
ANALISIS POLA PESERTA DIDIK BARU DI SDN 2 LOGEDE MENGGUNAKAN ALGORITMA APRIORI

Gregoria Mariana Selvi^{1*}, Haeni Budiati²

¹Program Studi Informatika, Fakultas Sains dan Komputer, UKRIM, Yogyakarta

²Program Studi Informatika, Fakultas Sains dan Komputer, UKRIM, Yogyakarta

E-mail: gregoria.m2042@student.ukrimuniversity.ac.id¹, heni@ukrimuniversity.ac.id²

Abstrak

Menggunakan algoritma Apriori, salah satu metode data mining yang digunakan untuk menemukan pola hubungan antara item dalam dataset yang besar, penelitian ini bertujuan untuk menganalisis pola siswa baru di SDN 2 Logede. Data siswa yang digunakan di sini, terdiri dari agama, jenis kelamin, tempat lahir, pekerjaan orang tua, pendidikan terakhir orang tua dan alamat tinggal. Penelitian ini melibatkan beberapa langkah seperti pengumpulan data, pra-pemrosesan data, penerapan algoritma Apriori, dan analisis hasil. Setelah data peserta didik baru dikumpulkan dari arsip sekolah, diproses untuk memastikan bahwa informasinya konsisten dan lengkap. Algoritma Apriori digunakan untuk menemukan hubungan antara sifat yang sering muncul bersama. Salah satu teknik data mining yang digunakan dalam penelitian ini adalah klasifikasi asosiasi role untuk mengolah data siswa di SDN 2 Logede. Penelitian ini mengolah data siswa menggunakan bahasa pemrograman python untuk mendapatkan pola dan aturan. Pola asosiasi adalah teknik data mining yang berguna untuk menemukan pola atau korelasi yang paling penting atau menarik dari kumpulan data yang besar. Algoritma Apriori, yang menggunakan teknik peraturan asosiasi untuk melakukan pencarian itemset sering, dapat menghasilkan pola siswa dari tahun 2018 hingga 2023 di Sekolah Negeri 2 Logede di Kabupaten Klaten. Dengan bantuan Algoritma Apriori, dapat disimpulkan bahwa jumlah siswa tertinggi akan terjadi pada tahun 2021. Penelitian ini menggunakan 73 data jumlah siswa dari tahun 2018 hingga 2023. Hasil pengujian menunjukkan bahwa nilai *Support* tertinggi adalah 1,0 dan nilai *Confiden* tertinggi adalah 1,0.

Abstract

Using the Apriori Algorithm, one of the data mining methods used to find patterns of relationships between items in large datasets, this study aims to analysis the new student pattern in Logede's SDN 2. The student data used here consists of religion, gender, place of birth, parent's job, parent's last education, and residence address. This research involves several steps, such as data collection, data pre-processing, Apriori Algorithm application, and results analysis. Once new student data is collected from the school archive, it is processed to ensure that the information is consistent and complete. Apriori Algorithms are used to find relationships between properties that often appear together. One of the data mining techniques used in this study is the classification of association patterns to process student data in SDN 2 Logede. An

Info Naskah:

Naskah masuk: 30 Juni 2024

Direvisi: 13 Juli 2024

Diterima: 17 Juli 2024

Keywords:

Apriori Algorithms;
Data Mining;
Student Patterns;
SDN 2 Logede;
Association Role;
Education;

association pattern is a data mining technique that is useful to find the most important or interesting pattern or correlation in a large data set. The Apriori Algorithm, which uses association regulation techniques to conduct frequent item searches, can generate student patterns from 2018 to 2023 at Logede State School 2 in Klaten district. With the help of the apriority algorithm, it can be concluded that the highest number of students will occur in 2021. The test results showed that the highest support value is 1.0 and the maximum confidence value is 1.0.

