

BAB 4

PEMBAHASAN

4.1 Tahapan *Preprocessing Data*

Tahap *preprocessing data* melibatkan serangkaian langkah untuk membersihkan, menyiapkan, dan menyusun data sebelum dilakukan analisis lebih lanjut. Ini mencakup menghapus inkonsistensi, menangani data yang tidak lengkap, dan menghapus duplikasi dalam data awal. Proses data *preprocessing* pada penelitian ini terdiri dari *data selection*, *data cleaning*, dan *data transformation*. Berikut adalah penjelasan proses yang terdapat pada tahap *data preprocessing* adalah sebagai berikut:

4.1.1 *Data Selection*

Tahapan *data selection* adalah tahap penting dalam penelitian. Data berasal dari jawaban kuesioner yang disebar di wilayah pekanbaru. Atribut yang dimasukkan dalam pengumpulan data adalah gejala-gejala kecanduan internet yang dirasakan oleh responden. Kuesioner tersebut terdiri dari 20 pertanyaan di mana setiap pertanyaan menyoroti aspek kecanduan internet yang dirasakan oleh responden terdapat pada Tabel 3.1. Berikut tabel 4.1 adalah data yang digunakan dalam penelitian ini.

Tabel 4.1 Seleksi Data

No	P1	P2	P3	P4	...	P20
1	Kadang - Kadang	Tidak Pernah	Kadang - Kadang	Sangat Sering	...	Sangat Sering
2	Kadang - Kadang	Kadang - Kadang	Sering	Kadang - Kadang	...	Jarang
3	Kadang - Kadang	Kadang - Kadang	Sering	Jarang	...	Sering
4	Jarang	Jarang	Jarang	Jarang	...	Sering
5	Selalu	Sering	Kadang - Kadang	Jarang	...	Selalu
6	Kadang - Kadang	Jarang	Jarang	Kadang - Kadang	...	Kadang - Kadang
7	Sangat Sering	Sering	Sering	Kadang - Kadang	...	Jarang



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

8	Kadang - Kadang	Kadang - Kadang	Jarang	Kadang - Kadang	...	Kadang - Kadang
9	Sering	Kadang - Kadang	Sering	Kadang - Kadang	Tidak Pernah
10	Selalu	Kadang - Kadang	Kadang - Kadang	Kadang - Kadang	...	Kadang - Kadang
11	Selalu	Jarang	Sering	Jarang	...	Kadang - Kadang
12	Kadang - Kadang	Jarang	Sering	Sangat Sering	...	Kadang - Kadang
...
510	Selalu	Sering	Selalu	Selalu	...	Tidak pernah

4.1.2 Data Cleaning

Data cleaning adalah proses pembersihan data dari duplikasi dan kesalahan pengisian. Ini dilakukan untuk mempermudah proses dan hasil dari analisis *clustering* dalam penelitian ini. Pada tahap ini juga dilakukan proses pemeriksaan data ganda, apabila ada data pengisian kuesioner yang ganda maka akan dihapus. Proses pembersihan data pada penelitian ini yaitu pembersihan kesalahan pengisian. Pada penelitian ini terdapat proses *cleaning* pada data yang kosong sebelumnya 511 data menjadi 510 data karena terdapat pertanyaan yang tidak dijawab oleh responden.

4.1.3 Data Transformation

Tahap *data transformation* adalah proses mengubah format data menjadi data yang sesuai dengan kebutuhan dalam penelitian. Setelah membersihkan data dari jawaban yang tidak relevan, langkah selanjutnya adalah transformasi data. Proses ini melibatkan perubahan data menjadi bentuk numerik. Data yang awalnya berupa huruf akan dikonversi menjadi nilai sesuai dengan angka yang telah ditetapkan dalam penelitian sebelumnya, yang terdapat dalam tabel 4.2.

