

BAB 4

HASIL PENELITIAN

4.1 RINGKASAN HASIL PENELITIAN

Hasil pelaksanaan penelitian tentang hasil yang didapatkan dari analisis pengaruh media sosial terhadap prestasi belajar mahasiswa Fakultas Teknik dan Teknologi Informasi (FTTI) di Universitas Jenderal Achmad Yani Yogyakarta. Dibutuhkan data penelitian berupa kuesioner yang disebar untuk mahasiswa untuk mengetahui pengaruh yang ditimbulkan dari penggunaan media sosial dan juga nilai IPK dari mahasiswa yang sudah mengisi kuesioner dalam mengetahui prestasi akademik. Pengambilan data kuesioner dilakukan mulai tanggal 10 Mei – 18 Mei 2024 yang disebar kepada responden dengan yang mengisi sebanyak 107 dari 4 Prodi. Selanjutnya data responden yang sudah terkumpulkan dilakukan proses preprocessing untuk pembersihan dan *normalisasi* data. Hasil pengelompokan data menggunakan metode Elbow terdapat 3 kelompok anggota dengan cluster 0 memiliki 3 anggota dengan hasil kurang berpengaruh, cluster 1 memiliki 47 anggota dengan hasil berpengaruh dan di cluster 2 memiliki nilai 57 anggota dengan hasil sangat berpengaruh. Evaluasi data dengan menggunakan 2 metrik yaitu metrik Silhouette Score memiliki nilai 0,196 dan metrik Davies-Bouldin Index dengan nilai 1,490. Data yang dihasilkan dari *clustering* akan diimplementasi dalam sistem website sederhana menampilkan data dari pengaruh media sosial dan prestasi akademik. Adapun tampilan 5 menu utama di sistem yaitu *Dashboard, Preprocessing data, Elbow Method, Clustering, dan Testing*.

4.2 KUESIONER MEDIA SOSIAL

Dalam penelitian ini sebelum menyebar kuesioner ke semua responden dapat dianalisis terlebih dahulu dengan variabel, populasi dan sampel, instrumen pertanyaan, teknik instrumen. Agar data dari penelitian dinyatakan valid dan layak untuk proses *clustering*.

4.2.1 Variabel Penelitian

Variabel adalah objek dari inti penelitian yang berfokus pada masalah yang ada dan kemudian ditarik kesimpulan (Priadana & Sunarsih, 2021).

1. Variabel *independent* (variasi bebas) adalah variabel yang dipengaruhi oleh penyebab dari perubahan variabel dependen (terikat). Adapun variabel bebas dari penelitian ini adalah prestasi belajar (X).
2. Variabel *dependent* (Variabel terikat) adalah variabel yang dipengaruhi dari akibat variabel bebas. Adapun variabel dari penelitian ini adalah media sosial (Y).

4.2.2 Populasi dan Sampel

1. Populasi merupakan semua wilayah bagian diukur kuantitas dan karakter tertentu yang telah ditentukan untuk dipelajari dengan hasil akhir berupa kesimpulan (Amin *et al.*, 2023). Populasi dalam penelitian adalah mahasiswa FTTI di Unjaya yang berjumlah 107 mahasiswa dari 4 prodi ditunjukkan pada Tabel 4.1.

Tabel 4.1 Populasi Mahasiswa

No	Program Studi	Angkatan			
		2020	2021	2022	2023
1	Informatika	11	9	9	8
2	Sistem Informasi	8	8	8	7
3	Teknologi Informasi	4	5	1	9
4	Teknik Industri	8	4	4	4
Jumlah Total		107			

2. Sampel adalah setengah sumber data penelitian dari jumlah populasi yang memiliki karakteristik yang dimiliki (Amin *et al.*, 2023). Dalam pengujiannya dari sampel data sebanyak 70 FTTI di Unjaya yang menggunakan media sosial.

4.2.3 Instrumen Pertanyaan Penelitian

Penyusunan pertanyaan penelitian menggunakan variabel sebagai indikator dalam kegiatan dalam penggunaan media sosial ditunjukkan di Tabel 4.2.

Tabel 4.2 Instrumen Pertanyaan Penelitian

No	Variabel	Indikator Pertanyaan	No butir pertanyaan
1	Frekuensi Media Sosial	Penggunaan media sosial	1, 2, 3, 4, 5
2		Media sosial sebagai alat komunikasi interaktif	6,7
3		Alasan penggunaan media sosial	8, 9, 10
4		Sarana media sosial terhadap prestasi belajar	11, 12, 13

Keterangan :

Skala 1: Responden **Sangat Setuju (SS)** dengan pertanyaan yang diberikan

Skala 2: Responden **Setuju (S)** dengan pertanyaan yang diberikan

Skala 3: Responden **Netral (N)** dengan pertanyaan yang diberikan

Skala 4: Responden **Kurang Setuju (KS)** dengan pertanyaan yang diberikan

Skala 5: Responden **Tidak Setuju (TS)** dengan pertanyaan yang diberikan

4.2.4 Teknik Instrumen Penelitian

Penelitian menggunakan metode kuantitatif dalam pengumpulan data berupa kuesioner atau angket. Dalam melakukan penelitian dibutuhkan cek validasi dan reliabilitas untuk keberhasilan karena menjadi penentuan angket atau kuesioner valid atau reliabel dan layak untuk disebarkan ke responden sebagai alat ukur dalam penelitian (Ardiansyah *et al.*, 2023).

1. Uji validitas

Validitas menurut KBBI yaitu uji terhadap penelitian untuk menguji kebenaran kuesioner. Hasil uji validasi ditunjukkan dalam semua butir pertanyaan valid dengan nilai $r_{hitung} > r_{tabel}$ (Ali *et al.*, 2022). Adapun kriteria dalam perhitungan r_{tabel} dengan indikator menggunakan df dengan rumus:

$$df = n - 2$$

(3)

Keterangan n adalah jumlah sampel, data responden 70 maka $70 - 2$. Maka hasil 68 dari nilai r tabel adalah 0,235, dapat ditunjukkan pada Tabel 4.3.

Tabel 4.3 Uji Validitas

PERTANYAAN	R HITUNG	KETERANGAN
P1	0,602	VALID
P2	0,647	VALID
P3	0,594	VALID
P4	0,644	VALID
P5	0,682	VALID
P6	0,582	VALID
P7	0,493	VALID
P8	0,458	VALID
P9	0,447	VALID
P10	0,661	VALID
P11	0,606	VALID
P12	0,743	VALID
P13	0,762	VALID

Tabel 4.3 menunjukkan nilai dari seluruh item memenuhi syarat validitas. Berdasarkan hasil pengujian validitas sesuai output SPSS Versi 25.0 tahun 2024. Hasil menunjukkan bahwa seluruh item memiliki r hitung $>$ r tabel dari 0,235. Jadi, dinyatakan valid dan layak semua item dalam kuesioner untuk dipakai dalam penelitian selanjutnya.

2. Uji Reliabilitas

Reliabilitas berdasarkan KBBI adalah suatu yang bersifat keakuratan dari alat ukur untuk memastikan kuesioner masih layak untuk dijadikan penelitian. Uji Reliabilitas dinyatakan reliabel apabila $\text{cronbach's } \alpha > 0,6000$ atau $\text{cronbach's } \alpha > r \text{ tabel}$ *cronbach alpha* (Anggraini *et al.*, 2022) dapat ditunjukkan pada Tabel 4.4.

Tabel 4.4 Uji Reliabilitas

Reliability Statistics	
Cronbach's	
Alpha	N of Items
.852	13

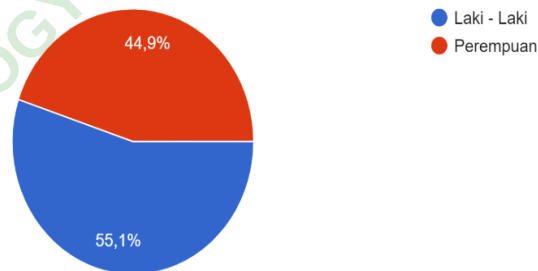
Tabel 4.4 menunjukkan seluruh item memenuhi syarat reliabilitas. Berdasarkan hasil uji reliabilitas sesuai output SPSS Versi 25.0 tahun 2024. Hasil menunjukkan seluruh item memiliki nilai *cronbach alpa* 0,852 > 0,600. Jadi, semua item kuesioner reliabel dan layak dipakai untu proses penelitian.

