Jurnal Ilmiah Teknik Informatika dan Komunikasi
Volume. 6, Nomor. 3, November 2026
E-ISSN : 2827-7945 ; P-ISSN : 2827-8127; Hal. 410-425
DOI: https://doi.org/10.55606/juitik.v6i3.2786
Tersedia: https://journal.sinov.id/index.php/juitik

Implementasi *Data Mining* dengan Algoritma Apriori untuk Mendukung Penentuan Penerima Bantuan Keluarga Harapan (PKH)

Arnoldus Yansen Sangu Ate[1*], Herman Huki Ratu[2], Karolus Wulla Rato[3]
[1-3]Program Teknik Informatika, Universitas Stella Maris Sumba, Indonesia
**Penulis Korespondensi: arnoldusysate@gmail.com*

Abstract*. The distribution of the Family Hope Program (Program Keluarga Harapan/PKH) requires social and economic information that can be examined consistently to support the review of prospective beneficiary households. This study proposes the use of Data Mining with the Apriori algorithm to identify association patterns among household characteristics relevant to PKH. Based on the research proposal, the planned data include household income, number of dependents, employment status, housing condition, education level, asset ownership, and relevant PKH components. The analysis stages comprise data selection, validation, transformation of attributes into categorical items, frequent-itemset generation, and association-rule formation using support and confidence. The resulting association patterns are expected to provide additional information for examining beneficiary records and supporting transparent, consistent, and data-informed decisions. The proposal does not yet provide a verified dataset or empirical Apriori output; therefore, numerical results and final association rules must be reported after data collection and processing. Apriori is positioned as a decision-support tool and does not replace official eligibility verification or decisions by authorized institutions.*

Keywords*: Apriori Algorithm; Association Rules; Beneficiary Eligibility; Data Mining; Family Hope Program (PKH).*

Abstrak.. Penyaluran Program Keluarga Harapan (PKH) memerlukan informasi sosial dan ekonomi yang dapat ditelaah secara konsisten untuk mendukung peninjauan keluarga calon penerima bantuan. Penelitian ini mengusulkan penggunaan *Data Mining* dengan algoritma Apriori untuk menemukan pola asosiasi antar karakteristik rumah tangga yang relevan dengan PKH. Berdasarkan proposal, data yang direncanakan mencakup pendapatan keluarga, jumlah tanggungan, status pekerjaan, kondisi rumah, tingkat pendidikan, kepemilikan aset, serta komponen PKH yang relevan. Tahapan analisis meliputi seleksi data, validasi, transformasi atribut menjadi item kategoris, pembentukan frequent itemset, dan penyusunan aturan asosiasi menggunakan nilai support dan confidence. Pola asosiasi yang dihasilkan diharapkan menjadi informasi tambahan untuk meninjau data calon penerima dan mendukung pengambilan keputusan yang transparan, konsisten, serta berbasis data. Proposal belum memuat dataset terverifikasi maupun keluaran empiris Apriori; oleh karena itu, hasil numerik dan aturan asosiasi final harus dilaporkan setelah data dikumpulkan dan diproses. Apriori ditempatkan sebagai alat pendukung keputusan, bukan pengganti verifikasi kelayakan resmi oleh instansi berwenang.

Kata Kunci: Algoritma Apriori; Association Rule; *Data Mining*; Kelayakan Penerima; Program Keluarga Harapan (PKH).

1. LATAR BELAKANG

Kemiskinan berdampak pada pemenuhan kebutuhan dasar, pendidikan, kesehatan, dan kesejahteraan keluarga. Pemerintah Indonesia menjalankan Program Keluarga Harapan (PKH) sebagai salah satu program bantuan sosial bagi keluarga yang memenuhi ketentuan. Pelaksanaan program membutuhkan pengelolaan data yang baik agar informasi keluarga dapat ditelaah secara sistematis. Peninjauan data calon penerima bantuan memerlukan informasi sosial ekonomi, seperti pendapatan, jumlah tanggungan, kondisi rumah, pendidikan, dan pekerjaan. Darmawan et al. (2022) menunjukkan relevansi *Data Mining* untuk mengidentifikasi pola pada data kesejahteraan sosial. Karakteristik keluarga tersebut dapat dianalisis bersama untuk mengetahui kombinasi yang sering muncul dalam dataset. Atribut

Naskah Masuk: 14 September 2026; Revisi: 30 September 2026; Diterima: 26 Oktober 2026; Terbit: 01 November 2026

yang dikaji meliputi pendapatan keluarga, jumlah tanggungan, kondisi tempat tinggal, pendidikan, pekerjaan, kepemilikan aset, sanitasi, serta komponen keluarga yang relevan. Penggunaan atribut akhir harus disesuaikan dengan kolom yang benar-benar tersedia dan diizinkan untuk digunakan dari Dinas Sosial Kabupaten Sumba Barat Daya. Pemeriksaan atribut secara terpisah belum tentu memperlihatkan kombinasi karakteristik yang sering muncul bersamaan. Karena itu, penelitian ini menggunakan association rule mining untuk mengidentifikasi pola hubungan antaritem dalam dataset. *Data Mining* merupakan pendekatan untuk menemukan pola dan informasi dari kumpulan data. Salah satu tekniknya adalah association rule mining, yang mengidentifikasi kemunculan bersama antaritem. Penelitian Darmawan et al. (2022) dan Aprianti et al. (2017) menjadi rujukan penerapan pendekatan berbasis pola pada data kesejahteraan atau bantuan sosial. Algoritma Apriori menemukan kombinasi item yang sering muncul (frequent itemset) dengan menyeleksi itemset berdasarkan minimum support. Itemset yang memenuhi batas tersebut kemudian digunakan untuk membentuk aturan asosiasi, yang dinilai menggunakan confidence.

