

# Penerapan Algoritma K-Means dan Naïve zBayes Dalam Menganalisa Tingkat Kepuasan Mahasiswa Terhadap Pelayanan Kampus

Johan Leonardo Naibaho

Universitas Harapan Medan, Medan, Indonesia  
e-mail: Johanleonardo17219@gmail.com

**Abstrak**— Penelitian ini bertujuan untuk menganalisis tingkat kepuasan mahasiswa terhadap pelayanan kampus menggunakan dua algoritma data mining, yaitu *K-Means* dan *Naive Bayes*. Algoritma *K-Means* digunakan untuk mengelompokkan mahasiswa berdasarkan tingkat kepuasan mereka, sementara algoritma *Naive Bayes* digunakan untuk memprediksi tingkat kepuasan mahasiswa berdasarkan atribut-atribut tertentu seperti fasilitas, layanan akademik, dan dukungan administratif. Data yang digunakan dalam penelitian ini diperoleh melalui survei yang disebarakan kepada mahasiswa Hasil pengelompokan dengan *K-Means* menunjukkan bahwa terdapat beberapa kelompok utama yang menggambarkan tingkat kepuasan yang berbeda di kalangan mahasiswa. Sementara itu, hasil prediksi dengan *Naive Bayes* menunjukkan tingkat akurasi yang cukup tinggi dalam mengklasifikasikan tingkat kepuasan mahasiswa. Dengan demikian, kombinasi kedua algoritma ini dapat memberikan wawasan yang mendalam bagi manajemen kampus dalam meningkatkan kualitas pelayanan dan kepuasan mahasiswa. Penelitian ini juga memberikan rekomendasi untuk peningkatan pelayanan berdasarkan temuan dari analisis data.

**Kata Kunci** : Data Mining, K-Means, Naive Bayes, Clustering, klasifikasi

**Abstract**— This study aims to analyze the level of student satisfaction with campus services using two data mining algorithms, namely K-Means and Naive Bayes. K-Means algorithm is used to cluster students based on their satisfaction level, while Naive Bayes algorithm is used to predict student satisfaction level based on certain attributes such as facilities, academic services, and administrative support. The results of clustering with K-Means show that there are several main groups that describe different levels of satisfaction among students. Meanwhile, the prediction results with Naive Bayes showed a fairly high level of accuracy in classifying student satisfaction levels. Thus, the combination of these two algorithms can provide deep insights for campus management in improving service quality and student satisfaction. This research also provides recommendations for service improvement based on the findings from data analysis.

**Keywords**: Data Mining, Algoritma K-means, Naïve Bayes, Clustering, Classification

## 1. PENDAHULUAN

Teknologi Informasi adalah istilah yang merujuk pada penggunaan teknologi dalam pengumpulan, penyimpanan, pengolahan, dan penyebaran informasi. Ini mencakup segala sesuatu yang terkait dengan penggunaan perangkat keras, perangkat lunak, jaringan, dan infrastruktur untuk mengelola data dan informasi. Teknologi ini mencakup perangkat keras seperti komputer, server, dan jaringan, serta perangkat lunak yang berfungsi untuk mengelola sistem dan aplikasi. Perkembangan teknologi yang pesat telah melahirkan berbagai inovasi, seperti kecerdasan buatan (AI), komputasi awan (cloud computing), dan Internet of Things (IoT). AI mampu mengotomasi proses analisis data, sedangkan cloud computing memungkinkan penyimpanan data yang aman dan mudah diakses dari berbagai perangkat. IoT mendukung interkoneksi antarperangkat untuk mengumpulkan dan berbagi data secara real-time. Sehingga di zaman sekarang teknologi sudah sangat berkembang pesat dan hampir digunakan di berbagai aspek [1]. Data Mining adalah proses ekstraksi informasi yang berharga dan tersembunyi dari sekumpulan data besar. Dalam era digital ini, jumlah data yang dihasilkan terus meningkat dengan cepat, baik dari sumber-sumber online maupun offline. Oleh karena itu, teknik analisis data mining menjadi semakin penting untuk mengungkap pola, tren, dan pengetahuan yang terkandung dalam data tersebut. Data mining merupakan sebuah cara atau teknik yang dipergunakan untuk proses pengolahan data. Proses pengolahan data yang dilakukan pada data mining dilakukan pada data yang besar. Dimana data-data tersebut dikumpulkan terlebih dahulu hingga menjadi kumpulan-kumpulan data kemudian dilakukan proses pengolahan data. Pengolahan data yang dilakukan pada data mining bertujuan untuk menggali informasi-informasi penting ataupun pengetahuan yang tersimpan pada kumpulan data tersebut [2]. Dalam dunia pendidikan, data mining sudah sangat banyak digunakan untuk meneliti dan membuat perkembangan dalam dunia belajar. Universitas atau kampus sebagai penyedia layanan pendidikan tinggi, berupaya untuk menyediakan pelayanan yang terbaik bagi mahasiswa mereka. Tingkat kepuasan mahasiswa adalah salah satu ukuran penting untuk menilai efektivitas dan kualitas pelayanan kampus. Tingkat kepuasan beberapa mahasiswa itu berbeda-beda sehingga agar memudahkan dalam menganalisa tingkat kepuasan mahasiswa terhadap pelayanan kampus. Untuk membantu menganalisis tingkat kepuasan tersebut, algoritma K-Means sering digunakan sebagai metode clustering atau pengelompokan data. Algoritma K-Means merupakan algoritma clustering atau pengelompokan data, Algoritma K-Means dimulai dengan memilih secara acak K, K di sini merupakan banyaknya