4.2.5 Responden Penelitian

Untuk penjelasan secara rinci dan jelas tentang deskripsi data yang mengisi responden sebagai berikut:

1. Jenis Kelamin

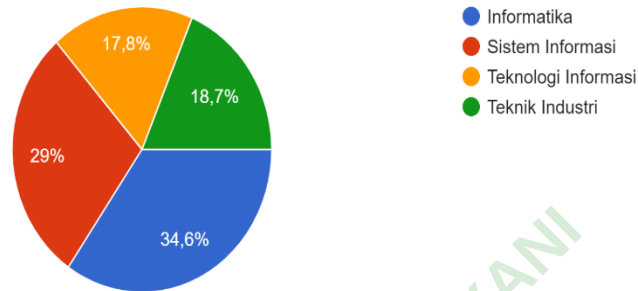
Jenis Kelamin
107 jawaban

**Gambar 4.1** Persentase Jenis Kelamin

Kesimpulan dari Gambar 4.1 bahwa kuesioner yang disebarakan pada mahasiswa FTTI sebanyak laki-laki 59 orang dengan persentase 55,1% dan perempuan 48 dengan persentase 44,9%. Jadi, Jumlah laki-laki dalam penelitian lebih banyak dibandingkan perempuan.

2. Program Studi

Program Studi
107 jawaban

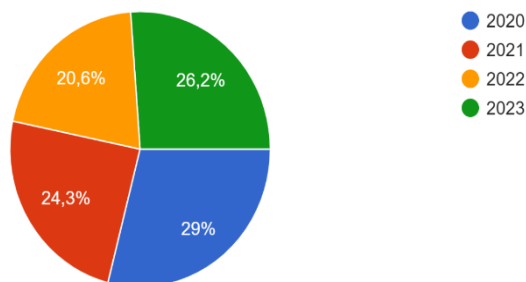


Gambar 4.2 Persentase Program Studi

Kesimpulan dari Gambar 4.2 bahwa kuesioner disebarikan pada mahasiswa FTTI dari program studi informatika sebanyak 37 responden dengan persentase 34,6%, program studi sistem informasi sebanyak 31 responden dengan persentase 29%, program studi teknologi informasi sebanyak 19 responden dengan persentase 17,8%, dan program studi Teknik industri sebanyak 20 responden dengan persentase 18,7%.

3. Tahun Angkatan

Tahun Angkatan
107 jawaban



Gambar 4.3 Persentase Tahun Angkatan

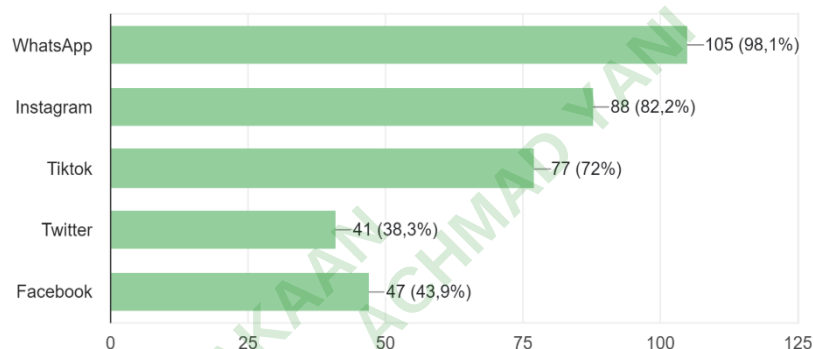
Kesimpulan dari Gambar 4.3 bahwa kuesioner yang disebarikan pada mahasiswa tiap angkatan FTTI dari tahun angkatan 2020 sebanyak 31 responden dengan persentase 29%, tahun angkatan 2021 sebanyak 26

mahasiswa dengan 24,3%, tahun angkatan 2022 sebanyak 22 mahasiswa dengan 20,6%, dan tahun angkatan 2023 sebanyak 28 mahasiswa dengan 26,2%.

4. Media Sosial yang digunakan

Sosial Media yang digunakan

107 jawaban



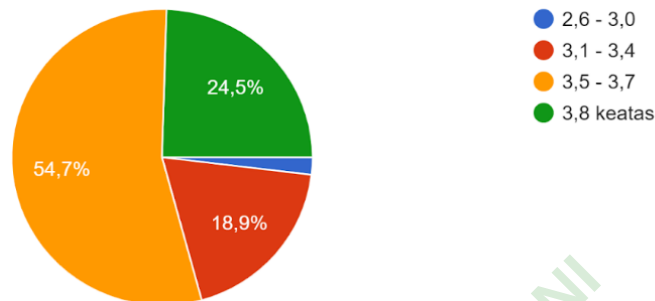
Gambar 4.4 Persentase Media Sosial

Kesimpulan dari Gambar 4.4 bahwa kuesioner yang disebar pada penggunaan media sosial yang sering digunakan mahasiswa yaitu WhatsApp dengan persentase 98,1%, Instagram dengan persentase 82,2%, Tiktok dengan persentase 72%, Twitter dengan persentase 38,3%, dan Facebook dengan persentase 43,9%. Media sosial sering digunakan yaitu WhatsApp, Instagram, Tiktok, Facebook, dan Twitter.

4.2.6 Nilai IPK

Nilai IPK mahasiswa FTTI) di Unjaya yang telah mengisi kuesioner atau angket. Data diperoleh dengan mengajukan ke pihak BAAK (Bagian Administrasi Akademik dan Kemahasiswaan) FTTI.

IPK
106 jawaban



Gambar 4.5 Persentase Nilai IPK

Kesimpulan dari Gambar 4.5 bahwa nilai IPK FTTI yaitu IPK 2,6 – 3,0 diperoleh sebanyak 1,9%, IPK 3,1 – 3,4 diperoleh sebanyak 18,9%, nilai IPK 3,5 – 3,7 diperoleh sebanyak 54,7%, dan nilai IPK 3,8 keatas diperoleh sebanyak 24,5%.

4.3 TAHAPAN PENGUMPULAN DATA

Pengambilan data yang peroleh dijalankan menggunakan jupyter notebook menggunakan Bahasa pemrograman python. Data yang didapatkan dari kuesioner berupa identitas responden berupa nama lengkap, NIM, jenis kelamin, program studi tahun angkatan dan juga 13 hasil jawaban responden dan nilai IPK berupa SKS dan nilai IPK dari 4 program studi.

4.3.1 Pre-processing Data

Langkah dalam *preprocessing* data untuk mendapatkan hasil algoritma K-Means:

1. Memasukan *Libraries*

Menjalankan perintah algoritma K-Means pada jupyter notebook digunakan *library* seperti *numpy*, *matplotlib*, *pandas*, *seaborn*, *clustering* K-Means, *preprocessing*, *metrik*. Kode yang digunakan dalam *libraries* dapat dilihat pada kode:

```
# import package
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sns
from sklearn.cluster import KMeans
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import silhouette_score,
davies_bouldin_score
```

2. Memasukan dataset

Dataset dalam penelitian ini digunakan berupa file *xlsx* yang akan di *import* menggunakan *library pandas* untuk menampilkan data. Kode yang digunakan dalam dataset dapat dilihat pada kode berikut ini:

```
# open dataset
data = pd.read_excel('Data Pengaruh media sosial.xlsx')
# menampilkan data
data.head()
```

Pemanggilan data yang berfungsi untuk menyimpan data file yang dapat ditampilkan berupa file excel ditunjukkan di Gambar 4.6.

