

Association Pattern Analysis of Recurrent Faults in PLN Distribution Networks Using the Apriori Algorithm

Dwina Kuswardani^{1*}, Muhamad Habibi²

^{1,2}Pendidikan Profesi Insinyur, Sekolah Pasca Sarjana, Institut Teknologi PLN, Jakarta, Indonesia

Email: ^{1*}dwina2580002@itpln.ac.id, ¹dwina@itpln.ac.id, ²muhamadhabibi99@gmail.com

(*: corresponding author)

Abstract - Repeated disturbances in power distribution networks can reduce system reliability and increase maintenance costs. However, maintenance activities are often performed reactively because recurring disturbance patterns have not been systematically identified from historical operational data. This study aims to identify recurring disturbance patterns in the distribution network of PT PLN (Persero) by applying the Apriori algorithm within the CRISP-DM framework. The dataset consisted of 29,446 historical disturbance records containing disturbance type, cause, and location. After the data preparation stage, which included data cleaning, standardization, and transformation, 10,274 valid transactions were obtained for association rule mining. The Apriori algorithm was implemented using a minimum support threshold of 0.5%, a minimum confidence threshold of 70%, and lift values greater than 1 as the evaluation criterion. The analysis produced 57 frequent 1-itemsets, 180 frequent 2-itemsets, 35 frequent 3-itemsets, and 64 association rules. The results indicate that Core FO Problem, FOC – Putus Core, and Putus Kabel FO are the most frequently occurring disturbance patterns. The strongest association rule, Banten and FOT – Software → Konfigurasi Di Router Problem, achieved a support of 0.993%, a confidence of 81.6%, and the highest lift value of 30.936, indicating a strong positive relationship between the antecedent and consequent. These findings demonstrate that the Apriori algorithm is effective in discovering recurring disturbance patterns from historical operational data and can provide maintenance-oriented knowledge to support inspection prioritization and preventive maintenance planning in power distribution networks.

Keywords: Apriori Algorithm, Association Rule Mining, CRISP-DM, Distribution Network, Preventive Maintenance, Power Distribution Disturbance.

Abstrak – Gangguan berulang pada jaringan distribusi PT PLN (Persero) dapat menurunkan keandalan sistem dan meningkatkan biaya pemeliharaan. Pemanfaatan data historis gangguan masih belum optimal untuk mengidentifikasi pola gangguan yang berulang sebagai dasar penyusunan strategi pemeliharaan preventif. Penelitian ini bertujuan mengidentifikasi pola asosiasi gangguan berulang pada jaringan distribusi menggunakan algoritma Apriori dengan pendekatan *Cross-Industry Standard Process for Data Mining* (CRISP-DM). Dataset yang digunakan terdiri atas 29.446 data historis gangguan yang memuat atribut jenis gangguan, penyebab gangguan, dan lokasi kejadian. Pada tahap data preparation dilakukan proses data cleaning, standarisasi, dan transformasi sehingga diperoleh 10.274 transaksi yang siap dianalisis. Proses *association rule mining* menggunakan parameter *minimum support* sebesar 0,5%, *minimum confidence* sebesar 70%, dan nilai lift lebih dari 1 sebagai kriteria evaluasi. Hasil penelitian menghasilkan 57 *frequent 1-itemset*, 180 *frequent 2-itemset*, 35 *frequent 3-itemset*, serta 64 *association rules* yang memenuhi parameter yang ditetapkan. Pola gangguan yang paling sering muncul adalah “Core FO Problem, FOC – Putus Core, dan Putus Kabel FO”, sedangkan aturan asosiasi dengan nilai lift tertinggi adalah “Banten - FOT – Software → Konfigurasi di Router Problem” dengan nilai support 0,993%, confidence 81,6%, dan lift 30,936. Hasil penelitian menunjukkan bahwa algoritma Apriori mampu mengidentifikasi pola keterkaitan gangguan berulang dari data historis sehingga dapat menghasilkan pengetahuan berbasis data (*maintenance-oriented knowledge*) yang berpotensi mendukung penentuan prioritas inspeksi dan penyusunan program pemeliharaan preventif pada jaringan distribusi PT PLN (Persero).

Kata Kunci: Algoritma Apriori, penambangan aturan asosiasi, CRISP-DM, Gangguan Jaringan Distribusi, Pemeliharaan Preventif.

1. PENDAHULUAN

Jaringan distribusi tenaga listrik merupakan bagian akhir dari sistem tenaga yang berfungsi menyalurkan energi listrik kepada pelanggan. Keandalan jaringan distribusi menjadi indikator penting dalam menjamin kontinuitas pelayanan, karena gangguan yang terjadi pada jaringan distribusi secara langsung memengaruhi frekuensi dan durasi pemadaman yang dialami pelanggan. Dalam praktik operasi PT PLN (Persero), gangguan berulang pada penyulang distribusi masih menjadi tantangan karena dapat meningkatkan nilai indeks keandalan, seperti *System Average Interruption Frequency Index* (SAIFI) dan *System Average Interruption Duration Index* (SAIDI), serta menambah biaya operasi dan pemeliharaan. Perkembangan digitalisasi sistem distribusi juga menyebabkan meningkatnya volume data gangguan yang tersimpan dalam sistem informasi operasi, sehingga diperlukan metode analisis yang mampu mengekstraksi informasi penting dari data historis tersebut untuk mendukung pengambilan keputusan berbasis data [1]–[4].

Gangguan berulang pada jaringan distribusi dapat disebabkan oleh berbagai faktor, seperti vegetasi yang menyentuh jaringan, gangguan akibat binatang, kegagalan isolator, kerusakan konduktor, pengaruh cuaca, maupun kondisi peralatan yang mengalami degradasi. Selama ini, analisis gangguan umumnya dilakukan berdasarkan

frekuensi kejadian atau evaluasi individual terhadap penyebab gangguan. Pendekatan tersebut belum mampu mengidentifikasi hubungan keterkaitan antar karakteristik gangguan yang sering muncul secara bersamaan (*co-occurrence*), sehingga peluang untuk menemukan pola gangguan berulang yang tersembunyi (*hidden patterns*) masih terbatas. Padahal, informasi mengenai hubungan antar penyebab, lokasi, waktu, maupun karakteristik gangguan dapat menjadi dasar dalam menentukan prioritas inspeksi dan strategi pemeliharaan yang lebih efektif [5]-[7].

Salah satu pendekatan yang dapat digunakan untuk mengungkap pola keterkaitan tersebut adalah *Association Rule Mining* (ARM). Teknik ARM bertujuan menemukan hubungan asosiatif antar item dalam suatu kumpulan data melalui pembentukan *frequent itemset* dan *association rules*. Algoritma Apriori merupakan salah satu algoritma ARM yang paling banyak digunakan karena mampu mengidentifikasi kombinasi item yang sering muncul berdasarkan parameter *minimum support* dan *minimum confidence*. Kekuatan hubungan antar item selanjutnya dapat dievaluasi menggunakan nilai *lift*, sehingga aturan asosiasi yang dihasilkan tidak hanya memiliki frekuensi kemunculan tinggi, tetapi juga menunjukkan tingkat keterkaitan yang bermakna. Dalam konteks jaringan distribusi tenaga listrik, pendekatan ini memungkinkan identifikasi pola gangguan berulang yang tidak mudah dikenali melalui analisis statistik konvensional [8]-[10].

Berbagai penelitian telah menerapkan teknik *data mining*, kecerdasan buatan, dan analitik data pada sistem tenaga listrik. Zhang *et al.* menerapkan *Association Rule Mining* untuk menganalisis gangguan pada *Distribution Terminal Unit* (DTU) dengan membentuk transaksi berdasarkan data alarm dan status operasi peralatan, sehingga aturan asosiasi yang dihasilkan berfokus pada identifikasi hubungan antar kejadian gangguan untuk mendukung proses diagnosis penyebab gangguan [8]. Sementara itu, Kuang *et al.* memanfaatkan *association rule* untuk meningkatkan kualitas data pengukuran pada jaringan distribusi pintar melalui deteksi dan pembersihan *outlier*, sehingga keluaran penelitian lebih diarahkan pada peningkatan kualitas data dibandingkan analisis pola gangguan [9]. Di sisi lain, berbagai penelitian tinjauan pustaka menunjukkan bahwa penerapan kecerdasan buatan, *predictive maintenance*, dan *condition monitoring* semakin berkembang dalam pengelolaan aset jaringan distribusi, namun umumnya berfokus pada klasifikasi gangguan, prediksi kegagalan, *fault diagnosis*, atau pemeliharaan berbasis kondisi [2], [6]-[7], [11]-[13].

