

Analisis Pola Diabetes Melitus Menggunakan Algoritma Apriori

Ade Yuliana¹, Fifi Devianti²

^{1,2} Program Studi Teknik Informatika- Politeknik TEDC Bandung

Jl. Politeknik-Pesantren KM2 Cibabat Cimahi Utara – Cimahi Jawa Barat - Indonesia

yulianaad@poltektedc.ac.id, fidvnty@gmail.com

Abstrak— Indonesia termasuk dalam 10 besar negara dengan jumlah penderita diabetes melitus tertinggi di dunia. Di Kota Cimahi, Puskesmas Cigugur Tengah juga mencatatkan kasus diabetes melitus yang cukup tinggi. Namun, pengelolaan data masih terbatas pada analisis sederhana, yang membuat kebijakan penanggulangan kurang efektif. Karena itu, dibutuhkan metode yang lebih baik agar kebijakan yang diambil lebih tepat sasaran dan berdampak nyata dalam mengurangi penyakit ini. Penelitian ini bertujuan untuk mengidentifikasi pola gejala pada penderita Diabetes Melitus (DM) di Puskesmas Cigugur Tengah menggunakan algoritma Apriori. Data yang digunakan terdiri dari 516 rekam medis pasien dengan 10 atribut gejala, seperti sering kencing, cepat lapar, kesemutan, dan usia ≥ 45 tahun. Melalui teknik data mining dengan nilai minimum support 3%, penelitian ini berhasil menemukan 5 (lima) kandidat itemset anamnesa. Berdasarkan analisis data dengan algoritma Apriori di RapidMiner, kombinasi gejala "Cepat Lapar," "Sering Kencing," dan "Kesemutan" memiliki frekuensi kemunculan tinggi (0.723) dan menjadi pola dominan dalam deteksi awal diabetes melitus, dengan gejala tunggal "Cepat Lapar" (0.959) dan "Sering Kencing" (0.864) sebagai indikator utama. Kombinasi ini, terutama jika ditambah gejala "Lemas" (support 0.516), menunjukkan keterkaitan yang kuat, sedangkan gejala "Sering Haus" dan usia ≥ 45 tahun berkontribusi lebih rendah pada pola ini. Kombinasi "Lemas" dan "Sering Haus" memiliki nilai lift tertinggi (1.383), menunjukkan hubungan signifikan dengan gejala lainnya, sehingga menguatkan potensi prediksi kondisi kesehatan terkait diabetes melitus.

Kata Kunci— Diabetes Melitus, Algoritma Apriori, RapidMiner, Pola Gejala, Puskesmas.

Abstract— Indonesia is included in the top 10 countries with the highest number of diabetes mellitus sufferers in the world. In Cimahi City, the Cigugur Tengah Health Center also recorded quite high cases of diabetes mellitus. However, data management is still limited to simple analysis, which makes response policies less effective. Therefore, better methods are needed so that the policies taken are more targeted and have a real impact in reducing this disease. This study aims to identify symptom patterns in Diabetes Mellitus (DM) patients at Cigugur Tengah Health Center using the Apriori algorithm. The rising prevalence of DM calls for in depth data analysis to enhance treatment and prevention strategies. This study analyzed 516 patient medical records with 10 symptom attributes, including frequent urination, constant hunger, tingling sensations, and age ≥ 45 years. Through data mining with a minimum support of 3%, this research identified 5 (five) symptom candidate itemsets. Based on the analysis of data from 516 patients

using the Apriori algorithm in RapidMiner, the combination of symptoms "Always Hungry," "Frequent Urination," and "Tingling" has a high occurrence frequency (0.723) and stands out as a dominant pattern for early detection of diabetes mellitus, with individual symptoms like "Always Hungry" (0.959) and "Frequent Urination" (0.864) serving as primary indicators. This combination, especially when "Fatigue" is added (support 0.516), shows strong correlation, while symptoms like "Always Thirsty" and age ≥ 45 contribute less to this pattern. The combination of "Fatigue" and "Always Thirsty" has the highest lift value (1.383), indicating a significant association with other symptoms, thus strengthening the predictive potential for diabetes-related health conditions.

Keywords: Diabetes Mellitus, Apriori Algorithm, RapidMiner, Symptom Patterns, Community Health Center

I. PENDAHULUAN

Diabetes Melitus atau DM adalah suatu penyakit kronis yang ditandai oleh kadar glukosa darah yang melebihi nilai normal secara menahun. Sebutan glukosa darah sering dikenal oleh masyarakat dengan gula darah. Penyakit diabetes melitus terbagi menjadi beberapa tipe utama, yaitu diabetes tipe 1 yang disebabkan karena tidak adanya produksi insulin, diabetes tipe 2 disebabkan akibat kurangnya efektifitas kerja insulin, diabetes gestasional yang terjadi saat kehamilan dan diabetes tipe lainnya yang disebabkan oleh obat, faktor penyakit lain dan sebagainya. Gejala yang sering dialami pasien diabetes melitus seperti sering kencing, cepat lapar, sering haus, kesemutan luka sulit sembuh, cepat lelah dan lain-lain. Faktor resiko diabetes melitus terbagi menjadi yang tidak dapat diubah seperti usia lebih dari 40 tahun, riwayat keluarga, kehamilan dengan gula darah tinggi, ibu riwayat melahirkan bayi dengan berat lebih dari 4 kg dan bayi dengan berat badan lahir kurang dari 2,5 kg (Kemenkes RI, 2019).

Menurut Survei Kesehatan Indonesia tahun 2023, prevalensi penderita diabetes melitus berdasarkan diagnosis dokter pada penduduk semua umur telah mencapai 877.531 pasien. Penderita diabetes melitus tertinggi berada di Provinsi Jawa Barat dengan total pasien sebanyak 156.977 jiwa dan berdasarkan diagnosis dokter pada penduduk usia diatas 15 tahun sebanyak 114.619 pasien (Kemkes, 2023). Berdasarkan literatur, prevalensi penderita diabetes melitus (DM) di kota Cimahi dari tahun 2020 hingga 2021 mengalami peningkatan, dari 9.716 kasus pada tahun 2020 menjadi 10.108 kasus pada

tahun 2021. Prevalensi tertinggi DM di Kota Cimahi terdapat di Puskesmas Cigugur Tengah, dengan 1.173 kasus pada tahun 2020 dan 1.235 kasus pada tahun 2021 (Ahmad et al., 2024).