	Timestamp	Nama Lengkap	NIM	Jenis Kelamin	Program Studi	Tahun Angkatan	SKS	Nilai IPK	Sosial Media yang digunakan	1. Saya menggunakan media sosial setiap hari	4. Saya mengakses media sosial sebelum kegiatan perkuliahan di kelas	5. Saya mengakses media sosial pada waktu istirahat	6. Saya menggunakan media sosial sebagai media untuk memperoleh informasi terbaru
0	2024-05-15 18:39:37.576	Junita Alviyanti	202102035	Perempuan	Informatika	2020	142	3.70	WhatsApp, Instagram, Tiktok	S ...	S	S	SS
1	2024-05-15 18:53:47.300	Feri Kristianto	202102031	Laki - Laki	Informatika	2020	142	3.67	WhatsApp, Instagram, Facebook	SS ...	S	SS	S
2	2024-05-15 18:55:12.927	Muhammad Azwar	202102039	Laki - Laki	Informatika	2020	148	3.67	WhatsApp, Instagram, Tiktok	SS ...	N	S	SS
3	2024-05-15 18:57:18.049	Ario Bara Riski	202102004	Laki - Laki	Informatika	2020	142	3.83	WhatsApp, Facebook	TS ...	TS	N	N
4	2024-05-15 19:10:12.050	Sulfa Nurohid	202102019	Laki - Laki	Informatika	2020	144	3.77	WhatsApp, Instagram, Tiktok, Facebook	SS ...	TS	S	SS

5 rows x 22 columns

Gambar 4.6 Menampilkan dataset

3. Menampilkan tipe data dari semua kolom yang ada dalam DataFrame

Untuk memastikan variabel pada tipe data di setiap kolom yang DataFrame termasuk jumlah non-null untuk mendapatkan gambaran struktur dataset. Kode yang digunakan dalam dataset dapat dilihat pada kode berikut ini:

```
# Menampilkan daftar semua kolom yang ada dalam dataset
data.columns.tolist()
```

Hasil dari proses tipe data dari semua kolom yang ada dalam DataFrame ditunjukkan di Gambar 4.7.



Column	Dtype	Non-Null Count
0 Timestamp	datetime64[ns]	107 non-null
1 Nama Lengkap	object	107 non-null
2 NIM	int64	107 non-null
3 Jenis Kelamin	object	107 non-null
4 Program Studi	object	107 non-null
5 Tahun Angkatan	int64	107 non-null
6 SKS	int64	107 non-null
7 Nilai IPK	float64	107 non-null
8 Sosial Media yang digunakan	object	107 non-null
9 1. Saya menggunakan media sosial setiap hari	object	107 non-null
10 2. Saya mengakses media sosial di rumah lebih dari 1 Jam/hari	object	107 non-null
11 3. Saya mengakses media sosial saat berada di kampus	object	107 non-null
12 4. Saya mengakses media sosial sebelum kegiatan perkuliahan di kelas	object	107 non-null
13 5. Saya mengakses media sosial pada waktu istirahat	object	107 non-null
14 6. Saya menggunakan media sosial sebagai media untuk memperoleh informasi terbaru	object	107 non-null
15 7. Saya menggunakan media sosial sebagai sarana untuk melakukan diskusi materi perkuliahan	object	107 non-null
16 8. Saya lebih sering mengakses media sosial dibandingkan belajar	object	107 non-null
17 9. Ketika saya merasakan bosan di dalam proses pembelajaran di kelas, saya mengakses media sosial	object	107 non-null
18 10. Saya menggunakan media sosial sebagai motivasi giat belajar	object	107 non-null
19 11. Penggunaan media sosial dapat mempermudah saya untuk lebih fokus kepada mata kuliah yang diberikan dosen	object	107 non-null
20 12. Penggunaan media sosial dapat memberikan saya pengaruh yang besar dalam nilai akademik	object	107 non-null
21 13. Penggunaan media sosial berpengaruh positif dalam prestasi belajar saya	object	107 non-null

dtypes: datetime64[ns](1), float64(1), int64(3), object(17)
memory usage: 18.5+ KB

Gambar 4.7 Menampilkan hasil tipe data di DataFrame

4. Memilih kolom yang relevan untuk proses analisis

Memilih kolom yang relevan berupa hasil kuesioner dan nilai IPK.

Kode yang digunakan dalam dataset dapat ditunjukkan sebagai berikut:

```
# Kolom yang relevan untuk analisis
cols_to_keep = [
    'Nilai IPK',
    '1. Saya menggunakan media sosial setiap hari',
    '2. Saya mengakses media sosial di rumah lebih dari 1
      Jam/hari',
    '3. Saya mengakses media sosial saat berada di kampus',
    '4. Saya mengakses media sosial sebelum kegiatan
      Perkuliahan di kelas',
    '5. Saya mengakses media sosial pada waktu istirahat',
    '6. Saya menggunakan media sosial sebagai media untuk
      Memperoleh informasi terbaru',
    '7. Saya menggunakan media sosial sebagai sarana untuk
      melakukan diskusi materi perkuliahan',
    '8. Saya lebih sering mengakses media sosial
      Dibandingkan belajar',
    '9. Ketika saya merasakan bosan di dalam proses
      pembelajaran di kelas, saya mengakses media
      sosial',
    '10. Saya menggunakan media sosial sebagai motivasi
      giat belajar',
    '11. Penggunaan media sosial dapat mempermudah saya
      untuk lebih fokus kepada mata kuliah yang
      diberikan dosen',
    '12. Penggunaan media sosial dapat memberikan saya
      pengaruh yang besar dalam nilai akademik',
    '13. Penggunaan media sosial berpengaruh positif dalam
      prestasi belajar saya'
]
data = data[cols_to_keep]
data
```

Hasil data yang akan diproses dalam processing berupa nilai IPK dengan jawaban kuesioner ditunjukkan di Gambar 4.8.

	Nilai IPK	1. Saya menggunakan media sosial setiap hari	2. Saya mengakses media sosial di rumah lebih dari 1 Jam/hari	3. Saya mengakses media sosial saat berada di kampus	4. Saya mengakses media sosial sebelum kegiatan perkuliahan di kelas	5. Saya mengakses media sosial pada waktu istirahat	6. Saya menggunakan media sosial sebagai media untuk memperoleh informasi terbaru	7. Saya menggunakan media sosial sebagai sarana untuk melakukan diskusi materi perkuliahan	8. Saya lebih sering mengakses media sosial dibandingkan belajar	9. Ketika saya merasakan bosan di dalam proses pembelajaran di kelas, saya mengakses media sosial	10. Saya menggunakan media sosial sebagai motivasi giat belajar	Pengg media memper saya lebih kepada: kulial dib
0	3.70	S	SS	S	S	S	SS	SS	N	N	N	
1	3.67	SS	SS	S	S	SS	S	KS	S	S	N	
2	3.67	SS	SS	SS	N	S	SS	N	S	N	S	
3	3.83	TS	TS	TS	TS	N	N	S	N	N	N	
4	3.77	SS	SS	S	TS	S	SS	SS	N	S	S	
...	...	...	...	...	...	...	...	...	...	...	...	
102	3.45	S	S	S	S	S	S	S	S	S	KS	
103	3.84	SS	SS	S	S	S	SS	N	S	KS	N	
104	3.83	SS	SS	S	S	S	S	S	KS	TS	N	
105	3.66	SS	S	S	S	S	S	S	N	N	S	
106	3.62	SS	S	N	KS	S	S	SS	N	N	N	

107 rows x 14 columns

Activate Window
Go to Settings to activate Windows Defender

Gambar 4.8 Menampilkan hasil Pemilih Kolom

5. Mengubah skala likert

Mengubah nilai skala *likert* dari huruf ke numerik untuk memudahkan *algoritma* untuk menganalisis *clustering* di data numerik, kode ditunjukkan sebagai berikut:

```
# konversi string ke integer
likert_scale = {'SS': 1, 'S': 2, 'N': 3, 'KS': 4, 'TS': 5}
data_x = data.replace(likert_scale)
data_x
```

Hasil dari pengubahan data menjadi numerik untuk pertanyaan di kuesioner yang disebar ke mahasiswa, dapat ditunjukkan di Gambar 4.9.

