

Ekstraksi Perilaku Pasien Pada Kunjungan Poliklinik Rumah Sakit Menggunakan *FP-Growth*

Aziz Mu'min¹, Ivana Valentina², Vanessa Metayani³, Vardina Nava⁴, Mewati Ayub⁵, Hapnes Toba⁶, Swat Lie Liliawati⁷

^{1,2,3,4,5,6,7}Magister Ilmu Komputer, Universitas Kristen Maranatha

E-mail: ¹2279807@maranatha.ac.id, ²2279006@maranatha.ac.id, ³2279803@maranatha.ac.id, ⁴2279809@maranatha.ac.id, ⁵mewati.ayub@maranatha.ac.id, ⁶hapnestoba@it.maranatha.edu, ⁷swat.ll@maranatha.ac.id,

ABSTRACT

KEYWORDS:

Association Rules; FP-Growth; Machine Learning; Patient Visit Behavior.

The implementation of a hospital management information system (SIMRS) in a hospital can provide new knowledge in managing the hospital and allow hospital management to obtain large amounts of patient data regarding patient visits. One of the challenges in using big data in hospitals is the extraction of patient behavior in visiting polyclinics in hospitals. This patient visit behavior is a very important factor for hospital management to make the right decisions. In this study using the association rules method to extract patient visit data in order to produce good information and understand the behavior of patient visits in the hospital. The results of this study show that the association rules method can extract patient visit data and produce strong association rules on patient visit behavior.

ABSTRAK

KATA KUNCI:

Association Rules; FP-Growth; Machine Learning; Perilaku Kunjungan Pasien.

Penerapan sistem informasi manajemen rumah sakit (SIMRS) pada sebuah rumah sakit dapat memberikan pengetahuan baru dalam melakukan pengelolaan rumah sakit dan memungkinkan manajemen rumah sakit untuk memperoleh data pasien dalam jumlah besar mengenai kunjungan pasien. Salah satu tantangan dalam menggunakan big data di rumah sakit adalah ekstraksi perilaku pasien dalam melakukan kunjungan ke poliklinik di rumah sakit. Perilaku kunjungan pasien ini merupakan faktor yang sangat penting bagi pihak manajemen rumah sakit untuk mengambil keputusan yang tepat. Dalam penelitian ini menggunakan metode association rules untuk mengekstrak data kunjungan pasien agar dapat menghasilkan informasi yang baik dan dapat dipahami perilaku kunjungan pasien di rumah sakit. Hasil penelitian ini menunjukkan bahwa dengan metode association rules dapat mengekstraksi data kunjungan pasien dan menghasilkan aturan asosiasi yang kuat pada perilaku kunjungan pasien.

PENDAHULUAN

Rumah sakit adalah suatu institusi pelayanan kesehatan yang menyelenggarakan pelayanan kesehatan perorangan secara paripurna atau menyeluruh yang menyediakan pelayanan rawat jalan, rawat inap dan gawat darurat [1]. Sistem informasi rumah sakit memiliki peranan penting dalam pelayanan klinis dan administratif. Rumah sakit yang menggunakan sistem berbasis elektronik (SIMRS) itu dirancang untuk mengintegrasikan fungsi utama rumah sakit ke dalam satu sistem terpadu yang disimpan dalam pusat *database*. SIMRS ini memungkinkan mendapatkan data mengenai kunjungan pasien pada setiap harinya secara terperinci. Salah satu upaya dalam pembangunan kesehatan adalah dengan peningkatan mutu pelayanan kesehatan di rumah sakit.

Ekstraksi data adalah proses di mana individu mengambil data dari sumbernya. Perusahaan biasanya mengekstrak data jika mereka berharap untuk memindahkannya ke area lain, menemukan informasi di dalam

data atau memperbarui data dengan informasi atau *metadata* tambahan. Dalam penelitian ini khususnya ekstraksi data dilakukan secara manual dengan cara mengambil data secara manual yang bersumber dari SIMRS kemudian diekspor secara manual melalui format CSV yang mana data yang diambil berisi transaksi kunjungan pasien rumah sakit selama periode tahun 2022.