Nilai support dan confidence membantu menggambarkan frekuensi pola dan proporsi kemunculan konsekuen, tetapi tidak membuktikan hubungan sebab-akibat atau menentukan kelayakan bantuan. Dalam penelitian ini, algoritma Apriori diterapkan pada data keluarga yang berkaitan dengan PKH untuk mengidentifikasi pola asosiasi antaratribut. Hasilnya diharapkan menjadi informasi pendukung bagi peninjauan data oleh petugas, bukan penentu otomatis status penerima bantuan. Lokasi penelitian ini adalah Dinas Sosial Kabupaten Sumba Barat Daya. Data yang digunakan perlu berasal dari sumber yang sah dan sesuai izin instansi, kemudian diproses berdasarkan atribut yang tersedia untuk menghasilkan frequent itemset dan aturan asosiasi. Penelitian terdahulu memperlihatkan penggunaan Apriori pada data kesejahteraan dan bantuan sosial. Penelitian ini memusatkan perhatian pada penerapan Apriori terhadap dataset PKH di Dinas Sosial Kabupaten Sumba Barat Daya. Kontribusi yang diharapkan adalah penyajian pola asosiasi secara terstruktur untuk membantu eksplorasi dan peninjauan data. Hasil analisis diharapkan membantu petugas memahami kombinasi karakteristik keluarga yang muncul bersama. Keputusan akhir mengenai penerima PKH tetap mengikuti ketentuan resmi dan proses verifikasi oleh instansi yang berwenang. Penelitian ini bertujuan menerapkan algoritma Apriori untuk mengidentifikasi pola asosiasi antaratribut pada data PKH di Dinas Sosial Kabupaten Sumba Barat Daya serta menyajikan aturan asosiasi sebagai informasi pendukung peninjauan data.

2. KAJIAN TEORITIS

Data Mining

Data Mining merupakan proses analisis data untuk menemukan pola, hubungan, dan informasi yang bermanfaat. Dalam penelitian ini, *Data Mining* digunakan untuk mengeksplorasi keterkaitan karakteristik sosial ekonomi keluarga pada dataset PKH. Penerapan *Data Mining* pada data kesejahteraan sosial oleh Darmawan et al. (2022) serta kajian Aliyah et al. (2023) menjadi rujukan bagi penggunaan analisis pola dalam penelitian ini. Dataset dianalisis berdasarkan atribut yang benar-benar tersedia, seperti pendapatan, jumlah tanggungan, kondisi rumah, kepemilikan aset, sanitasi, dan komponen keluarga. Daftar final atribut harus disamakan dengan data aktual yang diperoleh dari instansi.

Algoritma Apriori

Algoritma Apriori merupakan salah satu metode *Data Mining* yang digunakan untuk menemukan frequent itemset atau kombinasi item yang sering muncul secara bersamaan dalam suatu kumpulan data. Algoritma ini bekerja secara bertahap dengan membentuk kandidat kombinasi item, menghitung frekuensi kemunculannya, kemudian mempertahankan kombinasi yang memenuhi nilai minimum support. Kombinasi tersebut selanjutnya dapat digunakan untuk membentuk aturan asosiasi berdasarkan nilai confidence (Darmawan et al., 2022). Secara umum, tahapan algoritma Apriori meliputi: Pembentukan itemset awal, yaitu mengidentifikasi atribut atau item yang terdapat dalam dataset. Perhitungan support, yaitu menghitung frekuensi kemunculan setiap item atau kombinasi item. Seleksi frequent itemset, yaitu memilih kombinasi item yang memenuhi minimum support. Pembentukan kandidat berikutnya, yaitu menggabungkan itemset yang memenuhi persyaratan untuk membentuk kombinasi dengan jumlah item lebih banyak. Pembentukan aturan asosiasi, yaitu menyusun hubungan antaritem dan menghitung confidence untuk menilai kekuatan hubungan tersebut. Dalam penelitian ini, algoritma Apriori digunakan untuk menemukan pola keterkaitan antara karakteristik keluarga, seperti penghasilan, jumlah tanggungan, kondisi tempat tinggal, dan komponen keluarga. Pola yang ditemukan dapat menjadi informasi pendukung dalam proses evaluasi data calon penerima bantuan PKH.

Program Keluarga Harapan (PKH)

Program Keluarga Harapan (PKH) merupakan program bantuan sosial bersyarat yang ditujukan kepada keluarga miskin dan rentan untuk mendukung peningkatan kesejahteraan. Program ini berkaitan dengan pemenuhan kebutuhan dasar, terutama dalam bidang pendidikan dan kesehatan. Pelaksanaannya diharapkan dapat membantu keluarga penerima manfaat meningkatkan kualitas hidup dan mengurangi risiko kemiskinan antargenerasi (Kementerian

Sosial Republik Indonesia, 2022). Berdasarkan proposal penelitian, atribut yang dianalisis untuk menemukan pola penerima bantuan meliputi penghasilan keluarga, jumlah tanggungan, kondisi rumah, luas bangunan, jenis lantai dan dinding rumah, kepemilikan aset, sumber air bersih, sanitasi, serta keberadaan komponen keluarga seperti anak usia sekolah, balita, ibu hamil, lansia, dan penyandang disabilitas. Status kepesertaan dalam basis data kesejahteraan sosial juga dapat menjadi bagian dari data yang ditelaah. Atribut tersebut digunakan sebagai variabel analisis untuk mengetahui pola keterkaitan karakteristik keluarga dalam dataset. Algoritma Apriori membantu menemukan kombinasi atribut yang sering muncul bersamaan, sehingga hasilnya dapat digunakan sebagai informasi pendukung evaluasi calon penerima PKH, bukan sebagai pengganti verifikasi dan ketentuan resmi pemerintah.

Atribut Penelitian

Proposal mengusulkan atribut sosial ekonomi dan kondisi keluarga, antara lain pendapatan keluarga, jumlah tanggungan, kondisi rumah, luas bangunan, jenis lantai, jenis dinding, sumber air bersih, fasilitas sanitasi, kepemilikan aset, status DTKS, keberadaan ibu hamil, anak usia dini, anak sekolah pada jenjang SD/SMP/SMA, lansia, dan penyandang disabilitas berat.

Tabel 1 Atribut .

No.	Atribut	Kategori
1	Pendapatan keluarga	Rendah, sedang, tinggi
2	Jumlah tanggungan	Sedikit (1–2), sedang (3–4), banyak (>4)
3	Kondisi rumah	Layak, kurang layak, tidak layak
4	Luas bangunan rumah	Sempit, sedang, luas
5	Lantai dan dinding rumah	Tanah/semen/keramik; bambu-kayu/semi permanen/permanen
6	Sumber air dan sanitasi	Kategori kelayakan sesuai data lapangan
7	Kepemilikan aset	Tidak ada, sedikit, banyak
8	Komponen keluarga	Ibu hamil, balita, anak sekolah, lansia, disabilitas
9	Status DTKS	Terdaftar, tidak terdaftar

Penelitian Terdahulu

Penelitian terdahulu digunakan sebagai bahan perbandingan untuk mengetahui penerapan metode *Data Mining* dalam analisis penerima bantuan sosial. Ringkasan penelitian terdahulu disajikan pada Tabel 2 berikut.

