



# ANALISIS FAKTOR DOMINAN KEBERHASILAN AKADEMIK MAHASISWA PENERIMA BEASISWA MENGGUNAKAN PRINCIPAL COMPONENT ANALYSIS

**Nur Mufidah<sup>\*1</sup>, Dadang Syarif Sihabudin Sahid<sup>2</sup>, Satria Perdana Arifin<sup>3</sup>, Wahyu Ari Sandi<sup>4</sup>**

<sup>1,2,3,4</sup>Teknologi Informasi, Politeknik Caltex Riau, 28265, Indonesia

nurmufidah30@gmail.com, dadang@pcr.ac.id, satria@pcr.ac.id, batosayary@gmail.com

<sup>\*</sup>Penulis Koresponden

## ABSTRAK

Keberhasilan akademik mahasiswa penerima beasiswa dipengaruhi oleh berbagai faktor non-akademik yang saling berkaitan dan sering kali menimbulkan redundansi data sehingga sulit diidentifikasi secara akurat. Kondisi ini menimbulkan kebutuhan akan metode analisis yang mampu menyederhanakan variabel tanpa menghilangkan informasi penting. Penelitian ini bertujuan mengidentifikasi faktor dominan yang memengaruhi keberhasilan akademik mahasiswa penerima beasiswa menggunakan Principal Component Analysis (PCA) sebagai solusi untuk mengatasi korelasi tinggi antarvariabel dan multikolinearitas dalam dataset. Sebanyak 262 mahasiswa mengisi kuesioner berisi 54 item non-akademik yang telah divalidasi melalui uji validitas dan reliabilitas. Hasil PCA menunjukkan bahwa 38 komponen utama mampu menjelaskan 95,08% variansi total, menandakan efektivitas reduksi dimensi dalam merangkum informasi penting. Analisis loading matrix mengungkapkan lima faktor dominan, yaitu kondisi psikologis, metode pembelajaran, kesesuaian jurusan, motivasi belajar, dan kondisi keuangan. Hasil ini memberikan pemahaman yang lebih terstruktur mengenai faktor-faktor kunci yang memengaruhi performa akademik, sekaligus menjadi dasar pengambilan keputusan dalam pembinaan dan monitoring penerima beasiswa.

**Kata kunci:** *Analisis Faktor, Beasiswa, Faktor Keberhasilan Akademik, Principal Component Analysis*

## ABSTRACT

*The academic success of scholarship recipients is influenced by various interrelated non-academic factors that often lead to data redundancy, making accurate identification of dominant factors challenging. This condition creates the need for an analytical method capable of simplifying variables without losing essential information. This study aims to identify the dominant factors affecting the academic success of scholarship recipients using Principal Component Analysis (PCA) as a solution to address high inter-variable correlations and multicollinearity within the dataset. A total of 262 students completed a questionnaire consisting of 54 non-academic items that had previously undergone validity and reliability testing. The PCA results show that 38 principal components explain 95.08% of the total variance, indicating the effectiveness of dimensionality reduction in summarizing essential information. The loading matrix analysis reveals five dominant factors—psychological conditions, learning methods, major suitability, learning motivation, and financial conditions. These findings provide a more structured understanding of key factors influencing academic performance and serve as a foundation for decision-making in the development and monitoring of scholarship programs.*

**Keywords:** *Academic Success Factors, Factor Analysis, Principal Component Analysis, Scholarship*

---

### Histori Artikel

Diserahkan: 18 Nov 2025

Diterima setelah Revisi: 2 Des 2025

Diterbitkan: 11 Des 2025

---

### 1. PENDAHULUAN

Keberhasilan akademik mahasiswa penerima beasiswa merupakan salah satu aspek penting dalam mengevaluasi efektivitas program bantuan pendidikan yang dirancang oleh institusi pendidikan tinggi. Beasiswa tidak hanya berfungsi sebagai dukungan finansial, tetapi juga sebagai bentuk investasi jangka panjang untuk meningkatkan kualitas sumber daya manusia [1]. Namun, upaya mempertahankan

performa akademik yang konsisten sering kali menghadapi berbagai kendala, termasuk faktor-faktor non-akademik yang memengaruhi proses belajar mahasiswa [2].

Berbagai penelitian sebelumnya menunjukkan bahwa keberhasilan akademik tidak hanya ditentukan oleh kemampuan kognitif, tetapi juga oleh variabel psikologis, sosial, dan finansial yang membentuk cara mahasiswa merespons lingkungan belajarnya [3], [4]. Motivasi belajar, misalnya, berperan penting dalam membangun ketekunan, effort, serta persistensi dalam menghadapi tantangan akademik [5]. Selain itu, kondisi psikologis seperti stres, kecemasan, dan tekanan akademik dapat memberikan dampak langsung terhadap kemampuan mahasiswa dalam berkonsentrasi, menyerap informasi, dan mengambil keputusan akademik [6], [7].

