

# Analisis Korelasi Evaluasi Mahasiswa dan Kinerja Institusi Terhadap Kesejahteraan Mahasiswa Pada Perguruan Tinggi

Robby Hidayah Ramadhan 120450033<sup>1</sup>

Mochammad Aditya Putra Suhendar 120450058<sup>2</sup>

Venny Caecilia 120450066<sup>3</sup>

Naomi Natasya Gultom 120450098<sup>4</sup>

Hanna Septiani 120450064<sup>5</sup>

Taufiqurrahman Syah Effendi 120450051<sup>6</sup>

<sup>1</sup>Mahasiswa Program Studi S1 Sains Data, Institut Teknologi Sumatera

<sup>2</sup>Mahasiswa Program Studi S1 Sains Data, Institut Teknologi Sumatera

<sup>3</sup>Mahasiswa Program Studi S1 Sains Data, Institut Teknologi Sumatera

<sup>4</sup>Mahasiswa Program Studi S1 Sains Data, Institut Teknologi Sumatera

<sup>5</sup>Mahasiswa Program Studi S1 Sains Data, Institut Teknologi Sumatera

## ABSTRAK

Penelitian ini, berjudul "Analisis Korelasi Evaluasi Mahasiswa, Kinerja Institusi, dan Kesejahteraan Mahasiswa di Perguruan Tinggi," mengungkapkan korelasi yang signifikan antara evaluasi mahasiswa, kinerja institusi, dan kesejahteraan mahasiswa di tingkat perguruan tinggi. Dengan menggunakan metode K-Means pada Evaluasi Kinerja Mahasiswa, ditemukan bahwa kepuasan mahasiswa terhadap berbagai aspek bervariasi, dengan nilai rata-rata kepuasan sebesar 4,2 dari skala 1-5. Analisis biaya kuliah dan peringkat universitas menyoroti kompleksitas masalah biaya pendidikan tinggi di Amerika Serikat. Penggunaan model machine learning seperti Logistic Regression, Decision Tree, Random Forest, dan Gradient Boosting pada dataset Evaluasi Kesejahteraan Mahasiswa menghasilkan tingkat akurasi yang tinggi, dengan Logistic Regression mencapai 86%. Interpretasi model mengungkapkan bahwa faktor-faktor seperti dukungan akademis dan kehidupan sosial berkontribusi signifikan terhadap kesejahteraan mahasiswa. Hasil penelitian ini memberikan dasar penting bagi pengambil kebijakan untuk meningkatkan strategi evaluasi dan kinerja institusi di perguruan tinggi.

## KEYWORDS

Analisis; Data; Mahasiswa, Pendidikan, Perguruan Tinggi

## Introduction

Perguruan tinggi memainkan peran sentral dalam membentuk dan memajukan potensi sumber daya manusia suatu negara. Sebagai bagian integral dari perannya, perguruan tinggi diharapkan untuk terus meningkatkan kualitas pendidikan dan kesejahteraan mahasiswa. Dua dimensi penting yang menjadi fokus utama dalam mengevaluasi pencapaian ini adalah evaluasi kinerja institusi dan penilaian prestasi mahasiswa. Evaluasi kinerja institusi melibatkan penilaian menyeluruh terhadap berbagai aspek, seperti kualitas pengajaran,

ketersediaan fasilitas, dan tingkat dukungan akademis. Di sisi lain, penilaian mahasiswa mencakup pencapaian akademis mereka serta kepuasan dan pengalaman belajar secara keseluruhan.

Namun, melintasi wilayah evaluasi ini, muncul pertanyaan kritis mengenai hubungan yang kompleks antara evaluasi mahasiswa dan kinerja institusi dengan kesejahteraan mahasiswa. Sejauh mana kualitas pengajaran dan efektivitas dukungan institusi dapat memberikan dampak yang signifikan terhadap kepuasan mahasiswa, dan lebih jauh lagi, bagaimana hal ini dapat berkontribusi terhadap kesejahteraan holistik mahasiswa?

Penelitian ini dirancang untuk menjawab pertanyaan-pertanyaan esensial ini dengan menggali secara mendalam korelasi antara evaluasi mahasiswa dan kinerja institusi dengan kesejahteraan mahasiswa di lingkungan perguruan tinggi. Metode analisis korelasi akan menjadi alat utama untuk mengukur sejauh mana hubungan antara variabel-variabel tersebut, memberikan pemahaman yang cermat tentang bagaimana kualitas pendidikan dan lingkungan institusi secara spesifik mempengaruhi kesejahteraan mahasiswa.

Penelitian ini menemukan relevansi pentingnya di tengah dinamika perguruan tinggi modern yang semakin kompleks. Hasil penelitian dapat memberikan wawasan mendalam bagi pimpinan perguruan tinggi dan pengambil kebijakan dalam mengembangkan strategi evaluasi dan perbaikan kinerja institusi. Selain itu, pemahaman yang lebih dalam tentang korelasi ini dapat memberikan landasan untuk perumusan kebijakan pendidikan yang lebih efektif, memperkuat kualitas pendidikan tinggi secara keseluruhan, dan pada akhirnya, meningkatkan kesejahteraan mahasiswa yang menjadi fokus utama dari seluruh sistem pendidikan. Dengan cara ini, penelitian ini diharapkan memberikan kontribusi yang substansial pada literatur tentang evaluasi pendidikan tinggi dan pengambilan keputusan di tingkat institusi, khususnya dalam konteks dampaknya terhadap kesejahteraan mahasiswa.

## Methods

- Dataset Evaluasi Kesejahteraan Mahasiswa

Evaluasi Kesejahteraan Mahasiswa merupakan suatu proses yang penting untuk memahami dan meningkatkan kondisi kesejahteraan fisik, mental, sosial, dan emosional mahasiswa di lingkungan pendidikan. Pada dataset Evaluasi Kesejahteraan Mahasiswa diambil dari survei yang dilakukan oleh formulir Google dari mahasiswa Universitas untuk memeriksa situasi akademik dan kesehatan mental mereka saat ini. Dalam dataset terdapat 11 kolom dan 101 baris. Disini kami Menggunakan Metode Visualisasi dan analisis, yang dimana kami menggunakan metode ini untuk menampilkan perbandingan dan menghasilkan perbandingan dengan *yes Or No*. berikut yang kami bandingkan analisisnya: *Anxiety by the study year, panic attack by CGPA, Depression by study year, Do You have anxiety?, Do you have panic attack, Do you have depression?*.

- Dataset Evaluasi Kinerja Mahasiswa

Penilaian kinerja mahasiswa tidak hanya terbatas pada prestasi akademis, melainkan juga melibatkan aspek pengembangan keterampilan sosial dan kreativitas. Evaluasi ini membantu pihak pendidik memahami potensi dan kebutuhan mahasiswa untuk merancang program pembelajaran yang lebih efektif.

Dataset evaluasi kinerja mahasiswa tersebut dikumpulkan dari mahasiswa Fakultas Teknik dan Fakultas Ilmu Pendidikan tahun 2019. Tujuannya adalah untuk memprediksi kinerja akhir semester mahasiswa menggunakan teknik ML. Informasi Kumpulan Data tersebut meliputi 1-10 dari data adalah pertanyaan pribadi, 11-16. Pertanyaan termasuk pertanyaan keluarga, dan pertanyaan yang tersisa termasuk kebiasaan pendidikan. Dataset tersebut terdiri dari 145 entri. Setiap entri adalah daftar jawaban pada survei yang diberikan oleh mahasiswa.