Berdasarkan observasi awal peneliti di Puskesmas Cigugur Tengah selama 2 (dua) hari, terdapat 425 kasus penyakit Diabetes Melitus dalam periode Januari hingga Mei 2024, yang terdiri dari 4 (empat) jenis penyakit diabetes melitus yaitu diabetes tipe 1, diabetes tipe 2, diabetes gestational dan diabetes tipe lainnya. Peningkatan prevalensi diabetes ini memerlukan perhatian khusus karena berdampak signifikan pada kesehatan masyarakat dan sistem kesehatan secara keseluruhan. Namun, belum adanya pengolahan data penyakit diabetes melitus yang memadai menyebabkan Puskesmas Cigugur Tengah masih menggunakan analisis sederhana dalam mengambil kebijakan penanganan penyakit ini. Penanganan khusus dalam pengelolaan data penyakit diabetes melitus sangat diperlukan agar kebijakan yang diambil untuk upaya penanggulangan penyakit ini bisa lebih efektif dan tepat sasaran.

Dalam hal ini, Teknik pengelolaan data seperti data mining dapat menyelesaikan permasalahan diatas. Data Mining bertujuan untuk menjelaskan, menkonfirmasi dan mengeksplorasi suatu penelitian. Data mining memiliki beberapa Teknik yang sering digunakan seperti *clustering*, *classification*, *association*, *regression* dan lain-lain (Medina, 2021).

Pada penelitian ini peneliti menggunakan Teknik *association rule* khususnya algoritma Apriori untuk menentukan pola penyakit diabetes melitus. Algoritma Apriori merupakan algoritma yang sangat terkenal untuk menemukan pola frekuensi tinggi. Pola frekuensi tinggi adalah pola-pola item di dalam suatu database yang memiliki frekuensi atau *support* di atas ambang batas tertentu yang disebut dengan istilah minimum *support*. Pola frekuensi tinggi ini digunakan untuk menyusun aturan asosiasi dan juga beberapa teknik data mining lainnya (Rahayu et al., 2024).

Peneliti mengidentifikasi faktor-faktor yang berkaitan dengan munculnya diabetes melitus menggunakan data primer dan data sekunder. Dimana data primer berupa data rekam medis dari Puskesmas Cigugur Tengah periode April hingga September 2024.

Setelah pelaksanaan penelitian, hasil dari penelitian ini dapat memiliki berbagai aplikasi praktis, mulai dari sosialisasi dan penyuluhan kesehatan, pembuatan kebijakan kesehatan, hingga penelitian epidemiologi lebih lanjut, memberikan dampak yang luas dan signifikan.

Berdasarkan uraian latar belakang tersebut, peneliti tertarik mengangkat topik penelitian yang berjudul “Analisis Pola Penyakit Diabetes Melitus Menggunakan Algoritma Apriori”.

II. METODE PENELITIAN

Terdapat 8 (delapan) tahapan yang dilakukan penelitian ini, diantaranya:

A. Menentukan Masalah

Dalam hal ini, penulis mengidentifikasi bahwa pada Puskesmas Cigugur Tengah terdapat peningkatan penyakit diabetes melitus dan belum adanya pengelolaan atau pencegahan terhadap penyakit tersebut. Identifikasi permasalahan menjadi langkah awal dalam penelitian ini, yang telah diuraikan dalam latar belakang mengenai penggunaan algoritma apriori untuk mengungkap pola penyakit diabetes melitus di Puskesmas Cigugur Tengah.

B. Menentukan Tujuan dan Ruang Lingkup

Menentukan tujuan dan ruang lingkup adalah apa yang akan dicapai oleh penelitian serta batasan-batasan yang penulis lakukan agar penelitian tidak terlalu luas dan menjadi lebih efektif. Tahap ini telah tercantum dalam tujuan dan manfaat penelitian serta batasan masalah (Sugianto & Sukmawati, 2023).

C. Mencari Literatur

Mencari literatur dilakukan untuk mencari berbagai sumber yang berkaitan tentang penyakit diabetes, data mining, association rule, dan algoritma apriori. Metode yang digunakan yaitu, studi Pustaka dengan mencari data/informasi melalui sumber buku, internet, jurnal yang berkaitan dengan penelitian yang dilakukan. Data-data tersebut selanjutnya dikumpulkan oleh penulis dengan menggunakan metode studi Pustaka dan membentuk beberapa konsep sebagai berikut :

1. Konsep diabetes melitus
2. Konsep data mining
3. Konsep association rule
4. Konsep Apriori
5. Konsep RapidMiner

D. Pengumpulan Data

Penulis akan menggunakan beberapa metode pengumpulan data sebagai berikut :

a. Data Primer

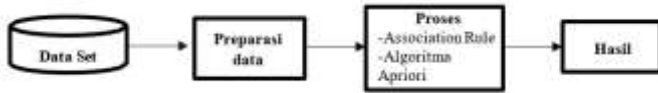
Teknik yang digunakan dalam pengumpulan data primer yaitu observasi. Dalam Penelitian ini peneliti melakukan observasi selama 2 (dua) hari yaitu pada tanggal 23 Juni – 24 Juni 2024 di Puskesmas Cigugur Tengah Jl. Abdul Halim No.199, Karangmekar, Kec. Cimahi Tengah, Kota Cimahi, Jawa Barat 40522. Pada penelitian tersebut peneliti memperoleh data primer yang berupa data rekam medis periode April – September 2024 sebanyak 516 data.

b. Data Sekunder

Dalam konteks penelitian ini, data sekunder yang digunakan mencakup informasi yang dikumpulkan dari berbagai sumber terpercaya, termasuk jurnal ilmiah, makalah akademis, artikel-artikel dari berbagai publikasi, serta buku yang berfokus pada topik terkait. Khususnya, sumber-sumber tersebut berkaitan dengan bidang data mining, aturan asosiasi (association rule), dan algoritma apriori, yang relevan dan mendukung tujuan serta kerangka penelitian ini.

