Analisis Algoritma Apriori untuk Mendukung Kesalahan Suku Kata pada Hasil Tes Literasi Siswa Sekolah Dasar

Mochammad Ilham Aziz[1]*, Atika Nur Fadilla[2], Anis Sholikhah[3], Saifulloh Azhar[4], Muhammad Oktoda Noorrohman[5]

[1,4,5]Program Studi Informatika, Institut Teknologi Al Mahrusiyah, Kota Kediri, Indonesia
[2]Program Studi Keperawatan, Sekolah Tinggi Ilmu Kesehatan Husada, Jombang, Indonesia
[3]Program Studi Sistem Informasi, Institut Teknologi Al Mahrusiyah, Kota Kediri, Indonesia
Email: mochammadilhamaziz@itama.ac.id[1], fadillaatika8@gmail.com[2], anissholikhah164@gmail.com[3] , azharnian@gmail.com[4], moktodan01@gmail.com[5]

Abstrak

Penelitian ini bertujuan untuk menganalisis pola kesalahan pengucapan suku kata pada hasil tes literasi siswa sekolah dasar dengan menerapkan algoritma Apriori sebagai salah satu teknik *data mining* dalam menemukan pola keterkaitan antar kesalahan. Melalui pendekatan kuantitatif deskriptif, data diperoleh dari hasil tes membaca siswa yang menunjukkan berbagai jenis kesalahan pada suku kata tertentu. Data kemudian diolah menggunakan algoritma Apriori untuk menghasilkan aturan asosiasi dengan memperhitungkan nilai *support*, *confidence*, dan lift guna mengidentifikasi hubungan antar kesalahan pengucapan yang paling signifikan. Hasil penelitian menunjukkan bahwa terdapat hubungan kuat antara kesalahan pengucapan beberapa suku kata, seperti Hon, De, dan Dak, dengan kesalahan pada suku kata Yam, yang memiliki nilai *confidence* di atas 50%. Temuan ini mengindikasikan adanya pola sistematis dalam kesalahan fonologis siswa yang dapat digunakan sebagai indikator kesulitan membaca dan mengenali bunyi bahasa. Kesimpulannya, penerapan algoritma Apriori efektif dalam mengungkap pola tersembunyi pada data kesalahan literasi siswa dan dapat menjadi alat diagnostik pendukung bagi guru untuk merancang strategi pembelajaran fonetik serta intervensi literasi yang lebih tepat sasaran.

Kata Kunci: Algoritma Apriori; Literasi Dasar; Kesalahan Membaca; Suku Kata*; Data Mining*

ABSTRACT

This study aims to analyze syllable mispronunciation patterns in elementary school students' literacy test results by applying the Apriori algorithm as a data mining technique to discover patterns of association between errors. Using a descriptive quantitative approach, data were obtained from students' reading test results, which showed various types of errors in certain syllables. The data were then processed using the Apriori algorithm to generate association rules by calculating support, confidence, and lift values to identify the most significant relationships between pronunciation errors. The results showed a strong correlation between mispronunciations of several syllables, such as Hon, De, and Dak, and errors in the syllable Yam, which had a confidence value above 50%. This finding indicates a systematic pattern in students' phonological errors that can be used as an indicator of difficulty reading and recognizing language sounds. In conclusion, the application of the Apriori algorithm is effective in revealing hidden patterns in students' literacy error data and can be a supporting diagnostic tool for teachers in designing more targeted phonics learning strategies and literacy interventions.

Keywords: *Apriori Algorithm; Basic Literacy; Reading Errors; Syllables; Data Mining*

1. PENDAHULUAN

Kemampuan membaca permulaan merupakan fondasi penting dalam pengembangan literasi siswa sekolah dasar. Namun, berbagai penelitian menunjukkan bahwa kesalahan dalam mengenali huruf, menggabungkan suku kata, serta membaca kata secara utuh masih menjadi kendala utama dalam proses belajar membaca. (Oktari & Dwianto, 2025) menemukan bahwa kesulitan membaca siswa kelas 1 SDN Sukamaju Way Kanan Lampung disebabkan oleh faktor internal, seperti kurangnya kesadaran fonologis dan rendahnya minat membaca, serta faktor eksternal, seperti minimnya dukungan orang tua dan lingkungan belajar yang kurang kondusif. Pola kesalahan tersebut bersifat berulang dan sistematis, terutama pada pengucapan serta penggabungan suku kata, sehingga diperlukan pendekatan analisis yang mampu mengidentifikasi hubungan antar jenis kesalahan secara objektif dan berbasis data.