Di sisi lain, kualitas metode pembelajaran yang diberikan dosen juga merupakan faktor signifikan. Penggunaan strategi pembelajaran yang tidak sesuai dengan gaya belajar mahasiswa dapat menurunkan efektivitas proses pembelajaran secara keseluruhan [8]. Kesesuaian jurusan juga menjadi aspek penting, karena mahasiswa yang menempuh program studi sesuai minat dan kemampuan cenderung memiliki tingkat kepuasan belajar yang lebih tinggi serta performa akademik yang lebih stabil [9].

Selain faktor-faktor tersebut, kondisi finansial mahasiswa tetap memegang peranan besar, meskipun mereka telah menerima beasiswa. Beberapa penelitian menemukan bahwa tekanan ekonomi keluarga dapat memengaruhi fokus belajar mahasiswa dan menyebabkan distraksi yang berdampak pada pencapaian akademik [10]. Situasi ini memperlihatkan bahwa keberhasilan akademik merupakan fenomena multidimensional yang dipengaruhi oleh beragam variabel non-akademik yang saling berkaitan.

Permasalahan utama yang sering muncul dalam menganalisis faktor-faktor tersebut adalah banyaknya indikator yang saling berkorelasi, sehingga sulit untuk menentukan variabel mana yang paling dominan memengaruhi performa akademik. Korelasi tinggi antarvariabel dapat menyebabkan multikolinearitas yang berakibat pada bias interpretasi ketika menggunakan metode analisis tradisional seperti regresi linier [11]. Karena itu, metode statistik yang mampu menyederhanakan variabel kompleks dengan tetap mempertahankan informasi penting sangat diperlukan.

*Principal Component Analysis* (PCA) merupakan salah satu teknik analisis multivariat yang dapat digunakan untuk mengatasi masalah tersebut[12]. PCA bekerja dengan metode multivariat yang mengurangi kompleksitas data sambil mempertahankan kovarians data, sejalan dengan pernyataan bahwa PCA memetakan variabel asli ke komponen utama yang mewakili variansi terbesar [13]. Metode ini banyak digunakan dalam penelitian pendidikan, psikologi, dan data mining untuk mengidentifikasi pola laten dari variabel yang kompleks [14], [15]. PCA juga efektif dalam mengurangi redundansi data dan memperjelas struktur faktor dominan yang sulit diidentifikasi melalui metode statistik biasa [16].

Dalam konteks mahasiswa penerima beasiswa, penggunaan PCA menjadi penting untuk mengidentifikasi faktor esensial yang paling berpengaruh terhadap keberhasilan akademik. Pengelola beasiswa perlu memahami faktor-faktor dominan tersebut untuk meningkatkan strategi pembinaan seperti pemberian mentoring, konseling, pemantauan perkembangan mahasiswa, serta modifikasi kebijakan akademik [17]. Beberapa penelitian dalam ranah pendidikan tinggi juga menekankan pentingnya pendekatan analitis berbasis data untuk perbaikan program beasiswa agar lebih efektif dan tepat sasaran [18], [19].

Penelitian ini dilakukan untuk mengidentifikasi faktor dominan penentu keberhasilan akademik mahasiswa penerima beasiswa melalui pendekatan PCA terhadap 54 variabel non-akademik yang telah divalidasi. Dengan menggunakan sampel sebanyak 262 mahasiswa, penelitian ini diharapkan memberikan kontribusi akademik berupa pemetaan faktor yang lebih terstruktur, serta membantu institusi pendidikan dalam mengembangkan program intervensi dan pendampingan mahasiswa yang lebih sesuai kebutuhan.

## 2. METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan metode analisis *Principal Component Analysis* (PCA). PCA dipilih karena mampu mereduksi sekumpulan variabel non-akademik yang saling berkorelasi menjadi beberapa komponen utama yang dapat mewakili variansi terbesar dalam dataset tanpa kehilangan informasi penting.

## 2.1 PENGUMPULAN DATA

Data diperoleh melalui kuesioner yang disusun berdasarkan sembilan faktor non-akademik, yaitu: motivasi belajar, kesesuaian jurusan, kemampuan adaptasi, manajemen waktu, metode pembelajaran, kondisi keuangan, kondisi psikologis, kondisi fisik, dan sarana belajar. Total terdapat 54 butir pernyataan menggunakan skala Likert 1–5. Instrumen divalidasi oleh psikolog untuk memastikan kesesuaian isi, setelah itu dilakukan pra-survei kepada 30 mahasiswa untuk menilai keterbacaan item. Kuesioner kemudian disebarluaskan kepada mahasiswa penerima beasiswa Politeknik Caltex Riau dengan total 262 responden yang memenuhi kriteria. Jumlah sampel memenuhi ketentuan perhitungan menggunakan rumus Slovin.

