

IMPLEMENTASI ALGORITMA FOLD-GROWTH UNTUK MENEMUKAN POLA KELULUSAN MAHASISWA

Muhammad Hasbi Assidiqqi¹, Alwis Nazir², Iwan Iskandar³, Jasril⁴
Fitri Insani⁵

Teknik Informatika Fakultas Sains dan Teknologi, UIN Sultan Syarif Kasim Riau
¹11950115130@students.uin-suska.ac.id, ²alwis.nazir@uin-suska.ac.id,
³iwan.iskandar@uin-suska.ac.id, ⁴jasril@uin-suska.ac.id,
⁵fitri.insani@uin-suska.ac.id

Diterima pada 26 November 2022; disetujui pada 30 Januari 2023; dan diterbitkan pada 3 Maret 2023.

Abstrak

Kelulusan mahasiswa tepat waktu merupakan hal yang penting untuk sebuah program studi dan perguruan tinggi. Menurut Peraturan BAN-PT Nomor 23 Tahun 2022 tentang Instrumen Pemantauan dan Evaluasi Peringkat Akreditasi Perguruan Tinggi, presentase kelulusan mahasiswa tepat waktu idelanya adalah $\geq 37,5\%$ untuk perguruan tinggi akademik dan $\geq 47,5\%$ untuk perguruan tinggi vokasi. Kelulusan mahasiswa tepat waktu pada Program Studi Teknik Informatika UIN Sultan Syarif Kasim Riau setiap tahunnya masih dibawah standar yang telah ditetapkan oleh BAN-PT. Pada tahun 2021 jumlah mahasiswa yang lulus tepat waktu hanya sebesar 12%. Dengan menggunakan data mahasiswa berupa jalur masuk serta data kelulusan berupa lama studi dan IPK, dilakukan penggalian pola kelulusan menggunakan algoritma *FOLD-Growth*. Algoritma *FOLD-Growth* merupakan algoritma gabungan dari *FOLDARM* dan *FP-Growth*. Dengan 274 data mahasiswa, nilai minimum *support* 20%, minimum *confidence* 50% menghasilkan dua pola dengan nilai *lift ratio* terkuat adalah 3,61, sedangkan dengan minimum *support* 10% dan minimum *confidence* 50%, menghasilkan tiga pola dengan *lift ratio* terkuat adalah 7,03. Hasil dari penelitian ini dapat digunakan sebagai pedoman dalam memperbaiki tingkat kelulusan mahasiswa tepat waktu, dengan cara memberikan kuota penerimaan mahasiswa baru yang lebih besar pada jalur yang menghasilkan kelulusan tepat waktu, sehingga akan meningkatkan kelulusan mahasiswa tepat waktu.

Kata Kunci: Kelulusan mahasiswa, *Association Rule*, *lift ratio*, *FOLD-Growth*, *FP-Growth*

Abstract

Timely graduation of students is important for a study program and higher education institution. According to BAN-PT Regulation Number 23 of 2022 concerning Instruments for Monitoring and Evaluation of Higher Education Accreditation Ratings, the ideal percentage of student graduation on time is $\geq 37.5\%$ for academic tertiary institutions and $\geq 47.5\%$ for vocational tertiary institutions. The timely graduation of Informatics Engineering Study Program students at UIN Sultan Syarif Kasim Riau every year is still below the standards set by BAN-PT. In 2021 the number of students who graduate on time is only 12%. By using student data in the form of acceptance and graduation data in the form of length of study and GPA, excavation of graduation patterns is carried out using the FOLD-Growth algorithm. The FOLD-Growth algorithm is a combined algorithm of FOLDARM and FP-Growth. With 274 student data, a minimum support value of 20%, and a minimum confidence of 50%, it produces two patterns with the strongest lift ratio value of 3.61, while with a minimum support of 10% and a minimum confidence of 50%, it produces three patterns. with the strongest lift ratio of 7.03. The results of this study can be used as a guideline for increasing student graduation rates on time, by providing a larger intake quota for new students on a pathway that results in on-time graduation, so that it will increase student graduation on time.

Keywords: Student graduation, *Association Rule*, *lift ratio*, *FOLD-Growth*, *FP-Growth*

1. Pendahuluan

Perguruan tinggi dan lembaga pendidikan berperan penting untuk melakukan pengembangan pendidikan

dalam upaya membangun bangsa [1, 2]. Perguruan tinggi mengharuskan mahasiswa untuk melaksanakan tridarma perguruan tinggi sebagai syarat menyelesaikan studinya.

