

Analisis Pola Akses Internet Penduduk Indonesia berdasarkan Karakteristik Provinsi menggunakan ARM

Analysis of Internet Access Patterns among the Indonesian Population based on Provincial Characteristics using ARM

¹Frans Budi Partogi Manik*, ²Angga Bayu Santoso

^{1,2}Program Studi Sistem Informasi, Fakultas Teknik dan Ilmu Komputer, Universitas Teknokrat Indonesia

^{1,2}Jl. Zainal Abidin Pagaralam No.9-11 Labuhan Ratu, Bandar Lampung, Indonesia

*e-mail: frans_budi@teknokrat.ac.id, anggabayu@teknokrat.ac.id

(received: 30 August 2026, revised: 7 September 2026, accepted: 8 September 2026)

Abstrak

Pemanfaatan internet di tanah air terus menunjukkan tren peningkatan, tetapi ciri khas akses internet di tiap provinsi tetap bervariasi, dipengaruhi oleh kepemilikan perangkat digital, penggunaan komputer, dan lokasi akses. Sebagian besar analisis yang dilakukan hingga kini masih bersifat deskriptif, sehingga belum mampu menyingkap hubungan antara indikator akses internet dengan cara yang menyeluruh. Tujuan dari penelitian ini adalah untuk mengenali pola interaksi antarindikator akses internet masyarakat Indonesia berdasarkan karakteristik provinsi melalui metode ARM menggunakan ARM. Data untuk penelitian ini diambil dari publikasi Badan Pusat Statistik (BPS) mengenai indikator pemanfaatan teknologi informasi dan internet across seluruh provinsi di Indonesia. Proses penelitian meliputi pengumpulan data, preprocessing, integrasi data, transformasi, diskretisasi, pembentukan data transaksi, penerapan ARM, dan evaluasi aturan asosiasi menggunakan nilai *support*, *confidence*, dan *lift*. Penelitian ini diharapkan mampu menghasilkan aturan asosiasi yang mencerminkan hubungan antarindikator akses internet di masing-masing provinsi, sehingga dapat memberikan wawasan mengenai karakteristik digital masing-masing wilayah di Indonesia. Hasil penelitian diharapkan menjadi referensi bagi pemerintah dalam perumusan kebijakan pemerataan infrastruktur digital serta mendukung kemajuan penggunaan Data Mining dalam analisis data statistik resmi.

Kata kunci: ARM, ARM, badan pusat statistik, data mining, akses internet.

Abstract

Internet Usage in Indonesia continues to show an upward trend, but the characteristics of internet access in each province vary, influenced by digital device ownership, computer Usage, and access location. Most analyses conducted to date have been descriptive in nature, and thus have not been able to reveal the relationships between internet access indicators in a comprehensive manner. The objective of this study is to identify patterns of interaction among internet access indicators for the Indonesian population based on provincial characteristics using the ARM method with the Apriori algorithm. The data for this study were drawn from a publication by the Central Statistics Agency (BPS) on indicators of information technology and internet Usage across all provinces in Indonesia. The research process included data collection, preprocessing, data integration, transformation, discretization, transaction data formation, application of the Apriori algorithm, and evaluation of association rules using support, confidence, and lift values. This study is expected to generate association rules that reflect the relationships among internet access indicators in each province, thereby providing insights into the digital characteristics of each region in Indonesia. The research results are expected to serve as a reference for the government in formulating policies for the equitable distribution of digital infrastructure and to support the advancement of data Mining in the analysis of official statistical data.

Keywords: ARM, A Priori Algorithm, central bureau of statistics, data mining, internet access.

1 Pendahuluan

Sekarang ini, internet sudah menjadi bagian penting dalam kehidupan sehari-hari masyarakat. Banyak orang menggunakan internet untuk mencari informasi, berkomunikasi, belajar, berbisnis, dan lain-lain. Di Indonesia, penggunaan internet juga semakin meningkat, tetapi masih ada perbedaan dalam hal akses dan pemanfaatan internet di setiap wilayah. Hal ini disebabkan oleh perbedaan kondisi sosial, ekonomi, pendidikan, dan infrastruktur di setiap daerah. Oleh karena itu, kita perlu menganalisis data tentang penggunaan internet untuk memahami pola hubungan antarindikator yang terkait dengan akses internet [1].

Beberapa penelitian telah dilakukan untuk menganalisis pola penggunaan internet menggunakan metode ARM. Penelitian Wibowo et al. menggunakan data Survei Sosial Ekonomi Nasional (Susenas) di Provinsi Jawa Timur untuk mengidentifikasi perbedaan pola penggunaan internet berdasarkan gender. Hasil penelitian menunjukkan bahwa ada kesamaan dan perbedaan pola aktivitas internet antara pengguna laki-laki dan perempuan. Penelitian ini juga menemukan bahwa ada hubungan antara beberapa aktivitas internet, seperti penggunaan layanan keuangan, berita, surat elektronik, dan barang serta jasa [2].

Penelitian lain oleh Hopipah et al. menggunakan metode *Modified Apriori* dengan teknik hashing untuk menganalisis pola akses pengunjung *Website* Universitas Singaperbangsa Karawang. Hasil penelitian menunjukkan bahwa metode asosiasi dapat digunakan untuk memahami perilaku pengguna berdasarkan data akses. Namun, penelitian ini masih terbatas pada aktivitas pengunjung suatu *Website* dan belum menganalisis karakteristik akses internet masyarakat berdasarkan wilayah administratif [3].

Penelitian Mansyur dan Ridho menerapkan *Association Rule Apriori* pada data permintaan publikasi Badan Pusat Statistik. Hasil penelitian menunjukkan bahwa pengguna cenderung mengakses publikasi yang sama dengan tahun berbeda secara bersamaan. Penelitian ini juga menunjukkan bahwa hasil *Association Rule* dapat digunakan sebagai dasar untuk meningkatkan kualitas layanan *Website* BPS. Berdasarkan penelitian terdahulu, ARM telah terbukti dapat digunakan untuk menemukan pola hubungan dari data penggunaan internet, aktivitas *Website*, dan data layanan statistik. Namun, penelitian terdahulu masih memiliki ruang pengembangan, terutama dalam penggunaan unit analisis wilayah yang lebih luas. Oleh karena itu, penelitian ini dilakukan untuk menganalisis pola akses internet penduduk Indonesia berdasarkan karakteristik provinsi menggunakan ARM [4][5].

Penelitian ini menggunakan data dari publikasi Badan Pusat Statistik yang memuat indikator pemanfaatan teknologi informasi dan internet pada provinsi di Indonesia. Data akan melalui tahapan preprocessing, transformasi, dan diskretisasi sebelum dibentuk menjadi transaksi yang dapat dianalisis menggunakan ARM. Aturan yang dihasilkan selanjutnya dievaluasi berdasarkan *support*, *confidence*, dan *lift* untuk memperoleh hubungan antarindikator yang memiliki keterkaitan kuat.

Tujuan penelitian ini adalah mengidentifikasi pola hubungan antarindikator akses internet berdasarkan karakteristik provinsi di Indonesia. Penelitian ini berbeda dengan penelitian sebelumnya yang lebih banyak berfokus pada perilaku *individu*, akses *Website*, atau permintaan publikasi. Penelitian ini menggunakan provinsi sebagai unit analisis dan memanfaatkan data statistik resmi BPS untuk memperoleh pola asosiasi pada skala nasional. Hasil penelitian diharapkan dapat memberikan gambaran mengenai keterkaitan karakteristik akses internet antarprovinsi dan menjadi informasi pendukung dalam memahami kondisi serta pemerataan pemanfaatan teknologi informasi di Indonesia.

