

ANALISIS POLA KONSUMSI ENERGI BANGUNAN BERDASARKAN DATA SENSOR LINGKUNGAN BERBASIS ALGORITMA APRIORI

Jadiaman Parhusip^{1*}, Naufal Ihsan Sriyanto², Tegar Pratama³, Zhykwa Cerryl Mavanudin⁴, Darryl Sandi Augustin⁵

¹²³⁴⁵Teknik Informatika, Universitas Palangka Raya; Jalan Yos Sudarso, Kota Palangka Raya, Kalimantan Tengah 73112.

Keywords:

frequent pattern mining; Apriori algorithm; energy consumption; association rules; environmental sensors

Corespondent Email:

parhusip.jadiaman@it.upr.a
c.id

Abstrak. Konsumsi energi listrik pada bangunan sangat dipengaruhi oleh faktor-faktor lingkungan seperti suhu dan kelembapan yang dapat diidentifikasi melalui analisis pola data. Penelitian ini bertujuan untuk menemukan hubungan antara kondisi lingkungan dengan tingkat konsumsi energi menggunakan algoritma Apriori. Dataset yang digunakan adalah *Tetuan City Power Consumption* yang terdiri atas 52.416 entri dengan sembilan atribut sensor. Data diproses melalui tahap pra-pemrosesan, diskritisasi, dan penerapan algoritma Apriori dengan parameter minimum *support* sebesar 5% dan minimum *confidence* sebesar 30%. Hasil analisis menunjukkan bahwa kondisi kelembapan tinggi dan suhu rendah berkorelasi kuat dengan konsumsi energi rendah dengan nilai *lift* 1,973 dan *confidence* 65,8%. Sebaliknya, kondisi suhu tinggi dan kelembapan rendah menjadi faktor yang memicu lonjakan konsumsi energi dengan nilai *lift* 1,857 dan *confidence* 61,8%. Temuan ini memberikan wawasan penting dalam pengembangan sistem manajemen energi bangunan yang efisien serta dapat dijadikan dasar dalam perancangan sistem peringatan dini untuk mengantisipasi beban puncak berdasarkan data sensor lingkungan.



Copyright © [JITET](#) (Jurnal Informatika dan Teknik Elektro Terapan). This article is an open access article distributed under terms and conditions of the Creative Commons Attribution (CC BY NC)

Abstract. Electric energy consumption in buildings is highly influenced by environmental factors such as temperature and humidity, which can be identified through pattern analysis. This study aims to discover association rules between environmental conditions and energy consumption levels using the Apriori algorithm. The dataset used is the *Tetuan City Power Consumption* dataset, consisting of 52,416 entries and nine sensor attributes. The data underwent preprocessing, discretization, and Apriori implementation with a minimum support of 5% and a minimum confidence of 30%. The analysis results show that high humidity and low temperature are strongly correlated with low energy consumption, achieving a lift value of 1.973 and a confidence of 65.8%. Conversely, high temperature and low humidity were found to be the main triggers for increased energy consumption, with a lift value of 1.857 and a confidence of 61.8%. These findings provide valuable insights for developing efficient building energy management systems and early warning mechanisms to anticipate peak load conditions based on environmental sensor data.

1. PENDAHULUAN

Energi listrik merupakan salah satu kebutuhan mendasar dalam kehidupan modern. Peningkatan jumlah bangunan dan intensitas penggunaan perangkat elektronik menyebabkan konsumsi energi mengalami pertumbuhan yang signifikan setiap tahunnya. Jika tidak dikelola dengan baik, kondisi tersebut berpotensi menimbulkan kelebihan beban pada jaringan listrik, peningkatan biaya operasional, serta dampak lingkungan berupa peningkatan emisi karbon. Oleh karena itu, analisis pola konsumsi energi menjadi langkah penting dalam mendukung kebijakan efisiensi energi dan keberlanjutan lingkungan.

Dalam konteks manajemen energi bangunan, pemanfaatan pendekatan analitik berbasis data (*data-driven analysis*) menjadi semakin relevan. Melalui penerapan teknik data mining dan *machine learning*, berbagai penelitian berupaya mengungkap pola tersembunyi antara variabel lingkungan seperti suhu, kelembapan, dan intensitas radiasi matahari dengan tingkat konsumsi energi. Pendekatan ini tidak hanya memungkinkan prediksi kebutuhan energi secara lebih akurat, tetapi juga membantu dalam mengidentifikasi perilaku penggunaan energi yang tidak efisien.

Penelitian oleh Ashouri et al. [1] menunjukkan bahwa metode data mining mampu mengembangkan sistem rekomendasi penghematan energi berdasarkan kombinasi faktor lingkungan dan karakteristik bangunan. Sementara itu, Mishra et al. [2] mengusulkan model *Long Short-Term Memory* (LSTM) yang memanfaatkan data historis dan faktor cuaca untuk memprediksi konsumsi energi bangunan dengan tingkat akurasi yang tinggi. Pendekatan berbasis aturan asosiasi juga terbukti efektif dalam mengungkap pola keterkaitan antar variabel, seperti yang ditunjukkan oleh Triyanto et al. [3] yang berhasil menerapkan algoritma Apriori untuk menemukan pola kombinasi itemset yang signifikan pada data transaksi penjualan. Keandalan metode ini dalam mengidentifikasi hubungan tersembunyi menjadi dasar penerapannya pada domain lain, termasuk analisis pola konsumsi energi.

Lebih lanjut, Nazeriye et al. [4] menerapkan pendekatan *data mining* berbasis *clustering* dan algoritma klasifikasi untuk menganalisis pengaruh karakteristik bangunan dan faktor lingkungan terhadap efisiensi serta konsumsi

energi pada skala besar. Temuan mereka menegaskan bahwa algoritma data mining mampu mengungkap hubungan kompleks antara variabel lingkungan dan perilaku konsumsi energi, sekaligus menghasilkan pengetahuan yang lebih detail melalui pemodelan berbasis klaster. Pendekatan serupa dapat diterapkan pada dataset *Tetuan City Power Consumption* untuk mengidentifikasi keterkaitan antara variabel lingkungan (suhu, kelembapan, dan intensitas cahaya) dengan total konsumsi energi listrik pada berbagai zona bangunan di Kota Tetuan.