Tabel 4.2 Bobot Kriteria

Kuesioner	Keterangan	Bobot Nilai
TP	Tidak Pernah	0
J	Jarang	1
KK	Kadang Kadang	2
S	Sering	3
SS	Sangat Sering	4
S	Selalu	5

Transformasi data berperan penting dalam memastikan bahwa data yang digunakan untuk analisis atau pemodelan relevan dengan tujuan analisis. Berikut ini adalah hasil dari proses transformasi data penelitian pada tabel 4.3.

Tabel 4.3 Transformasi Data

No	P1	P2	P3	P4	...	P20
1	2	0	2	4	...	4
2	2	2	3	2	...	1
3	2	2	3	1	...	3
4	1	1	1	1	...	3
5	5	3	2	1	...	5
6	2	1	1	2	...	2
7	4	3	3	2	...	1
8	2	2	1	2	...	2
9	3	2	3	2	...	0
10	5	2	2	2	...	2
11	5	1	3	1	...	2
12	2	1	3	4	...	2
...	2	1	2
510	5	3	5	5	...	0

4.2 Tahapan Proses *Fuzzy C-Means*

Pada penelitian ini penulis mengambil data hasil kuisisioner sebanyak 510 data dari kuisisioner yang disebar di wilayah Pekanbaru dan telah dilakukan tahap *preprocessing data*. Data tersebut akan dikelompokkan menggunakan metode *Fuzzy C-Means Clustering* (FCM) dengan jumlah *cluster* sebanyak 2 *cluster*. Adapun pada *cluster* ke-0 di kategorikan dengan tingkat kecanduan rendah, pada *cluster* ke-1 dikategorikan dengan tingkat kecanduan tinggi.

Proses *clustering* dengan menggunakan Algoritma *Fuzzy C-Means* sebagai berikut:

1. Menentukan matriks partisi awal U dalam bentuk matriks dengan ukuran $n \times m$ (dimana n adalah jumlah sampel data yaitu 510, dan m adalah jumlah atribut setiap data berjumlah 20). X_{ij} = data sampel ke- i ($i=1,2,\dots,n$), atribut ke- j ($j=1,2,\dots,m$).
2. Menentukan nilai parameter awal
 - a. Jumlah *cluster* (c) : 2
 - b. Pangkat (w) : 2
 - c. Maksimum iterasi ($MaxIter$) : 100
 - d. Error terkecil yang diharapkan (s) : 10^{-4}
 - e. Fungsi obyektif awal (P_0) : 0
 - f. Iterasi awal (t) : 1
3. Membangkitkan bilangan acak μ_{ik} , $i=1,2,\dots,n$; $k=1,2,\dots,c$; sebagai elemen-elemen matriks partisi awal (U) terdapat pada Tabel 4.4. Nilai-nilainya berkisar antara 0 hingga 1, dengan memberikan nilai secara acak. Jumlah baris matriks sesuai dengan jumlah data, dan jumlah kolomnya sesuai dengan jumlah *cluster*. Setiap kolom matriks berisi jumlah data yang dihasilkan secara acak.

Tabel 4.4 Partisi Awal

No	Partisi Awal	
1	0,6	0,4
2	0,3	0,7
3	0,8	0,2
4	0,2	0,8
5	0,7	0,3
6	0,2	0,8
7	0,3	0,7
8	0,4	0,6
9	0,9	0,1
10	0,7	0,3
...
508	0,2	0,8



Hak Cipta Dilindungi Undang-Undang

1. Diarangi mengutip sebagian atau seluruhnya atau menyalin dalam bentuk apa pun tanpa izin UIN Suska Riau.
2. Diarangi mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

509	0,5	0,5
510	0,5	0,5

4. Menentukan pusat cluster (V)

Pada iterasi pertama, Perhitungan dengan menggunakan Rumus 3.2. Dengan 2 pusat cluster V_{kj} , dengan $k = 1,2$; dan $j = 1,2$. Tabel 4.5 proses menentukan pusat cluster ke-1 dan dan Tabel 4.6 untuk cluster ke-2.

