

BAB II

TINJAUAN PUSTAKA

2.1 ISOLATION FOREST

Algoritma *Isolation Forest* digunakan untuk mendeteksi anomali dalam jaringan komputer. Algoritma ini beroperasi berdasarkan prinsip mengisolasi titik data dan telah terbukti efektif dalam mengidentifikasi anomali dengan memanfaatkan karakteristik outlier yang terisolasi, yang sering diproses lebih cepat daripada titik data biasa. Metodologi penelitian ini memerlukan pengumpulan data jaringan yang mencakup lalu lintas normal dan anomali. Data ini kemudian diproses terlebih dahulu untuk menghilangkan gangguan diformat untuk analisis. Algoritma *Isolation Forest* kemudian digunakan untuk mengembangkan model deteksi anomali. Model tersebut dilatih menggunakan sebagian data berlabel dan kemudian dievaluasi dengan data yang tidak diketahui untuk menilai keakuratan dan efektivitasnya dalam mendeteksi anomali. Temuan menunjukkan bahwa algoritma *Isolation Forest* berhasil mengidentifikasi anomali jaringan dengan menghasilkan skor anomali untuk setiap sampel. Algoritma ini bekerja dengan "mengisolasi" pengamatan melalui pemilihan fitur secara acak dan nilai pembagi yang berada di antara nilai maksimum dan minimum dari fitur yang dipilih. Karena partisi rekursif dapat diilustrasikan dengan struktur pohon, jumlah partisi yang diperlukan untuk mengisolasi sampel sesuai dengan panjang jalur dari simpul akar ke simpul terminal. Panjang jalur ini, yang dirata-ratakan di seluruh hutan pohon acak, berfungsi sebagai ukuran kenormalan dan fungsi keputusan. Pemisahan acak

menghasilkan jalur yang lebih pendek untuk anomali, jadi ketika hutan pohon acak secara kolektif menghasilkan panjang jalur yang lebih pendek untuk sampel tertentu, hal itu menunjukkan kemungkinan yang lebih tinggi bahwa sampel tersebut merupakan anomali, yang menghasilkan akurasi yang tinggi.

Isolation Forest merupakan teknik deteksi anomali berdasarkan struktur pohon yang menggunakan metode khusus untuk menentukan titik data yang tidak biasa, yang disebut sebagai anomali. Berbeda dengan teknik lain yang mencoba memodelkan distribusi data, *Isolation Forest* mengisolasi titik data secara langsung. Konsep yang mendasarinya adalah bahwa anomali dapat diisolasi lebih mudah daripada data normal karena cenderung terletak jauh dari kumpulan data. Algoritme tersebut membangun beberapa pohon keputusan secara acak, di mana, pada setiap tahap, sebuah fitur dipilih secara acak dan data dibagi menurut nilai fitur tersebut. Titik data yang diisolasi lebih cepat yang memerlukan lebih sedikit pemisahan diidentifikasi sebagai anomali. Karena efisiensinya, *Isolation Forest* sering diterapkan pada kumpulan data besar dengan banyak dimensi dan sangat efektif untuk deteksi anomali di berbagai bidang, termasuk keamanan siber, deteksi penipuan, dan pemeliharaan prediktif (Carudin., 2024).

Tahapan-tahapan *Outlier Algoritma Isolation Forest*

1. Pengumpulan Data: Kumpulkan dan bersihkan data yang relevan.
2. Pemilihan Fitur: Pilih fitur yang akan digunakan untuk deteksi anomali.
3. Pembentukan Model: Inisialisasi model *Isolation Forest* dengan parameter seperti jumlah pohon, `max_samples`, dan tingkat kontaminasi (`contamination`).

4. Training Model: Latih model dengan data untuk mempelajari pola normal dan deteksi anomali.
5. Skor Anomali: Hitung skor anomali untuk setiap titik data, dengan nilai mendekati 1 menunjukkan anomali.
6. Evaluasi Model: Evaluasi performa model menggunakan ROC curve, confusion matrix, atau tuning threshold.
7. Deteksi Anomali: Gunakan model untuk mendeteksi anomali di dataset baru atau dalam aplikasi real-time.
8. Pemantauan dan Tuning: Monitor dan, jika perlu, update model berdasarkan perubahan pola data.

2.2 Beban Puncak

Beban puncak listrik mengacu pada penggunaan listrik maksimum yang tercatat selama periode waktu tertentu, seperti harian, mingguan, atau bulanan. Puncak ini biasanya terjadi antara pukul 17.00 dan 22.00. Puncak ini merupakan momen ketika permintaan listrik konsumen mencapai level tertingginya dalam jangka waktu tertentu, baik diukur dalam jam, hari, minggu, bulan, atau bahkan tahun.