Dalam konteks ini, algoritma Apriori menjadi salah satu teknik *data mining* yang potensial untuk diterapkan di bidang pendidikan. (Zakur & Flaih, 2023) menjelaskan bahwa algoritma Apriori berfungsi untuk menemukan pola keterkaitan (*association rules*) dari kumpulan data besar dengan mengukur tingkat kemunculan dan kekuatan hubungan antar item menggunakan nilai *support* dan *confidence*. Pendekatan ini dinilai efektif karena mampu mengungkap hubungan tersembunyi antar variabel pembelajaran, termasuk pola kesalahan siswa, tanpa memerlukan data berlabel. Selain itu, pengembangan *Hybrid Apriori Algorithm* terbukti mempercepat proses pencarian aturan serta meningkatkan akurasi hasil analisis, sehingga relevan digunakan untuk mengidentifikasi pola kesalahan fonologis atau suku kata pada hasil tes literasi siswa sekolah dasar.

Algoritma Apriori dapat diimplementasikan dalam berbagai aplikasi, dan aturan asosiasi juga dapat digunakan untuk memantau perilaku pelanggan yang ditemukan di departemen ritel. Algoritma Apriori juga menemukan tren pelanggan berdasarkan kelompok barang yang sama yang dibeli (Adri et al., 2021). Algoritma Apriori adalah salah satu metode penambangan data (*data mining*) yang paling populer. Algoritma ini dapat menemukan set item yang sering muncul dari set data transaksi dan menurunkan aturan asosiasi. Aturan tersebut adalah pengetahuan yang ditemukan dalam basis data. Karena kombinasi yang eksplosif, tidak mudah untuk menemukan item yang sering muncul (item dengan frekuensi lebih besar dari atau sama dengan set item minimum yang didukung yang ditentukan oleh pengguna). Setelah memperoleh satu set item yang sering diperoleh, lanjutkan saja operasi dan hasilkan aturan asosiasi dengan keyakinan yang lebih tinggi dari atau sama dengan keyakinan minimum yang ditentukan

pengguna. Algoritma saat ini dalam basis data simulasi didefinisikan dalam artikel ini, dan aturan asosiasi dengan berbagai nilai keyakinan ditemukan(Nurcahyo et al., 2023).

Penambangan data disebut penemuan pengetahuan dalam basis data, yang merupakan ekstraksi non-sepele dari informasi yang sebelumnya tidak diketahui dan berpotensi berguna dari data dalam basis data (Aziz et al., 2023). Salah satu bidang penelitian yang paling signifikan dan menarik adalah penambangan data. Salah satu mesin pencari penambangan data yang paling umum digunakan adalah Basis Data Transaksi Hukum Pertambangan. Berdasarkan algoritma Apriori dan strategi algoritma yang ditingkatkan, banyak algoritma untuk menambang aturan asosiasi telah diusulkan, tetapi sebagian besar algoritma tidak memperhatikan struktur basis data (Yannuansa et al., 2021). Teknik yang direkomendasikan termasuk transfer ke basis data, dan teknik transposisi khusus ini telah ditingkatkan lebih lanjut. Metode ini akan dapat mengurangi jumlah pemindaian pengumpulan data dan kemudian menggunakan lebih sedikit waktu untuk membuat aturan asosiasi (Aziz, 2025). Tujuan yang ingin dicapai melalui penelitian ini adalah memperoleh pemahaman terkait tingkat kesalahan siswa sekolah dasar dalam membaca suku kata berdasarkan buku panduan tes membaca suku kata yang ada. Dengan demikian, tim INOVASI dapat mengembangkan rencana untuk menghasilkan buku panduan tes siswa yang lebih baik.

Menurut Zakur & Flaih, (2023) algoritma Apriori merupakan salah satu metode paling populer dalam teknik *association rule mining* yang digunakan untuk menemukan pola keterkaitan (*association rules*) dari kumpulan data yang besar melalui parameter utama berupa nilai *support* dan *confidence*. Nilai *support* menggambarkan frekuensi kemunculan *itemset* dalam dataset, sedangkan *confidence* menunjukkan kekuatan hubungan kondisional antar item yang diuji. Melalui pendekatan ini, peneliti dapat mengekstraksi hubungan tersembunyi antar variabel dalam data yang rumit dan beragam, mencakup aspek kesehatan, pendidikan, serta kebiasaan manusia dalam berinteraksi. Zakur & Flaih, (2023) menjelaskan bahwa algoritma Apriori memiliki keunggulan karena bersifat *unsupervised*, mampu menghasilkan aturan yang mudah dipahami, dan dapat diterapkan pada data yang tidak berlabel, sehingga fleksibel untuk berbagai konteks penelitian. Dalam konteks pendidikan, penerapan algoritma ini dapat membantu guru dan peneliti dalam mengidentifikasi pola kesalahan belajar siswa secara sistematis, misalnya dalam menganalisis kesalahan membaca atau pengucapan suku kata, untuk kemudian digunakan sebagai dasar penyusunan intervensi pembelajaran berbasis data.

