunakan menunjukkan kinerja yang baik sepanjang proses klasifikasi mulai dari prapemrosesan hingga evaluasi. Dengan nilai ROC-AUC diatas 0,81 seluruh model menunjukkan kemampuan yang baik dalam membedakan antara pelanggan yang churn dan yang tidak churn. Selain itu, XGBoost menghasilkan perkiraan probabilitas yang akurat dan berhasil mengidentifikasi pelanggan berisiko tinggi.
2. XGBoost dan Gradient Boosting adalah algoritma yang menunjukkan kinerja terbaik dalam penelitian ini. XGBoost mencapai akurasi sebesar 80,27% dan ROC-AUC sebesar 0,8415, sedangkan Gradient Boosting mencapai 80,13% dan 0,8411. Kedua algoritma berbasis boosting ini unggul untuk prediksi churn.

Menurut hasil analisis interpretabilitas menggunakan Feature Importance dan Permutation Importance, penyebab utama penurunan tenure, Total Charges, Monthly Charges, jenis Contract, Internet Service Fiber optic, serta Payment Method Electronic check. Kemungkinan churn meningkat dengan kontrak bulanan dan langganan yang singkat. Hasil ini menunjukkan bahwa model itu akurat dan dapat menemukan faktor-faktor utama yang menyebabkan churn. Model optimal (XGBoost) kemudian diintegrasikan ke dalam prototipe sistem prediksi berbasis web menggunakan Streamlit. Prototipe ini membantu perusahaan menemukan pelanggan berisiko tinggi untuk rencana retensi yang lebih terarah

5.2 Saran

Berdasarkan hasil penelitian ini, terdapat beberapa saran untuk penelitian selanjutnya:

1. Untuk meningkatkan kinerja model, penelitian selanjutnya dapat melakukan penyesuaian hiperparameter yang lebih mendalam atau menerapkan strategi optimisasi seperti Grid Search atau Bayesian Optimization.
2. Teknik seperti SMOTE atau strategi resampling lainnya dapat digunakan untuk lebih memperbaiki ketidak seimbangan kelas dan mengevaluasi dampaknya terhadap peningkatan recall pada kelas churn.
3. Untuk membuat model lebih realistis sesuai dengan kondisi nyata perusahaan telekomunikasi, penelitian ini dapat diperluas dengan memanfaatkan data real- time atau statistik industri.

Penelitian selanjutnya juga dapat meneliti teknik interpretabilitas lain seperti SHAP atau LIME untuk memahami kontribusi setiap fitur dengan lebih baik.