

BAB I

PENDAHULUAN

1.1 Latar Belakang

Keandalan sistem distribusi tenaga listrik merupakan faktor krusial dalam mendukung pembangunan nasional dan peningkatan kesejahteraan masyarakat. Listrik yang sebelumnya dianggap sebagai kebutuhan sekunder kini telah menjadi kebutuhan primer yang menopang hampir seluruh aktivitas kehidupan modern, mulai dari sektor rumah tangga, industri, hingga pelayanan publik [1]. Dalam sistem distribusi tenaga listrik, keandalan pasokan harus terus dijaga karena gangguan atau kegagalan dapat terjadi akibat kerusakan peralatan, gangguan eksternal, kondisi lingkungan, maupun kesalahan manusia [2]. Di Indonesia, PT PLN (Persero) terus menghadapi tantangan dalam menjaga keandalan jaringan distribusi yang semakin kompleks untuk menyalurkan tenaga listrik dari sumber daya besar ke konsumen, seiring dengan gangguan yang sering terjadi baik secara internal seperti kerusakan komponen maupun eksternal seperti gangguan oleh pohon, hewan, serta kondisi cuaca ekstrem [3]. Gangguan listrik yang tidak terencana dapat menimbulkan kerugian ekonomi yang signifikan, kondisi ini menunjukkan bahwa pendekatan prediksi gangguan yang lebih adaptif dan proaktif sangat dibutuhkan.

Selama ini, pemeliharaan jaringan distribusi PLN masih banyak mengandalkan pendekatan *corrective maintenance* dan *preventive maintenance*. *Corrective maintenance* dilakukan setelah gangguan terjadi, sementara *preventive maintenance* dilaksanakan berdasarkan jadwal tertentu tanpa mempertimbangkan kondisi aktual peralatan [4]. Data pemeliharaan historis ini menjadi peluang besar untuk mengimplementasikan *predictive maintenance* yang memanfaatkan pembelajaran mesin untuk memprediksi risiko gangguan sebelum terjadi kegagalan sistem [5]. Metode pembelajaran mesin seperti K-Nearest Neighbor (K-NN) telah menunjukkan potensi tinggi dalam analisis prediktif berbagai domain. K-NN adalah algoritma klasifikasi berbasis perhitungan jarak yang menempatkan objek baru pada kelas tertentu berdasarkan kedekatan dengan data historis [6]. Keunggulan K-

NN antara lain kesederhanaan implementasi, tidak membutuhkan asumsi distribusi data, serta efektif untuk data multivariat.

Berdasarkan penelitian sebelumnya yang dikutip dalam jurnal [7], hasil penelitian ini menghasilkan nilai akurasi data terhadap klasifikasi penjualan produk pada aplikasi RapidMiner sebesar 57.83%. Berdasarkan hasil nilai akurasi dapat disimpulkan bahwa model K-Nearest Neighbor mampu memberikan prediksi sebesar 57,83% dari keseluruhan data yang diuji. Dari 100 kasus yang diuji, model K-Nearest Neighbor memberikan hasil yang benar sekitar 58 kali, dan salah sekitar 42 kali. Akurasi data digunakan sebagai salah satu matriks dalam mengevaluasi performa suatu model terutama dalam konteks klasifikasi. akurasi data saja tidak selalu memberikan gambaran hasil yang lengkap mengenai performa suatu model, maka dari itu diperlukan pertimbangan matriks lain seperti *precision*, *recall* dan *F1-score*.

Pada penelitian yang dilakukan oleh Rahmawati & Indriyanti dengan judul “Penerapan Algoritma K-Nearest Neighbor untuk Prediksi Penjualan Produk Digital” juga menggunakan K-NN untuk memprediksi penjualan produk digital pada PT.Global Indo Multimedia dan menghasilkan akurasi 71%-90%, yang menegaskan bahwa performa K-NN sangat dipengaruhi oleh karakteristik data serta parameter model yang digunakan. Sebagai pembanding, penelitian ini turut mengintegrasikan metode Random Forest yang merupakan algoritma ensemble berbasis pohon keputusan. Random Forest bekerja dengan membangun sejumlah pohon keputusan secara acak kemudian menggabungkan hasilnya melalui mekanisme voting untuk menghasilkan prediksi akhir yang memiliki akurasi tinggi [8], keunggulan Random Forest terletak pada kemampuannya menangani overfitting, kemampuannya mengukur tingkat kepentingan setiap fitur dalam proses klasifikasi, dan menghasilkan akurasi prediksi yang baik [9]. Salah satunya penelitian oleh [10] yang menghasilkan akurasi prediksi sebesar 83%.

Meskipun beberapa penelitian tersebut menunjukkan keberhasilan penerapan K-NN maupun Random Forest, sebagian besar masih terbatas pada domain penjualan produk atau klasifikasi masalah pelanggan, dan belum menyentuh aspek prediksi risiko gangguan pada infrastruktur jaringan distribusi listrik secara

spesifik. Prediksi risiko gangguan jaringan listrik memiliki karakteristik dan tantangan yang berbeda, karena melibatkan data teknis pemeliharaan yang kompleks, variasi jenis gangguan yang beragam, serta dampak yang sangat signifikan jika terjadi kesalahan prediksi. Oleh karena itu, diperlukan penelitian khusus yang fokus pada penerapan K-NN maupun Random Forest untuk memprediksi risiko gangguan jaringan listrik dengan memanfaatkan data pemeliharaan historis yang dimiliki PLN.