Nilai IPK	1. Saya menggunakan media sosial setiap hari	2. Saya mengakses media sosial di rumah lebih dari 1 Jam/hari	3. Saya mengakses media sosial saat berada di kampus	4. Saya mengakses media sosial sebelum kegiatan perkuliahan di kelas	5. Saya mengakses media sosial pada waktu istirahat	6. Saya menggunakan media sosial sebagai media untuk memperoleh informasi terbaru	7. Saya menggunakan media sosial sebagai sarana untuk melakukan diskusi materi perkuliahan	8. Saya lebih sering mengakses media sosial dibandingkan belajar	9. Ketika saya merasakan bosan di dalam proses pembelajaran di kelas, saya mengakses media sosial	10. Saya menggunakan media sosial sebagai motivasi giat belajar	Pengg media memper saya lebih kepada kulial dib
0	3.70	2	1	2	2	2	1	1	3	3	3
1	3.67	1	1	2	2	1	2	4	2	2	3
2	3.67	1	1	1	3	2	1	3	2	3	2
3	3.83	5	5	5	5	3	3	2	3	3	3
4	3.77	1	1	2	5	2	1	1	3	2	2
...	...	...	...	...	...	...	...	...	...	...	...
102	3.45	2	2	2	2	2	2	2	2	2	4
103	3.84	1	1	2	2	2	1	3	2	4	3
104	3.83	1	1	2	2	2	2	2	4	5	3
105	3.66	1	2	2	2	2	2	2	3	3	2
106	3.62	1	2	3	4	2	2	1	3	3	3

107 rows x 14 columns

Gambar 4.9 Menampilkan Hasil Mengubah Skala Likert

6. Mengubah variabel dari DataFrame ke Array

Pada penelitian ini data yang dibutuhkan berupa hasil kuesioner dan nilai IPK yang berbentuk data numerik. Agar dataset dapat digunakan untuk menghitung nilai K dalam bentuk *array* ditunjukkan pada kode berikut:

```
# mengubah variabel yang sebelumnya berbentuk data frame
menjadi array
x_array = np.array(data_x)
x_array
```

Hasil dari memilih kolom – kolom yang penting agar dapat dikelola oleh algoritma k-means dapat dilihat pada Gambar 4.10.

```
Out[7]: array([[3.7 , 2. , 1. , ..., 5. , 4. , 3. ],
               [3.67, 1. , 1. , ..., 4. , 3. , 2. ],
               [3.67, 1. , 1. , ..., 5. , 2. , 2. ],
               ...,
               [3.83, 1. , 1. , ..., 3. , 4. , 4. ],
               [3.66, 1. , 2. , ..., 3. , 2. , 3. ],
               [3.62, 1. , 2. , ..., 3. , 3. , 4. ]])
```

Gambar 4.10 Hasil Mengubah variabel ke Array

7. Mengecek kolom yang telah diubah menjadi numerik

Pengecekan diperlukan untuk memastikan semua kolom di DataFrame dapat diubah ke tipe data numerik agar tetap konsisten dalam persiapan analisis data dalam clustering ditunjukkan di kode berikut:

```
# informasi kolom yang telah diubah menjadi numerik
data_x = data_x.apply(pd.to_numeric, errors='coerce')
print(data_x.dtypes)
```

Hasilnya dapat ditunjukkan pada Gambar 4.11.

Nilai IPK		
1. Saya menggunakan media sosial setiap hari		float64
2. Saya mengakses media sosial di rumah lebih dari 1 Jam/hari		int64
3. Saya mengakses media sosial saat berada di kampus		int64
4. Saya mengakses media sosial sebelum kegiatan perkuliahan di kelas		int64
5. Saya mengakses media sosial pada waktu istirahat		int64
6. Saya menggunakan media sosial sebagai media untuk memperoleh informasi terbaru		int64
7. Saya menggunakan media sosial sebagai sarana untuk melakukan diskusi materi perkuliahan		int64
8. Saya lebih sering mengakses media sosial dibandingkan belajar		int64
9. Ketika saya merasakan bosan di dalam proses pembelajaran di kelas, saya mengakses media sosial		int64
10. Saya menggunakan media sosial sebagai motivasi giat belajar		int64
11. Penggunaan media sosial dapat mempermudah saya untuk lebih fokus kepada mata kuliah yang diberikan dosen		int64
12. Penggunaan media sosial dapat memberikan saya pengaruh yang besar dalam nilai akademik		int64
13. Penggunaan media sosial berpengaruh positif dalam prestasi belajar saya		int64
dtype: object		

Gambar 4.11 Kolom Numerik

8. Standarisasi data

standarisasi atau *normalisasi* tujuannya untuk scaling data agar tetap seimbang dalam perhitungan jarak antara 0-1. Standarisasi dilakukan apabila angka yang besar di variabel, sebelum melakukan perhitungan di analisis *clustering* yang lebih akurat ditunjukkan pada kode berikut:

```
# proses standarisasi agar scalling data berkisar 0-1
scaler = MinMaxScaler()
x_scaled = scaler.fit_transform(x_array)
x_scaled
```

Hasil dari proses standarisasi ditunjukkan di Gambar 4.12.

```
Out[9]: array([[0.82014388, 0.25, 0., ..., 1., 0.75,
0.5],
[0.79856115, 0., 0., ..., 0.75, 0.5,
0.25],
[0.79856115, 0., 0., ..., 1., 0.25,
0.25],
...,
[0.91366906, 0., 0., ..., 0.5, 0.75,
0.75],
[0.79136691, 0., 0.25, ..., 0.5, 0.25,
0.5],
[0.76258993, 0., 0.25, ..., 0.5, 0.5,
0.75]])
```

Gambar 4.12 Menampilkan Standarisasi Data

4.3.2 Clustering K-Means

Langkah dalam *Clustering* K-Means data untuk mendapatkan hasil yang akan diolah menggunakan algoritma K-Means

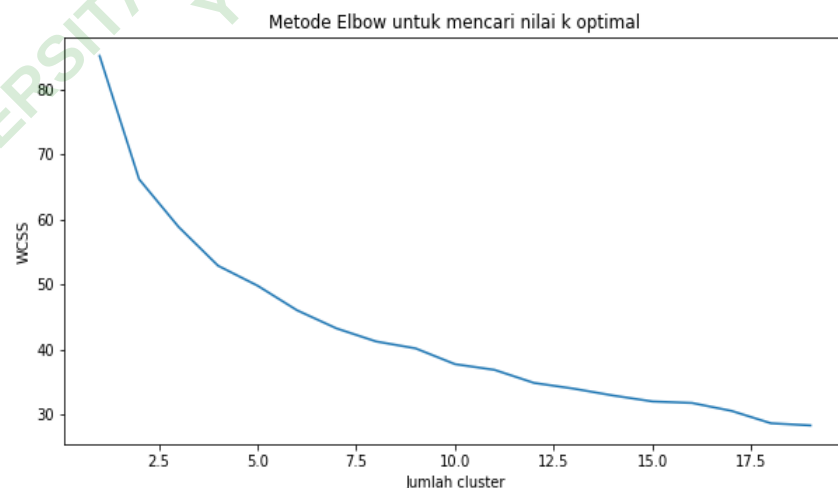
1. Menentukan nilai K dengan metode Elbow

Selain penggunaan metode *Silhouette Score*, dan *Davies-Bouldin Index*, metode Elbow untuk mencari nilai optimal K dengan WCSS yang digunakan untuk evaluasi hasil cluster ditunjukkan di kode sebagai berikut:

```
# Menentukan jumlah klaster yang optimal menggunakan metode elbow
wcss = []
for i in range(1, 20):
    kmeans = KMeans(n_clusters=i, init='k-means++',
                    random_state=42)
    kmeans.fit(x_scaled)
    wcss.append(kmeans.inerti)

plt.figure(figsize=(10, 5))
plt.plot(range(1, 20), wcss)
plt.title('Metode Elbow untuk mencari nilai k optimal')
plt.xlabel('Jumlah cluster')
plt.ylabel('WCSS')
plt.show()
```

Hasil dari perhitungan metode Elbow untuk menentukan nilai k di dapatkan dilihat pada Gambar 4.13.



Gambar 4.13 Metode Elbow

Dari Gambar 4.13, penggunaan Metode Elbow untuk optimal K digunakan untuk cluster terbaik semakin kecil nilainya WCSS semakin baik memiliki sumbu X menyatakan jumlah cluster dan sumbu Y menyatakan WCSS. Dengan menggunakan metode Elbow mendapatkan 3 nilai K optimal karena nilai K=3 ke atas mengalami penurunan yang tidak signifikan dan membentuk suku pada grafiknya.

2. Menentukan jumlah cluster

Berdasarkan nilai yang didapatkan maka pengelompokan pengaruh media sosial yang sering digunakan mahasiswa FTTI di Unjaya dapat dikelompokkan ke dalam 3 kelompok dapat ditunjukkan di kode sebagai berikut:

```
# menentukan dan mengkonfigurasi fungsi kmeans
optimal_k = 3
kmeans = KMeans(n_clusters=optimal_k, init='k-means++',
random_state=123)
# menentukan kluster dari data
kmeans.fit(x_scaled)
```

3. Mencari nilai pusat dari tiap cluster

Setelah cluster data diketahui, maka mencari nilai tiap pusat menggunakan *centroid* untuk mengetahui rata-rata di titik semua data dalam cluster, kode ditunjukkan pada kode sebagai berikut:

```
# mencari nilai pusat dari masing masing cluster
print(kmeans.cluster_centers_)
```

Hasil dari nilai pusat di setiap cluster ditunjukkan di Tabel 4.5.

