

ABSTRAK

ARI AHMAD DAHRIL, 2026. Prediksi Pelanggan Potensial Penunggak Tagihan Listrik menggunakan Algoritma *Gaussian Naive Bayes* pada Data PLN (Dibimbing oleh Rizki Yusliana Bakti, S.T., M.T. dan Ir. Muhammad Faisal, S.Si.,M.T.,Ph.D.,IPM).

Penunggakan pembayaran tagihan listrik pascabayar merupakan salah satu permasalahan yang berdampak pada arus kas PT Perusahaan Listrik Negara (Persero). Mekanisme penanganan tunggakan yang masih bersifat reaktif menyebabkan biaya operasional yang tinggi sehingga diperlukan pendekatan prediktif berbasis data untuk mengidentifikasi pelanggan berisiko sejak dini. Penelitian ini menerapkan algoritma *Gaussian Naive Bayes* untuk mengklasifikasikan tingkat risiko pelanggan listrik pascabayar PLN UP3 Makassar Selatan ke dalam tiga kategori, yaitu Rendah, Sedang, dan Tinggi, berdasarkan pola historis pembayaran. Data riwayat pembayaran diagregasi menjadi 670 data pelanggan dengan delapan atribut prediktor. Tahapan penelitian meliputi preprocessing data, standarisasi menggunakan *StandardScaler*, seleksi fitur dengan *SelectKBest (ANOVA F-test)*, pembagian data 80:20 menggunakan *Stratified Split*, optimasi *hyperparameter var_smoothing* melalui *Bayesian Optimization*, serta evaluasi akhir menggunakan *5-Fold Stratified Cross Validation*. Hasil pengujian menunjukkan model memperoleh akurasi sebesar 85,07%, dengan *precision*, *recall*, dan *F1-score* tertinggi pada kelas Rendah (0,90; 0,84; 0,87) dan performa lebih rendah pada kelas Tinggi (0,74; 0,57; 0,64) akibat jumlah data yang tidak seimbang. Atribut jumlah pembayaran tepat waktu, jumlah keterlambatan, dan status pembayaran bulan terakhir merupakan fitur yang paling berpengaruh dalam proses klasifikasi. Hasil ini menunjukkan bahwa *Gaussian Naive Bayes* efektif digunakan sebagai model pendukung *Early Warning System* bagi PLN dalam mengidentifikasi pelanggan yang berpotensi menunggak.

Kata Kunci: Tunggakan Listrik, *Gaussian Naive Bayes*, Klasifikasi, Risiko Pelanggan, Data Mining.