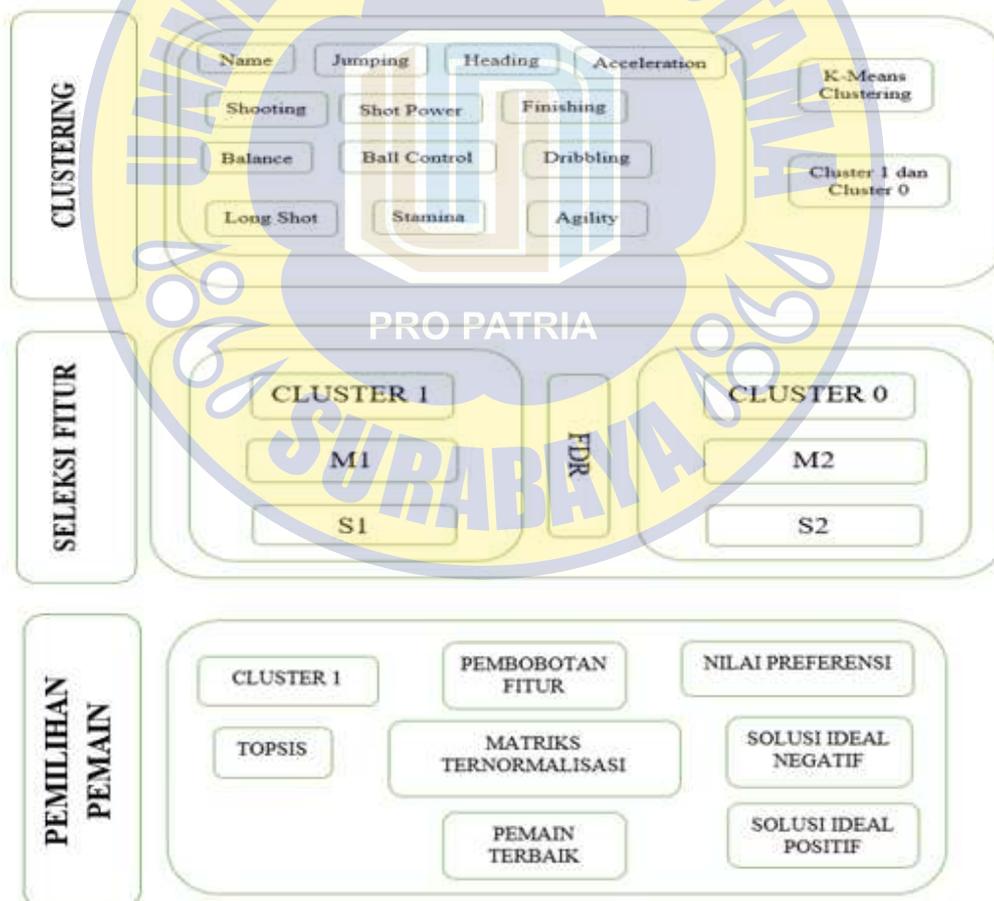


BAB 3 METODOLOGI PENELITIAN

Pada bab ini akan dijelaskan tentang rancangan penelitian, teknik pengumpulan data, populasi dan sampel data, pengembangan instrumen, teknik yang digunakan untuk menganalisa data serta rencana kerja.

3.1 Rancangan Penelitian

Pada bagian ini akan dijelaskan rancangan sistem pemrosesan data seperti pada bagan dibawah ini:



Gambar 3. 1 Desain sistem

3.2 Pengumpulan Data dan Alat

Data yang digunakan berasal dari database sebuah game terbitan EA Sports yaitu FIFA 2019. Dalam penelitian ini data yang digunakan adalah data pemain depan atau penyerang dari setiap tim yang ada pada Liga Primer Inggris. Untuk jumlah data yang digunakan sebanyak 275 *record* dengan 13 atribut pada setiap pemain. Pada penelitian ini penulis menggunakan program Python yang telah disusun menggunakan algoritma K-Means dan metode penunjang keputusan TOPSIS untuk melakukan analisa terhadap data pemain