Pola kesalahan membaca siswa bersifat sistematis dan berulang, terutama pada aspek penggabungan suku kata dan pelafalan yang kurang tepat. Pola semacam ini sangat relevan

untuk dianalisis menggunakan algoritma Apriori, yang mampu mengidentifikasi keterkaitan antar jenis kesalahan suku kata secara kuantitatif. Dengan demikian, hasil penelitian (Oktari & Dwianto, 2025) dapat dijadikan dasar empiris bagi penerapan *data mining* untuk menemukan aturan asosiasi (*association rules*) antara jenis kesalahan membaca siswa sekolah dasar.

Beberapa pendekatan telah diusulkan untuk mengatasi keterbatasan algoritma apriori dan meningkatkan kinerjanya. Prosedur yang disarankan, yang mengurangi operasi pemangkasan daftar objek kandidat, diperkenalkan. Alih-alih memindai seluruh set data, Anda dapat mengatasi masalah seperti data yang buruk atau data yang berlebihan, daripada mencari set data lengkap, ini berfokus pada pencarian relevansi dalam set data yang difilter, sehingga meningkatkan konsistensi data (Fanani et al., 2024). Dengan membagi set data menjadi partisi horizontal, teknik lain juga dapat digunakan untuk mengatasi pembatasan yang berlebihan pada ukuran penyimpanan *itemset* yang sering dan tidak sering. Dibandingkan dengan algoritma Apriori, rencana ini mengurangi ukuran setiap transaksi dan mengurangi jumlah waktu yang dibutuhkan untuk menyelesaikannya (Wijaya et al., 2022).

Menurut Butsianto et al., (2025) untuk menambang pola yang sering terjadi, mereka datang dengan algoritma yang disempurnakan. Algoritma ini menggunakan basis data yang ditransposisikan dan membuat beberapa modifikasi pada algoritma yang mirip dengan Apriori. Manfaat utama dari teknik yang diusulkan adalah bahwa basis data disimpan dalam bentuk yang ditransposisikan, dan basis data diurutkan dan dikurangi dalam setiap iterasi dengan membuat ID transaksi untuk setiap pola. Teknologi yang diusulkan menghemat banyak ruang komputasi dan menghemat banyak waktu di mesin. Hasil eksperimen diberikan dengan data sintetis, dan hasilnya dibandingkan dengan algoritma Apriori klasik. Dengan cara ini, teknik yang diusulkan sangat berguna dalam mendeteksi set model data yang besar(Noorrohman et al., 2025).

Menurut Mahfudzi Mabruri (2023) telah membandingkan hasil dengan algoritma Apriori tradisional ketika mencari pola-pola ini dalam basis data besar dengan data sintetis. Inilah sebabnya mengapa teknik yang diusulkan sangat berguna dalam menemukan model berdasarkan set data yang besar. Algoritma Apriori dapat menemukan tren pelanggan berdasarkan serangkaian produk yang sering dibeli. Banyak industri telah memasang perangkat lunak penambangan data yang sukses. Dalam industri ritel, penambangan data dapat digunakan dalam kegiatan pemasaran untuk menargetkan pelanggan yang menguntungkan berdasarkan imbalan. Jika teknologi penambangan data digunakan dalam kegiatan pasar, maka industri ritel

akan mendapatkan, mempertahankan, dan mencapai kesuksesan yang lebih besar di pasar yang sangat kompetitif ini (Rianti et al., 2023).

Menurut Yulianto et al., (2024) Mengusulkan pendekatan baru untuk mengurangi jumlah literasi yang dihasilkan oleh jumlah kandidat yang dipindai, sehingga mengatasi masalah pemborosan waktu dalam memindai seluruh basis data serta algoritma apriori. Sebelum membuat seri material total (Li), gunakan seri tersebut untuk mendapatkan nomor transaksi dengan akun pendukung terkecil antara material X dan Y. Ulangi langkah-langkah ini hingga tidak ada detail umum lainnya yang ditemukan. Untuk menghasilkan akun dukungan algoritma kandidat dengan hasil yang lebih optimal dan waktu pengerjaan yang lebih efisien. Tujuan dari hasil analisis diharapkan dapat memberikan gambaran empiris tentang pola kesalahan membaca dan mengucap suku kata, sehingga bermanfaat bagi guru dijadikan dasar dalam merancang strategi pembelajaran fonetik dan intervensi literasi yang lebih tepat sasaran. Dengan demikian, penerapan algoritma Apriori tidak hanya membantu memahami data kesalahan siswa secara kuantitatif, tetapi juga berkontribusi dalam meningkatkan efektivitas pembelajaran literasi di tingkat sekolah dasar.