***Penulis korespondensi:**

Gregoria Mariana Selvi

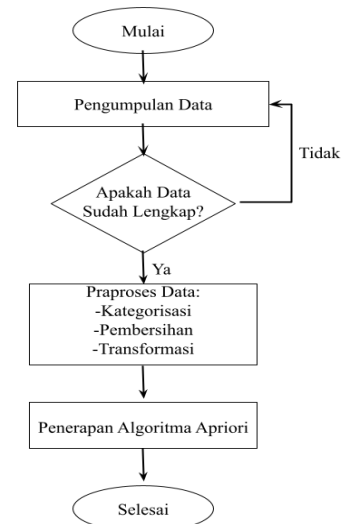
E-mail: gregoria.m2042@student.ukrimuniversity.ac.id

1. Pendahuluan

Teknologi telah mempengaruhi bidang pendidikan, dan sistem pendidikan di Indonesia menggunakan sistem zonasi untuk memperhitungkan lokasi tempat tinggal siswa. Namun, zonasi belum sepenuhnya mengatasi kesenjangan antar sekolah dari segi kualitas, sarana, dan prasarana. Hal ini dapat menurunkan semangat dan daya saing siswa. Selain itu, jumlah siswa baru masuk menjadi tolak ukur keberhasilan sekolah. Faktor-faktor seperti demografi wilayah atau zonasi mempengaruhi jumlah siswa baru yang masuk ke suatu sekolah. Sistem zonasi pendidikan di Indonesia memiliki tantangan dalam mengatur distribusi siswa baru di setiap sekolah. Dalam penelitian ini, menggunakan Algoritma Apriori dapat membantu dalam memahami pola migrasi siswa baru di wilayah tertentu dengan mempertimbangkan faktor seperti jarak, infrastruktur, dan preferensi[1]. Melalui analisis pola menggunakan algoritma Apriori, sekolah dapat meningkatkan efisiensi dalam penempatan siswa baru berdasarkan pola yang ditemukan dari data historis sebelumnya. Hal ini dapat membantu mengurangi ketimpangan distribusi siswa di antara sekolah-sekolah dan memberikan aksesibilitas yang lebih baik terhadap pendidikan yang berkualitas. Algoritma Apriori adalah salah satu algoritma data mining yang sering digunakan untuk menemukan pola asosiasi dalam data. Penelitian ini menganalisis asosiasi menggunakan algoritma Apriori untuk menemukan pola-pola yang sering muncul dan memiliki hubungan erat dengan siswa baru, dengan mencari itemset yang sering muncul dalam transaksi berdasarkan tingkat support dan tingkat confidence. Dengan memahami pola-pola ini, sekolah dapat meningkatkan efisiensi penempatan siswa baru dan menciptakan lingkungan pendidikan yang lebih merata dan berkualitas[2]. Tujuan dari penelitian ini adalah mengidentifikasi pola-pola yang muncul secara konsisten untuk melihat pola asosiasi dalam setiap transaksi yang ada dan menganalisis faktor-faktor yang mempengaruhi pola-pola asosiasi dalam setiap transaksi[3]

2. Metode

Dalam penelitian ini, metode yang digunakan adalah[7]:



Gambar 2.1 Flowchart Tahapan Penelitian

1. Pengumpulan Data

Teknik pengumpulan data yang tepat dan instrumen penelitian yang baik sangat membantu dalam mendapatkan data yang valid dan akurat (Ardiansyah et al., 2023). Dan pada penelitian ini dapat akan diperoleh dari sekolah yang bersangkutan dan berpatok pada data BPS Kabupaten Klaten. [4]

2. Pra Proses Data

a. Kategorisasi Data

Tujuan utama dari kategorisasi data adalah untuk membuat data lebih mudah diurus, dianalisis, dan dipahami. Kategori ini dapat dibentuk berdasarkan berbagai kriteria, seperti jenis data, sifat data, atau tujuan analisis data.

b. Pembersihan Data

Pembersihan data, juga dikenal sebagai pembersihan data, adalah proses mengidentifikasi, memperbaiki, dan menghapus kesalahan, inkonsistensi, atau ketidakakuratan dalam kumpulan data.

c. Transformasi Data

Proses mengubah atau mengubah data dari satu bentuk ke bentuk lain dikenal sebagai transformasi data.

3. Penerapan Algoritma Apriori

Penerapan Algoritma Apriori dilakukan dengan melakukan perhitungan secara manual atau *source code* menggunakan Python dan untuk memverifikasi hasil perhitungan pada penelitian ini.

