



## Model Optimalisasi Seleksi Penerimaan Beasiswa Perguruan Tinggi Swasta Menggunakan K-Means dan TOPSIS

Munawir Fikri Al-akbari<sup>1\*</sup>, Muhamad Arief Munandar<sup>2</sup>, Gandung Triyono<sup>3</sup>  
<sup>1,2,3</sup>Universitas Budi Luhur, Indonesia

\*email: 2311601914@student.budiluhur.ac.id

---

### Info Artikel

Dikirim: 22 Juli 2025  
Diterima: 1 Desember 2025  
Diterbitkan: 9 Desember 2025

### Kata kunci:

K-Means;  
Pendidikan Tinggi;  
Seleksi Beasiswa;  
Sistem Pendukung Keputusan;  
TOPSIS.

---

### ABSTRAK

Seleksi penerima beasiswa di perguruan tinggi swasta kerap menghadapi permasalahan subyektivitas dan kurangnya pendekatan yang sistematis. Untuk mengatasi hal tersebut, penelitian ini mengusulkan pendekatan hybrid menggunakan *K-Means Clustering* dan *Technique for Order of Preference by Similarity to Ideal Solution* (TOPSIS) guna mengoptimalkan proses seleksi beasiswa. Data mahasiswa yang mencakup aspek akademik (IPK), sosio-ekonomi (penghasilan dan pekerjaan orang tua, jumlah tanggungan keluarga), serta non-akademik (prestasi dan keaktifan organisasi) dianalisis menggunakan algoritma K-Means untuk mengelompokkan mahasiswa dengan karakteristik serupa. Validasi *Silhouette Score* menghasilkan empat kluster optimal dengan skor 0,1683. Distribusi mahasiswa menunjukkan mayoritas berada pada kluster menengah (61,2%), sementara kluster unggul dan khusus memiliki proporsi lebih kecil, masing-masing 21,3% dan 17,5%. Selanjutnya, metode TOPSIS diterapkan untuk menghasilkan pemeringkatan antar kluster. Kluster 4 memperoleh skor tertinggi (5,4535), diikuti oleh kluster 3 (0,6359), kluster 1 (0,6014), dan kluster 2 (0,5807). Analisis kontribusi atribut mengungkapkan bahwa IPK merupakan faktor dominan (48,61%–52,26%), diikuti oleh penghasilan orang tua (16,15%–19,59%) dan jumlah tanggungan keluarga (11,36%–12,09%). Hasil penelitian ini menunjukkan bahwa model K-Means dan TOPSIS mampu menyeimbangkan pendekatan merit-based dan need-based dalam seleksi beasiswa. Model yang dikembangkan memberikan dasar yang objektif untuk alokasi kuota beasiswa, serta berkontribusi pada pengembangan sistem pendukung keputusan yang lebih transparan, efisien, dan akuntabel.

---

## 1. PENDAHULUAN

Program beasiswa di Perguruan Tinggi Swasta (PTS) menjadi salah satu instrumen strategis bagi penerima dan pemberi beasiswa. Beasiswa tidak hanya berfungsi sebagai bantuan finansial, tetapi juga sebagai alat untuk menjangkau mahasiswa berprestasi dan meningkatkan reputasi institusi [1]. Di Indonesia, PTS memiliki kontribusi signifikan dalam memperluas akses pendidikan ini, sehingga optimalisasi program beasiswa menjadi sangat relevan. Meskipun demikian, proses seleksi beasiswa seringkali menghadapi tantangan signifikan. Jumlah pendaftar yang besar dengan kriteria penilaian yang beragam menjadikan proses seleksi secara manual menjadi tugas yang sangat rumit dan memakan waktu [2]. Ketergantungan pada metode konvensional ini juga berisiko tinggi terhadap subyektivitas. Proses pengambilan keputusan dengan metode konvensional ini rentan terhadap ketidaktepatan sasaran dan kurangnya transparansi dalam pengambilan

keputusan [3]. Akibatnya, beasiswa mungkin tidak sampai kepada kandidat yang paling layak dan membutuhkan.

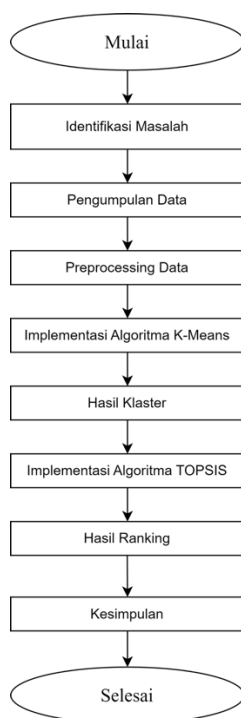
Untuk mengatasi permasalahan tersebut, pemanfaatan teknologi Sistem Pendukung Keputusan (SPK) menawarkan solusi yang menjanjikan. Penelitian ini mengusulkan sebuah model yang mengintegrasikan dua metode komputasi. Tahap pertama menggunakan algoritma *K-Means Clustering*. Pendekatan ini, seperti yang ditunjukkan dalam penelitian sebelumnya [4], efektif untuk melakukan pengelompokan (segmentasi) pendaftar ke dalam beberapa kluster berdasarkan kemiripan atribut. Pengelompokan ini bertujuan menyederhanakan dan memetakan pendaftar ke dalam kategori yang lebih homogen. Setelah pendaftar dikelompokkan, tahap selanjutnya adalah menerapkan metode *Technique for Order of Preference by Similarity to Ideal Solution* (TOPSIS) pada setiap kluster. TOPSIS merupakan metode pengambilan keputusan multi-kriteria yang efektif untuk melakukan perankingan, sebagaimana telah terbukti dalam berbagai studi kasus seleksi mahasiswa dan karyawan terbaik [5][6]. Metode ini bekerja dengan prinsip bahwa alternatif terbaik harus memiliki jarak terpendek dari solusi ideal positif dan terjauh dari solusi ideal negatif. Penggunaan kombinasi *K-Means* dan TOPSIS dapat menciptakan sistem yang lebih akurat, menjadi inti dari model yang diusulkan untuk menghasilkan sistem yang lebih objektif dan efisien [7].