Meskipun penelitian-penelitian tersebut menunjukkan efektivitas *Association Rule Mining* dan teknik analitik lainnya pada sistem tenaga listrik, masih terdapat beberapa keterbatasan. Pertama, penelitian Zhang *et al.* [8] membangun transaksi berdasarkan alarm dan status operasi *Distribution Terminal Unit*, sehingga aturan asosiasi yang dihasilkan merepresentasikan hubungan antar kejadian peralatan dan belum menggambarkan pola gangguan berulang pada jaringan distribusi. Kedua, atribut yang dianalisis pada penelitian terdahulu lebih berorientasi pada data perangkat (*device-oriented data*), sedangkan penelitian ini membentuk transaksi dari data historis gangguan yang mengintegrasikan atribut operasional, seperti jenis gangguan, penyebab gangguan, lokasi kejadian, dan sehingga mampu merepresentasikan pola gangguan berulang secara lebih komprehensif. Ketiga, keluaran penelitian terdahulu umumnya digunakan untuk diagnosis gangguan atau peningkatan kualitas data, sedangkan penelitian ini menghasilkan *association rules* yang diinterpretasikan sebagai dasar penentuan prioritas inspeksi dan penyusunan program pemeliharaan preventif pada jaringan distribusi PT PLN (Persero).

Dengan demikian, kontribusi penelitian ini tidak hanya terletak pada penerapan algoritma Apriori pada objek yang berbeda, tetapi juga pada perancangan transaksi, pemilihan atribut, serta interpretasi aturan asosiasi yang secara langsung mendukung pengambilan keputusan pemeliharaan berbasis data. *State of the Art* penelitian disajikan pada Tabel 1 untuk memperlihatkan posisi penelitian ini dibandingkan penelitian terdahulu.

Tabel 1. *State of the Art* Penelitian Terdahulu

Peneliti	Metode	Objek	Kontribusi	Research Gap
Zhang <i>et al.</i> (2021)	Association Rule Mining	Distribution Terminal Unit	Analisis penyebab gangguan	Belum menganalisis gangguan berulang pada penyulang distribusi
Kuang <i>et al.</i> (2021)	Association Rule	Data pengukuran smart grid	Pembersihan <i>outlier</i>	Tidak membahas pola gangguan
Mahmoud <i>et al.</i> (2021)	<i>Systematic Review</i>	Smart Grid	<i>Predictive Maintenance</i>	Tidak menggunakan ARM
Zhou <i>et al.</i> (2023)	<i>Review</i>	Asset Management	<i>Condition Monitoring</i>	Tidak menghasilkan aturan asosiasi
Rezapour <i>et al.</i> (2023)	<i>Review</i>	Fault Location	AI untuk lokasi gangguan	Tidak mengidentifikasi <i>frequent itemset</i>
Cao <i>et al.</i> (2025)	<i>Review</i>	Fault Diagnosis	AI untuk diagnosis gangguan	Tidak menerapkan Apriori pada data historis gangguan

Berdasarkan kajian penelitian terdahulu, penerapan *Association Rule Mining* dan algoritma Apriori pada bidang ketenagalistrikan masih memiliki ruang pengembangan. Zhang *et al.* menerapkan algoritma Apriori untuk menganalisis alarm pada *Distribution Terminal Unit* (DTU) sehingga berfokus pada identifikasi hubungan antar

kejadian alarm peralatan distribusi, namun belum memanfaatkan data historis gangguan jaringan sebagai dasar penyusunan strategi pemeliharaan. Sementara itu, Kuang *et al.* menggunakan aturan asosiasi untuk meningkatkan kualitas data pengukuran pada jaringan distribusi pintar melalui deteksi *outlier*, dengan fokus pada peningkatan kualitas data, bukan pada analisis pola gangguan berulang. Di sisi lain, berbagai penelitian mengenai *fault diagnosis*, *predictive maintenance*, dan *asset management* pada sistem distribusi tenaga listrik lebih banyak memanfaatkan pendekatan *machine learning*, *deep learning*, atau kecerdasan buatan untuk klasifikasi, prediksi, dan penentuan lokasi gangguan, tanpa mengeksplorasi hubungan asosiasi antar karakteristik gangguan menggunakan teknik *Association Rule Mining*.

Dengan demikian, kesenjangan penelitian (*research gap*) terletak pada masih terbatasnya penelitian yang mengintegrasikan data historis gangguan jaringan distribusi sebagai dataset transaksi untuk membangun *frequent itemset* dan *association rules* menggunakan algoritma Apriori sebagai dasar penyusunan rekomendasi pemeliharaan berbasis data. Padahal, identifikasi pola keterkaitan antar penyebab gangguan, lokasi, maupun karakteristik gangguan yang berulang berpotensi memberikan informasi yang lebih komprehensif dalam menentukan prioritas inspeksi dan pemeliharaan preventif dibandingkan analisis yang hanya didasarkan pada frekuensi kejadian. Oleh karena itu, penelitian ini menerapkan algoritma Apriori dalam kerangka CRISP-DM untuk mengidentifikasi pola asosiasi gangguan berulang pada jaringan distribusi PT PLN (Persero), kemudian mengevaluasi setiap aturan menggunakan nilai *support*, *confidence*, dan *lift* sehingga menghasilkan informasi yang dapat dimanfaatkan sebagai dasar pengambilan keputusan pemeliharaan berbasis data.

Kebaruan (*novelty*) penelitian ini tidak terletak pada penggunaan algoritma Apriori maupun kerangka kerja CRISP-DM, karena keduanya merupakan metode yang telah mapan dan banyak digunakan dalam berbagai penelitian *data mining*. Kontribusi penelitian ini terletak pada penyusunan strategi pembentukan transaksi berorientasi pemeliharaan (*maintenance-oriented transaction construction strategy*) untuk data historis gangguan jaringan distribusi PT PLN (Persero). Berbeda dengan pendekatan *Association Rule Mining* pada domain transaksi komersial, penelitian ini mentransformasikan data operasional gangguan menjadi dataset transaksi yang merepresentasikan keterkaitan antar atribut teknis, seperti jenis gangguan, komponen yang mengalami gangguan, penyebab gangguan, dan karakteristik gangguan lainnya. Strategi pembentukan transaksi tersebut memungkinkan algoritma Apriori mengidentifikasi pola keterkaitan antar komponen gangguan yang sebelumnya sulit diperoleh melalui analisis frekuensi kejadian secara konvensional.

Selain itu, penelitian ini tidak berhenti pada pembentukan *frequent itemset* dan *association rules*, tetapi menginterpretasikan aturan asosiasi menjadi pengetahuan pemeliharaan (*maintenance-oriented knowledge*) yang mendukung pengambilan keputusan berbasis data. Setiap aturan yang memenuhi nilai *minimum support*, *minimum confidence*, dan *lift* dianalisis untuk menghasilkan rekomendasi teknis dalam penentuan prioritas inspeksi, penyusunan program pemeliharaan preventif, serta pengelolaan aset jaringan distribusi. Dengan demikian, kontribusi penelitian ini tidak hanya berupa penerapan algoritma Apriori pada data gangguan PT PLN (Persero), tetapi juga menghadirkan pendekatan sistematis untuk mentransformasikan data operasional menjadi pengetahuan yang dapat dimanfaatkan sebagai dasar data-driven maintenance decision support dalam pengelolaan jaringan distribusi tenaga listrik.