Algoritma *FP Growth* merupakan algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (*frequent item set*) dalam sebuah data di dalam pencarian *frequent item set* dalam sebuah kumpulan data dengan membangkitkan struktur *prefix-tree* atau sering disebut dengan *FP-Tree*. Dari hasil pengujian dapat disimpulkan bahwa penerapan *data mining* menggunakan Algoritma *FP Growth* bisa digunakan untuk menganalisa perilaku kunjungan pasien[5].

Association rules adalah salah satu *task data mining* deskriptif yang bertujuan untuk menemukan aturan asosiatif antara *item-item* data. Langkah utama yang perlu dalam *association rules* adalah mengetahui seberapa sering kombinasi item muncul dalam *database*, yang disebut sebagai *frequent patterns*. Suatu aturan asosiatif dapat diketahui dengan dua parameter *support* yaitu persentase kombinasi item dalam *database* dan *confidence* yaitu kuatnya hubungan antar item dalam aturan asosiatif[3].

Tujuan dari penelitian ini adalah menganalisis perilaku kunjungan pasien terhadap poliklinik di rumah sakit, sehingga dapat membantu manajemen dalam mengambil keputusan guna meningkatkan kualitas pelayanan di rumah sakit, dengan diketahuinya perilaku kunjungan pasien harapannya manajemen dapat menemukan formula dalam pembuatan paket kunjungan berdasarkan perilaku kunjungan pasien yang paling banyak dipilih atau dijalani oleh pasien.

METODE PENELITIAN

A. Association Rules

Metode *Association Rules* adalah sebuah metode dalam *data mining* yang sangat efektif untuk menemukan hubungan antara variabel dalam semua *database* yang sangat besar. *Association Rules mining* adalah teknik *data mining* untuk menemukan aturan asosiatif antara suatu kombinasi *item*. Analisis asosiasi dikenal juga sebagai salah satu teknik *data mining* yang menjadi dasar dari salah satu teknik *data mining* lainnya. Secara khusus, salah satu tahap analisis asosiasi yang menarik perhatian banyak peneliti untuk menghasilkan algoritma yang efisien yaitu analisis pola frekuensi tinggi (*frequent pattern mining*). *Association Rules* adalah aturan yang harus memenuhi batas *minimum support* dan *minimum confident*. *Support* dari suatu *Association Rules* adalah presentasi kombinasi *item* tersebut dalam data, di mana jika mempunyai *item X* dan *item Y* maka *support* adalah proporsi dari transaksi dalam data yang mengandung *X* dan *Y*.

Confident dari *association rules* adalah ukuran ketepatan suatu *rule*, yaitu presentasi transaksi dalam data yang mengandung *X* dan mengandung *Y*. Dengan adanya *confident* kita dapat mengukur kuatnya hubungan antar *item* dalam *association rules*.

B. Data Collection

Data yang dipergunakan pada penelitian ini berasal dari data laporan transaksi kunjungan pasien yang berasal dari sistem informasi management rumah sakit (SIMRS) pada Rumah Sakit Unggul Karsa Medika dalam kurun waktu Januari sampai dengan Desember Tahun 2022, yang menggambarkan secara lengkap transaksi kunjungan pasien terhadap departemen atau poliklinik yang ada di Rumah Sakit Unggul Karsa Medika. Dalam rentang waktu tersebut terdapat 129.113 kunjungan pasien terhadap departemen atau poliklinik yang terekam melalui SIMRS. SIMRS mencatat *tgl_kunjungan*, *no_reg*, *no_rm*, *jk*, *penjamin*, *usia*, *kec*, *kab*, *dept*, *dokter*, *jml_treat* dan *jml_resep*. Seperti terlihat pada Tabel 1.