2 Tinjauan Literatur

Pada penelitian sebelumnya Menggunakan ARM untuk mengidentifikasi pola penggunaan internet berdasarkan gender dengan data Susenas di Jawa Timur. Hasilnya menunjukkan adanya perbedaan kombinasi aktivitas internet antara laki-laki dan perempuan, sehingga ARM mampu menemukan pola yang tidak terlihat melalui analisis deskriptif. Namun, penelitian tersebut terbatas pada satu provinsi dan karakteristik *individu*, sehingga belum menjelaskan pola akses internet berdasarkan karakteristik wilayah pada cakupan nasional [6].

Penelitian mengenai kesenjangan digital menunjukkan bahwa akses internet berkaitan dengan berbagai karakteristik *individu*, rumah tangga, dan wilayah. Kondisi tersebut menunjukkan bahwa akses internet bersifat *multidimensional* dan tidak cukup dijelaskan melalui satu indikator. Namun, penelitian sebelumnya lebih banyak menganalisis perbedaan atau hubungan karakteristik terhadap

<http://sistemasi.ftik.unisi.ac.id>

tingkat akses, sedangkan penelitian ini berfokus pada identifikasi kombinasi indikator yang muncul secara bersamaan melalui ARM [7].

Perkembangan ARM juga menunjukkan penerapannya pada data numerik dan *multidimensional*. Kajian Numerical ARM menunjukkan bahwa data numerik dapat dianalisis melalui transformasi dan diskretisasi untuk menemukan pola asosiasi. Pendekatan tersebut relevan dengan penelitian ini karena indikator akses internet BPS berbentuk data numerik yang perlu ditransformasikan sebelum digunakan dalam pembentukan aturan asosiasi. Penerapan ARM dalam bidang teknologi digital menunjukkan kemampuannya dalam menemukan hubungan beberapa karakteristik secara bersamaan. Pendekatan tersebut lebih sesuai untuk penelitian eksploratif yang bertujuan menemukan pola daripada menguji hubungan sebab-akibat atau melakukan prediksi. Namun, sebagian penelitian masih menggunakan *individu* atau aktivitas pengguna sebagai unit analisis dan belum banyak mengkaji pola indikator digital pada tingkat wilayah administratif [8].

Studi ARM mengenai internet di Indonesia masih sangat menitikberatkan pada sifat-sifat pengguna atau aktivitas situs *Web*. Wibowo et al. Menganalisis pola penggunaan berdasarkan gender, sedangkan penelitian *Web Usage Mining* berfokus pada perilaku pengunjung *Website*. Penelitian ARM pada data statistik BPS juga telah diterapkan pada bidang lain, tetapi belum secara khusus mengidentifikasi hubungan antarindikator akses internet pada seluruh provinsi Indonesia.

Berdasarkan penelitian terdahulu, terdapat kesenjangan pada penggunaan ARM untuk menganalisis indikator akses internet pada tingkat wilayah nasional. Penelitian sebelumnya lebih banyak berfokus pada *individu*, aktivitas digital, atau wilayah terbatas, sementara perkembangan ARM pada data numerik menunjukkan peluang untuk menganalisis data statistik melalui diskretisasi. Dengan demikian, masih diperlukan penelitian yang mengidentifikasi pola keterkaitan antarindikator akses internet menggunakan data statistik resmi pada seluruh provinsi Indonesia [9].

Artikel ini berfokus pada identifikasi pola keterkaitan antarindikator akses internet berdasarkan karakteristik provinsi menggunakan Association Rule Mining (ARM) dengan algoritma Apriori. Data berasal dari publikasi BPS dan dianalisis menggunakan *support*, *confidence*, dan *lift*. Berbeda dengan penelitian terdahulu yang berfokus pada individu atau aktivitas *Website*, penelitian ini menggunakan provinsi sebagai unit analisis dan mencakup seluruh Indonesia. Penelitian ini diarahkan untuk menemukan kombinasi indikator yang muncul secara bersamaan, sehingga dapat memberikan informasi mengenai karakteristik digital antarwilayah dan mendukung pemahaman mengenai pemerataan akses internet di Indonesia.

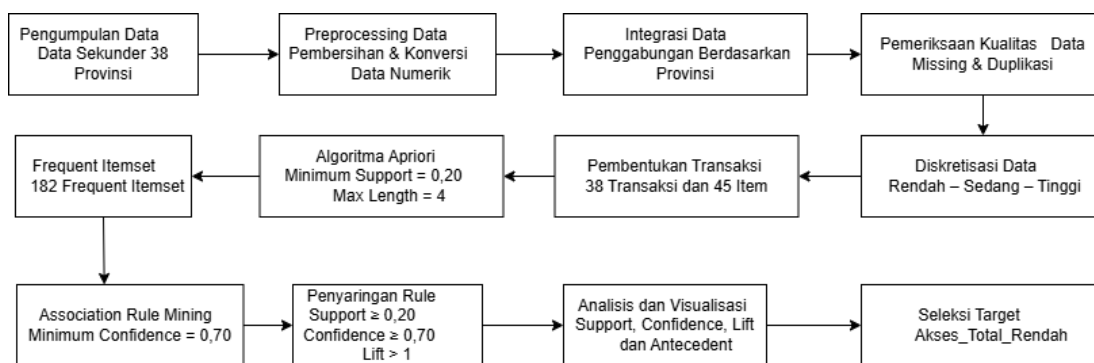
3 Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan memanfaatkan data sekunder yang menggambarkan karakteristik penduduk dan akses internet pada tingkat provinsi di Indonesia. Unit analisis yang digunakan adalah 38 provinsi, dengan data yang mencakup karakteristik pendidikan, kelompok umur, status sekolah, akses internet, dan kondisi rumah tangga. Data dari beberapa kelompok indikator diintegrasikan berdasarkan nama provinsi sehingga menghasilkan dataset yang terdiri atas 38 provinsi dan 21 variabel. Metode analisis yang digunakan adalah Association Rule Mining (ARM) dengan algoritma Apriori untuk mengidentifikasi pola keterkaitan antara karakteristik penduduk dan kondisi akses internet berdasarkan kemunculan beberapa atribut secara bersamaan pada suatu provinsi. Secara umum, penelitian dilakukan melalui beberapa tahapan, yaitu preprocessing dan integrasi data, pemeriksaan kualitas data, diskretisasi data numerik, pembentukan transaksi, penerapan algoritma Apriori, pembentukan Association Rule, serta analisis dan interpretasi pola yang diperoleh [10][11].

3.1 Tahapan Penelitian

Tahapan dalam studi ini mencakup pengumpulan informasi, pengolahan awal data, penggabungan data, evaluasi mutu data, pengubahan data menjadi diskrit, pembentukan transaksi, penerapan ARM, penciptaan aturan asosiasi, penyaringan aturan, serta analisis dan visualisasi temuan penelitian.

Gambar 1
Alur Tahapan penelitian



3.2 Pengumpulan Data

Studi ini memanfaatkan data sekunder dari publikasi Badan Pusat Statistik (BPS) yang menyajikan informasi tentang karakteristik populasi dan akses internet per provinsi di Indonesia. Data yang diambil mencakup berbagai kelompok variabel, yaitu tingkat pendidikan, kategori usia, status pendidikan, akses internet, dan keadaan rumah tangga. Data tersebut diperoleh dari tabel statistik BPS yang menyampaikan informasi akses internet sesuai dengan provinsi, pendidikan, kategori usia, partisipasi dalam pendidikan, serta klasifikasi daerah. Data dari beberapa kelompok indikator tersebut selanjutnya diterapkan sebagai dasar untuk membentuk dataset penelitian dan diorganisasikan berdasarkan Provinsi, sehingga diperoleh total 38 provinsi dengan 21 variabel. Berikut adalah variabel yang dipakai dalam penelitian ini. Adapun variabel yang digunakan dapat dilihat pada Tabel 1[12][13].