2. METODE

Asosiasi Pertambangan *(Role of Mining Association)* berupaya menemukan aturan umum yang menentukan hubungan item-item yang sering muncul dalam basis data dan memiliki dua kriteria utama, yaitu poin dukungan dan kepercayaan (Aziz, 2025). Set item yang sering digunakan didefinisikan sebagai set item yang nilai dukungannya lebih besar atau sama dengan ambang batas minimum, dan aturan yang sering didefinisikan sebagai aturan yang nilai keyakinannya mencapai atau melampaui batas minimum. Ambang batas ini secara tradisional diasumsikan dapat dipakai untuk mengekstrak *item* yang kerap muncul. Pedoman Asosiasi Pertambangan adalah mengidentifikasi semua aturan yang memiliki tingkat dukungan dan kepercayaan lebih tinggi daripada tingkat dukungan dan kepercayaan minimum.

Algoritma Apriori yang diusulkan oleh R. Agrawal dkk. Pada tahun 1993, sebuah algoritma penambangan data satu dimensi dan satu lapis untuk aturan Asosiasi Boolean muncul. Ide utamanya adalah menggunakan metode rekursif untuk mencari lapis demi lapis. Oleh karena itu algoritma ini bisa lebih fokus pada suatu hal yang rekursif (Fey, 2023).

Algoritma Apriori adalah aturan asosiasi dalam penambangan data. Aturan yang menunjukkan hubungan antara beberapa properti biasanya disebut sebagai analisis kedekatan atau analisis keranjang pasar. Analisis Asosiasi atau Ekstraksi Aturan Asosiasi adalah teknik

penambangan data yang digunakan untuk menemukan aturan portofolio program. Pentingnya asosiasidapat ditentukan oleh dua tolok ukur, yaitu dukungan dan kepercayaan. Persentase campuran item dalam database ini disebut dukungan, dan kekuatan hubungan antara item dalam aturan kemitraan disebut kepercayaan.

2.1. Mendukung

Dukungan mode terkait adalah persentase portofolio proyek dalam basis data. Dukungan (A → B) = P (AUB) seperti pada persamaan 1. Hasilnya adalah rasio, karena rasio tersebut baik jika persentase populasinya baik. Rasio jumlah kejadian A dan B dalam semua kejadian yang diberikan dikenal sebagai Hukum A → B.

$$Support\ (A \rightarrow B) = \frac{\sum Transactions\ A\ and\ B}{\sum Transaction} \text{.....(1)}$$

2.2. Kepercayaan Diri

Kepercayaan aturan asosiasi merupakan ukuran keakuratan suatu aturan, yaitu penyajian transaksi dalam suatu database yang berisi A dan berisi B. *Database* tersebut yang nantinya akan terolah dengan nilai *Confidence.* Dengan keyakinan, frekuensi hubungan antara unsur-unsur hukum perkumpulan dapat dievaluasi. Keyakinan = P (B | A) seperti pada persamaan 2.

$$Confidence = \frac{\sum Transactions\ A\ and\ B}{\sum Transaction\ A} \text{.....(2)}$$

Bagian penelitian dari artikel ini adalah mengubah nilai dukungan minimum dan memberikan aturan asosiasi yang berbeda. Jika nilai dukungan minimum tinggi, aturan dapat difilter dengan lebih akurat. Konsep penambangan data pada asosiasi dan *itemset*, diinyatakan sebagai:

Sebuah A → B menyatakan Jika A benar, maka B juga benar.

Contohnya: “Jika Deepavali tiba, penjualan kembang api akan meningkat”.

A = Minggu Deepavali (festival cahaya India).

B = Penjualan kembang api meningkat.

Asosiasi Pertambangan Aturan ingin menemukan aturan umum yang menentukan Untuk setiap aturan, jika A →B→B→A, maka A dan B diberi nama "*itemset* yang menarik". Contohnya:

a. Orang membeli kaus kaki dan sepatu .
b. Orang membeli sepatu listrik dan sebagainya.

2.3. Transaksi Meja

Analisis keranjang belanja dilakukan untuk satu transaksi produk. Misalnya, pelanggan yang membeli kaus kaki lebih cenderung membeli sepatu. Pola hubungan tersebut dapat dinyatakan sebagai aturan asosiasi (Kaus Kaki → Sepatu).

2.4. Transaksi Basis Data

Mengumpulkan semua transaksi penjualan dikenal sebagai populasi. Populasi tersebut mewakili transaksi dalam suatu rekaman data. Setiap transaksi direpresentasikan oleh tupel data sebagaimana dijelaskan pada Tabel 1.

