

IMPLEMENTASI TEKNIK ASSOCIATION RULE MINING UNTUK MENGETAHUI POLA MINAT KERJA MAHASISWA SESUAI BIDANG KEAHLIAN

Dwi Winarti¹, Wulan Andang²

^{1,2} Fakultas Ilmu Komputer, Universitas Dharma Indonesia Jl. Lintas Sumatera Km 18 Koto Baru
email: dwiwinarti@undhari.ac.id, wulanap2@gmail.com

Abstract

The Association Rule Mining technique will brings out the students' interest and talent patterns in the subjects that they get in college. From the academic scores that have been classified into 3 major subjects, namely those which related to programming, design, and network courses, this is adjusted to the vision and mission of the Computer Science Faculty of Informatics Engineering Study Program. The association technique is in the form of most emerging rules or strong rules. Combined with the Fp-Growth algorithm to make it easier to find patterns from a set of classified data. Fp- Growth will create a Tree structure to strengthen generating large data sets. The samples are taken from several subjects in the S1 Informatics Engineering study program. They are grouped into programming, graphic design and networking subjects. The Fp-Growth algorithm can show that the association rule can form a pattern of students' interest and talent towards courses and their relationship to work according to their field of expertise. It is expected to be able to predict the job that will be obtained after graduating from the college.

Keywords: association rule mining, interest in work, prediction

Abstrak

Alumni Teknik Informatika 45 % bekerja tidak sesuai dengan bidang keahlian yang diperoleh selama menempuh pendidikan diperguruan tinggi. Ada beberapa faktor diantaranya: pertama, alumni masih belum percaya kemampuannya sendiri atau belum menguasai progam studi yang mereka pilih, kedua ketersediaan lapangan pekerjaan. Teknik *Association Rule Mining* akan membantu memunculkan pola minat dan bakat mahasiswa terhadap mata kuliah yang mereka dapat ketika kuliah. Nilai akademik diklasifikasikan kedalam 3 mata kuliah jurusan yaitu yang berkaitan dengan matakuliah *progammer*, desain, dan jaringan hal ini di sesuaikan dengan visi dan misi progam studi Teknik Informatika Fakultas Ilmu Komputer. Teknik asosiasi berupa *rule-rule* yang paling banyak muncul atau *strong rule*, dikombinasikan dengan algoritma *Fp-Growth* sehingga mempermudah menemukan pola dari sekumpulan data yang sudah diklasifikasikan. *Fp- Growth* akan membentuk struktur *Tree* untuk memperkuat dalam memunculkan himpunan data yang besar. Pengambilan sampel dari matakuliah pada progam studi S1 Teknik Informatika. Data sampel dikelompokkan dalam matakuliah *programming*, desain grafis dan jaringan. Algoritma *Fp-Growth* dapat menunjukkan bahwa *rule* asosiasi dapat membentuk sebuah pola minat dan bakat mahasiswa terhadap mata kuliah dan keterkaitannya dengan pekerjaan sesuai bidang keahlian. Hasil penelitian dapat memprediksi pekerjaan yang sesuai setelah lulus dari perguruan tinggi.

Kata Kunci : Association, rule mining, minat kerja, prediksi

1. PENDAHULUAN

Banyak permasalahan yang terjadi di bidang pendidikan, sebagai contoh mahasiswa Teknik Informatika Fakultas Ilmu Komputer Universitas Dharma Indonesia. Berdasarkan data alumni 45% dibuktikan bahwa Alumni Teknik Informatika bekerja tidak sesuai

dengan bidang keahlian yang diperoleh selama menempuh pendidikan diperguruan tinggi. Ada beberapa faktor diantaranya: pertama, alumni masih belum percaya kemampuannya sendiri atau belum menguasai progam studi yang mereka pilih, kedua ketersediaan lapangan pekerjaan. Bahkan, mahasiswa yang mendapatkan IPK diatas 3,5

masih ragu dengan kemampuannya dalam memilih bidang pekerjaan. Untuk membuktikan data tersebut diambil sebagian sampel dari data mahasiswa Teknik Informatika. Teknik *rule asosiasi* ini dapat membentuk pola sehingga dapat terlihat seberapa besar keterkaitan satu item dengan item lainnya, yaitu keterkaitan matakuliah yang paling diminati dengan bidang pekerjaan yang akan dipilih nantinya didukung dengan nilai *confidence* dan *support* akan dihasilkan *Strong rule* yang menghasilkan kelinieritasnya suatu program studi terhadap pekerjaan. Hasil penelitian ini dapat menjadi referensi mahasiswa dalam memilih pekerjaan yang tepat sesuai dengan skill yang dimiliki.