3.3 Bahan dan Data

Data yang digunakan dalam penelitian ini memiliki 2 tipe yaitu string dan integer untuk tampilan data dan sampelnya ditunjukkan pada tabel dan gambar di bawah ini:

Tabel 3. 1 Atribut Pemain

Attribut	Keterangan	Jenis Data
<i>Name</i>	Nama Pemain	String
<i>Acceleration</i>	Akselarasi Pemain	Integer
<i>Sprint Speed</i>	Kecepatan Lari	Integer
<i>Dribbling</i>	Menggiring Bola	Integer
<i>Agility</i>	Kelincahan	Integer
<i>Balance</i>	Keseimbangan Badan	Integer
<i>Ball Control</i>	Kontrol Bola	Integer

<i>Shooting</i>	Tendangan	Integer
<i>Finishing</i>	Penyelesaian Peluang	Integer
<i>Shot Power</i>	Kekuatan Tendangan	Integer
<i>Heading</i>	Sundukan	Integer
<i>Jumping</i>	Lompatan	Integer
<i>Long Shot</i>	Tendangan Jarak Jauh	Integer
<i>Stamina</i>	Stamina	Integer

Tabel 3. 2 Sampel Data Pemain

Player Name	Acceleration	Sprint Speed	Dribbling	Heading	Jumping	Stamina
Eden Hazard	93	89	94	61	56	83
Sergio Agüero	89	80	90	76	82	77
Harry Kane	68	73	82	86	78	90
Paul Pogba	71	79	86	74	85	91
Mohamed Salah	95	91	90	60	68	85
Alexis Sanchez	87	79	90	74	88	84
Romelu Lukaku	77	91	76	89	71	84
Leroy Sane	93	97	87	73	64	80
Roberto Firmino Barbosa de Oliveira	79	77	87	78	82	91
Sadio Mane	96	93	88	63	76	85
Pierre Emerick Aubameyang	93	95	80	79	79	76
Raheem Sterling	96	92	88	38	58	79
Riyad Mahrez	89	83	91	49	62	77

3.4 Teknik Analisis Data

3.4.1 K-Means Clustering

Berdasarkan atribut yang dimiliki oleh setiap pemain, semua data yang ada akan dilakukan pengelompokan dengan metode *K-Means Clustering*. Semua data tersebut akan dibagi menjadi beberapa cluster. Setiap cluster terdapat titik pusat

yang mejadi parameter bagi data yang lain. tahapan selajutnya dilakukan pengujian apakah ada perubahan pada setiap cluster yang ada. Bila tidak ada perubahan proses selanjutnya adalah seleksi alternatif terbaik melalui pembobotan dan pengujian dengan metode *TOPSIS*.

Prosedur untuk proses perhitungan *K-Means*:

1. Penentuan Jumlah Kluster
2. Mengalokasikan data ke dalam cluster secara acak
3. Menghitung centroid / rata-rata dari data yang ada pada setiap cluster
4. Mengalokasikan data pada centroid terdekat
5. Kembali pada tahap ke 3 apabila masih ada data yang berpindah pada cluster lain ataupun perubahan nilai centroidnya.

Titik pusat atau *centroid* dapat ditentukan secara acak. Lalu selanjutnya centroid dapat didapatkan melalui perhitungan formulasi

$$v_{ij} = \text{SUM} (k=0 \text{ to } N_i) (x_{kj}) / N_i$$

Dimana:

I_k : Indeks cluster

J : Indeks Variabel

V_{ij} : Centroid / rata – rata cluster ke-i untuk variabel ke-j

X_k : Nilai data ke-k yang ada dalam cluster untuk variabel ke-j

N_i : Jumlah data yang menjadi anggota cluster ke- i

Perhitungan antara jarak data dan centroid dapat dilakukan dengan persamaan *Euclidean Distance Space*

$$d_{ij} = \sqrt{\sum_{k=1}^p \{x_{ik} - x_{jk}\}^2}$$

Dimana:

d_{ij} : Jarak antara objek I dan J

P: Dimensi data

x_{ik} : Koordinat dari objek I pada dimensi K

x_{jk} : Koordinat dari objek J pada dimensi K

3.4.2 FDR

Setelah seluruh data melewati proses klastering dan menempati dua klaster yang telah ditentukan, tahapan selanjutnya kami akan melakukan proses uji keterkaitan antar variabel dengan metode FDR (*Fisher's discriminant ratio*) proses ini dilakukan dengan cara menghitung nilai keterkaitan antar variabel yang ada dengan metode statistika. Pengujian dilakukan hingga didapatkan sejumlah variabel dengan nilai varian atau perbedaan antar variabel yang tinggi yang nantinya fitur tersebut akan digunakan dalam proses TOPSIS.