Tabel 1. Transaksi Data

Id	Transaksi
TX1	Dasi, Sepatu, Kaus Kaki
TX2	Sabuk, Kemeja, Sepatu, Dasi, Kaus Kaki
TX3	Dasi, Sepatu
TX4	Kaus Kaki, Sepatu, Ikat Pinggang

Aturan asosiasi Dasi → Kejutan menunjukkan bahwa Dasi adalah pendahulu dari aturan, sedangkan Kejutan merupakan hasil yang tak terelakkan dari aturan tersebut. Semua aturan asosiasi memiliki tingkat dukungan dan kepercayaan masing-masing. Persentase dukungan yang memenuhi suatu aturan, atau dengan kata lain dukungan dalam aturan R, merupakan nilai rasio.

3. HASIL DAN PEMBAHASAN

3.1 Analisis Data

Data yang digunakan dalam penelitian ini mencakup data utama tentang hasil tes siswa. Subjek penelitian ini adalah siswa sekolah dasar dari 4 kabupaten di Kalimantan Utara, yaitu Tanjung Pallas Timur, Tanjung Saleh, Mali Barat (Mali), dan Mary Marina Timur (Mary Marina Timur). Data ini diambil dari data tim INOVASI tahun 2018 dengan jumlah 200 data. Ini merupakan rencana kerja sama antara pemerintah Australia dan Indonesia. INOVASI bekerja sama secara langsung dengan Kementerian Pendidikan dan Kebudayaan untuk mencari cara meningkatkan hasil belajar siswa di sekolah-sekolah di berbagai wilayah Indonesia, terutama dalam hal literasi dan numerasi. Secara keseluruhan, temuan mengenai faktor penyebab kesalahan membaca menegaskan bahwa guru perlu berperan sebagai pendidik sekaligus diagnostik literasi. Guru tidak hanya mengajarkan cara membaca, tetapi juga menganalisis mengapa siswa melakukan kesalahan dan bagaimana memperbaikinya melalui intervensi yang tepat. Strategi pembelajaran yang diferensiatif, multisensori, dan berbasis data hasil evaluasi akan membantu meningkatkan kemampuan membaca siswa secara signifikan.

Tabel 2. Suku Kata

No	*Syllables*
1	Su
2	Sa
3	Hon
4	Gi
5	De
6	Dak
7	Yam
8	Po
9	Ma
10	Tu

Sumber: Data Tes dari tim INOVASI, 2018

Tahap 2 menerapkan metode dengan menggunakan algoritma Apriori untuk memprediksi tingkat kesalahan pada kata-kata yang sering salah diucapkan oleh siswa. Tujuan utama dari algoritma Apriori adalah memperoleh aturan asosiasi yang sesuai dengan batas minimum *support* dan *confidence* yang telah ditetapkan.

3.2 Hasil Tes Siswa SD

Penelitian ini menggunakan data hasil tes literasi membaca permulaan dari siswa sekolah dasar kelas awal. Data terdiri atas rekaman kesalahan pengucapan suku kata yang diidentifikasi selama proses membaca kata sederhana. Setiap kesalahan diubah menjadi item dalam dataset. Sebagai data acuan untuk analisis tingkat kesalahan suku kata yang sering salah dan untuk perbaikan di masa mendatang, seperti pada Tabel 3.

Tabel 3. Hasil Tes Siswa

No	*Student Test Results*
1	su sa hon gi de dak yam po ma tu
2	su sa hon gi de dak yam po ma tu
3	su sa hon gi de dak yam
…	….
198	hon de dak yam
199	hon dak yam
200	gi de yam po ma tu

3.3 Penentuan Itemset

Penentuan *itemset* dilakukan dengan mengacu pada nilai *support* yang dimiliki oleh setiap item. *Support* minimum = 50%. Nilai *support* dapat ditentukan dengan persamaan 3. Kemudian akan didapatkan nilai dukungan untuk setiap item. Setiap item yang ada terdapat nilai yang dibutuhkan pada pengolahan tahap berikutnya. Hasil dari pengolahan awal dapat dlihat pada Tabel 4.

$$Support\ (A) = \frac{\sum Syllables\ that\ contain\ value\ (A)}{\sum Syllables} \text{x } 100\% \quad \text{……..…..(3)}$$

Tabel 4 . Nilai Dukung Setiap Item

Itemset	*Support*
Su	48%
Sa	44%
Hon	76%
Gi	46%
De	74%
Dak	74%
Yam	70%
Po	46%
Ma	52%
Tu	46%