Studi yang telah selesai ditandai dengan pemberian ijazah kelulusan [3]. Kelulusan mahasiswa tepat waktu yang tinggi merupakan kesuksesan dalam membangun strategi pendidikan dan penerapan sistem belajar mengajar [4]. Angka kelulusan tepat waktu yang tinggi juga merupakan salah satu penilaian akreditasi perguruan tinggi [5, 6].

Menurut Peraturan BAN-PT Nomor 23 Tahun 2022 tentang Instrumen Pemantauan dan Evaluasi Peringkat Akreditasi Perguruan Tinggi, presentase kelulusan mahasiswa tepat waktu pada sebuah program studi idealnya adalah $\geq 37,5\%$ untuk perguruan tinggi akademik, dan $\geq 47,5\%$ untuk perguruan tinggi vokasi [5]. Pada Program Studi Teknik Informatika UIN Sultan Syarif Kasim Riau masih banyak mahasiswa yang lulus tidak tepat waktu. Berdasarkan hasil wawancara dengan Ketua Program Studi Teknik Informatika, yaitu Bapak Iwan Iskandar, S.T., M.T., jumlah kelulusan mahasiswa Teknik Informatika tepat waktu setiap tahunnya masih dibawah standar yang sudah ditetapkan oleh BAN-PT. Pada tahun 2021 kelulusan mahasiswa tepat waktu Teknik Informatika UIN Sultan Syarif Kasim Riau hanya sebesar 12%.

Dengan menggunakan data mahasiswa berupa jalur masuk dan data kelulusan berupa lama masa studi dan IPK, dapat diperoleh keterkaitan diantara kedua variable data itu mempergunakan *data mining*. Dimana *data mining* adalah proses menggali dan mengelola data mentah dalam suatu wadah yang besar untuk mendapatkan pengetahuan baru (*knowledge discovery*) [4]. Pengetahuan baru yang diperoleh dari proses *data mining* memungkinkan dilakukannya pemanfaatan seperti pengembangan strategi pendidikan, atau pengelolaan sumber daya dengan lebih baik. Teknik dalam *data mining* yang dapat digunakan salah satunya adalah teknik asosiasi. Teknik asosiasi adalah metode untuk mencari suatu pola yang mirip atau kombinasi item yang sering muncul dalam suatu transaksi [7].

Penelitian tentang pencarian pola kelulusan mahasiswa sebelumnya pernah dilakukan oleh [8] menggunakan algoritma *FOLD-Growth* dengan 95 data, minimum *support* 4%, minimum *confidence* 30% menghasilkan 7 *rule*. Sedangkan menggunakan minimum *support* 4%, minimum *confidence* 60% menghasilkan 4 *rule*. Menggunakan data yang sama, dengan minimum *support* 8%, minimum *confidence* 30% menghasilkan 4 *rule*. Sedangkan menggunakan minimum *support* 8%, minimum *confidence* 60% menghasilkan 3 *rule*. Penelitian serupa juga pernah dilakukan oleh [9] dengan menggunakan data mahasiswa dan data kelulusan mahasiswa Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Gunung Djati Bandung dari tahun 2006 sampai 2009, minimum *support* 20%, minimum *confidence* 50% menghasilkan 8 *rule*. Selain itu, penelitian tentang analisis kinerja

algoritma *FOLD-Growth* dan *FP-growth* pada penggalian pola asosiasi pernah dilakukan oleh [10] yang melibatkan dataset sintetik, dari penelitian tersebut menghasilkan sejumlah pola yang bervariasi, dimana jumlah dan panjang dari pola dipengaruhi oleh ukuran dataset, nilai minimum *support* dan minimum *confidence*. Secara umum dapat disimpulkan bahwa durasi eksekusi, skalabilitas, realibilitas dan utilitas memori *FOLD-Growth* lebih baik dibandingkan dengan *FP-Growth*.

Dengan menggunakan teknik asosiasi dan algoritma *FOLD-Growth* penelitian ini dilakukan untuk mendapatkan informasi berupa pola keterkaitan antara jalur masuk terhadap tingkat kelulusan berupa lama masa studi dan IPK pada Program Studi Teknik Informatika UIN Sultan Syarif Kasim Riau. Hasil dari penelitian ini dapat digunakan sebagai pedoman dalam memperbaiki tingkat kelulusan mahasiswa tepat waktu, dengan cara memberikan kuota penerimaan mahasiswa baru yang lebih besar pada jalur yang menghasilkan kelulusan tepat waktu, sehingga akan meningkatkan kelulusan mahasiswa tepat waktu.

