

Implementasi *FP-Growth* dan *Fuzzy Tsukamoto* untuk Menentukan Persentase Kuota Jalur Masuk Perguruan Tinggi

Nafa Khairunnisa^{1*}, Jumadi², Ichsan Taufik³

^{1,2,3}Fakultas Sains dan Teknologi, Teknik Informatika, UIN Sunan Gunung Djati, Bandung, Indonesia

E-mail: ^{1*}1207050088@student.uinsgd.ac.id, ²jumadi@uinsgd.ac.id, ³ichsan@uinsgd.ac.id

(* : corresponding author)

Abstrak

Setiap universitas berusaha mencapai atau mempertahankan akreditasi unggul. Mahasiswa yang lulus dengan predikat memuaskan berperan penting dalam menentukan akreditasi tersebut. Menurut BAN-PT 2021, universitas dianggap unggul jika memiliki mahasiswa dengan masa studi maksimal 4,5 tahun dan IPK rata-rata $\geq 3,25$. Salah satu cara untuk mempertahankannya adalah dengan mengelola persebaran kuota jalur masuk Penerimaan Mahasiswa Baru (PMB) secara optimal. Penelitian ini bertujuan untuk menyelidiki bagaimana variabel IPK, masa studi, dan jalur masuk pada data wisudawan berhubungan satu sama lain. Untuk mendapatkan pola asosiasi antar variabel tersebut maka digunakanlah *FP-Growth*. Selanjutnya, persentase distribusi kuota dihitung menggunakan *Fuzzy Tsukamoto*. Metode ini diuji dengan *Mean Absolute Error* (MAE) dan diperoleh nilai MAE sebesar 5.585. Dari penelitian ini didapatkan akurasi model sebesar 94.42% dan nilai presisi 62.5%, yang menandakan bahwa metode yang digunakan efektif dalam membantu menentukan distribusi kuota yang optimal untuk PMB. Dengan demikian, hasil ini dapat digunakan untuk mendukung kebijakan universitas dalam menentukan persebaran kuota yang lebih tepat guna mendukung pencapaian akreditasi unggul.

Kata kunci: *FP-Growth*, *fuzzy Tsukamoto*, KDD, *python*, *RapidMiner*

Abstract

Every university strives to achieve or maintain excellent accreditation. Students who graduate with a satisfactory predicate play an important role in determining the accreditation. According to BAN-PT 2021, a university is considered excellent if it has students with a maximum study period of 4.5 years and an average GPA ≥ 3.25 . One way to maintain it is to optimally manage the distribution of quotas for the New Student Admission (PMB) entrance pathway. This study aims to investigate how the variables of GPA, study period, and entry path in the data of graduates relate to each other. To get the association pattern between these variables, *FP-Growth* is used. Furthermore, the percentage of quota distribution is calculated using *Fuzzy Tsukamoto*. This method was tested with *Mean Absolute Error* (MAE) and obtained an MAE value of 5.585. From this research, the accuracy of the model is 94.42% and the precision value is 62.5%, which indicates that the method used is effective in helping determine the optimal quota distribution for PMB. Thus, these results can be used to support university policies in determining a more appropriate quota distribution to support the achievement of superior accreditation.

Keywords: *FP-Growth*, *fuzzy Tsukamoto*, KDD, *python*, *RapidMiner*

1. PENDAHULUAN

Wisuda menjadi agenda tahunan yang krusial bagi setiap universitas. Lulusan dengan predikat memuaskan dapat berkontribusi terhadap peningkatan akreditasi institusi Pendidikan terkait. Menurut pedoman BAN-PT tahun 2021, universitas dinilai unggul jika memiliki mahasiswa dengan masa studi maksimal 4,5 tahun dan IPK rata-rata $\geq 3,25$ [1]. Upaya untuk mencapai kriteria ini dapat dimulai dari tahap penerimaan mahasiswa baru dengan memfokuskan seleksi calon mahasiswa berpotensi unggul. Untuk mendukung hal tersebut, analisis data wisudawan memiliki urgensi dalam mengevaluasi hubungan antara jalur masuk PMB, masa studi, dan IPK. Analisis ini juga diperlukan untuk mengidentifikasi pengaruh jalur masuk terhadap keberhasilan akademik mahasiswa. Namun, pengaruh ini masih memerlukan bukti berbasis data.

Banyak jalur masuk yang bisa ditempuh mahasiswa untuk masuk ke dalam sebuah universitas. Terdapat delapan jalur masuk jenjang S1 di Universitas Islam Negeri (UIN) Sunan Gunung Djati Bandung, yaitu Seleksi Nasional Berdasarkan Prestasi (SNBP), Seleksi Nasional Berdasarkan Tes (SNBT), Seleksi Mandiri, Prestasi, Ujian Masuk Bersama Lainnya, Penelusuran

Minat dan Kemampuan (PMDK), Program Kerjasama Perusahaan/Institusi/Pemerintah, dan Program Internasional.

Data mining menjadi pendekatan yang relevan untuk menemukan pola dan hubungan dalam dataset wisudawan. *Data mining* adalah cabang dari *data science* yang digunakan untuk mengidentifikasi pola tersembunyi, tren, atau hubungan dalam data [2]. Pada penelitian ini, teknik *data mining* diterapkan untuk menganalisis data wisudawan berdasarkan tiga variabel utama, yaitu jalur masuk PMB sebagai variabel pembeda berdasarkan karakteristik awal mahasiswa, IPK sebagai indikator kualitas akademik, dan masa studi sebagai indikator efisiensi waktu penyelesaian studi. Pola ini digunakan sebagai dasar untuk rekomendasi distribusi kuota PMB. Sebagai contoh, mahasiswa dengan IPK tinggi cenderung menyelesaikan studi lebih cepat. Sementara itu, jalur masuk PMB juga dapat memengaruhi IPK dan masa studi, karena setiap jalur memiliki persyaratan dan karakteristik yang berbeda.

Dalam penelitian ini, dua metode digunakan, yaitu *FP-Growth* untuk menemukan pola hubungan antar variabel dalam bentuk *Association Rules* dan *Fuzzy Tsukamoto* untuk mengevaluasi kelayakan pola tersebut sebagai dasar penentuan distribusi kuota jalur masuk. *FP-Growth* digunakan untuk menghitung kecenderungan atau pola yang sering muncul dalam jumlah transaksi antar variabel yang disebut dengan *Association Rule* [3]. Pada penelitian ini, *Fuzzy Tsukamoto* digunakan untuk memberikan nilai bobot persentase kuota berdasarkan aturan *fuzzy* yang dibentuk.