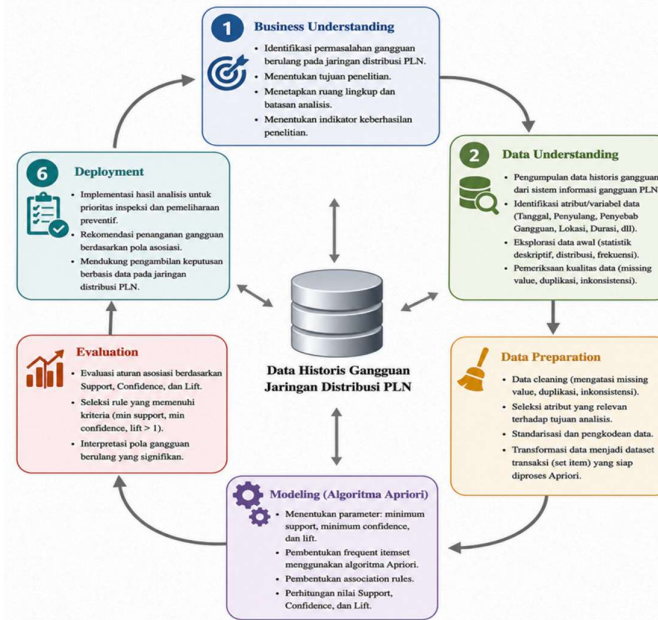
2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan *data mining* dengan framework *Cross Industry Standard Process for Data Mining* (CRISP-DM) [14] untuk menganalisis pola asosiasi gangguan berulang pada jaringan distribusi PT PLN (Persero). Framework CRISP-DM dipilih karena menyediakan tahapan yang sistematis mulai dari identifikasi permasalahan, pemahaman data, persiapan data, pembangunan model, evaluasi hingga implementasi hasil analisis. Pada penelitian ini, algoritma Apriori digunakan sebagai metode *Association Rule Mining* untuk memperoleh *frequent itemset* dan membentuk *association rules* berdasarkan nilai *minimum support*, *minimum confidence*, dan *lift*. Tahapan penelitian ditunjukkan pada Gambar 1.

2.1. Pemahaman Bisnis (*Business Understanding*)

Tahap *Business Understanding* bertujuan mengidentifikasi permasalahan utama yang dihadapi pada sistem distribusi tenaga listrik, yaitu tingginya frekuensi gangguan berulang yang berdampak terhadap penurunan keandalan jaringan. Selama ini evaluasi gangguan umumnya dilakukan berdasarkan jumlah kejadian atau jenis gangguan secara terpisah sehingga hubungan antar karakteristik gangguan belum dapat diketahui secara menyeluruh. Berdasarkan kondisi tersebut, penelitian ini bertujuan mengidentifikasi pola keterkaitan antar gangguan pada jaringan distribusi menggunakan algoritma Apriori sehingga diperoleh *association rules* yang menggambarkan hubungan antar karakteristik gangguan berdasarkan atribut yang terdapat pada data historis gangguan. Melalui proses pembentukan transaksi dari kombinasi atribut teknis yang relevan, penelitian ini berupaya mengungkap pola gangguan berulang yang tidak mudah dikenali melalui analisis frekuensi kejadian

secara konvensional. Informasi yang dihasilkan selanjutnya diinterpretasikan menjadi *maintenance-oriented knowledge* untuk mendukung penentuan prioritas inspeksi, penyusunan program pemeliharaan preventif, serta pengambilan keputusan berbasis data dalam pengelolaan jaringan distribusi PT PLN (Persero).



Gambar 1. Kerangka Penelitian Menggunakan CRISP-DM

No.	Id tiket	Id pelanggan	Id insiden	Nama pelanggan	---	Isi laporan	Penyebab	Penyebab detail	Nama QC penyebab	Nama QC tindakan	Terminasi GIATLONG	Terminasi Provinsi	---	Tanggal send NOC	Priority
1	EN3A6568	2,401E+10		NON PELANGGAN	----	GIS BUDIKEMULI AAN - TUTT JBBJ ULTG KARETA SID 1112020 17185 KELUHAN : DOWN	FOT - WIRING	UTP / STP PROBLEM	-	-	-	-	----	1/22/2024 14.29	NORMAL
2	EN3U560F	2,401E+10		NON PELANGGAN	----	0100013705 9 METRO QNQ PT PLN REPEATER BERAMBANG KAB JEMBRANA BALI	NETWORK	OTHER	-	-	-	-	----	1/3/2024 20.22	NORMAL
3	EN34Z3DN	2,401E+10		NON PELANGGAN	----	PERAPIHAN	NOT INCIDENT	PERAPIHAN LASTMILE	-	-	-	-	----	1/31/2024 11.59	NORMAL
4	EN33TMR	2,401E+10		NON PELANGGAN	----	PENGECEKAN PERANGKAT HARIF DAN RECTIP CI GI JATAKE	NOT INCIDENT	PENGECEKAN OK	-	-	-	-	----	1/15/2024 20.52	NORMAL
5	EN32828L	8,31E+15	2,401E+10	PT PLN (PERSERO) UNIT INDUK PUSAT PENGATUR BEBAN JAWA, MADURA DAN BALI	----	GANGGUAN JDA PENDO - PLTGU CILEGON KETERANGAN : GANGGUAN CORE DIE PELAPOR : BAPAK ADI PERANGKAT : SEDANG DI TANYAKAN	PS - POWER dan ENV	SOURCE POWER PROBLEM	PUTUS KABEL FO	PENYAMBUNGAN KABEL	5 9304833 - 15090.106 -10334655 7300	Banten	----	1/10/2024 14.12	NORMAL
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Gambar 2. Contoh Dataset Penelitian

2.2. Pemahaman Data (*Data Understanding*)

Sebelum digunakan dalam proses penambangan data, telah dilakukan tahap pemahaman data (*data understanding*) untuk memahami karakteristik, struktur, dan kualitas data historis gangguan jaringan distribusi. Dataset penelitian diperoleh dari data *Trouble Ticket System* (TIS) - *Network Operation Center* (NOC), Divisi Operasi dan pemeliharaan dari PT Indonesia Comment Plus PLN dengan populasi penelitian sebanyak 29.446 data dari seluruh kejadian gangguan pada jaringan distribusi yang tercatat selama periode Januari–Agustus 2024.

Selanjutnya dilakukan pemeriksaan konsistensi data untuk menghilangkan data duplikat, memperbaiki inkonsistensi penulisan atribut, serta menangani data yang tidak lengkap sebelum memasuki tahap *data preparation*.

Dataset terdiri atas sejumlah atribut yang relevan untuk analisis asosiasi, meliputi waktu kejadian, lokasi gangguan, jenis gangguan, penyebab gangguan, dan komponen yang mengalami gangguan seperti terlihat pada Gambar 2. Pada tahap ini juga dilakukan analisis distribusi *item* terhadap atribut-atribut tersebut untuk mengetahui frekuensi kemunculan setiap kategori. Analisis distribusi *item* diperlukan karena algoritma Apriori membentuk *frequent itemset* berdasarkan nilai *support* setiap *item*. *Item* dengan frekuensi kemunculan yang tinggi memiliki peluang lebih besar untuk membentuk aturan asosiasi, sedangkan *item* yang sangat jarang muncul umumnya tidak memenuhi ambang *minimum support*. Oleh karena itu, distribusi frekuensi jenis gangguan, penyebab gangguan, lokasi gangguan menjadi dasar dalam memahami karakteristik dataset sebelum proses pemodelan dilakukan.

2.3. Persiapan Data (*Data Preparation*)

Tahap persiapan data bertujuan menghasilkan dataset yang siap diproses menggunakan algoritma Apriori melalui serangkaian proses seleksi atribut, pembersihan data, standarisasi, dan transformasi data menjadi bentuk transaksi. Pada tahap ini dipilih atribut-atribut yang relevan dengan tujuan penelitian, yaitu mengidentifikasi pola asosiasi gangguan berulang pada jaringan distribusi PT PLN (Persero). Atribut yang digunakan terdiri atas jenis gangguan (atribut penyebab) yang menyatakan kategori gangguan menurut pelapor, penyebab gangguan (atribut nama QC penyebab) yang menyatakan penyebab menurut *quality control* (QC), lokasi gangguan (atribut nama QC provinsi) seperti di Gambar 2, sehingga didapat dataset gangguan yang disajikan pada Tabel 2.