E. Metode yang Digunakan

Pada tahap ini peneliti akan menjelaskan tentang proses dalam pengelolaan data pada penelitian ini melalui gambar 1. Berikut adalah proses yang dimaksud:



1. Dataset

Data set yang digunakan dalam penelitian ini adalah data rekam medis pasien penyakit diabetes melitus yang berjumlah 516 data dari 4 jenis penyakit diabetes melitus dari periode April sampai dengan September 2024 yang diambil langsung dari Puskesmas Cigugur Tengah. Pengambilan data rekam medis tersebut berupa file excel yang berisi tanggal daftar, nomor RM, NIK, nama, tanggal lahir, usia, jenis kelamin, golongan darah, kode ICD, diagnosa, kunjungan, kasus, poli, alamat, RT, RW, kecamatan, kelurahan, nomor BPJS, pembayaran, systole, diastole, tekanan nadi, respirasi, suhu, berat badan, tinggi badan, keluhan dan anamnesa. Data set tersebut selanjutnya akan dilakukan preparasi untuk menghilangkan atribut-atribut yang tidak digunakan untuk penelitian ini. Untuk atribut yang digunakan dalam penelitian ini yaitu nomor RM dan anamnesa.

2. Data Preparasi

a. Data Awal

Data rekam medis pasien ada 28 atribut yang belum dilakukan preparasi data seperti tanggal daftar, NIK, nama, tanggal lahir, jenis kelamin, golongan darah, kode ICD, diagnosa, kunjungan, kasus, poli, alamat, RT, RW, kecamatan, kelurahan, nomor BPJS, pembayaran, systole, diastole, tekanan nadi, respirasi, suhu, berat badan, tinggi badan dan keluhan.

b. Hasil Preparasi

Setelah data dipersiapkan, 28 atribut awal dipilih menjadi 3 atribut utama yang digunakan untuk analisis pola frekuensi tinggi. Atribut-atribut tersebut adalah nomor RM, usia dan anamnesa. Pemilihan atribut ini didasarkan pada dampaknya terhadap pola penyakit diabetes melitus. Adapun 3 (tiga) atribut tersebut yang sudah di preparasi data dapat dilihat pada tabel 1 sebagai berikut.

TABEL I
KATEGORI 3 ATRIBUT DATA REKAM MEDIS PASIEN

No.	Atribut	Keterangan
1	Nomor Rekam Medis	Berisikan nomor rekam medis pasien
2	Usia	Berikan usia pasien
3	Anamnesa	Berisikan gejala yang dialami pasien untuk tindakan asuhan

atribut yang digunakan dari data tersebut berupa atribut nomor rekam medis, usia dan anamnesa. Anamnesa berisi gejala klinis (sering kencing, cepat lapar, sering haus, lemas dan berat badan turun drastis) dan gejala tambahan (usia ≥ 45 tahun, luka sulit sembuh, mata kabur, kesemutan). *Field-field* pada tabel 2 akan peneliti gunakan sebagai indikator analisis untuk hasil penelitian.

TABEL II
TABEL YANG AKAN DIGUNAKAN UNTUK ANALISIS MENGGUNAKAN RAPIDMINER

ID	Usia ≥ 45 Tahun	Sering Kencing	Cepat Lapar	Sering Haus	Lemas	Berat Badan Turun Drastis	Luka Sulit Sembuh	Mata Kabur	Kesemutan
P1	1	1	1	0	1	0	0	0	1
P2	1	1	1	0	1	0	0	0	1
P3	1	1	1	0	1	0	0	0	1
P4	1	1	1	0	0	0	0	0	0
P5	0	0	0	0	0	0	0	0	0
P6	1	1	1	0	1	0	0	0	1
-	-	-	-	-	-	-	-	-	-
P516	1	0	0	0	0	0	0	0	0

Keterangan :
0 = Tidak Mengalami
1 = Ya Mengalami

c. Proses

Pada penelitian ini, peneliti akan menggunakan salah satu aplikasi data mining berupa *RapidMiner*. Peneliti akan menggunakan aplikasi tersebut untuk melakukan proses penentuan pola penyakit diabetes melitus. Untuk algoritma, peneliti menggunakan algoritma Apriori untuk menemukan pola frekuensi tinggi. tahapan algoritma apriori sebagai berikut:

- 1) Pembentukan kandidat itemset, kandidat k-itemset dibentuk dari kombinasi (k- 1)-itemset yang didapat dari iterasi sebelumnya dapat membentuk kandidat itemset.
- 2) Perhitungan support dari tiap kandidat k-itemset. Untuk mengukur jumlah transaksi yang mempunyai item, dibutuhkan support dari tiap-tiap kandidat yang didapatkan dengan meneliti database yang akan digunakan. Cara mencari support dapat dilakukan menggunakan perhitungan pada persamaan (1) dan persamaan (2).
- 3) Analisis pola frekuensi tinggi. Pola frekuensi tinggi ditetapkan dari kandidat k- itemset yang supportnya lebih besar dari minimum support.
- 4) Bila tidak didapat pola frekuensi tinggi tidak didapatkan lagi, maka keseluruhan proses akan dihentikan . dan jika tidak, maka k harus ditambah 1 kemudian proses kembali lagi ke tahap 1.

d. Hasil

Hasil yang akan didapatkan dalam penelitian ini berupa pola kombinasi itemset dan rules dari penyakit diabetes melitus dari pengujian dengan menggunakan aplikasi *RapidMiner*. Pola kombinasi yang paling tinggi supportnya dan pola kombinasi dengan item paling banyak. Hasil dari penelitian yang menggunakan algoritma Apriori untuk analisis pola penyakit diabetes melitus biasanya melibatkan penemuan pola-pola asosiasi yang signifikan antara faktor-faktor risiko, gejala, atau karakteristik pasien.