3. Hasil dan Pembahasan

Berikut ini adalah pembahasan penggunaan Algoritma Apriori dalam mengelola dataset yang berhasil diimplementasikan yaitu sebagai berikut[5]:

3.1 Menentukan Frequent Itemset

Setelah menganalisis satu persatu dari semua dataset yang digunakan berikut adalah Frequent Items awal untuk diimplementasi Algoritma Apriori sebagai berikut[6]:

Tabel 3.1 Frequent Itemset

No	Item
1	{Jenis Kelamin – Laki-Laki; Perempuan}
2	{Agama – Islam}
3	{Tempat Lahir – Klaten; Wonosobo; Wonogiri; Gunung Kidul; Blora; Boyolali; Jember; Tegal}
4	{Pekerjaan Orang Tua – Buruh; Petani; Wiraswasta; Karyawanswasta; Sopir; TNI; Pedagang}
5	{Ijazah Orang Tua – SD; SLTA; SLTP; SMP; SMK; STM; SMA; DI/DII}
6	{Alamat – Dukuh Bogem, Dukuh Druwak, Dukuh Lampar, Dukuh Ngemplak, Dukuh Noloprayan, Dukuh Pakem, Dukuh Paprigan, Dukuh Penyalahan, Dukuh Setran, Dukuh Sidokerto, Dukuh Tegalrejo, Dukuh Watesari}

3.2 Menghitung Support dan Confidence dan setiap Frequent Itemset

3.2.1 Menghitung Nilai Support

Menentukan nilai minimum confidence dan minimum support, diketahui Minimum Support = 0.3[8].

Dengan rumus Support sebagai berikut :

$$\text{Support (A)} = \frac{\text{Jumlah Transaksi mengandung A}}{\text{Total Transaksi}} \quad (1)$$

Berdasarkan rumus diatas berikut adalah hasil pengelaminasi itemset yang memenuhi nilai minimal support 0.3 sebagai berikut:

Tabel 3.2 Hasil Nilai Support

No	Item	Support
1	{JenisKelamin-LakiLaki}	0.5
2	{JenisKelamin-Perempuan}	0.4
3	{Agama-Islam}	1.0
4	{TempatLahir-Klaten}	0.9
5	{PekerjaanOrtu-Buruh}	0.7
6	{IjazahOrtu-SD}	0.5
7	{Alamat-Dukuh Druwak Desa Logede}	0.3
8	{Alamat-Dukuh Sidokerto Desa Logede}	0.5

3.2.2 Menghitung Nilai Confidence

Menentukan nilai minimum confidence dan minimum support, diketahui Minimum Confidence = 0.5[8].

Rumus Confidence sebagai berikut :

$$\text{Confidence } P(B | A) = \frac{\text{Jumlah Transaksi mengandung A dan B}}{\text{Total Transaksi mengandung A}} \quad (2)$$

Tabel 3.3 Hasil Nilai Confidence

No	Item	Confidence
1	{Agama-Islam}	1.0
2	{TempatLahir-Klaten}	1.0
3	{Agama-Islam}, {TempatLahir-Klaten}	0.9
4	{PekerjaanOrtu-Buruh}	1.0
5	{Agama-Islam}, {PekerjaanOrtu-Buruh}	0.7
6	{Agama-Islam}, {PekerjaanOrtu-Buruh}, {TempatLahir-Klaten}	0.7
7	{PekerjaanOrtu-Buruh}, {TempatLahir-Klaten}	1.0
8	{Agama-Islam}, {JenisKelamin-LakiLaki}	0.5
9	{Agama-Islam}, {Alamat-Dukuh Sidokerto Desa Logede}	0.5
10	{Agama-Islam}, {IjazahOrtu-SD}	0.5
11	{JenisKelamin-LakiLaki}	1.0
12	{Alamat-Dukuh Sidokerto Desa Logede}	1.0
13	{IjazahOrtu-SD}	1.0
14	{JenisKelamin-Perempuan}, {Agama-Islam}	1.0
15	{JenisKelamin-Perempuan}	1.0
16	{JenisKelamin-Perempuan}, {TempatLahir-Klaten}	1.0
17	{JenisKelamin-Perempuan}, {Agama-Islam}, {TempatLahir-Klaten}	1.0
18	{IjazahOrtu-SD}, {Agama-Islam}, {TempatLahir-Klaten}	0.8
19	{IjazahOrtu-SD}, {TempatLahir-Klaten}	0.8
20	{JenisKelamin-LakiLaki}, {TempatLahir-Klaten}	0.8
21	{JenisKelamin-Perempuan}, {Agama-Islam}, {PekerjaanOrtu-Buruh}	0.7
22	{JenisKelamin-LakiLaki}, {PekerjaanOrtu-Buruh}	0.6
23	{JenisKelamin-Perempuan}, {PekerjaanOrtu-Buruh}	0.7
24	{IjazahOrtu-SD}, {Agama-Islam}, {PekerjaanOrtu-Buruh}, {TempatLahir-Klaten}	0.6