2. Metodologi Penelitian

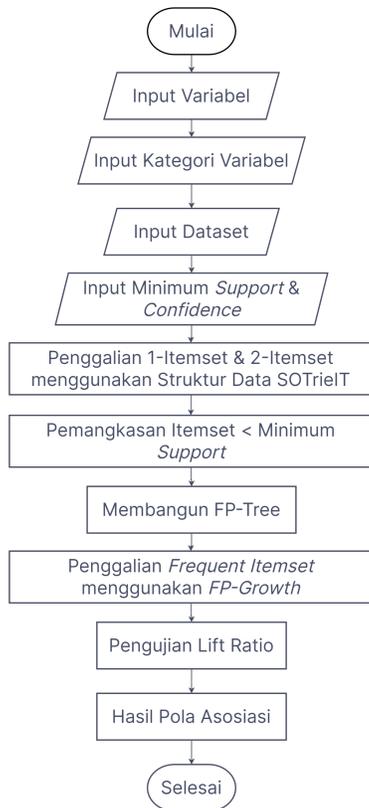
Metode yang dipergunakan yaitu metode eksperimen dengan tahapan proses *Knowledge Discovery in Database* (KDD). Tahap pertama yang dilakukan adalah studi literatur tentang *association rule* dan algoritma *FOLD-Growth*.

Association rule adalah sebuah aturan dalam mencari pola asosiasi. Aturan asosiasi menggunakan klausa *IF – THEN* [11]. Contoh: Jika membeli kopi, maka juga akan membeli gula. Terdapat dua parameter yang mendasari aturan asosiasi untuk dikatakan penting atau tidaknya suatu kombinasi item, yaitu nilai *support* (pendukung) dan nilai *confidence* (kepercayaan) [7]. *Support* merupakan frekuensi seberapa sering kombinasi item tersebut muncul. Nilai *support* menentukan apakah sebuah kombinasi item layak untuk dicari nilai *confidence*-nya. Sedangkan *confidence* merupakan parameter kuat atau tidaknya hubungan kombinasi item tersebut. Contoh: apabila membeli kopi, maka membeli gula, dengan *support* 60% dan *confidence* 80%. Artinya jika pelanggan membeli kopi, maka pelanggan juga akan membeli gula dengan kepastian 80% dan frekuensi kemunculan 60%. Untuk mencari nilai *support* dan *confidence* digunakan persamaan berikut [8].

$$Support(X) = \frac{\sigma(X)}{N} \quad (1)$$

$$Support(X, Y) = \frac{\sigma(X, Y)}{N} \quad (2)$$

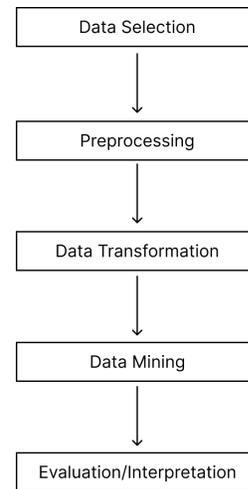
$$Confidence(X \rightarrow Y) = \frac{Support(X, Y)}{Support X} \quad (3)$$



Gambar 1. Flowchart Sistem

Langkah selanjutnya adalah mengumpulkan data. Data yang dikumpulkan meliputi data sekunder yakni data mahasiswa Teknik Informatika UIN Sultan Syarif Kasim Riau dan data kelulusannya dari tahun 2019-2022. Data tersebut diperoleh melalui koordinasi dengan devisa data pada Pusat Teknologi Informasi dan Pangkalan Data (PTIPD) UIN Sultan Syarif Kasim Riau. Langkah berikutnya adalah menganalisis dan merancang sistem yang akan dibangun terkait minimum spesifikasi dan komponen komponen apa saja yang akan dipergunakan, *Flowchart* dari sistem dapat diamati dalam Gambar 1.

Pada Gambar 1 menjelaskan alur dari sistem yang akan dibangun. Proses pertama adalah menginputkan variabel apa saja yang akan digunakan. Setelah variabel diinputkan, proses selanjutnya adalah menginputkan apa saja kategori atau *value* dari masing-masing variabel yang telah diinputkan sebelumnya. Proses ketiga adalah menginputkan *dataset* dalam format *csv*. Setelah *dataset* terinput, proses selanjutnya menginputkan nilai minimum *support* dan minimum *confidence*. Setelah diinputkan maka sistem akan melakukan penggalian *1-itemset* dan *2-itemset* dalam bentuk data *SOTriET*. Selanjutnya sistem akan melakukan pemangkasan *itemset* yang tidak *frequent*, setelah proses pemangkasan dilakukan maka akan dibentuk *FP-Tree*. Setelah proses membentuk *FP-Tree* selesai, sistem melakukan penggalian *frequent itemset* dan proses terakhir yaitu



Gambar 2. Tahapan Proses KDD

menampilkan hasil berupa pola.