Teknik yang dibuat dalam data mining adalah bagaimana menelusuri data yang ada untuk membangun sebuah model, kemudian menggunakan model tersebut agar dapat mengenali pola data yang lain yang tidak berada dalam basis data yang tersimpan. (Siska Haryanti, 2015). Proses mining program studi, merupakan proses mining untuk mengetahui hubungan tingkat kelulusan dengan program studi. Program studi digunakan untuk proses mining guna mengetahui hubungan tingkat kelulusan dengan program studi. (Kurniawani, 2016). Data Mining untuk menentukan korelasi *support* dan *confidence* jurusan siswa pada tingkat sekolah menengah terhadap indeks prestasi kumulatif (IPK) di perguruan tinggi sebagai solusi tepat pemilihan program studi pada perguruan tinggi (Jollyta et al., 2019). Analisis Data mining dengan algoritma *Fp-Growth* dalam mendukung strategi perguruan tinggi. (Winarti et al., 2018).

FP-Growth Algoritma *FP-Growth* adalah salah satu alternatif algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul. Karakteristik algoritma *FP-Growth* adalah struktur data yang digunakan adalah tree yang disebut *FP-Tree*. (Afuan et al., 2019). Keberhasilan prestasi akademik mahasiswa selama mengikuti pendidikan di perguruan tinggi dinilai dari penilaian mata kuliah, penilaian semester, penilaian akhir tahun akademik dan penilaian akhir program studi. (Ikhwan & Nofriansyah, 2015).

Metode *FP-Growth* dapat dibagi menjadi 3 tahapan utama yaitu sebagai berikut:

1. Tahap pembangkitan *conditional pattern base*.
2. Tahap pembangkitan *conditional FP-Tree*.
3. Tahap pencarian *frequent itemset*. Ketiga tahap tersebut merupakan langkah yang akan dilakukan untuk mendapat *frequent itemset*. Dan Input *FP-Tree* Tree Output (Pelawi, 2019).

Hasil *support* dan *confidence* pada data mining untuk market basket dapat dibatasi, diterapkan association rule, dimana minsup dan minconf dapat diatur sesuai dengan kebutuhan. Minsup adalah ambang batas *support*, sedangkan *minconf* adalah ambang batas *confidence* (Fitriyani, 2016). untuk memperkuat proses *Growth* maka akan dilakukan proses rule mining. *Association Rule Mining* menemukan *frequent itemsets*, yaitu *itemset* yang memiliki *support* yang lebih besar dari minimum *support*. (ELWANI, 2017). Untuk menentukan aturan asosiasi yang akan dipilih maka harus diurutkan berdasarkan perkalian *support* dengan *confidence*. (Irvan Firnando, 2015). *Support* dari suatu *association rule* adalah presentasi kombinasi item tersebut dalam *database*, dimana jika mempunyai item A dan item B maka *support* adalah proporsi dari transaksi dalam database yang mengandung A (Fauzy et al., 2016).

2. METODE PENELITIAN

Kerangka kerja penelitian, kerangka kerja ini merupakan langkah-langkah yang akan dilakukan dalam penyelesaian masalah yang akan dibahas. Adapun kerangka kerja penelitian ini dapat dilihat pada gambar 2.1:



Gambar 1. Kerangka Kerja Peneliti

1. Tahap *retrieve* data merupakan tahap pengambilan data atau tahap penempatan data. Pada tahap pemutakhiran, data yang tersimpan diperbaharui dan disesuaikan dengan peristiwa terbaru. Kemudian pada tahap *retrieving*, data yang tersimpan diakses dan diringkas kembali untuk diproses lebih lanjut atau untuk keperluan pembuatan laporan.
2. Analisis Korelasi
Pada Tahap ini dilakukan korelasi data menggunakan Progam *Microsoft Excell*.
3. Menganalisa dengan *Fp-Growth* dalam pencarian *frequent itemsets* untuk menghasilkan suatu *output*. *Fp-Growth* akan menghasilkan *frequent Itemset* tanpa harus melakukan *candidates generation*
4. Pembentukan *Fp-Tree* Mengolah data yang akan diproses dengan pembangunan *FP-Tree*. Data yang terpilih tersebut mempunyai beberapa kriteria yang merupakan syarat dalam pembangunan *FP-Tree*. Kriteria didapat dari *itemset*
5. Tahap *association rule* mining merupakan teknik data mining untuk menemukan aturan assosiatif antara suatu kombinasi item. Aturan assosiatif antara suatu kombinasi item akan diperoleh dari atribut yang diperoleh dari data alumni diantara nya nama, asal jurusan, dan pekerjaan yang dipilih Fungsi *association rule* mining digunakan untuk menemukan relasi atau korelasi diantara himpunan item. Nilai support sebuah item diperoleh dengan rumus berikut:

$$Support(A) = \frac{\text{jumlah transaksi mengandung } A}{\text{Total transaksi}} \quad (1)$$

Tahapan Asosiasi Analisis asosiasi dikenal juga sebagai salah satu teknik Data Mining yang menjadi dasar dari berbagai teknik Data Mining lainnya. (Arisandi, 2016). Salah satu tahap dari analisis asosiasi yang disebut analisis pola frekuensi tinggi (*frequent pattern mining*) menarik algoritma yang efisien.. (Hilmy & Andoko, 2016). Penting tidaknya suatu aturan assosiatif dapat diketahui dengan dua parameter, *support* (nilai penunjang) yaitu persentase kombinasi item tersebut dalam database dan *confidence*. (Kusumo et al., 2019). Berikut merupakan proses data mining.

1. Data *real world real world* bisa mengandung data yang tidak, terdapat *noise*, tidak konsisten. (Mulya, 2019)
2. Sistem mining dengan performa tinggi membutuhkan data yang berkualitas. Data preparation atau preparation menghasilkan dataset yang lebih sedikit daripada dataset yang asli, ini bias meningkatkan efisiensi dari Data Mining. Langkah ini mengandung:
 - a. Memilih data yang Relevan
Data yang ada dalam database sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk analisis yang akan diambil dari database. (Syahputra et al., 2018). Data yang relevan dilakukan oleh peneliti dengan mengambil sampel data mahasiswa yang aktif dan matakuliah yang diambil pada progam studi Teknik Informatika Fakultas Ilmu Komputer Universitas Dharmas Indonesia.
 - b. Mengurangi Data.
Proses pemangkasan yaitu hasil dari item yang telah dikombinasikan kemudian dipangkas dengan menggunakan minimum support (Domi Sepri, 2017). Dalam kasus ini peneliti mengelompokkan beberapa matakuliah diantaranya matakuliah umum, matakuliah berkarya, matakuliah jurusan(wajib), matakuliah pilihan. dari hasil klasifikasi matakuliah yang diambil adalah matakuliah jurusan (wajib). Hal ini dilakukan untuk

- mempermudah dalam pengolahan data.
3. Tahap selanjutnya adalah integrasi data. Integrasi data merupakan proses pengubahan atau penggabungan data ke dalam format yang sesuai untuk diproses dalam data mining. Dalam penelitian ini akan dicari nilai support dan confidence dari kombinasi atribut jenis kemampuan dengan atribut parameter jenis kemampuan. (Fana Wiza & Kuning, 2016)
 4. Data Integration and Transformation, untuk meningkatkan akurasi dan efisensi algoritma. Data yang digunakan dalam penelitian ini bernilai kategorikal untuk model klasifikasi, data ditransformasi ke dalam angka menggunakan software *Rapidminer*. (eka sabna, 2016) .
 6. Menguji hasil dengan Rapid Miner menggunakan sistem data dengan pengujian sistem *Rapid Miner* . Algoritma *Fp-Growth* akan melengkapi proses dalam menentukan *minimum support* dan syarat *minimum* dalam menentukan *confidence* dan akan terbentuk hasil yang berupa *Fp-Tree* dan *rule-rule* akan terlihat jelas. Hasil dari *rule-rule* tersebut sebagai *ekstrak* dari data yang diolah berupa pengetahuan baru (*Knowledge*).