25	{IjazahOrtu-SD},{PekerjaanOrtu-Buruh},{TempatLahir-Klaten}	0.6
26	{JenisKelamin-Perempuan},{Agama-Islam},{PekerjaanOrtu-Buruh},{TempatLahir-Klaten}	0.7
27	{JenisKelamin-Perempuan},{PekerjaanOrtu-Buruh},{TempatLahir-Klaten}	0.7
28	{JenisKelamin-LakiLaki},{PekerjaanOrtu-Buruh},{TempatLahir-Klaten}	0.6
29	{Alamat-Dukuh Druwak Desa Logede}	1.0
30	{Alamat-Dukuh Druwak Desa Logede},{TempatLahir-Klaten}	1.0

3..2.3 Menerapkan Aturan Asosiasi pada Algoritma Apriori

Setelah menentukan nilai minimum *support* dan *confidence* adalah menerapkan aturan asosiasi atau keterkaitan antara *role* terkuat dari nilai tertinggi *support* dan *confidence*[9].

```

#mengumpulkan aturan yang dihasilkan dalam dataframe
untuk pengujian model
#mengurut berdasarkan nilai confidence paling tinggi
#mengurut berdasarkan nilai support yang paling kuat
rules = association_rules(frequent_itemsets, metric="confidence", min_threshold= 1)
rules = rules.sort_values(["support"], ascending=False)
rules.head(10)

```

Gambar 3.1 Code Untuk Menghasilkan Aturan Asosiasi

Langkah selanjutnya adalah menjalankan *code* program dan ditampilkan dataframe dari setiap dataset yang saling berasosiasi dan di penelitian ini menampilkan 10 dataset dengan nilai *support* dan *confidence* tertinggi yang memiliki asosiasi yang tinggi[10].

	antecedents	consequents	antecedent support	consequent support	support	confidence
4	(Klaten)	(Islam)	0.504110	1.0	0.504110	1.0
9	(Buruh)	(Islam)	0.767329	1.0	0.767329	1.0
9	(Buruh Klaten)	(Islam)	0.712329	1.0	0.712329	1.0
3	(Dukuh Sidokerto Desa Logede)	(Islam)	0.506849	1.0	0.506849	1.0
5	(Laki - Laki)	(Islam)	0.506849	1.0	0.506849	1.0
7	(SD)	(Islam)	0.506849	1.0	0.506849	1.0
6	(Perempuan)	(Islam)	0.493151	1.0	0.493151	1.0
16	(Klaten, Dukuh Sidokerto Desa Logede)	(Islam)	0.479452	1.0	0.479452	1.0
18	(Perempuan, Klaten)	(Islam)	0.479452	1.0	0.479452	1.0
19	(Klaten, SD)	(Islam)	0.438356	1.0	0.438356	1.0

Gambar 3.2 Hasil Program Data Asosiasi

4. Kesimpulan

Berdasarkan perancangan analisis dan hasil implementasi penelitian pada program yang menggunakan Algoritma Apriori pada studi kasus yang menganalisis pola peserta didik baru di SDN 2 Logede, dapat disimpulkan bahwa:

Setiap peserta didik yang memiliki salah satu dari pola berikut dipastikan beragama Islam:

Peserta didik lahir di Kabupaten Klaten.