Tabel 2. Data Awal Historis Gangguan Jaringan Distribusi

No.	Jenis Gangguan	Penyebab	Lokasi
1	UTP / STP Problem	-	-
2	Other	-	-
3	Perapihan Lastmile	-	-
4	Pengecekan OK	-	-
5	Source Power Problem	Putus Kabel FO	Banten
6	Putus Protect	Putus Kabel FO	Jawa Barat
7	Putus Protect	Putus Kabel FO	Jawa Barat
8	Putus Protect	Putus Kabel FO	Jawa Barat
9	Odf_Pigtail Rusak	Core FO Problem	DKI Jakarta
....			
22946			

Kemudian dataset gangguan dilakukan distribusi frekuensi masing-masing atribut untuk memastikan data layak digunakan pada proses data mining. Atribut Jenis Gangguan merupakan klasifikasi teknis gangguan yang terdiri atas 27 kategori, antara lain FOC – Putus Core, PS – Power and *Environment*, FOT – Perangkat, *Not Incident*, serta kategori gangguan lainnya yang tercatat pada sistem operasional PT. PLN (Persero). Atribut penyebab gangguan memuat informasi penyebab utama gangguan, seperti Putus Kabel FO, *Core FO Problem*, *Battery/Rectifier Problem*, serta nilai "-" untuk data yang belum teridentifikasi secara spesifik. Sementara itu, atribut Lokasi Kejadian berisi informasi lokasi gangguan provinsi tempat gangguan terjadi. Gambaran karakteristik data yakni distribusi frekuensi masing-masing kategori Jenis Gangguan sebanyak 27 dan penyebab gangguan sebanyak 43 seperti terlihat pada Tabel 3 hingga 5. Terlihat bahwa Jenis Gangguan FOC - Putus Core (*Fiber Optic Cable Putus Core*) paling banyak terjadi. Sedangkan Penyebab Gangguan paling banyak disebabkan pada *Core FO Problem*.

Tabel 3. Data Distribusi Jenis Gangguan

Jenis Gangguan	Jumlah Kejadian
FOC - Putus Core	2733
FOC - Putus Kabel	2547
FOT – Perangkat	2192
FOC - Core Bending	551
PS - Power dan ENV	528
FOT – Software	521
FOT – Warring	233

Tabel 4. Distribusi Penyebab Gangguan

Penyebab	Jumlah Kejadian
Core FO Problem	299
Putus Kabel FO	2573
Router Problem	1354
Patch Cord Problem	539
Switch Problem	308
Konfigurasi di Router Problem	268
SFP Problem	197

Tabel 5. Distribusi Penyebab Gangguan

Penyebab	Jumlah Kejadian
Telicon Problem	24
AC Supply Power Lost di Lastmile	23
Adaptor Converter Problem	23
Putus Kabel FO Tipe ADSS	22
Vsat Problem	21
Cable Power Problem	20
Pigtail	15

Jenis Gangguan	Jumlah Kejadian	Penyebab	Jumlah Kejadian	Penyebab	Jumlah Kejadian
MSR	162	UTP / STP Problem	193	Adaptor Router	14
Network	110	AC Supply Power Lost di POP	177	Perangkat LNB VSAT	10
FOT – ONT	77	Konfigurasi di Switch Problem	175	Router Cisco Problem	9
FOC - OLT to Uplink	75	Modem Problem	160	Port Mikrotik Problem	8
FOC – FAT	59	Bandwidth Problem	129	Radio	8
3rd Party	42	MCB Problem	112	Attenuator Problem	7
FOC - OLT to FDT	40	Modul Problem	95	Mikrotik Problem	5
FOC - Konvensional	29	Converter Problem	79	Port Converter Problem	3
FOC - FDT to FAT	26	Battery / Rectifier Problem	76	Adaptor Switch Problem	2
FOT - Konvensional	21	Port Switch Problem	70	FO Bending	2
FOT - GPON OLT	17	Core Bending	67	Repeater	1
FOT - Site Blocking	17	Connector Bulkhead Problem	58		
FOC – FDT	15	Upperlevel Problem	49		
FOT – Server	15	CCTV Problem	39		
Ps - Power & Env	8	Konfigurasi Converter Problem	33		
FOC - DW Outdoor	5	Source Power Problem	28		
FOC - DW Indoor	2	Card Modul Problem	27		
FOT – STB	2	Konfigurasi Cross Connect Problem	27		
3rd Party - Indosat	1				
FOT - IPTV	1				

Sedangkan data distribusi atribut lokasi gangguan di seluruh Indonesia yang paling banyak mengalami gangguan di provinsi DKI Jakarta seperti terlihat pada Tabel 6.

Tabel 6. Data Distribusi Atribut Lokasi Gangguan

Provinsi	Jumlah Gangguan	Provinsi	Jumlah Gangguan
DKI Jakarta	2453	Riau	103
Banten	932	Sulawesi Utara	94
Jawa Barat	418	Jambi	78
Aceh	305	Lampung	73
Sumatera Selatan	240	Bengkulu	55
Sulawesi Selatan	179	Kepulauan Bangka Belitung	53
Jawa Timur	172	Sulawesi Barat	50
Sumatera Utara	158	Maluku	39
Kalimantan Barat	155	Nusa Tenggara Barat	32
Kalimantan Tengah	143	Gorontalo	28
Kalimantan Selatan	139	Papua	27
Sulawesi Tengah	139	Kalimantan Utara	22
Bali	135	Nusa Tenggara Timur	20
Jawa Tengah	129	Kepulauan Riau	5
Sulawesi Tenggara	114	Papua Pegunungan	4
Kalimantan Timur	108	Papua Tengah	2
Sumatera Barat	105	Maluku Utara	2

3.1 Hasil Data Preparation

Tahap *Data Preparation* dilakukan melalui proses pembersihan data (*data cleaning*), transformasi atribut, serta pembentukan format transaksi yang sesuai dengan kebutuhan algoritma Apriori. Proses *data cleaning* dilakukan untuk menghilangkan data yang tidak relevan terhadap tujuan penelitian, seperti data duplikat, data dengan atribut yang tidak lengkap (*missing values*) yakni nilai "-" dan "NOT INCIDENT" pada atribut Penyebab berjumlah sebanyak 13.959 data dan nilai "-" dan "Lainnya" pada atribut nama qc penyebab sejumlah 5.213 data. Setelah proses tersebut, jumlah data berkurang menjadi 10.274 data yang dinilai valid dan siap diproses pada tahap transformasi.

Setelah proses pembersihan selesai, kemudian dilakukan proses transformasi data menjadi dataset transaksi (*transaction dataset*) sebagai masukan bagi algoritma Apriori. Pada penelitian ini, satu transaksi didefinisikan sebagai satu kejadian gangguan yang direpresentasikan dalam bentuk kombinasi atribut teknis yang saling

berkaitan. Berbeda dengan transaksi pada analisis market basket yang merepresentasikan kumpulan produk dalam satu pembelian, transaksi pada penelitian ini dibentuk dari kombinasi informasi gangguan yang berasal dari satu kejadian operasional pada jaringan distribusi.

Model pembentukan transaksi (*fault component-based transaction model*) dikembangkan dengan menggabungkan atribut-atribut yang dianggap merepresentasikan karakteristik gangguan, yaitu jenis gangguan, penyebab gangguan, serta lokasi tempat gangguan terjadi. Setiap atribut diperlakukan sebagai sebuah item, sehingga satu transaksi merupakan himpunan item yang menggambarkan karakteristik lengkap dari satu kejadian gangguan sebanyak 10.274 data transaksi yang salah satu bagian disajikan pada Tabel 7.

Tabel 7. Data Transaksi

TID	Item 1 (Jenis Gangguan)	Item 2 (Penyebab)	Item 3 (Lokasi)
1	PS - Power dan Env	Putus Kabel Fo	Banten
2	FOC - Putus Core	Putus Kabel Fo	Jawa Barat
3	FOC - Putus Core	Putus Kabel Fo	Jawa Barat
4	FOC - Putus Core	Putus Kabel Fo	Jawa Barat
5	FOC - Core Bending	Core Fo Problem	DKI Jakarta
6	FOC - Putus Kabel	Putus Kabel Fo	Tidak Diketahui

Representasi transaksi tersebut memungkinkan algoritma Apriori melakukan pencarian *frequent itemset* dan membentuk *association rules* berdasarkan keterkaitan antar atribut gangguan, bukan hanya berdasarkan frekuensi kemunculan masing-masing atribut secara terpisah. Dengan pendekatan ini, pola hubungan antar komponen gangguan yang berulang dapat diidentifikasi dan selanjutnya diinterpretasikan sebagai pengetahuan pemeliharaan (*maintenance-oriented knowledge*) untuk mendukung penentuan prioritas inspeksi dan penyusunan strategi pemeliharaan *preventif* pada jaringan distribusi PT PLN (Persero).