## 2.2 ANALISA UJI VALIDITAS DAN RELIABILITAS

Setelah data didapatkan pada pelaksanaan pra-survei, Uji validitas dilakukan menggunakan metode Pearson Product Moment, di mana setiap skor item dikorelasikan dengan skor total pada variabelnya. Item dinyatakan valid apabila nilai  $r$  hitung lebih besar dari  $r$  tabel pada tingkat signifikansi 0,05 [11].

$$r = \frac{n \sum xy - \sum x \sum y}{\sqrt{(n \sum x^2 - (\sum x)^2)(n \sum y^2 - (\sum y)^2)}} \quad (1)$$

Diketahui:

$r$  = Pearson Product Moment

$x$  = Skor item

$y$  = Total skor variabel

$n$  = Jumlah responden

Berikut ini merupakan table uji validitas:

Tabel 1. Kriteria nilai uji Validitas

Nilai Korelasi ( $r$ hitung)	Interpretasi
$r > r$ tabel (0,361)	Item valid
$r \leq r$ tabel (0,361)	Item tidak valid

Sumber: <https://qmc.binus.ac.id/2014/11/01/>

Dimana, Jika  $r$  hitung lebih besar dari 0,361 maka item tersebut valid, artinya pertanyaan itu sesuai dan tepat untuk mengukur hal yang ingin diteliti. Apabila  $r$  hitung sama dengan atau lebih kecil dari 0,361 maka item tersebut tidak valid, artinya pertanyaannya kurang tepat atau tidak cukup kuat hubungannya dengan keseluruhan isi kuesioner, sehingga perlu diperbaiki atau dihapus. Berikut hasil dari uji validitas berdasarkan data yang diinputkan:

Valid: 53 item, Tidak Valid: 1 item

Item Tidak Valid:

	$r$	p_value
Q49(Kondisi Fisik)	0.167133	0.006699

Gambar 1. Hasil Uji Validitas

Gambar 1 merupakan hasil dari uji validasi yang sudah dilakukan. Dari total 54 item yang diuji, 53 item dinyatakan valid karena memenuhi kriteria signifikansi, sedangkan 1 item dinyatakan tidak valid. Item Q49 memiliki nilai korelasi sebesar 0,167 dan nilai signifikansi sebesar 0,006, yang berarti tidak signifikan pada tingkat kepercayaan 95% ( $p > 0,05$ ). Hal ini menunjukkan bahwa butir tersebut tidak cukup merepresentasikan konstruk kondisi fisik secara konsisten dengan item-item lainnya dalam faktor yang sama. Maka tindakan yang akan dilakukan dengan menghapus item tersebut agar menjaga kualitas dan konsistensi data yang digunakan.

Proses selanjutnya adalah uji reliabilitas, yang berfungsi untuk mengetahui konsistensi hasil pengukuran instrumen. Pada penelitian ini, reliabilitas diuji menggunakan Cronbach's Alpha, yang mengukur sejauh mana item-item dalam satu variabel saling berkaitan secara internal.

$$\alpha = \frac{\kappa}{\kappa - 1} \left( 1 - \frac{\sum \sigma_i^2}{\sigma_t^2} \right) \quad (2)$$

Diketahui:

$\alpha$  = Nilai Reliabilitas

$\kappa$  = Jumlah item dalam konstruk

$\sigma_t^2$  = Varians dari masing-masing item

$\sigma_t^2$  = Varians total skor

Tabel 2 merupakan kriteria uji reliabilitas sebagai berikut:

**Tabel 2. Kriteria Uji Reliabilitas**

Nilai Cronbach's Alpha	Interpretasi
$\geq 0.90$	Sangat reliabel
0.80 – 0.89	Reliabel
0.70 – 0.79	Cukup reliabel
0.60 – 0.69	Kurang reliabel

Sumber: <https://qmc.binus.ac.id/2014/11/01/>

Berikut ini hasil yang didapatkan dari pengujian reliabilitas yang telah dilakukan.

```
# Ambil hanya item yang valid
valid_cols = validity_df[validity_df['Valid']].index.tolist()

# Hitung Cronbach's Alpha untuk item valid saja
def cronbach_alpha(df_items):
    item_scores = df_items.values
    item_vars = item_scores.var(axis=0, ddof=1)
    total_var = item_scores.sum(axis=1).var(ddof=1)
    n_items = item_scores.shape[1]
    return n_items / (n_items - 1) * (1 - (item_vars.sum() / total_var))

# Terapkan hanya pada item valid
alpha = cronbach_alpha(df[valid_cols])
print(f"\nCronbach's Alpha (Item Valid Saja): {alpha:.3f}")

Cronbach's Alpha (Item Valid Saja): 0.981
```