Tabel 4.5 Perhitungan Pusat Cluster pada Iterasi Pertama Cluster ke-1

No	Derajat keanggotaan pada cluster ke-1	Data yang di cluster			$(\mu_{ik1})^2$	$(\mu_{ik1})^2 \cdot x_{i1}$...	$(\mu_{ik1})^2 \cdot x_{i20}$
		μ_{ik1}	x_{i1}	...				
1	0,6	2	...	0	0,37	0,75	...	0,00
2	0,3	2	...	2	0,12	0,24	...	0,24
3	0,8	2	...	2	0,57	1,14	...	1,14
4	0,2	1	...	1	0,05	0,05	...	0,05
5	0,7	5	...	3	0,56	2,79	...	1,67
6	0,2	2	...	1	0,05	0,10	...	0,05
7	0,5	4	...	3	0,23	0,91	...	0,68
8	0,4	2	...	2	0,13	0,26	...	0,26
9	0,9	3	...	2	0,80	2,39	...	1,60
10	0,7	5	...	2	0,46	2,29	...	0,92
11	0,5	1	...	2	0,28	1,42	...	0,57
12	0,2	2	...	2	0,04	0,08	...	0,08
13	0,5	2	...	2	0,29	0,57	...	0,57
...
508	0,6	3	...	0	0,31	0,92	...	0,00
509	1,0	2	...	2	0,90	1,81	...	1,81
510	0,3	5	...	0	0,11	0,56	...	0,00
$V_k = \frac{\sum_{i=1}^n ((u_{ik})^w * X_{ij})}{\sum_i (u_{ik})^w}$					165,12	404,00	...	216,28
						2,45	...	1,31

Table 4.6 Perhitungan Pusat Cluster pada Iterasi Pertama Cluster ke-2

No	Derajat keanggotaan pada cluster ke-1	Data yang di cluster				$(\mu_{ik1})^2$	$(\mu_{ik1})^2 \cdot x_{i1}$...	$(\mu_{ik1})^2 \cdot x_{i20}$
		x_{i1}	...	x_{i20}					
1	0,4	2	...	0	0,15	0,30	...	0,60	
2	0,7	2	...	2	0,43	0,86	...	0,43	
3	0,2	2	...	2	0,06	0,12	...	0,18	
4	0,8	1	...	1	0,59	0,59	...	1,76	
5	0,3	5	...	3	0,06	0,32	...	0,32	
6	0,8	2	...	1	0,60	1,20	...	1,20	
7	0,5	4	...	3	0,27	1,09	...	0,27	
8	0,6	2	...	2	0,41	0,82	...	0,82	
9	0,1	3	...	2	0,01	0,03	...	0,00	
10	0,3	5	...	2	0,10	0,52	...	0,21	
11	0,5	5	...	2	0,22	1,09	...	0,44	
12	0,8	2	...	2	0,65	1,29	...	1,29	
13	0,5	2	...	2	0,22	0,43	...	0,43	
...	
508	0,4	3	...	0	0,20	0,59	...	0,00	
509	0,0	2	...	2	0,00	0,00	...	0,00	
510	0,7	5	...	0	0,44	2,22	...	0,00	
$V_k = \frac{\sum_{i=1}^n ((u_{ik})^w * X_{ij})}{\sum_i (u_{ik})^w}$					171,55	442,65	...	226,87	
						2,58	...	1,32	

Pusat cluster v yang terbentuk pada iterasi pertama adalah :

$$V = \begin{bmatrix} 2,45 & 1,53 & \dots & 1,31 \\ 2,58 & 1,57 & \dots & 1,32 \end{bmatrix}$$

- Setelah mendapatkan pusat cluster, langkah berikutnya adalah menghitung fungsi objektif pada iterasi ke-t, PT dengan persamaan Rumus 3.3.