Tabel 4.5 Nilai Pusat Cluster

Pertanyaan Kuesioner	Cluster 0	Cluster 1	Cluster 2
1.Nilai IPK.	0,810	0,747	0,773
2. Saya menggunakan media sosial setiap hari.	0,833	0,042	0,078
3.Saya mengakses media sosial di rumah lebih dari 1 Jam/hari.	0,833	0,058	0,109

4. Saya mengakses media sosial saat berada di kampus.	0,833	0,159	0,285
5. Saya mengakses media sosial sebelum kegiatan Perkuliahan di kelas.	0,677	0,154	0,342
6. Saya mengakses media sosial pada waktu istirahat.	0,583	0,063	0,228
7. Saya menggunakan media sosial sebagai media untuk Memperoleh informasi terbaru.	0,889	0,035	0,157
8. Saya menggunakan media sosial sebagai sarana untuk melakukan diskusi materi perkuliahan.	0,556	0,099	0,385
9. Saya lebih sering mengakses media sosial dibandingkan belajar.	0,750	0,319	0,460
10. Ketika saya merasakan bosan di dalam proses pembelajaran di kelas, saya mengakses media sosial.	0,750	0,255	0,482
11. Saya menggunakan media sosial sebagai motivasi giat belajar.	0,833	0,196	0,451
12. Penggunaan media sosial dapat mempermudah saya untuk lebih fokus kepada mata kuliah yang diberikan dosen.	0,583	0,260	0,644
13. Penggunaan media sosial dapat memberikan saya pengaruh yang besar dalam nilai akademik.	0,750	0,228	0,508
14. Penggunaan media sosial berpengaruh positif dalam prestasi belajar saya.	0,750	0,202	0,508

4. Menampilkan hasil cluster

Setelah proses dalam clustering dengan data diubah ke numerik, sehingga data bisa olah untuk mencari jumlah tiap kluster untuk dikelompokkan dalam bentuk visualisasi data. Sebelum menampilkan hasil cluster data terlebih dahulu diberi label yang diisi nomor *cluster*, ditunjukkan di kode sebagai berikut:

```
# menampilkan hasil kluster
print(kmeans.labels_)
# menambahkan kolom cluster dalam dataframe data
data_x['cluster'] = kmeans.labels_
data_x.head(10)
```

Hasil dari perlabelan data cluster dengan menggunakan array dapat ditunjukkan pada Gambar 4.14.

Out[15]:

3. Saya mengakses media sosial saat berada di kampus	4. Saya mengakses media sosial sebelum kegiatan perkuliahan di kelas	5. Saya mengakses media sosial pada waktu istirahat	6. Saya menggunakan media sosial sebagai media untuk memperoleh informasi terbaru	7. Saya menggunakan media sosial sebagai sarana untuk melakukan diskusi materi perkuliahan	8. Saya lebih sering mengakses media sosial dibandingkan belajar	9. Ketika saya merasakan bosan di dalam proses pembelajaran di kelas, saya mengakses media sosial	10. Saya menggunakan media sosial sebagai motivasi giat belajar	11. Penggunaan media sosial dapat mempermudah saya untuk lebih fokus kepada mata kuliah yang diberikan dosen	12. Penggunaan media sosial dapat memberikan saya pengaruh yang besar dalam nilai akademik	13. Penggunaan media sosial berpengaruh positif dalam prestasi belajar saya	cluster
2	2	2	1	1	3	3	3	5	4	3	2
2	2	1	2	4	2	2	3	4	3	2	2
1	3	2	1	3	2	3	2	5	2	2	2
5	5	3	3	2	3	3	3	3	3	3	0
2	5	2	1	1	3	2	2	4	4	4	2
2	4	1	1	2	3	3	2	4	2	3	2
2	2	2	1	1	3	3	2	2	3	2	1
1	1	1	1	2	3	2	2	2	2	2	1
3	5	3	2	2	3	5	3	5	3	3	2
1	1	1	1	1	2	2	2	1	2	2	1

Gambar 4.14 Menampilkan hasil label cluster

5. Menghitung jumlah tiap cluster

Tiap cluster memiliki jumlah nilai untuk melihat perbedaan elemen cluster yang terbentuk ditunjukkan dikode sebagai berikut:

```
# menghitung jumlah cluster
for i in range(optimal_k):
    print(f"Jumlah di cluster {i}: {(data_x['cluster'] == i).sum()}")
```

Hasil pengelompokkan pengaruh media sosial yang digunakan mahasiswa FTTI memiliki jumlah cluster 0 sejumlah 3 anggota, cluster 1 sejumlah 47 anggota, dan cluster 2 sejumlah 57 anggota.

4.3.3 Evaluasi

Evaluasi digunakan untuk menghitung jumlah cluster yang terbentuk proses clustering K-Means dengan nilai dari suatu hal dalam pengukuran kegiatan dengan menganalisis dari hasil visualisasi menggunakan algoritma K-Means menggunakan

metrik Silhouette Score dan *Davies-Bouldin Index (DBI)* ditunjukkan dikode sebagai berikut:

```
# Evaluasi kinerja clustering
silhouette_avg = silhouette_score(x_scaled,
kmeans.fit_predict(x_scaled))
db_score = davies_bouldin_score(x_scaled,
kmeans.fit_predict(x_scaled))
ch_score = calinski_harabasz_score(x_scaled, kmeans.labels_)

print(f'Silhouette Score: {silhouette_avg}')
print(f'Davies-Bouldin Index: {db_score}')
```

Hasil perbandingan evaluasi *metrik Silhouette Score* memiliki nilai 0,196 menunjukkan bahwa *clustering* berada di objek yang berdekatan antara batas *cluster* yang salah sedangkan dengan menggunakan metode *Elbow* memiliki nilai optimal 3 karena nilai *inertia* mengalami penurunan yang signifikan sehingga data dikatakan cluster cenderung baik. Visualisasi dengan metrik Davies-Bouldin Index dengan nilai 1,490 memiliki nilainya antara di atas 1 dapat dikatakan cukup baik karena nilai DBI sangat baik.

4.3.4 Hasil

Visualisasi menggunakan plot data cluster yang jelas untuk mengetahui penyebaran dan pengelompokkan data dari pengaruh media sosial terhadap mahasiswa FTTH ditunjukkan di kode sebagai berikut:

```
# Centroids
centroids = kmeans.cluster_centers_

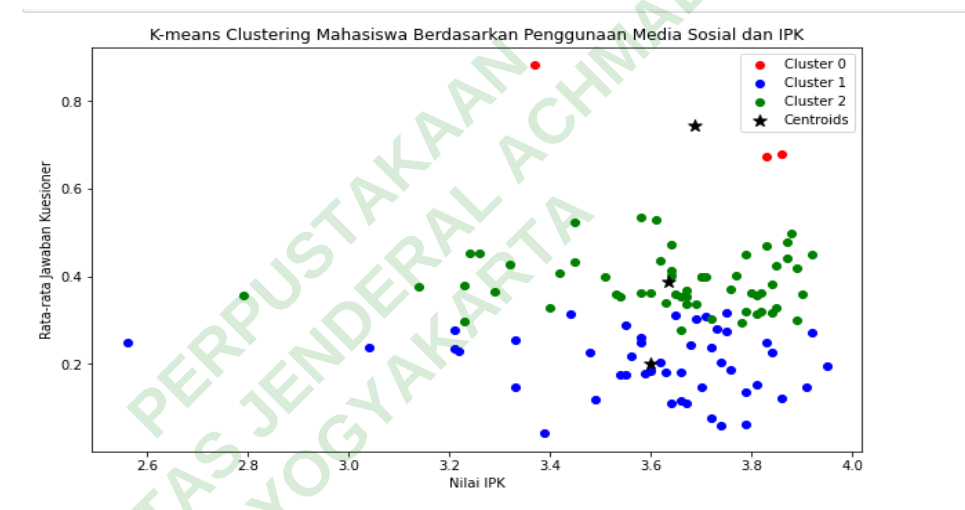
# Visualisasi hasil clustering berdasarkan IPK
plt.figure(figsize=(10, 6))

# Plotting clusters
colors = ['red', 'blue', 'green']
for i in range(3):
    cluster_data = data[data_x['cluster'] == i]
    # Calculate the mean of scaled features for each student in
    the cluster
    mean_scaled_featur x_scaled[data_x['cluster'] ==
    i].mean(axis=1)
    plt.scatter(cluster_data['Nilai IPK'], mean_scaled_features,
    color=colors[i], label=f'Cluster {i}')
```

```
# Plotting centroids
centroid_ipk = scaler.inverse_transform(centroids)[: , 0]
centroid_mean_features = centroids.mean(axis=1)
plt.scatter(centroid_ipk, centroid_mean_features, color='black',
            marker='*', s=100, label='Centroids')

plt.xlabel('Nilai IPK')
plt.ylabel('Rata-rata Jawaban Kuesioner')
plt.title('K-means Clustering Mahasiswa Berdasarkan Penggunaan
Media Sosial dan IPK')
plt.legend()
plt.show()
```

Visualisasi data untuk menampilkan penyebaran data clustering ditunjukkan pada Gambar 4.15.