Tabel 1. Sampel Data RAW Kunjungan Pasien

	<i>tgl_kunjungan</i>	<i>no_reg</i>	<i>no_rm</i>	<i>jk</i>	<i>penjamin</i>	<i>usia</i>	<i>kec</i>	<i>kab</i>	<i>dept</i>	<i>dokter</i>	<i>jml_treat</i>	<i>jml_resep</i>
1	01/01/2022	2201010001	00-00-00-00	Laki-laki	UMUM	1 year 2 months 27	null	null	INSTALASI FARMASI	null	0	1

days												
2	01/01/2022	2201010002	00-02-80-46	Perempuan	BPJS KESEHATAN	23 years 11 months 27 days	Batam Kota	Batam Kota	RAWAT INAP	dr. Handry Pangestu, Sp. PD.	4	47
3	01/01/2022	2201010003	00-00-00-00	Perempuan	UMUM	50 years 8 months 2 days	Batam Kota	Batam Kota	INSTALASI LABORATORIUM	dr. Hendra Subroto, Sp.PK.	0	0
4	01/01/2022	2201010004	00-00-00-00	Laki-laki	UMUM	31 years 8 months 11 days	Katampang	Kab. Bandung	INSTALASI LABORATORIUM	null.	0	0
5	01/01/2022	2201010005	00-00-00-00	Laki-laki	UMUM	26 years 3 months 7 days	null	null	INSTALASI FARMASI	null	0	0
...
129113	31/12/2022	2212310267	00-04-90-80	Perempuan	BPJS KESEHATAN	39 years 8 months 14 days	CIWIDEY	Kab. Bandung	INSTALASI GAWAT DARURAT	dr. Nikki Sabastian	0	0

C. Data Preparation

Tahap ini dilakukan dengan tujuan untuk verifikasi data kunjungan pasien dari SIMRS agar menghasilkan data kunjungan pasien dengan format yang sesuai dengan standar dan kebutuhan dalam penelitian ini. Tahapan ini perlu dilakukan karena pada proses *data mining* yang menghasilkan *raw data* kunjungan pasien ditemukan data dengan format yang tidak diharapkan, data yang kosong dan berbagai ketidaksesuaian data lainnya.

1. Memuat raw data dan trim data

Pada tahap pertama ini kami mengambil data dari SIMRS yang diperlukan dalam penelitian ini, data yang diambil dari SIMRS tersebut berupa *file excel* merupakan data yang besar dengan ukuran 15 Mb, untuk mendapatkan data yang kami butuhkan pada tahap ini kami merubah data tersebut menjadi *file* dengan format CSV (*comma-separated values*) agar dapat diolah pada *notebook*. Kemudian pada tahap ini kami melakukan proses memuat data mentah sekaligus dilakukan *cleaning* atau pembersihan data sesuai tipe datanya, dibersihkan dari data-data yang tidak mempunyai nomor rekam medis dalam hal ini yang memiliki nomor rekam medis 00-00-00-00 dan juga data-data yang memiliki data *null* pada salah satu kolomnya. Pada tahap ini juga kami melakukan kategorisasi untuk kolom usia, *jml_treat*, *jml_resep* sehingga menghasilkan data sebanyak 97.619 baris, seperti pada Tabel 2 berikut.

Tabel 2. Sample Data RAW Kunjungan Pasien sesuai kategori

	tgl_kunjungan	no_reg	no_rm	jk	penjamin	usia	kec	kab	dept	dokter	jml_treat	jml_resep
1	27/01/2022	2201270011	00-00-00-03	Laki-laki	UMUM	Dewasa	CANGKUAN G	KAB. BANDUN G	INSTALASI GAWAT DARURAT	dr. Yohanes Setyawan	Sedikit	Sedikit
2	18/07/2022	2207180420	00-00-00-04	Perempuan	UMUM	Dewasa	CILEUNYI	KAB. BANDUN G	INSTALASI GAWAT DARURAT	dr. Rudy Nohara	Sedikit	Sedikit
3	30/03/2022	2203300150	00-00-00-08	Laki-laki	UMUM	Dewasa	MARGAASIH	KAB. BANDUN G	KLINIK SYARAF	dr. Adnin Nugroho, Sp.N.	Sedikit	Sedikit
4	31/03/2022	2203310052	00-00-00-08	Laki-laki	UMUM	Dewasa	MARGAASIH	KAB. BANDUN G	INSTALASI LABORATORIUM	dr. Hendra Subroto, Sp.PK.	Tidak Ada Treatment	Tidak Ada Resep