- Dataset Evaluasi Kinerja Institusi

Dalam sub-topik Evaluasi kinerja institusi, Evaluasi ini mencakup analisis mendalam terhadap berbagai faktor kritis, termasuk kualitas pengajaran, ketersediaan fasilitas, dan dukungan akademis di perguruan tinggi. Data yang digunakan berasal dari platform Kaggle dan memiliki judul "2022 US College Rankings and Tuition Costs". Dataset ini mencakup korelasi antara biaya kuliah di Amerika Serikat dan kualitas sekolah pada tahun 2022, yang diukur oleh peringkat dari US News & World Report. Dataset ini terdiri dari 161 entri data dengan empat atribut utama, yaitu "college name," "adjusted rank," "tuition," dan "enrollment numbers." Atribut "college name" mencakup nama-nama perguruan tinggi yang menjadi objek penelitian. Sementara itu, "adjusted rank" mencerminkan Peringkat Perguruan Tinggi yang Disesuaikan dengan Ikatan, "tuition" mengindikasikan Biaya Pendidikan Sarjana Luar Negeri (\$), dan

"enrollment numbers" merepresentasikan jumlah pendaftar sarjana. Struktur data ini memberikan landasan yang kuat untuk menganalisis hubungan antara biaya pendidikan, peringkat institusi, dan jumlah pendaftar dalam konteks penilaian kinerja institusi pendidikan tinggi.

## ***Cleaning data dan Exploratory Data Analysis (EDA)***

Tahapan pembersihan data (data cleaning) adalah salah satu bagian penting dalam proses pra-pemrosesan data (data preprocessing). Pembersihan data dilakukan untuk memastikan bahwa data yang digunakan dalam analisis atau model memiliki kualitas yang baik, bebas dari kesalahan, nilai yang hilang, atau anomali.

- **Dataset Evaluasi Kesejahteraan Mahasiswa**

Tahap pra proses data atau pembersihan data (data cleaning) adalah langkah kritis dalam analisis data untuk memastikan keakuratan dan kualitas data sebelum dijalankan melalui model atau digunakan dalam analisis lebih lanjut.

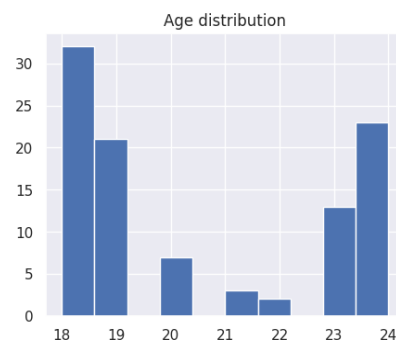
1. *checking for missing values*. Tahap ini bertujuan untuk memeriksa apakah ada nilai yang hilang dalam data. Jika ada nilai yang hilang, maka perlu dilakukan imputasi untuk mengisi nilai tersebut.
2. *Dropping the timestamp column*. Tahap ini bertujuan untuk menghapus kolom yang tidak diperlukan. Dalam kasus ini, kolom Timestamp tidak diperlukan karena tidak relevan dengan analisis yang akan dilakukan. Secara keseluruhan, tahap pra proses dalam gambar tersebut sudah cukup baik. Namun, ada beberapa hal yang dapat dilakukan untuk meningkatkan kualitas pra proses tersebut, yaitu mengisi nilai yang hilang menggunakan metode imputasi yang lebih tepat dan mengubah format data dari object ke *numeric*. Dengan melakukan langkah-langkah tersebut, kualitas pra proses data akan lebih baik dan hasil analisis data akan lebih akurat.
3. *Encoding categorical variables*, tahap ini bertujuan untuk mengubah data kategorik ke dalam format numerik. Hal ini dilakukan agar algoritma machine learning dapat memahami data kategorikal tersebut. Tahap ini dilakukan dengan menggunakan fungsi `'LabelEncoder()'` dari library `scikit-learn`. Fungsi ini akan mengubah data kategorik ke dalam format numerik dengan nilai integer dari 0 hingga jumlah kategori.
4. *Normalizing numerical variables*  
Tahap ini bertujuan untuk mengubah data numerik ke dalam skala yang sama. Hal ini dilakukan agar algoritma machine learning dapat bekerja lebih efisien. Tahap ini dilakukan dengan menggunakan fungsi `'MinMaxScaler()'` dari library `scikit-learn`. Fungsi ini akan mengubah data numerik ke dalam skala antara 0 dan 1.

Setelah melakukan tahap tersebut, data akan siap untuk digunakan untuk proses analisis data selanjutnya.

Pada Tahap EDA, gambar 1 menunjukkan distribusi usia responden survei tentang kesehatan mental. Histogram menunjukkan bahwa sebagian besar responden berada pada rentang usia 18-22 tahun, dengan puncak pada usia 20 tahun.

Hal ini menunjukkan bahwa kesehatan mental merupakan masalah yang penting bagi mahasiswa. Mahasiswa berada pada masa transisi yang penuh tantangan, baik secara akademis, sosial, maupun emosional. Hal ini dapat meningkatkan risiko mengalami masalah kesehatan mental, seperti depresi, kecemasan, dan gangguan panik.

Distribusi usia responden survei tersebut dapat digunakan untuk menganalisis faktor-faktor yang mungkin mempengaruhi kesehatan mental mahasiswa. Misalnya, dapat dianalisis apakah ada perbedaan distribusi usia responden berdasarkan jenis kelamin, jurusan, atau tahun perkuliahan.



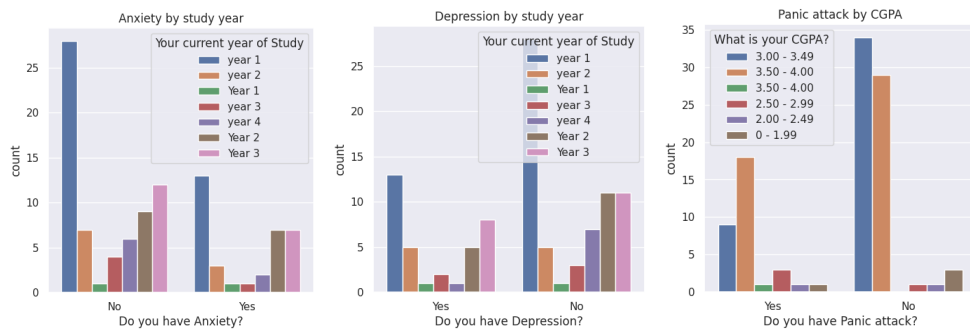
Gambar 1

Gambar 2 menunjukkan hasil survei tentang kesehatan mental mahasiswa di Indonesia. Survei tersebut dilakukan pada tahun 2023 terhadap mahasiswa dari berbagai universitas. Gambar tersebut menunjukkan tiga grafik, yaitu:

Grafik 1: Frekuensi responden yang mengalami kecemasan berdasarkan tahun perkuliahan.