Gambar 4.15 Visualisasi Metrik

Dari Gambar 4.15 untuk visualisasi hasil *clustering* dengan rata-rata kuesioner didapatkan dari mahasiswa yang mengisi jawaban dengan SS, S, N, KS, TS dengan menggunakan plot yang setiap kelompok diwakilkan dengan warna yang berbeda yaitu:

- Warna titik merah ● menunjukkan cluster 0 berjumlah 3
- Warna titik biru ● menunjukkan cluster 1 berjumlah 43
- Warna titik hijau ● menunjukkan cluster 2 berjumlah 57
- Tanda bintang ★ menunjukkan titik tengah *centroid*

Pembagian pengelompokkan data dengan pengaruh yang ditimbulkan penggunaan media sosial terhadap prestasi mahasiswa FTTI di Unjaya yaitu:

1. Cluster 0

- Penggunaan Media Sosial: rendah
- Rata-rata IPK: 3,7

Hasilnya kurang berpengaruh karena mahasiswa jarang menggunakan media sosial apalagi dengan nilai IPK tinggi sekitar 3,7

2. Cluster 1

- Penggunaan media sosial: tinggi
- Rata-rata IPK: 3,6

Hasil berpengaruh karena mahasiswa sering menggunakan media sosial ditandai dengan nilai IPK tinggi sekitar 3,6

3. Cluster 2

- Penggunaan media sosial: sedang
- Rata - rata IPK: 3,65

Hasil sangat berpengaruh karena mahasiswa lebih sering menggunakan media sosial nilai IPK cenderung tinggi

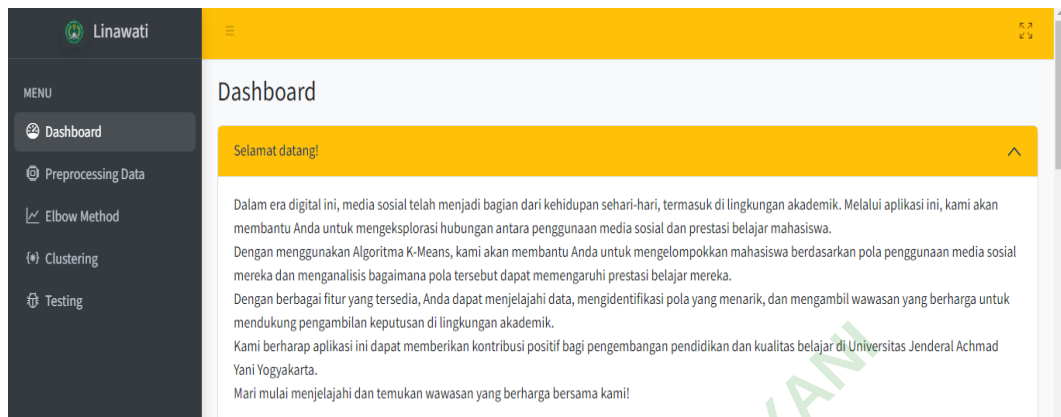
Dengan hasil visualisasi data dapat disimpulkan bahwa media sosial memberikan pengaruh terhadap prestasi belajar pada mahasiswa FTTI, hal ini ditandai dengan seringnya penggunaan media sosial dengan nilai IPK yang cukup tinggi sekitar 3,6 ke atas. Hasil dari clustering akan diimplementasikan dalam website untuk menampilkan data.

4.4 IMPLEMENTASI DESAIN ANTARMUKA

4.4.1 Halaman Dashboard

Pada menu dashboard menampilkan penjelasan awal dari sistem yang dibuat, dataset dari hasil pengisian kuesioner, distribusi jenis kelamin mahasiswa, dan distribusi program studi agar memudahkan pembaca melakukan pengujian dari sistem yang dibuat. Di bagian kiri ada menu navbar yang menghubungkan dengan fitur yang lain.

1. Tampilan home dapat ditampilkan pada Gambar 4.16.



Gambar 4.16 Halaman Dashboard Awal

2. Pada menu dashboard menampilkan tabel kuesioner mahasiswa yang didapatkan dari kuesioner mahasiswa dan nilai IPK ditampilkan pada Gambar 4.17.

Timestamp	Nama Lengkap	NIM	Jenis Kelamin	Program Studi	Tahun Angkatan	SKS	Nilai IPK	Sosial Media yang digunakan	1. Saya menggunakan media sosial setiap hari	2. Saya mengakses media sosial di rumah lebih dari 1 Jam/hari	3. Saya mengakses media sosial saat berada di kampus
2024-05-15 18:39:37.576	Junita Alviyanti	202102035	Perempuan	Informatika	2020	142	3.70	WhatsApp, Instagram, Tiktok	S	SS	S
2024-05-15 18:53:47.300	Feri Kristianto	202102031	Laki - Laki	Informatika	2020	142	3.67	WhatsApp, Instagram, Facebook	SS	SS	S
2024-05-15 18:55:12.927	Muhammad Azwar	202102039	Laki - Laki	Informatika	2020	148	3.67	WhatsApp, Instagram, Tiktok	SS	SS	SS

Gambar 4.17 Tampilan hasil kuesioner dan IPK

3. Pada menu dashboard menampilkan juga salah satu gambaran responden yang mengisi kuesioner berupa jenis kelamin dan prodi ditampilkan pada Gambar 4.18.

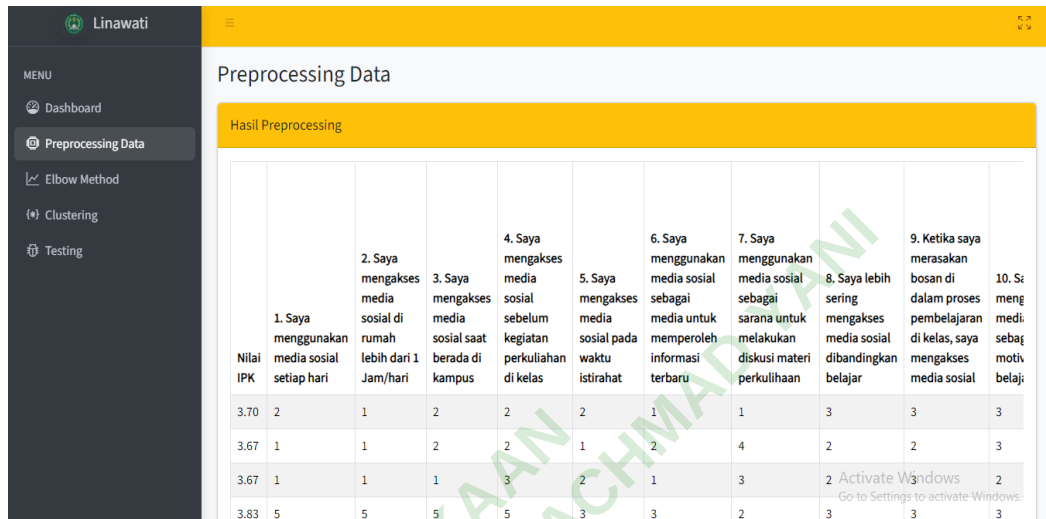


Gambar 4.18 Tampilan Gambaran Responden

4.4.2 Preprocessing Data

Halaman *preprocessing* data menampilkan data yang sudah dibersihkan dan diubah ke data numerik dan juga menampilkan tipe data agar dapat dilakukan proses clustering.

1. Data kuesioner dan nilai IPK yang dilakukan *preprocessing* akan dipilih kolom yang dibutuhkan dan diubah ke data numerik ditunjukkan di Gambar 4.19.