Berdasarkan hal tersebut, penelitian ini bertujuan untuk menganalisis pola hubungan antara kondisi lingkungan dan tingkat konsumsi energi menggunakan algoritma Apriori. Fokus penelitian mencakup tiga aspek utama, yaitu: (1) menemukan kombinasi kondisi lingkungan yang menjadi pemicu utama terjadinya beban puncak energi, (2) mengidentifikasi kondisi yang berkaitan dengan konsumsi energi rendah, dan (3) mengukur kekuatan hubungan antar variabel dengan metrik *support*, *confidence*, dan *lift*. Melalui pendekatan ini, diharapkan dapat diperoleh pemahaman yang lebih komprehensif mengenai perilaku konsumsi energi bangunan serta mendukung strategi efisiensi energi di masa mendatang.

2. TINJAUAN PUSTAKA

2.1 Data Mining dan Tahapan CRISP-DM

Data mining merupakan proses otomatisasi analisis data yang besar untuk menemukan pola, tren, dan hubungan antar variabel dalam data yang kompleks. Tujuan utama dari data mining adalah mengubah data mentah menjadi informasi yang berguna bagi pengambilan keputusan bisnis dan penelitian ilmiah. Dalam konteks penelitian ini, data mining digunakan untuk mengidentifikasi pola konsumsi energi bangunan dari data sensor lingkungan.

Model proses data mining yang umum digunakan adalah CRISP-DM (*Cross Industry Standard Process for Data Mining*). Tahapan utama dalam CRISP-DM meliputi: 1) *Business Understanding*, yaitu pemahaman terhadap tujuan bisnis dan konteks penelitian; 2) *Data Understanding*, meliputi pengumpulan dan eksplorasi data awal; 3) *Data Preparation*, termasuk pembersihan, normalisasi, serta fitur ekstraksi; 4) *Modeling*, tahap pembangunan model analitik; 5) *Evaluation*, evaluasi hasil

model berdasarkan kriteria tertentu; dan 6) *Deployment*, implementasi hasil untuk pengambilan keputusan nyata.[5]

Dalam penelitian ini, fokus berada pada tahap *Modeling* dan *Evaluation*, dimana model *association rule mining* menggunakan algoritma Apriori diimplementasikan untuk menemukan pola yang signifikan dalam data sensor energi. Tahapan ini penting karena hasil dari model akan digunakan untuk analisis pola konsumsi energi yang akurat dan bermanfaat untuk pengelolaan energi bangunan.

2.2 Eksplorasi Data dan Pra-pemrosesan

Eksplorasi data (*Exploratory Data Analysis/EDA*) sangat penting dalam memahami sebaran variabel, hubungan antar variabel, serta mendeteksi anomali atau outlier dalam dataset. Melalui EDA, karakteristik data seperti distribusi suhu, kelembapan, dan konsumsi energi dapat dianalisis sehingga proses pemodelan dapat lebih akurat dan efisien.

Pra-pemrosesan data meliputi langkah-langkah: 1) *Data Cleaning* yang berfungsi memperbaiki data yang hilang, duplikasi, atau inkonsistensi. 2) Normalisasi agar variabel dengan rentang berbeda dapat diperlakukan secara adil dalam algoritma. 3) Transformasi terutama diskritisasi variabel kontinu menjadi kategori yang memudahkan proses *association rule mining*. 4) *Feature Engineering* untuk menciptakan fitur baru yang relevan dan meningkatkan kualitas pemodelan. Tahapan ini sangat krusial karena data sensor energi biasanya memiliki variabilitas tinggi dan kemungkinan terdapat nilai-nilai ekstrim atau outlier yang dapat mengganggu proses pembelajaran algoritma.[6]

2.3 Association Rule Mining

Association Rule Mining digunakan untuk mengidentifikasi pola keterkaitan antar item dalam dataset dengan menghasilkan aturan implikasi yang menggambarkan hubungan antar elemen. Teknik ini bertujuan menemukan pola yang menarik secara efisien melalui analisis item yang sering muncul bersamaan. Evaluasi aturan dilakukan menggunakan metrik dasar seperti *support* untuk mengukur kekuatan kemunculan item dan *confidence* untuk menilai konsistensi kombinasi item dalam suatu transaksi.[7]

2.4 Algoritma Apriori

Algoritma Apriori merupakan salah satu metode *association rule learning* yang paling populer dalam bidang data mining. Algoritma ini pertama kali diperkenalkan untuk menemukan pola asosiasi antar item dalam basis data transaksi. Prinsip utama algoritma ini adalah *apriori property*, yaitu setiap subset dari *frequent itemset* juga harus *frequent*.

Dengan prinsip ini, algoritma dapat menyaring kombinasi item yang tidak memenuhi syarat secara efisien. Tahapan algoritma Apriori dimulai dengan menghitung frekuensi setiap item tunggal dalam dataset. Item yang memiliki frekuensi lebih rendah dari nilai minimum support akan dihapus. Selanjutnya, algoritma membentuk kombinasi 2-itemset, lalu 3-itemset, dan seterusnya, sambil menghitung nilai support dari masing-masing kombinasi. Proses ini berlanjut hingga tidak ada lagi kandidat itemset yang memenuhi kriteria.[8]

2.5 Metrik Evaluasi Aturan Asosiasi

Kualitas aturan asosiasi dalam data mining dievaluasi menggunakan tiga metrik utama, yaitu *support*, *confidence*, dan *lift*, yang menggambarkan frekuensi, kepercayaan, dan kekuatan hubungan antar itemset dalam dataset.

Support mengukur proporsi transaksi yang mengandung itemset tertentu dibandingkan jumlah total transaksi. *Support* menandai seberapa sering itemset tersebut muncul dalam dataset sehingga dapat dianggap signifikan untuk dianalisis lebih lanjut. Secara matematis, *support* dari itemset X adalah rasio jumlah transaksi yang mengandung X dengan total transaksi.

Confidence adalah ukuran probabilitas kepercayaan aturan asosiasi, yakni seberapa besar kemungkinan item *consequent* muncul jika *antecedent* sudah ada. *Confidence* dihitung sebagai perbandingan jumlah transaksi yang mengandung kedua *antecedent* dan *consequent* dengan jumlah transaksi yang mengandung *antecedent* saja.