$$FDR = \frac{(m_1 - m_2)^2}{(s_1^2 + s_2^2)} \quad [12]$$

Dimana:

FDR = *Fisher's discriminant ratio*

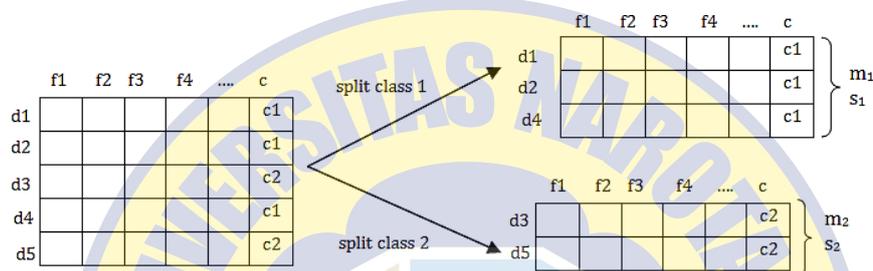
M1 = Mean dari kelas 1

M2 = Mean dari kelas 2

S1 = Varian dari kelas 1

S2 = Varian dari kelas 2

Sebelum dilakukan pengujian terhadap setiap variabel yang ada, pertama – tama data akan displit / dibagi menjadi dua bagian berdasarkan klaster yang ditempati dan kemudian akan diuji masing – masing berdasarkan klasternya.



Gambar 4.1 Proses pemisahan / split data [12]

3.4.3 TOPSIS

Sistem penunjang keputusan dapat digunakan untuk menentukan apakah seseorang pemain memenuhi kriteria atau tidak melalui performa[14]. Sehingga pada tahanan selanjutnya setelah seluruh pemain dikelompokkan berdasarkan acuan yang ditentukan, proses selanjutnya akan dilakukan pembobotan dengan mengukur setiap kriteria-kriteria yang ada. Dilanjutkan dengan perhitungan solusi ideal positif dan solusi ideal negatif Adapun langkah – langkah yang harus dilakukan pada metode Topsis ini adalah sebagai berikut:

1. Penggambaran alternatif dan kriteria kedalam sebuah matriks
2. Membuat matriks ternormalisasi dengan metode

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{k=1}^m x_{kj}^2}}, i = 1, 2, \dots, m, j = 1, 2, \dots, n$$

3. Menghitung matriks terbobot pada keputusan ternormalisasi
4. Menentukan solusi ideal positif dan solusi ideal negatif

$$A^+ = \{(\max V_{ij} \mid j \in J), (\min V_{ij} \mid j \in J'), \\ i = 1, 2, 3, \dots, m\} = \{V_1^+, V_2^+, \dots, V_n^+\}$$

$$A^- = \{(\max V_{ij} \mid j \in J), (\min V_{ij} \mid j \in J'), \\ i = 1, 2, 3, \dots, m\} = \{V_1^-, V_2^-, \dots, V_n^-\}$$

$$J = \{j = 1, 2, 3, \dots, n \text{ dan } j \text{ merupakan } \textit{benefit criteria}\}$$

$$J' = \{j = 1, 2, 3, \dots, n \text{ dan } j \text{ merupakan } \textit{cost criteria}\}$$

5. Perhitungan jarak dari satu alternative ke solusi ideal positif ataupun ke solusi ideal negatif

$$S_i^+ = \sqrt{\sum_{j=1}^n (V_{ij} - V_j^+)^2} \quad S_i^- = \sqrt{\sum_{j=1}^n (V_{ij} - V_j^-)^2}$$