Algoritma serupa digunakan dalam studi kasus sebelumnya. Penelitian yang dilakukan oleh Jafar dkk berhasil mengimplementasikan *FP-Growth* untuk menentukan pola pemesanan hotel dari data reservasi tahun 2017 dan 2018 [4]. Pola ini kemudian dapat dijadikan sebagai acuan pengembangan layanan hotel agar dapat sesuai dengan pola perilaku konsumen. Penelitian lainnya membandingkan dua algoritma asosiasi yaitu Apriori dengan *FP-Growth* untuk menentukan akurasi tata letak barang dari 8 *item* sebanyak 131 data [5]. Penelitian tersebut membuktikan bahwa *FP-Growth* lebih unggul karena dapat menghasilkan 6 rule dengan akurasi 128% menggunakan *lift ratio*. Pada tahun 2021, Annisa Almira dkk meneliti pola pencurian daya listrik untuk menentukan target operasi atau jenis pelanggaran potensial [6]. Hasilnya dari 10 data sampel berhasil menemukan 13 *association rules* yang tiga di antaranya memiliki nilai *confidence* 100%. *FP-Growth* juga pernah digunakan pada penelitian lain untuk mencari pola pemesanan makanan dan minuman di kedai distrik [7]. Data yang digunakan yaitu pembelian selama periode 1-31 Oktober sebanyak 2.360 transaksi. Penelitian ini berhasil menghasilkan 1 aturan dengan nilai *support* 0,161% dan nilai *confidence* 0.806%.

Selain *FP-Growth*, banyak penelitian sebelumnya yang menggunakan *Fuzzy Tsukamoto*. Pada tahun 2022, Calvin C. C. dkk melakukan penelitian menggunakan *Fuzzy Tsukamoto* untuk mengetahui pasokan BBM pada Pertashop sehingga dapat menurunkan terjadinya kekurangan stok [8]. Para peneliti tersebut menggunakan data penjualan dari tahun 2017-2020. *Fuzzy Tsukamoto* berhasil diimplementasikan dengan hasil akurasi sebesar 87%. Di tahun yang sama, sebuah penelitian menggunakan *Fuzzy Tsukamoto* untuk menentukan rekomendasi pemberian beasiswa di Universitas Kristen Indonesia Toraja (UKI Toraja) [9]. Dataset yang digunakan mencakup Indeks Prestasi (IP), tanggungan orang tua, penghasilan orang tua, catatan tambahan, dan surat rekomendasi. Hasilnya terdapat sekitar 192 *rule* terbentuk dan akurasi sebesar 86%. *Fuzzy Tsukamoto* digunakan pada penelitian dengan studi kasus berbeda pada tahun 2024. Studi ini menerapkan *Fuzzy Tsukamoto* untuk penentuan kelayakan siswa dapat mengikuti Seleksi Nasional Masuk Perguruan Tinggi Negeri (SNMPTN) di MAN 1 Kota Makassar dengan menggunakan dataset nilai pengetahuan, nilai keterampilan, prestasi, dan jumlah alumni yang lulus tahun sebelumnya [10]. Hasilnya menunjukkan bahwa akurasi *Fuzzy Tsukamoto* mencapai 76,6%.

Penelitian sebelumnya pernah menggabungkan algoritma asosiasi dengan *fuzzy* yaitu Apriori dengan *fuzzy normalization* untuk meningkatkan performa algoritma yang optimal [11]. Dengan *fuzzy normalization* di tahap *preprocessing* dapat menghasilkan rata-rata *confidence* 88,4% dan rata-rata *support* 6,8%, lebih baik dibandingkan tanpa *fuzzy*. Hal ini membuktikan bahwa algoritma asosiasi dengan *fuzzy* dapat digabungkan dan memiliki hasil yang signifikan.

Penelitian sebelumnya telah menunjukkan bahwa *FP-Growth* efektif dalam menemukan pola data, sementara *Fuzzy Tsukamoto* dapat membantu pengambilan keputusan berbasis aturan. Namun, penelitian sebelumnya hanya terbatas pada pola hubungan data atau keputusan berbasis aturan tanpa menghasilkan rekomendasi distribusi kuota yang mendukung kriteria akreditasi unggul. Selain itu, kombinasi *FP-Growth* dan *Fuzzy Tsukamoto* jarang ditemukan dalam penelitian yang secara spesifik menganalisis hubungan jalur masuk mahasiswa baru dengan kualitas akademik dan efisiensi studi. Oleh karena itu, penelitian ini menawarkan pendekatan baru dengan mengintegrasikan kedua metode tersebut untuk memberikan rekomendasi distribusi kuota yang lebih optimal.

Studi ini memberikan kontribusi baru dengan mengintegrasikan *FP-Growth* untuk menemukan pola hubungan antara jalur masuk, IPK, dan masa studi, serta *Fuzzy Tsukamoto* untuk mengevaluasi kelayakan pola tersebut. Kombinasi kedua metode ini memberikan rekomendasi distribusi kuota jalur masuk berbasis data yang mendukung peningkatan jumlah mahasiswa berkualitas. Pendekatan ini tidak hanya mendukung universitas dalam pengelolaan penerimaan mahasiswa baru secara strategis, tetapi juga berkontribusi pada pencapaian akreditasi unggul sesuai pedoman BAN-PT tahun 2021.

2. METODE PENELITIAN

Pendekatan *Knowledge Discovery in Database* (KDD) akan digunakan pada penelitian ini. Metode KDD sangat cocok digunakan dalam penelitian ini karena fokusnya pada eksplorasi dan analisis data besar untuk menemukan pola-pola atau hubungan yang berguna. KDD melibatkan berbagai tahap, seperti pemrosesan dan transformasi data, yang memungkinkan pengolahan data yang kompleks, seperti data wisudawan, secara efisien. Dengan menggabungkan teknik *FP-Growth* untuk menemukan pola asosiasi dan *Fuzzy Tsukamoto* untuk evaluasi, KDD dapat memberikan wawasan yang mendalam dalam menentukan persentase rekomendasi persebaran kuota jalur masuk PMB berdasarkan hubungan antar variabel yang ada.