cluster yang ingin dibentuk. Kemudian menetapkan nilai K secara random, untuk sementara nilai tersebut menjadi pusat dari cluster atau biasa disebut dengan centroid, mean atau “means” [3]. Selanjutnya, data akan dikelompokkan berdasarkan kemiripan nilai ke dalam cluster masing-masing. Dengan metode ini, kampus dapat mengidentifikasi kelompok mahasiswa dengan tingkat kepuasan yang berbeda. Sedangkan algoritma Naïve Bayes adalah metode klasifikasi yang didasarkan pada teorema Bayes, yang dikembangkan oleh ilmuwan Inggris, Thomas Bayes. Algoritma ini menggunakan pendekatan probabilitas dan statistik untuk melakukan klasifikasi data. Dalam algoritma ini, setiap atribut dan kelas memiliki nilai probabilitas masing-masing yang dihitung berdasarkan data yang tersedia. Proses klasifikasi pada Naïve Bayes berfokus pada pemanfaatan konsep statistik, di mana keputusan akhir ditentukan oleh probabilitas tertinggi dari kelas yang dituju. Algoritma ini sering digunakan dalam data mining karena kemampuannya yang cepat dan efektif dalam mengolah data berukuran besar, terutama untuk mengidentifikasi pola dan tren yang tersembunyi. Hasil akhir penentuan ataupun pengambilan keputusan berdasarkan dengan nilai probabilitas dari kelas yang dituju [4]. Berdasarkan pemaparan diatas, penulis menyusun skripsi dengan judul “Penerapan Algoritma K-Means dan Naïve Bayes Dalam Menganalisa Tingkat Kepuasan Mahasiswa Terhadap Pelayanan Kampus” di Universitas Harapan Medan.

## 2. METODE PENELITIAN

### 2.1. Data Mining

Data mining juga dapat diartikan sebagai serangkaian suatu proses dalam mencari atau menggali nilai tambah suatu data yang berupa pengetahuan yang selama ini tidak diketahui secara manual yang pengetahuannya dapat bermanfaat. Data mining adalah salah satu bentuk implementasi yang diterapkan untuk mencari sebuah model dan pola yang mampu melakukan prediksi pada suatu data berdasarkan data sebelumnya di periode waktu tertentu [5]. Saat ini data mining sudah sangat banyak digunakan, hampir di setiap aspek dan bidang sudah menerapkan data mining. Begitu juga dalam dunia pendidikan, terutama pada perguruan tinggi data mining dapat digunakan sebagai alat untuk mengukur tingkat kepuasan mahasiswa pada pelayanan kampus. Pengukuran tingkat kepuasan mahasiswa adalah sebuah kegiatan yang seharusnya dilakukan secara berkala untuk mengetahui berapa tingkat pelayanan yang telah diberikan. Dengan mengetahui tingkat kepuasan mahasiswa, pihak perguruan tinggi dapat memperbaiki layanan apabila kurang baik dan lebih meningkatkan pelayanan apabila tingkat kepuasan sudah baik [6].

### 2.2. Teknik Data Mining

Ketika data mining digunakan dalam mencari informasi ada beberapa teknik data mining yang digunakan, berikut beberapa teknik data mining yang sering digunakan [3]:

- a. *Association Discovery*
- b. *Clustering*
- c. *Sequential Discovery*
- d. *Classification*
- e. Neural Network

### 2.3. Algoritma K-Means

K-Means adalah metode clustering atau pengelompokkan data non-hierarki yang mempartisi data yang ada ke dalam bentuk dua atau lebih kelompok. K-Means mempartisi data ke dalam kelompok yang mempunyai karakteristik sama. K-means merupakan salah satu algoritma unsupervised learning dan termasuk dalam metode klusterisasi non hierarchical, yang berarti jumlah kluster tidak diketahui dan diinisiasikan di awal. K-means bertujuan untuk mengelompokkan tiap observasi pada data ke dalam suatu grup (atau kluster) menggunakan pengukuran jarak Euclidean, di mana observasi atau poin data yang berada dalam satu kluster adalah yang semirip mungkin, sedangkan observasi atau poin data yang berada dalam kluster yang berbeda adalah sebisa mungkin tidak serupa [7].

Berikut beberapa tahap dalam menggunakan Algoritma K-Means:

- 1).Menentukan K (nilainya bebas) sebagai jumlah cluster yang ingin dibentuk.
- 2).Membangkitkan nilai random untuk pusat cluster awal (centroid) sebanyak K.
- 3).Menghitung jarak setiap data input terhadap masing–masing centroid menggunakan rumus jarak Euclidean Distance hingga ditemukan jarak yang paling dekat dari setiap data dengan centroid. Berikut adalah persamaan Euclidian Distance:

$$d(xj - \mu j) = \sqrt{\sum (xi - \mu j)^2} \quad (1)$$

- 4).Mengklasifikasikan setiap data berdasarkan kedekatannya dengan centroid (jarak terkecil).  
 5).Memperbaharui nilai Nilai centroid baru di peroleh dari rata-rata cluster yang bersangkutan dengan menggunakan rumus:

$$\mu_j(t+1) = \frac{1}{n_{sj}} \sum_{j \in S_j} x_j$$

$\mu_j(t+1)$ :centroid baru pada iterasi ke-(t+1), $n_{sj}$ :banyak data pada cluster SJ (2)

- 6).Melakukan perulangan dari langkah 3 hingga 5, sampai anggota tiap cluster tidak ada yang berubah.

### 2.3 Algoritma Naïve Bayes

Naïve Bayes Classifier merupakan sebuah metoda klasifikasi yang berakar pada teorema Bayes. Metode pengklasifikasian dengan menggunakan metode probabilitas dan statistik yg dikemukakan oleh ilmuwan Inggris Thomas Bayes , yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes. Ciri utama dari Naïve Bayes Classifier ini adalah asumsi yg sangat kuat (naïf) akan independensi dari masing-masing kondisi / kejadian (Muqorobin, Bagoes Pakarti, 2022.).