2.4. Pemodelan (Modeling)

Tahap Modeling merupakan inti penelitian yang bertujuan mengidentifikasi pola asosiasi gangguan berulang menggunakan algoritma Apriori. Dataset transaksi hasil tahap persiapan data diproses untuk membentuk kandidat itemset, menentukan *frequent itemset* berdasarkan nilai *minimum support*, kemudian menghasilkan *association rules* yang memenuhi nilai *minimum confidence*. Selanjutnya setiap aturan dievaluasi menggunakan nilai lift untuk mengetahui tingkat kekuatan hubungan antar item. Implementasi model dilakukan menggunakan Platform Pengembangan: Google Colab dengan Library Pemrograman: Pandas dan Mlxtend, yang digunakan untuk pengolahan data, pembentukan model, serta analisis *association rule mining*. Proses pembentukan model ditunjukkan pada Gambar 3. Persamaan algoritma Apriori [15] yang digunakan adalah sebagai berikut, dimana *Support* dihitung menggunakan persamaan (1).

$$\text{Support}(A) = \frac{\text{Jumlah transaksi yang memuat } A}{\text{Total transaksi}} \quad (1)$$

Keterangan : A = item atau gangguan; Support (A) = tingkat kemunculan item A dalam seluruh transaksi. Sedangkan *Confidence* dihitung menggunakan persamaan (2).

$$\text{Confidence}(A \rightarrow B) = \frac{\text{support}(A \cup B)}{\text{support } A} \quad (2)$$

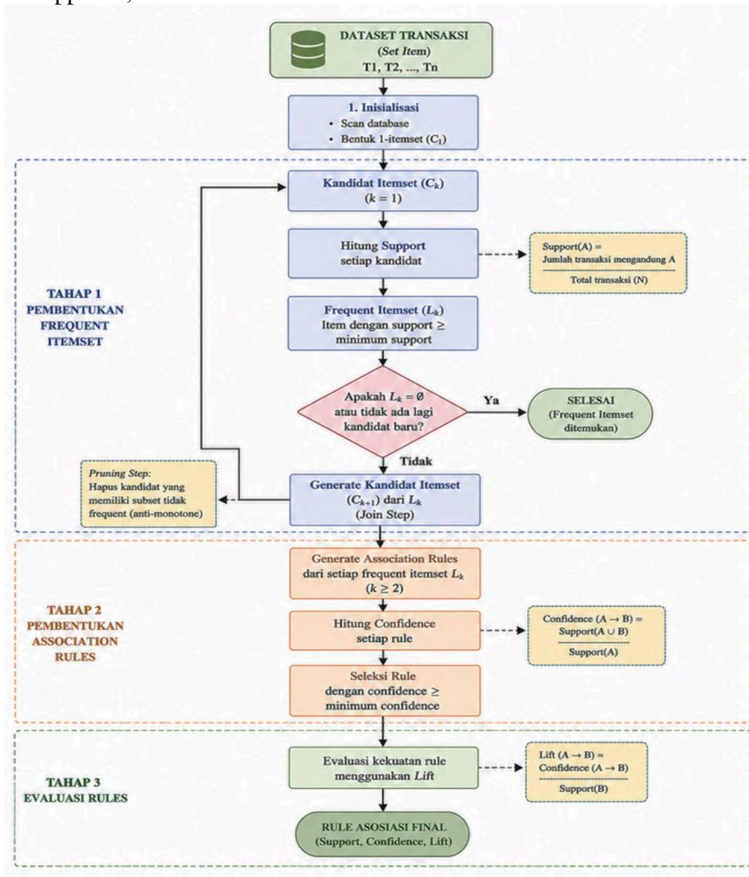
Keterangan: A = *antecedent*; B = *consequent*; Confidence = probabilitas kemunculan B ketika A terjadi. Parameter yang digunakan pada penelitian ditunjukkan pada Tabel 8.

Tabel 8. Parameter

Parameter	Nilai
Minimum Support	0,5 %
Minimum Confidence	70 %
Lift	> 1

Penentuan parameter dilakukan melalui pengujian batas secara sistenatis. Rentang kerja *minimum support* pada dataset ini adalah 0,0097% hingga 10%; batas bawah ditentukan oleh *support item* terendah dalam data, sedangkan batas atas dibuktikan secara empiris karena pada ambang 30% algoritma tidak lagi menghasilkan *frequent itemset*. Dengan demikian nilai 1% bukan merupakan batas terendah, melainkan satu titik di dalam rentang tersebut. Rentang *minimum confidence* pada dataset ini adalah 3,41% hingga 95,56%. Batas atas dibuktikan melalui pengujian pada ambang 98% yang menghasilkan nol aturan meskipun *frequent itemset* tetap terbentuk. Pada parameter terpilih sebagian besar aturan memiliki nilai *confidence* diatas 76% sehingga 65% - 74% tidak

mempresentasikan batas atas maupun sebaran *confidence* yang sebenarnya. Berdasarkan pengujian tersebut, ditetapkan minimum support 0,5% dan minimum *confidence*



Gambar 3. Alur Algoritma Apriori

2.5. Evaluasi (Evaluation)

Tahap evaluasi dilakukan untuk menilai kualitas aturan asosiasi yang dihasilkan algoritma Apriori. Aturan dianggap memenuhi kriteria apabila memiliki nilai *support* dan *confidence* di atas batas minimum yang telah ditentukan serta nilai *lift* lebih besar dari satu, yang menunjukkan adanya hubungan positif antar *item*.

$$Lift(A \rightarrow B) = \frac{confidence(A \rightarrow B)}{support B} \quad (3)$$

Perhitungan Lift menggunakan persamaan (3) dimana Nilai Lift lebih besar dari satu menunjukkan adanya hubungan positif antara *antecedent* dan *consequent*. Aturan asosiasi yang memenuhi kriteria kemudian diurutkan berdasarkan nilai lift tertinggi untuk memperoleh pola gangguan yang paling signifikan dan memiliki potensi digunakan sebagai dasar pengambilan keputusan pemeliharaan.

2.6. Deployment

Pada framework CRISP-DM, tahap *deployment* menjelaskan bagaimana hasil analisis berpotensi dimanfaatkan dalam konteks operasional. Pada penelitian ini, tahap *deployment* belum berupa implementasi langsung ke dalam sistem operasional PT PLN (Persero), melainkan difokuskan pada penyusunan rencana pemanfaatan hasil (*deployment plan*) berdasarkan association rules yang diperoleh. Aturan asosiasi yang memenuhi kriteria *minimum support*, *minimum confidence*, dan lift diinterpretasikan sebagai informasi yang dapat digunakan untuk mendukung proses analisis gangguan dan perencanaan pemeliharaan. Hasil penelitian ini berpotensi dimanfaatkan sebagai masukan dalam menentukan prioritas inspeksi jaringan, penyusunan program pemeliharaan preventif, identifikasi penyulang atau komponen yang memerlukan perhatian lebih, serta evaluasi pola gangguan berulang pada jaringan distribusi.

Dengan demikian, keluaran penelitian ini tidak diklaim sebagai sistem pendukung keputusan yang telah diterapkan, melainkan sebagai implikasi praktis berupa pengetahuan berbasis data (*maintenance-oriented*

knowledge) yang dapat menjadi referensi bagi pengembangan analisis gangguan dan pengambilan keputusan pemeliharaan pada penelitian maupun implementasi selanjutnya. Validasi operasional bersama praktisi PT PLN (Persero) serta integrasi ke dalam sistem pendukung keputusan menjadi bagian dari pengembangan yang dapat dilakukan pada penelitian mendatang.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Frequent Itemset

Proses pembentukan *frequent itemset* dilakukan menggunakan algoritma Apriori dengan nilai *minimum support* sebesar 0,5% dan *minimum confidence* 70%. Pemilihan nilai parameter *minimum support* tersebut berdasarkan pada hasil pengujian batas secara sistematis yang telah dijelaskan pada Tabel 8. Pada Tabel 9 menampilkan hasil sesuai parameter, dimana *frequent 1 itemset* berjumlah 57 data, *frequent 2 itemset* berjumlah 180 data dan *frequent 3 itemset* berjumlah 35 data. Berikut 10 data awal dari setiap *frequent itemset* seperti terlihat pada Tabel 9.