Nilai IPK	1. Saya menggunakan media sosial setiap hari	2. Saya mengakses media sosial di rumah lebih dari 1 Jam/hari	3. Saya mengakses media sosial saat berada di kampus	4. Saya mengakses media sosial sebelum kegiatan perkuliahan di kelas	5. Saya mengakses media sosial pada waktu istirahat	6. Saya menggunakan media sosial sebagai media untuk memperoleh informasi terbaru	7. Saya menggunakan media sosial sebagai sarana untuk melakukan diskusi materi perkuliahan	8. Saya lebih sering mengakses media sosial dibandingkan belajar	9. Ketika saya merasakan bosan di dalam proses pembelajaran di kelas, saya mengakses media sosial	10. Saya menggunakan media sosial sebagai motivasi belajar
3.70	2	1	2	2	2	1	1	3	3	3
3.67	1	1	2	2	1	2	4	2	2	3
3.67	1	1	1	3	2	1	3	2	2	2
3.83	5	5	5	5	3	3	2	3	3	3

Gambar 4.19 Tampilan Hasil Preprocessing Data

2. Pada menu home menampilkan informasi *preprocessing* berupa tipe data untuk mengecek setiap kolom ditampilkan di gambar 4.20.

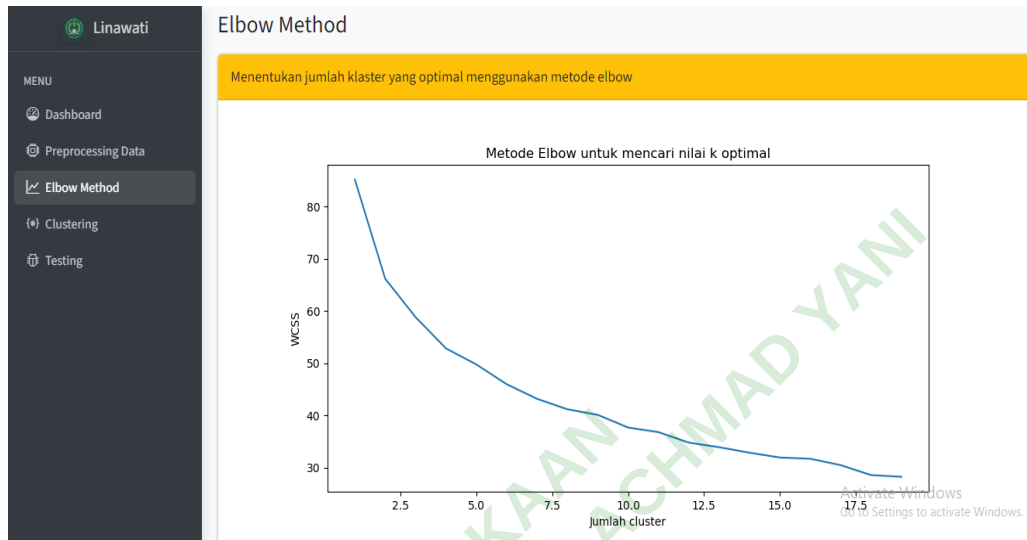


Kolom	Nilai Hilang	Tipe Data
Nilai IPK	0	float64
1. Saya menggunakan media sosial setiap hari	0	int64
2. Saya mengakses media sosial di rumah lebih dari 1 Jam/hari	0	int64
3. Saya mengakses media sosial saat berada di kampus	0	int64
4. Saya mengakses media sosial sebelum kegiatan perkuliahan di kelas	0	int64
5. Saya mengakses media sosial pada waktu istirahat	0	int64
6. Saya menggunakan media sosial sebagai media untuk memperoleh informasi terbaru	0	int64
7. Saya menggunakan media sosial sebagai sarana untuk melakukan diskusi materi perkuliahan	0	int64
8. Saya lebih sering mengakses media sosial dibandingkan belajar	0	int64
9. Ketika saya merasakan bosan di dalam proses pembelajaran di kelas, saya mengakses media sosial	0	int64
10. Saya menggunakan media sosial sebagai motivasi belajar	0	int64
11. Penggunaan media sosial dapat mempermudah saya untuk lebih fokus kepada mata kuliah yang diberikan dosen	0	int64
12. Penggunaan media sosial dapat memberikan saya pengaruh yang besar dalam nilai akademik	0	int64

Gambar 4.20 Tampilan Preprocessing Data untuk Tipe Data

4.4.3 Elbow Method

Menu elbow method untuk menampilkan grafik dari optimal nilai K dalam clustering menggunakan WCSS dapat ditampilkan pada Gambar 4.21.



Gambar 4.21 Tampilan Menu Elbow Method

4.4.4 Clustering

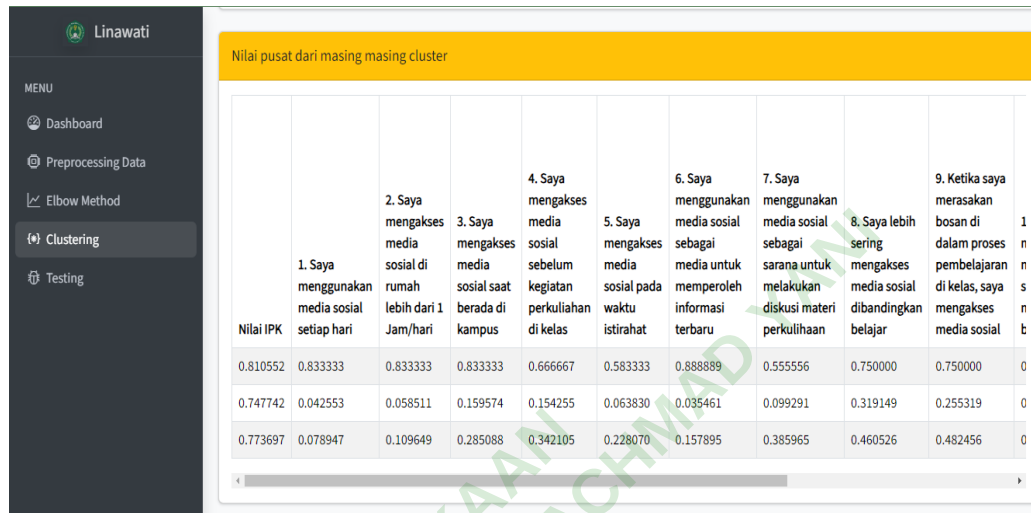
Pada menu halaman clustering ini menampilkan jumlah nilai K yang optimal, nilai *cluster* dari masing-masing *cluster*, hasil dari proses *cluster*, jumlah data di setiap cluster, dan visualisasi dari penyebaran kelompok clustering.

1. Pada tampilan pembuatan model terdapat model dari jumlah nilai K yang optimal dapat ditampilkan pada Gambar 4.22.



Gambar 4.22 Pembuatan Model K

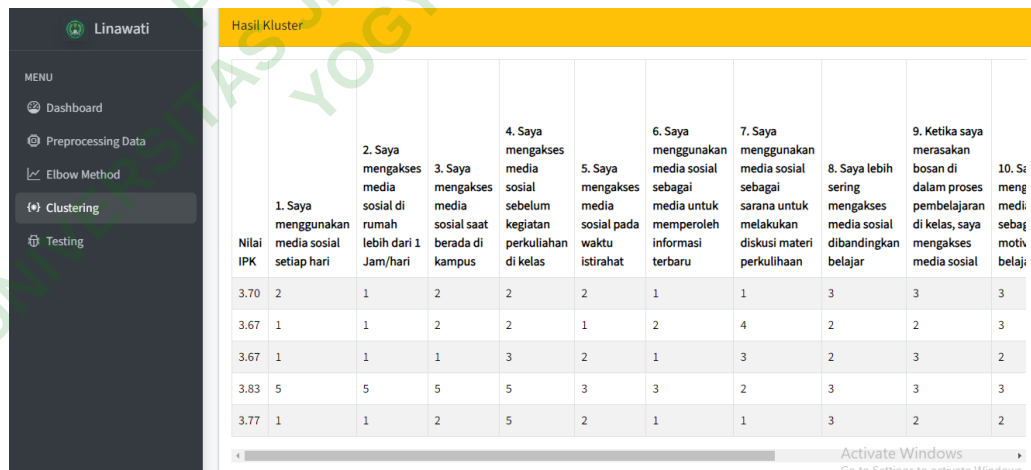
2. Tampilan di menu clustering terdapat hasil dari nilai pusat dari masing-masing clustering yang terdapat 3 pengelompokkan data ditampilkan pada Gambar 4.23.