Namun, belum banyak studi yang secara khusus menggabungkan *K-Means* dan TOPSIS secara integratif dalam konteks seleksi penerima beasiswa di Perguruan Tinggi Swasta (PTS). Penelitian-penelitian sebelumnya umumnya hanya menggunakan salah satu metode atau mengaplikasikannya pada domain yang berbeda. Oleh karena itu, penelitian ini bertujuan untuk mengembangkan model hybrid yang memanfaatkan *K-Means* sebagai tahap clustering dan TOPSIS sebagai tahap perankingan untuk menghasilkan rekomendasi penerima beasiswa yang lebih objektif, efisien, dan tepat sasaran di lingkungan PTS.

Penelitian ini menggunakan data primer mahasiswa dari salah satu perguruan tinggi swasta di Provinsi Riau, sebagai studi kasus dalam penerapan model. Dengan data tersebut, model yang dibangun diharapkan mampu mencerminkan kondisi nyata dan memberikan solusi yang aplikatif di lingkungan PTS. Tujuan dari penelitian ini adalah merancang dan membangun sebuah model *hybrid clustering* dan SPK yang mampu memberikan rekomendasi penerima beasiswa secara lebih optimal di lingkungan perguruan tinggi swasta berdasarkan ranking dari klusterisasi yang dilakukan. Dengan model ini, diharapkan pengelola dapat membuat keputusan yang lebih cepat dan berbasis data, sehingga program beasiswa dapat didistribusikan secara lebih adil dan tepat sasaran.

## 2. METODE PENELITIAN

Penelitian ini menggunakan data yang bersumber dari salah satu perguruan tinggi di provinsi Riau. Adapun tahapan penelitian ini dapat dilihat pada gambar 1.



Gambar 1. Tahapan Penelitian

Tahapan pertama pada penelitian ini yaitu melakukan identifikasi permasalahan yang terdapat pada distribusi beasiswa mahasiswa di salah satu perguruan tinggi swasta di provinsi Riau. Tahap selanjutnya adalah pengumpulan data yang diperlukan untuk menghasilkan distribusi beasiswa yang optimal. Setelah itu dilakukan pre-processing data yang bertujuan untuk melakukan eliminasi data yang tidak relevan, selain itu pada tahap ini dilakukan pemilihan atribut yang akan digunakan dan pembagian data yang akan digunakan sebagai data training dan data testing.

Tahap inti penelitian terdiri atas dua proses utama. Pertama, data mahasiswa dikelompokkan menggunakan algoritma *K-Means*. Algoritma ini dipilih karena memiliki kemampuan dalam mengelompokkan objek berdasarkan atribut multidimensi, sehingga mahasiswa dapat dikelompokkan ke dalam kluster yang lebih homogen dan representatif. Kedua, hasil klusterisasi dianalisis lebih lanjut dengan menerapkan algoritma TOPSIS. Metode ini dipilih karena efektif dalam menyelesaikan permasalahan pengambilan keputusan multikriteria secara objektif, dengan prinsip bahwa alternatif terbaik memiliki jarak terdekat dari solusi ideal positif dan terjauh dari solusi ideal negatif. Dengan demikian, TOPSIS mampu memberikan hasil pemeringkatan yang lebih transparan, akurat, dan adil. Hasil integrasi algoritma *K-Means* dan TOPSIS kemudian dievaluasi untuk memperoleh rekomendasi penerima beasiswa yang paling optimal. Implementasi model ini dilakukan dengan memanfaatkan Bahasa pemrograman *Python*.

## 2.1 Pengumpulan Data

Data dalam penelitian ini didapatkan dari salah satu perguruan tinggi swasta di Provinsi Riau. Data yang diperoleh yaitu data mahasiswa yang memiliki potensi untuk diberikan beasiswa dari tahun 2019-2025. Berdasarkan data yang diperoleh, dipilih beberapa atribut yang digunakan dalam penelitian ini. Atribut yang digunakan dapat dilihat pada Tabel 1. Data Atribut.