3. HASIL DAN PEMBAHASAN

Proses asosiasi akan diterapkan dalam menganalisa Minat mahasiswa terhadap mata kuliah . Proses ini disebut juga Retrieve data karena Pada tahap ini data diambil dari nilai pada beberapa matakuliah. Data yang diambil adalah data real dari mahasiswa Teknik Informatika yang Aktif. Dari retrieve data akan diambil mata kuliah yang ada pada kurikulum Teknik Informatika. Untuk mempermudah mengolah data data tersebut dikelompokkan menjadi 3 kelompok. Untuk bidang pekerjaan juga dikelompokkan menjadi 3 kelompok diantaranya adalah:

1. Kelompok Matakuliah Progam
2. Kelompok Matakuliah Jaringan
3. Kelompok Matakuliah Desain
4. Bidang Pekerjaan Programmer

5. Bidang Pekerjaan Administrator Jaringan
6. Bidang Pekerjaan Desain Grafis

dari data akan ditransformasikan kedalam data mentah dengan menggunakan kode digit yaitu 0 dan 1. Tahap ini merupakan tahap korelasi data untuk memudahkan transformasi maka atribut yang ada akan diberi pengkodean. Proses ini lebih mudah dalam pengimputan data ke *Tools Rapid Miner* korelasi data seperti pada gambar 2:

a1	a2	b1	b2	b3	b4	c1	c2	c3	c4	c5	d1	d2	d3	d4	d5
0	1	0	1	0	0	1	0	0	0	0	1	0	0	0	0
1	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0
1	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0
0	1	0	1	0	0	1	0	0	0	0	0	1	0	0	0
1	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0
0	1	0	0	0	1	1	0	0	0	0	0	0	1	0	0
1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1
0	1	0	1	0	0	1	0	0	0	0	0	1	0	0	0
1	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0
0	1	0	0	1	0	1	0	0	0	0	1	0	0	0	0
0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0
0	1	0	1	0	0	1	0	0	0	0	0	0	1	0	0
0	1	1	0	0	0	1	0	0	0	0	0	0	0	1	1
1	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0
1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1
0	1	1	0	0	0	1	0	0	0	0	1	0	0	0	0

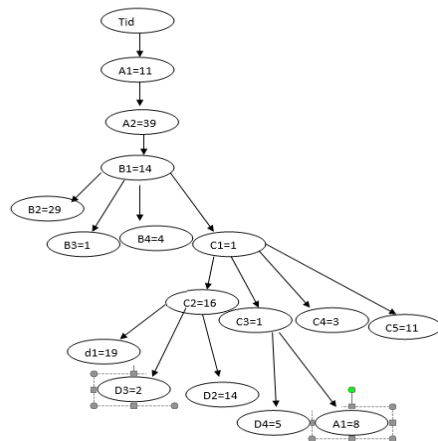
Gambar 2. Transformasi Data pada Aplikasi *Rapid Miner*

Selanjutnya, Metode *Fp-Growth* dilakukan berdasarkan dari hasil transformasi Data . frekuensi itemset yang paling banyak akan dijadikan acuan dalam menentukan nilai *support* dan *confidence*. Seperti pada tabel 1:

Tabel 1. Tabel Kemunculan

Item	Frekuensi
A1	20
C3	16
D2	15
B2	12
B3	10

Berikut merupakan tahap pembentukan *Fp-Tree* perancangan menggunakan *fp-tree* dimulai dari tid 0 hingga ke tid selanjutnya. Seperti pada gambar 3:



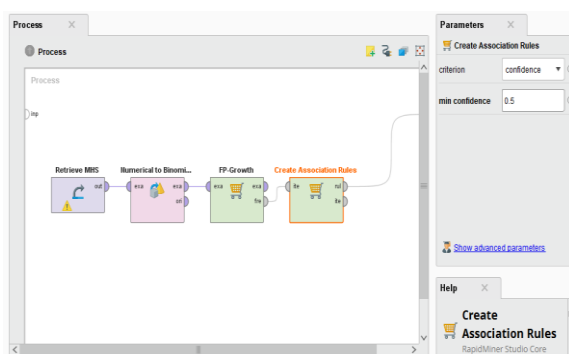
Gambar 3. pembentukan Fp-Tree

Tahap selanjutnya adalah mengevaluasi hasil dengan teknik *Association Mining*. Untuk menguji data, nilai *Support* dan *confidence* harus ditentukan terlebih dahulu. Hasil dari *support* dan *confidence* seperti pada tabel 2

Tabel 2. Hasil *Support* dan *Confidence* terpilih.