Tabel 9. Frequent Itemset

1 Itemset	Support (%)	2 Itemset	Support (%)	3 Itemset	Support (%)
Core Fo Problem	29,667	Core FO Problem + FOC - Putus Core	20,216	DKI Jakarta + Fot - Perangkat + Router Problem	8,351
FOC - Putus Core	27,205	FOC - Putus Kabel + Putus Kabel Fo	19,836	Core Fo Problem + DKI Jakarta + FOC - Putus Core	4,399
Putus Kabel Fo	25,725	Fot - Perangkat + Router Problem	11,515	FOC - Putus Kabel + Putus Kabel Fo + Tidak Diketahui	2,793
DKI Jakarta	25,492	DKI Jakarta + FOT - Perangkat	10,103	Core Fo Problem + FOC - Putus Core + Tidak Diketahui	2,647
FOC - Putus Kabel	25,375	DKI Jakarta + Router Problem	9,772	Banten + FOC - Putus Kabel + Putus Kabel Fo	1,956
FOT - Perangkat	21,851	Core Fo Problem + DKI Jakarta	8,03	DKI Jakarta + FOC - Putus Kabel + Putus Kabel Fo	1,723
Router Problem	13,237	DKI Jakarta + FOC - Putus Core	4,653	Banten + Core Fo Problem + FOC - Putus Core	1,616
Tidak Diketahui	10,901	FOC - Putus Core + Putus Kabel Fo	4,331	FOC - Putus Kabel + Jawa Barat + Putus Kabel Fo	1,47
Banten	9,694	Core Fo Problem + FOC - Putus Kabel	4,117	Core Fo Problem + DKI Jakarta + FOC - Putus Kabel	1,392
Jawa Barat	7,076	Core Fo Problem + Tidak Diketahui	3,485	Core Fo Problem + FOC - Putus Core + Jawa Barat	1,382

Secara keseluruhan, analisis *support* memperlihatkan bahwa *frequent 1-itemset* ditemui pada penyebab gangguan Core FO Problem dengan nilai *support* terbesar, pada *frequent 2-itemset* ditemui penyebab gangguan Core FO Problem bersamaan dengan jenis gangguan FOC - Putus Core dengan nilai *support* terbesar, sedangkan *frequent 3-itemset* ditemui di lokasi gangguan DKI Jakarta, jenis gangguan FOT – Perangkat dan penyebab gangguan Router Problem dengan nilai *support* terbesar.

3.2 Hasil Association Rule

Aturan asosiasi dibentuk berdasarkan nilai *minimum support* 0,5% dan nilai *minimum confidence* sebesar 70%. Beberapa aturan yang diperoleh memiliki nilai *confidence* dan *lift ratio* yang tinggi sehingga menunjukkan adanya hubungan yang kuat antar penyebab gangguan. Berdasarkan *frequent itemsets* yang memenuhi *minimum support*, dilakukan pembentukan aturan asosiasi menggunakan *minimum confidence* sebesar 70%. Hasil proses Apriori menghasilkan 64 *association rules* yang memenuhi seluruh parameter evaluasi seperti terlihat pada Tabel 10.

Tabel 10. Hasil Association Rule

No	K	Aturan Asosiasi (Association Rules)	Support %	Confidence %	Lift
1	3	Banten + FOT - Software → Konfigurasi Di Router Problem	0,993	81,6	30,936
2	3	Bandwidth Problem + Kalimantan Tengah → FOT - Software	0,584	100	19,644
3	2	Bandwidth Problem → FOT - Software	1,119	89,15	17,512
4	2	Battery / Rectifier Problem → PS - Power dan ENV	0,73	94,94	17,386
5	2	MCB Problem → PS - Power dan ENV	1,197	91,79	16,81
6	2	AC Supply Power Lost di POP → PS - Power dan ENV	1,606	91,16	16,695
7	3	Banten + Konfigurasi di Router Problem → FOT - Software	0,993	79,69	15,654
8	3	DKI Jakarta + MSR → Router Problem	0,925	90,48	6,835

No	K	Aturan Asosiasi (<i>Association Rules</i>)	Support %	Confidence %	Lift
9	3	DKI Jakarta + FOT - Perangkat → Router Problem	8,351	82,66	6,244
10	3	Banten + Router Problem → FOT - Perangkat	1,256	95,56	4,373
11	2	Modem Problem → FOT - Perangkat	1,499	91,67	4,195
12	3	Jawa Barat + Router Problem → FOT - Perangkat	1,139	89,31	4,087
13	2	Router Problem → FOT - Perangkat	11,515	86,99	3,981
14	3	Dki Jakarta + Router Problem → FOT - Perangkat	8,351	85,46	3,911
15	3	Bali + FOC - Putus Kabel → Putus Kabel FO	0,594	96,83	3,764
16	3	FOC - Putus Kabel + Sulawesi Tenggara. → Putus Kabel FO	0,876	95,74	3,722
17	3	MSR + Router Problem → DKI Jakarta	0,925	94,06	3,69
18	3	FOC - Putus Kabel + Sumatera Barat → Putus Kabel FO	1,353	93,92	3,651
19	3	FOC - Putus Kabel + Jawa Timur → Putus Kabel FO	0,857	92,63	3,601
20	3	Bali + Putus Kabel FO → FOC - Putus Kabel	0,594	91,04	3,588
21	3	Putus Kabel FO + Sulawesi Utara → Foc - Putus Kabel	1,217	89,93	3,544
22	3	FOC - Putus Kabel + Sumatera Utara → Putus Kabel FO	0,594	91,04	3,539
23	3	Jawa Timur + Putus Kabel FO → FOC - Putus Kabel	0,857	88,89	3,503
24	2	Converter Problem → FOT - Perangkat	0,603	76,54	3,503
25	3	FOC - Putus Kabel + Sulawesi Utara → Putus Kabel FO	1,217	89,29	3,471
26	3	Putus Kabel FO + Sulawesi Tenggara → FOC - Putus Kabel	0,876	87,38	3,444
27	3	Patch Cord Problem + Sumatera Selatan → FOC - Putus Core	0,944	93,27	3,428
28	3	FOC - Putus Kabel + Sulawesi Selatan → Putus Kabel FO	1,08	88,1	3,424
29	2	FOT - Ont → Dki Jakarta	0,652	87,01	3,413
30	3	Core FO Problem + Sumatera Selatan → FOC - Putus Core	0,886	92,86	3,413
31	3	Banten + Putus Kabel FO → FOC - Putus Kabel	1,956	84,45	3,328
32	3	Jawa Barat + Putus Kabel FO → FOC - Putus Kabel	1,47	84,36	3,324
33	3	Bali + Core FO Problem → FOC - Putus Core	0,506	89,66	3,296
34	3	Aceh + Core FO Problem → FOC - Putus Core	0,827	89,47	3,289
35	3	Putus Kabel FO + Sumatera Barat → FOC - Putus Kabel	1,353	82,74	3,261
36	3	Putus Kabel FO + Tidak Diketahui → FOC - Putus Kabel	2,793	82,71	3,259
37	3	FOC - Putus Kabel + Tidak Diketahui → Putus Kabel FO	2,793	83,19	3,234
38	2	Sfp Problem → FOT - Perangkat	1,421	70,19	3,212
39	3	Aceh + Putus Kabel FO → FOC - Putus Kabel	0,545	81,16	3,198
40	3	Dki Jakarta + FOC - Putus Core → Core FO Problem	4,399	94,56	3,187
41	3	Core FO Problem + Kalimantan Timur → FOC - Putus Core	0,565	86,57	3,182
42	3	Putus Kabel FO + Sumatera Utara → FOC - Putus Kabel	0,594	80,26	3,163
43	3	Putus Kabel FO + Sulawesi Tengah → FOC - Putus Kabel	1,11	79,17	3,12
44	3	FOC - Putus Kabel + Jawa Barat → Putus Kabel FO	1,47	79,47	3,089
45	2	Putus Kabel FO → FOC - Putus Kabel	19,836	77,11	3,039
46	2	FOC - Putus Kabel → Putus Kabel FO	19,836	78,17	3,039
47	3	FOC - Putus Kabel + Sulawesi Tengah → Putus Kabel FO	1,11	77,55	3,015
48	3	Dki Jakarta + Putus Kabel FO → FOC - Putus Kabel	1,723	76,29	3,007
49	3	FOC - Putus Core + Jawa Timur → Core FO Problem	0,857	88,89	2,996
50	3	Putus Kabel FO + Sulawesi Selatan → FOC - Putus Kabel	1,08	75,51	2,976
51	3	FOC - Putus Core + Jawa Barat → Core FO Problem	1,382	87,12	2,936
52	2	Router Problem → Dki Jakarta	9,772	73,82	2,896
53	3	Banten + FOC - Putus Core → Core FO Problem	1,616	85,57	2,884
54	3	Banten + FOC - Putus Kabel → Putus Kabel FO	1,956	74,17	2,883
55	3	Aceh + FOC - Putus Core → Core FO Problem	0,827	85	2,865
56	3	Core FO Problem + Sumatera Barat → FOC - Putus Core	1,012	77,61	2,853
57	3	FOT - Perangkat + Router Problem → Dki Jakarta	8,351	72,53	2,845
58	3	Core FO Problem + Tidak Diketahui → FOC - Putus Core	2,647	75,98	2,793
59	3	Core FO Problem + Jawa Timur → FOC - Putus Core	0,857	74,58	2,741
60	3	FOC - Putus Core + Tidak Diketahui → Core FO Problem	2,647	78,61	2,65
61	3	FOC - Putus Core + Sumatera Barat → Core FO Problem	1,012	77,61	2,616
62	2	Sumatera Selatan → FOC - Putus Core	2,083	71,1	2,613
63	2	FOC - Putus Core → Core FO Problem	20,216	74,31	2,505
64	2	FOC - OLT to UPLINK → Core FO Problem	0,526	71,05	2,395