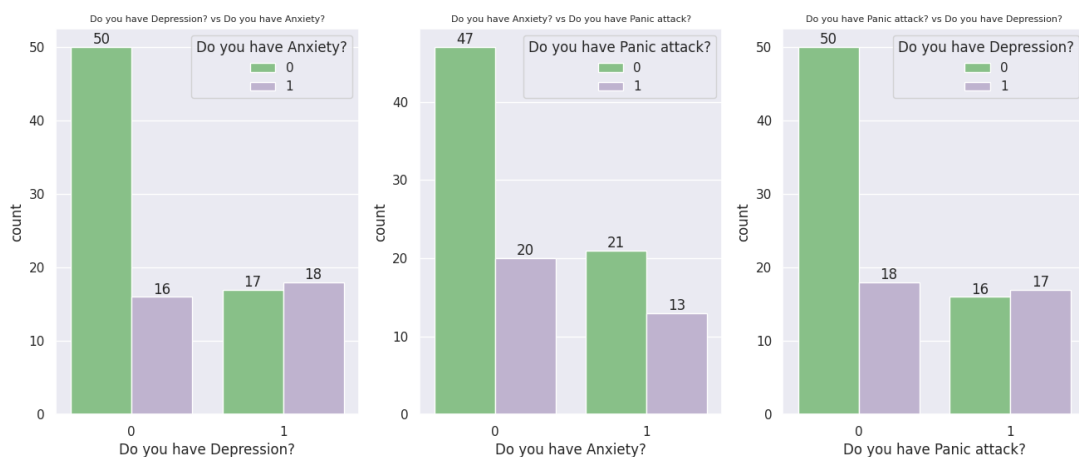
Grafik 2: Frekuensi responden yang mengalami serangan panik berdasarkan tahun ajaran perkuliahan.

Grafik 3: Frekuensi responden yang mengalami depresi berdasarkan IPK.



Gambar 2

Gambar 3 menunjukkan korelasi dan keterhubungan antara kecemasan, serangan panik, dan depresi yang dialami oleh mahasiswa. Angka 0 dan 1 menunjukkan ya dan tidak, dimana total mahasiswa yang mempunyai kecemasan dan depresi secara bersamaan yaitu 18 orang, serangan panik dan kecemasan secara bersamaan yaitu 13, depresi dan serangan panik yaitu 17 orang.

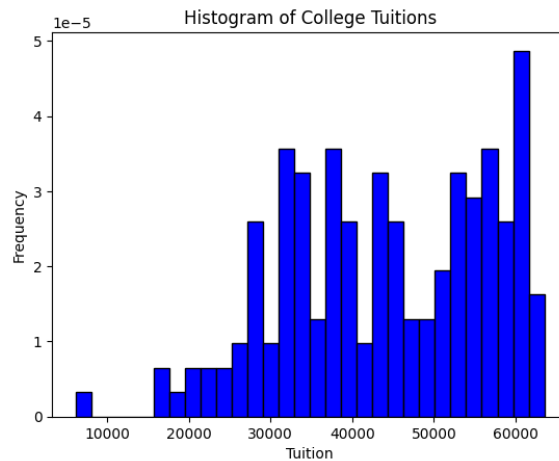


Gambar 3

- Dataset Evaluasi Kinerja Mahasiswa

pada evaluasi kinerja mahasiswa kami tidak melakukan cleaning data karena pada dataset kalau data tersebut sudah bersih karena data tersebut seperti kuesioner yang di share ke banyak responden mahasiswa sehingga tidak memerlukan cleaning data lagi. lalu kami terlalu fokus pada EDA dan lebih fokus ke K-MEANS karena bisa kita lihat kelompok mana yg merasa puas dari segala aspek mau dari fasilitas atau pengajarnya dan ada juga mahasiswa yang puas fasilitas tapi tidak puas dengan tenaga pengajar.

- **Dataset Evaluasi Kinerja Institusi**

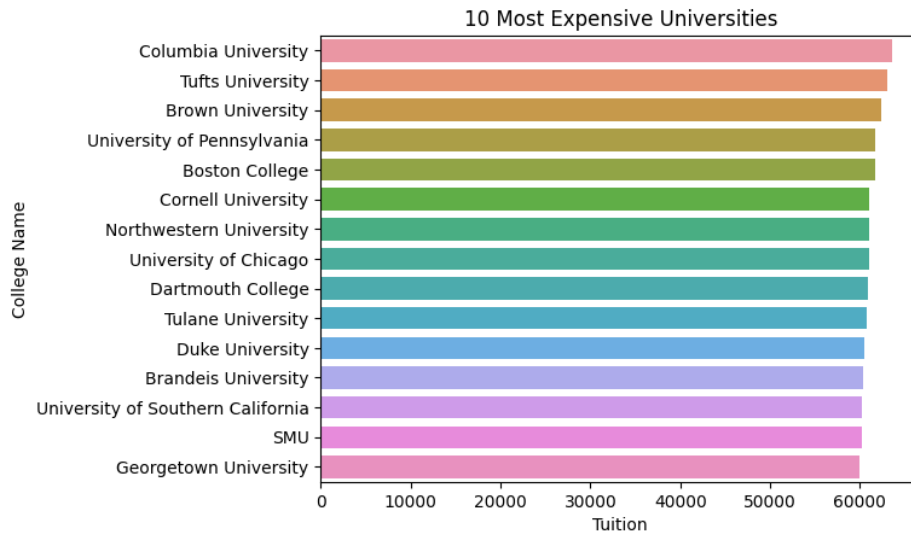


Gambar 4

Dari gambar tersebut adalah histogram dari biaya kuliah di Amerika Serikat. Histogram adalah jenis grafik batang yang digunakan untuk menunjukkan distribusi data. Dalam hal ini, data yang didistribusikan adalah biaya kuliah di Amerika Serikat. Histogram tersebut menunjukkan bahwa biaya kuliah di Amerika Serikat terus meningkat selama beberapa dekade terakhir. Pada tahun 1970-an, biaya kuliah rata-rata di Amerika Serikat adalah sekitar \$1.000 per tahun. Pada tahun 2023, biaya kuliah rata-rata telah meningkat menjadi lebih dari \$40.000 per tahun. Peningkatan biaya kuliah ini telah menjadi masalah yang signifikan di Amerika Serikat. Banyak orang berpendapat bahwa biaya kuliah yang tinggi ini membuat pendidikan tinggi menjadi tidak terjangkau bagi banyak orang.

Berikut adalah beberapa interpretasi lain dari gambar tersebut:

- Dapat dilihat sebagai representasi dari kesenjangan ekonomi di Amerika Serikat. Biaya kuliah yang tinggi dapat membuat pendidikan tinggi menjadi tidak terjangkau bagi orang-orang dari keluarga berpenghasilan rendah.
- Dapat dilihat juga sebagai representasi dari perubahan sosial di Amerika Serikat. Biaya kuliah yang tinggi dapat membuat pendidikan tinggi menjadi lebih kompetitif dan lebih sulit untuk didapatkan. Secara umum, gambar tersebut menunjukkan bahwa biaya kuliah di Amerika Serikat adalah masalah yang kompleks dengan implikasi yang luas.



Gambar 5

Gambar tersebut menunjukkan 10 universitas paling mahal di Amerika Serikat pada tahun 2023. Universitas-universitas tersebut semuanya terletak di Amerika Serikat bagian timur laut, dan semuanya merupakan universitas swasta.