Lift mengukur kekuatan hubungan antara *antecedent* dan *consequent* dibandingkan kejadian terpisah secara acak. Nilai *lift* lebih besar dari 1 menandakan adanya hubungan positif yang signifikan antara itemset, nilai 1 menunjukkan tidak ada asosiasi, dan kurang

dari 1 berarti hubungan negatif. Dengan demikian, *lift* sangat berguna untuk menilai relevansi aturan dari sudut pandang statistik independensi. Pemahaman dan penggunaan ketiga metrik tersebut sangat penting untuk memastikan aturan asosiasi yang ditemukan valid, berguna, dan tidak sekadar kebetulan statistik. Sebagai contoh, aturan dengan *confidence* tinggi tetapi *support* rendah mungkin kurang bermakna, sementara *lift* dapat menyoroti kekuatan asosiasi yang sesungguhnya.[9]

2.6 Penelitian Terkait

Beberapa penelitian telah menerapkan algoritma Apriori dan metode *frequent pattern mining* untuk menganalisis pola konsumsi energi serta data sensor lingkungan dan bangunan. Penelitian oleh Sunardiyo memodelkan prakiraan konsumsi energi listrik pada bangunan dengan mengaplikasikan data mining. Hasil penelitian menunjukkan bahwa algoritma Apriori efektif dalam mengidentifikasi pola konsumsi energi yang berulang dengan tingkat kepercayaan (*confidence*) yang tinggi. Namun, penelitian ini masih terbatas pada pengujian dataset dari satu jenis bangunan, sehingga generalisasi hasil perlu diuji pada konteks bangunan lain.[10]

Suswitaningrum melakukan analisis intensitas konsumsi energi listrik dan faktor lingkungan pada sebuah gedung. Penelitian ini menggunakan *frequent pattern mining* untuk melihat hubungan antara variabel lingkungan dan pola konsumsi energi. Hasilnya, ditemukan beberapa pola asosiasi yang signifikan antara suhu dan kelembapan dengan tingkat konsumsi energi. Keterbatasan penelitian ini adalah kurangnya integrasi dengan data sensor *real-time* yang dapat memberikan dinamika konsumsi secara lebih aktual.[11]

Atadjawa mengimplementasikan algoritma Apriori pada data penjualan alat-alat listrik untuk mengidentifikasi pola perilaku konsumen. Meskipun fokus utamanya pada penjualan produk, metodologi dan hasilnya memberikan gambaran yang berguna terkait analisis pola asosiasi dalam konteks konsumsi energi listrik. Penelitian ini membatasi kompleksitas dataset sehingga belum menguji skala data besar dengan variasi yang luas.[12]

Penelitian lain dari jurnal internasional menggunakan Apriori untuk analisis pola

konsumsi energi di gedung dan IoT, seperti studi yang menggunakan Apriori untuk mengidentifikasi hubungan antara parameter cuaca dan konsumsi energi bangunan. Studi tersebut menemukan bahwa kombinasi suhu, kelembapan, dan kecepatan angin memengaruhi tingkat konsumsi energi dengan tingkat *support* dan *confidence* yang signifikan.[13] Selain itu, riset yang memanfaatkan *frequent pattern mining* pada data sensor lingkungan industri juga berhasil meningkatkan akurasi prediksi dan identifikasi pola penggunaan energi, meskipun tantangan terkait volume data besar dan kecepatan pemrosesan masih menjadi kendala.[14] Menjawab tantangan tersebut, penelitian terbaru menerapkan pendekatan *big data* dan *association rule mining* dalam lingkungan IoT untuk menangani arus data yang masif. Metode ini terbukti mampu mengungkap pola tersembunyi (hidden patterns) dari sensor bangunan dengan efisiensi komputasi yang lebih tinggi, sehingga sangat relevan untuk strategi penghematan energi pada skala sistem yang lebih kompleks [15].

Berdasarkan penelitian-penelitian tersebut, penerapan algoritma Apriori pada data sensor lingkungan diharapkan mampu mengungkap pola konsumsi energi yang signifikan dan memberikan kontribusi untuk pengelolaan energi yang lebih efisien dan berkelanjutan.

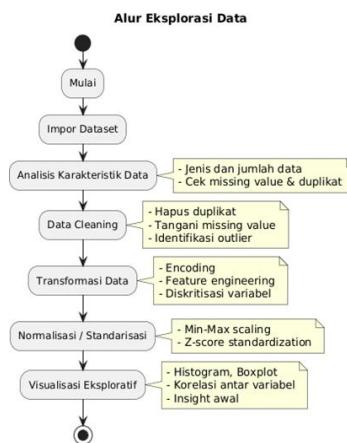
3. METODE PENELITIAN

3.1. Sumber Data

Penelitian ini menggunakan dataset *Tetuan City Power Consumption* yang tersedia secara publik. Dataset berisi data deret waktu konsumsi energi listrik pada tiga zona bangunan di Kota Tetuan, dilengkapi dengan variabel sensor lingkungan yang dicatat secara simultan. Total dataset terdiri atas 52.416 entri dengan 9 atribut yang mencakup informasi temporal dan pengukuran sensor. Data Eksplorasi dilakukan terhadap dataset *Tetuan City Power Consumption* yang berisi data deret waktu konsumsi energi listrik pada tiga zona bangunan di Kota Tetuan, disertai dengan atribut kondisi lingkungan seperti suhu, kelembapan, kecepatan angin, dan radiasi matahari.

Secara umum, tahapan eksplorasi data mencakup analisis karakteristik data, pembersihan data, transformasi dan normalisasi, hingga visualisasi eksploratif. Alur

proses eksplorasi data ditunjukkan pada **Gambar 1**.



Gambar 1. Alur Eksplorasi Data

Dalam tahap awal, dilakukan identifikasi karakteristik data meliputi jenis atribut, jumlah record, serta pemeriksaan nilai hilang (missing values) dan duplikasi. Hasil pemeriksaan menunjukkan bahwa dataset tidak mengandung data hilang maupun duplikasi. Selanjutnya dilakukan proses data cleaning untuk memastikan konsistensi nilai dan mendeteksi kemungkinan outlier. Nilai ekstrem yang ditemukan pada variabel radiasi dan konsumsi energi dianalisis lebih lanjut dan tetap dipertahankan karena merepresentasikan kondisi lingkungan yang valid.