Tabel 4.7 Perhitungan Fungsi Objektif

No	Kuadrat derajat			L1	L2	LT = L ₁ + L ₂
	x _{i1}	...	x _{i2}			
1	2	...	0	9,107	4,282	13,389
2	2	...	2	2,291	7,644	9,935
3	2	...	2	9,977	1,108	11,085
4	1	...	1	1,003	12,627	13,630
5	5	...	3	48,390	6,058	54,449
6	2	...	1	0,463	5,811	6,273
7	4	...	3	11,206	11,815	23,021
8	2	...	2	1,495	4,896	6,391
9	3	...	2	51,147	0,731	51,879
10	5	...	2	8,794	1,783	10,578
11	5	...	2	17,013	12,458	29,471
12	2	...	2	1,080	18,049	19,129
13	2	...	2	8,004	6,054	14,058
14	3	...	1	8,265	10,899	19,163
15	3	...	1	6,228	0,231	6,459
16	2	...	3	72,254	0,183	72,437
17	2	...	3	4,329	44,525	48,854
18	2	...	1	9,600	0,128	9,728
19	3	...	1	1,001	3,223	4,224
20	3	...	1	5,854	5,509	11,363
21	3	...	3	0,862	16,836	17,698
∴	∴	∴	∴	∴	∴	∴
508	3	...	0	3,479	2,237	5,717
509	2	...	2	4,225	0,011	4,236
510	5	...	0	14,624	57,935	72,560
Fungsi objektif = ∑						9328,202

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mempublikasikan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

6. Hitung perubahan dalam matriks partisi dengan: $i = 1,2,\dots,n$; dan $k = 1,2,\dots,c$ dengan menggunakan Rumus 3.4. Dalam proses pencarian hasil dari derajat keanggotaan, μ_{ik} dihitung melalui langkah-langkah tertentu. Tabel 4.8 merupakan perhitungan perubahan matriks partisi. Pertama-tama, nilai μ_{ik} diperoleh dengan mengalikan setiap nilai bobot dengan pusat *cluster*, menggunakan sistem perkalian matriks (baris dikali kolom). Selanjutnya, hasil tersebut dipangkatkan dengan $-1/\text{bobot}$ yang telah ditetapkan sebelumnya, dikurangi 1. Total nilai yang dihasilkan kemudian dibagi dengan jumlah total baris dalam setiap *cluster*. Perhitungan derajat keanggotaan dinotasikan dengan L_1 , yang merujuk pada perhitungan derajat keanggotaan *cluster* pertama untuk data 1 sampai n. L_2 awalnya merujuk pada perhitungan derajat keanggotaan *cluster* kedua untuk data 1 sampai n. Selanjutnya, nilai L_2 digabungkan dengan L_1 untuk menghasilkan LT , yang mencerminkan hasil penjumlahan L_1 dan L_2 .

Tabel 4.8 Perhitungan Perubahan Matriks Partisi

No.	L_1	L_2	LT	μ_{i1}	μ_{i2}
				L_1/LT	L_2/LT
1	0,0411	0,0352	0,0763	0,5385	0,4615
2	0,0513	0,0565	0,1078	0,4760	0,5240
3	0,0574	0,0535	0,1109	0,5175	0,4825
4	0,0543	0,0465	0,1009	0,5386	0,4614
5	0,0115	0,0106	0,0221	0,5213	0,4787
6	0,1108	0,1030	0,2138	0,5184	0,4816
7	0,0203	0,0231	0,0435	0,4680	0,5320
8	0,0858	0,0842	0,1699	0,5047	0,4953
9	0,0156	0,0155	0,0311	0,5011	0,4989
10	0,0522	0,0584	0,1105	0,4721	0,5279
11	0,0167	0,0175	0,0342	0,4886	0,5114
12	0,0357	0,0358	0,0715	0,4994	0,5006
13	0,0357	0,0358	0,0715	0,4994	0,5006
...
508	0,0885	0,0886	0,1771	0,4996	0,5004
509	0,2138	0,2189	0,4327	0,4940	0,5060
510	0,0076	0,0077	0,0153	0,4977	0,5023