Gambar 2 menjelaskan mengenai tahapan-tahapan yang ada pada *Knowledge Discovery in Database (KDD)*. Tahapan KDD yang pertama adalah *data selection*, yaitu menyeleksi atau memilih variabel data mana yang akan digunakan [12]. Tahapan ini bertujuan agar penelitian lebih berfokus pada variabel data yang paling berpengaruh sehingga lebih efisien dan efektif. Variabel data yang akan dipergunakan pada penelitian ini merujuk pada penelitian yang telah dilakukan sebelumnya oleh [8, 9]. Variabel yang digunakan pada data mahasiswa yakni NIM dan jalur masuk sedangkan data kelulusan berupa IPK dan masa studi. NIM hanya digunakan sebagai *primary key* atau *id* agar tidak terjadi duplikasi data.

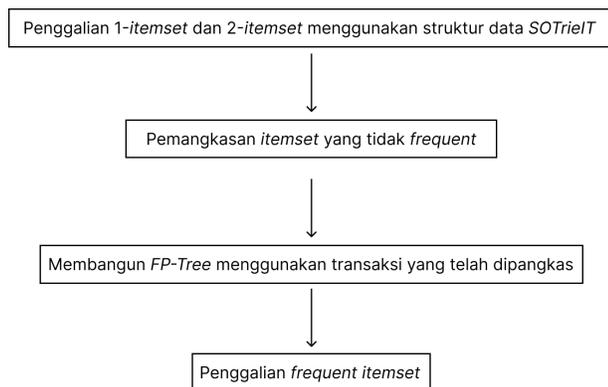
Tahapan kedua KDD yaitu *preprocessing*. Pada tahapan ini dilakukan pembersihan data (*cleaning*) dari data *noise*, *missing value*, *redundant*, *inconsistent* dan data yang tidak relevan agar tidak mengganggu jalannya program dan hasil yang didapatkan akurat [12]. Setelah data dibersihkan pada tahapan *preprocessing*. Tahapan ketiga pada KDD adalah *data transformation*. Dalam tahap ini dilakukan perubahan bentuk data sehingga sesuai dengan kebutuhan pengolahan algoritma yang digunakan [12]. Selain itu, dilakukannya proses transformasi data adalah untuk menyederhanakan data yang panjang. Pada tahapan *data transformation*, variabel NIM dihapus karena tidak akan digunakan pada proses asosiasi dan dilakukan perubahan data dengan mengkombinasikan IPK dan masa studi menjadi variabel kategori kelulusan seperti Tabel 1 berikut.

Tahapan keempat pada KDD yaitu *data mining*. pada tahapan *data mining* yaitu mengimplementasikan metode atau algoritma yang telah ditentukan [12]. Pada penelitian ini menggunakan teknik asosiasi dan juga algoritma *FOLD-Growth*.

Algoritma *FOLD-Growth* adalah algoritma guna

Kategori Kelulusan	Keterangan
A1	Masa studi ≤ 4 tahun dan IPK 3,50 - 4,00
A2	Masa studi ≤ 4 tahun dan IPK 3,00 - 3,49
A3	Masa studi ≤ 4 tahun dan IPK 2,50 - 2,99
A4	Masa studi ≤ 4 tahun dan IPK 2,00 - 2,49
B1	Masa studi > 4 tahun dan IPK 3,50 - 4,00
B2	Masa studi > 4 tahun dan IPK 3,00 - 3,49
B3	Masa studi > 4 tahun dan IPK 2,50 - 2,99
B4	Masa studi > 4 tahun dan IPK 2,00 - 2,49

Tabel 1. Ketentuan Transformati Data Kategori Kelulusan



Gambar 3. Tahapan Algoritma FOLD-Growth

mencari pola asosiasi atau hubungan keterkaitan dengan menerapkan struktur data *SOTrieIT* (*Support Order-Trie Itemset*). Pada algoritma *FOLD-Growth* menggabungkan antara dua algoritma, yaitu *FOLDARM* dan *FP-Growth*. Pemilihan algoritma *FOLD-Growth* dikarenakan algoritma ini mampu menggabungkan kelebihan dari algoritma *FOLDARM* dan *FP-Growth*, yaitu mempunyai kinerja cepat dalam pengukuran *frequent itemset* maksimum (*Kmax*) yaitu kecil ($Kmax \leq 10$) dan memiliki kerja cepat pada saat *frequent itemset* maksimum (*Kmax*) adalah besar ($Kmax > 10$). Selain itu, algoritma *FOLD-Growth* memiliki durasi eksekusi, skalabilitas, realibilitas dan utilitas memori lebih baik dibandingkan dengan *FP-Growth* [10]. Tahapan algoritma *FOLD-Growth* dapat dilihat pada Gambar 3.