Dalam upaya mengatasi kompleksitas data teknis tersebut, penelitian ini mengintegrasikan metode *Failure Mode and Effects Analysis* (FMEA) sebagai instrumen pelabelan risiko yang objektif. FMEA digunakan untuk mengonversi data mentah pemeliharaan menjadi nilai *Risk Priority Number* (RPN) berdasarkan parameter *Severity*, *Occurrence*, dan *Detection* guna mengidentifikasi titik keausan kritis pada jaringan distribusi [11]. Hasil dari analisis FMEA inilah yang kemudian menjadi landasan bagi algoritma K-Nearest Neighbors (K-NN) dan Random Forest untuk membangun serta membandingkan model klasifikasi dengan melakukan pelabelan tingkat risiko gangguan ke dalam kategori tinggi, sedang, atau rendah. Fokus penelitian mencakup kelengkapan data, pola variabel pemeliharaan, serta konsistensi data historis di PT PLN UP3 Lhokseumawe yang mencakup data pemeliharaan gardu.

Dengan demikian penelitian yang berjudul “Penerapan Metode K-Nearest Neighbor dan Random Forest dalam Memprediksi Risiko Gangguan Jaringan Listrik Berdasarkan Data Pemeliharaan PLN” diharapkan dapat membantu memprediksi potensi gangguan jaringan sehingga proses pemeliharaan menjadi lebih efektif. Selain itu, pemanfaatan metode pembelajaran mesin dalam bidang ketenagalistrikan mendukung pengolahan dan pemanfaatan data secara cerdas sebagai dasar pengambilan keputusan dalam keandalan sistem. Dengan memanfaatkan pola dari data historis pemeliharaan, sistem yang diusulkan diharapkan mampu memberikan rekomendasi yang lebih akurat dan objektif dibandingkan pendekatan konvensional, sehingga risiko gangguan dapat diminimalkan, alokasi sumber daya pemeliharaan menjadi lebih tepat sasaran, serta kontinuitas pasokan listrik kepada pelanggan tetap terjaga secara berkelanjutan.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, rumusan masalah dari penelitian yaitu:

1. Bagaimana menerapkan metode K-Nearest Neighbors (K-NN) dan Random Forest dalam memprediksi risiko gangguan jaringan listrik berdasarkan data pemeliharaan PLN?
2. Bagaimana keakuratan metode K-Nearest Neighbors (K-NN) dan Random Forest dalam memprediksi kategori risiko gangguan jaringan listrik (tinggi, sedang, rendah) berdasarkan data yang digunakan?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah tersebut, penelitian ini bertujuan untuk:

1. Menerapkan metode K-Nearest Neighbors (K-NN) dan Random Forest untuk memprediksi risiko gangguan jaringan listrik berdasarkan data pemeliharaan PLN
2. Mengukur dan menganalisis tingkat akurasi metode K-Nearest Neighbors (K-NN) dan Random Forest dalam memprediksi kategori risiko gangguan jaringan listrik.

1.4 Batasan Masalah

Untuk menghindari meluasnya materi pembahasan tugas akhir ini, maka pembahasan akan dibatasi mencakup hal-hal berikut:

1. Data yang digunakan berasal dari PLN UP3 Lhokseumawe, dengan rentang waktu pengambilan data pada tahun 2025.
2. Data pemeliharaan yang digunakan hanya mencakup data pemeliharaan gardu
3. Prediksi risiko gangguan diklasifikasikan ke dalam tiga kategori: risiko tinggi, sedang, dan rendah.
4. Model hanya mengevaluasi data historis pemeliharaan tanpa mempertimbangkan faktor eksternal, seperti cuaca atau kejadian alam lainnya.

1.5 Manfaat Penelitian

Manfaat dari penelitian adalah sebagai berikut:

1. Manfaat Praktis

Penelitian ini diharapkan dapat membantu PT PLN (Persero) UP3 Lhokseumawe dalam mengoptimalkan strategi pemeliharaan jaringan distribusi listrik dengan mengubah pendekatan dari *reactive maintenance* menjadi *predictive maintenance*. Sistem prediksi risiko gangguan ini dapat membantu memprioritaskan jadwal pemeliharaan berdasarkan tingkat risiko (tinggi, sedang, rendah), sehingga alokasi sumber daya (tenaga kerja, material, dan anggaran) menjadi lebih efisien. Dengan kemampuan memprediksi potensi gangguan lebih awal, diharapkan dapat mengurangi frekuensi pemadaman listrik tidak terencana, menurunkan biaya perbaikan darurat, serta meningkatkan keandalan sistem distribusi yang berdampak langsung pada peningkatan kualitas layanan kepada masyarakat

2. Manfaat Teoritis

Penelitian ini memberikan kontribusi pada pengembangan ilmu pengetahuan dengan memperkaya literatur penerapan algoritma K-Nearest Neighbor (K-NN) dan Random Forest dalam konteks sistem distribusi tenaga listrik, yang selama ini lebih banyak diterapkan pada domain retail, kesehatan, atau finansial. Selain itu, penelitian ini dapat menjadi basis untuk penelitian lanjutan dalam hal komparasi algoritma machine learning, integrasi dengan teknologi IoT untuk monitoring real-time di lingkungan industri ketenagalistrikan. Hasil penelitian juga dapat digunakan sebagai studi kasus pembelajaran bagi mahasiswa Sistem Informasi dan Teknik Elektro dalam memahami penerapan praktis data mining pada permasalahan industri.