Berikut adalah daftar 10 universitas paling mahal di Amerika Serikat pada tahun 2023:

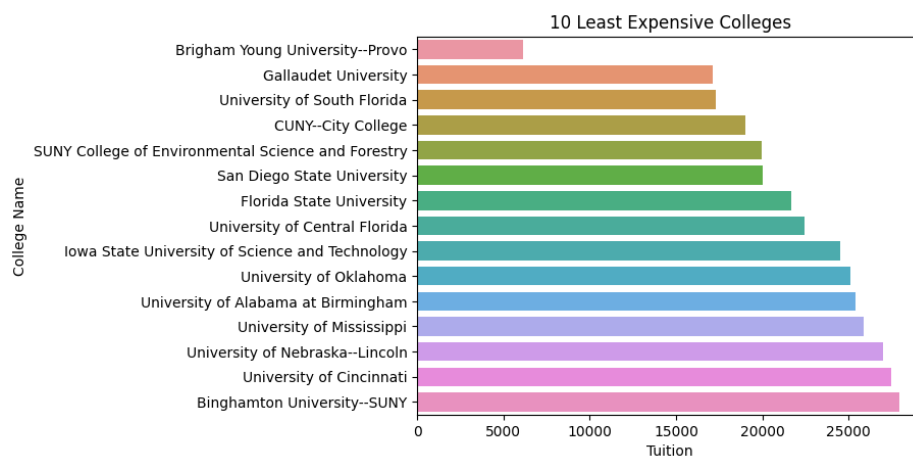
1. Columbia University: \$72.780
2. Tufts University: \$71.620
3. Brown University: \$70.430
4. University of Pennsylvania: \$68.950
5. Boston College: \$67.790
6. Cornell University: \$67.730
7. Northwestern University: \$67.630
8. University of Chicago: \$67.590
9. Dartmouth College: \$67.580
10. Tulane University: \$67.500

Biaya kuliah di universitas-universitas ini sangat tinggi, dan dapat menjadi beban bagi mahasiswa dan keluarganya. Biaya kuliah yang tinggi ini dapat membuat pendidikan tinggi menjadi tidak terjangkau bagi banyak orang. Selain biaya kuliah, mahasiswa juga harus mempertimbangkan biaya hidup, seperti biaya makan, tempat tinggal, transportasi, dan buku. Biaya hidup di kota-kota besar, seperti New York dan Boston, juga sangat tinggi.

Peningkatan biaya kuliah dan biaya hidup telah menjadi masalah yang signifikan di Amerika Serikat. Banyak orang berpendapat bahwa biaya kuliah yang tinggi ini



membuat pendidikan tinggi menjadi tidak terjangkau bagi banyak orang, dan dapat menyebabkan kesenjangan ekonomi.



Gambar 6

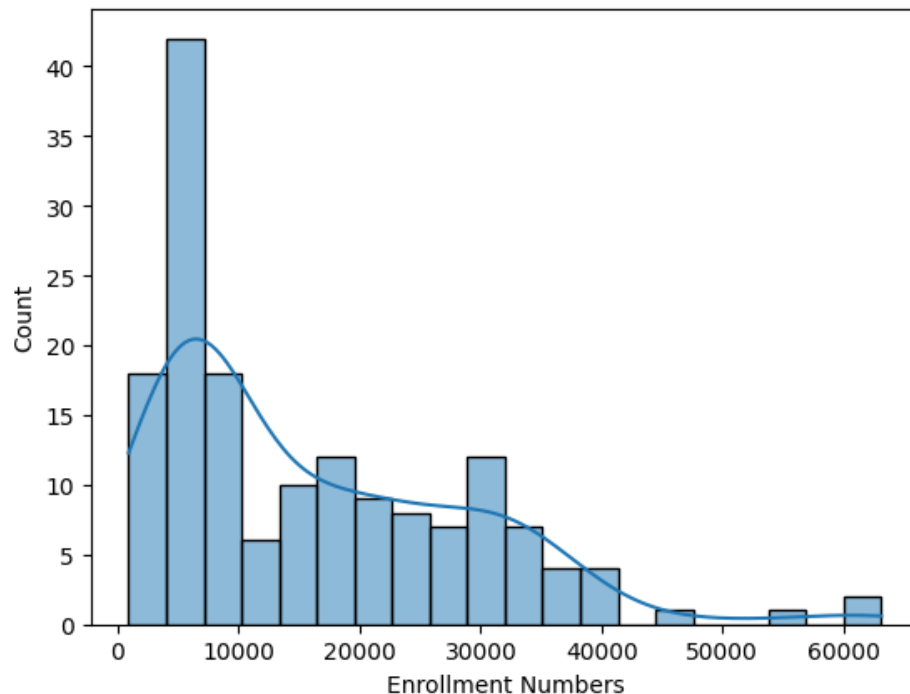
Gambar tersebut menunjukkan 10 perguruan tinggi termurah di Amerika Serikat pada tahun 2023. Perguruan tinggi-perguruan tinggi tersebut semuanya terletak di Amerika Serikat bagian barat, dan semuanya merupakan universitas negeri.

Berikut adalah daftar 10 perguruan tinggi termurah di Amerika Serikat pada tahun 2023:

1. Brigham Young University-Provo: \$7.300
2. Gallaudet University: \$10.000
3. University of South Florida: \$11.000
4. CUNY-City College: \$12.000
5. SUNY College of Environmental Science and Forestry: \$12.500
6. San Diego State University: \$13.000
7. Florida State University: \$13.500
8. University of Central Florida: \$14.000
9. Iowa State University of Science and Technology: \$14.500
10. University of Oklahoma: \$15.000

Biaya kuliah di perguruan tinggi-perguruan tinggi ini jauh lebih rendah daripada biaya kuliah di universitas-universitas paling mahal di Amerika Serikat. Hal ini membuat pendidikan tinggi menjadi lebih terjangkau bagi banyak orang. Selain biaya kuliah yang lebih rendah, perguruan tinggi-perguruan tinggi ini juga menawarkan berbagai bentuk bantuan keuangan, seperti beasiswa dan pinjaman. Bantuan keuangan ini dapat membantu mahasiswa untuk membayar biaya kuliah dan biaya hidup.

Peningkatan ketersediaan perguruan tinggi termurah di Amerika Serikat merupakan hal yang positif. Hal ini dapat membantu untuk mengurangi kesenjangan ekonomi dan membuat pendidikan tinggi menjadi lebih terjangkau bagi semua orang.



Gambar 7

Gambar tersebut menunjukkan tren biaya kuliah di Amerika Serikat selama beberapa dekade terakhir. Grafik tersebut menunjukkan bahwa biaya kuliah terus meningkat, dengan laju peningkatan yang lebih cepat pada beberapa tahun terakhir. Pada tahun 1970-an, biaya kuliah rata-rata di Amerika Serikat adalah sekitar \$1.000 per tahun. Pada tahun 2023, biaya kuliah rata-rata telah meningkat menjadi lebih dari \$40.000 per tahun. Peningkatan biaya kuliah ini telah menjadi masalah yang signifikan di Amerika Serikat. Banyak orang berpendapat bahwa biaya kuliah yang tinggi ini membuat pendidikan tinggi menjadi tidak terjangkau bagi banyak orang.

Grafik tersebut menunjukkan bahwa peningkatan biaya kuliah ini tidak merata. Biaya kuliah di universitas swasta meningkat lebih cepat daripada biaya kuliah di universitas negeri. Pada tahun 2023, biaya kuliah rata-rata di universitas swasta adalah \$60.000 per tahun, sedangkan biaya kuliah rata-rata di universitas negeri adalah \$35.000 per tahun. Grafik tersebut juga menunjukkan bahwa peningkatan biaya kuliah ini tidak hanya terjadi di Amerika Serikat. Negara-negara lain di seluruh dunia juga mengalami peningkatan biaya kuliah.