F. Proses dan Hasil Penelitian

Pada penelitian kali ini peneliti menggunakan metode eksperimen. Peneliti menggunakan metode untuk mengetahui pengaruh algoritma apriori terhadap menentukan pola penyakit diabetes melitus dimana nantinya akan ada pengujian – pengujian terhadap data untuk mendapatkan nilai support dan confidence terbaik pada data tersebut.

Penelitian ini menggunakan perangkat lunak RapidMiner dan algoritma apriori untuk mengolah data. Melalui RapidMiner, peneliti menghitung dan menganalisis nilai support dan confidence dari data yang diolah.

G. Analisis

Analisis ini bertujuan untuk memberikan rekomendasi yang berbasis data mengenai langkah-langkah yang dapat diambil untuk menanggulangi atau meminimalkan dampak penyakit diabetes melitus di wilayah yang sedang diteliti.

H. Penarikan Kesimpulan

Pada penelitian ini penarikan kesimpulan akan menyimpulkan hasil dari proses penentuan pola penyakit diabetes melitus menggunakan algoritma apriori yang telah dipilih.

III. HASIL DAN PEMBAHASAN

A. Hasil Penelitian

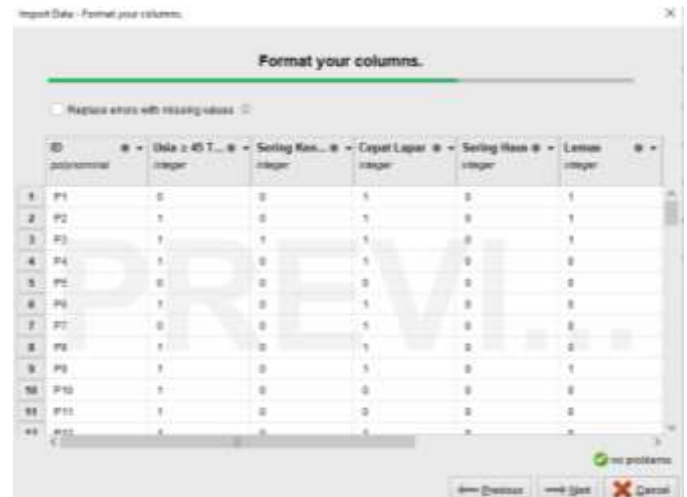
Bagian hasil penelitian mencakup tiga tahap, yaitu pengolahan data awal, pengujian algoritma Apriori berdasarkan frekuensi, serta hasil analisis data menggunakan algoritma Apriori.

1) Proses pengolahan data awal

Proses pengelolaan data awal pada pengujian ini menggunakan perangkat lunak RapidMiner. Tahapan ini dilakukan dengan memasukkan data yang sudah dipreparasi sebelumnya dengan cara import data pada perangkat lunak RapidMiner. Datanya berupa data rekam medis pasien penyakit diabetes melitus yang berjumlah 516 data dari 4 jenis penyakit diabetes melitus dari periode April-September 2024. Data tersebut memiliki 28 atribut yang belum dilakukan preparasi data seperti tanggal daftar, NIK, nama, tanggal lahir, jenis kelamin, golongan darah, kode ICD, diagnosa, kunjungan, kasus, poli, alamat, RT, RW, kecamatan, kelurahan, nomor BPJS, pembayaran, systole, diastole, tekanan nadi, respirasi, suhu, berat badan, tinggi badan dan keluhan. Lalu akan dipreparasi agar dapat diolah pada RapidMiner, dimana dari 28 atribut awal dipreparasi sehingga dipilih hanya 3 atribut utama yang digunakan untuk analisis pola frekuensi tinggi. Atribut-atribut tersebut adalah nomor RM, usia dan anamnesa.

Berdasarkan data atribut yang telah dipreparasi, atribut yang berjumlah 3 akan diuraikan kembali hingga menjadi 10 atribut. Atribut yang diurai adalah anamnesa yaitu menjadi gejala klinis (sering kencing, cepat lapar, sering haus, lemas dan berat badan turun drastis) dan gejala tambahan (luka sulit sembuh, mata kabur, kesemutan). Selain itu, usia jika

dikelompokkan menjadi dibawah dan diatas 45 tahun. Selanjutnya akan disesuaikan formatnya agar dapat diolah pada RapidMiner, yaitu Nomor RM (ID) dengan format Polynominal, lalu Usia ≥ 45 Tahun, Sering Kencing, Cepat Lapar, Sering Haus, Lemas, Berat Badan Turun Drastis, Luka Sulit Sembuh, Mata Kabur dan Kesemutan dengan format Binominal. Untuk lebih jelasnya dapat dilihat pada gambar 1 berikut.



Gbr. 1 Format Data

Adapun data preparasi yang telah diimport pada RapidMiner dapat dilihat pada gambar 2 berikut.



Gbr. 2 Preparasi data yang berhasil diimport

2) Pengujian Algoritma Apriori dengan Frekuensi

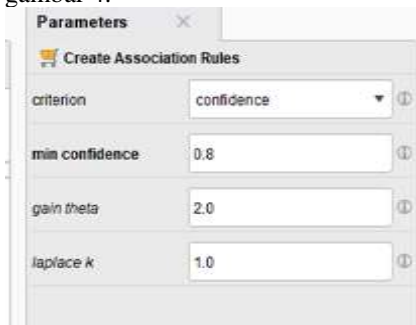
Setelah proses tersebut penulis membutuhkan pengujian data menggunakan perangkat lunak RapidMiner untuk untuk menentukan frekuensi pola penyakit diabetes melitus dengan asosiasi menggunakan algoritma Apriori dengan memasukkan data preparasi agar data tersebut dapat digunakan sebagai data yang akan diolah atau diuji. Selanjutnya mengimport data preparasi tadi ke dalam RapidMiner. Kemudian masukkan data preparasi tersebut ke dalam halaman process. Lalu mencari discretize by frequency, FP-Growth (Associations) dan create association rules di operator dan tarik ke halaman process. Kemudian sambungkan data preparasi ke discretize by frequency, lalu discretize by frequency ke numerical to

binominal, kemudian numerical to binominal ke FP-Growth (Associations), selanjutnya FP-Growth (Associations) ke create association rules, dan terakhir create association rules sambungkan ke result sehingga hasil dapat dilihat pada Gambar 3.