2.1. Metode Pengumpulan Data

Teknik sekunder digunakan untuk mengumpulkan data, yaitu melalui permohonan dataset wisudawan ke-96 dan ke-97 kepada bagian akademik UIN Sunan Gunung Djati Bandung. Data yang digunakan berupa data lulusan jenjang S1 UIN Sunan Gunung Djati. Proses pengumpulan berlangsung dari tanggal 2 hingga 20 Agustus 2024, dengan total 2.000 data wisudawan. Selain itu, data jumlah jalur masuk S1 PMB tahun 2024 juga dikumpulkan untuk keperluan evaluasi.

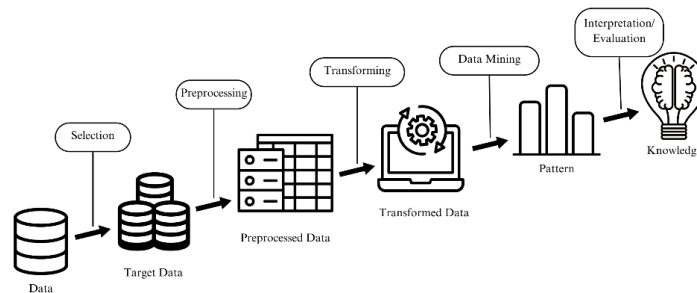
2.2. Kebutuhan *Hardware* dan *Software*

Adapun alat yang digunakan untuk melakukan penelitian ini ditunjukkan pada tabel 1. *RapidMiner* digunakan untuk mengimplementasikan *FP-Growth*. Kemudian bahasa pemrograman *Python* digunakan pada *Google Colab* untuk membersihkan data dan menghitung hasil kelayakan oleh *Fuzzy Tsukamoto*.

Tabel 1. Kebutuhan *Hardware* dan *Software*

No.	Komponen	Keterangan
1.	<i>Hardware</i>	Laptop
2.	<i>Software</i>	RapidMiner
		Python
		Google Colab

2.3. Knowledge Discovery in Database (KDD)



Gambar 1. Tahapan KDD

Gambar 1 menunjukkan tahapan proses KDD. Metode KDD adalah salah satu metode *data mining* yang digunakan untuk mengekstraksi wawasan dari *database*, yang termasuk proses pemrosesan dan transformasi data. KDD terdiri dari 5 tahapan yang berurutan[12], yaitu:

a. Data Selection

Tujuan dari proses ini adalah untuk memilih dan memeriksa kumpulan data yang telah dikumpulkan. Kumpulan data yang dipilih harus relevan dan memenuhi persyaratan analisis yang akan digunakan. Data yang dipilih harus mencakup semua elemen penelitian yang diperlukan [4].

b. Pre-Processing

Pre-processing meliputi proses *cleaning data* dengan menghapus data-data yang tidak diperlukan, termasuk data duplikasi dan *missing value*. *Missing value* dihapus dari data sehingga data menjadi konsisten. Data untuk analisis harus bersih dan siap diolah, sehingga dapat menghasilkan hasil analisis yang akurat. Proses pembersihan ini menggunakan *Python*.

c. Data Transformation

FP-Growth merupakan teknik *data mining* yang memerlukan format data khusus sebelum diaplikasikan. Misalnya, *data numerical* seperti IPK dan masa studi diubah menjadi data *nominal* menggunakan proses *binning*, kemudian dikonversi lagi menjadi data *binominal*. Transformasi ini penting untuk mempermudah proses analisis.

d. Data Mining

Kemudian peneliti melakukan *data mining* untuk menghasilkan informasi yang bermanfaat dengan menggunakan berbagai metode untuk mengidentifikasi pola-pola dalam data. Pada penelitian ini, *FP-Growth* diimplementasikan menggunakan tools *RapidMiner* untuk mempermudah proses perhitungan dan meningkatkan efisiensi analisis. *Minimum support* dan *minimum confidence* diatur pada nilai tertentu untuk memunculkan delapan jalur masuk PMB dalam hasil analisis.

e. Knowledge Interpretation/Evaluation

Selanjutnya yaitu proses interpretasi dan pengambilan keputusan berdasarkan pola yang ditemukan. Hasil dari *FP-Growth* berupa *frequent itemsets* dan *association rules* yang diambil sebagian berdasarkan jalur masuk yang ada. *Association rules* ini akan digunakan sebagai aturan pada metode *Fuzzy Tsukamoto*, yang juga memanfaatkan aturan dari BAN-PT. *Output* dari *Fuzzy Tsukamoto* berupa kelayakan data, yang menentukan apakah data tersebut dapat dijadikan dasar rekomendasi kuota jalur masuk. Hasil akhirnya adalah perhitungan persentase yang jika dijumlahkan menghasilkan total 100%.

Evaluasi diperlukan untuk mengetahui konsistensi pola dan informasi yang ditemukan dengan data aktual. Pada penelitian ini, hasil persentase dari *Fuzzy Tsukamoto* dibandingkan dengan persentase PMB UIN Sunan Gunung Djati tahun 2024. Akurasi dihitung berdasarkan selisih antara persentase hasil analisis dan persentase aktual jalur masuk.

2.4. Frequent Pattern Growth (FP-Growth)

FP-Growth adalah algoritma asosiasi yang dinamis dan efektif dalam mencari pola, hubungan, atau kecenderungan dari data transaksi. Algoritma ini dikembangkan oleh Han dan Kamber pada tahun 2000 dan telah digunakan dalam berbagai aplikasi, seperti analisis penjualan,

persediaan produk, dan analisis kecelakaan lalu lintas. Algoritma ini memungkinkan pencarian aturan asosiasi secara efektif, bahkan pada dataset besar dan kompleks. *FP-Growth* dapat memberikan hasil yang lebih akurat dalam menemukan aturan asosiatif, karena menggunakan struktur *tree* yang memungkinkan pengelompokan item yang sering dibeli secara bersamaan. Tahapan *FP-Growth* terdiri dari beberapa langkah yang berikut.