**Gambar 2. Hasil Uji Reliabilitas**

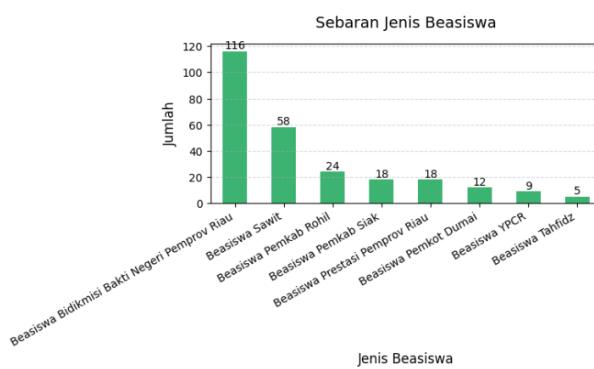
Pada hasil uji reliabilitas pada gambar 2 dapat disimpulkan bahwa hasil tergolong sangat reliabel karena hasilnya  $>0,90$ . Sehingga dapat dipercaya untuk mengukur variabel-variabel yang diteliti secara konsisten.

### 2.3 PENYEBARAN KUESIONER KE RESPONDEN UTAMA

Setelah melalui tahap validasi dan reliabilitas, kuesioner akan disebar kepada responden utama melalui google form. Untuk mendapatkan data sampel maka akan dilakukan perhitungan dengan menggunakan rumus slovin sebagai berikut:

$$n = \frac{N}{1 + N \cdot e^2} = \frac{586}{1 + 586 \cdot 0.05^2} = \frac{586}{2,465} = 237.79 \quad (3)$$

Berdasarkan rumus slovin, dapat disimpulkan bahwa data sampel yang akan digunakan yaitu minimal 238 mahasiswa. Pengumpulan Data dimulai dari penyebaran kuesioner kepada mahasiswa penerima beasiswa di Politeknik Caltex Riau. Pengisian kuesioner ini didapatkan sebanyak 262 mahasiswa pengisi kuesioner. Durasi Pengumpulan kuesioner ini dilakukan selama 23 hari, dimulai dari tanggal 17 Mei - 9 Juni 2025. Seluruh data telah diverifikasi kelengkapannya dan memenuhi kriteria untuk dianalisis untuk analisis selanjutnya. Sebaran responden dapat dilihat hasil grafik pada gambar 3:



**Gambar 3. Sebaran berdasarkan jenis beasiswa**

## 2.4 INSTRUMEN PENELITIAN

Berikut ini merupakan faktor yang memengaruhi keberhasilan akademik menurut beberapa jurnal [1], [2] dan [6]:

- i) Presensi Kehadiran  
Kehadiran yang teratur dalam kelas dapat meningkatkan interaksi dengan dosen dan mahasiswa lain, serta memperluas kesempatan untuk memahami materi dengan lebih baik.
- ii) Motivasi belajar  
Dalam konteks pendidikan tinggi, motivasi belajar sangat penting untuk membantu mahasiswa mencapai kesuksesan akademik dan mengatasi tantangan yang dihadapi. Pada Teori Hierarki kebutuhan Maslow, Abraham Maslow mengemukakan bahwa manusia memiliki lima Tingkat kebutuhan, mulai dari kebutuhan fisiologis hingga aktualisasi diri. Dalam konteks belajar, mahasiswa memiliki kebutuhan dasar terpenuhi cenderung memiliki motivasi lebih tinggi untuk mencapai prestasi akademik.
- iii) Kesesuaian Jurusan  
Kesesuaian jurusan merupakan tingkat kecocokan antara minat, bakat, kemampuan dan latar akademik seseorang dengan bidang studi yang dipilih. Kesesuaian jurusan juga memengaruhi kepuasan karir di masa depan. Mahasiswa yang memilih jurusan yang tepat cenderung lebih siap untuk menghadapi tantangan di bidang yang mereka pilih.
- iv) Kemampuan adaptasi  
Adaptasi dalam lingkungan akademik penting karena dapat membantu mahasiswa menyesuaikan diri dengan perubahan besar dari sekolah ke perguruan tinggi.
- v) Manajemen waktu  
Manajemen waktu adalah proses perencanaan dan pengorganisasian bagaimana seseorang membagi waktu mereka antara berbagai aktivitas untuk meningkatkan produktivitas dan efisiensi. Mengelola waktu dengan efektif memungkinkan mahasiswa untuk memprioritaskan tugas dan mencapai tujuan akademik dengan lebih baik.
- vi) Metode Pembelajaran  
Model pembelajaran diskusi dapat meningkatkan rasa percaya diri dan kemampuan berpikir kritis. Metode pembelajaran yang inovatif dan interaktif dapat meningkatkan keterlibatan dan pemahaman mahasiswa terhadap materi Pelajaran.
- vii) Kondisi keuangan  
Kondisi keuangan mahasiswa dapat menjadi faktor yang memengaruhi akademik. keuangan yang positif dan kemampuan akademik juga dapat memengaruhi perilaku pengelolaan keuangan mahasiswa. Mahasiswa dengan sikap keuangan yang baik cenderung memiliki perilaku pengelolaan keuangan yang lebih baik, yang dapat mendukung prestasi akademik mereka.
- viii) Kondisi Psikologis  
Kondisi psikologis juga menjadi aspek internal yang memengaruhi perilaku dan Keputusan seseorang. Dalam konteks ini, Faktor Psikologis ini meliputi stress, kelelahan, empati dan keterampilan sosial.
- ix) Kondisi Fisik  
Kondisi fisik yang baik akan berdampak pada kemampuan seseorang untuk belajar dan berpartisipasi dalam kegiatan akademik. Aktivitas fisik teratur dapat meningkatkan kinerja otak dan memperbaiki kemampuan kognitif, yang mana dapat meningkatkan prestasi akademik.
- x) Sarana Belajar  
Menurut penelitian sebelumnya [20], sarana belajar adalah peralatan dan perlengkapan yang secara langsung digunakan dalam proses Pendidikan, khususnya proses belajar mengajar, seperti meja, kursi serta alat-alat media pembelajaran. Fasilitas belajar yang memadai yang didapatkan oleh mahasiswa tersebut untuk mendukung proses pembelajaran yang lebih efektif.