Tahap berikutnya adalah transformasi data, yang meliputi pembuatan atribut turunan dari kolom waktu (*DateTime*) seperti *Month* dan *Hour*, serta penggabungan konsumsi energi dari tiga zona menjadi satu atribut *Total Power Consumption*. Seluruh atribut numerik yang relevan kemudian didiskritisasi menggunakan metode berbasis kuantil dengan tiga kategori, yaitu rendah, sedang, dan tinggi, untuk menjaga keseimbangan distribusi data. Proses normalisasi diterapkan untuk menyetarakan skala antarvariabel numerik dengan menggunakan metode *Min-Max Scaling* pada rentang [0,1]. Tahapan ini dilakukan agar perbedaan satuan pengukuran tidak memengaruhi analisis korelasi dan pengenalan pola.

Sebagai langkah akhir, dilakukan visualisasi eksploratif untuk memahami pola distribusi dan hubungan antarvariabel. Visualisasi menggunakan histogram, boxplot, dan *heatmap*

korelasi membantu dalam mengidentifikasi tren umum serta keterkaitan antara suhu, kelembapan, dan konsumsi energi listrik. Hasil eksplorasi ini memberikan pemahaman awal yang menjadi dasar dalam proses analisis lebih lanjut menggunakan algoritma Apriori.

3.3. Pra-Pemrosesan Data

Tahap pra-pemrosesan data dilakukan untuk mentransformasi data numerik menjadi kategori yang dapat diproses oleh algoritma Apriori. Variabel suhu dibagi menjadi tiga kategori (rendah, sedang, tinggi) menggunakan metode diskritisasi berbasis kuantil (*qcut* dengan parameter $q=3$). Hal yang sama dilakukan pada variabel kelembapan dan konsumsi listrik total. Metode kuantil dipilih untuk memastikan distribusi data yang seimbang pada setiap kategori, yaitu sepertiga dari total data untuk masing-masing interval. Pendekatan ini secara inheren mempartisi data kontinu ke dalam tiga kategori dengan jumlah populasi data yang kurang lebih setara, sehingga menghasilkan nilai *support* yang mendekati 33,3% untuk setiap kondisi tunggal.

3.4. Parameter Algoritma

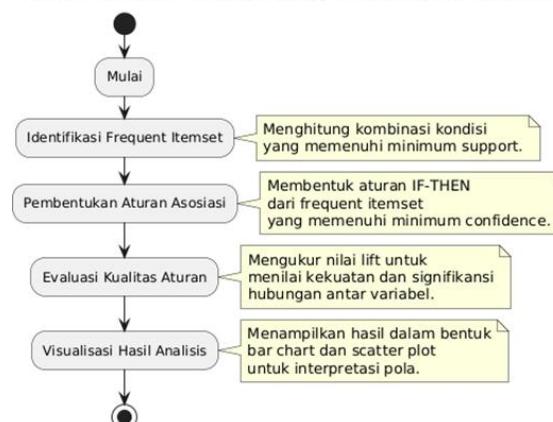
Tahap ini menjelaskan konfigurasi parameter yang digunakan pada proses penambangan pola menggunakan algoritma Apriori. Algoritma Apriori dipilih karena kemampuannya dalam mengidentifikasi pola keterkaitan antarvariabel secara efisien berdasarkan prinsip *frequent itemset mining*. Tujuan utama dari proses ini adalah menemukan aturan asosiasi (*association rules*) yang menggambarkan hubungan antara variabel suhu, kelembapan, dan konsumsi energi listrik.

Dalam penelitian ini, tiga parameter utama digunakan untuk mengukur kualitas dan kekuatan aturan asosiasi, yaitu *support*, *confidence*, dan *lift*. Nilai minimum *support* ditetapkan sebesar 0,3, yang berarti hanya kombinasi kondisi yang muncul pada minimal 30% dari total data yang akan dipertimbangkan sebagai *frequent itemset*. Penentuan nilai ini dimaksudkan untuk menyeimbangkan antara banyaknya aturan yang dihasilkan dan relevansi hubungan antarvariabel, sehingga pola yang ditemukan tetap signifikan namun tidak terlalu jarang muncul dalam data. Selanjutnya, minimum *confidence* ditetapkan sebesar 0,7, yang menunjukkan bahwa setiap aturan yang

terbentuk harus memiliki tingkat keyakinan minimal 70% agar dianggap valid. Parameter ini memastikan bahwa setiap aturan asosiasi yang dihasilkan memiliki reliabilitas tinggi dan dapat diinterpretasikan secara konsisten. Sementara itu, nilai *lift* digunakan sebagai metrik tambahan untuk menilai kekuatan hubungan antarvariabel. Nilai *lift* yang lebih besar dari 1 menandakan bahwa kemunculan variabel pada bagian *antecedent* secara positif memengaruhi kemunculan variabel pada bagian *consequent*, sehingga hubungan tersebut bersifat asosiatif dan signifikan. Aturan dengan nilai *lift* tertinggi dianggap paling menarik karena menunjukkan adanya hubungan yang kuat dan tidak terjadi secara kebetulan.

Hasil *frequent itemset* dan aturan asosiasi divisualisasikan menggunakan diagram batang dan scatter plot untuk mempermudah interpretasi hubungan antarvariabel. Visualisasi ini juga digunakan untuk mengidentifikasi aturan dengan nilai *lift* dan *confidence* tertinggi, yang dianggap mewakili pola konsumsi energi paling signifikan pada dataset yang dianalisis. Analisis dilakukan dalam beberapa tahapan. Pertama, identifikasi *frequent itemset* dengan menghitung frekuensi kemunculan setiap kombinasi kondisi yang memenuhi minimum *support*. Kedua, pembentukan aturan asosiasi dari *frequent itemset* yang memenuhi minimum *confidence*. Ketiga, evaluasi kualitas aturan menggunakan metrik *lift* untuk mengidentifikasi hubungan yang signifikan. Keempat, visualisasi hasil menggunakan bar chart dan scatter plot untuk memudahkan interpretasi pola yang ditemukan.

Alur Tahapan Analisis Data Menggunakan Algoritma Apriori



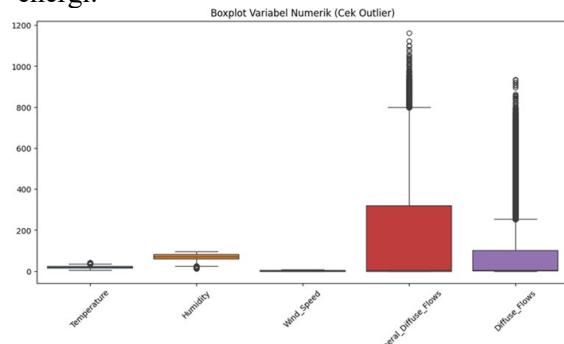
Gambar 2. Alur Tahapan Analisis

Proses analisis dilakukan melalui empat tahapan utama, sebagaimana ditunjukkan pada **Gambar 2**. Setiap tahap memiliki peran spesifik dalam menghasilkan pola hubungan yang signifikan antara variabel suhu, kelembapan, dan konsumsi energi listrik.