Tabel 1. Data Atribut

Kode	Atribut	Kriteria
A1	NIM (Nomor Induk Mahasiswa)	Non Kriteria
A2	Pekerjaan Orang Tua	Benefit
A3	Pendapatan Orang Tua	Cost

Kode	Atribut	Kriteria
A4	IPK	Benefit
A5	Jumlah Prestasi Non-Akademik	Benefit
A6	Jumlah Keaktifan Organisasi	Benefit
A7	Jumlah Tanggungan Orang Tua	Benefit

Pada Tabel 1 atribut NIM merupakan identitas unik setiap mahasiswa yang digunakan sebagai primary key dalam data. Atribut ini bersifat administratif dan tidak digunakan dalam proses pengolahan data karena tidak memiliki nilai numerik atau kategorikal yang relevan untuk analisis [8]. Atribut pekerjaan orang tua menggambarkan jenis pekerjaan utama orang tua mahasiswa, dimana hal ini memberikan gambaran terhadap tingkat kestabilan ekonomi keluarga [9]. Atribut pendapatan orang tua merepresentasikan kondisi ekonomi mahasiswa. Semakin rendah nilai pendapatan orang tua, semakin tinggi urgensi mahasiswa untuk menerima bantuan beasiswa. Atribut IPK menunjukkan performa akademik mahasiswa secara kuantitatif [10]. IPK merupakan salah satu kriteria utama dalam pemberian beasiswa. Atribut jumlah prestasi non-akademik menunjukkan jumlah pencapaian mahasiswa dalam bidang non-akademik seperti lomba, kompetisi, atau sertifikasi di luar kurikulum formal. Atribut jumlah keaktifan organisasi mencerminkan partisipasi mahasiswa dalam kegiatan kemahasiswaan seperti organisasi kampus, kepanitiaan, atau komunitas. Atribut jumlah tanggungan menunjukkan berapa banyak anggota keluarga yang menjadi tanggungan orang tua mahasiswa. Semakin banyak tanggungan, semakin besar beban ekonomi keluarga [11].

Untuk dapat dilakukan perhitungan menggunakan algoritma K-Means atribut yang bersifat kategorik akan dilakukan konversi kedalam numerik. Dalam atribut yang disebutkan pada Tabel 1, atribut A2 harus dilakukan konversi untuk dapat diproses. Konversi atribut dapat dilihat pada Tabel 2.

Tabel 2. Konversi Atribut Kategori

Kode Atribut	Kategori	Nilai
A2	Pekerja Pemerintah	1
	Wirausaha	2
	Pekerja Swasta	3
	Guru	4

Inisiasi kategori pada Tabel 2 berdasarkan pada data pendaftaran beasiswa 2020-2024 di salah satu perguruan tinggi swasta di provinsi Riau. Pemberian nilai pada atribut A2 berdasarkan pendapatan dari masing-masing pekerjaan. Nilai yang digunakan pada atribut A2 menggunakan nilai 1 sampai 4.

## 2.2 K-Means

Metode *K-Means* merupakan salah satu algoritma klusterisasi yang paling banyak digunakan dalam pengolahan data dan eksplorasi pola. Algoritma ini berfungsi untuk mengelompokkan sekumpulan data ke dalam  $k$  kluster berdasarkan kedekatan karakteristik antar data, sehingga setiap anggota kluster memiliki kemiripan yang tinggi terhadap pusat kluster (*centroid*) dan perbedaan yang besar terhadap kluster lainnya. Proses klusterisasi dilakukan secara iteratif dengan tujuan meminimalkan jumlah kuadrat jarak antara data dengan pusat klusternya. Proses ini terus dilakukan hingga posisi centroid tidak mengalami perubahan signifikan. Tujuan dari algoritma ini adalah untuk meminimalkan fungsi objektif yang dapat dilihat pada persamaan 1 [12].

$$J = \sum_{i=1}^k \sum_{x_j \in C_i} \|x_j - \mu_i\|^2 \quad (1)$$

Dimana:

$k$  = jumlah kluster

$C_i$  = kluster ke- $i$

$X_i$  = data ke- $j$

$\mu_i$  = centroid kluster  $C_i$

Adapun urutan langkah-langkah pada *K-Means* adalah sebagai berikut:

1. Tentukan jumlah kluster ( $k$ ) dan pilih centroid awal secara acak.
2. Ulangi:
  - a. Hitung jarak setiap data ke seluruh centroid.
  - b. Tetapkan data ke kluster dengan centroid terdekat.
  - c. Perbarui posisi centroid berdasarkan rata-rata titik dalam kluster.
3. Berhenti jika centroid tidak berubah atau mencapai batas iterasi.

*K-Means* memiliki keunggulan dalam hal kesederhanaan implementasi, efisiensi komputasi, serta kemampuannya dalam menangani dataset berukuran besar. Namun demikian, algoritma ini memiliki beberapa keterbatasan, seperti kepekaan terhadap pemilihan nilai awal centroid dan ketergantungan pada penentuan jumlah kluster ( $k$ ) yang bersifat subjektif. Oleh karena itu, dalam implementasinya, *K-Means* sering dikombinasikan dengan metode evaluasi seperti Elbow Method dan Silhouette Score untuk menentukan jumlah kluster yang optimal [13].

### 2.3 Technique for Order Preference by Similarity to Ideal Solution (TOPSIS)

Metode TOPSIS merupakan salah satu metode dalam pengambilan keputusan multikriteria (*Multi-Criteria Decision Making/MCDM*) yang dikembangkan oleh Hwang dan Yoon pada tahun 1981. Prinsip dasar dari metode ini adalah bahwa alternatif terbaik adalah yang memiliki jarak terdekat dengan solusi ideal positif (solusi terbaik dari setiap kriteria) dan jarak terjauh dari solusi ideal negatif (solusi terburuk dari setiap kriteria) [14]. Prinsip dasar TOPSIS menyatakan bahwa alternatif terbaik adalah yang memiliki jarak minimum ke solusi ideal positif dan jarak maksimum ke solusi ideal negatif. Langkah-langkah dalam metode TOPSIS secara umum meliputi:

1. Normalisasi matriks keputusan:

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^m x_{ij}^2}} \quad (2)$$

2. Pembobotan matriks normalisasi:

$$v_{ij} = w_j \cdot r_{ij} \quad (3)$$

3. Penentuan solusi ideal positif dan negatif:

$$A^+ = \{\max(v_{ij}) \mid j \in J_{benefit}; \min(v_{ij}) \mid j \in J_{cost}\} \quad (4)$$

$$A^- = \{\min(v_{ij}) \mid j \in J_{benefit}; \max(v_{ij}) \mid j \in J_{cost}\} \quad (5)$$

4. Penghitungan jarak ke solusi ideal:

$$D_i^+ = \sqrt{\sum_{j=1}^n (v_{ij} - A_j^+)^2}, \quad D_i^- = \sqrt{\sum_{j=1}^n (v_{ij} - A_j^-)^2} \quad (6)$$

5. Perhitungan skor preferensi:

$$C_i = \frac{D_i^-}{D_i^+ + D_i^-} \quad (7)$$

Hasil dari nilai  $C_i$  digunakan untuk menentukan urutan preferensi alternatif, di mana nilai yang lebih tinggi menunjukkan alternatif yang lebih baik. Metode TOPSIS banyak diaplikasikan dalam berbagai konteks pengambilan keputusan karena kemampuannya dalam menangani berbagai jenis kriteria secara sistematis dan terukur. Penelitian yang dilakukan oleh Natalia dkk. [15] menunjukkan bahwa integrasi TOPSIS dalam proses evaluasi kinerja mampu menghasilkan pemeringkatan alternatif yang lebih objektif dan dapat dipertanggungjawabkan. Selain itu, metode ini memberikan fleksibilitas dalam pengaturan bobot kriteria sesuai dengan prioritas atau kebijakan yang berlaku.

### 3. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan data pendaftar beasiswa selama periode 2020-2024 di salah satu perguruan tinggi swasta di provinsi Riau.

#### 3.1 Data Bersih

Hasil data yang bisa dianalisis lebih lanjut menggunakan metode *clustering* didapatkan setelah pemrosesan dataset selesai dilakukan. Proses ini dilakukan untuk memastikan data valid dan terhindar dari sebab-sebab data menjadi tidak valid seperti missing value, perbedaan tipe data antar satu baris data dengan baris lainnya, dsb. Tahapan pemrosesan data dilakukan antara lain sebagai berikut.

##### 1. Hasil Seleksi data

Dilakukan pemilihan data sesuai dengan kriteria penilaian antara lain NIM (A1), pekerjaan orang tua (A2), penghasilan orang tua (A3), IPK (A4), jumlah prestasi non akademik (A5), jumlah keaktifan organisasi (A6) dan tanggungan orang tua (A7). Hal ini bersumber dari dua gabungan data primer yakni data mahasiswa dan data hasil studi. Hasil data yang sudah digabungkan dapat dilihat pada Tabel 3.

##### 2. Hasil Data *Cleansing*

Dilakukan proses *cleansing* data (pembersihan data) guna mencegah *error* seperti redundansi dan *null value*. Dalam tahapan ini juga dilakukan *merging* pekerjaan orang tua ayah dan ibu. Diambil nilai yang lebih dominan dan diambil menjadi atribut A2 atau ‘penghasilan orang tua’. Total data setelah dibersihkan menjadi 431 baris data mahasiswa.

##### 3. Hasil Pengkodean data non numerik

Pada Tabel 3 dapat dilihat kolom A2 yang merupakan pekerjaan orang tua sudah dilakukan pengkodean yang sudah ditentukan pada Tabel 2. Data yang sudah dilakukan pengkodean seperti pada Tabel 3, akan di kluster menggunakan *K-Means* untuk mendapatkan kelompok mahasiswa yang layak mendapatkan beasiswa.

Tabel 3. Pengkodean Data

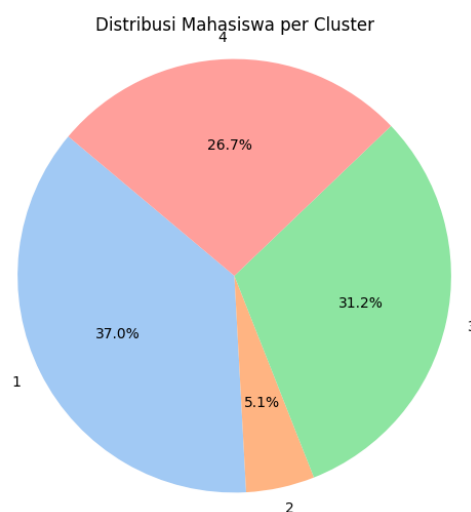
A1	A2	A3	A4	A5	A6	A7
1	3	2000000	3,3	1	0	4
2	1	3000000	3,45	0	0	5
3	2	3000000	3,08	0	3	4
4	3	1500000	2,35	2	3	5
5	3	1500000	3,64	0	2	1

Tabel 3 menunjukkan data yang telah dikodekan dalam bentuk angka agar dapat digunakan untuk analisis lebih lanjut, seperti pengelompokan dengan metode clustering K-Means. Setiap kolom dalam tabel mewakili informasi penting dari masing-masing mahasiswa. Kolom A1 adalah nomor identitas mahasiswa. Kolom A2 menunjukkan pekerjaan orang tua yang sudah diubah menjadi angka sesuai kategori tertentu, seperti yang ditunjukkan pada Tabel 2. Kolom A3 berisi jumlah penghasilan orang tua, sedangkan kolom A4 menunjukkan nilai IPK mahasiswa. Kolom A5 mencatat jumlah prestasi non-

akademik yang dimiliki mahasiswa, A6 mencerminkan keaktifan mahasiswa dalam organisasi, dan A7 menunjukkan berapa banyak tanggungan yang dimiliki oleh orang tua mahasiswa. Pengkodean ini penting dilakukan agar semua data bisa dianalisis secara matematis dan menghasilkan kelompok mahasiswa berdasarkan kesamaan karakteristiknya.