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Matriks partisi baru (U_1) untuk iterasi pertama sebagai berikut:

$$U = \begin{bmatrix} 0,5385 & 0,4615 \\ 0,4760 & 0,5240 \\ 0,5175 & 0,4852 \\ 0,5386 & 0,4614 \\ 0,5123 & 0,4787 \\ 0,5184 & 0,4816 \\ 0,4680 & 0,5320 \\ 0,5047 & 0,4953 \\ 0,5011 & 0,4989 \\ 0,4721 & 0,5279 \\ \dots & \dots \\ 0,4977 & 0,5023 \end{bmatrix}$$

7. Mengecek kriteria perhitungan

Jika ($|P^t - P^{t-1}| \xi$) atau ($t > \maxIter$) maka berhenti:

Jika tidak: $t = t+1$, ulangi langkah ke-4. Proses berhenti pada iterasi ke-23.

Dengan nilai $7051,815 - 7051,815 = 0,0000601581 > 3,176554658$.

Kemudian langkah berikutnya adalah menentukan setiap *cluster* yang terdapat pada Tabel 4.9. Nilai keanggotaan baru yang lebih besar akan masuk kedalam *cluster* tersebut. *Cluster* terbagi 2, *cluster* 0 (tingkat kecanduan rendah) dengan jumlah 233 record. Sedangkan *cluster* 1 (tingkat kecanduan tinggi) dengan jumlah 277 record.

Tabel 4.9 Menentukan Cluster

NO	Keanggotaan Baru		Cluster
	L_1/LT	L_2/LT	
1	0,541136381	0,458863619	0
2	0,546441538	0,453558462	0
3	0,602273391	0,397726609	0
4	0,498597141	0,501402859	1
5	0,5537235	0,4462765	0
6	0,489895699	0,510104301	1
7	0,581461188	0,418538812	0
8	0,60709292	0,39290708	0
9	0,559713576	0,440286424	0
10	0,601726037	0,398273963	0
...

508	0,388807678	0,611192322	1
509	0,399295159	0,600704841	1
510	0,49999941	0,50000059	1

4.3 Implementasi *Fuzzy C-Means* menggunakan Python

Untuk menjalankan algoritma *clustering fuzzy c-means* menggunakan bahasa pemrograman Python, beberapa pustaka yang diperlukan seperti *pandas*, *numpy*, *matplotlib*, dan *skfuzzy* yang terdapat pada Gambar 4.1.

```

import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
import skfuzzy as fuzz
from sklearn.decomposition import PCA
from seaborn import scatterplot as scatter
  
```

Gambar 4.1 Program Python *Import Library*

Selanjutnya, dataset yang digunakan dalam penelitian ini tersedia dalam format file *.csv* dan akan diimpor menggunakan pustaka *pandas*. Langkah berikutnya adalah memeriksa data untuk memastikan bahwa variabel-variabel dalam data, baik nama kolom maupun tipe data agar tidak mengalami kesalahan. Variabel yang akan digunakan untuk *clustering* adalah nilai pada 20 atribut pertanyaan. Sebelum dilakukan perhitungan analisis *cluster*, nilai-nilai yang akan digunakan untuk perhitungan harus ditentukan, nilai-nilai tersebut adalah sebagai berikut:

- a. Jumlah *cluster* (c) : 2
- b. Pangkat (w) : 2
- c. Maksimum iterasi (MaxIter) : 100
- d. Error terkecil yang diharapkan (s) : 10^{-4}
- e. Fungsi obyektif awal (P_0) : 0
- f. Iterasi awal (t) : 1