Tahapan pertama algoritma *FOLD-Growth* yaitu penggalan *1-itemset* dan *2-itemset* yang dilakukan dengan membaca data yang telah di transformasi sebelumnya, kemudian akan dibangkitkan kemungkinan *1-itemset* dan *2-itemset* yang selanjutnya disusun kedalam struktur data *SOTrieIT*. Tahapan kedua adalah

pemangkasan *itemset* yang tidak *frequent*, yaitu dengan menghitung nilai *support* menggunakan persamaan 1 untuk *1-itemset* dan persamaan 2 untuk *2-itemset*, kemudian memangkas *item* yang tidak mencukupi minimum *support* yang telah ditentukan sebelumnya. Tahapan ketiga membangun *FP-Tree* mempergunakan berbagai transaksi yang sudah dipangkas. Tahapan keempat yaitu penggalan *frequent itemset* dengan menghitung nilai *confidence* menggunakan persamaan 3. Jika nilai *confidence* besar dari minimum *confidence* yang telah ditetapkan, maka *frequent itemset* tersebut memenuhi *association rule* sehingga dapat digunakan sebagai pola.

Tahapan kelima pada KDD adalah evaluasi atau interpretasi hasil. Pada tahapan ini akan dilakukan evaluasi melalui pengujian pola menggunakan pengujian *lift ratio*. *Lift ratio* yaitu proses pengujian hasil untuk mendapatkan nilai kevalidan dari hasil tersebut [13, 14]. Proses *lift ratio* dilakukan dengan membandingkan antara nilai *confidence* dan nilai *benchmark confidence*. Apabila hasil dari pengujian $lift\ ratio \geq 1$, maka pola tersebut kuat [14]. Pengujian *lift ratio* dapat dilakukan menggunakan persamaan 4 dan *benchmark confidence* menggunakan persamaan 5. Selanjutnya adalah menginterpretasikan hasil. Interpretasi adalah proses mendefinisikan atau menafsirkan hasil yang diperoleh [12].

$$Lift\ Ratio = \frac{Confidence\ X,Y}{Benchmark\ Confidence\ X,Y} \quad (4)$$

$$Benchmark\ Confidence = \frac{Nc}{N} \quad (5)$$

Keterangan:

Nc = jumlah transaksi dengan item yang menjadi consequent, *N* = jumlah transaksi basis data.

3. Hasil dan Pembahasan

3.1 Perhitungan Manual

Semakin tinggi nilai minimum *support* dan minimum *confidence* maka akan semakin sedikit pola yang dihasilkan [8]. Merujuk pada penelitian [9] pada perhitungan manual dilakukan dengan minimum *support* 20%, minimum *confidence* 50%. dan menggunakan 30 sampel data yang sudah melalui tahapan *data selection*, *preprocessing* dan *data transformation*. 30 sampel data dapat dilihat pada Tabel 2.

Kemudian akan dibangkitkan kemungkinan *1-itemset* dan *2-itemset* yang disusun kedalam struktur data *SOTrieIT*. Pada data yang ada, *1-itemset* yang dibangkitkan adalah jalur masuk dan *2-itemset* adalah kombinasi antara jalur masuk dengan kategori kelulusan. Pembangkitan *itemset* kedalam struktur *SOTrieIT* dapat dilihat pada Gambar 4.

Tahapan selanjutnya dari algoritma *FOLD-Growth* adalah membaca setiap *itemset* yang telah dibangkitkan

Jalur Masuk	Kategori Kelulusan
SNMPTN	B2
PMB Jalur Undangan	B1
PMB Jalur Undangan	B1
PMB Jalur Undangan	B1
SNMPTN	B2
SBMPTN	B1
PMB Jalur Undangan	B1
SBMPTN	B2
SBMPTN	B1
PMB Jalur Undangan	B2
SNMPTN	B1
SBMPTN	B2
SNMPTN	B2
PMB Jalur Undangan	B2
PMB Ujian Mandiri	B2
SPAN-PTKIN	B2
SPAN-PTKIN	B2
SNMPTN	B2
SBMPTN	B2
PMB Jalur Undangan	B2
SBMPTN	B1
SBMPTN	B2
PMB Jalur Undangan	B2
SBMPTN	B2
PMB Ujian Mandiri	B2
PMB Jalur Undangan	B3
SNMPTN	B2
PMB Jalur Undangan	B2
PMB Jalur Undangan	B1
PMB Jalur Undangan	B2

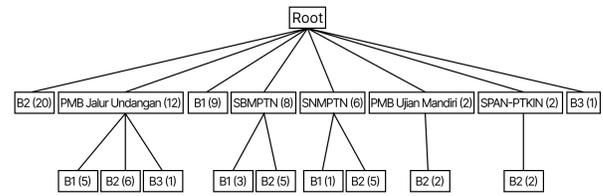
Tabel 2. Sampel Data

dan menghitung nilai *support* mempergunakan persamaan 1 untuk *1-itemset* dan persamaan 2 untuk *2-itemset*. Hasil dari perhitungan nilai *support 1-itemset* dapat diamati dalam Tabel 3 dan *2-itemset* pada Tabel 4.