## 2.5 PRINCIPAL COMPONENT ANALYSIS (PCA)

*Principal Component Analysis (PCA)* digunakan dalam penelitian ini sebagai teknik reduksi dimensi, dengan tujuan menyederhanakan kompleksitas data yang memiliki banyak variabel tanpa

menghilangkan informasi penting[21]. Hal ini penting untuk mengurangi risiko overfitting, meningkatkan efisiensi model, serta memperjelas interpretasi terhadap pola data. Adapun tahapan yang dilakukan dalam penerapan PCA adalah sebagai berikut:

i) Inisialisasi dan Eksekusi PCA

PCA dijalankan dengan parameter `n_components=0.95`, artinya hanya komponen utama yang secara kumulatif mampu menjelaskan  $\geq 95\%$  total variansi data yang akan dipertahankan. Penerapan PCA dengan threshold 95% variansi bertujuan untuk menyederhanakan data, mengurangi risiko overfitting, dan mempercepat proses pelatihan model machine learning, tanpa menghilangkan informasi penting yang terkandung dalam data asli.

```
# Inisialisasi dan Jalankan PCA
pca = PCA(n_components=0.95) # Ambil 1 komponen untuk ≥95% variansi
X_pca = pca.fit_transform(X_scaled)
```

Gambar 4. Inisialisasi PCA dan Eksekusi PCA

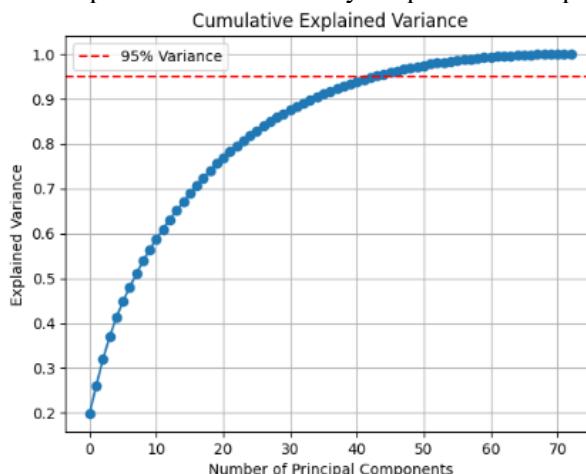
ii) Evaluasi Variansi dan Visualisasi

Variansi yang dijelaskan oleh masing-masing komponen utama dihitung, kemudian divisualisasikan dalam bentuk *cumulative explained variance plot*.

```
# Evaluasi variansi
explained_var = pca.explained_variance_ratio_
cum_var = np.cumsum(explained_var)
```

Gambar 5. Evaluasi Variansi

Setelah menjalankan PCA, langkah berikutnya adalah mengevaluasi seberapa besar variansi yang berhasil dijelaskan oleh setiap komponen utama. Pada Gambar 7 menjelaskan mengenai variabel `explained_var` berisi proporsi variansi yang dijelaskan oleh masing-masing komponen utama hasil PCA. Selanjutnya, `cum_var` menghitung nilai kumulatif variansi yang dijelaskan, sehingga dapat diketahui berapa banyak komponen yang diperlukan untuk mencapai persentase variansi tertentu, disini mencapai 95%. Visualisasinya dapat kita lihat pada Gambar 6.