Tabel 1 Variabel yang digunakan

Variabel	Keterangan
Tidak_Punya	Penduduk yang tidak memiliki akses/kondisi sesuai indikator
SD	Penduduk dengan tingkat pendidikan SD
SMP	Penduduk dengan tingkat pendidikan SMP
SMA	Penduduk dengan tingkat pendidikan SMA
D1-D3	Penduduk dengan tingkat pendidikan D1–D3
S1_Ke_Atas	Penduduk dengan tingkat pendidikan S1 ke atas
Umur 5-12	Penduduk kelompok umur 5-12 tahun
Umur 13-15	Penduduk kelompok umur 13-15 tahun
Umur 16-18	Penduduk kelompok umur 16-18 tahun
Umur 19-24	Penduduk kelompok umur 19-24 tahun
Umur 25 Ke-Aatas	Penduduk kelompok umur 25 tahun ke atas
Tidak Belum Sekolah	Penduduk yang tidak/belum bersekolah
Masih Bersekolah	Penduduk yang masih bersekolah
Tidak Sekolah Lagi	Penduduk yang tidak bersekolah lagi
Akses Perkotaan	Akses internet di wilayah perkotaan
Akses Pedesaan	Akses internet di wilayah pedesaan
Akses Total	Akses internet total
RT Perkotaan	Akses internet rumah tangga di wilayah perkotaan
RT Pedesaan	Akses internet rumah tangga di wilayah pedesaan
RT_Total	Akses internet rumah tangga total

3.3 Preprocessing Data

Langkah-langkah pra-pemrosesan dilaksanakan untuk menjamin mutu dan kecocokan data sebelum diimplementasikan dalam proses Association Rule Mining. Proses pra-pemrosesan dalam penelitian ini mencakup pembersihan data, integrasi data, pemeriksaan nilai yang hilang dan duplikat, serta konversi data menjadi format numerik.

a. Pembersihan Data

Data yang berasal dari beberapa lembar pertama-tama diseleksi berdasarkan variabel-variabel yang relevan dalam penelitian ini. Nama-nama provinsi dibersihkan dengan menghapus spasi yang tidak perlu dan menyamakan penulisan nama provinsi. Data agregat untuk tingkat nasional dengan label Indonesia dihapus karena fokus analisis penelitian adalah pada tingkat provinsi. Selanjutnya, variabel yang digunakan dalam analisis diubah menjadi bentuk numerik agar dapat diterapkan pada tahap pengolahan selanjutnya.

b. Penggabungan Data

Data dari beberapa kategori indikator seperti pendidikan, umur, status pendidikan, akses internet, dan kondisi rumah ditransformasikan dengan cara menggabungkan berdasarkan atribut Provinsi melalui proses merge. Hasil penggabungan ini menciptakan dataset yang berisi 38 provinsi dan 21 variabel, menjadikan semua informasi yang dibutuhkan untuk analisis terkumpul dalam satu dataset.

c. Evaluasi Nilai Hilang dan Duplikasi

Setelah penggabungan data, dilakukan evaluasi terhadap nilai yang hilang dan duplikat untuk memastikan bahwa setiap provinsi berfungsi sebagai unit analisis. Hasil evaluasi mengungkapkan keberadaan 38 provinsi yang berbeda tanpa adanya duplikasi. Terdapat satu nilai yang hilang dalam variabel Akses_Pedesaan dan satu nilai yang hilang dalam variabel RT_Pedesaan. Nilai yang hilang tersebut tidak dihitung sebagai item saat data diubah menjadi transaksi, sehingga tidak berdampak pada penghapusan seluruh data provinsi.

Oleh karena itu, proses pra-pengolahan dalam penelitian ini tidak melibatkan penanganan pencila menggunakan IQR ataupun normalisasi data, karena kedua metode tersebut tidak diterapkan dalam pengolahan data. Setelah fase pra-pengolahan selesai, data berlanjut ke tahap transformasi atau diskretisasi, di mana data numerik diubah menjadi kategori Rendah, Sedang, dan Tinggi sebagai bentuk data yang digunakan dalam proses ARM.

3.4 ARM

Association Rule Mining (ARM) adalah teknik yang digunakan untuk mengidentifikasi pola hubungan atau keterkaitan antaritem dalam sebuah dataset. Dalam studi ini, ARM diterapkan untuk mendeteksi pola keterkaitan karakteristik masyarakat menurut provinsi dengan kategori akses internet. Algoritma yang diterapkan untuk proses penemuan pola adalah Apriori, yang beroperasi dengan menemukan frequent itemset berdasarkan minimum *support*, lalu membentuk Association Rule sesuai dengan minimum *confidence*. Rule yang dihasilkan kemudian dievaluasi menggunakan *support*, *confidence*, dan lift untuk menilai kekuatan pola keterkaitan yang berhasil ditemukan. Setelah penjelasan ini, kita akan membahas 3.5 Transformasi Data, sebab data numerik Anda harus diubah menjadi kategori sebelum bisa digunakan sebagai item dalam ARM. Proses ini sudah digambarkan dalam kode Anda melalui diskretisasi menjadi kategori Rendah, Sedang, dan Tinggi[14][15].

3.5 Transformasi Data

Proses transformasi data dilaksanakan untuk merubah data angka menjadi kategori yang relevan untuk kegiatan ARM. Metode quantile discretization dipakai untuk membagi setiap variabel menjadi tiga kategori, yaitu Rendah, Sedang, dan Tinggi. Diskretisasi diterapkan pada variabel karakteristik populasi dan variabel target Akses_Total dengan menggunakan fungsi *qcut*, yang membagi nilai menjadi tiga kuantil.

Pengelompokan kategori ini dilakukan berdasarkan distribusi nilai masing-masing variabel, sehingga kategori Rendah merepresentasikan nilai dengan posisi relatif lebih rendah, kategori Sedang

<http://sistemasi.ftik.unisi.ac.id>

merepresentasikan nilai di posisi tengah, dan kategori Tinggi mencerminkan nilai dengan posisi relatif lebih tinggi. Hasil dari proses transformasi pada variabel Akses_Total menunjukkan 13 provinsi berada dalam kategori Rendah, 12 provinsi dalam kategori Sedang, dan 13 provinsi berada dalam kategori Tinggi.

Secara teoritis, pembagian kuantil menjadi tiga kategori dapat digambarkan sebagai:

Nilai kuantil 33,33% diterapkan sebagai batas minimum dalam pengelompokan data, seperti yang terlihat pada Persamaan (1).

$$Q_{33} = \text{KUANTIL KE} - 33,33\% \quad (1)$$

Nilai kuantil ke-66,67% digunakan sebagai limit maksimum dalam pengelompokan data, seperti yang terlihat pada Persamaan (2).

$$Q_{67} = \text{KUANTIL KE} - 66,67\% \quad (2)$$

Dengan pembagian kategori sebagai berikut :

Kategori Rendah diidentifikasi berdasarkan nilai data yang nilainya sama dengan atau lebih kecil dibandingkan kuantil ke-33,33%, seperti yang dijelaskan dalam Persamaan (3).

$$\text{Rendah} \leq Q_{33} \quad (3)$$

Kategori ditentukan berdasarkan nilai-nilai data yang terletak antara kuantil ke-33,33% dan kuantil ke-66,67%, seperti yang terlihat pada Persamaan (4).

$$Q_{33} < \text{SEDANG} \leq Q_{67} \quad (4)$$

Kategori Tinggi ditetapkan berdasarkan nilai-nilai data yang melebihi kuantil ke-66,67%, seperti yang terlihat pada Persamaan (5).

$$\text{TINGGI} > Q_{67} \quad (5)$$

3.6 ARM

Tahap dari ARM dilaksanakan setelah data numerik diubah menjadi kategori Rendah, Sedang, dan Tinggi, serta disusun menjadi transaksi. ARM berfungsi untuk mengidentifikasi *frequent itemset*, yaitu gabungan item yang muncul dengan frekuensi memenuhi ambang batas minimum *support* yang telah ditentukan. Dalam penelitian ini, setiap provinsi dianggap sebagai satu transaksi, sehingga terdapat 38 transaksi dengan 45 item yang diperoleh dari transformasi data.