t

5	05/04/2022	2204050050	00-00-00-08	Laki-laki	UMUM	Dewasa	MARGAASIH	KAB. BANDUNG	KLINIK THT KL	dr. Henny Widyastuti, Sp. THT-KL.	Sedikit	Sedikit
...
97619	31/12/2022	2212310267	00-04-90-80	Perempuan	BPJS KESEHATAN	Dewasa	CIWIDEY	Kab. Bandung	INSTALASI GAWAT DARURAT	dr. Nikki Sabastian	Tidak Ada Treatment	Tidak Ada Resep

2. Mengambil data yang diperlukan

Data yang diperlukan dalam penelitian ini adalah data kunjungan pada pasien poliklinik sehingga pada tahap ini kami melakukan pemilahan data dengan membuang data departemen INSTALASI GAWAT DARURAT sehingga data yang tadinya ada 97.619 baris berubah menjadi 85.501 baris setelah menghilangkan data departemen yang tidak diperlukan.

3. Mencari sebaran data

Dari jumlah data kunjungan dalam setahun kami melakukan pemilahan kembali data yang akan diproses dalam penelitian ini. Untuk mendapatkan data yang akurat, kami mencari sebaran data yang paling banyak pada setiap bulannya. Dari tahap ini, ditemukan data terbanyak selama tahun 2022, yaitu pada bulan Desember sebanyak 16.905 baris data.

4. Pemilihan kolom data yang diperlukan

Setelah mencari sebaran data terbanyak, selanjutnya kami memilih kolom yang diperlukan. Tujuan dari penelitian ini adalah untuk mencari pola kunjungan pasien yang digambarkan dengan data kunjungan antar departemen. Untuk lebih fokus pada tujuan tersebut, kami hanya menggunakan dan memilih kolom-kolom yang diperlukan untuk proses selanjutnya. Kolom-kolom data tersebut adalah no_rm, tgl_kunjungan, dept.

5. Menentukan *Grouping* data berdasarkan no_rm

Untuk mempermudah pemrosesan data dalam penelitian ini, kami mengurutkan data berdasarkan nomor rekam medis sebagai penanda pasien yang ada dalam data ini. Selain itu, data pasien yang dipilih adalah yang memiliki data kunjungan lebih dari 2, agar data departemen terpilih dan terurut untuk setiap pasiennya.

D. *Pemilihan data kunjungan dari data kunjungan pasien*

Filter data kunjungan pasien dalam penelitian ini kami lakukan dengan cara mengambil data kunjungan pasien pada bulan tertentu, yang kami ambil dalam penelitian ini adalah bulan di mana terdapat paling banyak data kunjungan dibanding bulan-bulan lainnya. Pada penelitian ini kami menggunakan data dengan sebaran terbanyak yaitu pada bulan Desember 2022 dengan 16.905 baris data, kemudian kami menentukan hanya ada tiga kolom yang kami gunakan agar lebih fokus pada tujuan dalam penelitian ini, yaitu untuk mengamati pola kunjungan pasien ke setiap departemen yang digambarkan melalui kolom data no_rm, tgl_kunjungan, dept.

E. *Analisis data menggunakan association rules*

1. Membuat *item set*

Data kunjungan pasien ini diproses menggunakan algoritma *FP Growth* untuk mendapatkan hasil *item set*. Dalam penelitian ini minimum *support* yang digunakan adalah sebesar 0.4 dan nilai *confident* sebesar 0.5 data *item set* kemudian diproses dengan *association rules* untuk mendapatkan *rules* pada data ini.