No	Jika (A)	Maka (B)	<i>Support</i> %	<i>Confidence</i> %
1	A1	B2	10/25= 0.4	10/20= 0.5
2	D2	B2	9/25= 0.36	9/15= 0.6
3	A3	C3	13/25= 0.52	13/20= 0.65

Tahap terakhir merupakan tahap pengujian menggunakan aplikasi *Rapid Miner*. Hasil nilai *support* dan *confidence* akan diuji kembali pada *tools* ini untuk menghasilkan *Lift Ratio* 1. seperti gambar 4:



Gambar 4. Input Nilai *Support* dan *Confidence*

pemilihan strong rule yang mempunyai nilai *Lift Ratio* 1 dan dibagi dengan jumlah sampel data . berikut merupakan tampilan *lift ratio* dari aplikasi Rapid Miner seperti pada gambar 5:

-1.200	0	1
-0.686	0.034	1.111
-0.686	0.020	1.061
-0.829	0.051	1.136
-0.629	0.038	1.123
-0.771	0.051	1.136
-0.514	-0.024	0.933
-0.743	-0.024	0.955
-0.686	0	1
-0.457	0	1

Gambar 5. *lift ratio*

Nilai *Lift Ratio* merupakan hasil dari beberapa proses Rule asosiasi. *strong rule* yang diambil adalah *Lift Ratio* yang bernilai 1. Presentase *Rule* yang sering muncul akan dijadikan sebagai hasil. Didukung dengan nilai *Support* dan *confidence*.

Berikut *strong rule* yang terbentuk seperti pada tabel 3.

Tabel 3. Pembentukan Rule Asosiasi

No	Rule Asosiasi	Kesimpulan
1	A1,B2	Jika yang diminati kelompok matakul jaringan maka bidang pekerjaan yang sesuai adalah Administrator Jaringan didukung dengan nilai <i>support</i> 40% dan nilai <i>confidence</i> 50%
2	D2, B2	Jika yang diminati kelompok matakuliah Desain/ Multimedia maka pekerjaan yang sesuai adalah Desain Grafis didukung dengan nilai <i>support</i> 36% dan nilai <i>Confidence</i> 60%
3	A3, C3	Jika yang diminati

kelompok mata kuliah pemograman maka bidang pekerjaan yang sesuai adalah pemograman didukung dengan nilai Support 52% dan confidence 65%

4. SIMPULAN

Dari penelitian dapat disimpulkan bahwa mahasiswa teknik informatika mempunyai minat yang tinggi dan nilai yang baik pada beberapa matakuliah yang sudah diklasifikasikan 3 kelompok yaitu: mata kuliah dibidang desain, progamming, dan jaringan. dari hasil yang didapat menunjukan bahwa mahasiswa teknik informatika dapat bekerja sesuai bidang keahlian yang dimiliki terkait dengan skill yang didapatkan ketika menempuh pendidikan diperguruan tinggi.

5. UCAPAN TERIMA KASIH

Terimakasih disampaikan kepada Kementerian Riset Tekonologi dan Pendidikan Tinggi atas dana Penelitian Dosen Pemula Tahun 2020