Dengan $i = 1, 2, 3, \dots, m$

Dengan $i = 1, 2, 3, \dots, m$

6. Menghitung preferensi untuk setiap alternative yang ada kemudian dilakukan perangkingan dari nilai C_i^+

$$C_i^+ = \frac{S_i^-}{S_i^+ + S_i^-}$$

Dimana $0 < C_i^+ < 1$ dan $i = 1, 2, 3, \dots, m$

3.5 Rencana Kerja

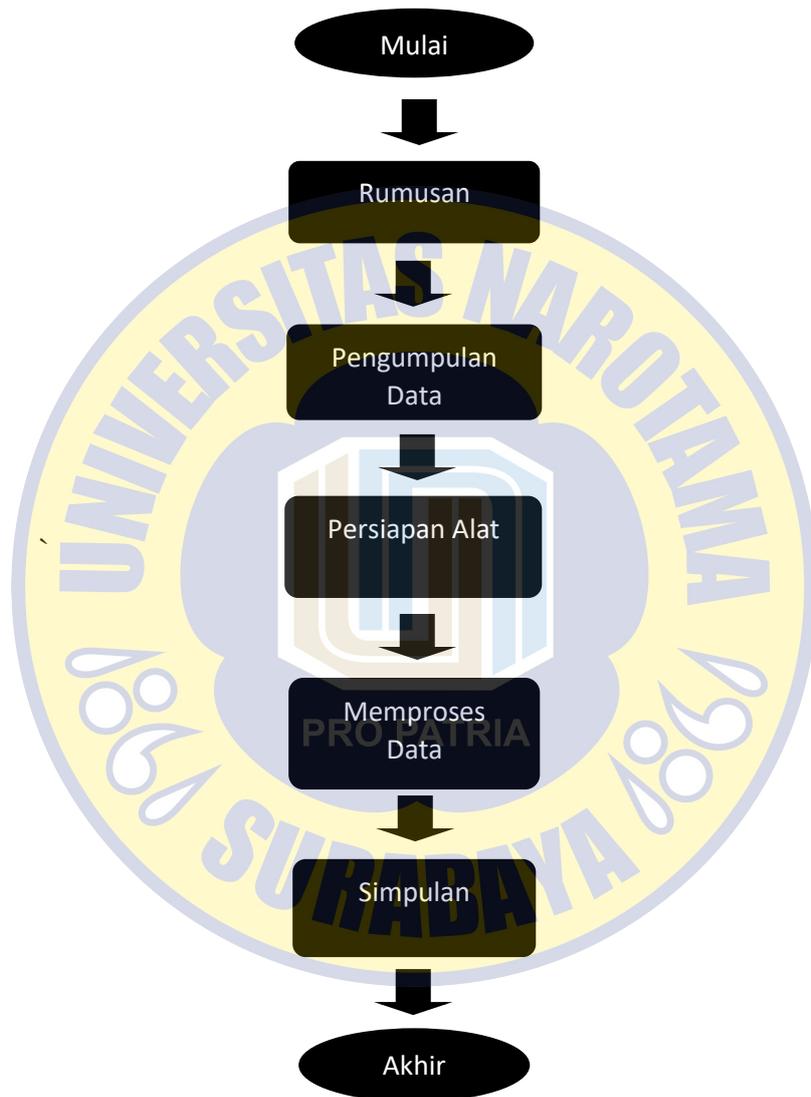
Rancangan rencana kerja yang akan dilukan dalam melakukan penelitian ini mulai dari tahap awal pengumpulan data, *pre processing*, penyesuaian data, pengolahan data hingga tahap akhir penyajian data akan dijelaskan pada table berikut.

PROSES	MARET				APRIL				MEI			
	M1	M2	M3	M4	M1	M2	M3	M4	M1	M2	M3	M4
Perencanaan	■											
Pengumpulan Data		■										
Pre Processing			■									
Penyesuaian Data				■								
Data Processing					■							
Pengelompokan Data								■				
Pengujian Data									■			
Analisa										■		
Hasil												■
Simpulan												■

Gambar 3.4.1 Jadwal Pelaksanaan

3.5 Tahapan Penelitian

Tahapan penelitian yang dilakukan dijelaskan dengan *flowchart* berikut:



3.5.1 *Flowchart* Metodologi Penelitian