2. Analisis frekuensi *item set*

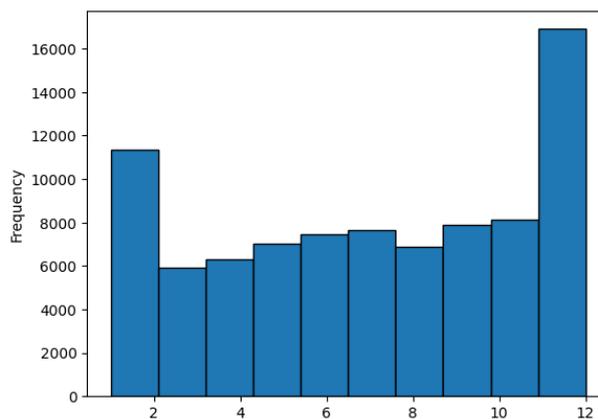
Pada tahap ini data yang diambil sebagai *item set* adalah data pasien dengan kunjungan lebih dari satu kali dalam satu bulan yang berarti *item set* yang diambil adalah pasien yang memiliki dua kunjungan atau lebih. Analisis ini digunakan untuk melihat pola dari kunjungan pasien tersebut.

HASIL DAN PEMBAHASAN

Tahapan selanjutnya merupakan tahapan data *mining* dengan menggunakan *association rules*. Data kunjungan pasien digunakan untuk melakukan analisis aturan asosiasi dan dipergunakan untuk memverifikasi kelayakan metode yang diusulkan dalam penelitian ini. Selain itu, penelitian ini menggunakan *item set* dari data kunjungan pasien yang paling sering muncul dalam rentang waktu satu bulan.

A. Pemilihan sebaran data kunjungan pasien

Oleh karena pola dari data kunjungan pasien dengan diagnosis yang sama diidentifikasi dengan asumsi gambaran dalam rentang satu bulan, maka dalam penelitian ini kami memilih sebaran data perbulan yang memiliki data terbanyak dalam tahun 2022. Kami melakukan pemrosesan data dan menemukan bulan yang memiliki data kunjungan sebanyak 16.905 baris data.



Gambar 1. Sebaran data kunjungan pasien dalam satu tahun (Tahun 2022)

Berdasarkan data kunjungan pasien yang digunakan terdapat nomor rekam medis pasien yang berkunjung ke beberapa departemen sehingga dapat diidentifikasi sebagai *item set* kunjungan pasien, pada data yang digunakan tercatat ada sebanyak 16.905 baris data pada bulan Desember 2022, data ini yang digunakan untuk menemukan pola kunjungan pasien dengan ketentuan pasien memiliki kunjungan lebih dari 2 kunjungan dalam satu bulan.

B. FP Growth

```
[23] !pip install fpgrowth_py

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Collecting fpgrowth_py
  Downloading fpgrowth_py-1.0.0-py3-none-any.whl (5.6 kB)
Installing collected packages: fpgrowth_py
Successfully installed fpgrowth_py-1.0.0

[24] from fpgrowth_py import fpgrowth
freqItemSet, rules = fpgrowth(result_array, minSupRatio=0.4, minConf=0.5)
#0.4. 0.5
print(len(freqItemSet))
print(len(rules))

101
100
```

Gambar 2. Penggunaan FP Growth dengan library python (fpgrowth_py)

Selanjutnya setelah ditemukan sebaran data, kami mencoba mencari *item set* dan *rules* dengan menggunakan salah satu *library* pada *python* yang mana pada tahap ini kami menentukan *minimum support* dengan nilai 0.4 dan *minimum confident* dengan nilai 0.5, sehingga dapat ditemukan *item set* sebanyak 101 dan *rules* sebanyak 100.