- a. Pembangkitan *Conditional Pattern Base*: Untuk menemukan pola yang sering muncul, *FP-Growth* menghasilkan *Conditional Pattern Base* (CPB), yang berisi semua itemset yang terjadi dalam dataset.
- b. Pembangkitan *Conditional FP-Tree*: *FP-Growth* menghasilkan *Conditional FP-Tree* (CFPT) dari CPB. CFPT digunakan untuk mencari *itemset* yang sering muncul dan mengurangi ukuran dataset.
- c. Pencarian *Frequent Itemset*: *FP-Growth* menggunakan CFPT untuk menemukan *itemset* yang sering muncul dan relevan untuk aturan asosiasi
- d. Pencarian Nilai *Support* dan *Confidence*: *FP-Growth* menghitung nilai *support* dan *confidence* untuk menilai kekuatan aturan asosiasi yang ditemukan.

2.5. Fuzzy Tsukamoto

Konsep *fuzzy logic* memanfaatkan derajat keanggotaan dalam suatu himpunan untuk mengatasi ketidakpastian sehingga keanggotaan tidak hanya bersifat benar atau salah [13]. Ada tiga sistem *inference fuzzy* yaitu Tsukamoto, Mamdani, dan Sugeno [13]. *Fuzzy Tsukamoto* adalah metode *fuzzy logic* yang digunakan untuk membuat keputusan dengan menggunakan aturan yang tidak pasti dan data yang tidak pasti. *Fuzzy Tsukamoto* ini memiliki beberapa keunggulan diantaranya yaitu efisiensi dalam pengambilan keputusan, kemampuan menghasilkan *output* yang jelas dari data tidak pasti, aplikasi yang luas dalam berbagai bidang, dan dapat mengatasi *taboo range* [14].

Pada *Fuzzy Tsukamoto*, setiap implikasi berupa yang direpresentasikan ke dalam himpunan *fuzzy* yaitu aturan IF-THEN *membership function* monoton [15]. Model tersebut ditunjukkan pada gambar 2.

If (X is A) and (Y is B) Then (Z is C)

Gambar 2. Model Umum *Fuzzy Tsukamoto*

Berikut tahapan *Fuzzy Tsukamoto* [15]:

- a. Fuzzifikasi, yaitu mengkonversi data yang tidak pasti menjadi *fuzzy set* untuk membuat keputusan berdasarkan aturan-aturan yang tidak pasti.
- b. *Inference Engine*, dalam sistem logika *fuzzy* adalah komponen yang menerapkan aturan-aturan *fuzzy* untuk memproses data input dan menghasilkan output. Pada tahap ini, aturan berbentuk "IF-THEN" diterapkan untuk mengevaluasi variabel input menggunakan *membership function*. Hasil dari evaluasi aturan-aturan ini kemudian digabungkan untuk menghasilkan keputusan yang lebih tepat.
- c. Defuzzifikasi, mengkonversi output *fuzzy* menjadi output yang jelas dan dapat diinterpretasikan. Defuzzifikasi dilakukan dengan menggunakan proses *Weighted Average Method* pada persamaan (1) untuk menghasilkan nilai tegas.

$$z = \frac{\sum a_i z_i}{\sum a_i} \quad (1)$$

Keterangan:

z = *Weighted Average Method*

a_i = Pembobot untuk periode i

z_i = Nilai untuk periode i

3. HASIL DAN PEMBAHASAN

3.1 Data Selection

Data baru ditambahkan yaitu tanggal masuk kuliah yang diatur secara *default* tanggal 1 September, tahunnya diambil dari digit kedua dan ketiga Nomor Induk Mahasiswa (NIM). Selanjutnya, masa studi mahasiswa dihitung dengan mengurangi tanggal daftar wisuda dari tanggal masuk. Dengan demikian, terdapat tiga atribut yang digunakan untuk penelitian ini yaitu IPK, masa studi, dan jalur masuk.

3.2 Pre-processing

Data wisudawan jenjang S1 UIN Sunan Gunung Djati digunakan. *Cleaning data* dilakukan dengan membersihkan data-data yang tidak diperlukan, termasuk data duplikasi dan data kosong. Dari 2000 data, diperoleh 1814 data bersih. Data yang sudah bersih dapat dilihat di tabel 2.

Tabel 2. Data yang Sudah Dibersihkan

Jalur Masuk	ipk	masa_studi_tahun
Seleksi Nasional Berdasarkan Tes (SNBT)	3.8	4.9
Seleksi Nasional Berdasarkan Tes (SNBT)	3.59	4.1
Seleksi Mandiri	3.6	4.1
...
Seleksi Nasional Berdasarkan Tes (SNBT)	3.54	4.1

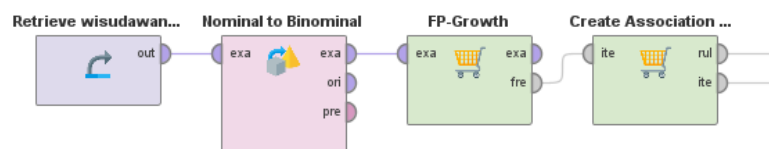
3.3 Data Transformation

Pada proses ini, transformasi data dilakukan dengan *binning data* untuk menyesuaikan format yang dapat diimplementasikan dalam *FP-Growth*. Algoritma *FP-Growth* dapat dijalankan jika variabel bertipe data binominal. IPK dan lama studi dalam tahun yang merupakan data numerik dikonversi ke format nominal. Variabel IPK dikonversi menjadi Rendah, Sedang, dan Tinggi. Nilai IPK 0-2.75 dikategorikan Rendah, 2.75-3.5 dikategorikan Sedang, dan 3.5-4.0 dikategorikan Tinggi. Sedangkan variabel masa_studi_tahun diubah menjadi Cepat, Normal, dan Lama. Masa studi 0-3.5 tahun dikategorikan cepat, 3.5-4.0 tahun dikategorikan normal, dan 4-8 tahun dikategorikan lama. Pada tahap ini, variabel IPK dan lama studi tahun sudah bertipe data nominal. Variabel Jalur Masuk sudah berupa tipe nominal jadi tidak perlu dikonversi lagi. Data yang sudah ditransformasi ini bisa dilihat pada tabel 3.

Tabel 3. Data Hasil Transformasi

Jalur Masuk	ipk	lama_studi_tahun
Seleksi Nasional Berdasarkan Tes (SNBT)	Tinggi	Lama
Seleksi Nasional Berdasarkan Tes (SNBT)	Tinggi	Lama
Seleksi Mandiri	Tinggi	Lama
...
Seleksi Nasional Berdasarkan Tes (SNBT)	Tinggi	Lama

Selanjutnya, data yang sudah ditransformasi tersebut dimasukkan ke dalam *software RapidMiner* untuk diolah.