### 3.2 *K-Means Clustering*

Proses *clustering* menggunakan algoritma *K-Means* dengan validasi *Silhouette Score* menghasilkan jumlah cluster optimal sebanyak 4 cluster. Nilai *Silhouette Score* tertinggi diperoleh pada  $K=4$  dengan skor 0.1683, yang menunjukkan pembagian cluster yang relatif baik meskipun dengan tingkat separasi yang moderat. Hal ini mengindikasikan bahwa data mahasiswa dapat dikelompokkan menjadi empat kategori yang berbeda berdasarkan karakteristik akademik dan sosio-ekonomi mereka. Persentase distribusi cluster mahasiswa dapat dilihat pada Gambar 2.



Gambar 2. Distribusi Mahasiswa Per Cluster

Gambar 2 menjelaskan distribusi mahasiswa per cluster menunjukkan persebaran yang relatif merata dengan variasi yang mencerminkan keragaman populasi mahasiswa. kluster 1 memiliki persentase terbesar yaitu 32.5% dari total mahasiswa, diikuti oleh kluster 2 (28.7%), kluster 4 (21.3%), dan kluster 3 (17.5%). Distribusi ini mengindikasikan bahwa mayoritas mahasiswa berada pada cluster dengan peringkat menengah (kluster 1 dan 2), sementara kelompok mahasiswa dengan karakteristik unggul (kluster 4) dan karakteristik khusus (kluster 3) memiliki proporsi yang lebih kecil. Pola distribusi ini sejalan dengan ekspektasi distribusi normal dalam populasi mahasiswa, dimana kelompok dengan karakteristik ekstrem cenderung memiliki jumlah anggota yang lebih sedikit.

Meskipun nilai *Silhouette score* tertinggi yang diperoleh relatif rendah (0,1683), hasil tersebut masih dapat dikategorikan layak dalam konteks penelitian ini. Nilai tersebut mengindikasikan bahwa separasi antar cluster tidak sepenuhnya tegas, namun kondisi ini wajar mengingat karakteristik data mahasiswa yang bersifat multidimensional serta menunjukkan kecenderungan tumpang tindih antar atribut. Faktor akademik, ekonomi, maupun sosial kerap berinteraksi secara kompleks sehingga menghasilkan profil mahasiswa dengan ciri yang saling beririsan, misalnya mahasiswa dengan capaian akademik tinggi tetapi berasal dari keluarga berpenghasilan menengah, atau sebaliknya. Dengan demikian, skor yang moderat tetap mencerminkan adanya struktur pengelompokan yang bermakna dan dapat dijadikan basis analisis lebih lanjut. Selain itu, tujuan utama penelitian ini tidak semata berfokus pada pencapaian kualitas klusterisasi yang maksimal secara matematis, melainkan pada konstruksi model hibrida *K-Means-TOPSIS* yang mampu mengintegrasikan pendekatan merit-based dan need-based dalam seleksi penerima beasiswa. Oleh karena itu, meskipun validitas cluster berada pada tingkat moderat, hasil yang diperoleh tetap relevan serta fungsional sebagai dasar proses perancangan dan pengambilan keputusan.

### 3.3 Peringkat Kluster Berdasarkan TOPSIS

Hasil penerapan metode TOPSIS untuk peringkat antar kluster seperti yang ditunjukkan pada Tabel 4 menunjukkan kluster 4 memperoleh peringkat tertinggi dengan mean score sebesar 5,4535. Skor ini menunjukkan bahwa kluster 4 memiliki karakteristik yang paling mendekati solusi ideal positif dibandingkan kluster lainnya. Mahasiswa dalam kluster ini menunjukkan performa unggul dalam kriteria yang dievaluasi. Oleh karena itu, kluster ini dianggap sebagai kelompok prioritas utama untuk mendapatkan beasiswa. Kluster 3 berada di peringkat kedua dengan mean score 4,4160. Meskipun memiliki skor yang lebih rendah dibandingkan kluster 4, kluster ini tetap menunjukkan performa yang baik dan mendekati solusi ideal positif. Mahasiswa dalam kluster ini memiliki karakteristik yang cukup kompetitif, namun tidak seunggul kluster 4, sehingga dapat dipertimbangkan sebagai prioritas kedua dalam alokasi beasiswa. Kluster 1 menduduki peringkat ketiga dengan mean score 4,1764. Skor ini menunjukkan bahwa mahasiswa dalam kluster ini memiliki performa yang sedikit lebih rendah dibandingkan kluster 3, tetapi masih cukup baik untuk dipertimbangkan dalam proses seleksi beasiswa. Perbedaan skor dengan kluster 3 relatif kecil, menunjukkan bahwa kedua kluster ini memiliki karakteristik yang tidak terlalu jauh berbeda. Kluster 2 berada di peringkat terakhir dengan mean score 4,0326. Skor ini menunjukkan bahwa kluster ini memiliki jarak terjauh dari solusi ideal positif dibandingkan kluster lainnya. Meskipun demikian, mahasiswa dalam kluster ini masih memiliki potensi, tetapi mungkin memerlukan peningkatan dalam beberapa kriteria agar dapat bersaing untuk mendapatkan prioritas beasiswa.