Kemudian data diolah menggunakan panggilan fungsi *c-means* terdapat pada Gambar 4.2. Fungsi ini melakukan proses *clustering Fuzzy C-Means* pada data yang diberikan dari pustaka *Fuzzy Logic Toolbox* di Python.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

```

cncr, u, u0, d, jm, p, fpc =
fuzz.cluster.cmeans(X.T, c=c, m=w, error=error,
maxiter=maxIter, init=None)

# Get the cluster membership probabilities
labels = np.argmax(u, axis=0)
print("Number of iterations:", p)
    
```

Gambar 4.2 Program Python Jumlah Iterasi

Algoritma akan berhenti iterasi saat nilai kesalahan antara iterasi kurang dari error. Hasil dari proses *fuzzy c-means* didapatkan proses hingga iterasi ke-23 dengan jumlah 510 data.

```

# Count the number of data points in each
cluster
cluster_counts =
data['cluster_FCM'].value_counts()

# Display the counts for each cluster
print(cluster_counts)
    
```

Gambar 4.3 Program Python Pembagian Cluster

Gambar 4.3 merupakan proses dalam mencari jumlah *cluster*. Jumlah data yang dihasilkan dari proses *clustering* terbagi menjadi 2 *cluster*, *cluster 0* yang merupakan kategori kecanduan internet dengan tingkat kecanduan rendah, dengan total 233 *record*. Sementara itu, *cluster 1*, yang menggambarkan kecanduan internet pada tingkat kecanduan tinggi, memiliki jumlah data terbanyak yaitu 277 *record*.

```

# Add cluster labels to the original
dataset
data['cluster_FCM'] = labels
# Display the dataset with cluster labels
    
```

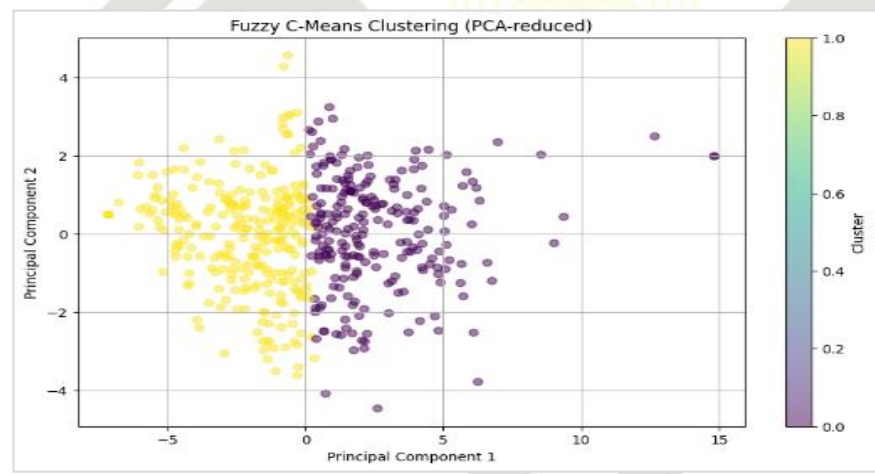
Gambar 4.4 Program Python Menampilkan Tabel

Gambar 4.4 adalah program untuk menampilkan tabel dari setiap *cluster* yang telah ditentukan berdasarkan data yang tersimpan di labels, Tabel 4.10 merupakan hasil pengelompokan data setiap *cluster*.

Gambar 4.10 Hasil Pengelompokan Cluster

No	P1	P2	P3	P4	P5	P6	P7	P20	Cluster
0	2	0	2	4	1	1	0	1	0
1	2	2	3	2	1	2	1	2	0
2	2	2	3	1	2	2	3	1	0
3	1	1	1	1	1	1	2	3	1
4	5	3	2	1	1	3	5	4	0
...
505	2	1	1	0	0	1	1	1	1
506	2	1	1	0	0	0	0	0	1
507	3	2	1	1	2	1	1	1	1
508	2	2	1	1	1	1	2	1	1
509	5	3	5	5	5	3	3	5	0

Gambar 4.5 menampilkan hasil tampilan pengolahan data dengan pembagian menjadi 2 kelompok menggunakan algoritma *Fuzzy C-Means*.



