# Penerapan Metode K-Nearest Neighbor Dan Random Forest Dalam Memprediksi Risiko Gangguan Jaringan Listrik Berdasarkan Data Pemeliharaan Perusahaan Listrik Negara (PLN)

## ABSTRAK

Gangguan pada jaringan distribusi listrik dapat menurunkan keandalan sistem dan kualitas pelayanan kepada pelanggan. Strategi pemeliharaan konvensional yang bersifat korektif dan preventif seringkali tidak cukup untuk memprioritaskan aset dengan risiko gangguan lebih tinggi. Penelitian ini menerapkan algoritma K-Nearest Neighbor (K-NN) dan Random Forest untuk memprediksi kategori risiko gangguan jaringan listrik berdasarkan data pemeliharaan historis dari PLN UP3 Lhokseumawe. Data melalui tahap preprocessing meliputi pembersihan, transformasi, dan normalisasi, kemudian dilabeli menggunakan metode Failure Mode and Effects Analysis (FMEA) ke dalam tiga kategori risiko, yaitu rendah, sedang, dan tinggi. Model dievaluasi menggunakan confusion matrix, precision, recall, dan F1-score. Hasil pengujian menunjukkan model K-NN mencapai akurasi 97,6% dengan rata-rata precision 95,20%, recall 98,00%, dan F1-score 96,40%, sedangkan model Random Forest mencapai akurasi 97,6% dengan rata-rata precision, recall, dan F1-score sebesar 96,2%. Analisis kepentingan fitur pada Random Forest mengidentifikasi Pengukuran, Temuan Tingkat 1, dan Inspeksi Tingkat 1 sebagai prediktor paling berpengaruh. Kedua algoritma terbukti efektif dalam mengidentifikasi tingkat risiko gangguan dan diharapkan dapat mendukung pemeliharaan prediktif berbasis data guna meningkatkan keandalan sistem distribusi tenaga listrik.

**Kata Kunci**: Prediksi risiko gangguan, K-Nearest Neighbor,Random Forest, Pemeliharaan prediktif, Sistem distribusi daya, PLN

# *Application of the K-Nearest Neighbor and Random Forest Methods to Predict the Risk of Power Grid Disruptions Based on Maintenance Data from the State Electricity Company (PLN)*

## *ABSTRACT*

*Disruptions to the electricity distribution network can reduce system reliability and customer service quality. Conventional corrective and preventive maintenance strategies are often insufficient to prioritize assets with a higher risk of disruption. This study applies the K-Nearest Neighbor (K-NN) and Random Forest algorithms to predict category the risk of power grid disruptions based on historical maintenance data from PLN UP3 Lhokseumawe. The data underwent preprocessing stages including cleaning, transformation, and normalization, then were labeled using the Failure Mode and Effects Analysis (FMEA) method into three risk categories: low, medium, and high. The model was evaluated using a confusion matrix, precision, recall, and F1-score. Test results showed that the K-NN model achieved 97.6% accuracy with an average precision of 95.20%, a recall of 98.00%, and an F1-score of 96.40%, while the Random Forest model achieved 97.6% accuracy with an average precision, recall, and F1-score of 96.2%. Feature importance analysis in Random Forest identified Measurements, Level 1 Findings, and Level 1 Inspections as the most influential predictors. Both algorithms proved effective in identifying the level of disruption risk and are expected to support data-driven predictive maintenance to improve the reliability of the power distribution system.*

***Keywords:*** *Disruption risk prediction, K-Nearest Neighbor, Random Forest, Predictive maintenance, Power distribution system, PLN*