Setelah dilakukan penghitungan nilai *support*, Langkah berikutnya yaitu melakukan pemangkasan *itemset* yang tidak memenuhi nilai minimum *support* yang sudah ditetapkan sebelumnya. Hasil dari *1-itemset* dan *2-itemset* setelah dilakukan pemangkasan dapat diamati dalam Tabel 5 dan Tabel 6.

Tahapan selanjutnya pada algoritma *FOLD-Growth* adalah membangun *FP-Tree* dari sisa *itemset* yang telah dipangkas sebelumnya. *FP-Tree* yang terbentuk dari proses sebelumnya digambarkan seperti Gambar 5.

Setelah *FP-Tree* terbentuk, tahapan selanjutnya adalah melakukan penggalan *frequent itemset* menggunakan *FP-Growth*. Pada tahapan ini terbagi menjadi dua fase, yang pertama adalah *conditional pattern base* yang dilakukan dengan mencari *suffix pattern* dan *prefix path* dari *FP-Tree*. Fase kedua pada penggalan *frequent itemset* menggunakan *FP-Growth*



Gambar 4. Itemset dalam Bentuk Struktur SOTrieIT

1-Itemset	Count	Support
B2	20	0,67
PMB Jalur Undangan	12	0,40
B1	9	0,30
SBMPTN	8	0,27
SNMPTN	6	0,20
SPAN-PTKIN	2	0,07
PMB Ujian Mandiri	2	0,07
B3	1	0,03

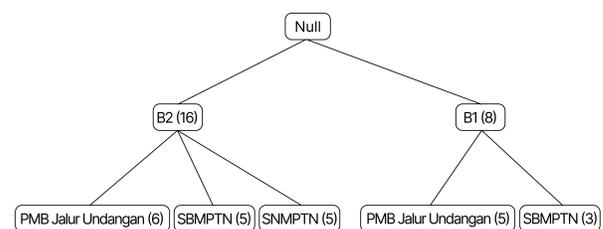
Tabel 3. Hasil Perhitungan Support 1-Itemset

adalah *conditional FP-Tree* dengan cara menjumlahkan *support itemset* dalam *prefix path* pada masing masing *suffix*. *Conditional pattern base* dan *conditional FP-Tree* dapat dilihat pada Tabel 7 dan Tabel 8.

Tahapan selanjutnya adalah menghitung nilai *confidence* dari *frequent itemset* yang terbentuk menggunakan persamaan 3. Hasil perhitungan nilai *confidence* dapat dilihat pada Tabel 9.

Tahapan berikutnya adalah melakukan pemangkasan *frequent itemset* yang tidak memenuhi nilai minimum *confidence* yang telah ditetapkan sebelumnya. Hasil dari pemangkasan *frequent itemset* yang tidak memenuhi minimum *confidence* dapat diamati dalam Tabel 10.

Setelah mendapatkan pola yang memenuhi nilai minimum *confidence*, selanjutnya dilakukan pengujian menggunakan *lift ratio* dengan perhitungan seperti persamaan 4. Jika pola yang dihasil mempunyai nilai *lift ratio* ≥ 1 . Maka pola tersebut kuat. Hasil pengujian lift



Gambar 5. FP-Tree yang Terbentuk dari Itemset

2-Itemset	Count	Support
PMB Jalur Undangan, B1	5	0,17
PMB Jalur Undangan, B2	6	0,20
PMB Jalur Undangan, B3	1	0,03
SBMPTN, B1	3	0,10
SBMPTN, B2	5	0,17
SNMPTN, B1	1	0,03
SNMPTN, B2	5	0,17
PMB Ujian Mandiri, B2	2	0,07
SPAN-PTKIN, B2	2	0,07

Tabel 4. Hasil Perhitungan Support 2-Itemset

1-Itemset	Count	Support
B2	20	0,67
PMB Jalur Undangan	12	0,40
B1	9	0,30
SBMPTN	8	0,27
SNMPTN	6	0,20

Tabel 5. Hasil Pemangkasan 1-itemset

ratio dapat diamati dalam Tabel 11.