Gbr. 3 Pengujian Algoritma Apriori Dengan Frekuensi

Kemudian menyesuaikan parameter dengan criterion yaitu confidence dan min. confidence yaitu 0,8 atau 80% Dapat dilihat pada gambar 4.



Gbr. 4 Penyesuaian Nilai Minimum Confidence

Lalu menyesuaikan parameter untuk penentuan item set nya, dengan mengatur nilai min. support sebesar 0.03 atau 3%. Dapat dilihat pada gambar 5.



Gbr.5 Penyesuaian Nilai Minimum Support

3) Hasil Pengujian Data dengan Algoritma Apriori

Pengujian algoritma apriori menggunakan RapidMiner dapat dilakukan setelah proses data preparasi selesai, memastikan data yang akan digunakan sudah memenuhi persyaratan algoritma asosiasi. Dalam tahap ini, semua operator yang dibutuhkan harus dihubungkan secara sistematis dan teliti, satu per satu, untuk menjamin aliran proses berjalan tanpa kesalahan. Selain itu, penyesuaian format data serta parameter asosiasi perlu dilakukan agar setiap langkah dapat dieksekusi dengan akurat sesuai dengan kebutuhan pengujian. Setiap parameter dan operator berperan penting dalam memastikan hasil pengujian dapat diinterpretasikan dengan tepat, khususnya dalam konteks analisis frekuensi item.

Setelah proses pengaturan selesai dan algoritma diproses, output yang dihasilkan berupa data frekuensi setiap item yang terekam dalam dataset. Frekuensi per item ini sangat berguna untuk memahami pola umum yang sering muncul dalam data, yang kemudian divisualisasikan pada gambar 6 sebagai hasil analisis awal. Selain itu, asosiasi antar item akan menghasilkan aturan-aturan atau association rules yang dapat diamati pada gambar 7.

Item	Frequency
Apple	1000
Banana	1000
Orange	1000
Apple, Banana	1000
Apple, Orange	1000
Banana, Orange	1000
Apple, Banana, Orange	1000

Gbr.6 Hasil Data Frekuensi Per Item

Rule	Support	Confidence	Lift
Apple > Banana	1000	1000	1.000
Apple > Orange	1000	1000	1.000
Banana > Orange	1000	1000	1.000
Apple, Banana > Orange	1000	1000	1.000

Gbr.7 Hasil Data Frekuensi Asosiasi Rules

B. Pembahasan

Bagian ini terdiri dari dua tahap pembahasan, yaitu analisis pola dengan frekuensi tinggi, serta pembentukan dan evaluasi pemodelan aturan asosiasi. Fokus utama pembahasan adalah hasil pengolahan data menggunakan RapidMiner.

1) Analisis Pola Frekuensi Tinggi

Analisis pola frekuensi tinggi pada algoritma Apriori bertujuan untuk menemukan asosiasi atau hubungan antar

item yang sering muncul secara bersamaan dalam data. Dalam konteks Analisis Pola Penyakit Diabetes Melitus, algoritma Apriori digunakan untuk mengidentifikasi kombinasi faktor risiko atau gejala yang sering muncul secara bersamaan pada pasien diabetes.

Berdasarkan data 516 pasien dengan 10 atribut yang telah diuji pada RapidMiner dengan menyesuaikan nilai minimum support sebesar 0,03 atau 3% dan nilai minimum confidence sebesar 0,8 atau sebesar 80%, maka membentuk 5 kandidat itemset mengenai anamnesanya.

Adapun hasil kandidat k-1 (1 itemset) dengan nilai minimum support sebesar 0,03 atau 3% setelah diuji pada RapidMiner adalah menghasilkan 6 jenis anamnesa. Dapat dilihat pada tabel 3.

TABEL II
HASIL KANDIDAT K-1 (1 ITEMSET)

Item Set	Jumlah	Support
Cepat Lapar	21	0,959
Sering Kencing	70	0,864
Kesemutan	89	0,828
Lemas	145	0,719
Usia ≥ 45 Tahun	343	0,335
Sering Haus	8	0,016

Hasil kandidat k-2 (2 itemset) dengan nilai minimum support sebesar 0,03 atau 3% dan nilai minimum confidence sebesar 0,8 atau sebesar 80% setelah diuji pada RapidMiner adalah menghasilkan 14 jenis anamnesa. Dapat dilihat pada tabel 4.

TABEL IV
HASIL KANDIDAT K-2 (2 ITEMSET)

Item Set	Jumlah	Support
Cepat Lapar, Sering Kencing	11	0,845
Cepat Lapar, Kesemutan	10	0,806
Cepat Lapar, Lemas	14	0,705
Cepat Lapar, Usia ≥ 45 Tahun	20	0,333
Cepat Lapar, Sering Haus	0	0,016
Sering Kencing, Kesemutan	25	0,740
Sering Kencing, Lemas	21	0,624
Sering Kencing, Usia ≥ 45 Tahun	56	0,308
Sering Kencing, Sering Haus	1	0,014
Kesemutan, Lemas	25	0,595
Kesemutan, Usia ≥ 45 Tahun	66	0,291
Kesemutan, Sering Haus	2	0,012
Lemas, Usia ≥ 45 Tahun	123	0,293
Lemas, Sering Haus	4	0,008

Hasil kandidat k-3 (3 itemset) dengan nilai minimum support sebesar 0,03 atau 3% dan nilai minimum confidence sebesar 0,8 atau sebesar 80% setelah diuji pada RapidMiner adalah menghasilkan 16 jenis anamnesa. Dapat dilihat pada tabel 5.