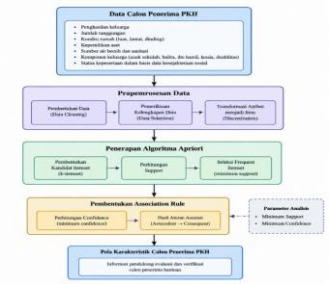
Tabel 2. Penelitian Terdahulu.

No.	Peneliti dan Tahun	Metode	Objek Penelitian	Hasil Penelitian
1	Kurniadi, Nuraeni, dan Jaelani (2022)	Fuzzy Mamdani	Prediksi calon penerima Program Keluarga Harapan (PKH)	Menghasilkan sistem prediksi calon penerima PKH dengan tingkat akurasi sebesar 87%.
2	Septasari dan Abdul Aziz (2024)	Artificial Neural Network (ANN) dan Classification and	Penentuan penerima PKH di Kabupaten Pringsewu	Membandingkan kinerja ANN dan CART; ANN menunjukkan performa lebih baik pada penelitian tersebut.

		Regression Tree (CART)		
3	Chan (2024)	Association Rule dengan algoritma Apriori	Pola penerima bantuan sosial di Kota Medan	Menemukan pola keterkaitan antaratribut bantuan sosial dengan nilai confidence hingga 100%.
4	Marlika, Aswi, dan Annas (2026)	Naive Bayes	Klasifikasi penerima bantuan PKH	Menerapkan klasifikasi untuk menganalisis penerima PKH dan mengevaluasi kinerja model.
5	Ufriani, Jasmir, dan Arvita (2023)	K-Means Clustering	Penentuan prioritas penerima bantuan sosial PKH di Kelurahan Kampung Singkep	Mengelompokkan data penerima ke dalam tiga kelompok prioritas berdasarkan hasil clustering.

Kerangka Pemikiran Penelitian

Penelitian ini berangkat dari kebutuhan untuk menganalisis data calon penerima PKH secara sistematis. Data keluarga yang tersedia diolah melalui tahapan prapemrosesan dan pengelompokan nilai atribut, kemudian dianalisis menggunakan algoritma Apriori untuk menghasilkan frequent itemset dan aturan asosiasi. Alur kerangka pemikiran penelitian dapat dijelaskan sebagai berikut:



Gambar 1. Kerangka Pemikiran Penelitian.

3.METODE PENELITIAN

Jenis Penelitian

Penelitian ini menggunakan metode penelitian kuantitatif dengan pendekatan *Data Mining*. Pendekatan kuantitatif digunakan untuk mengolah data calon penerima Program Keluarga Harapan (PKH) berdasarkan atribut yang berkaitan dengan kondisi sosial dan ekonomi keluarga. Data tersebut dianalisis menggunakan algoritma Apriori untuk menemukan pola hubungan antaratribut yang dapat mendukung proses penentuan calon penerima bantuan. Penelitian ini berfokus pada penerapan algoritma Apriori melalui proses pencarian frequent itemset dan pembentukan association rule. Pola yang diperoleh selanjutnya dianalisis untuk

mengetahui kombinasi karakteristik keluarga yang sering muncul dalam data calon penerima PKH.

Lokasi dan Waktu Penelitian

Penelitian dilaksanakan di Dinas Sosial Kabupaten Sumba Barat Daya, Nusa Tenggara Timur. Data yang dianalisis berupa data administrasi yang berkaitan dengan keluarga dan Program Keluarga Harapan (PKH), yang penggunaannya harus memperoleh izin dari instansi terkait. Waktu penelitian mencakup pengumpulan dan pemeriksaan data, prapemrosesan, penerapan algoritma Apriori, analisis hasil, serta penyusunan laporan. Periode penelitian perlu diisi sesuai waktu pelaksanaan sebenarnya.

Jenis dan Sumber Data

Jenis Data

Jenis data yang digunakan dalam penelitian ini adalah data kuantitatif dan data kategorikal. Data kuantitatif dapat berupa nilai penghasilan keluarga atau jumlah tanggungan, sedangkan data kategorikal berupa kondisi rumah, kepemilikan aset, kondisi sanitasi, serta komponen pendidikan dan kesehatan. Sebelum diproses menggunakan algoritma Apriori, data akan diubah ke dalam bentuk kategori atau item yang sesuai dengan kebutuhan analisis.

Sumber Data

Sumber data penelitian terdiri atas: Data primer, apabila peneliti melakukan pengumpulan data secara langsung melalui observasi, wawancara, atau kuesioner kepada pihak terkait. Data sekunder, berupa dokumen administrasi atau data calon penerima PKH yang diperoleh dari instansi berwenang sesuai dengan izin penggunaan data. Data yang digunakan harus relevan dengan tujuan penelitian dan dijaga kerahasiaannya, terutama informasi identitas serta kondisi sosial ekonomi keluarga.

Teknik Pengumpulan Data

Pengumpulan data merupakan tahapan penting dalam penelitian karena kualitas data yang diperoleh akan memengaruhi hasil analisis dan ketepatan pola yang dihasilkan. Pada penelitian ini, data yang digunakan berupa data calon penerima dan penerima Program Keluarga Harapan (PKH) yang diperoleh dari instansi terkait,. Teknik pengumpulan data yang digunakan dalam penelitian ini meliputi:

Observasi

Observasi dilakukan untuk mengetahui proses pendataan dan pengelolaan informasi calon penerima PKH pada lokasi penelitian. Kegiatan ini bertujuan memahami atribut yang tersedia serta permasalahan yang terjadi dalam pengolahan data.

Wawancara

Wawancara dilakukan kepada petugas atau pihak terkait untuk memperoleh informasi mengenai proses pendataan, atribut yang digunakan, dan kendala dalam pengelolaan data calon penerima bantuan.