4. HASIL DAN PEMBAHASAN

4.1 Eksplorasi Data

Eksplorasi data bertujuan untuk memahami karakteristik dan struktur awal dari dataset *Tetuan City Power Consumption*, yang berisi informasi konsumsi energi listrik pada tiga zona bangunan beserta data sensor lingkungan (suhu, kelembapan, kecepatan angin, dan intensitas cahaya). Analisis ini dilakukan untuk memperoleh gambaran menyeluruh mengenai sebaran nilai, potensi anomali, serta hubungan antarvariabel yang relevan terhadap konsumsi energi.

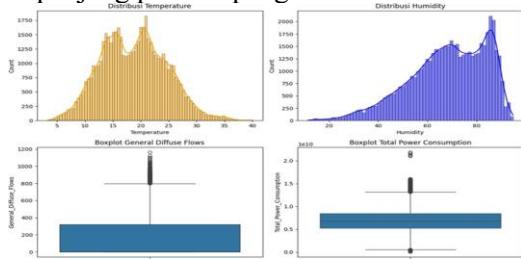


Gambar 3. Boxplot Identifikasi Outlier

Gambar 3 menampilkan boxplot dari variabel numerik utama pada dataset *Tetuan City Power Consumption*, yang meliputi suhu (*Temperature*), kelembapan (*Humidity*), kecepatan angin (*Wind Speed*), *General Diffuse Flows*, dan *Diffuse Flows*. Dari visualisasi terlihat bahwa sebagian besar variabel memiliki sebaran nilai yang relatif rapat tanpa outlier yang signifikan, kecuali dua variabel terakhir yaitu *General Diffuse Flows* dan *Diffuse Flows*. Kedua variabel tersebut menunjukkan banyak outlier dengan nilai ekstrim yang jauh dari rentang interkuartil (IQR), mengindikasikan adanya kondisi lingkungan ekstrem dengan intensitas radiasi atau difusi cahaya yang sangat tinggi.

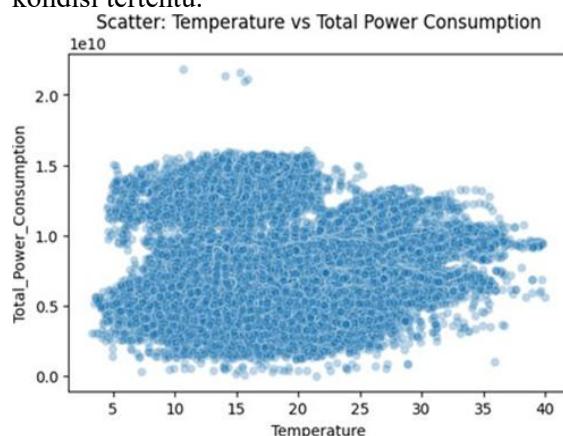
Kehadiran outlier pada variabel tersebut tidak dihapus karena dianggap merepresentasikan fenomena aktual pada sistem sensor energi di Kota Tetuan, bukan

kesalahan pencatatan data. Sementara itu, variabel suhu, kelembapan, dan kecepatan angin memiliki distribusi yang lebih stabil, menunjukkan bahwa pengukuran terhadap kondisi lingkungan cenderung konsisten sepanjang periode pengamatan.

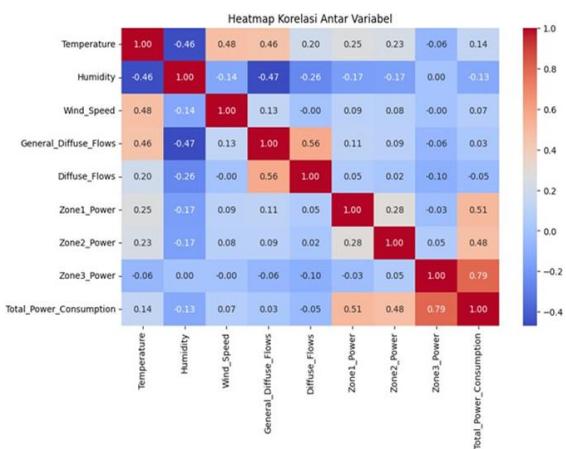


Gambar 4. Visualisasi Data

Grafik distribusi pada **Gambar 4** menunjukkan bahwa variabel suhu memiliki sebaran mendekati normal dengan rata-rata sekitar 18 - 20°C, sedangkan kelembapan cenderung miring ke kanan (*right-skewed*) yang menandakan dominasi kondisi lembab. Nilai konsumsi energi menunjukkan variasi besar dengan beberapa titik ekstrem, menggambarkan adanya periode lonjakan pemakaian daya pada kondisi tertentu.



Gambar 5. Scatter Plot *Temperature* vs *Total Power*



Gambar 6. Heatmap Korelasi

Scatter Plot Temperature vs Total Power menunjukkan pola menyebar, tanpa korelasi linear yang kuat. Hal ini mengindikasikan bahwa konsumsi daya tidak hanya dipengaruhi oleh suhu. Sedangkan, Heatmap Korelasi memperlihatkan bahwa *Total Power Consumption* paling kuat berkorelasi dengan *Zone 3 Power Consumption* (0.79) dan lumayan dengan *Zone 1 & 2*. Korelasi dengan faktor cuaca seperti *Temperature*, *Humidity*, dan *Wind Speed* cenderung lemah.

4.2 Frequent Itemset

Analisis menghasilkan 15 *frequent itemset* utama dengan nilai *support* berkisar antara 14,2% hingga 33,4%. Itemset dengan nilai *support* tertinggi adalah kelembapan sedang dan suhu sedang, masing-masing 33,4%. Kondisi tunggal lainnya seperti kelembapan rendah, suhu rendah, konsumsi total tinggi, rendah, sedang, serta suhu tinggi memiliki nilai *support* yang hampir seimbang di angka 33,3%. Fenomena di mana sembilan itemset teratas yang semuanya merupakan kondisi tunggal memiliki nilai *support* yang mendekati 33,3% merupakan konsekuensi metodologis yang dapat diprediksi dari tahap pra-pemrosesan data. Hal ini disebabkan oleh penerapan metode diskritisasi berbasis kuantil yang secara inheren mempartisi data kontinu ke dalam tiga interval dengan jumlah populasi data yang kurang lebih setara untuk setiap intervalnya.