Rumus dasar menggunakan algoritma Naïve Bayes adalah sebagai berikut:

$$P(X|Y) = \frac{P(Y|X) \times P(X)}{P(Y)}$$
(3)

- 1). $P(X|Y)$  = Kemungkinan bersyarat. X terjadi jika Y benar
- 2). $P(Y|X)$  = Kemungkinan bersyarat. Y terjadi jika X benar
- 3). $P(X)$  = Kemungkinan X tanpa syarat
- 4). $P(Y)$  = Kemungkinan Y tanpa syarat Kondisi X dan Y tidak boleh sama

## 3.HASIL DAN PEMBAHASAN

### 3.1 Data Awal

Data kuisisioner adalah sekumpulan informasi yang dikumpulkan dari responden melalui pertanyaan-pertanyaan terstruktur yang dirancang untuk mengumpulkan data mengenai pendapat, pengalaman, atau karakteristik tertentu. Dalam konteks analisis tingkat kepuasan mahasiswa terhadap pelayanan kampus, kuisisioner biasanya berisi pertanyaan yang berkaitan dengan berbagai aspek pelayanan yang diberikan oleh universitas, seperti fasilitas akademik, administrasi, layanan perpustakaan, dukungan teknis, dan lainnya.

**Tabel 1.** Kategori Pertanyaan

No.	Kode	Kategori	Max Score
1	X <sub>1</sub>	Ketersediaan Fasilitas	10
2	X <sub>2</sub>	Kualitas Pengajaran	15
3	X <sub>3</sub>	Layanan Administratif	10
4	X <sub>4</sub>	Sumber Daya	10
5	X <sub>5</sub>	Lingkungan Kampus	10
6	X <sub>6</sub>	Ekstrakurikuler	10

**Tabel 2.** Jenis Pertanyaan

No.	Kategori	Pertanyaan
1	Ketersediaan Fasilitas	Bagaimana Anda menilai ketersediaan ruang kelas di kampus?
2	Ketersediaan Fasilitas	Bagaimana Anda menilai ketersediaan fasilitas laboratorium di kampus?
3	Kualitas Pengajaran	Seberapa jelas dosen menjelaskan materi kuliah?
4	Kualitas Pengajaran	Bagaimana dosen mengelola waktu selama perkuliahan?

5	Kualitas Pengajaran	Seberapa baik dosen dalam memberikan contoh yang relevan dengan materi kuliah?
6	Layanan Administratif	Bagaimana penilaian Anda terhadap kecepatan pelayanan di kantor administrasi?
7	Layanan Administratif	Bagaimana penilaian Anda terhadap keramahan dan kesopanan staf administrasi?
8	Sumber Daya	Bagaimana pendapat Anda tentang layanan konseling dan bimbingan akademik yang disediakan di kampus?
9	Sumber Daya	Apakah Anda puas dengan ketersediaan perpustakaan dan koleksi buku di kampus?
10	Lingkungan Kampus	Bagaimana pendapat Anda mengenai kebersihan lingkungan kampus ini
11	Lingkungan Kampus	Seberapa efektif sistem keamanan kampus dalam menjaga keamanan dan ketertiban?
12	Ekstrakurikuler	Seberapa baik kualitas kegiatan yang disediakan oleh ekstrakurikuler kampus?
13	Ekstrakurikuler	Bagaimana pendapat Anda mengenai fasilitas yang tersedia untuk mendukung kegiatan ekstrakurikuler?

### 3.2 Penerapan Algoritma K-Means

Algoritma K-Means merupakan model centroid atau titik tengah yang menggunakan centroid untuk membuat cluster. Centroid berupa nilai yang digunakan untuk menghitung jarak suatu objek data terhadap centroid. Tahap ini merupakan tahapan untuk mempersiapkan data yang diperlukan untuk proses Data Mining. Tujuannya adalah agar data yang digunakan benar-benar sesuai dengan permasalahan yang akan dipecahkan, dapat dijamin kebenarannya, dan dalam format yang sesuai atau tepat.

Tabel berikut ini adalah sampel data kuisioner yang diperoleh:

**Tabel 3.** Tabel Normalisasi Data

No.	Nama Mahasiswa	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>
1	Johan	8	11	6	6	10	6
2	Elfrans Arsendy Tampubolon	8	11	8	8	6	3
3	ILFAN DANIEL BARUS	8	6	6	6	8	2
4	Alex prastio sitepu	6	9	6	6	6	6
5	Viktor Alex Sander Aritonang	6	12	8	4	8	4
6	Herdiansyah	6	8	10	5	7	3
7	Johanes	2	3	2	2	2	2
8	Moratama Tampubolon	8	11	7	6	7	4
9	Benny Siswanto Gea	6	11	6	6	6	4
10	Irfan Syahputra	6	11	6	6	6	2
11	Amat Giovani	6	10	7	7	7	2
12	Adrian Yusri	6	9	6	8	6	4
13	Muhammad Teguh Gunawan	7	12	8	7	7	2