Gambar 3. Desain Model Operator pada *RapidMiner*

Gambar 3 menunjukkan desain model operator pada *RapidMiner*. Pada desain tersebut, terdapat 4 operator yang digunakan yaitu, *retrieve data wisudawan*, *Nominal to Binominal*, *FP-Growth*, dan *Create Association Rules*. Operator yang digunakan selanjutnya yaitu *Nominal to*

Binominal yang berfungsi untuk mengonversi tipe data dari nominal ke binominal. Pada *FP-Growth* terdapat parameter *minimum support*. *Minimum support* diatur ke 0,015 atau 1,5%. Ini artinya frekuensi kemunculan *itemset* diatur ke tingkat 0,015 untuk minimalnya dan maksimalnya 1. Parameter *minimum confidence* diatur ke 0,3 atau setara 30% pada operator *Create Association Rules*, yang menunjukkan bagaimana dua item berhubungan dalam kondisi kemiripan data tertentu.

3.4 Data Mining

Desain model yang sudah dibuat kemudian dijalankan. Masing-masing operator yang digunakan memiliki kegunaan. *Retrieve* data wisudawan berfungsi untuk menyimpan dan memanggil data yang akan digunakan. Operator *Nominal to Binominal* untuk mengonversi format tipe data nominal menjadi binominal agar dapat diolah pada *FP-Growth*. Operator *FP-Growth* digunakan untuk pengolahan *minimum support*. Kemudian yang terakhir terdapat operator *Create Association Rules* untuk pengolahan nilai *minimum confidence*.

3.5 Knowledge Interpretation/Evaluation

Size	Support	Item 1	Item 2	Item 3
1	0.839	ipk = Tinggi		
1	0.710	masa_studi_tahun = N...		
1	0.525	Jalur Masuk = Seleksi ...		
1	0.290	masa_studi_tahun = L...		
1	0.149	ipk = Sedang		
1	0.148	Jalur Masuk = Ujian M...		
1	0.141	Jalur Masuk = Seleksi ...		
1	0.110	Jalur Masuk = Seleksi ...		
1	0.045	Jalur Masuk = Prestasi		
1	0.019	Jalur Masuk = Penelus...		
1	0.012	ipk = Rendah		
1	0.008	Jalur Masuk = Progra...		
1	0.004	Jalur Masuk = Progra...		
2	0.660	ipk = Tinggi	masa_studi_tahun = N...	
2	0.428	ipk = Tinggi	Jalur Masuk = Seleksi ...	

Gambar 4. Hasil *Frequent Itemset*

Pada Gambar 4 menunjukkan adanya 69 *frequent itemset* yang berhasil dibentuk. Didapatkan nilai *support* terkecil sebesar 0.004 dan nilai tertinggi sebesar 0.839. *Support* terkecil 0.004 berarti *itemset* paling jarang ditemukan hanya muncul di 0,4% dari total data, sedangkan *support* tertinggi 0.839 berarti *itemset* paling sering ditemukan muncul di 83,9% dari total data. Data memiliki variasi pola dari yang jarang hingga sering muncul, dengan pola teratas muncul di lebih dari separuh dataset.

```

Association Rules
[Jalur Masuk = Prestasi] --> [masa_studi_tahun = Lama] (confidence: 0.305)
[ipk = Sedang] --> [masa_studi_tahun = Normal] (confidence: 0.306)
[Jalur Masuk = Seleksi Mandiri] --> [masa_studi_tahun = Lama] (confidence: 0.310)
[ipk = Sedang, Jalur Masuk = Ujian Masuk Bersama Lainnya] --> [masa_studi_tahun = Normal] (confidence: 0.310)
[masa_studi_tahun = Lama] --> [ipk = Tinggi, Jalur Masuk = Seleksi Mandiri] (confidence: 0.319)
[masa_studi_tahun = Lama, Jalur Masuk = Prestasi] --> [ipk = Sedang] (confidence: 0.320)
[ipk = Rendah] --> [masa_studi_tahun = Normal] (confidence: 0.333)
[masa_studi_tahun = Lama, Jalur Masuk = Ujian Masuk Bersama Lainnya] --> [ipk = Sedang] (confidence: 0.345)
[masa_studi_tahun = Lama] --> [ipk = Sedang] (confidence: 0.357)
[Jalur Masuk = Penelusuran Minat dan Kemampuan (PMDK)] --> [ipk = Tinggi, masa_studi_tahun = Lama] (confidence: 0.381)
[ipk = Rendah] --> [Jalur Masuk = Seleksi Mandiri, masa_studi_tahun = Lama] (confidence: 0.381)
[ipk = Sedang, Jalur Masuk = Seleksi Nasional Berdasarkan Prestasi (SNBP)] --> [masa_studi_tahun = Normal] (conf: 0.400)
[ipk = Tinggi] --> [masa_studi_tahun = Normal, Jalur Masuk = Seleksi Mandiri] (confidence: 0.400)
[Jalur Masuk = Seleksi Mandiri, masa_studi_tahun = Lama] --> [ipk = Sedang] (confidence: 0.403)
[masa_studi_tahun = Lama, Jalur Masuk = Seleksi Nasional Berdasarkan Prestasi (SNBP)] --> [ipk = Sedang] (confid: 0.403)
[ipk = Tinggi, Jalur Masuk = Penelusuran Minat dan Kemampuan (PMDK)] --> [masa_studi_tahun = Lama] (confidence: 0.439)
[ipk = Sedang] --> [Jalur Masuk = Seleksi Mandiri, masa_studi_tahun = Lama] (confidence: 0.439)
[Jalur Masuk = Penelusuran Minat dan Kemampuan (PMDK)] --> [masa_studi_tahun = Lama] (confidence: 0.457)
[masa_studi_tahun = Normal] --> [ipk = Tinggi, Jalur Masuk = Seleksi Mandiri] (confidence: 0.473)
[ipk = Tinggi, masa_studi_tahun = Normal] --> [Jalur Masuk = Seleksi Mandiri] (confidence: 0.508)
[masa_studi_tahun = Normal] --> [Jalur Masuk = Seleksi Mandiri] (confidence: 0.510)
[ipk = Tinggi] --> [Jalur Masuk = Seleksi Mandiri] (confidence: 0.511)
[masa_studi_tahun = Normal, ipk = Sedang] --> [Jalur Masuk = Seleksi Mandiri] (confidence: 0.518)
[ipk = Tinggi, masa_studi_tahun = Normal] --> [Jalur Masuk = Seleksi Mandiri] (confidence: 0.518)
    
```