Itemset kombinasi dua kondisi menunjukkan nilai *support* yang lebih rendah, mencerminkan probabilitas gabungan yang secara matematis lebih kecil. Kombinasi konsumsi total rendah dengan suhu rendah memiliki *support* tertinggi di antara kombinasi

dua kondisi, yaitu 17,6%, diikuti oleh konsumsi total rendah dengan kelembapan tinggi sebesar 16,8%, kelembapan rendah dengan suhu tinggi sebesar 16,7%, dan suhu tinggi dengan konsumsi total tinggi sebesar 16,1%. Distribusi lengkap *frequent itemset* ditunjukkan pada Tabel 1.

Tabel 1. *Frequent Itemset* dengan Minimum *Support* 5%

No	Itemset	Support (%)
1	Kelembapan Sedang	33,4
2	Suhu Sedang	33,4
3	Kelembapan Rendah	33,3
4	Suhu Rendah	33,3
5	Konsumsi Total Tinggi	33,3
6	Konsumsi Total Rendah	33,3
7	Konsumsi Total Sedang	33,3
8	Suhu Tinggi	33,3
9	Kelembapan Tinggi	33,2
10	Konsumsi Total Rendah, Suhu Rendah	17,6
11	Konsumsi Total Rendah, Kelembapan Tinggi	16,8
12	Kelembapan Rendah, Suhu Tinggi	16,7
13	Suhu Tinggi, Konsumsi Total Tinggi	16,1
14	Kelembapan Tinggi, Suhu Rendah	14,2
15	Kelembapan Rendah, Konsumsi Total Sedang	14,2

4.3 Aturan Asosiasi

Dari *frequent itemset* yang ditemukan, dihasilkan 10 aturan asosiasi dengan nilai *lift* tertinggi. Aturan dengan performa terbaik adalah kelembapan tinggi dan suhu rendah mengimplikasikan konsumsi total rendah, dengan *support* 9,3%, *confidence* 65,8%, dan *lift* 1,973. Nilai *lift* mendekati 2 menunjukkan bahwa probabilitas konsumsi energi rendah hampir dua kali lipat lebih besar ketika kedua kondisi lingkungan tersebut terjadi bersamaan dibandingkan dengan probabilitas normal

Aturan kedua yang signifikan adalah kelembapan rendah dan konsumsi total tinggi mengimplikasikan suhu tinggi, dengan *support* 8,2%, *confidence* 61,8%, dan *lift* 1,857. Aturan ini mengidentifikasi kondisi suhu tinggi dan kelembapan rendah sebagai pemicu utama lonjakan konsumsi energi. Kondisi lingkungan yang panas dan kering menciptakan ketidaknyamanan termal yang mendorong penggunaan sistem pendingin ruangan secara intensif. Meskipun aturan ini tidak secara langsung memprediksi konsumsi, nilai *lift* sebesar 1,857 mengonfirmasi bahwa ketiga elemen dalam itemset ini memiliki hubungan sinergis yang kuat. Tabel 2 menampilkan 10 aturan asosiasi dengan nilai *lift* tertinggi.

Tabel 2. Aturan Asosiasi dengan Minimum *Confidence* 30%

No	Antecedent	Consequent	Support (%)	Confidence (%)
1	Kelembapan Tinggi, Suhu Rendah	Konsumsi Total Rendah	9,3	65,8
2	Kelembapan Rendah, Suhu Tinggi, Konsumsi Total Tinggi	Suhu Tinggi	8,2	61,8
3	Konsumsi Total Rendah, Kelembapan Tinggi	Suhu Rendah	9,3	55,6
4	Konsumsi	Kelembapa	9,3	53,2

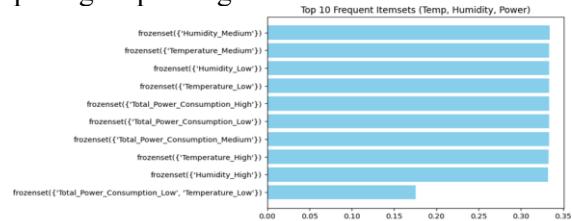
		Total Rendah, Suhu Rendah	n Tinggi		
5	Suhu Tinggi, Konsumsi Total Sedang	Kelembapan Rendah	7,1	53,3	
6	Konsumsi Total Rendah, Kelembapan Sedang	Suhu Rendah	5,7	53,1	
7	Suhu Rendah	Konsumsi Total Rendah	17,6	52,7	
8	Konsumsi Total Rendah	Suhu Rendah	17,6	52,7	
9	Suhu Tinggi, Konsumsi Total Tinggi	Kelembapan Rendah	8,2	51,2	
10	Konsumsi Total Rendah, Suhu Sedang	Kelembapan Tinggi	6	50,5	

Aturan-aturan lain dengan nilai *lift* di atas 1,5 menunjukkan asosiasi positif yang kuat antara kombinasi kondisi suhu dan kelembapan dengan tingkat konsumsi energi. Sebagai contoh, aturan ketiga menunjukkan bahwa konsumsi total rendah dengan kelembapan tinggi mengimplikasikan suhu rendah dengan *confidence* 55,6% dan *lift* 1,668. Aturan keempat memperlihatkan bahwa konsumsi total rendah dengan suhu rendah berkorelasi dengan kelembapan tinggi (*confidence* 53,2%, *lift* 1,602). Selain itu, suhu rendah secara individual berkorelasi dengan konsumsi total rendah dengan *confidence* 52,7% dan *lift* 1,580, menunjukkan bahwa faktor suhu memiliki pengaruh signifikan terhadap konsumsi energi meskipun tanpa mempertimbangkan kelembapan.