14	Chairul fakhri	6	11	6	6	6	2
15	MHD zhanistra	8	12	8	7	7	2
16	Ahmad Bukhori	6	9	6	5	6	4
17	Sri ulina Sinulingga	6	9	6	7	6	5
18	Windi halimardani	6	9	6	6	6	4
19	Safira Nurul lita	8	12	8	7	7	4
20	Ucok jatule	8	9	6	6	8	4
21	Nuar Adrian	8	12	8	8	9	4
22	Tommy abadi putra Marpaung	6	9	6	8	8	2
23	Wahyu hidayat ritonga	7	12	8	7	9	4
24	Anwar Rajab Nasution	6	9	6	6	6	4
25	Arma dani	6	9	6	6	6	3
26	Ilham fadli Harahap	8	10	8	7	6	4
27	Dimas pratama putra	5	9	6	4	6	2
28	Zul hamil yahya Koto	8	12	6	6	6	4
29	Jihan Nabila Fathia	8	12	8	7	7	4
30	Algazali Raihan Batubara	6	12	7	7	8	4
31	M.Hadit Haikal	8	12	8	7	8	4
32	Mellia cristanty br sinuhaji	8	12	8	8	8	6
33	Akbar abdul fattah	8	12	8	8	8	4
34	Surya darma	8	12	8	8	8	3
35	Fahri akbar	8	12	8	8	8	4
36	Mutiara putri	8	12	8	8	8	2
37	Gusti afrida	8	12	8	8	8	2
38	Seftiana baru	8	12	8	8	8	2
39	Raja pandapotan lingga	8	12	8	7	8	2
40	Rizky al amin	8	12	8	8	8	2
41	Thorriq hidayat	8	12	8	8	8	2
42	Dimas Gilang permata	8	12	8	8	8	2
43	Budi abdillah	8	12	8	8	8	2
44	Lazuardi	8	12	8	8	8	2
45	Raja alfarouqih	8	9	6	6	7	2
46	Brikman toyo	8	12	8	8	8	4
47	Fanny thania	8	12	8	8	8	2
48	Mohammad farhan	8	12	8	8	8	8
49	Abdillah Wijaya	8	12	8	8	8	8
50	Zhaky	8	12	8	8	8	8
51	Della rizkynta	8	12	8	8	8	8
52	Ilham pratama	8	12	6	7	8	2
53	Ahmad adi	8	11	7	7	7	4
54	M zain	8	12	8	8	8	6
55	Hazarul rahmat	8	9	6	6	6	6
56	Mohammad yasir	6	9	6	6	6	6
57	Laskamana	4	6	6	5	6	4
58	Lutfi andi	4	6	4	4	4	4
59	Muhammad habibi	4	6	5	5	4	4
60	Moh fadil	6	9	6	6	6	6

Tahap ini dilakukan penerapan algoritma *K-Means* dengan rumus : $d(x,y)= ||x-y||=$

$$\sqrt{\sum_{i=1}^n (x_i - y_i)^2}; i = 1,2,3, \dots, n$$

Dari 25 iterasi yang dilakukan sesuai dengan perhitungan algoritma *K-Means* maka diperoleh hasil akhir sebagai

berikut.

**Tabel 4.** Hasil Klastering

No.	Nama	Cluster	Keterangan
1	ILFAN DANIEL BARUS	Cluster-1	Kurang
2	Johanes	Cluster-1	Kurang
3	Ahmad Bukhori	Cluster-1	Kurang
4	Arma dani	Cluster-1	Kurang
5	Dimas pratama putra	Cluster-1	Kurang
6	Laskamana	Cluster-1	Kurang
7	Lutfi andi	Cluster-1	Kurang
8	Muhammad habibi	Cluster-1	Kurang
9	Johan	Cluster-2	Cukup Baik
10	Elfrans Arsendy Tampubolon	Cluster-2	Cukup Baik
11	Alex prastio sitepu	Cluster-2	Cukup Baik
12	Viktor Alex Sander Aritonang	Cluster-2	Cukup Baik
13	Herdiansyah	Cluster-2	Cukup Baik
14	Moratama Tampubolon	Cluster-2	Cukup Baik
15	Benny Siswanto Gea	Cluster-2	Cukup Baik
16	Irfan Syahputra	Cluster-2	Cukup Baik
17	Amat Giovani	Cluster-2	Cukup Baik
18	Adrian Yusri	Cluster-2	Cukup Baik
19	Muhammad Teguh Gunawan	Cluster-2	Cukup Baik
20	Chairul fakhri	Cluster-2	Cukup Baik
21	MHD zhanistra	Cluster-2	Cukup Baik
22	Sri ulina Sinulingga	Cluster-2	Cukup Baik
23	Windi halimardani	Cluster-2	Cukup Baik
24	Safira Nurul lita	Cluster-2	Cukup Baik
25	Ucok jatule	Cluster-2	Cukup Baik
26	Nuar Adrian	Cluster-2	Cukup Baik
27	Tommy abadi putra Marpaung	Cluster-2	Cukup Baik
28	Wahyu hidayat ritonga	Cluster-2	Cukup Baik
29	Anwar Rajab Nasution	Cluster-2	Cukup Baik
30	Ilham fadli Harahap	Cluster-2	Cukup Baik
31	Zul hamil yahya Koto	Cluster-2	Cukup Baik
32	Jihan Nabila Fathia	Cluster-2	Cukup Baik
33	Algazali Raihan Batubara	Cluster-2	Cukup Baik
34	M.Hadit Haikal	Cluster-2	Cukup Baik
35	Mellia cristanty br sinuhaji	Cluster-2	Cukup Baik
36	Akbar Abdul Fattah	Cluster-2	Cukup Baik
37	Surya darma	Cluster-2	Cukup Baik
38	Fahri akbar	Cluster-2	Cukup Baik
39	Mutiara putri	Cluster-2	Cukup Baik
40	Gusti Afrida	Cluster-2	Cukup Baik
41	Seftiana baru	Cluster-2	Cukup Baik
42	Raja Pandapotan Lingga	Cluster-2	Cukup Baik
43	Rizky al amin	Cluster-2	Cukup Baik
44	Thorriq Hidayat	Cluster-2	Cukup Baik
45	Dimas Gilang Permata	Cluster-2	Cukup Baik