Implementasi ARM menggunakan nilai minimum *support* sebesar 0,20. Angka ini digunakan untuk menentukan kombinasi item yang cukup sering muncul di seluruh transaksi. Dengan total 38 provinsi, nilai *support* 0,20 menunjukkan bahwa kombinasi item harus muncul di sekitar 8 provinsi atau lebih agar memenuhi ambang batas minimum *support*. Di samping itu, panjang kombinasi item dibatasi maksimal hingga 4 item ($\text{max_len} = 4$) untuk mengurangi kompleksitas kombinasi yang dihasilkan dan memastikan pola yang diperoleh tetap bisa dianalisis.

Nilai *support* berfungsi untuk mengukur proporsi transaksi yang mencakup kombinasi item tertentu. *Support* dari suatu *itemset* dapat dirumuskan sebagai:

Nilai *support* digunakan untuk menilai rasio transaksi yang mencakup suatu itemset X, seperti yang terlihat pada Persamaan (6).

$$\text{Support}(x) = \frac{\text{jumlah Transaksi yang mengandung } X}{\text{jumlah seluruh transaksi}} \quad (6)$$

Dengan cara ini, semakin tinggi nilai *support* dari suatu itemset, semakin sering kombinasi tersebut muncul dalam data yang diteliti. Melalui penggunaan algoritma Apriori dengan minimum *support* 0,20, ditemukan 182 frequent itemsets yang memenuhi kriteria minimum *support*. Frequent itemsets tersebut selanjutnya dimanfaatkan sebagai dasar untuk menyusun Association Rule pada tahap berikutnya.

3.7 Pembentukan Association Rule

Proses pembentukan Association Rule dimulai setelah diperoleh 182 frequent itemsets dari algoritma Apriori. Frequent itemsets tersebut dimanfaatkan untuk menghubungkan *antecedent* dan *consequent*. *Antecedent* adalah satu atau lebih item yang berperan sebagai kondisi dalam suatu rule, sedangkan *consequent* adalah item yang muncul sebagai hasil dari rule tersebut. Dalam studi ini, pembentukan Association Rule dilakukan dengan menetapkan minimum *confidence* sebesar 0,70 agar rule yang dihasilkan memiliki tingkat kepercayaan yang memadai [16][17].

Nilai *confidence* digunakan untuk menilai seberapa besar kemungkinan terjadinya *consequent* Y ketika ada *Antecedent* X dalam sebuah transaksi, seperti yang dijelaskan dalam Persamaan (7)

$$Confidence(X \rightarrow Y) = \frac{Support(X \cup Y)}{Support(X)} \quad (7)$$

Keterangan

- X = *Antecedent* ;
- Y = *consequent* ;
- *Support* (X ∪ Y) = *support* kombinasi *Antecedent consequent* ;
- *Support* (X) = *Antecedent consequent* ;

Sebagai contoh,rule:

Umur_19_24 → tinggi Akses_Total_Rendah

menunjukkan bahwa kategori Umur_19_24_Tinggi berfungsi sebagai *Antecedent*, sedangkan Akses_Total_Rendah berperan sebagai *consequent*. Aturan ini kemudian dievaluasi berdasarkan *support*, *confidence*, dan *lift* untuk mengetahui seberapa sering munculnya serta seberapa kuat keterkaitannya.

Setelah aturan asosiasi terbentuk, penyaringan dilakukan dengan syarat nilai *support* ≥ 0,20, *confidence* ≥ 0,70, dan *lift* > 1. Tujuan dari proses penyaringan ini adalah untuk mendapatkan aturan dengan frekuensi kemunculan yang memadai, tingkat *confidence* yang tinggi, dan menunjukkan hubungan positif antara *Antecedent* dan *consequent*. Aturan yang memenuhi syarat kemudian diurutkan berdasarkan *lift*, *confidence*, dan *support* untuk memudahkan pemilihan dan interpretasi aturan.

Selain penyaringan berdasarkan tiga parameter tersebut, juga dilakukan pemeriksaan terhadap *consequent* untuk memastikan bahwa aturan yang dipakai relevan dengan tujuan penelitian, yaitu menganalisis pola karakteristik provinsi yang berkaitan dengan akses internet. Pada tahap akhir, aturan yang digunakan berfokus pada Akses_Total_Rendah sebagai *consequent*, sehingga dihasilkan 25 aturan asosiasi yang menjadi dasar untuk analisis lebih lanjut. Dengan demikian, aturan asosiasi yang diperoleh tidak hanya ditentukan oleh nilai statistiknya, tetapi juga berdasarkan sejauh mana kesesuaiannya dengan tujuan dari penelitian tersebut.

3.8 Evaluasi dan Penyaringan Association Rule

Evaluasi Association Rule dilakukan dengan memanfaatkan tiga metrik, yaitu *support*, *confidence*, dan *lift*. *Support* menggambarkan proporsi transaksi yang mengandung kombinasi item, *confidence* menunjukkan tingkat kepastian *consequent* ketika *antecedent* muncul, sedangkan *lift* mencerminkan kekuatan hubungan antara *antecedent* dan *consequent*. Rumus yang digunakan adalah:

Nilai *support* untuk suatu aturan X mengarah ke Y dihitung dengan cara membandingkan jumlah transaksi yang mencakup kombinasi X dan Y dengan total seluruh transaksi, seperti yang terlihat pada Persamaan (8).

$$Support(X \rightarrow Y) = \frac{count(X \cup Y)}{N} \quad (8)$$

Nilai *confidence* untuk aturan X mengarah ke Y dihitung dengan membandingkan *support* dari kombinasi X dan Y dengan *support* dari *antecedent* X, seperti yang terlihat dalam Persamaan (9).

$$Confidence(X \rightarrow Y) = \frac{Support(X \cup Y)}{Support(X)} \quad (9)$$

Nilai *lift* berfungsi untuk menilai seberapa kuat hubungan antara *antecedent* X dan *consequent* Y dengan membandingkan nilai *confidence* terhadap dukungan *consequent*, seperti yang terlihat pada Persamaan (10).

$$Lift(X \rightarrow Y) = \frac{Confidence(X \rightarrow Y)}{Support(Y)} \quad (10)$$

Rule yang diterapkan dalam penelitian memenuhi syarat *support* ≥ 0,20, *confidence* ≥ 0,70, dan *lift* > 1. Nilai *lift* > 1 menunjukkan adanya asosiasi positif antara *antecedent* dan *consequent*. Setelah

melalui proses penyaringan dan pemilihan rule sesuai dengan tujuan penelitian, didapatkan 25 Association Rule dengan Akses_Total_Rendah sebagai fokus analisis[18][19].

3.9 Visualisasi Association Rule

Visualisasi Association Rule diterapkan dengan menggunakan grafik sebar untuk menggambarkan hubungan antara nilai *confidence* dan lift, sedangkan nilai *support* diwakili oleh ukuran titik. Sumbu horizontal menggambarkan nilai *confidence*, sumbu vertikal menunjukkan nilai lift, dan ukuran titik mewakili nilai *support* dari setiap rule. Garis acuan pada lift = 1 berfungsi sebagai batas untuk mengidentifikasi rule yang menunjukkan asosiasi positif. Rule yang terletak di atas garis ini memiliki nilai lift lebih dari 1, yang menandakan adanya asosiasi positif antara *antecedent* dan *consequent*. Visualisasi ini berguna untuk mempermudah identifikasi rule berdasarkan kekuatan asosiasi serta tingkat *confidence*-nya[20].