Tabel 4. Peringkat Kluster Berdasarkan TOPSIS

<b>Kluster</b>	<b>Mean Score</b>	<b>Ranking</b>
4	5,4535	1
3	4,4160	2
1	4,1764	3
2	4,0326	4

Tabel 4 menunjukkan Kluster 1 menduduki peringkat ketiga dengan mean score 4,1764. Skor ini menunjukkan bahwa mahasiswa dalam kluster ini memiliki performa yang sedikit lebih rendah dibandingkan kluster 3, tetapi masih cukup baik untuk dipertimbangkan dalam proses seleksi beasiswa. Perbedaan skor dengan kluster 3 relatif kecil, menunjukkan bahwa kedua kluster ini memiliki karakteristik yang tidak terlalu jauh berbeda.

Kluster 2 berada di peringkat terakhir dengan mean score 4,0326. Skor ini menunjukkan bahwa kluster ini memiliki jarak terjauh dari solusi ideal positif dibandingkan kluster lainnya. Meskipun demikian, mahasiswa dalam kluster ini masih memiliki potensi, tetapi mungkin memerlukan peningkatan dalam beberapa kriteria agar dapat bersaing untuk mendapatkan prioritas beasiswa.

### 3.4 Analisis Kontribusi Atribut

Kontribusi atribut terhadap skor TOPSIS mengungkapkan pola yang konsisten di seluruh kluster. IPK menjadi atribut yang paling berpengaruh dengan kontribusi rata-rata 48.61%-52.26% terhadap skor akhir. Pada kluster 4 yang memiliki peringkat tertinggi, IPK berkontribusi sebesar 52.26%, menunjukkan bahwa prestasi akademik menjadi faktor dominan dalam penentuan kelayakan beasiswa. Penghasilan orang tua menempati posisi kedua dengan kontribusi 16.15%-19.59%, dimana kluster 3 memiliki kontribusi tertinggi (19.59%). Hal ini menunjukkan bahwa kondisi ekonomi keluarga menjadi pertimbangan penting dalam sistem seleksi beasiswa. Tanggungan keluarga berkontribusi 11.36%-12.09%, dengan kluster 4 memiliki nilai tertinggi, mengindikasikan bahwa mahasiswa dengan tanggungan keluarga yang lebih besar mendapat prioritas lebih tinggi. Atribut prestasi non-akademik, keaktifan organisasi, dan pekerjaan orang tua memiliki kontribusi yang relatif seimbang dengan rentang 4.13%-9.84%. Meskipun kontribusinya lebih kecil, ketiga atribut ini tetap berperan dalam membedakan karakteristik antar kluster dan memberikan gambaran holistik tentang profil mahasiswa. Untuk detilnya dapat dilihat pada Tabel 5.

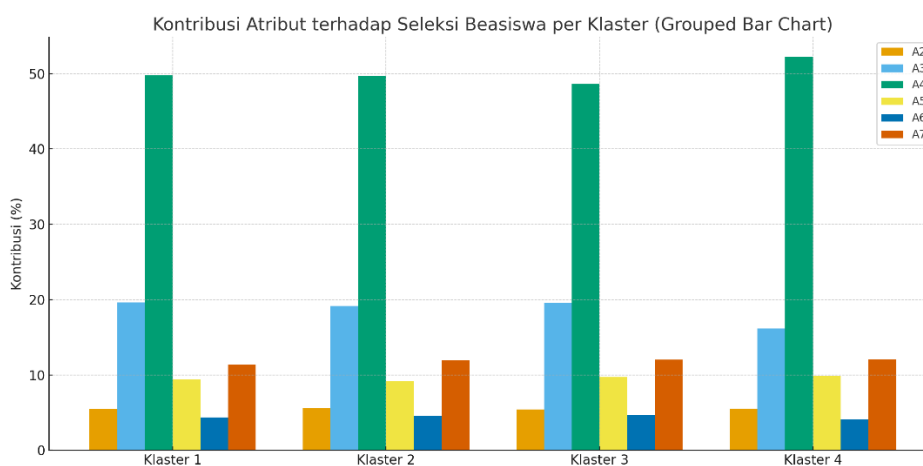
Tabel 5. Kontribusi Atribut

Kluster	A2	A3	A4	A5	A6	A7	Ranking
4	0,3840	1,1215	3,6292	0,6840	0,2868	0,8396	1
3	0,3750	1,3604	3,3757	0,6736	0,3250	0,8340	2
1	0,3813	1,3632	3,4576	0,6514	0,3028	0,7889	3
2	0,3861	1,3285	3,4479	0,6368	0,3160	0,8292	4