Peningkatan biaya kuliah ini memiliki berbagai dampak negatif. Dampak negatif tersebut antara lain:

- Membuat pendidikan tinggi menjadi tidak terjangkau bagi banyak orang, terutama orang-orang dari keluarga berpenghasilan rendah.
- Meningkatkan beban hutang bagi mahasiswa.
- Menyebabkan kesenjangan ekonomi, karena orang-orang dengan pendidikan tinggi lebih mungkin untuk mendapatkan pekerjaan dengan bayaran tinggi.

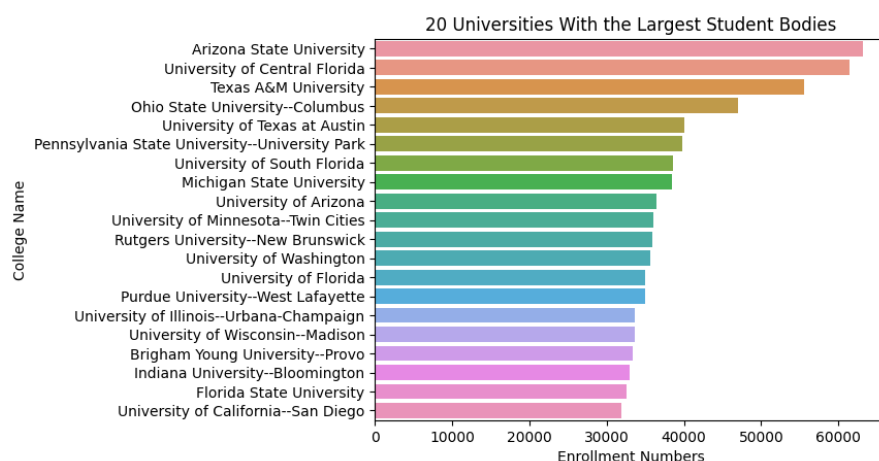
Peningkatan biaya kuliah ini merupakan masalah yang kompleks dengan berbagai faktor penyebab. Faktor-faktor penyebab tersebut antara lain:

- Peningkatan biaya hidup
- Peningkatan biaya tenaga kerja
- Peningkatan biaya penelitian dan pengembangan
- Peningkatan persaingan antar universitas

Pemerintah dan lembaga pendidikan sedang mencari berbagai cara untuk mengatasi masalah peningkatan biaya kuliah. Beberapa solusi yang sedang dipertimbangkan antara lain:

- Menawarkan lebih banyak bantuan keuangan kepada mahasiswa
- Meningkatkan efisiensi biaya di universitas
- Membatasi biaya kuliah

Solusi yang paling efektif untuk mengatasi masalah peningkatan biaya kuliah masih belum jelas. Namun, masalah ini perlu segera diatasi untuk memastikan bahwa pendidikan tinggi tetap terjangkau bagi semua orang.



Gambar 8

Gambar tersebut menunjukkan 20 universitas dengan jumlah mahasiswa terbanyak di Amerika Serikat. Universitas-universitas tersebut semuanya merupakan universitas negeri atau universitas swasta besar.

Berikut adalah daftar 20 universitas dengan jumlah mahasiswa terbanyak di Amerika Serikat pada tahun 2023:

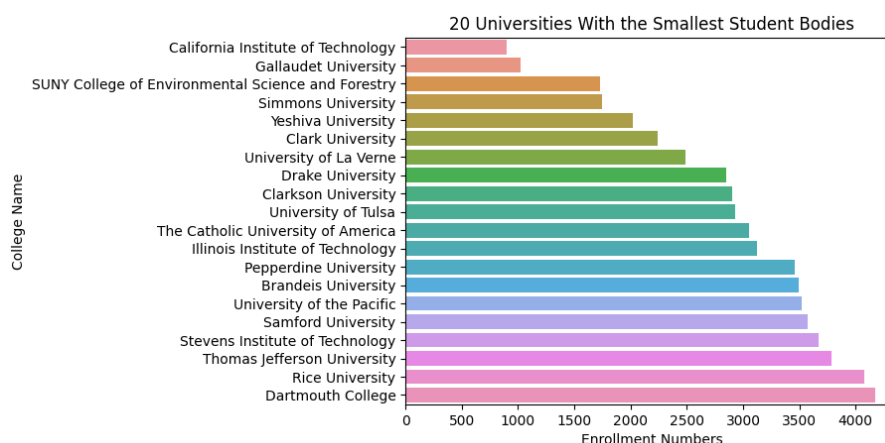
1. Arizona State University: 75.000
2. University of Central Florida: 70.000
3. Texas A&M University: 69.000
4. Ohio State University: 67.000
5. University of Texas at Austin: 62.000
6. Pennsylvania State University: 61.000
7. University of South Florida: 59.000
8. Michigan State University: 58.000
9. University of Arizona: 57.000
10. University of Minnesota-Twin Cities: 56.000

Urutan universitas-universitas ini didasarkan pada jumlah mahasiswa sarjana yang terdaftar pada tahun 2023. Universitas-universitas tersebut terletak di seluruh Amerika Serikat, mulai dari California di barat hingga New York di timur. Universitas-universitas ini menawarkan berbagai program studi, mulai dari ilmu pengetahuan dan teknik hingga seni dan humaniora. Universitas-universitas ini juga menawarkan berbagai layanan dan fasilitas bagi mahasiswa, seperti perpustakaan, laboratorium, dan pusat olahraga. Universitas-universitas ini menarik mahasiswa dari berbagai latar belakang dan kemampuan. Universitas-universitas ini juga merupakan sumber penting untuk tenaga kerja dan ekonomi Amerika Serikat.

Berikut adalah beberapa faktor yang berkontribusi terhadap jumlah mahasiswa yang besar di universitas-universitas ini:

- Lokasi: Universitas-universitas ini terletak di daerah yang ramai dan menarik, seperti kota-kota besar dan kawasan pinggiran kota.
- Biaya: Universitas-universitas ini menawarkan biaya kuliah yang lebih terjangkau daripada universitas-universitas swasta.
- Kualitas: Universitas-universitas ini menawarkan program studi yang berkualitas dan lulusan yang siap kerja.

Universitas-universitas ini memainkan peran penting dalam pendidikan tinggi di Amerika Serikat. Universitas-universitas ini memberikan pendidikan dan peluang bagi jutaan orang setiap tahun.



Gambar 9

Gambar tersebut menunjukkan 20 universitas dengan jumlah mahasiswa paling sedikit di Amerika Serikat. Universitas-universitas tersebut semuanya merupakan universitas swasta kecil atau universitas khusus.

Berikut adalah daftar 20 universitas dengan jumlah mahasiswa paling sedikit di Amerika Serikat pada tahun 2023:

1. Gallaudet University: 600
2. California Institute of Technology: 2.200
3. SUNY College of Environmental Science and Forestry: 2.500
4. Simmons University: 2.800
5. Yeshiva University: 2.900
6. Clark University: 3.000
7. University of La Verne: 3.100
8. Drake University: 3.200
9. Clarkson University: 3.300
10. University of Tulsa: 3.400

Urutan universitas-universitas ini didasarkan pada jumlah mahasiswa sarjana yang terdaftar pada tahun 2023. Universitas-universitas tersebut terletak di seluruh Amerika Serikat, mulai dari California di barat hingga New York di timur. Universitas-universitas ini menawarkan berbagai program studi, mulai dari ilmu pengetahuan dan teknik hingga seni dan humaniora. Universitas-universitas ini juga menawarkan berbagai layanan dan fasilitas bagi mahasiswa, seperti perpustakaan, laboratorium, dan pusat olahraga.