Gambar 5. Hasil *Association Rules*

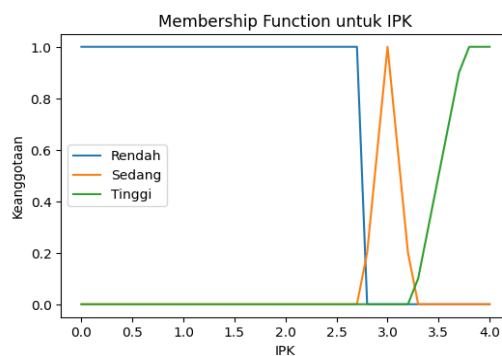
Pada Gambar 5 menunjukkan hasil *association rules* yang dibuat. Analisis menghasilkan 83 *association rules*, yaitu aturan hubungan antar *item* yang menunjukkan pola keterkaitan dalam data. Dengan nilai *confidence* terendah sebesar 0.305 dan tertinggi sebesar 1. *Confidence* terendah 0.305 berarti bahwa aturan dengan kepercayaan paling rendah hanya benar di 30,5% kasus, sedangkan *confidence* tertinggi 1 berarti aturan paling kuat benar di 100% kasus. Selain itu, nilai lift tertinggi yang terdeteksi yaitu 2.703 dan terendah bernilai 0.563.

Karena data terlalu sedikit, hanya ada satu pola atau hubungan yang ditemukan yang melibatkan jalur masuk Program Internasional. Pola yang muncul hanya menghubungkan jalur masuk ini dengan variabel IPK tanpa melibatkan masa studi, yang menunjukkan bahwa informasi dari jalur ini kurang kaya untuk analisis lebih mendalam. Akibatnya, algoritma *FP-Growth* tidak dapat menghasilkan aturan yang relevan untuk jalur ini.

Selanjutnya memasuki tahap *Fuzzy Tsukamoto* untuk menghitung kelayakan tiap data dijadikan pertimbangan menghitung persentase rekomendasi. Metode ini dilakukan dengan *Google Colab* menggunakan bahasa *Python*. Berikut tahapannya:

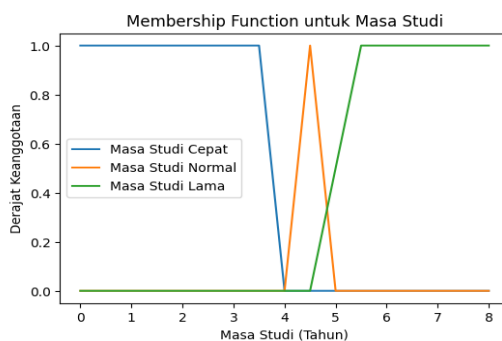
a. Fuzzifikasi

Proses fuzzifikasi dilakukan pada setiap variabel input, yaitu IPK dan masa_studi_tahun. Domain untuk IPK adalah 0 hingga 4, dengan interval 0.1. Variabel IPK memiliki tiga himpunan fuzzy yang menunjukkan nilai rendah (rentang nilai dari 0 hingga 2.75), sedang (rentang nilai dari 2.75 hingga 3.25), dan tinggi (rentang nilai dari 3.25 hingga 4.00). Rentang nilai ini disesuaikan dengan aturan BAN-PT tahun 2021. Tiga himpunan tersebut didefinisikan dengan *membership function* trapezoidal dan segitiga untuk masing-masing kategori yang ditunjukkan pada Gambar 6.



Gambar 6. Grafik *Membership Function* untuk IPK

Sementara itu, domain untuk masa_studi_tahun adalah 0 hingga 8 tahun. Variabel masa_studi_tahun memiliki tiga himpunan *fuzzy* yang menggambarkan tiga masa, yaitu cepat (rentang nilai dari 0 hingga 4.00), normal (rentang nilai dari 4.00 hingga 5.00), dan lama (rentang nilai dari 5.00 hingga 8.00). Rentang nilai ini disesuaikan dengan aturan BAN-PT tahun 2021. Gambar 7 menunjukkan tiga himpunan tersebut didefinisikan dengan *membership function* trapezoidal dan segitiga untuk masing-masing kategori.



Gambar 7. Grafik *Membership Function* untuk Masa Studi

b. *Inference Engine*

Inferensi *fuzzy* menggunakan hasil *FP-Growth* dengan nilai *threshold lift* ≥ 1.130 dan *confidence* teratas dari tiap jalur masuknya. Rules yang dibuat pun disesuaikan dengan aturan BAN-PT tahun 2021. *Rules* tersebut disajikan pada tabel 4. Setiap rules menghasilkan *crisp output* “Layak” atau “Tidak Layak”. Output ini bertujuan untuk memisahkan data mana yang layak dipertimbangkan dan tidak untuk perhitungan persentase.

Tabel 4. *Rules* untuk *Fuzzy Tsukamoto*

Rules	Jalur Masuk	Kondisi
R1	Seleksi Nasional Berdasarkan Prestasi (SNBP)	IF IPK Tinggi AND (Masa Studi Cepat OR Masa Studi Normal) THEN Layak IF (IPK Sedang OR Masa Studi Lama) THEN Tidak Layak
R2	Seleksi Nasional Berdasarkan Tes (SNBT)	IF IPK Tinggi AND (Masa Studi Cepat OR Masa Studi Normal) THEN Layak IF (IPK Sedang OR Masa Studi Lama) THEN Tidak Layak
R3	Seleksi Mandiri	IF IPK Tinggi AND (Masa Studi Cepat OR Masa Studi Normal) THEN Layak IF (IPK Rendah OR Masa Studi Lama) THEN Tidak Layak
R4	Prestasi	IF IPK Tinggi AND (Masa Studi Cepat OR Masa Studi Normal) THEN Layak IF (IPK Sedang OR Masa Studi Lama) THEN Tidak Layak
R5	Ujian Masuk Bersama Lainnya	IF IPK Tinggi AND (Masa Studi Cepat OR Masa Studi Normal) THEN Layak IF (IPK Rendah OR Masa Studi Lama) THEN Tidak Layak
R6	Program Kerjasama Perusahaan/Institusi/Pemerintah	IF IPK Tinggi AND (Masa Studi Cepat OR Masa Studi Normal) THEN Layak IF (IPK Sedang OR Masa Studi Lama) THEN Tidak Layak
R7	Penelusuran Minat dan Kemampuan (PMDK)	IF IPK Tinggi AND (Masa Studi Cepat OR Masa Studi Normal) THEN Layak IF (IPK Sedang OR Masa Studi Lama) THEN Tidak Layak
R8	Program Internasional	IF Masa Studi Cepat OR Masa Studi Normal THEN Layak IF Masa Studi Lama THEN Tidak Layak