Gambar 4.5 Visual Plot Cluster

Visualisasi hasil *clustering* menggunakan *scatter plot* menampilkan data yang terkelompok menjadi dua *cluster* utama yang dapat dibedakan melalui perbedaan warna. Warna ungu mewakili *cluster* 0 yang memiliki tingkat kecanduan rendah, sementara warna kuning mewakili *cluster* 1 yang memiliki tingkat kecanduan tinggi. Penjelasan gambar adalah sebagai berikut :

- a. *Cluster* ungu, yang mewakili tingkat kecanduan rendah, menunjukkan bahwa kelompok data ini memiliki nilai-nilai rendah atau karakteristik

Hak Cipta Dilindungi Undang-Undang
 1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

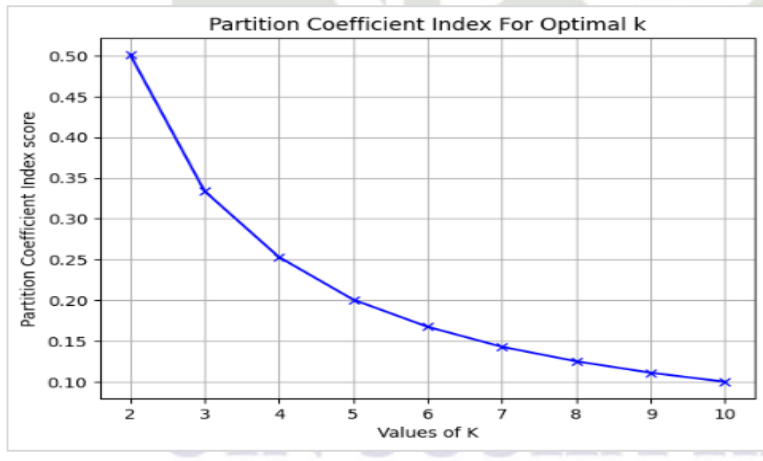
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

yang kurang dominan dalam dataset. Ini mencakup data yang memiliki nilai rendah dalam fitur-fitur yang dianalisis atau memiliki karakteristik yang kurang signifikan dalam permasalahan yang diteliti.

- b. *Cluster* kuning yang mewakili tingkat kecanduan tinggi, menunjukkan bahwa kelompok data ini memiliki nilai-nilai tinggi atau karakteristik yang lebih dominan dalam dataset. Ini menandakan bahwa data dalam *cluster* kuning memiliki nilai tinggi dalam fitur-fitur yang dipertimbangkan atau memiliki karakteristik yang sangat relevan atau signifikan dalam permasalahan yang diteliti.

Pemisahan yang jelas antara kedua *cluster* menunjukkan bahwa algoritma *clustering* berhasil mengelompokkan data dengan baik. Setiap *cluster* memiliki pola atau karakteristik yang khas, dan secara visual terlihat bahwa data dalam setiap *cluster* berada dalam wilayah yang terpisah dengan jelas dari *cluster* lainnya. Ini mengindikasikan bahwa algoritma *clustering* berhasil dalam mengidentifikasi pola atau struktur dalam data dan mengelompokkannya berdasarkan kesamaan karakteristiknya.