1. Peserta didik mempunyai orang tua yang bekerja sebagai buruh.

2. Peserta didik berdomisili di alamat Dukuh Sidokerto Desa Logede.
 3. Peserta didik berjenis kelamin laki-laki atau perempuan.
 4. Peserta didik mempunyai orang tua dengan ijazah terakhir pendidikan SD.
- Selain itu, kombinasi dari beberapa poin lainnya, seperti:

1. Lahir di Klaten dan berdomisili di Dukuh Sidokerto Desa Logede.
2. Lahir di Klaten dan mempunyai orang tua dengan ijazah pendidikan terakhir SD.
3. Berjenis kelamin perempuan atau laki-laki dan lahir di Klaten.

Juga dipastikan bahwa peserta didik tersebut beragama Islam.

Ucapan Terimakasih

Penulis berterima kasih kepada semua pihak yang berkontribusi dalam penelitian ini kepada kedua orang tua, dosen pembimbing pertama dan kedua, dosen penguji, pihak sekolah SDN 2 Logede dan semua orang yang terlibat dalam penelitian yang tidak dapat disebutkan satu persatu.

Daftar Pustaka:

- [1] Tasyirifiah, T., & Pitaloka, A. F. (2023). Dampak Pemberlakuan Sistem Zonasi terhadap Daya Saing Pelajar di Indonesia. *AI-DYAS*, 2(2), 381–391. <https://doi.org/10.58578/aldyas.v2i2.1200>
- [2] Anas, A., & Zebua, A. J. (2024). *Penerapan Aturan Asosiasi Pada Pengunjung Perpustakaan Universitas Graha Karya Muara Bulian*. 18(1), 27–34.
- [3] Dwi, K., & Kristoforus, H. (2023). *Pedagogi : Jurnal Pendidikan dan Pembelajaran Potret Beban Pendidikan Siswa dan Orangtua dalam Menghadapi*. 3(2), 44–54.
- [4] Ardiansyah, A., Zy, A. T., & Nugroho, A. (2023). Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Obat (Studi Kasus Klinik Pratama Keluarga Kesehatan). *JISAMAR (Journal of ...)*, 7(3), 777–788. <https://doi.org/10.52362/jisamar.v7i3.1163>
- [5] Fitriah, Imam Riadi, & Herman. (2023). Analisis Data Mining Sistem Inventory Menggunakan Algoritma Apriori. *Decode: Jurnal Pendidikan Teknologi Informasi*, 3(1), 118–129. <https://doi.org/10.51454/decode.v3i1.13>
- [6] Reisyar, K. E., & Harman, R. (2023). Implementasi Algoritma Apriori Dalam Pemodelan Pola Pembelian Produk Pada Edukits Btc. *Computer and Science Industrial Engineering (COMASIE)*, 9(3). <https://doi.org/10.33884/comasiejournal.v9i3.7656>
- [7] Saputra, Z., Sartika, D., & Haviz Irfani, M. (2024). RESOLUSI : Rekayasa Teknik Informatika dan Informasi Prediksi Calon Mahasiswa Penerima KIP Pada Universitas Indo Global Mandiri menggunakan Algoritma Decision Tree. *Jurnal Rekayasa Teknik Informatika Dan Informasi*, 43(3), 231–240.

<https://docs.python.org/3.13/tutorial/index.html>

[8]Mulya, M. F., Rismawati, N., & Alifi, R. R. (2019). Analisis Dan Implementasi Data Mining Menggunakan Algoritma Apriori Untuk Meningkatkan Penjualan Pada Kantin Universitas Tanri Abeng. *Faktor Exacta*, 12(3), 210. <https://doi.org/10.30998/faktorexacta.v12i3.4541>

[9]Nirmala, V. (2023). Peran Sekolah dalam Menanggulangi Kenakalan Remaja Pada SMK Bina Sriwijaya. *Jurnal Pendidikan Indonesia (PJPI) E-ISSN:*, 1(1), 65–76. <https://doi.org/10.00000/pjpi.v1n12023>

[10]Lokal, S. B., & Wahyudi, W. (2023). *Penggunaan Algoritma Apriori untuk Data Mining Penjualan*. 3(2)