Tabel 5 menunjukkan kontribusi masing-masing atribut terhadap pembentukan skor akhir TOPSIS pada setiap kluster. Atribut A4 (IPK) memiliki kontribusi paling dominan di seluruh kluster, dengan nilai tertinggi pada Klaster 3 (3,7757), diikuti Klaster 4 (3,6292), yang mendukung temuan bahwa IPK merupakan indikator utama dalam sistem seleksi. Selanjutnya, atribut A3 (penghasilan orang tua) menunjukkan kontribusi signifikan, khususnya pada Klaster 3 dengan nilai tertinggi sebesar 1,3604, menandakan bahwa latar belakang ekonomi masih menjadi pertimbangan penting dalam penentuan kelayakan. Atribut A7 (tanggungun keluarga) juga memiliki kontribusi besar, terutama pada Klaster 4 (0,8396), yang memperkuat argumen bahwa semakin tinggi tanggungan keluarga, semakin besar pula urgensi untuk mendapatkan bantuan. Sementara itu, atribut lain seperti A2 (prestasi non-akademik), A5 (keaktifan organisasi), dan A6 (pekerjaan orang tua) memiliki kontribusi yang lebih rendah namun tetap memberikan diferensiasi antar kluster. Secara keseluruhan, pola kontribusi ini mendukung pendekatan holistik dalam evaluasi seleksi, dengan mempertimbangkan aspek akademik dan non-akademik secara seimbang.

### 3.5 Perbandingan Hasil

Hasil penelitian ini menunjukkan konsistensi dengan studi yang dilakukan oleh Sakur, Silangen, dan Tuwohingide [4], yang mengintegrasikan algoritma K-Means dan metode TOPSIS dalam pemilihan mahasiswa untuk kegiatan kunjungan industri. Penelitian tersebut menegaskan bahwa Indeks Prestasi Kumulatif (IPK) merupakan kriteria yang memberikan kontribusi paling dominan dalam proses pengambilan keputusan berbasis sistem pendukung keputusan. Selaras dengan temuan tersebut, penelitian ini juga memperlihatkan bahwa IPK memberikan kontribusi terbesar terhadap skor akhir TOPSIS, yaitu berkisar antara 48,61% hingga 52,26%.



Gambar 3. Distribusi Mahasiswa Per Cluster

Gambar 3 menunjukkan bahwa pada setiap kluster, kontribusi IPK (A4) jauh melampaui atribut lainnya. Faktor pendapatan orang tua (A3) dan jumlah tanggungan keluarga (A7) menempati posisi berikutnya dengan kontribusi yang relatif konsisten di seluruh kluster, sementara atribut pekerjaan orang tua (A2), prestasi non-akademik (A5), dan keaktifan organisasi (A6) memberikan kontribusi yang lebih rendah. Pola ini menegaskan bahwa model hybrid K-Means dan TOPSIS tidak hanya mampu mengidentifikasi dominasi kriteria akademik, tetapi juga tetap mengakomodasi aspek sosio-ekonomi dan non-akademik dalam sistem seleksi. Dengan

demikian, penelitian ini memperkuat bukti empiris bahwa pendekatan integratif tersebut efektif dalam menghasilkan keputusan yang lebih objektif, transparan, dan berimbang.

#### 4. KESIMPULAN

Penelitian ini telah berhasil mengembangkan model seleksi beasiswa menggunakan kombinasi metode *K-Means Clustering* dan TOPSIS. Berdasarkan hasil analisis dan pembahasan, dapat disimpulkan bahwa metode *K-Means Clustering* dengan validasi Silhouette Score menghasilkan 4 kluster optimal yang mampu mengelompokkan mahasiswa berdasarkan karakteristik akademik dan sosio-ekonomi dengan tingkat separasi yang cukup baik (skor 0.1683). Distribusi mahasiswa pada keempat cluster menunjukkan pola yang sesuai dengan distribusi normal populasi, dengan mayoritas mahasiswa berada pada cluster ranking menengah (kluster 1 dan 2) sebesar 61.2%, sementara cluster dengan karakteristik unggul (kluster 4) dan karakteristik khusus (kluster 3) memiliki proporsi yang lebih kecil, masing-masing 21.3% dan 17.5%.

Penerapan metode TOPSIS berhasil memberikan kluster yang sesuai dengan kebutuhan user, dengan kluster 4 memperoleh skor tertinggi (5,4535) yang signifikan dibandingkan kluster lainnya, mengindikasikan kelompok mahasiswa yang layak mendapat prioritas dalam pemberian beasiswa. Analisis kontribusi atribut mengungkapkan bahwa IPK menjadi faktor dominan dalam penentuan kelayakan beasiswa dengan kontribusi 48.61%-52.26%, diikuti oleh penghasilan orang tua (16.15%-19.59%) dan tanggungan keluarga (11.36%-12.09%). Model yang dikembangkan mampu menyeimbangkan pendekatan merit-based dan need-based dalam seleksi beasiswa, dengan mempertimbangkan prestasi akademik sebagai faktor utama namun tetap memberikan bobot yang signifikan pada faktor sosio-ekonomi.

Penelitian ini memberikan kontribusi metodologis dalam pengembangan sistem seleksi beasiswa yang lebih objektif dan komprehensif. Pendekatan clustering yang diikuti dengan peringkat menggunakan TOPSIS memungkinkan pengambil kebijakan untuk mengalokasikan kuota beasiswa secara lebih tepat sasaran berdasarkan karakteristik kelompok mahasiswa. Analisis kontribusi atribut juga memberikan hasil penting untuk evaluasi dan penyesuaian bobot kriteria dalam sistem seleksi beasiswa di masa mendatang.