Dokumentasi

Dokumentasi dilakukan dengan mengumpulkan dokumen atau data yang relevan dengan penelitian. Data yang diperoleh kemudian diperiksa kelengkapan, konsistensi, dan kesesuaiannya dengan kebutuhan penerapan algoritma Apriori.

Prapemrosesan Data (Data Preprocessing)

Sebelum dianalisis dengan algoritma Apriori, data mentah terlebih dahulu melalui beberapa tahapan prapemrosesan: *Data Cleaning*: Membersihkan data dari *missing values*, data ganda, dan data yang tidak konsisten. *Data Selection*: Memilih atribut yang relevan dengan kriteria PKH, yaitu pendapatan keluarga, jumlah tanggungan, kondisi rumah, jenis lantai, jenis dinding, kepemilikan aset, serta keberadaan komponen keluarga (ibu hamil, balita, anak sekolah, lansia, disabilitas). *Data Transformation / Diskretisasi***:** Mengonversi data kontinu menjadi variabel kategoris berupa item transaksi.

Algoritma Apriori dan Association Rule

Algoritma Apriori mencari kombinasi itemset yang sering muncul (*frequent itemset*) dengan memenuhi batas ambang *Minimum Support* dan *Minimum Confidence*.

Association Rule

Association Rule atau aturan asosiasi merupakan teknik analisis data untuk mengidentifikasi hubungan antaritem yang muncul secara bersamaan dalam dataset. Hasil analisis biasanya disajikan dalam bentuk aturan \(A \rightarrow B\), yang menunjukkan bahwa kemunculan item A memiliki keterkaitan dengan kemunculan item B. Analisis ini umumnya menggunakan dua ukuran utama, yaitu support dan confidence. Aliyah et al. (2023) menggunakan algoritma Apriori untuk menganalisis pola penduduk yang berhak menerima bantuan. Penelitian Chan (2025) juga menerapkan association rule dengan algoritma Apriori

pada data penerima bantuan sosial. Studi tersebut relevan sebagai rujukan penerapan aturan asosiasi dalam analisis karakteristik penerima bantuan.

Support

Support digunakan untuk mengukur seberapa sering suatu item atau kombinasi item muncul dibandingkan dengan keseluruhan data. Rumus support untuk satu item adalah:

$$Support(A) = \frac{\text{Jumlah data yang mengandung } A}{\text{Jumlah seluruh data}} \times 100\% \quad \text{.....(i)}$$

Sementara itu, support untuk kombinasi dua item A dan B dihitung dengan rumus:

$$Support(A,B) = \frac{\text{Jumlah data yang mengandung } A \text{ dan } B}{\text{Jumlah seluruh data}} \times 100\% \quad \text{.....(ii)}$$

Nilai support digunakan untuk menyeleksi kombinasi item yang cukup sering muncul sehingga layak dianalisis lebih lanjut.

Confidence

Confidence digunakan untuk mengukur proporsi data yang mengandung A dan B terhadap seluruh data yang mengandung A. Rumusnya adalah:

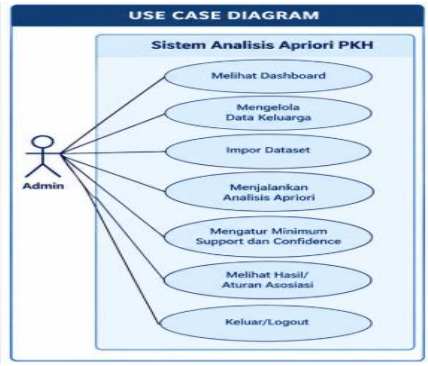
$$Confidence(A \rightarrow B) = \frac{\text{Jumlah data yang mengandung } A \text{ dan } B}{\text{Jumlah data yang mengandung } A} \times 100\% \quad \text{.....(iii)}$$

Support dan confidence digunakan bersama-sama untuk menilai pola asosiasi. Namun, nilai yang tinggi tidak secara otomatis membuktikan hubungan sebab-akibat atau menentukan bahwa suatu keluarga pasti berhak menerima bantuan.

Perancangan Sistem

Use Case Diagram

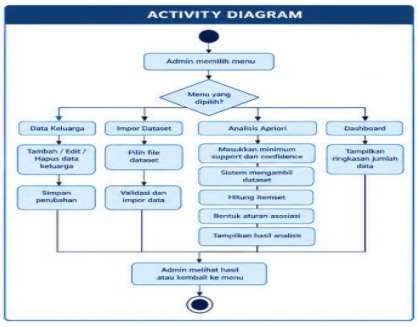
Use Case Diagram menggambarkan interaksi antara aktor dan fungsi yang disediakan oleh sistem. Berdasarkan fitur aplikasi, aktor utama adalah Admin/Pengelola Data.



Gambar 2. *Use Case Diagram.*

Activity Diagram Analisis Apriori

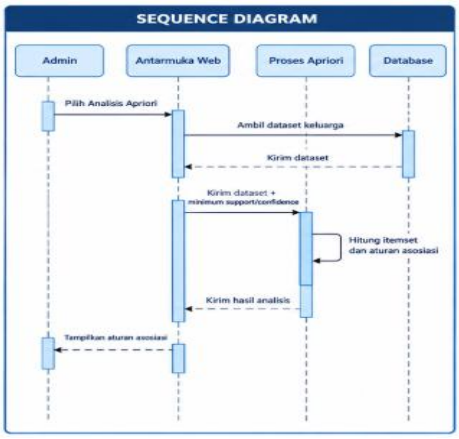
Activity Diagram menjelaskan alur aktivitas pengguna saat menjalankan analisis.



Gambar 3. *Activity Diagram*.

Sequence Diagram

Sequence Diagram menggambarkan urutan komunikasi antara admin, antarmuka web, proses algoritma Apriori, dan basis data.



Gambar 4. *Sequence Diagram.*

4. HASIL DAN PEMBAHASAN

Hasil Implementasi Sistem

Implementasi algoritma Apriori untuk analisis pola asosiasi data keluarga pada Program Keluarga Harapan (PKH) diwujudkan dalam sistem berbasis web menggunakan PHP dan MySQL. Sistem ini dirancang untuk memudahkan pengelolaan dataset keluarga, proses impor data, pelaksanaan analisis *Data Mining*, serta penyajian hasil berupa aturan asosiasi. Sistem memiliki beberapa menu utama, yaitu Dashboard, Data Keluarga, Impor Dataset, Analisis Apriori, dan Keluar. Setiap menu mendukung proses pengolahan data hingga diperoleh pola asosiasi dari atribut yang terdapat pada dataset.