Gambar 6. Visualisasi Variansi

Gambar 6. menunjukkan grafik *Cumulative Explained Variance* dimana hubungan antara jumlah *principal components* dengan persentase variansi kumulatif yang berhasil dijelaskan oleh komponen-komponen tersebut. Pada grafik, terdapat garis horizontal merah di level 0.95 (atau 95%), yang menandai ambang batas variansi yang ingin dicapai. Kurva biru menunjukkan bahwa semakin banyak komponen yang digunakan, maka semakin tinggi pula variansi kumulatif yang dijelaskan. Terlihat bahwa sebanyak 36 komponen utama diperlukan untuk mencapai  $\geq 95\%$  total variansi data awal.

iii) *Loading Matrix*

Setelah dilakukan transformasi data menggunakan *Principal Component Analysis (PCA)*, hasil reduksi dimensi disimpan ke dalam sebuah *DataFrame* agar memudahkan proses analisis lebih lanjut. Komponen utama hasil PCA dinamai secara berurutan sebagai PC1, PC2, dan seterusnya, sesuai jumlah total komponen.

```
# Buat loading matrix
loading_df = pd.DataFrame(
    pca.components_,
    columns=X.columns,
    index=[f'PC{i+1}' for i in range(pca.n_components_)]
)

# Tampilkan hanya 1 fitur paling dominan untuk tiap PC
for i in range(pca.n_components_):
    pc = loading_df.loc[i]
    fitur_dominan = pc.abs().idxmax()
    nilai_loading = pc[fitur_dominan]
    print(f"PC{i+1} + Fitur dominan: {fitur_dominan} (loading = {nilai_loading:.3f})")
```

Gambar 7. Code Loading Matrix

Pada Gambar 7 selanjutnya, dilakukan ekstraksi loading matrix, yaitu matriks yang menunjukkan bobot kontribusi masing-masing variabel asli terhadap setiap komponen utama. Matriks ini dibentuk dengan mentransposisi matriks `components_` dari hasil PCA, sehingga setiap baris merepresentasikan variabel asli, dan kolomnya mewakili komponen utama.

### 3. HASIL DAN PEMBAHASAN

*Principal Component Analysis* (PCA) digunakan dalam penelitian ini sebagai teknik reduksi dimensi untuk menyederhanakan data dengan jumlah fitur yang banyak tanpa kehilangan informasi penting. Pada penelitian ini, dilakukan eksperimen dengan berbagai threshold dengan variansi yang berbeda. Threshold yang dilakukan dimulai dari 95%, 90%, 85%, 80%, 75% dan 70%. Rentang threshold PCA sebesar 95% hingga 70% dipilih untuk mengevaluasi pengaruh jumlah informasi yang dipertahankan terhadap performa model. Threshold 95% digunakan sebagai batas atas karena mampu menangkap hampir seluruh variansi data, sementara 70% sebagai batas bawah yang masih dianggap layak dalam literatur PCA tanpa kehilangan terlalu banyak informasi penting.

Pemilihan rentang ini memungkinkan analisis yang seimbang antara efisiensi komputasi (semakin sedikit komponen) dan akurasi prediksi (semakin tinggi informasi yang dipertahankan). Hasil pengujian menunjukkan bahwa performa terbaik dicapai pada threshold 95%, di mana model tetap sederhana namun sangat informatif. Untuk lebih jelas, dapat dilihat pada tabel dibawah ini untuk perbandingan hasil antar *threshold* yang sudah dilakukan.

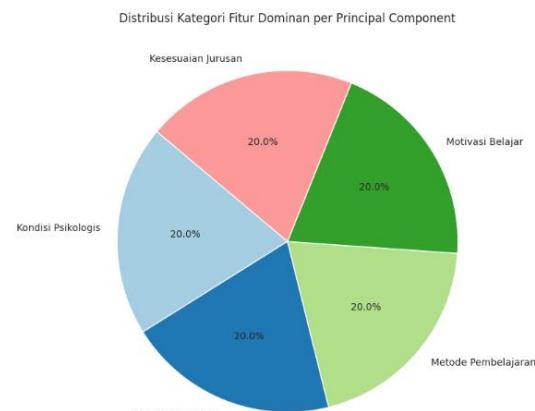
Tabel 3. Perbandingan hasil PCA

Threshold Variansi (%)	Jumlah Komponen Utama	Proporsi Variansi (%)
95	38	0.9508
90	29	0.9022
85	22	0.8513
80	17	0.8069
75	12	0.7540
70	8	0.7023

Tabel 3 menunjukkan bahwa semakin tinggi *threshold* variansi, semakin banyak komponen utama yang diperlukan, namun semakin besar pula informasi data yang dipertahankan. Threshold 95% dipilih karena mampu menjaga 95.08% variansi hanya dengan 38 komponen, jauh lebih efisien dibandingkan 53 atribut awal.