3.10 Interpretasi Association Rule

Interpretasi dilakukan terhadap 25 Association Rule yang telah memenuhi syarat $support \geq 0,20$, $confidence \geq 0,70$, dan $lift > 1$, dengan Akses_Total_Rendah sebagai *consequent*. Proses interpretasi mempertimbangkan kombinasi karakteristik pada *antecedent* serta nilai *support*, *confidence*, dan lift dari setiap rule. Rule yang memiliki nilai lift lebih tinggi menunjukkan asosiasi yang lebih kuat, sedangkan *confidence* yang tinggi menggambarkan tingkat kepastian *consequent* ketika *antecedent* muncul. Di samping itu, frekuensi kemunculan karakteristik dalam *antecedent* digunakan untuk mengidentifikasi karakteristik yang paling umum muncul dalam pola akses internet rendah. Dari hasil pemrosesan, Tidak_Punya_Rendah dan Umur_5_12_Rendah masing-masing muncul dalam 8 rule, sementara Umur_19_24_Tinggi muncul dalam 5 rule. Hasil interpretasi digunakan untuk menjelaskan pola hubungan karakteristik penduduk di tingkat provinsi yang berhubungan dengan kategori Akses_Total_Rendah. Interpretasi ini bersifat asosiasi, bukan hubungan sebab-akibat, sehingga rule digunakan untuk mempresentasikan pola keterkaitan yang terungkap dalam data penelitian.

4. Hasil Dan Pembahasan

4.1 Hasil Pengolahan dan Transformasi Data

Hasil dari tahap preprocessing mengindikasikan bahwa informasi dari berbagai indikator berhasil disatukan berdasarkan variabel Provinsi. Setelah melakukan langkah-langkah pembersihan data serta menghilangkan data agregat nasional, didapatkan 38 provinsi dan 21 variabel yang menjadi komponen dataset penelitian. Analisis data menunjukkan adanya 38 provinsi yang berbeda dan tidak terdapat provinsi yang terduplikasi. Dalam pemeriksaan untuk nilai yang hilang, teridentifikasi satu nilai kosong pada variabel Akses_Pedesaan dan RT_Pedesaan. Adapun hasil pengolahan data dapat dilihat pada Tabel 2.

Tabel 2 Hasil pengolahan awal data

Komponen	Hasil
Jumlah Provinsi	38
Jumlah Variabel	21
Provinsi duplikat	0
Missing value Akses Pedesaan	1
Missing value RT Pedesaan	1

Setelah tahap pra-pemrosesan, data angka diubah menggunakan metode diskretisasi kuantil menjadi tiga klasifikasi, yaitu Rendah, Sedang, dan Tinggi. Perubahan ini dilakukan pada variabel demografis penduduk serta Akses Total yang berperan sebagai variabel target. Hasil dari kategorisasi Akses Total mengungkapkan bahwa terdapat 13 provinsi yang termasuk dalam kategori Rendah, 12 provinsi berada dalam kategori Sedang, dan 13 provinsi dikategorikan sebagai Tinggi. Adapun Distribusi Kategori dapat dilihat pada Tabel 3.

Tabel 3 Distribusi kategori akses total

Kategori	Jumlah Provinsi	Persentase
http://sistemasi.ftik.unisi.ac.id		

Rendah	13	34,21%
Sedang	12	31,58%
Tinggi	13	34,21%
Total	38	100%

Distribusi ini menunjukkan bahwa distribusi kategori Akses_Total cukup merata. Situasi ini memungkinkan ketiga kategori tetap digunakan dalam pencarian pola, walaupun pada proses penyaringan akhir hanya kategori tertentu yang menghasilkan aturan yang memenuhi kriteria penelitian. Proses tersebut menghasilkan 38 transaksi dan 45 item yang selanjutnya digunakan dalam penerapan ARM. Adapun Hasil pembentukan transaksi digunakan dapat dilihat pada Tabel 4.

Tabel 4 Hasil pembentukan transaksi

komponen	Hasil
Jumlah transaksi	38
Jumlah item	45
Unit transaksi	Provinsi
Bentuk data	Item kategorikal

Dengan cara ini, pengolahan awal telah menghasilkan data yang terorganisir untuk proses Association Rule Mining, di mana masing-masing provinsi digambarkan sebagai satu transaksi yang terdiri dari beberapa item kategorikal. Data tersebut kemudian dimanfaatkan pada tahap selanjutnya untuk mencari frequent itemset dengan menggunakan algoritma Apriori.

4.2 Hasil Penerapan ARM

Penerapan algoritma Apriori dilaksanakan pada data transaksi yang telah dibuat pada fase sebelumnya. Dalam tahap ini, setiap provinsi diwakili sebagai satu transaksi dengan jumlah 38 transaksi dan 45 item. Algoritma Apriori digunakan dengan minimum *support* sebesar 0,20 dan panjang kombinasi item maksimum 4 item.

Nilai minimum *support* sebesar 0,20 menunjukkan bahwa kombinasi item harus muncul dalam setidaknya 20% dari total transaksi agar dapat diakui sebagai frequent itemset. Dengan total transaksi sebanyak 38, batas ini setara dengan sekitar 8 provinsi. Dari segi matematis, *support* mengukur proporsi transaksi yang mengandung suatu itemset.

Nilai sokongan dipakai untuk menilai peratusan transaksi yang mengandung itemset X berbanding dengan keseluruhan transaksi, seperti yang dinyatakan dalam Persamaan (11).

$$Support(X) = \frac{count(X)}{N} \quad (11)$$

dengan (Count(X) merupakan jumlah transaksi yang mengandung *itemset* (X), sedangkan (N) merupakan jumlah seluruh transaksi. Adapun hasil penerapan yang digunakan dapat dilihat pada Tabel 5.

Tabel 5. Parameter dan hasil penerapan association rule mining (ARM)

Parameter	Nilai
Jumlah transaksi	38
Jumlah item	45
Minimum <i>support</i>	0,20
Panjang <i>itemset</i> maksimum	4
Jumlah <i>frequent itemsets</i>	182

Hasil penerapan Apriori menunjukkan bahwa terdapat 182 kombinasi item yang memenuhi nilai minimum *support* 0,20. Frequent itemsets tersebut selanjutnya menjadi dasar dalam pembentukan Association Rule dengan menggunakan minimum *confidence* sebesar 0,70.

<http://sistemasi.ftik.unisi.ac.id>

4.3 Hasil Evaluasi dan Seleksi Association Rule

Berdasarkan 182 frequent itemsets yang didapat melalui algoritma Apriori, langkah selanjutnya adalah membangun dan menilai Association Rule dengan minimum *confidence* sebesar 0,70. Rule tersebut dinilai menurut *support*, *confidence*, dan *lift*. Kriteria yang diterapkan untuk mendapatkan rule yang sesuai dengan tujuan penelitian mencakup $support \geq 0,20$, $confidence \geq 0,70$, dan $lift > 1$. Rule yang memenuhi ketiga kriteria ini kemudian diseleksi berdasarkan *consequent* untuk memastikan bahwa pola yang ditelaah berkaitan dengan akses internet yang menjadi fokus. Adapun kriteria evaluasi yang digunakan dapat dilihat pada Tabel 6

Tabel 6 Kriteria evaluasi association rule

Parameter	kriteria
Minimum <i>support</i>	$\geq 0,20$
Minimum <i>confidence</i>	$\geq 0,70$
Minimum <i>lift</i>	> 1
<i>Consequent</i>	Akses Total Rendah
Jumlah <i>rule</i> final	25

Proses seleksi akhir menghasilkan 25 Association Rule dengan Akses_Total_Rendah sebagai *consequent*. Rule tersebut kemudian diurutkan berdasarkan nilai *lift*, *confidence*, dan *support* untuk menemukan pola yang memiliki keterkaitan lebih kuat. Dengan cara ini, 25 rule tersebut menjadi fondasi untuk analisis lebih lanjut mengenai karakteristik provinsi yang berhubungan dengan kondisi akses internet yang rendah.