Halaman Login

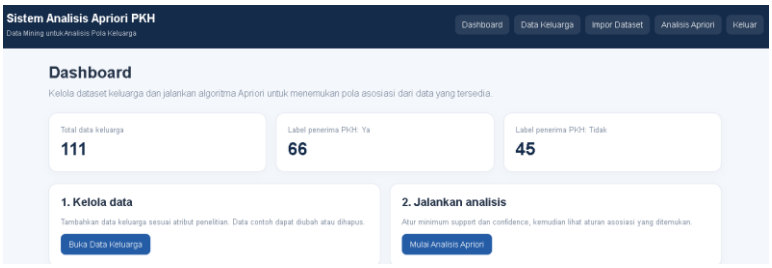
Halaman login merupakan halaman awal yang digunakan oleh Admin untuk masuk ke Sistem Analisis Apriori PKH. Admin harus memasukkan username dan password yang telah terdaftar agar dapat mengakses fitur-fitur sistem.



Gambar 5. Halaman Login .

Halaman Dashboard

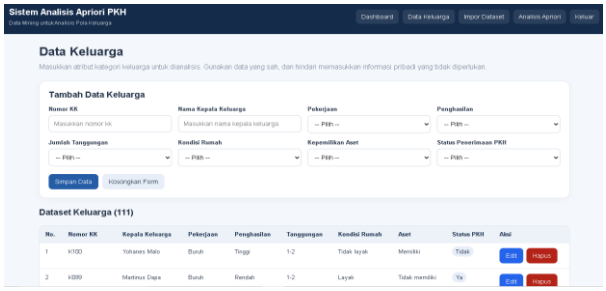
Halaman dashboard merupakan halaman utama yang menampilkan ringkasan data keluarga yang tersimpan dalam sistem. Berdasarkan tampilan aplikasi, dashboard menunjukkan total sebanyak 111 data keluarga, dengan 66 data berlabel penerima PKH "Ya" dan 45 data berlabel "Tidak". Dashboard juga menyediakan akses menuju pengelolaan data keluarga dan proses analisis Apriori. Dengan adanya halaman ini, pengguna dapat mengetahui ringkasan dataset dan mengakses fungsi utama sistem secara langsung.



Gambar 6. Halaman Dashboard.

Halaman Data Keluarga

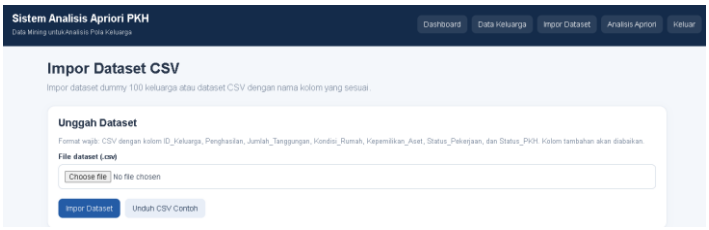
Halaman data keluarga digunakan untuk mengelola dataset yang akan dianalisis menggunakan algoritma Apriori. Pengelolaan data mencakup penambahan, perubahan, penghapusan, dan peninjauan data sesuai dengan fitur yang tersedia pada aplikasi. Data keluarga menjadi masukan utama dalam proses analisis. Sebelum diproses, data perlu diperiksa untuk memastikan kelengkapan atribut, konsistensi kategori, dan kesesuaian format. Data yang digunakan dalam penelitian juga harus dianonimkan untuk menjaga kerahasiaan informasi keluarga.



Gambar 7. Halaman Data Keluarga.

Halaman Impor Dataset

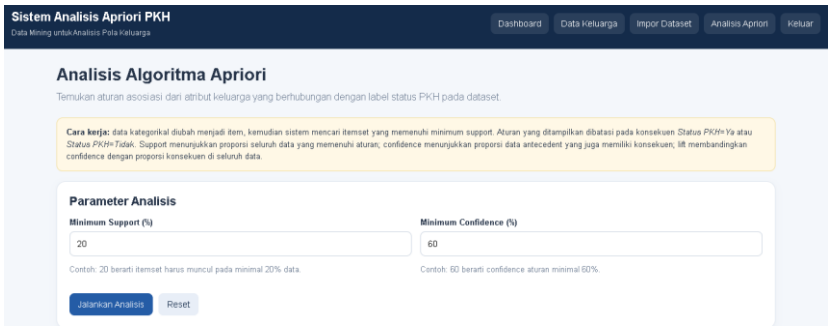
Halaman impor dataset digunakan untuk memasukkan data keluarga dari berkas CSV ke dalam basis data sistem. Fitur ini membantu pengguna memindahkan data dari berkas tabular ke aplikasi tanpa perlu memasukkan setiap baris secara manual. Proses impor perlu memastikan bahwa struktur kolom CSV sesuai dengan atribut yang telah ditentukan. Validasi juga diperlukan untuk mengidentifikasi data kosong, format yang tidak sesuai, dan kemungkinan duplikasi.



Gambar 8. Halaman Impor Dataset.

Halaman Analisis Apriori

Halaman analisis Apriori digunakan untuk menjalankan proses pencarian pola asosiasi berdasarkan dataset yang tersedia. Pengguna menentukan nilai minimum support dan minimum confidence sebagai parameter analisis. Sistem kemudian memproses data untuk membentuk frequent itemset dan menghasilkan aturan asosiasi yang memenuhi kriteria. Hasil analisis selanjutnya digunakan untuk mengidentifikasi keterkaitan antaratribut dalam data keluarga.



Gambar 9. Halaman Analisis.

Hasil Analisis Algoritma Apriori

Hasil Aturan Asosiasi

Berdasarkan implementasi algoritma Apriori pada sistem berbasis web, diperoleh sejumlah aturan asosiasi yang menggambarkan keterkaitan antara karakteristik keluarga dan status PKH. Atribut yang terlihat dalam hasil analisis meliputi penghasilan, jumlah tanggungan, kepemilikan aset, pekerjaan, dan kondisi rumah. Setiap aturan dievaluasi menggunakan nilai support, confidence, dan lift. Nilai-nilai tersebut digunakan untuk mengetahui frekuensi kemunculan pola, proporsi transaksi yang memenuhi konsekuen, serta kekuatan asosiasi dibandingkan dengan frekuensi dasar konsekuen.