4.4 Visualisasi dan Interpretasi

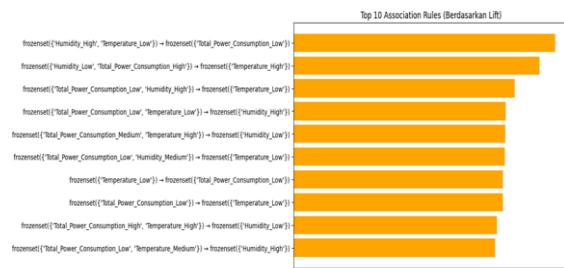
Visualisasi hasil analisis disajikan dalam bentuk grafik untuk memudahkan interpretasi pola yang ditemukan. **Gambar 7** menunjukkan distribusi nilai *support* untuk 10 *frequent itemset* teratas. Dari grafik tersebut, terlihat

bahwa kelembapan sedang (*Humidity_Medium*) menjadi item dengan frekuensi paling tinggi, disusul oleh suhu sedang (*Temperature_Medium*) dan kelembapan rendah (*Humidity_Low*). Hal ini menunjukkan bahwa kondisi lingkungan di dataset Tetuan City mayoritas berada pada suhu sedang dan kelembapan sedang hingga rendah. Menariknya, kombinasi item seperti konsumsi total rendah dengan suhu rendah juga masuk ke dalam daftar 10 besar, meskipun nilai *support*nya lebih kecil dibanding item tunggal. Fakta ini memberikan informasi penting bahwa terdapat hubungan alami antara kondisi suhu rendah dan rendahnya konsumsi daya listrik, yang masuk akal secara kontekstual karena pada suhu rendah kebutuhan penggunaan perangkat pendingin relatif lebih kecil.



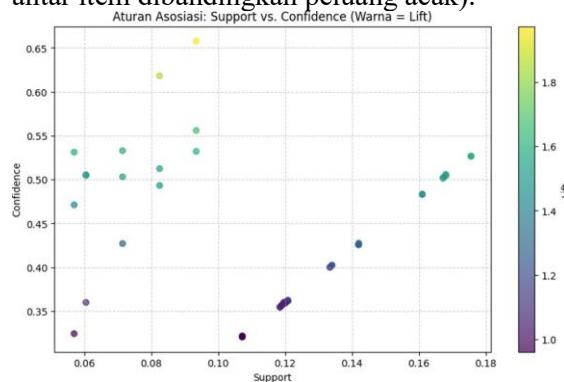
Gambar 7. Distribusi nilai *support* untuk 10 *frequent itemset* teratas

Gambar 7 menampilkan 10 aturan asosiasi dengan nilai *lift* tertinggi. *Lift* dipilih karena bukan hanya menggambarkan frekuensi bersama, tetapi juga menilai seberapa kuat hubungan antar item dibandingkan dengan peluang acak. Aturan dengan *lift* lebih besar dari 1 dianggap signifikan karena menunjukkan adanya asosiasi positif. Sebagai contoh, aturan kelembapan tinggi dan suhu rendah mengimplikasikan konsumsi total rendah memiliki *lift* mendekati angka 2. Artinya, peluang konsumsi listrik rendah menjadi dua kali lebih tinggi jika kondisi kelembapan tinggi dan suhu rendah terjadi bersamaan. Secara logis, aturan ini sesuai dengan kenyataan bahwa dalam kondisi udara lembab dan dingin, kebutuhan perangkat elektronik untuk pengatur suhu berkurang, sehingga konsumsi listrik juga menurun.



Gambar 8. Aturan asosiasi dengan nilai *lift* tertinggi

Untuk memperoleh pemahaman yang komprehensif terhadap karakteristik dari himpunan aturan asosiasi yang dihasilkan, digunakan visualisasi scatter plot seperti ditunjukkan pada **Gambar 9**. Metode ini dipilih karena kemampuannya untuk memetakan tiga metrik evaluasi utama *support*, *confidence*, dan *lift* secara simultan ke dalam satu representasi grafis, sehingga memungkinkan analisis multidimensi. Visualisasi ini memberikan gambaran yang lebih detail mengenai kualitas aturan asosiasi dengan menggunakan tiga dimensi: sumbu X merepresentasikan *support* (seberapa sering aturan terjadi di dataset), sumbu Y merepresentasikan *confidence* (seberapa besar kepercayaan bahwa konsekuensi benar jika anteseden terpenuhi), dan warna merepresentasikan *lift* (seberapa kuat hubungan antar item dibandingkan peluang acak).



Gambar 9. Hubungan *support*, *confidence*, dan *lift* untuk aturan asosiasi

Dari grafik terlihat bahwa sebagian besar aturan memiliki *support* di rentang 0,06 hingga 0,18, yang berarti aturan-aturan tersebut muncul dalam 6% hingga 18% dari total data. *Confidence* bervariasi antara 0,30 hingga 0,65, menunjukkan tingkat kepastian sedang. Artinya aturan tidak selalu deterministik, tetapi tetap cukup mumpuni untuk dijadikan dasar

interpretasi. Warna titik menunjukkan variasi nilai *lift*. Titik dengan warna hijau terang hingga kuning menandakan aturan dengan *lift* tinggi ($\geq 1,5$), sedangkan warna biru gelap menunjukkan *lift* mendekati atau di bawah 1 (kurang signifikan). Menariknya, terdapat beberapa aturan dengan *support* sedang (sekitar 0,10 hingga 0,15) namun *confidence* relatif tinggi ($>0,60$). Aturan seperti ini tergolong stabil karena cukup sering terjadi sekaligus memiliki tingkat kepastian yang lebih tinggi dibanding aturan lain. Sebaliknya, terdapat juga aturan dengan *support* rendah tetapi *lift* sangat tinggi, yang berarti meski pola jarang muncul, ketika terjadi maka hubungan antar variabel sangat kuat.

4.5 Implikasi Praktis

Temuan penelitian ini memiliki implikasi penting untuk manajemen energi bangunan. Aturan asosiasi yang mengidentifikasi kondisi hemat energi optimal dapat digunakan sebagai acuan dalam pengaturan sistem pendingin ruangan. Ketika sensor mendeteksi kombinasi suhu rendah dan kelembapan tinggi, sistem dapat mengurangi atau mematikan perangkat pendingin untuk menghemat energi tanpa mengorbankan kenyamanan penghuni. Kombinasi suhu rendah dan kelembapan tinggi merepresentasikan kondisi termal yang nyaman bagi penghuni, sehingga meminimalkan atau meniadakan penggunaan sistem pendingin ruangan yang merupakan kontributor utama beban listrik.