3.4 Kombinasi Dua Itemset

Dalam proses pembuatan kombinasi dua *itemset*, nilai *support* ditentukan berdasarkan perhitungan pada Tabel 4 dengan jumlah *support* minimum = 50%. Untuk menghitung nilai *support*, dapat menggunakan persamaan 4. Hasil dari proses pengolahan kombinasi dua *itemset* tersebut dapat dilihat pada Tabel 5.

$$Support\ (A,B) = \frac{\sum Syllables\ that\ contain\ value\ (A\ and\ B)}{\sum Syllables} \text{ x } 100\% \quad \ldots\ldots\ldots\ldots (4)$$

Tabel 5. Kombinasi Kandidat 2 *Itemset*

No	*Itemset*	*Amount*	*Support*
1	Hon, De	112	56%
2	Hon, Dak	136	68%
3	Hon, Yam	128	64%
4	Hon, Ma	96	48%
5	De, Dak	112	56%
6	De, Yam	96	48%
7	De, Ma	100	50%
8	Dak, Yam	132	66%
9	Dak, Ma	96	48%
10	Yam, Ma	96	48%

Dengan nilai dukungan minimum yang telah ditentukan, yaitu 50%. Kandidat yang tidak memenuhi persyaratan dukungan minimum untuk kombinasi 2 program akan dieliminasi. Setelah mendapatkan hasil dari eliminasi akan dibentuk kombinasi 2 program dapat dilihat pada Tabel 6.

Tabel 6. Kombinasi 2 *Itemset*

No	*Itemset*	*Amount*	*Support*
1	Hon, De	112	56%
2	Hon, Dak	96	68%
3	Hon, Yam	128	64%
4	De, Dak	112	56%
5	De, Ma	100	50%
6	Dak, Yam	132	66%

3.5 Kombinasi Tiga Itemset

Pada proses pembentukan kombinasi tiga *itemset*, perhitungan nilai *support* yang disajikan pada Tabel 6 didasarkan pada ketentuan nilai minimum *support* sebesar 50%. Nilai inilah yang nanti akan diolah. Adapun nilai *support* tersebut dapat dihitung menggunakan persamaan 5. Hasil dari proses pengolahan kombinasi tiga *itemset* tersebut dapat dilihat pada Tabel 7.

$$Support\ (A, B\ and\ C) = \frac{\sum Syllables\ that\ contain\ value\ (A, B\ and\ C)}{\sum Syllables} \text{x } 100\% \quad \text{.....(5)}$$

Tabel 7. Kombinasi Kandidat 3 *Itemset*

No	*Itemset*	*Amount*	*Support*
1	Hon, De, Dak	108	54%
2	Hon, De, Yam	128	64%
3	Hon, Dak, Yam	128	64%
4	Hon, De, Ma	100	50%
5	Hon, Dak, Ma	96	48%
6	Hon, Yam, Ma	92	46%

Dengan nilai dukungan minimum yang telah ditentukan, yaitu 50%. Kandidat yang tidak memenuhi persyaratan dukungan minimum untuk kombinasi tiga program akan dieliminasi. Setelah mendapatkan hasil dari eliminasi akan dibentuk kombinasi tiga program dapat dilihat pada Tabel 8.

Tabel 8. Kombinasi Tiga *Itemset*

No	*Itemset*	*Amount*	*Support*
1	Hon, De, Dak	108	54%
2	Hon, De, Yam	128	64%
3	Hon, Dak, Yam	128	64%
4	Hon, De, Ma	100	50%

3.6 Kombinasi Empat Itemset

Dalam proses pembuatan kombinasi empat *itemset*, nilai *support* ditentukan berdasarkan perhitungan pada Tabel 8 dengan jumlah *support* minimum Adalah 50%. Nilai inilah yang nanti akan diolah. Perhitungan nilai *support* dapat dilakukan dengan persamaan 6. Hasil dari proses pengolahan kombinasi empat *itemset* tersebut dapat dilihat pada Tabel 9.

$$Support\ (A, B, C\ and\ D) = \frac{\sum Syllables\ that\ contain\ value\ (\text{A, B, C and D})}{\sum Syllables} \text{x} 100\% \quad \text{.....(6)}$$

Tabel 9. Kombinasi Kandidat Empat *Itemset*

No	*Itemset*	*Amount*	*Support*
1	Hon, De, Dak, Yam	104	52%
2	Hon, De, Dak, Ma	68	34%
3	Hon, De, Yam, Ma	88	44%
4	Hon, Dak, Yam, Ma	72	36%

Aturan pertama memiliki *support* 52%, artinya 52% siswa yang salah mengucapkan suku kata Hon, De, dan Dak juga salah mengucapkan Yam. Nilai tersebut menunjukkan bahwa kesalahan pada tiga suku kata tersebut meningkatkan kemungkinan kesalahan pada suku kata Yam sebesar 52% dibandingkan kejadian acak. Pola ini mengindikasikan bahwa kesalahan siswa tidak terjadi secara terpisah, melainkan memiliki keterkaitan fonologis dan kognitif yang dapat dipetakan melalui algoritma Apriori. Dengan demikian, hasil teknis algoritma menunjukkan bahwa pola kesalahan membaca siswa bersifat sistematis dan dapat diidentifikasi secara kuantitatif melalui *association rule mining*.