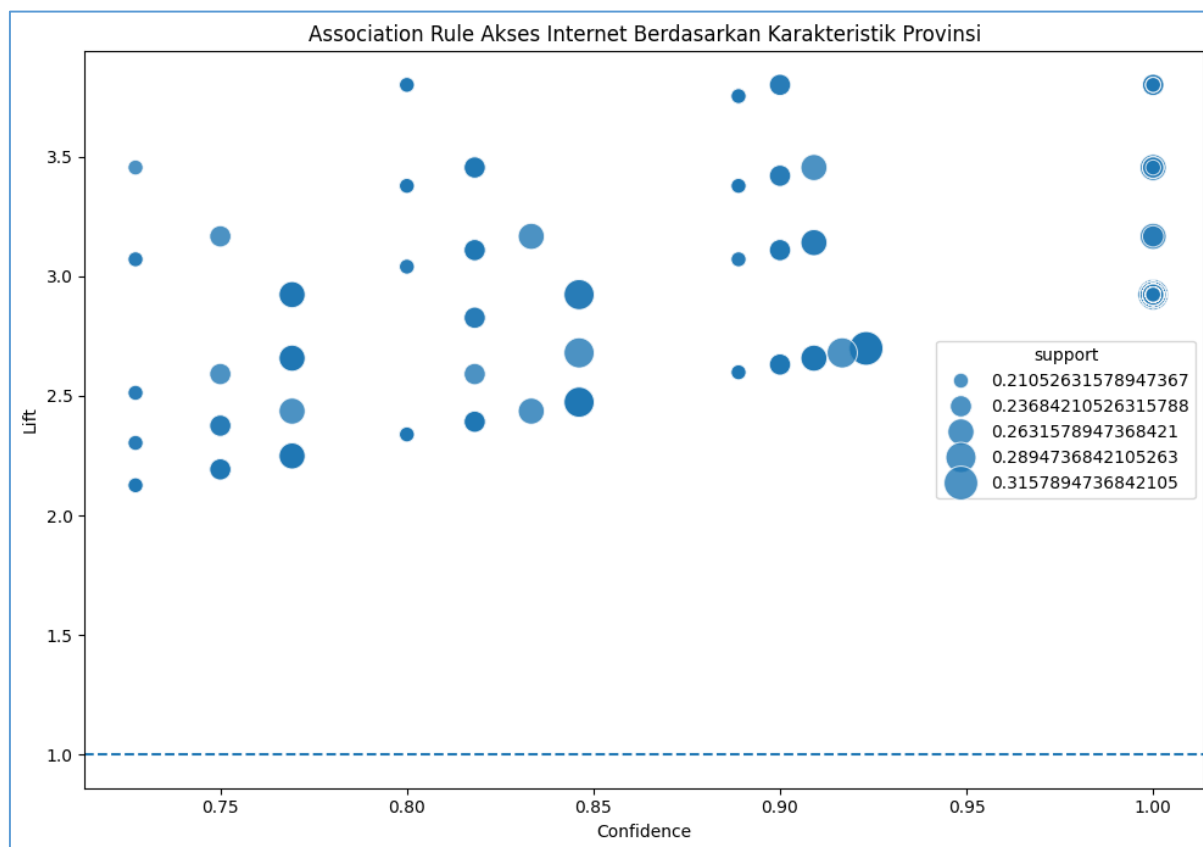
Nilai *support*, *confidence*, dan *lift* digunakan untuk menilai mutu masing-masing aturan. *Support* mencerminkan seberapa sering kombinasi *Antecedent* dan *consequent* muncul dalam dataset, *confidence* mencerminkan seberapa besar kepastian *consequent* ketika *Antecedent* hadir, sementara *lift* menunjukkan kekuatan hubungan antara keduanya. Aturan dengan nilai *lift* yang lebih dari satu menandakan adanya hubungan positif. Setelah melakukan evaluasi dan pemilihan, didapatkan 25 aturan yang sesuai dengan kriteria penelitian dan selanjutnya dipakai dalam analisis karakteristik *Antecedent* serta pembahasan pola penggunaan internet.

4.4 Visualisasi dan Analisis Association Rule

Hasil Association Rule yang telah memenuhi syarat selanjutnya ditampilkan menggunakan scatter plot. Dalam visualisasi ini, *confidence* dijadikan sumbu X, *lift* menjadi sumbu Y, dan *support* diwakili oleh ukuran titik. Garis horizontal pada $lift = 1$ berfungsi sebagai batas untuk menandakan rule yang memiliki keterkaitan positif. Rule yang terletak di atas garis tersebut memiliki nilai *lift* lebih dari 1, yang menunjukkan adanya asosiasi positif antara *antecedent* dan *consequent*. Visualisasi tersebut bertujuan untuk mempermudah analisis distribusi 25 rule berdasarkan nilai *confidence*, *lift*, dan *support*.

Visualisasi tersebut bertujuan untuk mempermudah analisis distribusi 25 aturan berdasarkan nilai *confidence*, *lift*, dan *support*. Aturan yang menduduki posisi lebih tinggi memperlihatkan nilai *lift* yang lebih besar, sedangkan posisi lebih ke kanan menandakan tingkat *confidence* yang lebih tinggi. Ukuran titik yang lebih besar menggambarkan nilai *support* yang lebih tinggi. Dengan demikian, visualisasi ini dapat berfungsi sebagai alat bantu dalam mengidentifikasi aturan yang memiliki tingkat *confidence* dan kekuatan asosiasi yang cukup tinggi.

Berdasarkan Gambar 2 di bawah, setiap titik melambangkan satu Association Rule yang diperoleh dari analisis. Sumbu horizontal menggambarkan nilai *confidence*, sedangkan sumbu vertikal



Gambar 2. Visualisasi association rule berdasarkan confidence, lift, dan support

merepresentasikan nilai lift. Ukuran titik mencerminkan nilai *support*, sehingga semakin besar ukuran titik tersebut, semakin tinggi nilai *support*-nya. Garis putus-putus pada lift = 1 berfungsi sebagai batas untuk menunjukkan adanya asosiasi positif. Semua Association Rule terletak di atas garis ini, sehingga semua rule yang ditampilkan memiliki lift lebih besar dari 1 dan menunjukkan asosiasi positif antara *antecedent* dan *consequent*.

4.5 Pembahasan

4.5.1 Pola Karakteristik yang Berkaitan dengan Akses Internet Rendah

Berdasar pada 25 aturan asosiasi yang diperoleh, terdapat sejumlah karakteristik yang muncul paling sering dalam bagian *Antecedent* dengan Akses_Total_Rendah sebagai *consequent*. Karakteristik Tidak_Punya_Rendah dan Umur_5_12_Rendah muncul masing-masing dalam 8 aturan, sementara Umur_19_24_Tinggi muncul di 5 aturan. Selanjutnya, SD_Rendah dan S1_Ke_Atas_Tinggi masing-masing terlihat dalam 3 aturan, sementara Umur_16_18_Tinggi muncul dalam 1 aturan. Adapun frekuensi karakteristik digunakan dapat dilihat pada Tabel 7.

Tabel 7 Frekuensi karakteristik pada antecedent association rule

Karakteristik	Jumlah Rule
Tidak_Punya_Rendah	8
Umur_5_12_Rendah	8
Umur_19_24_Tinggi	5
SD_Rendah	3
S1_Ke_Atas_Tinggi	3
Umur_16_18_Tinggi	1

Berdasarkan Tabel 7, Tidak_Punya_Rendah dan Umur_5_12_Rendah adalah karakteristik yang paling sering muncul pada *Antecedent*, masing-masing terdapat di 8 aturan. Ini mengindikasikan bahwa kedua karakteristik ini paling sering ditemui bersamaan dengan pola Akses_Total_Rendah dalam aturan asosiasi yang dihasilkan. Di sisi lain, Umur_19_24_Tinggi juga menunjukkan frekuensi yang cukup tinggi dengan 5 aturan. Temuan ini menunjukkan pola hubungan berdasarkan kemunculan atribut, sehingga tidak bisa ditafsirkan sebagai hubungan sebab-akibat antara karakteristik tersebut dan rendahnya akses internet.