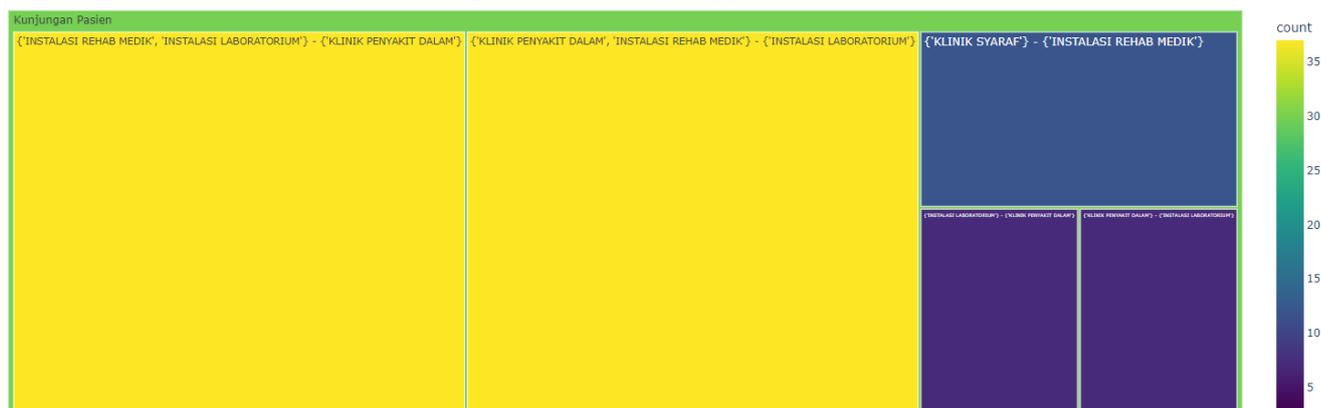
C. Association rules

Association rules adalah metode yang efektif untuk dapat melihat hubungan antara variabel dengan jumlah data yang besar. Data kunjungan pasien yang didapatkan dari proses sebelumnya dimana *item set* dan *rules* telah ditemukan datanya kemudian diproses kembali untuk menemukan gambaran *item set* dan *rules* yang diproses sebelumnya. Pola perilaku kunjungan pasien yang dihasilkan dan diamati pada penelitian ini, Hasil *association rules* yaitu jika A maka B, yaitu perilaku kunjungan pasien terbanyak ke beberapa departemen maka diperoleh pola kunjungan pasien. Dari hasil proses ini ditemukan *rules* yang paling banyak adalah {'INSTALASI REHAB MEDIK', 'KLINIK PENYAKIT DALAM'} - {'INSTALASI LABORATORIUM'} dengan jumlah 37, kemudian {'INSTALASI REHAB MEDIK', 'INSTALASI LABORATORIUM'} - {'KLINIK PENYAKIT DALAM'} dengan jumlah 37 dan {'KLINIK SYARAF'} - {'INSTALASI REHAB MEDIK'} dengan jumlah 12. Ini merupakan 3 *rules* yang paling banyak pada penelitian ini seperti yang digambarkan pada gambar 3.

index	rule	count
0	{'KLINIK SYARAF'} - {'INSTALASI REHAB MEDIK'}	12
1	{'INSTALASI REHAB MEDIK', 'KLINIK PENYAKIT DALAM'} - {'INSTALASI LABORATORIUM'}	37
2	{'INSTALASI REHAB MEDIK', 'INSTALASI LABORATORIUM'} - {'KLINIK PENYAKIT DALAM'}	37
3	{'KLINIK PENYAKIT DALAM'} - {'INSTALASI LABORATORIUM'}	7
4	{'INSTALASI LABORATORIUM'} - {'KLINIK PENYAKIT DALAM'}	7

Gambar 3. Daftar rules terbanyak pada kunjungan pasien

Terdapat 5 pola perilaku kunjungan pasien yang dihasilkan dan diamati pada penelitian ini, hasil dari *association rules* kemudian selanjutnya diurutkan berdasarkan jumlah *rules* atau perilaku kunjungan pasien terbanyak agar data dapat dipahami dengan mudah oleh pengguna.