4.3.1 Pengujian *Cluster* menggunakan *Partition Coefesient Index*



Gambar 4.6 Hasil Penentuan *Cluster*

Partition Coefficient Index adalah salah satu metrik evaluasi yang digunakan dalam pengujian dan validasi hasil dari algoritma *clustering* seperti *Fuzzy C-Means* (FCM). Tujuan pengujian menggunakan PCI adalah untuk

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

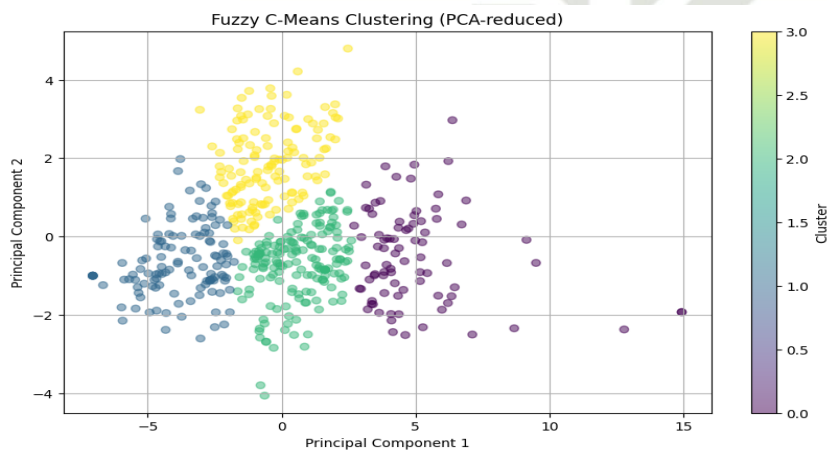
mengukur seberapa baik *cluster* yang dihasilkan oleh algoritma *clustering* dapat membagi data. Dalam penelitian ini, dilakukan eksplorasi terhadap jumlah *cluster* secara berurutan mulai dari 2 hingga 10. Hasil perhitungan uji validitas menggunakan indeks validitas *Partition* (PCI) menunjukkan bahwa *cluster* 2 mendapat perolehan terbaik, dengan nilai sebesar 0.5078 yang terdapat pada Gambar 4.6. Angka ini dikategorikan sebagai baik karena dalam skala PCI, semakin mendekati angka 1, semakin baik hasilnya.

Kemudian data di uji menggunakan validitas kuisiner young dengan 4 *cluster* didapatkan hasil sebagai berikut pada Tabel 4.11.

Tabel 4.11 Pengujian dengan 4 Cluster

No.	Jumlah
C_0	83
C_1	121
C_2	173
C_3	133

Pengujian dilakukan dengan menggunakan 4 *cluster* didapatkan bahwa *cluster* 0 berjumlah 83 *record*, *cluster* 1 berjumlah 121 *record*, *cluster* 2 berjumlah 173, dan *cluster* 3 berjumlah 133 *record*. Visualisasi hasil *clustering* dapat lihat pada Gambar 4.7.



Gambar 4.7 Hasil Penentuan dengan 4 Cluster



Visualisasi hasil *clustering* menggunakan *scatter plot* menampilkan data yang terkelompok menjadi 4 *cluster* utama yang dapat dibedakan melalui perbedaan warna. *Cluster 0* (*cluster* ungu, yang mewakili tingkat kecanduan normal, menunjukkan bahwa kelompok data ini memiliki nilai-nilai rendah atau karakteristik yang kurang dominan dalam dataset. Ini mencakup data yang memiliki nilai rendah dan memiliki karakteristik yang kurang signifikan dalam permasalahan yang diteliti). *Cluster 1* (*Cluster* biru mewakili tingkat kecanduan ringan. Data dalam *cluster* ini menunjukkan tanda-tanda kecanduan, namun dengan tingkat yang lebih rendah dibandingkan dengan *cluster-cluster* lainnya). *Cluster 2* (*Cluster* hijau mewakili tingkat kecanduan sedang. Data dalam *cluster* ini menunjukkan tanda-tanda kecanduan yang lebih signifikan dibandingkan dengan tingkat ringan, tetapi masih belum mencapai tingkat kecanduan yang dianggap berat). *Cluster 3* (*Cluster* kuning mewakili tingkat kecanduan berat. Data dalam *cluster* ini menunjukkan tingkat kecanduan yang signifikan, dengan karakteristik atau nilai yang mencerminkan keterlibatan yang kuat dalam perilaku penggunaan).

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.