c. Defuzzifikasi

Defuzzifikasi adalah langkah akhir dalam sistem logika *fuzzy* yang bertujuan mengonversi *output* dari tahap inferensi menjadi nilai bilangan nyata. Dari perhitungan *Weighted Average Method* dengan *cutoff* $z \geq 0.5$, didapatkan hasil kelayakan tiap jalur masuk. Tabel 5 menunjukkan hasil kelayakan tersebut.

Tabel 5. Hasil Kelayakan Tiap Jalur Masuk

Jalur Masuk	Layak	Tidak Layak
Seleksi Nasional Berdasarkan Prestasi (SNBP)	131	69
Seleksi Nasional Berdasarkan Tes (SNBT)	167	88
Seleksi Mandiri	623	329
Prestasi	54	28
Ujian Masuk Bersama Lainnya	205	63
Program Kerjasama Perusahaan/Institusi/Pemerintah	7	8
Penelusuran Minat dan Kemampuan (PMDK)	18	17
Program Internasional	7	0

Berdasarkan tabel 5, dapat dihitung persentase kelayakan berdasarkan proporsi nilai Layak Kategori terhadap total keseluruhan Layak menggunakan rumus persentase umum pada persamaan (2).

$$Persentase\ Layak = \frac{Layak\ Kategori}{\sum Seluruh\ Layak} \times 100\% \quad (2)$$

Berikut adalah proses perhitungannya:

- a. Seleksi Nasional Berdasarkan Prestasi (SNBP)

$$\frac{131}{1212} \times 100\% = 10.81\%$$
- b. Seleksi Nasional Berdasarkan Tes (SNBT)

$$\frac{167}{1212} \times 100\% = 13.78\%$$
- c. Seleksi Mandiri

$$\frac{623}{1212} \times 100\% = 51.40\%$$
- d. Prestasi

$$\frac{54}{1212} \times 100\% = 4.46\%$$
- e. Ujian Masuk Bersama Lainnya

$$\frac{205}{1212} \times 100\% = 16.91\%$$
- f. Program Kerjasama Perusahaan/Institusi/Pemerintah

$$\frac{7}{1212} \times 100\% = 0.58\%$$
- g. Penelusuran Minat dan Kemampuan (PMDK)

$$\frac{18}{1212} \times 100\% = 1.49\%$$
- h. Program Internasional

$$\frac{7}{1212} \times 100\% = 0.58\%$$

Persentase ini dijadikan rekomendasi untuk pihak universitas dalam menentukan distribusi kuota tiap jalur masuk. Secara keseluruhan, distribusinya yaitu 10.81% untuk jalur masuk SNBP, 13.78% untuk jalur masuk SNBT, 51.40% untuk jalur masuk Seleksi Mandiri, 4.46% untuk jalur masuk Prestasi, 16.91% untuk jalur masuk Ujian Masuk Bersama Lainnya, 0.58% untuk jalur masuk Program Kerjasama Perusahaan/Institusi/Pemerintah, 1.49% untuk jalur masuk PMDK, dan 0.58% untuk jalur masuk Program Internasional.

Selanjutnya, proses evaluasi dimulai dengan menghitung selisih antara persentase rekomendasi dengan persentase aktual yang diambil dari data jalur masuk PMB tahun 2024. Selisih yang lebih kecil menunjukkan bahwa *Fuzzy Tsukamoto* memiliki kinerja model yang baik.

Tabel 6. Selisih Persentase Tiap Jalur Masuk

Jalur Masuk	Persentase PMB tahun 2024 (%)	Persentase Rekomendasi (%)	Selisih Absolut (%)
Seleksi Nasional Berdasarkan Prestasi (SNBP)	7.79	10.81	3.02
Seleksi Nasional Berdasarkan Tes (SNBT)	11.87	13.78	1.91
Seleksi Mandiri	37.63	51.4	13.77
Prestasi	1.21	4.46	3.25
Ujian Masuk Bersama Lainnya	29.32	16.91	12.41
Program Kerjasama Perusahaan/Institusi/Pemerintah	0.26	0.58	0.32
Penelusuran Minat dan Kemampuan (PMDK)	11.41	1.49	9.92
Program Internasional	0.5	0.58	0.08
Jumlah			44.68
Rata-rata (<i>Mean Absolute Error</i>)			5.585

Berdasarkan tabel 6, didapatkan bahwa selisih tiap jalur masuk <14%. Nilai selisih terendah 0.08 pada jalur masuk Program Internasional menunjukkan bahwa perbedaan antara data rekomendasi dan data aktual hampir sama, dan nilai selisih tertinggi 13.77 diperoleh oleh jalur masuk Seleksi Mandiri menunjukkan bahwa terdapat perbedaan cukup besar antara rekomendasi dan data aktual.

Evaluasi dilakukan dengan metode *Mean Absolute Error* (MAE) untuk mengukur rata-rata kesalahan prediksi secara absolut. Nilai MAE yang didapatkan yaitu 5.585%. Data persentase yang digunakan menunjukkan selisih yang relatif kecil dan tidak memiliki *outlier* signifikan. Oleh karena itu, MAE sangat relevan untuk memberikan evaluasi yang adil terhadap kesalahan prediksi. Kemudian dari nilai MAE ini dihitung tingkat akurasi model menggunakan persamaan (3).

$$\text{Akurasi} = 1 - \frac{\text{MAE}}{100} = 1 - \frac{5.585}{100} = 0.9442 \quad (3)$$

Berdasarkan perhitungan di atas, hasilnya menunjukkan bahwa tingkat akurasi model mencapai 94.42%, yang mengindikasikan performa *Fuzzy Tsukamoto* dalam memprediksi kuota jalur masuk.