Tahap selanjutnya adalah interpretasi hasil. Interpretasi adalah proses mendefinisikan atau menafsirkan hasil yang diperoleh [12]. Berdasarkan hasil dari tahapan sebelumnya, maka didapatkan satu pola yaitu, jika mahasiswa masuk melalui PMB jalur Undangan, maka mahasiswa tersebut berada pada kategori kelulusan B2, yaitu lulus lebih dari 4 tahun dan IPK berada pada range 3.00-3.49, dengan frekuensi kemunculan 20% dan keyakinan 50%.

3.2 Perhitungan Sistem

Sistem bangun menggunakan database MySQL dan bahasa pemrograman PHP dengan mengikuti perhitungan manual. Alur sistem dapat diamati dalam Gambar 1. Setelah menginputkan variabel, kategori variabel dan dataset, selanjutnya adalah menginputkan nilai minimum support dan confidence dan sistem akan berjalan melakukan proses penggalan 1-itemset dan 2-itemset menggunakan struktur SOTrieIT beserta dengan nilai support itemset. Selanjutnya sistem akan melakukan pemangkasan itemset yang nilai support-nya tidak mencukupi nilai minimum support yang telah ditentukan sebelumnya. Selanjutnya sistem akan membangun FP-Tree dari itemset yang tersisa dan dilakukan penggalan frequent itemset beserta nilai confidence-nya. Selanjutnya sistem akan melakukan pemangkasan frequent itemset yang nilai confidence-nya tidak memenuhi nilai minimum confidence yang telah ditentukan. Selanjutnya sistem akan melakukan pengujian itemset yang tersisa menggunakan lift ratio.

Berdasarkan data mahasiswa dan kelulusan pada Program Studi Teknik Informatika UIN Sultan Syarif

2-Itemset	Count	Support
PMB Jalur Undangan, B2	6	0,20

Tabel 6. Hasil Pemangkasan 2-itemset

Suffix	Conditional Pattern Base	Support
PMB Jalur Undangan	{B2:6}	0,20

Tabel 7. Conditional Pattern Base

Frequent itemset	Frekuensi	Support
PMB Jalur Undangan, B2	6	0,20

Tabel 8. Conditional FP-Tree

Frequent itemset	Frekuensi	Support	Confidence
PMB Jalur Undangan, B2	6	0,20	0,50

Tabel 9. Hasil Perhitungan Nilai Confidence

Frequent itemset	Frekuensi	Support	Confidence
PMB Jalur Undangan, B2	6	0,20	0,50

Tabel 10. Hasil Pemangkasan Confidence

Frequent itemset	Frequent Item Consequent	Confidence	Benchmark Confidence	Lift Ratio
PMB Jalur Undangan, B2	6	0,50	0,20	2,5

Tabel 11. Hasil Pengujian Lift Ratio

Kasim Riau dari tahun 2019-2023 yang telah melalui tahapan data selection, preprocessing dan data transformation, dengan jumlah data sebanyak 274 data, didapatkan dua pola. Pola pertama, jika jalur masuk SNMPTN maka kategori kelulusannya adalah B2 atau lulus lebih dari 4 tahun dan IPK berada di range 3.00-3.49, dengan frekuensi kemunculan 20.8%, nilai kepastian 75% dan nilai lift ratio 3.61. Pola kedua, jika jalur masuk SBMPTN maka kategori kelulusannya adalah B2 atau lulus lebih dari 4 tahun dan IPK berada di range 3.00-3.49, dengan frekuensi kemunculan 22.6%, kepastian 79,22% dan nilai lift ratio 3.56. Hasil dari implementasi algoritma FOLD-Growth tersebut dapat dilihat pada Tabel 12.

No.	Pola	Support	Confidence	Lift Ratio
1.	SNMPTN → B2	57/274 = 20.8%	57/76 = 75%	3.61
2.	SBMPTN → B2	61/274 = 22.26%	61/77 = 79.22%	3.56

Tabel 12. Hasil Implementasi *FOLD-Growth*

4. Kesimpulan dan Saran

Berdasarkan penelitian yang sudah dibuat, algoritma *FOLD-Growth* dapat melakukan penggalian pola kelulusan mahasiswa dengan variabel data yang sedikit. Dengan 274 data, minimum support 20%, minimum confidence 50% menghasilkan dua pola dengan lift ratio terkuat adalah 3.61, sedangkan dengan minimum support 10% dan minimum confidence 50%, menghasilkan tiga pola dengan lift ratio terkuat adalah 7,03.