Nilai *lift* lebih besar dari satu menunjukkan bahwa hubungan antar item tidak terjadi secara acak, tetapi memiliki tingkat asosiasi yang kuat. Rule yang dihasilkan berjumlah 64 aturan asosiasi yang memenuhi ambang minimum *support* dan minimum *confidence* yang telah ditetapkan. Rentang nilai *support* berada antara 0,506% hingga 20,216%, dari keseluruhan transaksi. Nilai *confidence* berkisar antara 70,19% hingga 100%, yang secara probabilistik menunjukkan bahwa ketika suatu *antecedent* terjadi, maka peluang munculnya *consequent* berada pada kisaran 70,19% hingga 100%. Seluruh aturan memiliki nilai *lift* antara 2,395 hingga 30,936. Nilai *lift* yang

lebih besar dari 1 menandakan bahwa hubungan antarvariabel bersifat positif. Aturan dengan nilai lift tertinggi adalah: Banten + FOT - Software → Konfigurasi di Router Problem (Support 0,993; Confidence 81,6; Lift 30,836).

3.3 Pembahasan

Hasil penelitian menunjukkan bahwa algoritma Apriori mampu mengidentifikasi pola asosiasi gangguan berulang pada jaringan distribusi PT PLN (Persero) berdasarkan data historis gangguan. Dari proses pembentukan *frequent itemset* menggunakan *minimum support* sebesar 0,5% diperoleh 57 *frequent 1-itemset*, 180 *frequent 2-itemset*, dan 35 *frequent 3-itemset* yang selanjutnya menghasilkan 64 *association rules* yang memenuhi nilai *minimum confidence* sebesar 70% serta memiliki nilai lift lebih besar dari satu. Kondisi tersebut menunjukkan bahwa hubungan antar item yang terbentuk bukan merupakan hubungan acak, tetapi memiliki tingkat keterkaitan yang bermakna sehingga dapat digunakan sebagai dasar analisis gangguan berulang.

Analisis *frequent itemset* menunjukkan bahwa gangguan pada infrastruktur fiber optic merupakan pola yang paling sering muncul. Hal ini terlihat dari tingginya nilai *support* pada item Core Fo Problem (29,667%), FOC – Putus Core (27,205%), Putus Kabel FO (25,725%), dan FOC – Putus Kabel (25,375%). Tingginya nilai *support* tersebut menunjukkan bahwa gangguan pada komponen fiber optic merupakan karakteristik yang dominan dalam data historis gangguan sehingga komponen tersebut layak menjadi prioritas dalam kegiatan inspeksi maupun pemeliharaan preventif. Interpretasi ini didasarkan pada *persentase support* setiap item yang ditampilkan pada hasil *frequent itemset*, sehingga penggunaan istilah “dominan” didukung oleh data kuantitatif penelitian.

Association rule dengan nilai lift tertinggi adalah “Banten + FOT – Software → Konfigurasi di Router Problem” dengan support 0,993%, confidence 81,6%, dan lift 30,936. Nilai lift yang jauh lebih besar dari satu menunjukkan bahwa kemunculan kedua kondisi tersebut memiliki hubungan yang sangat kuat dibandingkan apabila keduanya muncul secara independen. Dengan demikian, apabila gangguan perangkat lunak FOT ditemukan pada wilayah Banten, pemeriksaan konfigurasi router dapat diprioritaskan sebagai bagian dari proses penelusuran penyebab gangguan.

Selain itu, aturan “Router Problem → FOT – Perangkat” memiliki *support* sebesar 11,515%, *confidence* 86,99%, dan lift 3,981. Hasil tersebut menunjukkan bahwa sebagian besar gangguan router diikuti oleh gangguan pada perangkat FOT sehingga kedua komponen memiliki hubungan yang erat. Informasi ini dapat dimanfaatkan sebagai dasar penyusunan prosedur inspeksi yang tidak hanya berfokus pada satu komponen, tetapi juga mencakup pemeriksaan perangkat lain yang memiliki keterkaitan berdasarkan aturan asosiasi yang diperoleh.

Temuan penelitian ini sejalan dengan penelitian Zhang et al. yang menunjukkan bahwa *association rule Mining* mampu mengidentifikasi hubungan antar kejadian gangguan pada Distribution Terminal Unit. Namun demikian, penelitian ini memiliki perbedaan pada pembentukan transaksi dan tujuan analisis. Zhang et al. membangun transaksi berdasarkan alarm dan status operasi perangkat sehingga aturan asosiasi digunakan untuk mendukung *fault diagnosis*, sedangkan penelitian ini membentuk transaksi dari data historis gangguan yang mengintegrasikan atribut jenis gangguan, penyebab gangguan, dan lokasi kejadian sehingga aturan asosiasi lebih diarahkan untuk mendukung penyusunan strategi pemeliharaan berbasis data.

Demikian pula dengan penelitian Kuang et al., *association rule* digunakan untuk meningkatkan kualitas data pengukuran melalui deteksi *outlier*, sedangkan pada penelitian ini *association rule* dimanfaatkan untuk menghasilkan pengetahuan mengenai pola gangguan berulang sebagai dasar penentuan prioritas inspeksi dan *preventive maintenance*.

Secara keseluruhan, hasil penelitian menunjukkan bahwa algoritma Apriori mampu menghasilkan informasi yang tidak hanya menggambarkan frekuensi gangguan, tetapi juga hubungan antar karakteristik gangguan yang terjadi secara bersamaan. Informasi tersebut berpotensi dimanfaatkan sebagai dasar penyusunan prioritas inspeksi, penentuan fokus pemeliharaan preventif, serta pengambilan keputusan pemeliharaan berbasis data pada jaringan distribusi PT PLN (Persero). Namun demikian, interpretasi hasil tetap bergantung pada kualitas pembentukan transaksi dan konsistensi data historis yang digunakan sebagai masukan algoritma Apriori.

3.4 Implikasi Praktis bagi PLN

Temuan penelitian memiliki implikasi langsung terhadap pengelolaan jaringan distribusi PT PLN (Persero). Aturan asosiasi yang diperoleh dapat dimanfaatkan untuk: (a) Mengidentifikasi penyebab gangguan dominan pada lokasiterte gangguanntu; (b) Menentukan prioritas inspeksi dan pemeliharaan berdasarkan pola gangguan historis; (c) Mendukung penyusunan program *preventive maintenance* dan *condition-based maintenance*; (d) Meningkatkan efektivitas pengelolaan aset distribusi melalui pemanfaatan data historis; (e) Menjadi masukan dalam pengambilan keputusan untuk meningkatkan keandalan jaringan distribusi.

4. KESIMPULAN

Penelitian ini berhasil menerapkan algoritma Apriori dengan pendekatan CRISP-DM untuk mengidentifikasi pola asosiasi gangguan berulang pada jaringan distribusi PT PLN (Persero) berdasarkan data historis gangguan. Setelah melalui tahap data preparation, diperoleh 10.274 transaksi yang digunakan sebagai masukan pada proses association rule mining. Dengan menggunakan parameter minimum support sebesar 0,5%, minimum confidence sebesar 70%, dan nilai lift lebih besar dari 1, algoritma Apriori menghasilkan 64 aturan asosiasi yang menunjukkan adanya hubungan antar jenis gangguan, penyebab gangguan, dan lokasi kejadian.