DAFTAR PUSTAKA

- AFUAN, L., ASHARI, A., & SUYANTO, Y. (2019). *Query Exxpasion in Information Retrieval using Frequent Pattern (FP) Growth Algorithm for Frequent Itemset Search and Association Rules Mining*. 10(2), 263–267.
- ARISANDI, D. (2016). Analisa Pola Pemilihan Program Studi Bagi Calon Mahasiswa Di Universitas Abdurrah Menggunakan Association Rule. *Rabit : Jurnal Teknologi Dan Sistem Informasi Univrab*, 1(1), 1–5. <https://doi.org/10.36341/rabit.v1i1.8>
- DOMI SEPRI, M. A. (2017). Analisa dan Perbandingan Metode Algoritma Apriori dan Fp-Growth untuk Mencari Pola Daerah Strategis pengenalan Kampus Studi Kasus di STKIP ADZKIA Padang. *JURNAL SISTEM INFORMASI KAPUTAMA*, 1(1), 1–9.
- EKA SABNA, M. (2016). Penerapan Data Mining Untuk Memprediksi Prestasi Akademik Mahasiswa Berdasarkan Dosen , Motivasi , Kedisiplinan , Ekonomi ,. *CorelIT*, 2(2), 1–4.
- ELWANI. (2017). PENEELwani. (2017). PENENTUAN ATURAN ASOSIASI PADA TRANSAKSI PEMINJAMAN BUKU MENGGUNAKAN ALGORITMA FP-GROWTH. *Informatika Manajemen Dan Komputer*, 9(1), 1–11.
- NTUAN ATURAN ASOSIASI PADA TRANSAKSI PEMINJAMAN BUKU MENGGUNAKAN ALGORITMA FP-GROWTH. *Informatika Manajemen Dan Komputer*, 9(1), 1–11.
- FANA WIZA, & KUNING, L. (2016). Pemodelan Pola Hubungan Kemampuan Lulusan. *Teknologi Informasi Dan Komunikasi Digital Zone*, 7, 2–7.
- FAUZY, M., SALEH W, K. R., & ASROR, I. (2016). Penerapan Metode Association Rule Menggunakan. *Jurnal Ilmiah Teknologi Informasi Terapan*, II(2), 1–7.
- FITRIYANI, F. (2016). Implementasi Algoritma Fp-Growth Menggunakan Association Rule Pada Market Basket Analysis. *Jurnal Informatika*, 2(1), 1–10. <https://doi.org/10.31311/ji.v2i1.85>
- HILMY, N. F., & ANDOKO, B. S. (2016). Rancang Bangun Aplikasi Data Mining Analisis Tingkat Kelulusan Menggunakan Algoritma Fp-Growth (Studi Kasus Di Politeknik Negeri Malang). *Informatika Polinema*, 1–5.
- IKHWAN, A., & NOFRIANSYAH, D. (2015). Penerapan Data Mining dengan Algoritma Fp-Growth untuk Mendukung Strategi Promosi Pendidikan (Studi Kasus Kampus STMIK Triguna Dharma). *Ilmiah Saintikom*, 1–16.
- IRVAN FIRNANDO, D. (2015).

- Implementasi Algoritma Apriori Dan Forecasting. *Prosiding SENTIA*, 7(3), 2085–2347.
- JOLLYTA, D., EFENDI, S., ZARLIS, M., & MAWENGKANG, H. (2019). Optimasi Cluster Pada Data Stunting: Teknik Evaluasi Cluster Sum of Square Error dan Davies Bouldin Index. *Prosiding Seminar Nasional Riset Information Science (SENARIS)*, 1(September), 918. <https://doi.org/10.30645/senaris.v1i0.100>
- KURNIAWAN, H., & INFORMASI, J. S. (2016). Aplikasi datamining untuk memprediksi tingkat kelulusan mahasiswa menggunakan algoritma apriori di ibi darmajaya bandar lampung. *02(01)*, 79–93.
- KUSUMO, H., SEDIYONO, E., & MARWATA, M. (2019). Analisis Algoritma Apriori Untuk Mendukung Strategi Promosi Perguruan Tinggi. *Walisongo Journal of Information Teknologi*, 1(1), 1–12.
- MULYA, D. P. (2019). Analisa Dan Implementasi Association Rule Dengan Algoritma Fp-Growth. *Teknologi Dan Sistem Informasi Bisnis*, 1(1), 1–11.
- PELAWI, B. P. (2019). Penerapan Algoritma Fp-Growth Dalam Menemukan Hubungan Data Prestasi Akademik Dengan Etika Mahasiswa (Study Kasus : POLTEKKES KEMENKES RI Medan). *Jurikom*, 6(2), 2–8.
- SISKA HARYANTI, A. S. (2015). Implementasi Data Mining Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma C4 . 5. *Media Infotama*, 11(2), 1–9.
- Syahputra, T., Halim, J., & Sintho, E. P. (2018). Penerapan Data Mining Dalam Menentukan Pilihan Jurusan Bidang Studi SMA Menggunakan Metode. *Penerapan Data Mining Dalam Menentukan Pilihan Jurusan Di Bidang Studi SMA Menggunakan Metode Clustering Dengan Teknik Single Linkage JURTEKSI*, IV(2), 1–4.
- WINARTI, D., STUDI, P., INFORMATIKA, T., KOMPUTER, F. I., INDONESIA, U. D., BARAT, S., & ASOSIASI, R. (2018). Analisis Data Mining Dengan Algoritma Fp-Growth Dalam Mendukung Strategi Promosi. *SIMTIKA*, 1(1), 27–31.