TABEL V
HASIL KANDIDAT K-3 (3 ITEMSET)

Item Set	Jumlah	Support
Cepat Lapar, Sering Kencing, Kesemutan	9	0,723

Item Set	Jumlah	Support
Cepat Lapar, Sering Kencing, Lemas	8	0,616
Cepat Lapar, Sering Kencing, Usia ≥ 45 Tahun	10	0,308
Cepat Lapar, Sering Kencing, Sering Haus	0	0,014
Cepat Lapar, Kesemutan, Lemas	8	0,585
Cepat Lapar, Kesemutan, Usia ≥ 45 Tahun	10	0,289
Cepat Lapar, Kesemutan, Sering Haus	0	0,012
Cepat Lapar, Lemas, Usia ≥ 45 Tahun	13	0,293
Cepat Lapar, Lemas, Sering Haus	0	0,008
Sering Kencing, Kesemutan, Lemas	13	0,523
Sering Kencing, Kesemutan, Usia ≥ 45 Tahun	19	0,275
Sering Kencing, Kesemutan, Sering Haus	1	0,012
Sering Kencing, Lemas, Usia ≥ 45 Tahun	18	0,271
Sering Kencing, Lemas, Sering Haus	1	0,008
Kesemutan, Lemas, Usia ≥ 45 Tahun	23	0,252
Kesemutan, Lemas, Sering Haus	2	0,008

Hasil kandidat k-4 (4 itemset) dengan nilai minimum support sebesar 0,03 atau 3% dan nilai minimum confidence sebesar 0,8 atau sebesar 80% setelah diuji pada RapidMiner adalah menghasilkan 9 jenis anamnesa. Dapat dilihat pada tabel 6.

TABEL VI
HASIL KANDIDAT K-4 (4 ITEMSET)

Item Set	Jumlah	Support
Cepat Lapar, Sering Kencing, Kesemutan, Lemas	7	0,516
Cepat Lapar, Sering Kencing, Kesemutan, Usia ≥ 45 Tahun	9	0,275
Cepat Lapar, Sering Kencing, Kesemutan, Sering Haus	0	0,012
Cepat Lapar, Sering Kencing, Lemas, Usia ≥ 45 Tahun	7	0,271
Cepat Lapar, Sering Kencing, Lemas, Sering Haus	0	0,008
Cepat Lapar, Kesemutan, Lemas, Usia ≥ 45 Tahun	8	0,252
Cepat Lapar, Kesemutan, Lemas, Sering Haus	0	0,008
Sering Kencing, Kesemutan, Lemas, Usia ≥ 45 Tahun	12	0,240
Sering Kencing, Kesemutan, Lemas, Sering Haus	1	0,008

Hasil kandidat k-5 (5 itemset) dengan nilai minimum support sebesar 0,03 atau 3% dan nilai minimum confidence sebesar 0,8 atau sebesar 80% setelah diuji pada RapidMiner adalah menghasilkan 2 jenis anamnesa. Dapat dilihat pada tabel 7.

TABEL VII
HASIL KANDIDAT K-5 (5 ITEMSET)

Item Set	Jumlah	Support
Cepat Lapar, Sering Kencing, Kesemutan, Lemas, Usia ≥ 45 Tahun	7	0,240
Cepat Lapar, Sering Kencing, Kesemutan, Lemas, Sering Haus	0	0,008

24. Jika pasien merasa kesemutan, lemas dan sering haus, maka dia akan cenderung merasa sering kencing dengan nilai *support* 0,008 dan nilai *confidence* 1.
25. Jika pasien merasa sering kencing, kesemutan, lemas dan berusia ≥ 45 tahun, maka dia akan cenderung merasa cepat lapar dengan nilai *support* 0,240 dan nilai *confidence* 1.
26. Jika pasien merasa lemas dan sering haus, maka dia akan cenderung merasa cepat lapar, sering kencing dan kesemutan dengan nilai *support* 0,008 dan nilai *confidence* 1.
27. Jika pasien merasa cepat lapar, lemas dan sering haus, maka dia akan cenderung merasa sering kencing dan kesemutan dengan nilai *support* 0,008 dan nilai *confidence* 1.
28. Jika pasien merasa sering kencing, lemas dan sering haus, maka dia akan cenderung merasa cepat lapar dan kesemutan dengan nilai *support* 0,008 dan nilai *confidence* 1.
29. Jika pasien merasa cepat lapar, sering kencing, lemas dan sering haus, maka dia akan cenderung merasa kesemutan dengan nilai *support* 0,008 dan nilai *confidence* 1.
30. Jika pasien merasa kesemutan, lemas dan sering haus, maka dia akan cenderung merasa cepat lapar dan sering kencing dengan nilai *support* 0,008 dan nilai *confidence* 1.
31. Jika pasien merasa cepat lapar, kesemutan, lemas dan sering haus, maka dia akan cenderung merasa sering kencing dengan nilai *support* 0,008 dan nilai *confidence* 1.
32. Jika pasien merasa sering kencing, kesemutan, lemas dan sering haus, maka dia akan cenderung merasa cepat lapar dengan nilai *support* 0,008 dan nilai *confidence* 1.

Dari pembentukan aturan asosiasi yang dianalisis, ditemukan bahwa kombinasi gejala seperti sering kencing, merasa cepat lapar, kesemutan, lemas, dan sering haus, terutama pada pasien berusia ≥ 45 tahun, memiliki kecenderungan kuat (dengan *confidence* bernilai 1) untuk muncul bersamaan. Gejala-gejala tersebut saling berkaitan dan cenderung memunculkan gejala tambahan seperti merasa cepat lapar, sering kencing, atau kesemutan. Beberapa kombinasi gejala menunjukkan nilai *support* lebih tinggi, yang mengindikasikan frekuensi kemunculan pola tersebut dalam data, seperti aturan yang melibatkan usia ≥ 45 tahun dengan gejala lemas atau sering kencing yang lebih sering muncul bersama dengan perasaan cepat lapar.

3) Evaluasi Pemodelan Aturan Asosiasi

Evaluasi pemodelan data mining dengan algoritma Apriori dalam menentukan aturan asosiasi menggunakan lift ratio tertinggi muncul pada kombinasi kandidat itemset, setelah diproses dengan RapidMiner menghasilkan hasil yang dapat dilihat pada gambar 10 dan 11 berikut.

Gbr.10 Hasil Lift Ratio 1

Gbr.11 Hasil Lift Ratio 2

Berdasarkan gambar 10 dan 11 diketahui bahwa evaluasi pemodelan menggunakan *lift ratio* untuk beberapa *premises* dan *conclusion* dengan algoritma Apriori adalah sebagai berikut.