Nilai IPK	1. Saya menggunakan media sosial setiap hari	2. Saya mengakses media sosial di rumah lebih dari 1 Jam/hari	3. Saya mengakses media sosial saat berada di kampus	4. Saya mengakses media sosial sebelum kegiatan perkuliahan di kelas	5. Saya mengakses media sosial pada waktu istirahat	6. Saya menggunakan media sosial sebagai media untuk memperoleh informasi terbaru	7. Saya menggunakan media sosial sebagai sarana untuk melakukan diskusi materi perkuliahan	8. Saya lebih sering mengakses media sosial dibandingkan belajar	9. Ketika saya merasakan bosan di dalam proses pembelajaran di kelas, saya mengakses media sosial	10. Saya merasa lebih nyaman dengan media sosial
0.810552	0.833333	0.833333	0.833333	0.666667	0.583333	0.888889	0.555556	0.750000	0.750000	0
0.747742	0.042553	0.058511	0.159574	0.154255	0.063830	0.035461	0.099291	0.319149	0.255319	0
0.773697	0.078947	0.109649	0.285088	0.342105	0.228070	0.157895	0.385965	0.460526	0.482456	0

Gambar 4.23 Tampilan Nilai Pusat dari Masing Cluster

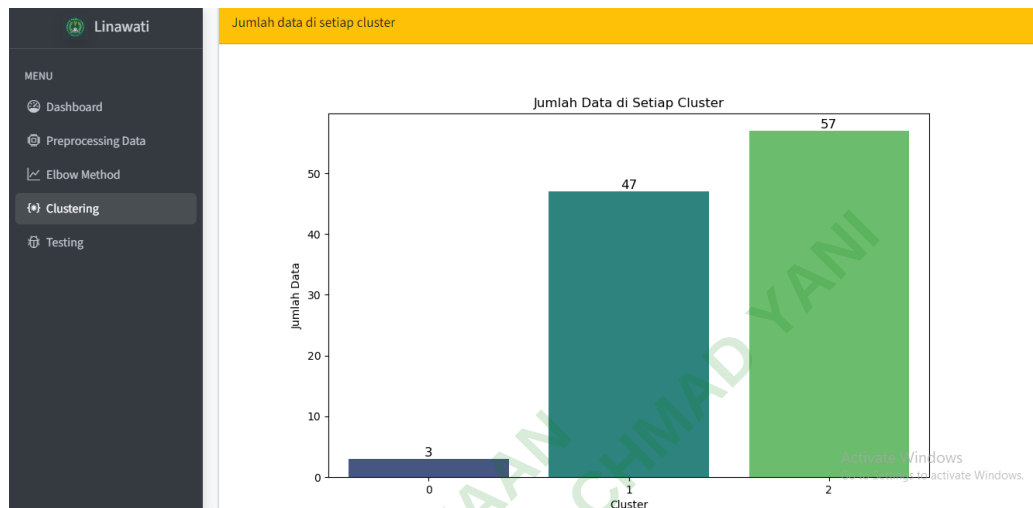
3. Tampilan di menu clustering terdapat hasil cluster. Hasil akhir dari proses clustering ditunjukkan di Gambar 4.24.



Nilai IPK	1. Saya menggunakan media sosial setiap hari	2. Saya mengakses media sosial di rumah lebih dari 1 Jam/hari	3. Saya mengakses media sosial saat berada di kampus	4. Saya mengakses media sosial sebelum kegiatan perkuliahan di kelas	5. Saya mengakses media sosial pada waktu istirahat	6. Saya menggunakan media sosial sebagai media untuk memperoleh informasi terbaru	7. Saya menggunakan media sosial sebagai sarana untuk melakukan diskusi materi perkuliahan	8. Saya lebih sering mengakses media sosial dibandingkan belajar	9. Ketika saya merasakan bosan di dalam proses pembelajaran di kelas, saya mengakses media sosial	10. Saya merasa lebih nyaman dengan media sosial
3.70	2	1	2	2	2	1	1	3	3	3
3.67	1	1	2	2	1	2	4	2	2	3
3.67	1	1	1	3	2	1	3	2	3	2
3.83	5	5	5	5	3	3	2	3	3	3
3.77	1	1	2	5	2	1	1	3	2	2

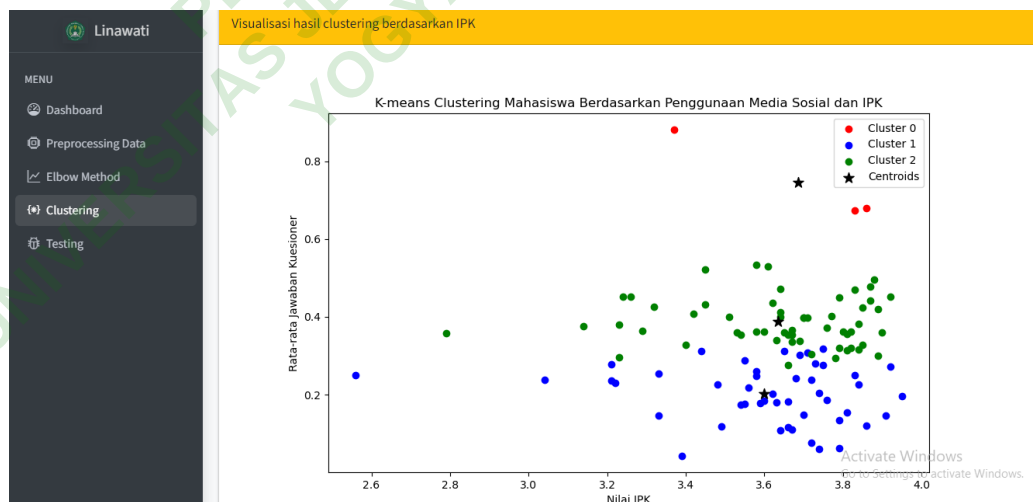
Gambar 4.24 Hasil Cluster

4. Tampilan di menu *clustering* terdapat jumlah data di setiap *cluster*, terdapat 3 *cluster* yang di gambarkan menggunakan diagram batang dan jumlah dari setiap *cluster* ditampilkan di Gambar 4.25.



Gambar 4.25 Diagram Batang Setiap Cluster

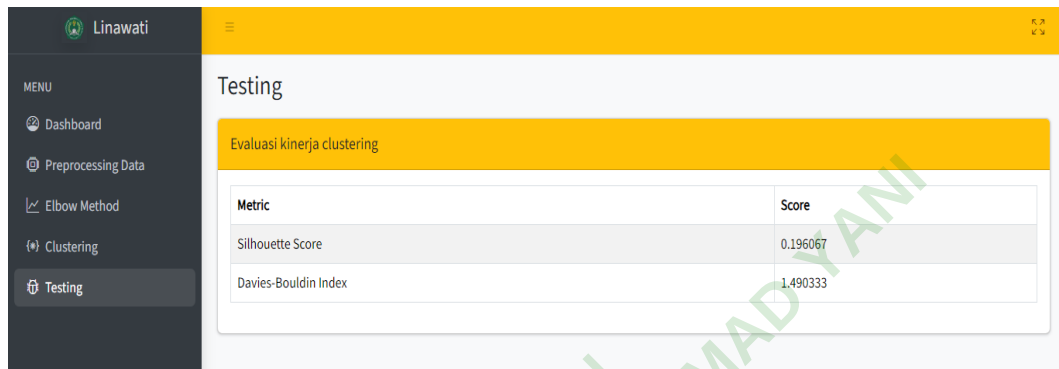
5. Tampilan di menu visualisasi *clustering* berdasarkan nilai IPK untuk menampilkan hasil sebaran 3 kelompok data ditampilkan di Gambar 4.26.



Gambar 4.26 Visualisasi Hasil Cluster

4.4.5 Testing

Pada menu *testing* menampilkan 2 pengujian dengan menggunakan 3 *metrik* yaitu menggunakan *algoritma* K-Means menggunakan *metrik* *Silhouette Score* dan *Davies-Bouldin Index* ditampilkan di Gambar 4.27.



The screenshot shows the Linawati application interface. On the left is a dark sidebar menu with options: Dashboard, Preprocessing Data, Elbow Method, Clustering, and Testing (which is highlighted). The main content area is titled 'Testing' and contains a yellow header 'Evaluasi kinerja clustering'. Below this is a table with two columns: 'Metric' and 'Score'.

Metric	Score
Silhouette Score	0.196067
Davies-Bouldin Index	1.490333

Gambar 4.27 Testing dengan metrik