PLN menggunakan dua istilah untuk membedakan waktu pemakaian listrik pelanggannya, yaitu Waktu Beban Puncak (WBP) dan Waktu Beban Non-Puncak (LWBP). Istilah WBP digunakan karena pada periode ini terjadi peningkatan pemakaian listrik secara serentak di sebagian besar pelanggan PLN di rumah-rumah mereka. Biasanya, periode WBP berlangsung dari pukul 17.00 hingga 22.00. Sebaliknya, Waktu Beban Non-Puncak mengacu pada periode pemakaian listrik

yang merupakan kebalikan dari Waktu Beban Puncak. Dalam praktiknya, pemakaian listrik selama WBP mungkin tidak secara konsisten berlangsung selama 5 jam penuh antara pukul 17.00 dan 22.00, tetapi dapat dipastikan bahwa kondisi WBP akan terjadi dalam rentang waktu tersebut.

Tingkat konsumsi listrik selama WBP cenderung signifikan dan secara konsisten memengaruhi penurunan tegangan secara keseluruhan. Ketika perangkat listrik dan elektronik beroperasi pada tegangan yang menurun, perangkat tersebut dipaksa bekerja lebih lama dan mengerahkan lebih banyak tenaga. Ini adalah salah satu faktor yang berkontribusi terhadap peningkatan penggunaan listrik selama WBP.(Anca., 2019)

Data beban puncak untuk Trafo II 20 MVA, bersumber dari PLN Palopo dan mencakup catatan penggunaan listrik dari tahun 2019 hingga 2023. Kumpulan data ini menangkap variasi dalam konsumsi listrik, dengan penekanan pada periode Waktu Beban Puncak (WBP) dan luar waktu beban puncak (LWBP). Kumpulan data ini menawarkan wawasan komprehensif tentang pola beban listrik bulanan.

2.3 Skor Anomali

Menghitung Anomaly Score Anomaly score adalah nilai yang menunjukkan kriteria anomali row data. Secara matematis, anomaly score didapat dari hasil perhitungan formula di bawah ini:

$$s(x, n) = 2 \frac{E(h(x))}{c(n)} \quad (1)$$

Ket.

$H(x)$: *path length*

$E(h(x))$: rata-rata *path length* dari setiap *isolation tree*.

$C(n)$: rata-rata $h(x)$ dari nilai variabel ke- n

Path length adalah jarak yang mengukur titik di mana data tidak dapat dibagi lagi dengan root node dari struktur algoritma pohon keputusan. Anomaly score sendiri memiliki rentang nilai dari -1 sampai dengan 1. Anomaly score yang makin mendekati -1 dianggap anomali dan makin mendekati nilai 1 dianggap normal (Ahmad Zulfikar dkk., 2023)

2.4 Gardu induk

Pusat pembangkit listrik umumnya di hubungkan dengan saluran transmisi udara yang menyalurkan tenaga listrik dari pusat pembangkit ke pusat-pusat konsumsi tenaga listrik, yaitu gardu gardu induk (GI). Proses penyaluran tenaga listrik dari pusat pembangkit ke pusat beban biasanya melibatkan gardu induk (GI) yang merupakan bagian integral dari sistem penyaluran (transmisi) tenaga listrik. Gardu induk ini mengubah tenaga listrik tegangan tinggi ke tingkat tegangan lain atau tegangan menengah, sekaligus melakukan pengukuran, pengawasan, operasi, dan pengaturan keamanan pada sistem tenaga listrik. Selain itu, GI juga mengelola penyaluran tenaga listrik ke gardu induk lain melalui saluran tegangan tinggi dan ke gardu induk distribusi melalui penyulang tegangan menengah. Gardu induk merupakan komponen integral dari jaringan transmisi dan distribusi listrik. Sistem transmisi melibatkan pemindahan energi listrik dari pembangkit listrik ke gardu induk, sedangkan sistem distribusi menyalurkan energi listrik dari gardu induk ke

konsumen akhir. Saat energi listrik mengalir melalui sistem transmisi atau distribusi, terjadi penurunan tegangan di sepanjang jaringan. Penurunan tegangan ini dapat disebabkan oleh beberapa faktor, termasuk panjang jaringan, luas penampang, jenis resistansi, sambungan yang buruk, arus yang mengalir melalui konduktor, dan faktor daya.(Muhammad Fahmi Hakim dkk., 2023)