1. *Premis* terkait kondisi "Sering Haus" dan "Cepat Lapar": Maka menghasilkan *conclusion*: Kondisi ini menghasilkan *lift ratio* sebesar 1,042. Meskipun nilai *lift ratio* sedikit di atas 1, ini menunjukkan hubungan lemah, mengindikasikan bahwa "Sering Haus" sedikit mempengaruhi kemungkinan "Cepat Lapar".
2. *Premis* "Sering Kencing, Usia ≥ 45 Tahun", "Kesemutan, Sering Haus", "Lemas, Sering Haus" hingga kombinasi lengkap dari beberapa gejala lainnya yang menghasilkan "Cepat Lapar": Maka menghasilkan *conclusion*: Semua *premis* ini memiliki *lift ratio* sebesar 1,042, menunjukkan bahwa pengaruh gejala-gejala ini terhadap "Cepat Lapar" masih minimal dan tidak menunjukkan hubungan signifikan.
3. *Premis* "Kesemutan, Sering Haus" hingga kombinasi lebih kompleks yang menghasilkan "Sering Kencing": Maka menghasilkan *conclusion*: *Lift ratio* sebesar 1,156 menunjukkan adanya pengaruh yang sedikit lebih kuat dari sebelumnya. Ini menandakan adanya hubungan yang lebih nyata, meskipun masih lemah antara gejala "Kesemutan, Sering Haus" dengan "Sering Kencing".
4. *Premis* "Kesemutan, Sering Haus" hingga "Kesemutan, Lemas, Sering Haus" yang menghasilkan kombinasi "Cepat Lapar, Sering Kencing": Maka menghasilkan *conclusion*: *Lift ratio* 1,183

menunjukkan hubungan lebih kuat dari sebelumnya, menyarankan bahwa kombinasi gejala tersebut lebih mungkin berasosiasi dengan "Cepat Lapar, Sering Kencing".

5. *Premis* "Lemas, Sering Haus", hingga kombinasi beberapa kondisi yang menghasilkan "Kesemutan": Maka menghasilkan *conclusion*: Dengan *lift ratio* 1,208, hubungan ini mulai lebih signifikan. Ini mengindikasikan bahwa "Lemas, Sering Haus" lebih mungkin memicu "Kesemutan".
6. *Premis* "Lemas, Sering Haus" hingga "Sering Kencing, Lemas, Sering Haus" yang menghasilkan "Cepat Lapar, Kesemutan": Maka menghasilkan *conclusion*: *Lift ratio* 1,240 menunjukkan pengaruh yang lebih kuat, mengindikasikan bahwa gejala-gejala tersebut memiliki hubungan nyata terhadap kondisi "Cepat Lapar, Kesemutan".
7. *Premis* "Lemas, Sering Haus" hingga kombinasi gejala yang menghasilkan "Sering Kencing, Kesemutan" dan kombinasi "Cepat Lapar, Sering Kencing, Kesemutan": Maka menghasilkan *conclusion*: *Lift ratio* 1,350 hingga 1,383 menunjukkan hubungan yang paling signifikan di antara semua kombinasi. Ini menunjukkan bahwa "Lemas, Sering Haus" merupakan prediktor yang lebih kuat untuk kondisi "Sering Kencing, Kesemutan".

Oleh karena itu dapat disimpulkan bahwa evaluasi pemodelan dengan algoritma Apriori menunjukkan bahwa kombinasi gejala seperti "Lemas" dan "Sering Haus" memiliki *lift ratio* yang lebih tinggi, terutama dalam kaitannya dengan kondisi "Kesemutan," "Sering Kencing," dan "Cepat Lapar." Kombinasi ini mengindikasikan adanya hubungan yang cukup signifikan dibandingkan dengan gejala lain yang memiliki *lift ratio* lebih rendah. Dengan *lift ratio* tertinggi mencapai 1.383, "Lemas" dan "Sering Haus" dapat dianggap sebagai indikator yang lebih kuat untuk memprediksi kondisi kesehatan terkait dibandingkan gejala lainnya.

IV. PENUTUP

A. Kesimpulan

Berdasarkan hasil pengujian serta hasil pembahasan yang di lakukan oleh peneliti tentang analisis pola penyakit diabetes melitus menggunakan algoritma Apriori, maka dapat di simpulkan beberapa hal sebagai berikut:

1. Berdasarkan pengolahan 516 data pasien dengan 10 atribut menggunakan *RapidMiner*, dengan nilai *minimum support* 0,03 (3%), ditemukan 5 kandidat *itemset* anamnesa. Hasil menunjukkan bahwa pola kombinasi gejala "Cepat Lapar," "Sering Kencing," dan "Kesemutan" adalah yang paling dominan (0.723), diikuti kombinasi dengan "Lemas" (0.516). Nilai kemunculan menurun drastis dengan tambahan gejala "Usia \geq 45 Tahun" atau "Sering Haus," yang kurang berkontribusi. Gejala tunggal "Cepat Lapar" (0.959) dan "Sering Kencing" (0.864) memiliki kemunculan tinggi, menjadikannya indikator utama, sementara "Sering Haus" memiliki kemunculan sangat rendah (0.008 hingga

0.016). Secara keseluruhan, kombinasi "Cepat Lapar," "Sering Kencing," dan "Kesemutan" berpotensi digunakan untuk deteksi awal atau diagnosis berdasarkan pola gejala yang sering muncul.