Pemilihan *threshold* ini bukan semata untuk menyederhanakan data, tetapi untuk menjaga keseimbangan antara efisiensi komputasi dan akurasi model. *Threshold* di bawah 90% memang lebih ringan secara komputasi, tetapi mengorbankan informasi penting yang berdampak pada performa model. Sebaliknya, menaikkan *threshold* di atas 95% hanya akan menambah kompleksitas tanpa peningkatan signifikan dalam hasil. Untuk memperjelas distribusi fitur yang paling dominan dalam membentuk masing-masing principal component, dilakukan visualisasi dalam bentuk diagram pie (gambar 8) berdasarkan kategori fitur dominan yang telah diidentifikasi sebelumnya.

Pada Gambar 8 menunjukkan lima kategori fitur dominan hasil PCA dengan threshold 95%, masing-masing berkontribusi sebesar 20% terhadap komponen utama, yaitu: Kondisi Psikologis, Metode Pembelajaran, Kesesuaian Jurusan, Motivasi Belajar dan Kondisi Keuangan. Distribusi ini menunjukkan bahwa keberhasilan akademik mahasiswa penerima beasiswa dipengaruhi secara seimbang oleh kombinasi faktor akademik dan non-akademik. Tidak ada satu kategori yang paling dominan, yang menandakan bahwa pendekatan evaluasi dan pembinaan perlu memperhatikan semua aspek secara holistik — bukan hanya dari sisi akademik, tetapi juga psikologis dan sosial ekonomi.



**Gambar 8. Visualisasi Fitur Dominan**

#### 4. KESIMPULAN

Penelitian ini menunjukkan bahwa PCA efektif dalam mereduksi dimensi data dan mengidentifikasi faktor dominan penentu keberhasilan akademik mahasiswa penerima beasiswa. Dari 54 variabel non-akademik, diperoleh 38 komponen utama dengan total variansi 95,08%. Faktor yang paling dominan adalah kondisi psikologis, metode pembelajaran, kesesuaian jurusan, motivasi belajar, dan kondisi keuangan. Temuan ini dapat digunakan sebagai dasar pengembangan strategi pembinaan mahasiswa dan evaluasi program beasiswa.

#### UCAPAN TERIMA KASIH

Ucapan terima kasih kepada seluruh pihak yang terlibat dalam penelitian ini, kepada mahasiswa yang telah menyisihkan waktu untuk mengisi kuesioner untuk penelitian ini.

#### DAFTAR PUSTAKA

- [1] Darsi and Rasmita, "Prestasi Belajar Mahasiswa ditinjau dari Dimensi Regulitas dalam Kurikulum Merdeka," *Indones. Inst. Corp. Learn. Stud.*, vol. 5, no. September, p. 5, 2024.
- [2] G. Fithria, S. Amanah, and M. Simanjuntak, "Faktor-Faktor yang Memengaruhi Keberhasilan Studi Penerima Bidikmisi di Universitas Sultan Ageng Tirtayasa," *J. Apl. Bisnis dan Manaj.*, vol. 8, no. 3, pp. 833–845, 2022, doi: 10.17358/jabm.8.3.833.
- [3] F. J. Hamu, D. Wea, and N. Setyaningtiyas, "Faktor-Faktor yang Mempengaruhi Kinerja Akademik Mahasiswa : Analisis Structural Equation Model," *J. Paedagogy*, vol. 10, no. 1, p. 175, 2023, doi: 10.33394/jp.v10i1.6473.
- [4] Suripto and A. S. Rahmanita, Rr Nurul Kirana, "Teknik pre-processing dan classification dalam data science," Master of Industrial Engineering BINUS. [Online]. Available: <https://mie.binus.ac.id/2022/08/26/teknik-pre-processing-dan-classification-dalam-data-science/>
- [5] Y. E. Kurniawati, "What is Exploratory Data Analysis?," School of Information Systems BINUS. [Online]. Available: <https://sis.binus.ac.id/2025/01/21/39476/>
- [6] N. Nurmatalasari and E. Purwanto, "Prediksi Performa Mahasiswa Menggunakan Model Regresi Logistik," *J. Deriv. J. Mat. dan Pendidik. Mat.*, vol. 9, no. 2, pp. 145–152, 2022, doi: 10.31316/jderivat.v9i2.2639.
- [7] R. K. Putri, M. Athoillah, and A. Haqiqiyah, "Analisis Faktor yang Mempengaruhi Ketepatan Kelulusan Mahasiswa dengan Algoritma Regresi Linear," *J. Lebesgue J. Ilm. Pendidik. Mat. Mat. dan Stat.*, vol. 5, no. 2, pp. 671–680, 2024, doi: 10.46306/lb.v5i2.571.
- [8] P. Bintoro, Ratnasari, E. Wihardjo, I. P. Putri, and A. Asari, *Pengantar Machine Learning*. Solok, Sumatera Barat: Pt. Mafy Media Literasi Indonesia, 2024. [Online]. Available: <https://repository.um.ac.id/5619/1/fullteks.pdf>
- [9] S. A. Syuhada, S. H. Hasanah, and P. S. Statistika, "Analisis Faktor-Faktor yang Mempengaruhi Tingkat Pengangguran di Indonesia dengan Pendekatan Principal Component Analysis (PCA) dan Analisis Profil," vol. 2, no. 1, pp. 1134–1150, 2025.
- [10] R. A. Zuama, M. A. Ghani, D. Gunawan, and A. L. Matihudin, "Implementasi Metode Waterfall Dalam Mengembangkan Sistem Informasi Ujian Online Dengan Fitur Proctoring," *INFORMATICS Educ. Prof. J. Informatics*, vol. 7, no. 2, p. 218, 2023, doi: <https://doi.org/10.35143/jkt.v11i2.6819>