Hasil Aturan Asosiasi

No.	Aturan Asosiasi	Support	Confidence	Lift	Interpretasi
1	Penghasilan=Rendah DAN Tanggungan=3 atau lebih → Status PKH=Ya	26,13%	85,29%	1,434	Di antara data yang memenuhi kondisi di kiri, 85,29% juga memiliki label Ya. Lift menunjukkan bahwa label tersebut lebih sering muncul dibanding proporsi dasar.
2	Aset=Tidak memiliki DAN Tanggungan=3 atau lebih → Status PKH=Ya	27,03%	83,33%	1,402	Di antara data yang memenuhi kondisi di kiri, 83,33% juga memiliki label Ya. Lift menunjukkan bahwa label tersebut lebih sering muncul dibanding proporsi dasar.
3	Aset=Tidak memiliki DAN Penghasilan=Rendah → Status PKH=Ya	26,13%	78,38%	1,318	Di antara data yang memenuhi kondisi di kiri, 78,38% juga memiliki label Ya. Lift menunjukkan bahwa label tersebut lebih sering muncul dibanding proporsi dasar.
4	Tanggungan=3 atau lebih → Status PKH=Ya	46,05%	78,13%	1,314	Di antara data yang memenuhi kondisi di kiri, 78,13% juga memiliki label Ya. Lift menunjukkan bahwa label tersebut lebih sering muncul dibanding proporsi dasar.
5	Pekerjaan=Buruh DAN Tanggungan=3 atau lebih → Status PKH=Ya	24,32%	77,14%	1,297	Di antara data yang memenuhi kondisi di kiri, 77,14% juga memiliki label Ya. Lift menunjukkan bahwa label tersebut lebih sering muncul dibanding proporsi dasar.
6	Aset=Tidak memiliki DAN Pekerjaan=Buruh → Status PKH=Ya	24,32%	75,00%	1,261	Di antara data yang memenuhi kondisi di kiri, 75,00% juga memiliki label Ya. Lift menunjukkan bahwa label tersebut lebih sering muncul dibanding proporsi dasar.
7	Rumah=Layak DAN Tanggungan=3 atau lebih → Status PKH=Ya	25,23%	73,68%	1,239	Di antara data yang memenuhi kondisi di kiri, 73,68% juga memiliki label Ya. Lift menunjukkan bahwa label tersebut lebih sering muncul dibanding proporsi dasar.
8	Penghasilan=Rendah → Status PKH=Ya	36,94%	69,49%	1,169	Di antara data yang memenuhi kondisi di kiri, 69,49% juga memiliki label Ya. Lift menunjukkan bahwa label tersebut lebih sering muncul dibanding proporsi dasar.
9	Pekerjaan=Buruh → Status PKH=Ya	33,33%	66,07%	1,111	Di antara data yang memenuhi kondisi di kiri, 66,07% juga memiliki label Ya. Lift menunjukkan bahwa label tersebut lebih sering muncul dibanding proporsi dasar.
10	Tanggungan=1-2 → Status PKH=Tidak	27,93%	65,96%	1,827	Di antara data yang memenuhi kondisi di kiri, 65,96% juga memiliki label Tidak. Lif menunjukkan bahwa label tersebut lebih sering muncul dibanding proporsi dasar.
11	Rumah=Tidak layak → Status PKH=Ya	26,13%	65,91%	1,108	Di antara data yang memenuhi kondisi di kiri, 65,91% juga memiliki label Ya. Lift menunjukkan bahwa label tersebut lebih sering muncul dibanding proporsi dasar.
12	Aset=Tidak memiliki → Status PKH=Ya	39,64%	64,71%	1,088	Di antara data yang memenuhi kondisi di kiri, 64,71% juga memiliki label Ya. Lift menunjukkan bahwa label tersebut lebih sering muncul dibanding proporsi dasar.

Gambar 10. Hasil Aturan Asosiasi Algoritma Apriori.

Analisis Nilai Support

Nilai support menunjukkan persentase transaksi dalam dataset yang mengandung itemset pada suatu aturan asosiasi. Berdasarkan gambar 10, aturan keempat, yaitu Tanggungan ≥ 3 → PKH=Ya, memiliki nilai support tertinggi sebesar 45,05%. Hal ini menunjukkan bahwa kombinasi tanggungan tiga orang atau lebih dan status PKH "Ya" ditemukan pada proporsi transaksi tersebut dalam dataset yang dianalisis. Aturan kedua belas, yaitu Tidak memiliki aset → PKH=Ya, memiliki support sebesar 39,64%, sedangkan aturan kedelapan, yaitu Penghasilan rendah → PKH=Ya, memiliki support sebesar 36,94%. Sebaliknya, aturan kelima dan keenam memiliki support masing-masing sebesar 24,32%. Pola tersebut muncul pada proporsi transaksi yang lebih kecil dibandingkan aturan keempat. Hasil ini menunjukkan bahwa frekuensi kemunculan setiap pola berbeda-beda. Namun, support yang tinggi tidak selalu menunjukkan bahwa aturan memiliki hubungan asosiasi yang paling kuat karena ukuran tersebut tidak mempertimbangkan kekuatan hubungan secara relatif.

Analisis Nilai Confidence

Nilai confidence digunakan untuk mengukur proporsi transaksi yang mengandung konsekuen di antara transaksi yang memenuhi anteseden suatu aturan. Aturan pertama, yaitu Penghasilan rendah DAN tanggungan ≥ 3 → PKH=Ya, memperoleh confidence tertinggi sebesar 85,29%. Artinya, di antara transaksi yang memenuhi kedua kondisi tersebut, 85,29% memiliki label PKH "Ya". Aturan kedua memiliki confidence sebesar 83,33%, diikuti aturan ketiga sebesar 78,38%. Hasil tersebut menunjukkan bahwa kombinasi penghasilan rendah dan jumlah tanggungan tiga orang atau lebih memiliki proporsi konsekuen "Ya" tertinggi di antara aturan yang ditampilkan. Sementara itu, aturan kesembilan, yaitu Pekerjaan buruh → PKH=Ya, memiliki confidence sebesar 66,67%. Dengan demikian, sekitar dua pertiga transaksi yang memenuhi kondisi pekerjaan buruh juga memiliki label PKH "Ya" pada dataset tersebut. Pada aturan dengan konsekuen berbeda, aturan kesepuluh, yaitu Tanggungan 1–2 → PKH=Tidak, menghasilkan confidence sebesar 65,96%. Temuan ini menunjukkan bahwa aturan asosiasi dapat dibentuk untuk kedua label status PKH, bukan hanya untuk label "Ya".