Dengan nilai dukungan minimum yang telah ditentukan, yaitu 50%. Kandidat yang tidak memenuhi persyaratan dukungan minimum dari empat kombinasi program studi akan dieliminasi. Setelah mendapatkan hasil dari eliminasi akan terbentuk kombinasi empat set program studi dapat dilihat pada Tabel 10.

Tabel 10. Kombinasi 4 Itemset

No	*Itemset*	*Amount*	*Support*
1	Hon, De, Dak, Yam	104	52%

3.7 Penetapan Aturan Asosiasi

Jika kombinasi *itemset* dengan nilai tertinggi sudah ditentukan, langkah berikutnya adalah membuat aturan asosiasi dengan syarat keyakinan paling sedikit 50%. Perhitungan keyakinan untuk A → B → C dapat dilakukan dengan persamaan 7. Selanjutnya maka akan mendapatkan nilai keyakinan dari kombinasi empat *itemset* yang dapat dilihat pada Tabel 11.

$$Confident = \frac{\sum Syllables\ that\ contain\ value\ (A, B\ and\ C)}{\sum Syllables\ contain\ value\ A} \text{x } 100\% \quad \text{.....(7)}$$

Tabel 11. Aturan Asosiasi

Association Rules	*Support*	*Confident*
Jika siswa salah dalam suku kata Hon, De, Dak maka siswa juga salah dalam mengucapkan suku kata Yam	52%	68%

Hasil penelitian menunjukkan bahwa kesalahan membaca siswa bukanlah sesuatu yang acak, melainkan terstruktur dan dapat diprediksi. Dengan memahami keterkaitan antar kesalahan suku kata, guru dapat menanamkan kesadaran fonologis secara lebih efektif,

membantu siswa mengoreksi kesalahannya sendiri, dan mempercepat perkembangan kemampuan membaca permulaan. Dengan demikian hasil yang didapat bisa digunakan untuk mengukur seberapa mahir literasi siswa. Ada hubungan yang kuat antara kesalahan siswa dalam mengucapkan suku kata Hon, De, dan Dak dengan kesalahan dalam mengucapkan suku kata Yam. Dengan kata lain, kesalahan pada tiga suku kata pertama menjadi indikator kemungkinan kesalahan pada suku kata Yam.

4. PENUTUP

Simpulan dan Saran

Artikel ini difokuskan pada pencarian pola tersembunyi dari berbagai objek yang umum digunakan dengan memanfaatkan teknik penambangan data. Oleh karena itu, algoritma apriori dapat menemukan pola-pola ini dari basis data besar, yang memainkan peran penting, sehingga berbagai departemen dapat membuat atau menentukan keputusan bisnis yang lebih baik. Sebelumnya, algoritma itu sendiri mampu menemukan atau memperoleh tren kesalahan suku kata berdasarkan sejumlah item yang sering dikembalikan. Tiga dari sepuluh suku kata yang dibaca siswa mengandung tiga huruf yang berakhiran konsonan: Hon, Dak, dan Yam. Berdasarkan temuan penelitian, kesalahan rata-rata siswa terletak pada konsonan, yang dapat disiapkan oleh guru dalam rencana pengajaran dan alat peraga agar guru dapat memberikan kombinasi yang tepat. Temuan ini menunjukkan adanya pola sistematis dalam kesalahan pengucapan suku kata siswa. Kesalahan tersebut tidak bersifat acak, tetapi saling berkaitan dan dapat dijadikan indikator kemampuan fonologis siswa. Hubungan ini memiliki implikasi penting dalam pembelajaran literasi dasar, karena memperlihatkan bahwa peningkatan literasi harus dimulai dari penguatan kemampuan mengenali dan melafalkan bunyi bahasa dengan benar. Penelitian selanjutnya disarankan untuk memperluas objek dan variabel dengan melibatkan lebih banyak variasi suku kata agar pola hubungan kesalahan pengucapan dapat diketahui secara lebih menyeluruh. Selain itu, penelitian dapat dilakukan pada jenjang pendidikan dan kelompok usia yang berbeda guna melihat perkembangan kemampuan fonologis dan literasi siswa.