Berdasarkan penelitian yang dilakukan, terdapat beberapa saran untuk penelitian berikutnya supaya hasil yang didapatkan lebih akurat, yaitu dengan mempergunakan jumlah data yang lebih banyak serta menggunakan nilai minimum support dan confidence yang lebih bervariasi.

Daftar Pustaka

- [1] Y. hendayana, "Dirjen dikti: Perguruan tinggi menjadi mata air bagi pembangunan bangsa," Feb 2021.
- [2] H. Indriyawati and T. Winarti, "Analisa data mining kemampuan lulusan dengan kebutuhan stakeholder menggunakan algoritma apriori (studi kasus universitas semarang)," *Proceeding SENDI U*, pp. 499–505, Jul. 2021. [Online]. Available: <https://www.unisbank.ac.id/ojs/index.php/sendu/article/view/8641>
- [3] A. Anas and A. J. Zebua, "Implementasi algoritma apriori untuk mendapatkan pola kelulusan mahasiswa," *Jurnal Ilmiah Media Sisfo*, vol. 16, no. 1, p. 54–61, Apr. 2022. [Online]. Available: <https://ejournal.unama.ac.id/index.php/mediasisfo/article/view/650>
- [4] I. Farida and S. W. H. L. Hendric, "Prediksi pola kelulusan mahasiswa menggunakan teknik data mining classification emerging pattern," *Petir*, vol. 12, no. 1, pp. 1–17, 2019.
- [5] P. BAN-PT. (2022) Peraturan badan akreditasi nasional perguruan tinggi nomor 1 tahun 2022 tentang mekanisme akreditasi untuk akreditasi yang dilakukan oleh badan akreditasi nasional perguruan tinggi. [Online]. Available: https://www.banpt.or.id/wp-content/uploads/2022/02/Peraturan-BAN-PT-Nomor-1-2022-tentang-/Mekanisme-Akreditasi_FINAL.pdf
- [6] M. Yatimah, "Implementasi data mining untuk prediksi kelulusan tepat waktu mahasiswa stimik esq menggunakan decision tree c4.5," *JUMANJI (Jurnal Masyarakat Informatika Unjani)*, vol. 5, no. 2, pp. 89–98, 2021. [Online]. Available: <https://jumanji.unjani.ac.id/index.php/jumanji/article/view/95>
- [7] W. T. Saputro, H. M. Jumasa, and M. Murhadi, "Algoritma apriori untuk menemukan pola aturan asosiasi," *INTEK : Jurnal Informatika dan Teknologi Informasi*, vol. 3, no. 1, pp. 9–16, May 2020. [Online]. Available: <https://jurnal.umpwr.ac.id/index.php/intek/article/view/167>
- [8] A. S. U. Dewi. (2021) Pola asosiasi jalur masuk terhadap kelulusan mahasiswa menggunakan algoritma fold-growth (studi kasus jurusan teknik informatika uin Maulana Malik Ibrahim Malang). [Online]. Available: <http://etheses.uin-malang.ac.id/32312/>
- [9] M. Irfan, "Analisa pola asosiasi jalur masuk terhadap kelulusan mahasiswa dengan menggunakan metode fold-growth (studi kasus fakultas sains dan teknologi)," *JURNAL ISTEK*, vol. Vol 9, No 2 (2015): ISTEK, 2015.
- [10] N. M. A. W. Rully Soelaiman, "Analisis kinerja algoritma fold-growth dan fp-growth pada penggalian pola asosiasi," *SNATI 2006 : Bidang Pengolahan Citra dan Sinyal*, vol. 2006, no. Snati, p. 13–18, 2006.
- [11] W. Sahara, S. D. Saragih, and A. P. Windarto, "Teknik asosiasi datamining dalam menentukan pola penjualan dengan metode apriori," *TIN: Terapan Informatika Nusantara*, vol. 2, no. 12, p. 684–689, 2022.
- [12] Y. Mardi, "Data mining : Klasifikasi menggunakan algoritma c4.5," *Edik Informatika*, vol. 2, no. 2, p. 213–219, 2017.
- [13] M. Tahir and N. Sitompul, "Penerapan algoritma fp-growth dalam menentukan kecenderungan mahasiswa mengambil mata kuliah pilihan," *Network Engineering Research Operation*, vol. 6, no. 1, pp. 56–63, 2021. [Online]. Available: <https://nero.trunojoyo.ac.id/index.php/nero/article/view/216>
- [14] L. zahrotun dan Tedy Setiadi dan Taufik Haryadi, "Aplikasi data mining untuk mencari pola asosiasi tracer study menggunakan algoritma foldarm," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 4, no. 1, pp. 37–43, 2018. [Online]. Available: <https://teknosi.fti.unand.ac.id/index.php/teknosi/article/view/352>