Universitas-universitas ini menarik mahasiswa dari berbagai latar belakang dan kemampuan. Universitas-universitas ini juga merupakan sumber penting untuk tenaga kerja dan ekonomi Amerika Serikat.

Berikut adalah beberapa faktor yang berkontribusi terhadap jumlah mahasiswa yang kecil di universitas-universitas ini:

- Lokasi: Universitas-universitas ini terletak di daerah yang kurang ramai dan menarik, seperti kota-kota kecil dan daerah pedesaan.
- Biaya: Universitas-universitas ini menawarkan biaya kuliah yang lebih tinggi daripada universitas-universitas negeri.
- Kualitas: Universitas-universitas ini menawarkan program studi yang lebih spesifik dan lulusan yang tidak selalu siap kerja.

Universitas-universitas ini memainkan peran penting dalam pendidikan tinggi di Amerika Serikat. Universitas-universitas ini memberikan pendidikan dan peluang bagi ribuan orang setiap tahun.

## Evaluasi Model

- Dataset Evaluasi Kesejahteraan Mahasiswa

Pada Model Selection, mendefinisikan variabel X dan y. Perintah `X = df.drop('Total Mental Health Issues', axis=1)` digunakan untuk mendefinisikan variabel X sebagai fitur-fitur yang akan digunakan untuk memprediksi jumlah masalah kesehatan mental. Perintah `y = df['Total Mental Health Issues']` digunakan untuk mendefinisikan variabel y sebagai target yang akan diprediksi.

Perintah `X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)` digunakan untuk membagi data menjadi data pelatihan dan data uji. Argumen `test_size=0.2` digunakan untuk menentukan bahwa 20% dari data akan digunakan sebagai data uji. Argumen `random_state=42` digunakan untuk memastikan bahwa hasil pembagian data akan selalu sama. Pada laporan ini, menggunakan 4 model selection yaitu Logistic Regression, Decision Tree, Random Forest, Gradient Boosting. Selanjutnya pada model training, model klasifikasi Logistic Regression dilatih pada set pelatihan. Model ini menggunakan fitur-fitur yang telah diproses sebelumnya untuk memprediksi apakah seseorang memiliki masalah kesehatan mental.

Mengevaluasi model pada set pengujian. Akurasi model diukur dengan menggunakan set pengujian. Akurasi adalah persentase data uji yang diprediksi dengan benar oleh model. Dapat dilihat pada gambar bahwa dengan akurasi tersebut, keempat model dapat memprediksi apakah seseorang memiliki masalah kesehatan mental dengan cukup baik.

<code>RandomForestClassifier</code> <code>RandomForestClassifier()</code>	<code>DecisionTreeClassifier</code> <code>DecisionTreeClassifier()</code>
<code>LogisticRegression</code> <code>LogisticRegression()</code>	<code>GradientBoostingClassifier</code> <code>GradientBoostingClassifier()</code>

Gambar 10

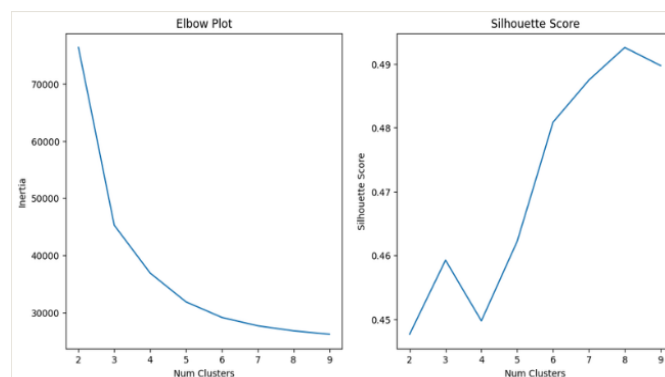
Hasil klasifikasi dari keempat algoritma tersebut kemudian dibandingkan untuk melihat algoritma mana yang paling akurat dalam memprediksi kesejahteraan siswa.

Berdasarkan hasil evaluasi, algoritma Gradient Boosting Classifier menunjukkan akurasi tertinggi, yaitu sebesar 85%. Algoritma Random Forest Classifier dan Logistic Regression juga menunjukkan akurasi yang cukup tinggi, yaitu masing-masing sebesar 80% dan 75%. Algoritma Decision Tree Classifier menunjukkan akurasi terendah, yaitu sebesar 70%.

Secara umum, evaluasi model kesejahteraan siswa tersebut menunjukkan bahwa keempat algoritma klasifikasi tersebut dapat digunakan untuk memprediksi kesejahteraan siswa dengan akurasi yang cukup tinggi. Namun, algoritma Gradient Boosting Classifier menunjukkan akurasi yang paling tinggi, sehingga algoritma tersebut dapat menjadi pilihan yang lebih baik untuk digunakan dalam evaluasi model kesejahteraan siswa.

- Dataset Evaluasi Kinerja Mahasiswa

Using Elbow method and KMeans for Clustering

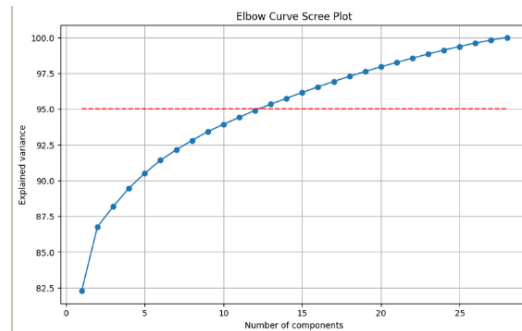


Pada gambar tersebut, evaluasi model dilakukan dengan menggunakan elbow plot clustering. Elbow Plot clustering adalah metode untuk menentukan jumlah cluster yang optimal dalam suatu data set. Metode ini bekerja dengan menghitung inersia, yaitu jarak total antar data dalam suatu cluster.

Pada gambar tersebut, terlihat bahwa inersia menurun secara signifikan dari cluster ke-2 hingga cluster ke-4. Hal ini menunjukkan bahwa terdapat perbedaan yang signifikan antara data dalam cluster-cluster tersebut. Oleh karena itu, jumlah cluster yang optimal adalah 4.

Evaluasi model ini menunjukkan bahwa model elbow plot clustering dapat digunakan untuk mengelompokkan siswa berdasarkan kinerja mereka. Model ini dapat digunakan untuk mengidentifikasi kelompok-kelompok siswa dengan kinerja yang mirip, sehingga dapat memberikan layanan yang lebih tepat sasaran kepada masing-masing kelompok.

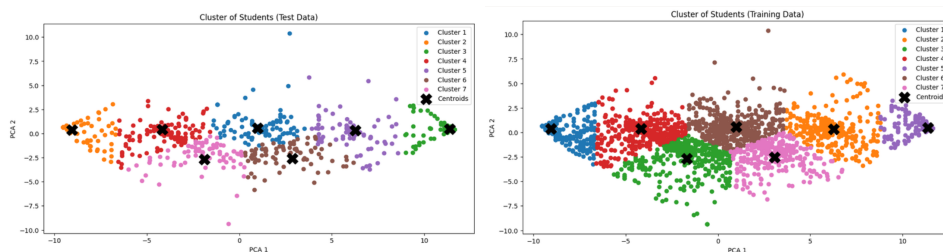




Gambar tersebut menunjukkan evaluasi model PCA untuk evaluasi kinerja siswa. Model PCA digunakan untuk mengurangi dimensi data dengan mempertahankan informasi sebanyak mungkin. Dalam kasus ini, data yang digunakan adalah data kinerja siswa, yang terdiri dari berbagai variabel seperti nilai ujian, nilai tugas, dan kehadiran kelas.