DAFTAR PUSTAKA

Adri, A., Rumlaklak, N. D., & Sina, D. R. (2021). Implementasi Algoritma Apriori Untuk Analisa Data Penjualan (Studi Kasus: Toko Ud. Suryani). *Jurnal Komputer Dan Informatika*, *9*(2), 182–188. https://doi.org/10.35508/jicon.v9i2.5132

Aziz, M. I. (2025). Implementasi Machine Learning Untuk Deteksi Anomali Kesehatan Janin Menggunakan Metode Ensemble Berbasis Decision Tree. *Jurnal Tecnoscienza*, *9*(2), 189–

198. https://doi.org/10.51158/zymd0m16

Aziz, M. I., Fanani, A. Z., & Affandy, A. (2023). Analisis Metode Ensemble Pada Klasifikasi Penyakit Jantung Berbasis Decision Tree. *Jurnal Media Informatika Budidarma*, *7*(1), 1–12. https://doi.org/10.30865/mib.v7i1.5169

Butsianto, S., Naya, C., & Muhammad, A. (2025). *BULLETIN OF COMPUTER SCIENCE RESEARCH Implementasi Algoritma Apriori dalam Menemukan Pola Asosiasi pada Data Penjualan Produk Retail*. *5*(5), 938–947. https://doi.org/10.47065/bulletincsr.v5i5.731

Fanani, A. J., Mutamaqin, M. I., & Aziz, M. I. (2024). Penerapan Strategi Pembelajaran Berbasis Masalah untuk Mengembangkan Kemampuan Berpikir Kritis dan Menurunkan Kecemasan Matematika pada Siswa SMA. *Dharma Pendidikan*, *19*(2), 156–163. https://doi.org/10.69866/dp.v19i2.532

Fey, F. P. (2023). Application of the Apriori Algorithm to Purchase Patterns. *The Indonesian Journal of Computer Science*, *12*(2), 553–561. https://doi.org/10.33022/ijcs.v12i2.3105

Mahfudzi Mabruri. (2023). *Implementasi Algoritma Apriori UntukMenentukan Rekomendasi Menu Pada RestoranPojok Indah Seafood Semarang*. *19*(x), 849–860.

Noorrohman, M. O., Aziz, M. I., Azhar, S., Pradana, S., Yulianto, R., Ratnasari, D., La, F., Ifanka, V., Asyiqien, M. Z., Teknologi, I., Mahrusiyah, A., & Kediri, K. (2025). *Monitoring Dashboard Using Linear Regression for Employee Performance*. *11*(11), 3149–3158.

Nurcahyo, R., Fanani, A. Z., Affandy, A., & Aziz, M. I. (2023). Peningkatan Algoritma C4. 5 Berbasis PSO Pada Penyakit Kanker Payudara. *Jurnal Media Informatika Budidarma*, *7*(4), 1758–1765. https://doi.org/10.30865/mib.v7i4.6841

Oktari, D. P., & Dwianto, A. (2025). Analisis Kesulitan Membaca Siswa Kelas 1 SDN Sukamaju Way Kanan Lampung. *Epistema*, *5*(2), 135–146. https://doi.org/10.21831/ep.v5i2.83102

Rianti, A., Majid, N. W. A., & Fauzi, A. (2023). CRISP-DM: Metodologi Proyek Data Science. *Prosiding Seminar Nasional Teknologi Informasi Dan Bisnis (SENATIB)*, 107–114. https://ojs.udb.ac.id/index.php/Senatib/article/view/3015

Wijaya, H. O. L., S, A. A. T., Armanto, A., & Sari, W. M. (2022). Prediksi Pola Penjualan Barang pada UMKM XYZ dengan Metode Algoritma Apriori. *Jurnal Sistem Komputer Dan Informatika (JSON)*, *3*(4), 432. https://doi.org/10.30865/json.v3i4.4200

Yannuansa, N., Safa'udin, M., & Aziz, M. I. (2021). Pemanfaatan Algoritma K-Means Clustering dalam Mengolah Pengaruh Hasil Belajar Terhadap Pendapatan Orang Tua Pada Mata Pelajaran Produktif. *Jurnal Tecnoscienza*, *6*(1), 43–55. https://doi.org/10.51158/tecnoscienza.v6i1.530

Yulianto, S. P. R., Fanani, A. Z., Affandy, A., & Aziz, M. I. (2024). Analisis Metode Smoote pada Klasifikasi Penyakit Jantung Berbasis Random Forest Tree. *Jurnal Media*

Informatika Budidarma, *8*(3), 1460. https://doi.org/10.30865/mib.v8i3.7712

Zakur, Y., & Flaih, L. (2023). Apriori Algorithm and Hybrid Apriori Algorithm in the Data Mining: A Comprehensive Review. *E3S Web of Conferences*, *448*. https://doi.org/10.1051/e3sconf/202344802021