Berdasarkan hasil analisis, gangguan yang paling sering muncul berkaitan dengan infrastruktur fiber optic, yaitu “Core Fo Problem, FOC – Putus Core, dan Putus Kabel FO”. Selain itu, aturan asosiasi dengan nilai lift tertinggi yaitu “Banten + FOT – Software → Konfigurasi Di Router Problem” menunjukkan hubungan asosiasi yang kuat antar komponen gangguan. Hasil tersebut menunjukkan bahwa algoritma Apriori mampu menggali pola keterkaitan gangguan yang tidak terlihat secara langsung dari data historis dan menghasilkan informasi berbasis data (*maintenance-oriented knowledge*).

Penelitian ini memberikan kontribusi berupa pendekatan analisis pola gangguan yang dapat digunakan sebagai referensi dalam menentukan prioritas inspeksi dan penyusunan strategi pemeliharaan preventif pada jaringan distribusi. Namun, hasil penelitian masih berupa rekomendasi berbasis pola data dan belum mencakup implementasi langsung pada sistem operasional. Penelitian selanjutnya dapat dilakukan melalui validasi bersama praktisi PT PLN (Persero), integrasi hasil aturan asosiasi ke dalam sistem pendukung keputusan, serta pengembangan model transaksi yang lebih representatif untuk meningkatkan kualitas analisis pola gangguan.

Hasil analisis menunjukkan bahwa infrastruktur Fiber Optic (FO) merupakan salah satu kelompok gangguan yang paling sering muncul berdasarkan nilai support pada frequent itemset, sehingga dapat menjadi salah satu fokus dalam analisis pola gangguan berulang. Selain itu, aturan asosiasi FOT – Software → MCB Problem menunjukkan adanya hubungan asosiasi positif antar kedua item tersebut berdasarkan nilai support, confidence, dan lift yang diperoleh. Namun, hasil association rule mining tidak menunjukkan hubungan sebab-akibat, sehingga pola tersebut diinterpretasikan sebagai keterkaitan antar kejadian gangguan yang dapat menjadi informasi tambahan dalam analisis gangguan.

Dari sisi praktis, hasil penelitian ini berpotensi memberikan informasi pendukung dalam menentukan prioritas inspeksi, penyusunan rencana pemeliharaan preventif, serta evaluasi pola gangguan berulang pada jaringan distribusi. Association rules yang dihasilkan dapat menjadi masukan dalam mengidentifikasi kombinasi gangguan dan komponen jaringan yang perlu mendapat perhatian lebih. Namun, pemanfaatan hasil penelitian dalam proses operasional masih memerlukan validasi lebih lanjut bersama pihak terkait atau domain expert PT PLN (Persero) sebelum diterapkan sebagai dasar pengambilan keputusan pemeliharaan.

Penelitian selanjutnya disarankan menggunakan dataset dengan periode pengamatan yang lebih panjang serta menambahkan atribut lain, seperti kondisi cuaca, beban penyulang, indeks keandalan (SAIFI, SAIDI, dan ENS), maupun kondisi aset distribusi. Selain itu, hasil algoritma Apriori dapat dibandingkan dengan metode *Frequent Pattern Growth* (FP-Growth) atau dikombinasikan dengan teknik *machine learning* dan *predictive maintenance* untuk meningkatkan kemampuan analisis pola gangguan serta mendukung pengembangan sistem pendukung keputusan pada jaringan distribusi tenaga listrik.

DAFTAR PUSTAKA

- [1] M. K. da Silva, R. S. Lopes, M. E. Fontana, and J. K. B. de Almeida, "A Data-Driven Approach to Maintenance-Preventable Causes of Failure Analysis in Power Distribution Systems," *Brazilian Journal of Operations & Production Management*, vol. 22, no. 3, Art. no. 2702, 2025, doi: 10.14488/BJOPM.2702.2025.
- [2] S. Barja-Martinez, M. Aragüés-Peñalba, Í. Munné-Collado, P. Lloret-Gallego, E. Bullich-Massagué, and R. Villafafila-Robles, "Artificial Intelligence Techniques for Enabling Big Data Services in Distribution Networks: A Review," *Renewable and Sustainable Energy Reviews*, vol. 150, Art. no. 111459, 2021, doi: 10.1016/j.rser.2021.111459.
- [3] M. Aghahadi, A. Bosisio, M. Merlo, A. Berizzi, A. Pegoiani, and S. Forciniti, "Digitalization Processes in Distribution Grids: A Comprehensive Review of Strategies and Challenges," *Applied Sciences*, vol. 14, no. 11, Art. no. 4528, 2024, doi: 10.3390/app14114528.
- [4] Kamgala, A. K., and Singh, C., "Reliability Assessment Framework for the Distribution System Including Distributed Energy Resources," *IEEE Transactions on Sustainable Energy*, vol. 12, no. 3, pp. 1539–1548, Jul. 2021, doi: 10.1109/TSTE.2021.3053911.
- [5] M. Klerx, J. Morren, and H. Slootweg, "Analyzing Parameters That Affect the Reliability of Low Voltage Cable Grids and Their Applicability in Asset Management," *IEEE Transactions on Power Delivery*, vol. 34, no. 4, pp. 1432–1441, Aug. 2019, doi: 10.1109/TPWRD.2019.2903928.
- [6] M. A. Mahmoud, N. R. M. Nasir, M. Gurunathan, P. Raj, and S. A. Mostafa, "The Current State of the Art in Research on Predictive Maintenance in Smart Grid Distribution Network: Fault's Types, Causes, and Prediction Methods—A Systematic Review," *Energies*, vol. 14, no. 16, Art. no. 5078, 2021, doi: 10.3390/en14165078.

- [7] N. Zhou, Y. Xu, S. Cho, and C. T. Wee, "A Systematic Review for Switchgear Asset Management in Power Grids: Condition Monitoring, Health Assessment, and Maintenance Strategy," *IEEE Transactions on Power Delivery*, vol. 38, no. 5, pp. 3296–3311, 2023, doi: 10.1109/TPWRD.2023.3272883.
- [8] X. Zhang, Y. Tang, Q. Liu, G. Liu, X. Ning, and J. Chen, "A Fault Analysis Method Based on Association Rule Mining for Distribution Terminal Unit," *Applied Sciences*, vol. 11, no. 11, Art. no. 5221, 2021, doi: 10.3390/app11115221.
- [9] H. Kuang, R. Qin, M. He, X. He, R. Duan, C. Guo, and X. Meng, "An Association Rules-Based Method for Outliers Cleaning of Measurement Data in the Distribution Network in Smart Grids," *Frontiers in Energy Research*, vol. 9, Art. no. 730058, 2021, doi: 10.3389/fenrg.2021.730058.
- [10] Q. Sun, "Enhancing Power Grid Data Analysis with Fusion Algorithms for Efficient Association Rule Mining in Large-Scale Datasets," *International Journal of Computers Communications & Control*, vol. 19, no. 3, 2024, doi: 10.15837/ijccc.2024.3.6232.
- [11] H. Rezapour, S. Jamali, and A. Bahmanyar, "Review on Artificial Intelligence-Based Fault Location Methods in Power Distribution Networks," *Energies*, vol. 16, no. 12, Art. no. 4636, 2023, doi: 10.3390/en16124636.
- [12] M. Bindi, M. C. Piccirilli, A. Luchetta, and F. Grasso, "A Comprehensive Review of Fault Diagnosis and Prognosis Techniques in High Voltage and Medium Voltage Electrical Power Lines," *Energies*, vol. 16, no. 21, Art. no. 7317, 2023, doi: 10.3390/en16217317.
- [13] Y. Cao, J. Tang, S. Shi, D. Cai, L. Zhang, and P. Xiong, "Fault Diagnosis Techniques for Electrical Distribution Network Based on Artificial Intelligence and Signal Processing: A Review," *Processes*, vol. 13, no. 1, Art. no. 48, 2025, doi: 10.3390/pr13010048.
- [14] D. B. Saputra, V. Atina, and F. E. Nastiti, "Penerapan Model CRISP-DM pada Prediksi Nasabah Kredit Menggunakan Algoritma Random Forest," *Idealis: Indonesia Journal Information System*, vol. 7, no. 2, pp. 240–247, Jul. 2024, doi.org/10.36080/idealis.v7i2.3244.
- [15] F. Prasetyo and H. Hasugian, "Analisis Pola Pembelian Produk Makanan Menggunakan Algoritma FP-Growth untuk Strategi Penjualan," *Jurnal Idealis*, vol. 7, no. 1, pp. 11–20, Jan. 2023, doi: org/10.36080/idealis.v7i1.3085.