46	Budi abdillah	Cluster-2	Cukup Baik
47	Lazuardi	Cluster-2	Cukup Baik
48	Raja alfarouqih	Cluster-2	Cukup Baik
49	Brikman toyo	Cluster-2	Cukup Baik
50	Fanny thania	Cluster-2	Cukup Baik
51	Ilham pratama	Cluster-2	Cukup Baik
52	Ahmad adi	Cluster-2	Cukup Baik
53	M zain	Cluster-2	Cukup Baik
54	Hazarul rahmat	Cluster-2	Cukup Baik
55	Mohammad yasir	Cluster-2	Cukup Baik
56	Moh fadil	Cluster-2	Cukup Baik
57	Mohammad farhan	Cluster-3	Sangat Baik
58	Abdillah Wijaya	Cluster-3	Sangat Baik
59	Zhaky	Cluster-3	Sangat Baik
60	Della rizkynta	Cluster-3	Sangat Baik

### 3.3 Penerapan Algoritma Naive Bayes

Naive Bayes adalah sekelompok algoritma klasifikasi berdasarkan Teorema Bayes dengan asumsi sederhana bahwa setiap fitur dalam dataset bekerja secara independen satu sama lain. Meskipun asumsi independensi ini jarang benar dalam situasi nyata, algoritma Naive Bayes tetap bekerja dengan baik dalam banyak kasus aplikasi dunia nyata, terutama untuk klasifikasi teks dan pengenalan pola. Berikut ini adalah data latih yang digunakan untuk pelatihan algoritma Naive Bayes.

**Tabel 5.** Data Uji

No	Responden	C1	C2	C3	C4	C5	C6
1	Fransko Gultom	8	12	8	8	8	8
2	Fachri Mufti	8	12	8	8	8	8
3	Sutrisna	8	12	8	8	8	8
4	Muhammad Arsyad	8	12	8	7	6	6
5	Roziardi	8	12	8	7	6	6
6	M.Ramulia	8	12	8	7	6	6

Tahapan pertama akan dilakukan konversi data sesuai dengan ketentuan yang telah ditentukan pada tabel konversi indikator:

**Tabel 6.** Data Konversi

No	Responden	C1	C2	C3	C4	C5	C6	Class
1	Fransko Gultom	Sangat Baik	?					
2	Fachri Mufti	Sangat Baik	?					
3	Sutrisna	Sangat Baik	?					
4	Muhammad Arsyad	Sangat Baik	Sangat Baik	Sangat Baik	Sangat Baik	Baik	Baik	?
5	Roziardi	Sangat Baik	Sangat Baik	Sangat Baik	Sangat Baik	Baik	Baik	?
6	M.Ramulia	Sangat Baik	Sangat Baik	Sangat Baik	Sangat Baik	Baik	Baik	?

a. Perhitungan Fransko Gultom

1). Menghitung jumlah class / label

$$P(Y = \text{Kurang}) = \frac{8}{60} = 0.13333333$$

$$P(Y = \text{Cukup Baik}) = 48/60 = 0.8$$

$$P(Y = \text{Sangat Baik}) = 4/60 = 0.066666667$$

2). Menghitung jumlah kasus yang sama dengan class yang sama.

Perhitungan Class C1

$$P(C1 = \text{Sangat Baik}, | Y = \text{Kurang}) = 1/8$$

$$P(C1 = \text{Sangat Baik}, | Y = \text{Cukup Baik}) = 31/48$$

$$P(C1 = \text{Sangat Baik}, | Y = \text{Sangat Baik}) = 4/4$$

Perhitungan Class C2

$$P(C2 = \text{Sangat Baik}, | Y = \text{Kurang}) = 0/8$$

$$P(C2 = \text{Sangat Baik}, | Y = \text{Cukup Baik}) = 27/48$$

$$P(C2 = \text{Sangat Baik}, | Y = \text{Sangat Baik}) = 4/4$$

Perhitungan Class C3

$$P(C3 = \text{Sangat Baik}, | Y = \text{Kurang}) = 0/8$$

$$P(C3 = \text{Sangat Baik}, | Y = \text{Cukup Baik}) = 27/48$$

$$P(C3 = \text{Sangat Baik}, | Y = \text{Sangat Baik}) = 4/4$$

Perhitungan Class C4

$$P(C4 = \text{Sangat Baik}, | Y = \text{Kurang}) = 0/8$$

$$P(C4 = \text{Sangat Baik}, | Y = \text{Cukup Baik}) = 19/48$$

$$P(C4 = \text{Sangat Baik}, | Y = \text{Sangat Baik}) = 4/4$$

Perhitungan Class C5

$$P(C5 = \text{Sangat Baik}, | Y = \text{Kurang}) = 1/8$$

$$P(C5 = \text{Sangat Baik}, | Y = \text{Cukup Baik}) = 25/48$$

$$P(C5 = \text{Sangat Baik}, | Y = \text{Sangat Baik}) = 4/4$$

Perhitungan Class C6

$$P(C6 = \text{Sangat Baik}, | Y = \text{Kurang}) = 0/8$$

$$P(C6 = \text{Sangat Baik}, | Y = \text{Cukup Baik}) = 0/48$$

$$P(C6 = \text{Sangat Baik}, | Y = \text{Sangat Baik}) = 4/4$$

3). Kalikan semua hasil variabel,

\*Class "Kurang"

$$= 1/8 * 0/8 * 0/8 * 0/8 * 1/8 * 0/8$$

$$= 0$$

\*Class "Cukup Baik"

$$= 31/48 * 27/48 * 27/48 * 19/48 * 25/48 * 0/48$$

$$= 0$$

\*Class "Sangat Baik"

$$= 4/4 * 4/4 * 4/4 * 4/4 * 4/4 * 4/4$$

$$= 1$$

4). Bandingkan hasil dari ketiga class.