4.5.2 Rule dengan Kekuatan Asosiasi Tertinggi

Berdasarkan 25 aturan asosiasi yang telah diperoleh, analisis berikutnya dilakukan dengan memperhatikan nilai *support*, *confidence*, dan *lift* guna mengidentifikasi aturan dengan kekuatan asosiasi paling tinggi. Salah satu aturan yang menunjukkan nilai *lift* tertinggi adalah Tidak_Punya_Rendah + Umur_19_24_Tinggi → Akses_Total_Rendah dengan *support* mencapai 21,05%, *confidence* 100%, dan *lift* sebesar 2,923. Nilai *confidence* sebesar 100% mengindikasikan bahwa dalam data penelitian, semua transaksi yang memiliki kombinasi Tidak_Punya_Rendah dan Umur_19_24_Tinggi juga menunjukkan Akses_Total_Rendah. Nilai *lift* sebesar 2,923 menandakan adanya hubungan positif antara kemunculan Akses_Total_Rendah pada kombinasi tersebut. Adapun asosiasi tertinggi digunakan dapat dilihat pada Tabel 8.

Tabel 8 Rule dengan kekuatan asosiasi tertinggi

<i>Antecedent</i>	<i>Consequent</i>	<i>Support (%)</i>	<i>Confidence (%)</i>	<i>Lift</i>
Tidak_Punya_Rendah + Umur_19_24_Tinggi	Akses_Total_Rendah	21,05	100,00	2,923
Umur_19_24_Tinggi	Akses_Total_Rendah	28,95	84,62	2,473

Rule Umur_19_24_Tinggi → Akses_Total_Rendah juga menunjukkan asosiasi yang kuat dengan support sebesar 28,95%, confidence sebesar 84,62%, dan lift sebesar 2,473. Angka-angka ini mengindikasikan bahwa kategori Akses_Total_Rendah memiliki frekuensi kemunculan yang signifikan dalam transaksi yang melibatkan Umur_19_24_Tinggi. Jika dibandingkan dengan rule yang hanya memiliki satu antecedent, penambahan Tidak_Punya_Rendah sebagai antecedent memberikan lift yang lebih tinggi, yakni 2,923. Ini menunjukkan bahwa kombinasi kedua atribut tersebut memiliki asosiasi yang lebih kuat dengan Akses_Total_Rendah dibandingkan ketika menggunakan Umur_19_24_Tinggi secara terpisah.

Temuan ini sebaiknya dipahami sebagai suatu keterkaitan atau pola asosiasi, bukan sebagai hubungan sebab-akibat. Dengan demikian, aturan ini menunjukkan bahwa kombinasi dari karakteristik tersebut sering kali muncul bersamaan dengan kategori Akses_Total_Rendah dalam data yang dianalisis dari provinsi tersebut.

4.5.3 Analisis Kombinasi Karakteristik

Analisis berikutnya dilakukan untuk mengevaluasi jumlah karakteristik yang terdapat di bagian *Antecedent* dengan tujuan untuk memahami pola kombinasi karakteristik terkait dengan Akses_Total_Rendah. Aturan yang terbentuk dapat mengandung satu atau beberapa karakteristik pada *Antecedent*. Salah satu pola yang menunjukkan asosiasi yang kuat adalah Umur_19_24_Tinggi → Akses_Total_Rendah, sementara kombinasi dua karakteristik ditunjukkan oleh Tidak_Punya_Rendah + Umur_19_24_Tinggi → Akses_Total_Rendah. Aturan dengan dua *Antecedent* tersebut memiliki *confidence* mencapai 100% dan *lift* 2,923, yang lebih tinggi daripada aturan dengan satu *Antecedent* Umur_19_24_Tinggi → Akses_Total_Rendah yang memiliki *confidence* sebesar 84,62% dan *lift* 2,473. Adapun perbandingan jumlah *Antecedent* yang digunakan dapat dilihat pada Tabel 9.

Tabel 9 Perbandingan rule berdasarkan jumlah *antecedent*

Jumlah <i>Antecedent</i>	Association Rule	Support (%)	Confidence (%)	Lift
1	Umur_19_24_Tinggi → Akses_Total_Rendah	28,95	84,62	2,473
2	Tidak_Punya_Rendah + Umur_19_24_Tinggi → Akses_Total_Rendah	21,05	100,00	2,923

Berdasarkan Tabel 9, penambahan elemen pada *antecedent* menghasilkan rule dengan nilai *confidence* dan *lift* yang lebih tinggi dalam contoh rule yang diperiksa. Ini menunjukkan bahwa penggabungan elemen tertentu dapat menciptakan pola asosiasi yang lebih kuat dibandingkan hanya menggunakan satu elemen. Namun, hasil tersebut tetap mencerminkan hubungan asosiasi berdasarkan pola kemunculan data dan tidak dapat dipakai untuk menyimpulkan hubungan sebab-akibat.

4.5.4 Temuan Utama Penelitian

Berdasarkan temuan dari Association Rule Mining dengan algoritma Apriori, terdapat 25 Association Rule yang dihasilkan dengan Akses_Total_Rendah sebagai *consequent*. Analisis tersebut mengindikasikan bahwa Tidak_Punya_Rendah dan Umur_5_12_Rendah merupakan atribut yang paling sering muncul pada *antecedent*, masing-masing sebanyak 8 rule, sementara Umur_19_24_Tinggi tercatat pada 5 rule. Rule yang menggabungkan Tidak_Punya_Rendah + Umur_19_24_Tinggi → Akses_Total_Rendah menunjukkan salah satu asosiasi yang paling kuat dengan *support* 21,05%, *confidence* 100%, dan *lift* 2,923, sementara rule Umur_19_24_Tinggi → Akses_Total_Rendah memiliki *support* 28,95%, *confidence* 84,62%, dan *lift* 2,473. Adapun ringkasan utama Association Rule yang digunakan dapat dilihat pada Tabel 10.

Tabel 10 Ringkasan temuan utama association rule

Aspek	Temuan
Jumlah rule final	25 rule
<i>Consequent</i>	Akses_Total_Rendah
Karakteristik paling sering muncul	Tidak_Punya_Rendah dan Umur_5_12_Rendah
Frekuensi masing-masing	8 Rule
Rule dengan <i>lift</i> tertinggi yang dibahas	Tidak_Punya_Rendah + Umur_19_24_Tinggi → Akses_Total_Rendah
<i>Support</i>	21,05%
<i>Confidence</i>	100%
<i>Lift</i>	2,923

Secara keseluruhan, penelitian ini memperlihatkan adanya hubungan antara berbagai karakteristik populasi dan kategori akses internet yang terbatas di tingkat provinsi. Temuan ini mengindikasikan bahwa metode ARM dapat dimanfaatkan untuk menemukan kombinasi karakteristik yang sering terkait dengan Akses_Total_Rendah. Temuan tersebut mencerminkan pola asosiasi, sehingga tidak dikatakan bahwa karakteristik tertentu secara langsung menyebabkan rendahnya akses internet.

5. Kesimpulan

Berdasarkan hasil studi mengenai Analisis Pola Akses Internet Penduduk Indonesia Berdasarkan Karakteristik Provinsi menggunakan Association Rule Mining (ARM), dapat disimpulkan bahwa penerapan ARM melalui algoritma Apriori dapat mengidentifikasi pola keterkaitan antarindikator akses internet berdasarkan karakteristik provinsi di Indonesia. Data yang dianalisis meliputi 38

provinsi dengan 21 variabel, yang diubah menjadi 38 transaksi dan 45 item, lalu diproses dengan minimum *support* 0,20 dan maksimum panjang itemset 4 sehingga menghasilkan 182 frequent itemsets. Selanjutnya, dengan minimum *confidence* 0,70 dan kriteria *support* $\geq 0,20$, *confidence* $\geq 0,70$, dan *lift* > 1 , dihasilkan 25 Association Rule dengan Akses_Total_Rendah sebagai *consequent*. Hasil penelitian menunjukkan bahwa Tidak_Punya_Rendah dan Umur_5_12_Rendah merupakan karakteristik yang paling sering muncul pada *antecedent*, masing-masing terdapat pada 8 rule. Rule Tidak_Punya_Rendah + Umur_19_24_Tinggi $\rightarrow$ Akses_Total_Rendah menunjukkan salah satu asosiasi terkuat dengan *support* 21,05%, *confidence* 100%, dan *lift* 2,923. Temuan tersebut memberikan gambaran mengenai keterkaitan karakteristik akses internet antarprovinsi dan dapat menjadi informasi pendukung dalam memahami kondisi serta pemerataan pemanfaatan teknologi informasi di Indonesia. Hasil penelitian merupakan pola asosiasi, bukan hubungan sebab-akibat, sehingga karakteristik yang terdapat dalam rule tidak dapat dianggap sebagai faktor yang secara langsung memengaruhi rendahnya akses internet.