Gambar 4. Daftar rules terbanyak pada kunjungan pasien dalam bentuk graph

Gambar 4 menggambarkan hasil analisis perilaku kunjungan pasien yang diurutkan mulai dari perilaku terbanyak sampai yang sedikit, data diatas telah menggambarkan 5 data terbanyak yang dapat digunakan oleh pihak manajemen rumah sakit untuk dapat meningkatkan pelayanan rumah sakit kepada pasien yang berkunjung.

KESIMPULAN

Berdasarkan hasil analisis, dapat ditarik kesimpulan bahwa metode *FP Growth* dan *association rules* yang digunakan dalam penelitian ini dapat menghasilkan informasi dan pemahaman yang baik dalam

mengetahui perilaku kunjungan pasien di rumah sakit. Data kunjungan pasien di rumah sakit banyak dihasilkan setiap harinya dan disertai karakteristik yang berbeda-beda pada setiap bulannya.

Dengan melakukan pemrosesan data menggunakan *FP Growth* untuk dapat menemukan *item set* dan *rules* yang terdapat pada data kunjungan pasien ini sangat membantu dalam menemukan *rules* yang dicari, cara yang dilakukan sangat sederhana dengan menggunakan salah satu *library* dalam *python* yaitu *fpgrowth_py*.

Metode *association rules* membantu dalam menemukan atau menerjemahkan *rules* yang ditemukan pada proses sebelumnya sehingga perilaku kunjungan pasien dapat terlihat dan dapat diketahui jumlah terbanyaknya, sehingga data hasil analisis tersebut dapat dipahami oleh manajemen rumah sakit untuk mengambil keputusan dalam upaya meningkatkan kualitas layanan kepada pasien di rumah sakit.

UCAPAN TERIMA KASIH

Terima kasih saya sampaikan kepada seluruh *civitas academica* Universitas Kristen Maranatha Bandung, khususnya para dosen pembimbing dan teman-teman perkuliahan khususnya teman satu kelompok tugas holistik perkuliahan yang telah mencurahkan ilmu dan pengalamannya hingga Jurnal ini dapat diselesaikan dengan baik.

DAFTAR PUSTAKA

1. Albertus Indarko Wiyogo, et al. "Ekstraksi Perilaku Komuter Pada Commuter Line Menggunakan Rule-Based Machine Learning." *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 9, no. 1, 2023, pp. 154-166.
2. "Ekstraksi Data: Pengertian, Jenis, dan Tips Melakukannya." *Aksaragama.com*, 18 April 2022, <https://aksaragama.com/ekstraksi-data/>. Accessed 18 June 2023.
3. "Implementasi Association Rules dengan Algoritma Apriori pada Dataset Kemiskinan." *Limits*, vol. 14, no. 2, 2017, p. 145.
4. Puteri Fannya, et al. "Quality of polyclinic services at public health centers in Padang city, Indonesia." *International Journal Of Community Medicine And Public Health*, vol. 5, no. 10, 2018, p. 4207. *International Journal Of Community Medicine And Public Health*, <https://www.ijcmph.com/index.php/ijcmph/article/view/3584>. Accessed 18 June 2023.
5. Ranga Yogasuwara. "Implementasi Algoritma Frequent Growth (FP-Growth) Menentukan Asosiasi Antar Produk." *Jurnal Sistem Komputer dan Informatika*, vol. 4, no. 1, 2022, p. 165.
6. Ruth Molly I, and Meyrolen Itaar. "Analisis Pemanfaatan Sistem Informasi Manajemen Rumah Sakit (SIMRS) Pada RRSUD DOK II Jayapura." *Journal of Software Engineering Ampera*, vol. 2, no. 2, 2021, pp. 95-101.
7. Undang-undang Republik Indonesia. *Undang-undang Nomor 44 Tahun 2009 tentang Rumah Sakit*. Republik Indonesia, 2009.