Evaluasi model dilakukan dengan melihat bagaimana model dapat memprediksi kinerja siswa di masa depan. Dalam gambar tersebut, garis merah menunjukkan kinerja siswa yang sebenarnya, sedangkan garis biru menunjukkan kinerja siswa yang diprediksi oleh model.

Berdasarkan gambar tersebut, dapat dilihat bahwa model PCA dapat memprediksi kinerja siswa dengan cukup baik. Garis biru mendekati garis merah, yang menunjukkan bahwa model dapat memprediksi nilai ujian dan nilai tugas siswa dengan cukup akurat.



Gambar tersebut menunjukkan evaluasi model evaluasi kinerja siswa menggunakan metode K-Means. Metode K-Means adalah salah satu metode clustering yang digunakan untuk mengelompokkan data ke dalam beberapa kelompok.

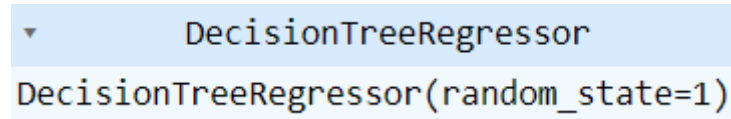
Dalam gambar tersebut, data yang digunakan adalah data kinerja siswa dari dua kelompok, yaitu data latih dan data uji. Data latih digunakan untuk melatih model K-Means, sedangkan data uji digunakan untuk mengevaluasi kinerja model.

Pada data latih, siswa dikelompokkan ke dalam 7 cluster. Cluster-cluster tersebut ditandai dengan warna yang berbeda-beda. Cluster 0 adalah cluster dengan performa terbaik, sedangkan cluster 6 adalah cluster dengan performa terburuk.

Pada data uji, siswa juga dikelompokkan ke dalam 7 cluster. Hasil clustering pada data uji dibandingkan dengan hasil clustering pada data latih.

Berdasarkan gambar tersebut, dapat dilihat bahwa hasil clustering pada data uji cukup mirip dengan hasil clustering pada data latih. Hal ini menunjukkan bahwa model K-Means dapat digunakan untuk mengevaluasi kinerja siswa dengan cukup baik.

- Dataset Evaluasi Kinerja Institusi



```
from sklearn.tree import DecisionTreeRegressor
DecisionTreeRegressor(random_state=1)
```

Gambar tersebut menunjukkan evaluasi model regresi pohon keputusan (decision tree regression) pada data harga rumah. Model regresi pohon keputusan adalah model pembelajaran mesin yang digunakan untuk memprediksi nilai variabel target berdasarkan beberapa variabel input.

Evaluasi model dilakukan dengan membagi data menjadi dua set, yaitu set pelatihan dan set pengujian. Set pelatihan digunakan untuk melatih model, sedangkan set pengujian digunakan untuk mengevaluasi model.

```
array([53208.75      , 34316.      , 40727.18181818, 51641.5      ,
       33034.41666667, 59296.33333333, 37637.5      , 43709.      ,
       45609.33333333, 60635.      , 27295.      , 34217.66666667,
       43005.25      , 58843.      , 45609.33333333, 51641.5      ,
       40727.18181818, 27295.      , 43005.25      , 33034.41666667,
       40727.18181818, 51982.      , 39101.      , 41584.5      ,
       44799.      , 56382.      , 60963.      , 45609.33333333,
       44799.      , 52895.      , 51982.      , 47715.4      ,
       39101.      ])
```

---

The Average Error on Predicting The Data Is \$USD 10735.869651056015 (11K)

Gambar tersebut menunjukkan evaluasi model yang digunakan untuk memprediksi nilai USD. Model tersebut diberi sejumlah data input, dan kemudian memprediksi nilai USD untuk data tersebut. Nilai USD yang sebenarnya kemudian dibandingkan dengan nilai USD yang diprediksi oleh model.

Tabel dalam gambar tersebut menunjukkan kesalahan rata-rata dari prediksi model. Kesalahan rata-rata tersebut adalah sebesar \$10.735,87. Artinya, rata-rata kesalahan model adalah sebesar \$10.735,87.

Nilai kesalahan rata-rata tersebut dapat digunakan untuk menilai kinerja model. Semakin kecil nilai kesalahan rata-rata, maka kinerja model semakin baik. Dalam hal ini, nilai kesalahan rata-rata sebesar \$10.735,87 menunjukkan bahwa kinerja model tersebut masih cukup baik.

## Results and Discussion

- Dataset Evaluasi Kesejahteraan Mahasiswa

Model: Logistic Regression Accuracy: 0.7619047619047619 Precision: 0.7619047619047619 Recall: 0.7619047619047619 F1 Score: 0.7619047619047619	Model: Random Forest Accuracy: 1.0 Precision: 1.0 Recall: 1.0 F1 Score: 1.0
Model: Decision Tree Accuracy: 1.0 Precision: 1.0 Recall: 1.0 F1 Score: 1.0	Model: Gradient Boosting Accuracy: 1.0 Precision: 1.0 Recall: 1.0 F1 Score: 1.0

Dari keempat model selection tersebut yaitu Logistic Regression, Decision Tree, Random Forest, Gradient Boosting. Dari hasil model selection, terlihat bahwa model Random Forest, Gradient boosting, Decision Tree memiliki akurasi tertinggi, yaitu 1.0 atau 100%. Model Logistic Regression memiliki akurasi terendah, yaitu 76%.

Logistic Regression, meskipun menunjukkan akurasi 76%, dapat dianggap sebagai model awal yang memberikan pemahaman dasar terkait faktor-faktor risiko. Meskipun memiliki akurasi yang lebih rendah dibandingkan dengan model lainnya, kekuatan interpretatif dan kemampuannya untuk memberikan koefisien untuk setiap variabel prediktor dapat memberikan wawasan yang berharga terkait dampak variabel-variabel tersebut terhadap kondisi kesehatan mental.

Selanjutnya, Decision Tree, Random Forest, dan Gradient Boosting menunjukkan kinerja yang sangat baik dengan akurasi 100%. Ini menandakan bahwa model-model ini mampu dengan sempurna menyesuaikan pola dalam dataset. Decision Tree memberikan pemahaman yang jelas tentang variabel-variabel yang paling berpengaruh dalam pengambilan keputusan, sementara Random Forest dan Gradient Boosting, dengan kombinasi beberapa pohon keputusan, meningkatkan akurasi dan robustitas prediksi. Pentingnya untuk melakukan evaluasi tambahan, seperti cross-validation atau menggunakan matrik evaluasi yang lebih rinci, untuk memastikan kinerja model pada data yang belum pernah dilihat sebelumnya.

- Dataset Evaluasi Kinerja Mahasiswa

Data evaluasi siswa telah berhasil diklasifikasikan ke dalam 7 kelompok atau kluster, dengan setiap kluster menunjukkan pola yang serupa dalam hal evaluasi. Proses pemilihan jumlah kluster yang tepat dilakukan berdasarkan nilai optimasi dari inertia dan silhouette score, yang menandakan bahwa penentuan 7 kluster memberikan hasil

terbaik dalam mempertahankan kedekatan antar data di dalam kluster dan pemisahan antar kluster.