Dari hasil yang diperoleh sebelumnya, maka nilai tertinggi berada pada Class "Sangat Baik" yaitu 1.

b. Perhitungan Fachri Mufti

1). Menghitung jumlah class / label

$$P(Y = \text{Kurang}) = 8/60 = 0.133333333$$

$$P(Y = \text{Cukup Baik}) = 48/60 = 0.8$$

$$P(Y = \text{Sangat Baik}) = 4/60 = 0.066666667$$

2). Menghitung jumlah kasus yang sama dengan class yang sama.

Perhitungan Class C1

$$P(C1 = \text{Sangat Baik}, | Y = \text{Kurang}) = 1/8$$

$$P(C1 = \text{Sangat Baik}, | Y = \text{Cukup Baik}) = 31/48$$

$$P(C1 = \text{Sangat Baik}, | Y = \text{Sangat Baik}) = 4/4$$

Perhitungan Class C2

$$P(C2 = \text{Sangat Baik}, | Y = \text{Kurang}) = 0/8$$

$$P(C2 = \text{Sangat Baik}, | Y = \text{Cukup Baik}) = 27/48$$

$$P(C2 = \text{Sangat Baik}, | Y = \text{Sangat Baik}) = 4/4$$

Perhitungan Class C3

$$P(C3 = \text{Sangat Baik}, | Y = \text{Kurang}) = 0/8$$

$$P(C3 = \text{Sangat Baik}, | Y = \text{Cukup Baik}) = 27/48$$

$$P(C3 = \text{Sangat Baik}, | Y = \text{Sangat Baik}) = 4/4$$

Perhitungan Class C4

$$P(C4 = \text{Sangat Baik}, | Y = \text{Kurang}) = 0/8$$

- $P(C4 = \text{Sangat Baik}, | Y = \text{Cukup Baik}) = 19/48$   
 $P(C4 = \text{Sangat Baik}, | Y = \text{Sangat Baik}) = 4/4$   
 Perhitungan Class C5  
 $P(C5 = \text{Sangat Baik}, | Y = \text{Kurang}) = 1/8$   
 $P(C5 = \text{Sangat Baik}, | Y = \text{Cukup Baik}) = 25/48$   
 $P(C5 = \text{Sangat Baik}, | Y = \text{Sangat Baik}) = 4/4$   
 Perhitungan Class C6  
 $P(C6 = \text{Sangat Baik}, | Y = \text{Kurang}) = 0/8$   
 $P(C6 = \text{Sangat Baik}, | Y = \text{Cukup Baik}) = 0/48$   
 $P(C6 = \text{Sangat Baik}, | Y = \text{Sangat Baik}) = 4/4$
- 3). Kalikan semua hasil variabel,
- \* Class “Kurang”  
 $= 1/8 * 0/8 * 0/8 * 0/8 * 1/8 * 0/8$   
 $= 0$
  - \* Class “Cukup Baik”  
 $= 31/48 * 27/48 * 27/48 * 19/48 * 25/48 * 0/48$   
 $= 0$
  - \* Class “Sangat Baik”  
 $= 4/4 * 4/4 * 4/4 * 4/4 * 4/4 * 4/4$   
 $= 1$
- 4). Bandingkan hasil dari ketiga class.  
 Dari hasil yang diperoleh sebelumnya, maka nilai tertinggi berada pada Class “Sangat Baik” yaitu 1.

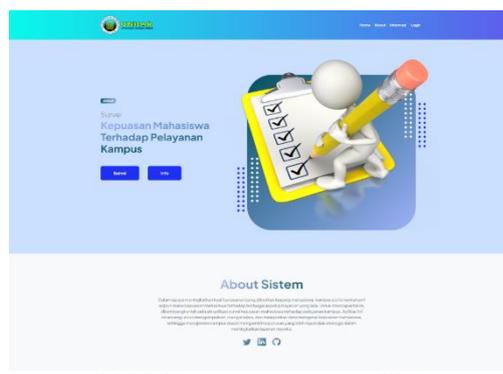
Lakukan perhitungan untuk Sutrisna, Muhammad Arsyad dan yang lainnya sesuai dengan tahapan diatas.

### 3.4 Implementasi Sistem

Untuk mencapai tujuan implementasi, fokus utama adalah menerapkan desain sistem yang telah disusun sebelumnya. Proses ini melibatkan pemilihan perangkat keras dan perangkat lunak yang tepat untuk mendukung simulasi aplikasi analisis tingkat kepuasan mahasiswa terhadap pelayanan kampus menggunakan algoritma K-Means dan Naive Bayes. Dengan demikian, implementasi ini bukan hanya sekedar menerapkan teknologi, tetapi juga berkontribusi pada perbaikan dan pertumbuhan berkelanjutan dari aplikasi analisis tingkat kepuasan mahasiswa terhadap pelayanan kampus yang menggunakan algoritma *K-Means* dan *Naive Bayes*.

#### 1. Halaman Menu Utama

Menu utama adalah tampilan awal ketika *user* memasuki sistem. Halaman ini berisi tampilan luar tentang aplikasi data mining untuk Universitas Harapan Medan.



**Gambar 1.** Form Menu Utama

#### 2. Halaman Isi Data Responden

Halaman Isi Data Responden merupakan formulir yang diperlukan oleh pengguna sebelum mereka dapat mengisi kuisioner.