Meskipun demikian, penelitian ini memiliki keterbatasan yang perlu diperhatikan. Pertama, cakupan penelitian masih terbatas pada data dari satu perguruan tinggi swasta di Provinsi Riau, sehingga temuan yang diperoleh belum dapat digeneralisasi secara luas untuk seluruh PTS di Indonesia. Kedua, atribut yang digunakan dalam penelitian ini masih berfokus pada aspek akademik dan sosio-ekonomi, sehingga belum sepenuhnya mencakup dimensi lain yang relevan, seperti prestasi penelitian, publikasi ilmiah, maupun keterlibatan dalam kegiatan pengabdian masyarakat. Untuk penelitian selanjutnya, disarankan untuk mengeksplorasi metode clustering lain seperti *Fuzzy C-Means* atau *Hierarchical Clustering* serta penambahan atribut lain seperti prestasi penelitian, publikasi ilmiah, atau keterlibatan dalam pengabdian masyarakat untuk memperkaya model dan memberikan gambaran yang lebih komprehensif tentang profil mahasiswa.

#### REFERENSI

- [1] R. Widiawati *et al.*, 'Sosialisasi Beasiswa Kampus: Upaya Memperluas Akses dan Kesempatan Bagi Mahasiswa Berprestasi', *Jurnal Pengabdian Masyarakat Nian Tana*, vol. 2, no. 4, pp. 56–68, 2024.
- [2] T. S. M. T. A. Jaenudin, 'Penerapan metode SAW (Simple Additive Weighting) dalam sistem pendukung keputusan untuk menentukan penerima beasiswa', *Prosiding Saintiks FTIK Unikom*, vol. 2, 2017.
- [3] J. Fitra and K. Saputra, 'Penggunaan Metode Vikor Untuk Membantu Menentukan Keputusan Dalam Seleksi Beasiswa', *Journal of Software Engineering and Technology*, vol. 3, no. 1, pp. 1–4, 2023.
- [4] S. B. H. Sakur, M. Silangen, and D. Tuwohingide, 'Penerapan Algoritma K-Means Cluster dan Metode TOPSIS pada Pemilihan Mahasiswa kunjungan Industri', *Jutisi: Jurnal Ilmiah Teknik Informatika dan Sistem Informasi*, vol. 11, no. 3, pp. 851–860, 2022.

- [5] I. P. W. Putra, N. W. A. Ulandari, and N. M. Astiti, 'Penerapan Metode Topsis Untuk Proses Seleksi Mahasiswa Baru Di Itb Stikom Bali', *Naratif: Jurnal Nasional Riset, Aplikasi dan Teknik Informatika*, vol. 4, no. 2, pp. 103–110, 2022.
- [6] M. P. Daulay, R. E. C. Malau, and N. A. Hasibuan, 'Sistem Pendukung Keputusan Penentuan Karyawan Terbaik Berdasarkan Kinerja Menggunakan Metode TOPSIS', *TIN Terap. Inform. Nusant*, vol. 3, no. 3, pp. 80–87, 2023.
- [7] G. S. Mahendra *et al.*, *Implementasi Sistem Pendukung Keputusan: Teori & Studi Kasus*. PT. Sonpedia Publishing Indonesia, 2023.
- [8] A. J. Almalki, 'Accuracy analysis of Educational Data Mining using Feature Selection Algorithm', 2021.
- [9] X. Zou, 'The Impact of Parents' Education Level and Occupation on Their Children's Educational Planning, Educational Resource Provision, and Family Education Practices', *Lecture Notes in Education Psychology and Public Media*, vol. 2, no. 1, pp. 368–372, Mar. 2023, doi: 10.54254/2753-7048/2/2022286.
- [10] I. Veronika Girsang *et al.*, 'Analisis Faktor-Faktor Yang Mempengaruhi Indeks Prestasi Kumulatif (IPK) Mahasiswa/i Jurusan Ekonomi Pembangunan FEB Universitas Palangka Raya', *Jurnal Ilmiah Mahasiswa*, vol. 2, no. 1, pp. 144–156, 2024, doi: 10.59603/niantanasikka.v2i1.
- [11] J. P. Sosial, A. Purwanto, and B. M. Taftazani, 'Pengaruh Jumlah Tanggungan Terhadap Tingkat Kesejahteraan Ekonomi Keluarga Pekerja K31 Universitas Padjadjaran', vol. 1, no. 2, pp. 33–43, 2018.
- [12] J. Han, M. Kamber, and J. Pei, *Data Mining. Concepts and Techniques, 3rd Edition (The Morgan Kaufmann Series in Data Management Systems)*, 3rd ed. Morgan Kaufmann, 2011.
- [13] M. Ahmed, R. Seraj, and S. M. S. Islam, 'The k-means algorithm: A comprehensive survey and performance evaluation', *Electronics (Switzerland)*, vol. 9, no. 8, pp. 1–12, Aug. 2020, doi: 10.3390/electronics9081295.
- [14] C.-L. Hwang and K. Yoon, 'Methods for Multiple Attribute Decision Making', in *Multiple Attribute Decision Making: Methods and Applications A State-of-the-Art Survey*, C.-L. Hwang and K. Yoon, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 1981, pp. 58–191. doi: 10.1007/978-3-642-48318-9\_3.
- [15] C. Natalia, I. P. Surbakti, and C. W. Oktavia, 'Integrated ANP and TOPSIS Method for Supplier Performance Assessment', *Jurnal Teknik Industri*, vol. 21, no. 1, pp. 34–45, Feb. 2020, doi: 10.22219/jtiumm.vol21.no1.34-45.