2.5 Penelitian Sebelumnya

Tabel 2.1 Penelitian Sebelumnya

No	Peneliti	Judul Penelitian	Hasil Penelitian
1.	Ahmad Zulfikar, Farhan Ariq Rahmani, Nurul Azizah(2023)	“Deteksi Anomali Menggunakan Isolation Forest Belanja Barang Persediaan Konsumsi Pada Satuan Kerja Kepolisian Republik Indonesia”	Hasil pengujian menunjukkan bahwa model isolation forest memiliki potensi untuk mendeteksi anomali yang terjadi pada data transaksi akun belanja persediaan (521811). Berdasarkan hasil uji coba beberapa nilai contamination parameter, nilai 0,3% merupakan nilai yang paling optimal dalam uji coba model yang dipilih dalam mendeteksi anomali. Terlepas dari valid atau tidaknya indikasi fraud pada data yang dideteksi sebagai anomali oleh

No	Peneliti	Judul Penelitian	Hasil Penelitian
			<p>model, hasil learning dari penelitian ini dapat dimasukkan sebagai fitur tambahan yang berfungsi sebagai early warning system pada aplikasi otorisasi belanja satker, seperti aplikasi sistem aplikasi keuangan tingkat instansi (SAKTI).</p>
2.	<p>Fei Tony Liu, Kai Ming Ting (2008).</p>	<p>“Isolation Forest”</p>	<p>Pada dasarnya, Isolation Forest adalah detektor anomali yang akurat dan efisien, terutama untuk basis data besar. Kemampuannya dalam menangani basis data dengan volume tinggi sangat diinginkan untuk aplikasi dunia nyata."</p>
3.	<p>Julien Lesouplea, edric Baudoinb, Marc Spigaib, Jean-Yves Tournereta (2021).</p>	<p>“Generalized Isolation Forest For Anomaly Detection”</p>	<p>Makalah ini menyajikan algoritma hutan isolasi baru yang dikenal sebagai hutan isolasi umum untuk deteksi anomali.</p>

No	Peneliti	Judul Penelitian	Hasil Penelitian
			<p>Algoritma ini memungkinkan penghilangan artefak tertentu yang terkait dengan hutan isolasi, sehingga menghasilkan pohon yang tidak memiliki cabang kosong—keterbatasan yang ditemukan dalam algoritma hutan isolasi yang diperluas (EIF). Eksperimen yang dilakukan pada kumpulan data sintetis dan tolok ukur memungkinkan kita untuk menilai kinerja metode yang diusulkan, yang menghasilkan hasil yang sebanding dengan yang dicapai dengan EIF. Meskipun demikian, algoritma yang diusulkan secara signifikan mengurangi waktu eksekusi dibandingkan dengan EIF dan memerlukan lebih sedikit</p>

No	Peneliti	Judul Penelitian	Hasil Penelitian
			parameter untuk disimpan (ambang batas pada setiap simpul untuk hutan isolasi umum versus vektor intersepsi untuk setiap simpul dalam EIF).
4.	Oktavia Triana (2024).	“Deteksi Anomali Jaringan Menggunakan Algoritma Isolation Forest”	Studi ini menunjukkan bahwa algoritma Isolation Forest merupakan alat yang sangat efektif dan efisien untuk mendeteksi anomali jaringan, dengan potensi aplikasi yang luas dalam keamanan siber dan manajemen jaringan. Hasilnya diharapkan dapat meningkatkan keamanan dan keandalan sistem informasi secara signifikan sekaligus membuka jalan bagi penelitian lebih lanjut dalam pengembangan metode deteksi anomali yang lebih canggih dan adaptif.

No	Peneliti	Judul Penelitian	Hasil Penelitian
5.	Gregorius Airlangga (2024).	“Advanced Machine Learning Techniques For Seismic Anomaly Detection In Indonesia: A Comparative Study Of Lof, Isolation Forest, And One - Class Svm”	Kesimpulannya, penerapan Faktor Outlier Lokal (LOF), Hutan Isolasi, dan SVM Kelas Satu untuk deteksi anomali seismik telah menghasilkan wawasan informatif mengenai kemampuan masing-masing algoritma dalam domain khusus ini. Uji H Kruskal-Wallis mengonfirmasi perbedaan signifikan dalam kinerja algoritma, dengan SVM Kelas Satu dan Hutan Isolasi menunjukkan kemampuan yang unggul untuk membedakan antara aktivitas seismik normal dan outlier. SVM Kelas Satu menunjukkan skor siluet tertinggi, yang menunjukkan
6.	Windy Oktavian Hengki Parulian	“Perkiraan Beban Puncak Rata-Rata	Untuk memperkirakan beban puncak rata-rata pada bulan atau