**Gambar 2.** Halaman Isi Data Responden

3. Halaman Isi Kuisisioner

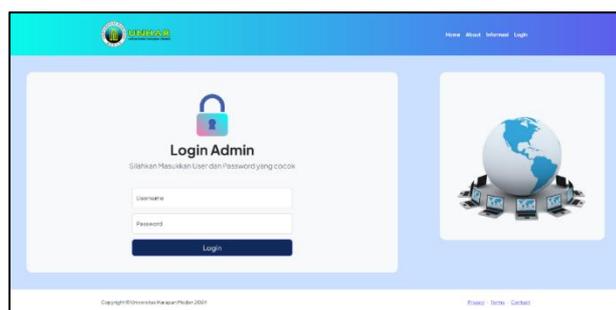
Halaman Isi Kuisisioner adalah tempat di mana responden atau pengunjung *web* diminta untuk menjawab kuisisioner yang disediakan. Berikut adalah desain halaman Isi Kuisisioner.



**Gambar 4.3** Halaman Isi Kuisisioner

4. Halaman *Login Admin*

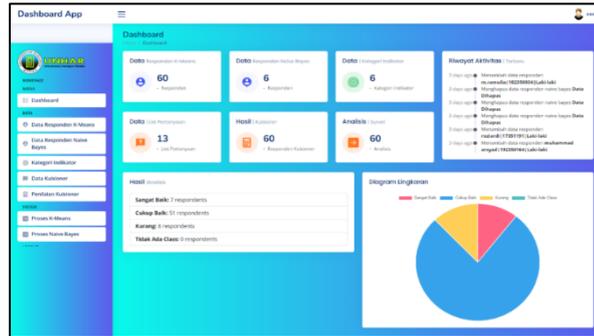
Dalam bagian ini, terdapat penekanan pada desain halaman *Login* yang dirancang khusus untuk pengguna dengan peran sebagai administrator *web*.



**Gambar 4.** *Form Login Admin*

5. Halaman Menu Utama *Admin*

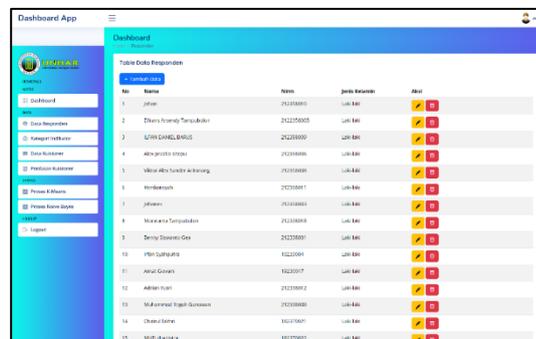
Menu Utama *Admin* dibuat untuk merancang bentuk desain halaman *web* yang akan digunakan oleh *admin*



Gambar 5. Halaman Menu Utama

6. Form Data Responden

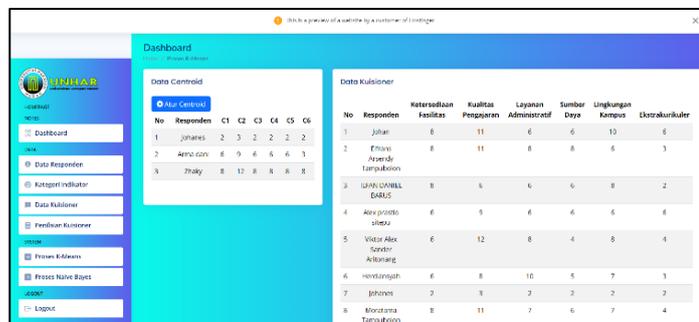
Halaman Data Responden dirancang untuk memberikan fungsionalitas melihat, data Data Responden yang tersimpan dalam *database*.



No	Nama	NPM	Jenis Kelamin	Aksi
1	Juhan	212220001	Laki-laki	[Edit] [Hapus]
2	Enisa Aranyo Tampubolon	212220002	Laki-laki	[Edit] [Hapus]
3	Idris GANGL BIRUS	212220003	Laki-laki	[Edit] [Hapus]
4	Alex pristo silexa	212220004	Laki-laki	[Edit] [Hapus]
5	Vikar Alex Sapin Alrifanang	212220005	Laki-laki	[Edit] [Hapus]
6	Hovlansyah	212220006	Laki-laki	[Edit] [Hapus]
7	Johnes	212220007	Laki-laki	[Edit] [Hapus]
8	Muhammad Tampubolon	212220008	Laki-laki	[Edit] [Hapus]
9	Benny Siregar Gie	212220009	Laki-laki	[Edit] [Hapus]
10	Irfan Supriatna	182220001	Laki-laki	[Edit] [Hapus]
11	Ahmad Givani	182220002	Laki-laki	[Edit] [Hapus]
12	Adrian Irfan	212220012	Laki-laki	[Edit] [Hapus]
13	Muhammad Iqbal Susanto	212220013	Laki-laki	[Edit] [Hapus]
14	Chandra	182220014	Laki-laki	[Edit] [Hapus]
15	Melki Ananta	182220015	Laki-laki	[Edit] [Hapus]