2. Berdasarkan hasil pengujian data yang telah dilakukan dengan *RapidMiner* menggunakan algoritma Apriori, maka diperoleh penentuan pola penyakit diabetes melitus menghasilkan analisis aturan asosiasi yang menunjukkan bahwa kombinasi gejala seperti sering kencing, merasa cepat lapar, kesemutan, lemas, dan sering haus, terutama pada pasien berusia \geq 45 tahun, memiliki kecenderungan yang kuat untuk muncul bersamaan, dengan confidence bernilai 1. Gejala-gejala ini saling berkaitan dan cenderung memunculkan gejala tambahan. Beberapa kombinasi gejala, terutama yang melibatkan usia \geq 45 tahun dengan gejala lemas dan sering kencing, menunjukkan nilai support lebih tinggi, menandakan frekuensi kemunculan pola tersebut. Selain itu, evaluasi pemodelan dengan algoritma Apriori mengungkap bahwa kombinasi "Lemas" dan "Sering Haus" memiliki *lift ratio* tertinggi (1.383), mengindikasikan hubungan yang signifikan dengan gejala "Kesemutan," "Sering Kencing," dan "Cepat Lapar." Hal ini menunjukkan bahwa kombinasi gejala ini dapat menjadi indikator yang lebih kuat untuk memprediksi kondisi kesehatan terkait dibandingkan dengan gejala lainnya.

B. Saran

Pada penelitian kali ini mungkin terdapat banyak kekurangan dalam pelaporan maupun hasil pengujian pada penelitian kali ini. Maka dari itu untuk pengembangan lebih lanjut maka di perlukan beberapa saran sebagai berikut:

1. Pada penelitian selanjutnya diharapkan untuk menggunakan algoritma lainnya seperti algoritma *FP-Growth (Frequent Pattern Growth)* dan *Eclat (Equivalence Class Transformation)* agar memperoleh hasil yang berbeda ataupun lebih baik sehingga dapat dibandingkan manakah algoritma yang cocok untuk menentukan aturan asosiasi.
2. Diharapkan pada penelitiannya berikut, dapat menggunakan nilai *minimum support* dan nilai *minimum confidence* yang lebih baik lagi, agar dapat menghasilkan kandidat *itemset* dan pola aturan asosiasi yang lebih beragam dan lebih baik.

REFERENSI

- [1] Ahmad, N., Irianto, G., & Nurhasanah, F. V. D. (2024). Determinan Risiko yang Berkaitan dengan Diabetes Melitus Di Puskesmas Cigugur Tengah Kota Cimahi. In *Jurnal Kesmas (Kesehatan Masyarakat) Khatulistiwa* (Vol. 11, Issue 1, pp. 1–9).
- [2] Aprillia, D. (2013). Belajar Data Mining dengan RapidMiner. *Innovation and Knowledge Management in Business Globalization: Theory & Practice*, Vols 1 and 2, 5(4), 1–5.
- [3] Buulolo, E. (2013). Implementasi Algoritma Apriori Pada Sistem. 71–83.
- [4] Nurzanah, S., Alam, S., & Iman, T. (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.
- [5] Febrinasari, P., Sholikah, A., & Pakha, D. (2020). *Buku Saku Diabetes Melitus untuk Awam*. Surakarta : UNS Press. Penerbitan Dan Pencetakan UNS (UNS Press), 1, 79.
- [6] Kemenkes RI. (2019). *Buku Pintar Kader Posbindu. Buku Pintar Kader Posbindu*, 1–65.
- [7] Kemkes. (2023). *Survei Kesehatan Indonesia. Kota Kediri Dalam Angka*, 1–68.
- [8] Linawati, N., Hadisaputro, S., & Mardiyono. (2021). Alternatif Layanan Komplementer Pemberian Saponin Terhadap Penurunan Kadar Gula Darah Puasa dan 2 Jam Postprandial Pada Penderita Diabetes Mellitus Tipe II.
- [9] Medina, I. (2021). Apa itu Data Mining dan Bagaimana Metodenya? - Dicoding Blog. <https://www.dicoding.com/blog/apa-itu-data-mining/>
- [10] Ritha, N. Suswaini, E., & Pebriadi, W. (2021). Penerapan Association Rule Menggunakan Algoritma Apriori Pada Poliklinik Penyakit Dalam (Studi Kasus: Rumah Sakit Umum Daerah Bintang). *Jurnal Sains Dan Informatika*, 7(2), 222–230.
- [11] Perkeni. (2021). *Pedoman Pengelolaan dan Pencegahan Diabetes Melitus Tipe 2 Di Indonesia*.
- [12] Aiman, Suryadin, H. Masita, Sari. (2022). Metodologi Penelitian Kuantitatif. In Yayasan Penerbit Muhammad Zaini.
- [13] Pradnyana, Agustini, K. (2022). Konsep Dasar Data Mining. *Konsep Data Mining*, 1, 1–16.
- [14] Qotrun, A. (2021). Pengertian Metode Observasi dan Contohnya - Gramedia Literasi. In Gramedia Blog. <https://www.gramedia.com/literasi/pengertian-metode-observasi-dan-contohnya/>
- [15] Rahayu, P., Sudipa, I., Suryani, Surachman, A., Ridwan, A., Darmawiguna, M., Sutoyo, M., Slamet, I., Harlina, S., & Sanjaya, M. (2024). *Buku Ajar Data Mining (Vol. 1, Issue January 2024)*. Pt. Sonpedia Publishing Indonesia.
- [16] Rahayu, W., Bernadus, N., & Datya, I. (2024). Penerapan Data Mining Dalam Mengetahui Pola Transaksi Pembelian Obat Menggunakan Algoritma Apriori Di Apotek Kharisma Farma Tiga. *J-Icon : Jurnal Informatika Dan Komputer*, 12(1), 44–55.
- [17] Salsabila, S. (2019). Modul Data Mining Association Rule. 9, 1–17.
- [18] Sangadji, W., & Ayu, M. (2018). Epidemiologi Penyakit Diabetes Melitus (DM). *Jurnal Universitas Esa Unggul*, 2(1), 35–42.
- [19] Sugianto, A., & Sukmawati, D. (2023). Penerapan Algoritma FP-Growth Untuk Mengetahui Pola Pada Data Transaksi Percetakan (Studi Kasus Java Printing Batujajar). 05(01), 20–26.
- [20] Tampubolon, K., Saragih, H., Reza, B., (2013). Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Alat-Alat Kesehatan. 93–106.
- [21] Yuliana, A. (2024). Pengantar Metodologi Penelitian Kualitatif. In CV. Gita Lentera (Vol. 5). CV. Gita Lentera. https://books.google.co.id/books?id=nMEXEQAAQBAJ&printsec=fro-ntcov_er&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false