10.51211/itbi.v7i2.2382.

- [11] M. M. Sanaky, “Analisis Faktor-Faktor Keterlambatan Pada Proyek Pembangunan Gedung Asrama Man 1 Tulehu Maluku Tengah,” *J. Simetrik*, vol. 11, no. 1, pp. 432–439, 2021, doi: 10.31959/js.v11i1.615.
- [12] M. Celestin *et al.*, “PRINCIPAL COMPONENT ANALYSIS FOR SIMPLIFYING MULTIVARIATE FINANCIAL DATA IN,” vol. 9, no. 02, pp. 171–179, 2025.
- [13] E. Elhaik, *Principal Component Analyses (PCA) - based findings in population genetic studies are highly biased and must be reevaluated*. Nature Publishing Group UK, 2022. doi: 10.1038/s41598-022-14395-4.
- [14] A. Hasanah, “Kesesuaian Minat Karir dengan Keputusan Memilih Jurusan di Perguruan Tinggi,” *J. Classr. Action Res.*, vol. 5, pp. 198–202, 2023.
- [15] F. Aryasatya, A. Katrina, R. F. Syabila, F. Siregar, F. Matematika, and P. Alam, “Identifikasi Faktor-Faktor Esensial dalam Hasil Evaluasi Siswa menggunakan Teknik Principal Component Analysis (PCA),” vol. 4, pp. 6423–6437, 2024.
- [16] A. Fadila *et al.*, “Pengaruh Motivasi Belajar terhadap Peningkatan Prestasi Akademik Siswa,” *WACANA J. Bahasa, Seni dan Pengajaran*, vol. 7, pp. 121–133, 2023.
- [17] Filan Firmansyah, Saputra Dwi Nurchaya, and Zuhana Realita Alfy, “Implementasi Algoritma Blowfish Pada Sistem Manajemen Surat Dengan Pendekatan Rational Unified Process Yang Ramah Lingkungan,” *JSiI (Jurnal Sist. Informasi)*, vol. 11, no. 2, pp. 1–6, 2024, doi: 10.30656/jsii.v11i2.9065.
- [18] L. S. Ihzaniah, A. Setiawan, and R. W. N. Wijaya, “Perbandingan Kinerja Metode Regresi K-Nearest Neighbor dan Metode Regresi Linear Berganda pada Data Boston Housing,” *Jambura J. Probab. Stat.*, vol. 4, no. 1, pp. 17–29, 2023, doi: 10.34312/jjps.v4i1.18948.
- [19] A. Rahmadhani, D. S. Sihabudin Sahid, and Y. D. Lulu Widyasari, “Implementasi SEM-Multiple Linear Regression dalam Prediksi Jumlah Pendaftaran Mahasiswa Baru di Perguruan Tinggi XYZ,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 9, no. 2, pp. 150–162, 2023, doi: 10.25077/teknosi.v9i2.2023.150-162.
- [20] N. Nurdiansyah, M. Muliadi, R. Herteno, D. Kartini, and I. Budiman, “Implementasi Metode Principal Component Analysis (PCA) dan Modified K-Nearest Neighbor pada Klasifikasi Citra Daun Tanaman Herbal,” *J. Mnemon.*, vol. 7, no. 1, pp. 1–9, 2024, doi: 10.36040/mnemonic.v7i1.6664.
- [21] H. Sariwati and I. Komputer, “Penerapan Algoritma Principal Component Analysis,” vol. 1, no. 5, pp. 1–16, 2024.