Gambar 6. Form Data Responden

7. Proses K-Means

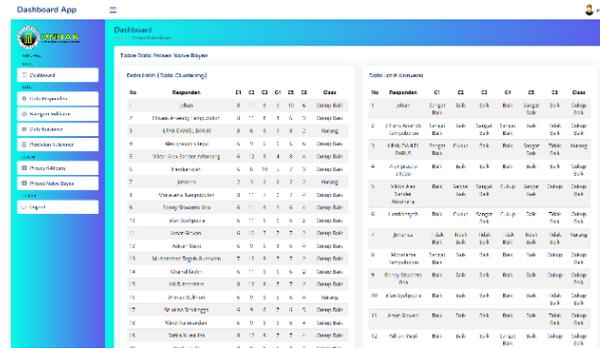


No	Responden	Ketertarikan Fasilitas	Kualitas Pelayanan	Sumber Daya	Lingkungan	Ukttrakurkuler
1	Juhan	8	11	6	6	10
2	Enisa Aranyo Tampubolon	8	11	8	8	6
3	Idris GANGL BIRUS	8	6	6	6	8
4	Alex pristo silexa	6	9	6	6	6
5	Vikar Alex Sapin Alrifanang	6	12	8	4	8
6	Hovlansyah	6	8	10	5	7
7	Johnes	2	8	2	2	2
8	Muhammad Tampubolon	8	11	7	6	7

Gambar 9. Halaman Proses K-Means

8. Proses Naive Bayes

Halaman Proses *Naive Bayes* difungsikan untuk melakukan proses algoritma *Naive Bayes* dari data klastering dari proses sebelumnya yang dijadikan sebagai data pelatihan. Output dari proses ini adalah klastering yang akan digunakan pada inputan metode *Naive Bayes*.



Gambar 10. Halaman Proses Naive Bayes

#### 4. KESIMPULAN

1. Algoritma *K-Means* dan Naïve Bayes dapat digunakan secara bersama untuk mengelompokkan mahasiswa berdasarkan tingkat kepuasan mereka terhadap pelayanan kampus. Algoritma *K-Means* memungkinkan pengelompokkan mahasiswa ke dalam beberapa kluster berdasarkan kemiripan jawaban mereka dalam kuisioner kepuasan. Dengan demikian, mahasiswa yang memiliki pola kepuasan yang sama akan dikelompokkan bersama. Sementara itu, algoritma Naïve Bayes dapat digunakan untuk mengklasifikasikan tingkat kepuasan mahasiswa berdasarkan faktor-faktor tertentu yang telah diidentifikasi dari data.
2. Berbagai faktor dapat mempengaruhi tingkat kepuasan mahasiswa terhadap pelayanan kampus. Beberapa faktor utama yang perlu diperhatikan meliputi kualitas fasilitas akademik (ruang kelas, laboratorium, perpustakaan), layanan administrasi (penerimaan, pendaftaran, bantuan keuangan), dukungan teknis (layanan IT, akses internet), kualitas pengajaran dan interaksi dengan dosen, kegiatan ekstrakurikuler. Dengan memahami faktor-faktor ini, universitas Harapan Medan dapat mengambil langkah-langkah yang tepat untuk meningkatkan pelayanan dan, pada akhirnya, meningkatkan tingkat kepuasan mahasiswa secara keseluruhan. Analisis data menggunakan algoritma *K-Means* dan *Naive Bayes* memungkinkan identifikasi faktor-faktor kunci ini, serta mengukur seberapa besar pengaruh masing-masing faktor terhadap kepuasan mahasiswa.

#### REFERENCES

- [1] A. Anugerah Pratama, “Pemanfaatan Teknologi Informasi dan Komunikasi (TIK) Oleh Rumah Tangga Untuk Kehidupan Sehari-hari,” *J. Inf. Technol. Soc.*, vol. 1, no. 2, 2023, [Online]. Available: <https://jits.unmuhbabel.ac.id/>
- [2] A. Rahmat and R. Permana, “Jurnal Teknorama (Informatika dan Teknologi El Rahma) ARTICLE HISTORY,” 2023.
- [3] M. Zunaidi and K. Erwansyah, “Penerapan Data Mining Untuk Menganalisa Tingkat Kepuasan Pelanggan Telkomsel Terhadap Sikap Pelayanan Caroline Officer Dengan Menggunakan Metode ...,” *J. Cyber Tech*, no. x, 2021, [Online]. Available: <https://ojs.trigunadharma.ac.id/index.php/jct/article/view/451>
- [4] A. Triayudi and G. Soepriyono, “Penerapan Data Mining Untuk Mengukur Kepuasan Mahasiswa Terhadap Pembelajaran dengan Menggunakan Algoritma Naïve Bayes,” *J. Comput. Syst. Informatics*, vol. 4, no. 1, pp. 39–44, 2022, doi: 10.47065/josyc.v4i1.2524.
- [5] I. Priyadi, J. Santony, and J. Na’am, “Data Mining Predictive Modeling for Prediction of Gold Prices Based on Dollar Exchange Rates, Bi Rates and World Crude Oil Prices,” *Indones. J. Artif. Intell. Data Min.*, vol. 2, no. 2, p. 93, Sep. 2019, doi: 10.24014/ijaidm.v2i2.6864.
- [6] D. A. R. Saragih, M. Safii, and ..., “Penerapan Data Mining Klasifikasi Tingkat Kepuasan Mahasiswa Terhadap Pelayanan Sistem Informasi di Program Studi Sistem Informasi,” *J. Inf. ...*, vol. 2, no. 2, pp. 173–177, 2021, [Online]. Available: <http://ejurnal.seminar-id.com/index.php/josh/article/view/628%0Ahttp://ejurnal.seminar-id.com/index.php/josh/article/download/628/425>
- [7] P. Alkhairi and A. P. Windarto, “Penerapan K-Means Cluster pada Daerah Potensi Pertanian Karet Produktif di Sumatera Utara,” *Semin. Nas. Teknol. Komput. Sains*, pp. 762–767, 2019.
- [8] M. Muqorobin, Bagoes Pakarti, “Sistem Prediksi Lama Studi Kuliah Menggunakan Metode Naive Bayes.” [Online]. Available: <https://jurnal.itbaas.ac.id/index.php/jikombis>