Selain itu, dilakukan penghitungan untuk mengetahui seberapa banyak prediksi yang berada dalam selisih $\pm 5\%$. Hasilnya menunjukkan bahwa nilai presisi mencapai 62.5% sebagaimana diperoleh dari persamaan (4).

$$\text{Presisi} = \frac{\text{Jumlah selisih} \leq 5}{8} = \frac{5}{8} = 0.625 \quad (4)$$

4. KESIMPULAN DAN SARAN

Berdasarkan hasil analisis menggunakan *software RapidMiner* dan *Python*, terdapat 83 aturan hubungan antar *item*, yang menunjukkan pola keterkaitan antar variabel IPK, masa studi, dan jalur masuk dalam data. *Frequent itemset* yang ditemukan sebanyak 69. Aturan hubungan ini digunakan sebagai aturan dalam sistem *Fuzzy Tsukamoto* bersamaan dengan aturan BAN-PT tahun 2021. Riset ini penting untuk memudahkan proses pengambilan keputusan dalam menentukan kuota penerimaan mahasiswa baru, sehingga kebijakan yang diambil lebih tepat, objektif, dan berdasarkan data.

Sistem *Fuzzy Tsukamoto* menunjukkan kinerja yang baik dalam mengevaluasi kelayakan mahasiswa berdasarkan data jalur masuk dan kriteria akademik, dengan nilai MAE mencapai 5.585, akurasi sebesar 94.42%, dan presisi 62.5%. Hal ini membuktikan bahwa model yang dirancang akurat dan dapat diandalkan untuk mendukung pengambilan keputusan dalam penentuan kelayakan data.

Penelitian ini menggunakan data wisudawan sebagai dasar untuk menentukan distribusi kuota jalur PMB, namun pendekatan ini memiliki keterbatasan karena tidak mencakup seluruh populasi mahasiswa, termasuk yang tidak lulus tepat waktu atau yang *drop out*. Untuk penelitian lebih lanjut, disarankan menggunakan data yang lebih komprehensif dengan mencakup seluruh mahasiswa, baik yang lulus maupun tidak, guna memberikan gambaran yang lebih menyeluruh dan akurat.

DAFTAR PUSTAKA

- [1] B. A. N. P. Tinggi, “Badan Akreditasi Nasional Perguruan Tinggi Jakarta 2020,” pp. 5–7, 2020.
- [2] C. Puspitasari, “Sekilas tentang Data Science, Data Mining, dan Machine Learning,” 2022. <https://binus.ac.id/malang/2022/05/sekilas-tentang-data-science-data-mining-dan-machine-learning/> (accessed Oct. 22, 2024).
- [3] I. M. Gufron, and U. Budiyanto, “Algoritma FP-Growth Untuk Mengkaji Pola Belanja Konsumen Pada Baby Shop By Netti,” *Seminar Nasional Mahasiswa Fakultas Teknologi Informasi*, vol. 1, no. 1, pp. 479–487, 2022.
- [4] J. Jafar, and N. Rahaningsih, “Menentukan Pola Reservasi Hotel dengan Algoritma Fp-Growth,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 1, pp. 540–546, 2023.
- [5] A. Anggrawan, M. Mayadi, and C. Satria, “Menentukan Akurasi Tata Letak Barang dengan Menggunakan Algoritma Apriori dan Algoritma FP-Growth,” *MATRIK J. Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 21, no. 1, pp. 125–138, 2021.
- [6] A. Almira, Suendri, and A. Ikhwan, “Implementasi Data Mining Menggunakan Algoritma

- FP-Growth pada Analisis Pola Pencurian Daya Listrik,” *Jurnal Informatika Universitas Pamulang*, vol. 6, pp. 442–448, 2021.
- [7] T. Dewi, R. Astuti, and Y. Arie Wijaya, “Analisis Pola Pembelian Makanan dan Minuman di Kedai Distrik Menggunakan Algoritma Fp-Growth,” *Jurnal Informatika dan Rekayasa Perangkat Lunak*, vol. 6, no. 1, pp. 238–243, 2024.
- [8] C. C. Citra, *et al.*, “Implementasi Metode Fuzzy Tsukamoto Dalam Menentukan Supply BBM Pada Pertashop,” *SKANIKA: Sistem Komputer dan Teknik Informatika*, vol. 5, no. 2, pp. 157–166, 2022.
- [9] P. Gloria, and E. Sedyono, “Perancangan Sistem Rekomendasi Pemberian Beasiswa dengan Metode Fuzzy Tsukamoto,” *ITA: Journal of Information Technology Ampera*, vol. 3, no. 2, pp. 124–147, 2022.
- [10] N. Rahmi, *et al.*, “Analisis Penerapan Fuzzy Tsukamoto dalam Penentuan Kelayakan Keikutsertaan Siswa pada SNMPTN,” *Jurnal Ilmiah SISFOTENIKA*, vol. 14, no. 2, pp. 185–196, 2024.
- [11] M. N. M. Br.Sipahutar, O. S. Sitompul, and S. Sutarman, “Association Rules Analysis Using Algorithm Apriori And Fuzzy Normalization,” in *2021 IEEE International Conference on Health, Instrumentation & Measurement, and Natural Sciences (InHeNce)*, 2021, pp. 1–5.
- [12] Q. Widayati, “Penerapan Data Mining Menggunakan Metode Teknik Classification (Qoriani Widayati),” *Jurnal Ilmiah Matrik*, vol. 20, no. 2, pp. 163–174, 2018.
- [13] A. Rindengan and A. . L. Yohanes, *Sistem Fuzzy*. 2019.
- [14] A. Burhanuddin, “Analisis Komparatif Inferensi Fuzzy Tsukamoto, Mamdani dan Sugeno Terhadap Produktivitas Padi di Indonesia,” *LEDGER: Journal Informatic and Information Technology*, vol. 2, no. 1, pp. 48–57, 2023.
- [15] N. Ratama, *Fuzzy Inference System Metode Tsukamoto dalam Melakukan Evaluasi Kinerja*. Eureka Media Aksara, 2023.