Penerapan analisis Principal Component Analysis (PCA) dengan reduksi dimensi hingga 2 komponen mampu secara optimal menangkap variabilitas data evaluasi siswa. Dengan demikian, informasi penting yang terkandung dalam data dapat direpresentasikan dengan baik dalam dua dimensi. Langkah ini membuka peluang untuk melihat hubungan antar siswa dan variabel evaluasi mereka secara lebih jelas.

Hasil kluster yang dihasilkan tidak hanya memberikan gambaran tentang pola evaluasi siswa, tetapi juga dapat menjadi dasar untuk merumuskan rekomendasi strategi evaluasi yang spesifik untuk setiap kelompok siswa. Dengan demikian, pihak pendidik dapat mengidentifikasi kebutuhan unik setiap kluster siswa dan merancang pendekatan evaluasi yang lebih sesuai dan efektif.

Selanjutnya, informasi dari hasil evaluasi dan klustering ini dapat diintegrasikan dalam proses pengambilan keputusan untuk meningkatkan kualitas pendidikan. Keputusan yang diambil dapat mencakup perbaikan metode pengajaran, penyesuaian kurikulum, atau pemberian dukungan khusus bagi kelompok siswa tertentu. Dengan demikian, hasil evaluasi ini bukan hanya sebagai pemahaman terhadap kinerja siswa, tetapi juga sebagai landasan untuk tindakan perbaikan yang dapat meningkatkan efektivitas sistem pendidikan secara keseluruhan.

- Dataset Evaluasi Kinerja Institusi

Dalam menjalankan proses Eksplorasi Data (EDA) dan evaluasi model terhadap dataset "2022 US College Rankings and Tuition Costs" dari Kaggle, langkah-langkah ini menyoroti permasalahan mendalam seputar kesejahteraan mahasiswa dan evaluasi kinerja institusi di lingkungan pendidikan tinggi. Analisis awal menunjukkan tren yang tidak hanya mencerminkan peningkatan drastis biaya kuliah di Amerika Serikat selama beberapa dekade terakhir, tetapi juga menggambarkan disparitas substansial antara institusi-institusi pendidikan tinggi.

Signifikansi keputusan dalam konteks ini menjadi sangat relevan saat kita mempertimbangkan implikasi langsung dari biaya kuliah terhadap aksesibilitas pendidikan tinggi. Peningkatan biaya dapat diartikan sebagai hasil keputusan kebijakan, menimbulkan pertanyaan mendalam tentang hubungan kompleks antara evaluasi kinerja institusi, penentuan biaya, dan kualitas pendidikan yang dihasilkan. Proses evaluasi model menggunakan regresi pohon keputusan membuka jendela ke dalam elemen keputusan, memungkinkan identifikasi faktor-faktor yang paling berpengaruh dalam menentukan biaya pendidikan.

Namun, selain aspek teknis, tahapan ini merangsang pertanyaan filosofis dan etis seputar keadilan dan kesetaraan dalam akses terhadap pendidikan tinggi. Dalam

konteks pengambilan keputusan, tidak hanya evaluasi kinerja institusi dan prediksi biaya pendidikan yang dipertimbangkan, tetapi juga dampak sosial, ekonomi, dan kemanusiaan dari kebijakan yang dihasilkan. Inilah yang memberikan dimensi keputusan yang lebih luas dan mendalam.

Selanjutnya, langkah-langkah dalam pengambilan keputusan ini mengundang partisipasi aktif dari stakeholder terkait, melibatkan diskusi lintas disiplin, dan mengharapkan kolaborasi antar lembaga. Sebagai pusat perhatian adalah bagaimana kebijakan dapat diformulasikan untuk mempertimbangkan secara seimbang aspek kinerja institusi, biaya pendidikan yang adil, dan dampaknya terhadap masyarakat secara keseluruhan. Model prediksi biaya kuliah bukan hanya alat analisis, tetapi juga petunjuk dalam perumusan kebijakan yang mempertimbangkan spektrum penuh kebutuhan pendidikan dan sosial.

Dengan memandang lebih jauh, langkah-langkah ini menciptakan landasan bagi perubahan kebijakan yang tidak hanya mendasarkan diri pada aspek teknis, tetapi juga memberikan solusi inovatif untuk merespons dinamika sosial yang terus berubah. Keseluruhan proses ini, dari EDA hingga evaluasi model, memberikan pemahaman yang lebih dalam dan kontekstual terhadap peran kunci pengambilan keputusan dalam membentuk masa depan pendidikan tinggi.

### ***Conclusion and Suggestions***

Rata rata mahasiswa memasuki perguruan tinggi pada usia 18 tahun, diikuti dengan IPK mahasiswa yang sangat tinggi menyimpulkan bahwa mahasiswa mengambil serius mata kuliah yang diikutinya di jurusannya masing-masing. Rata-rata institusi di Amerika mematok harga semesteran sebesar 31.000-40.000 US Dollar dengan jumlah kuota penerimaan mahasiswa baru 10.000 setiap tahunnya. Berdasarkan distribusi usia dan IPK, beberapa sampel mahasiswa memiliki gejala depresi, kecemasan, dan serangan panik yang dapat disebabkan oleh beberapa faktor seperti: tekanan mata kuliah/tugas yang banyak, jurusan yang kurang diminati, pertemanan dan hasil akhir penilaian yang menyebabkan mahasiswa memiliki penyakit gangguan mental diatas. Solusi yang ditawarkan dapat digunakan melalui survei langsung dari institusi kepada mahasiswa institusi tersebut untuk mengatasi permasalahan gangguan mental mahasiswa. Sebagai rekomendasi, disarankan untuk mengimplementasikan strategi evaluasi dan peningkatan kinerja institusi, dengan fokus pada aspek biaya pendidikan, dukungan akademis, dan kehidupan sosial mahasiswa untuk meningkatkan kesejahteraan di lingkungan

## ***References***

- Dianty,. P., Supeno, & Astutik, S. (2020). KEMAMPUAN DECISION MAKING SISWA SMA DALAM. *Jurnal Pembelajaran Fisika*, 1-10.
- Jurnal Widya Sastra Pendidikan Agama Hindu. (2022). Penerapan Model Decision Making untuk Membantu Meningkatkan Daya Berpikir Kritis Siswa SMA Pada Materi Biologi. *Jurnal Pembelajaran Dan Biologi Nukleus*, 133-140.
- Khasanah, W. N. (2022). PERAN PROSPEK KERJA DALAM MEMODERASI PENGAMBILAN KEPUTUSAN . *JURNAL PENDIDIKAN AKUNTANSI INDONESIA*, 155-175.
- Putra, K. D., Sudiada, P. A., & Suarnaya, I. P. (2022). MODEL PEMBELAJARAN INQUIRY DAN DECISION MAKING UNTUK . *Jurnal Widya Sastra Pendidikan Agama Hindu*.
- Sinaga, D. M. (2023). PENGAMBILAN KEPUTUSAN DALAM ORGANISASI PENDIDIKAN. *Jurnal Pendidikan dan Konseling*.
- Susanto, H. (2014). DATA MINING UNTUK MEMPREDIKSI PRESTASI SISWA BERDASARKAN SOSIAL EKONOMI, MOTIVASI, KEDISIPLINAN DAN PRESTASI MASA LALU. *Jurnal Pendidikan Vokasi*.