Referensi

- [1] P. Sustainability and P. E-mail, “Intersectional Digital Inequality In Indonesia * Agnieszka Choczyńska,” Vol. 6, No. 3, pp. 11–22, 2024.
- [2] A. D. Syarif, H. Rabbani, and A. Hakim, “Penerapan Metode PCA dan Apriori dalam Deteksi Pola Penggunaan Internet di kalangan Mahasiswa Sivitas Akademika dalam menunjang Aspek Edukasi maupun Personal . (Sangotra et al ., 2026). memperlancar transfer ilmu . (Lukianova et al ., 2024), interaksi ,” Vol. 5, No. November 2025, pp. 56–63, 2026.
- [3] H. S. Hopipah, J. H. Jaman, and U. Enri, “Web Usage Mining Guna Analisis Pola Akses Pengunjung Website dengan Association Rule,” SATIN - Sains dan Teknol. Inf., Vol. 7, No. 2, pp. 53–63, 2021, DOI: 10.33372/stn.v7i2.735.
- [4] F. Mansyur, F. Ridho, B. Majene, S. Barat, P. Studi, and K. Statistik, “Analisis Pola Permintaan Publikasi Data,” Vol. 07, No. 02, pp. 187–198, 2020.
- [5] P. H. Marpaung, N. Sahara, A. Rambe, and F. Suryani, “Penerapan Metode Association Rule menggunakan Algoritma Apriori pada Simulasi Prediksi Hujan Wilayah Sibolga,” J. Penelit. Multidisiplin Bangsa, Vol. 1, No. 3, pp. 117–122, 2024, DOI: 10.59837/jpnmb.v1i3.38.
- [6] A. H. Lubis, R. Muliono, N. Khairina, and N. Novita, “The Impact of k-Means on Association Rules Mining Algorithms Performance,” J. Comput. SCI. Inf. Technol. Telecommun. Eng., Vol. 5, No. 2, pp. 640–653, 2024, DOI: 10.30596/jcositte.v5i2.20907.
- [7] J. K. Online, A. M. Setyanti, S. A. Prestianawati, M. Fawwaz, and K. Lumpur, “Rural - Urban Disparities in Digital Technology Access in Indonesia : Evidence From Susenas,” Vol. 6, No. 2, pp. 186–199, 2025, DOI: 10.15408/jko.v6i2.46031.
- [8] R. M. T. Nursasongka, I. Fahrurrozi, U. Y. Oktiawati, U. Taufiq, U. Farooq, and G. Alfian, “Utilizing Association Rule Mining for Enhancing Sales Performance in Web-based Dashboard Application,” Indones. J. Electr. Eng. Comput. SCI., Vol. 36, No. 2, pp. 1105–1113, 2024, DOI: 10.11591/ijeecs.v36.i2.pp1105-1113.
- [9] C. A. Gemawaty and Y. I. Syuhardi, “Pemetaan Kesenjangan Digital Provinsi di Indonesia menggunakan Algoritma K-Means berdasarkan Indikator Teknologi Informasi dan Komunikasi Tahun 2024,” Vol. 4, No. 3, pp. 646–652, 2026.
- [10] D. Dwiputra, A. M. Widodo, H. Akbar, and G. Firmansyah, “Evaluating the Performance of Association Rules in Apriori and FP-Growth Algorithms: Market Basket Analysis to Discover Rules of Item Combinations,” J. World SCI., Vol. 2, No. 8, pp. 1229–1248, 2023, DOI: 10.58344/jws.v2i8.403.
- [11] W. Windary and A. H. Hasugian, “Data Mining of Rural Digital Technology Adoption Factors using Apriori Algorithm,” J. Inf. Syst. Technol. Res., Vol. 4, No. 3, pp. 224–233, 2025, DOI: 10.55537/jistr.v4i3.1324.
- [12] B. Almahenzar and A. W. Wijayanto, “Analisis Intensitas Hujan Provinsi Jawa Barat Tahun 2020 menggunakan Association Rule Apriori dan FP-Growth,” J. Syst. Comput. Eng. , Vol. 3, No. 2, pp. 258–271, 2022.
- [13] E. Primanda and S. I. Oktora, “Analisis Ketertinggalan Desa di Provinsi Papua dan Papua Barat menggunakan Association Rule Mining,” Statistika, Vol. 24, No. 1, pp. 102–114, 2024, DOI: 10.29313/statistika.v24i1.2302.

- [14] D. A. Putri, Y. Maulita, and H. Sembiring, “Penerapan Algoritma *Apriori* mengetahui Pola Tindakan Kriminal berdasarkan Wilayah (Studi Kasus : Polsek Sunggal),” *Bridg. J. Publ. Sist. Inf. dan Telekomun.*, Vol. 2, No. 4, pp. 32–46, 2024, DOI: 10.62951/bridge.v2i4.200.
- [15] E. R. D. Suradjiman, M. Soleh, and M. A. Rahman, “Sistem Rekomendasi Produk Obat pada Apotek Juang Jaya menggunakan Algoritma *Apriori*,” *J. Ilmu Pengetah. dan Teknol.*, Vol. 7, No. 1, pp. 33–46, 2023, DOI: 10.31543/jii.v7i1.229.
- [16] I. Khoiruzzidan and L. Iswari, “Student Grade Association Analysis with *Apriori* Algorithm,” *J. Sains, Nalar, dan Apl. Teknol. Inf.*, Vol. 3, No. 2, pp. 50–57, 2024, DOI: 10.20885/snati.v3.i2.31.
- [17] A. D. Arsita and I. Afriantoro, “Analisis Pola Penjualan di Toko Retail menggunakan Algoritma *Apriori* dengan *Rapid Minner* Abstrak Pendahuluan yang sangat besar dalam Persaingan Ekonomi . Dalam hal ini yaitu Persaingan Pengusaha- Baik . Pengusaha tentunya tidak ingin merasakan Penurunan h,” Vol. 10, No. 2, pp. 734–753, 2024.
- [18] M. M. Jannah, Y. A. Iskandar, and F. E. Zulvia, “Association Rules Mining for Designing Product Bundling Strategy (Study Case at a Cafe in Jakarta),” *J. Logistik Indones.*, Vol. 7, No. 2, pp. 162–171, 2023, DOI: 10.31334/logistik.v7i2.3466.
- [19] D. Anggraini, H. Harliana, and T. Prabowo, “Implementasi Association Rule melalui Algoritma *Apriori* pada Analisis Data Transaksi Penjualan,” *Ilk. J. Comput. SCI. Appl. Informatics*, Vol. 5, No. 3, pp. 200–208, 2023, DOI: 10.28926/ilkomnika.v5i3.589.
- [20] F. S. Saad, N. Azizah, and Z. Yani, “Penerapan Algoritma *Apriori* untuk Menentukan Pola Permintaan Bahan Makanan berdasarkan Data Transaksi,” *J. Pustaka AI (Pusat Akses Kaji. Teknol. Artif. Intell.*, Vol. 5, No. 2, pp. 388–397, 2023, DOI: 10.55382/jurnalpustakaai.v5i2.1181.