No	Peneliti	Judul Penelitian	Hasil Penelitian
	Hasibuan \ Janter Napitupulu Joslan Sinaga (2023).	Pada Penyulang Batunadua Gardu Induk Padang Sidempuan Dengan Metode Arima (Auto Regressive Integrated Moving Average)”	tahun tertentu, diperlukan data beban puncak listrik pada bulan atau tahun yang sama. Misalnya, jika kita ingin memperkirakan beban puncak pada bulan November 2022, kita memerlukan data beban puncak yang akurat dari bulan November tahun sebelumnya.
7.	Fifi D. Rosalina , Yuniar Farida, Abdulloh Hamid (2016).	“Metode Logika Fuzzy Sebagai Evaluasi Distribusi Daya Listrik Berdasarkan Beban Puncak Pembangkit Tenaga Listrik”	Dalam penelitian ini, model logika fuzzy yang menggunakan metode Mamdani digunakan untuk mengevaluasi daya listrik dalam sistem pembangkitan daya, dengan mempertimbangkan tiga faktor yang memengaruhi: kapasitas terpasang, daya mampu, dan produksi. Namun, seperti yang ditunjukkan oleh data pada Tabel 2, aplikasi fuzzy menghasilkan

No	Peneliti	Judul Penelitian	Hasil Penelitian
			nilai beban puncak dari 10 sampel data, dengan tingkat kesalahan tertinggi mencapai 60% dan terendah sebesar 1,39%.
8.	Yayu Triwulan, Nasrun Hariyanto, Sabat Anwari (2013).	“Peramalan Beban Puncak Listrik Jangka Pendek Menggunakan Metode Jaringan Syaraf Tiruan”	Hasil peramalan beban listrik dengan menggunakan Jaringan Syaraf Tiruan (JST) Backpropagation menunjukkan bahwa metode ini menghasilkan peramalan yang lebih mendekati data target dan memberikan wawasan yang signifikan dibandingkan dengan metode koefisien beban PLN. Analisis statistik menunjukkan bahwa rata-rata galat peramalan yang dihasilkan dengan menggunakan metode koefisien beban PLN adalah 1,85%, sehingga menghasilkan tingkat akurasi

No	Peneliti	Judul Penelitian	Hasil Penelitian
			<p>sebesar 98,15%. Sebaliknya, rata-rata galat peramalan dengan menggunakan algoritma JST Backpropagation hanya 0,12%, sehingga mencapai tingkat akurasi sebesar 99,88%. Oleh karena itu, dapat dibuktikan bahwa peramalan beban dengan pendekatan JST menghasilkan hasil peramalan yang lebih akurat.</p>
9.	Afriansyah, Bakhtiar, Thalib Bini (2019).	<p>“Studi Pengaruh Beban Puncak Terhadap Susut Umur Transformator di PT PLN (Persero) Rayon Daya”</p>	<p>Beban tinggi pada transformator berkontribusi pada penurunan masa pakainya. Di antara empat transformator yang dianalisis, dua unit menunjukkan suhu lilitan yang melampaui standar IEC 354. Transformator pertama (PT.IBT.007), yang beroperasi pada 90,3% dari kapasitas 200 kVA-nya (dengan pendinginan</p>

No	Peneliti	Judul Penelitian	Hasil Penelitian
			<p>ONAN), mencatat suhu lilitan 98,13°C, yang mengakibatkan pengurangan masa pakai sebesar 122% dan perkiraan sisa masa pakai sekitar 15 tahun. Transformator kedua (PT.IBT.SPTC), yang beroperasi pada 91,5% dari kapasitasnya dengan suhu lilitan 99,2°C, juga menghadapi pengurangan masa pakai sebesar 122%, sehingga sisa masa pakainya sekitar 12 tahun.</p>
10.	Dessy Putri Wahyuningtyas, et. al. 2014.	“Studi Pengaruh Sedimen Terhadap Beban Puncak Pada Plta Sengguruh”	<p>Hasil cross-sectional dihasilkan, Echosounding dilakukan di 50 lokasi, termasuk 17 titik di sepanjang Sungai Brantas yang mengalir ke waduk Sengguruh, 31 titik di sepanjang Sungai Lesti yang mengarah ke waduk Sengguruh, dan 2 titik di dekat</p>

No	Peneliti	Judul Penelitian	Hasil Penelitian
			intake waduk. Penelitian ini menganalisis 3 bagian: bagian SB 2, bagian SL 2, dan bagian SB 2 SL 2. Data echosounding yang digunakan dalam penelitian ini berasal dari tahun 2008, 2011, 2012, dan 2015. Selain itu, analisis ini memberikan wawasan berharga tentang endapan sedimen yang ada di dasar waduk.