Analisis Nilai Lift

Nilai lift digunakan untuk mengevaluasi kekuatan asosiasi dengan membandingkan frekuensi kemunculan bersama suatu itemset terhadap frekuensi yang diharapkan apabila anteseden dan konsekuen independen. Dari 12 aturan yang ditampilkan, aturan kesepuluh memperoleh nilai lift tertinggi sebesar 1,627. Aturan tersebut menghubungkan Tanggungan 1–2 dengan Status PKH=Tidak. Nilai ini menunjukkan bahwa kedua kondisi tersebut memiliki asosiasi positif dalam dataset yang dianalisis. Aturan pertama memiliki lift sebesar 1,434, sedangkan aturan kedua memiliki lift sebesar 1,402. Kedua nilai tersebut juga menunjukkan asosiasi positif antara kombinasi karakteristik keluarga dan label PKH "Ya". Seluruh aturan yang ditampilkan mempunyai nilai lift lebih besar dari 1. Berdasarkan ukuran ini, setiap aturan menunjukkan asosiasi positif dalam dataset. Meskipun demikian, nilai lift tidak membuktikan hubungan sebab-akibat dan tidak menjamin bahwa pola yang sama akan ditemukan pada populasi lain.

Pembahasan Aturan Asosiasi

Hasil analisis memperlihatkan bahwa aturan terbaik bergantung pada ukuran yang digunakan. Aturan pertama memiliki confidence tertinggi sebesar 85,29%, sedangkan aturan keempat mempunyai support tertinggi sebesar 45,05%. Adapun aturan kesepuluh memperoleh lift tertinggi sebesar 1,627. Perbedaan tersebut menunjukkan pentingnya mempertimbangkan support, confidence, dan lift secara bersamaan ketika menginterpretasikan aturan asosiasi. Jika hanya menggunakan confidence, aturan dengan konsekuen yang sering muncul dapat terlihat

kuat, padahal kekuatan asosiasinya perlu diperiksa dengan lift. Temuan lain yang terlihat adalah adanya aturan dengan konsekuen Status PKH=Tidak. Aturan kesepuluh memperlihatkan keterkaitan antara jumlah tanggungan satu sampai dua orang dan label PKH "Tidak". Namun, hasil ini hanya menggambarkan pola yang ditemukan pada dataset, bukan bukti bahwa jumlah tanggungan tersebut menyebabkan keluarga tidak menerima bantuan. Secara keseluruhan, hasil penelitian menunjukkan bahwa algoritma Apriori dapat digunakan untuk mengeksplorasi keterkaitan antaratribut dalam data keluarga. Sistem memberikan informasi yang lebih terstruktur melalui penyajian aturan asosiasi beserta nilai evaluasinya. Hasil analisis dapat digunakan sebagai bahan eksplorasi data, tetapi tidak boleh dijadikan satu-satunya dasar untuk menetapkan kelayakan penerima PKH.

5. KESIMPULAN

Berdasarkan hasil implementasi dan analisis *Data Mining* menggunakan algoritma Apriori pada sistem analisis data keluarga Program Keluarga Harapan (PKH), dapat disimpulkan bahwa sistem berbasis web menggunakan PHP dan MySQL telah menyediakan fasilitas pengelolaan data keluarga, impor dataset, serta penyajian aturan asosiasi berdasarkan nilai support, confidence, dan lift. Berdasarkan hasil analisis yang ditampilkan oleh sistem, terdapat 12 aturan asosiasi yang menghubungkan karakteristik keluarga dengan status PKH "Ya" dan "Tidak". Aturan dengan nilai confidence tertinggi adalah penghasilan rendah dan jumlah tanggungan tiga orang atau lebih terhadap status PKH "Ya", dengan confidence sebesar 85,29%, support sebesar 26,13%, dan lift sebesar 1,434. Sementara itu, aturan dengan nilai lift tertinggi adalah jumlah tanggungan satu sampai dua orang terhadap status PKH "Tidak", dengan confidence sebesar 65,96% dan lift sebesar 1,627. Hasil tersebut menunjukkan bahwa algoritma Apriori dapat digunakan untuk mengidentifikasi pola keterkaitan antaratribut dalam dataset keluarga. Sistem yang dikembangkan dapat membantu proses eksplorasi data secara lebih terstruktur. Namun, aturan asosiasi yang dihasilkan hanya menunjukkan keterkaitan statistik dalam dataset dan tidak membuktikan hubungan sebab-akibat maupun menentukan kelayakan resmi penerima bantuan PKH.

DAFTAR REFERENSI

Agrawal, R., & Srikant, R. (1994). Fast algorithms for mining association rules in large databases. *Proceedings of the 20th International Conference on Very Large Data Bases*, 487–499. https://www.vldb.org/conf/1994/P487.PDF

Aprianti, W., Hafizd, K. A., & Rizani, M. R. (2017). Implementasi association rules dengan algoritma Apriori pada dataset kemiskinan. *Limits: Journal of Mathematics and Its Applications, 14*(2), 145–155. https://doi.org/10.12962/limits.v14i2.2933

Chan, F. R. (2025). Analisis pola penerima bansos Kota Medan dengan metode association rule menggunakan algoritma Apriori. *TECHSI: Jurnal Teknik Informatika, 16*(1), 9–27. https://doi.org/10.29103/techsi.v16i1.20454

Darmawan, I. A., Randy, M. F., Yunianto, I., Mutoffar, M. M., & Salis, M. T. P. (2022). Penerapan data mining menggunakan algoritma Apriori untuk menentukan pola golongan penyandang masalah kesejahteraan sosial. *Sebatik, 26*(1), 223–230. https://doi.org/10.46984/sebatik.v26i1.1622

Hamdala, T., & Esabella, S. (2020). Rancang bangun aplikasi monitoring dan evaluasi penyaluran dana sosial di Dinas Sosial Kabupaten Sumbawa berbasis web. *Jurnal JINTEKS, 2*(2), 105–111.

Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and techniques* (3rd ed.). Morgan Kaufmann. https://doi.org/10.1016/C2009-0-61819-5

Julfani, L., & Putra, I. M. (2024). Dampak Program Keluarga Harapan (PKH) terhadap peningkatan kesejahteraan keluarga miskin di Kelurahan Kerasaan I Kecamatan Pematang Bandar Kabupaten Simalungun. *Al-Kharaj: Jurnal Ekonomi, Keuangan & Bisnis Syariah*, *6*(1), 591–608.

Kementerian Sosial Republik Indonesia. (n.d.). *Program Keluarga Harapan (PKH).* https://kemensos.go.id/

Kurniadi, D., Nuraeni, F., & Jaelani, D. (2022). Implementasi logika fuzzy Mamdani pada sistem prediksi calon penerima Program Keluarga Harapan. *Jurnal Algoritma*, *19*(1), 160–171. https://doi.org/10.33364/algoritma/v.19-1.1016

Marlika, N., Aswi, & Annas, S. (2026). Penerapan algoritma Naive Bayes untuk klasifikasi penerima bantuan Program Keluarga Harapan (PKH). *VARIANSI: Journal of Statistics and Its Application on Teaching and Research, 8*(1), 1–7. https://doi.org/10.35580/variansiunm418

Masroni, Alkadri, S. P. A., & Insani, R. W. S. (2023). Sistem pendukung keputusan rekomendasi penerima bantuan iuran BPJS Kesehatan menggunakan metode ROC dan SMART. *Jurnal Fasilkom, 13*(3), 496–503.

Moruk, E., Martanto, M., & Hayati, U. (2024). Algoritma K-Means untuk clustering penerima Program Keluarga Harapan di Nanaenoe Nusa Tenggara Timur. *JATI (Jurnal Mahasiswa Teknik Informatika), 8*(2), 1681–1687. https://doi.org/10.36040/jati.v8i2.9148

Nabila, K., Suharso, P., & Hartanto, W. (2021). Implementasi Program Bantuan Pangan Non Tunai (BPNT) di Desa Paleran Kecamatan Umbulsari Kabupaten Jember. *Jurnal Pendidikan Ekonomi: Jurnal Ilmiah Ilmu Pendidikan, Ilmu Ekonomi, dan Ilmu Sosial, 15*(2). https://doi.org/10.19184/jpe.v15i2.21327

Nurzanah, S. C., Alam, S., & Hermanto, T. I. (2022). Analisis association rule untuk identifikasi pola gejala penyakit hipertensi menggunakan algoritma Apriori (Studi kasus: Klinik Rafina Medical Center). *JIKO (Jurnal Informatika dan Komputer), 5*(2), 132–141.

Purnama, A. I., Aziz, A., & Wiguna, A. S. (2020). Penerapan data mining untuk mengklasifikasi penerima bantuan PKH Desa Wae Jare menggunakan metode Naïve Bayes. *Kurawal: Jurnal Teknologi, Informasi dan Industri, 3*(2), 173–180. https://doi.org/10.33479/kurawal.v3i2.348

Putra, R. F., Zebua, R. S. Y., Budiman, B., Rahayu, P. W., Bangsa, M. T. A., Zulfadhilah, M., Choirina, P., Wahyudi, F., & Andiyan, A. (2023). *Data mining: Algoritma dan penerapannya*. PT. Sonpedia Publishing Indonesia.

Rosmayati, I., Wahyuningsih, W., Harahap, E. F., & Hanifah, H. S. (2023). Implementasi data mining pada penjualan kopi menggunakan algoritma Apriori. *Jurnal Algoritma, 20*(1), 99–107. https://doi.org/10.33364/algoritma/v.20-1.1259

Septasari, D., & Abdul Aziz, R. (2024). Perbandingan metode Artificial Neural Network (ANN) dan Classification and Regression Tree (CART) dalam menentukan penerima bantuan Program Keluarga Harapan (PKH). *Jurnal Algoritma, 21*(2), 437–448. https://doi.org/10.33364/algoritma/v.21-2.1794

Sinaga, S., & Husein, A. M. (2019). Penerapan algoritma Apriori dalam data mining untuk memprediksi pola pengunjung pada objek wisata Kabupaten Karo. *Jurnal Teknologi dan Ilmu Komputer Prima, 2*(1), 320–325. https://doi.org/10.34012/jutikomp.v2i1.461

Takdirillah, R. (2020). Penerapan data mining menggunakan algoritma Apriori terhadap data transaksi sebagai pendukung informasi strategi penjualan. *Edumatic: Jurnal Pendidikan Informatika, 4*(1), 37–46. https://doi.org/10.29408/edumatic.v4i1.2081

Tan, P.-N., Steinbach, M., & Kumar, V. (2006). *Introduction to data mining*. Pearson Addison Wesley.

Ufriani, S., Jasmir, & Arvita, Y. (2023). Penerapan algoritma clustering K-Means untuk menentukan prioritas penerima bantuan dana sosial PKH di Kelurahan Kampung Singkep. *Jurnal Informatika dan Rekayasa Komputer (JAKAKOM), 3*(1), 342–350. https://doi.org/10.33998/jakakom.2023.3.1.726

Wahidah, A., Widayani, A., Wardhani, S., Rachmawati, I., & Latifah, N. (2022). Prosedur penyaluran bantuan sosial di era pandemi COVID-19 pada Dinas Sosial Kabupaten Blitar. *Journal of Management Studies, 16*(1), 51–63.

Wijaya, A., Faqih, A., Solihudin, D., Rohmat, C. L., & Permana, S. E. (2023). Penerapan association rules menggunakan algoritma Apriori untuk identifikasi pola pembelian. *JATI (Jurnal Mahasiswa Teknik Informatika), 7*(6), 3871–3878.

Zai, C. (2022). Implementasi data mining sebagai pengolahan data. *Portaldata.org, 2*(3), 1–12.