Sebaliknya, aturan yang mengidentifikasi pemicu beban puncak dapat diintegrasikan dengan sistem peringatan dini. Dengan memanfaatkan data ramalan cuaca, sistem dapat mengantisipasi lonjakan permintaan energi ketika prakiraan menunjukkan periode suhu tinggi dan kelembapan rendah. Hal ini memungkinkan manajemen beban proaktif dan alokasi sumber daya yang lebih efisien untuk menjaga stabilitas pasokan energi. Dalam konteks manajemen energi, pola-pola dengan karakteristik ini dapat digunakan untuk membuat sistem peringatan dini atau strategi efisiensi energi yang lebih efektif.

5. KESIMPULAN DAN SARAN

- a. Hasil penelitian menunjukkan bahwa algoritma Apriori efektif dalam menemukan pola penting konsumsi energi listrik pada dataset *Tetuan City*. Kelembapan tinggi dan suhu rendah cenderung menghasilkan konsumsi energi rendah (*lift* = 1,973; *confidence* = 65,8%), sedangkan suhu tinggi dan kelembapan rendah memicu konsumsi tinggi (*lift* = 1,857; *confidence* = 61,8%).
- b. Kelebihan metode ini adalah kemampuannya menghasilkan aturan yang jelas dan mudah diterapkan. Pola-pola tersebut dapat mendukung pengelolaan energi bangunan, seperti pengaturan beban otomatis berbasis kondisi lingkungan serta penerapan sistem peringatan dini untuk mengantisipasi kenaikan konsumsi.
- c. Kekurangan penelitian terletak pada penggunaan tunggal algoritma Apriori yang membutuhkan proses komputasi lebih besar pada dataset besar. Selain itu, aturan asosiasi belum diuji menggunakan model prediktif, sehingga kemampuan generalisasi terhadap data baru belum dapat dipastikan.
- d. Pengembangan ke depan dapat mencakup penggunaan algoritma lain seperti *FP-Growth* untuk menilai efisiensi komputasi dan kualitas pola. Integrasi dengan model prediksi berbasis *machine learning* juga disarankan untuk meningkatkan akurasi perkiraan konsumsi energi di masa mendatang.

DAFTAR PUSTAKA

- [1] M. Ashouri, H. Rezaie, and M. Hammad, "Development of building energy saving advisory: A data mining approach," *Energy Build.*, vol. 163, pp. 128-140, Mar. 2018, doi: 10.1016/j.enbuild.2018.01.034.
- [2] S. Mishra et al., "DECODE: Data-driven Energy Consumption Prediction leveraging Historical Data and Environmental Factors in Buildings," *Energy Build.*, vol. 288, p. 113950, 2024, doi: 10.1016/j.enbuild.2024.113950.
- [3] T. Triyanto, H. Yulianti, and M. Ikhwani, "Penerapan Data Mining Terhadap Data Penjualan Dengan Menggunakan Algoritma Apriori Pada Toko Citra Utama," *Journal of Information System, Applied, Management, Accounting and Research (JISAMAR)*, vol. 8, no. 2, pp. 401-415, May 2024, doi: 10.5236/jisamar.v8i2.1485.
- [4] M. Nazeriye, A. Haeri, and F. Martínez-Álvarez, "Analysis of the Impact of Residential Property and Equipment on Building Energy Efficiency and Consumption—A Data Mining Approach," *Applied Sciences*, vol. 10, no. 10, p. 3589, 2020, doi: 10.3390/app10103589.
- [5] P. Chapman, J. Clinton, R. Kerber, T. Khabaza, T. Reinartz, C. Shearer, and R. Wirth, *CRISP-DM 1.0: Step-by-Step Data Mining Guide*. CRISP-DM Consortium (NCR Systems Engineering, DaimlerChrysler, SPSS, OHRA), 2000.
- [6] R. Maulid, "Teknik Analisis Data CRISP-DM dalam Data Mining," *DQLab Blog*, 2022. [Online]. Available: <https://dqlab.id/teknik-analisis-data-crisp-dm-dalam-data-mining>.
- [7] M. Fitriani, G. F. Nama, and Mardiana, "Implementasi Association Rule Dengan Algoritma Apriori Pada Data Peminjaman Buku UPT Perpustakaan Universitas Lampung Menggunakan Metodologi CRISP-DM," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 10, no. 1, 2022.
- [8] IBM, "Apa algoritma Apriori?" IBM, 2024. [Online]. Tersedia: <https://www.ibm.com/id-id/think/topics/apriori-algorithm>.
- [9] DQLab, "Algoritma Association Rule Learning dalam Data Science," 2024. [Online]. Tersedia: <https://dqlab.id/algoritma-association-rule-learning-dalam-data-science>
- [10] S. Sunardiyo, H. Ananta, R. D. M. Putri, R. T. Yofanda, and R. M. Yofatama, "Model Prakiraan Konsumsi Energi Listrik di Bangunan Gedung Kampus Berbasis Artificial Neural Network Backpropagation," in *Konservasi Alam*, vol. 3, pp. 149–194, 2022, doi: 10.15294/ka.v1i3.153.
- [11] R. P. Atadjawa, T. Haryanti, and L. Kurniawati, "Penerapan Asosiasi Algoritma Apriori Pada Data Penjualan Alat-Alat Listrik dan Teknik," *METIK*, vol. 5, no. 2, pp. 71–76, 2021, doi: 10.47002/metik.v5i2.290.
- [12] E. Suswitaningrum, N. Hudallah, R. D. M. Putri, and B. Sunarko, "Analisis Intensitas Konsumsi Energi Listrik dan Peluang Penghematan Energi Listrik pada Gedung C Kantor Sekretariat Daerah Kabupaten Semarang," *J. ELTIKOM: J. Tek. Elektro, Teknol. Inf. dan Komput.*, vol. 6, no. 1, pp. 26–39, Jun. 2022, doi: 10.31961/eltikom.v6i1.545.
- [13] C. A. Abukhadra et al., "Analysis and Prediction of Energy Consumption Using

- Apriori Algorithm," IJCSNS Int. J. Comput. Sci. Netw. Secur., vol. 23, no. 3, 2023.
- [14] K. Almalki dan A. Amine, "Improving Sensor Data Accuracy Using Frequent Pattern Mining," IEEE Sensors Journal, vol. 20, no. 5, pp. 2345-2352, 2024.
- [15] M. Dolores, C. Fernandez-Basso, J. Gómez-Romero, and M. J. Martin-Bautista, "A big data association rule mining based approach for energy building behaviour analysis in an IoT environment," *Scientific Reports*, vol. 13, no. 1, Art. no. 19810, Nov. 2023, doi: 10.1038/s41598-023-47056-1.