

# Penerapan *Data Mining* Untuk Pengelompokan Posisi Pemain Sepak Bola Menggunakan Algoritma *K-Means Clustering*

Mochamad Ali Akbar, Fety Fatimah, Jejen Jaenudin  
Program Studi Teknik Informatika, Fakultas Teknik, Universitas Ibn Khaldun,  
e-mail: [aliakbar020496@gmail.com](mailto:aliakbar020496@gmail.com),

## **Abstrak**

Banyak faktor yang mempengaruhi kemenangan dalam olahraga sepak bola salah satunya yaitu penempatan posisi yang tepat atau sesuai untuk setiap pemainnya, pada posisi depan, tengah dan belakang. Jumlah data yang digunakan yaitu 28 data pemain. *K-Means Clustering* akan mengelompokkan data kedalam cluster tertentu yang memiliki kemiripan data sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu cluster yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan kedalam kelompok yang lain. Hasil dari proses perhitungan clustering menggunakan 3 cluster, cluster 1 masuk kedalam cluster pemain belakang dengan karakteristik nilai rata-rata menguasai bola sebesar 2,00 perjuangan yang khas sebesar 2,11 dan kepemimpinan sebesar 2,11 kemudian cluster 2 masuk kealam kelompok pemain tengah dengan nilai karakteristik penyampaian bola (*passing*) sebesar 3,0 dan penyaluran bola sebesar 2,66 dan cluster 3 masuk kealam kelompok pemain depan dengan karakteristik kecepatan sebesar 1,875. Masing-masing nilai rata-rata dihitung berdasarkan atribut paling tinggi dari tiap cluster. Hasil dari penggunaan aplikasi Weka, cluster 0 berjumlah 13 (46%), cluster 1 berjumlah 8 (29%) dan cluster 2 (25%) berjumlah 7.

**Kata Kunci**—*K-Means, Clustering, Weka, Posisi Pemain, Sepak Bola*

## **Abstract**

Many factors that affect sports football victory in one of them, namely the right position or placement appropriate to each of the players, on the position of front, middle and back. The amount of data used, namely data 28 players. *K-Means Clustering* will classify data into clusters that have the semblance of data so that data that has the same characteristics are grouped into a single cluster, and data that has different characteristics are grouped into other groups. The result of the process of calculation of clustering using 3 clusters, cluster 1 entry into a cluster of players back to the characteristics of the average value of possession of the typical struggle of 2.00 2.11 and 2.11 leadership of then cluster 2 incoming kealam group midfielder with a value of characteristics of delivery ball (*passing*) of 3.0 and channeling the ball of 2.66 and cluster 3 incoming kealam group forward with characteristic speed of 1.875. The respective average value calculated based on the highest attributes of each cluster. The result of the use of application Weka, cluster 0 totaled 13 (46%), cluster 1 totaled 8 (29%) and cluster 2 (25%) amounted to 7.

**Keywords**—*K-Means, Clustering, Weka, Position Players, Football*

## **PENDAHULUAN**

Perkembangan teknologi informasi semakin melesat. Teknologi informasi merupakan suatu teknologi yang digunakan untuk mengolah data, memproses, mendapatkan, menyusun dan menyimpan data dalam berbagai cara untuk menghasilkan informasi yang berkualitas juga relevan, akurat dan tepat waktu. Teknologi informasi ini pun sudah merambak hingga

kecabang olah raga salah satunya yaitu sepak bola. Dalam olah raga sepak bola teknologi informasi ini sangat berguna dalam pengolahan data pemain sepak bola seperti posisi pemain dan skil masing-masing pemain. Adapun metode yang dipakai dan masuk kedalam teknologi informasi yaitu metode data mining.

Data mining mengacu pada proses untuk menambang (*mining*) pengetahuan dari sekumpulan data yang sangat besar. Sebenarnya

data mining merupakan suatu langkah dalam knowledge discovery in databases (KDD). Dalam proses mencari pola atau informasi data terpilih dengan menggunakan teknik atau metode tertentu. Pola-pola ini dikenali oleh perangkat tertentu yang dapat memberikan suatu analisa data yang berguna dan berwawasan yang kemudian dapat dipelajari lebih teliti, yang memungkinkan menggunakan perangkat pendukung keputusan yang lainnya. [1]

*Clustering* (Klasterisasi) adalah proses mengelompokkan atau penggolongan objek berdasarkan informasi yang diperoleh dari data yang menjelaskan hubungan antar objek dengan prinsip untuk memaksimalkan kesamaan antar anggota satu kelas dan meminimumkan kesamaan antar kelas/*cluster*. *Clustering* dalam data mining berguna untuk menemukan pola distribusi di dalam sebuah dataset yang berguna untuk proses analisa data. [2]

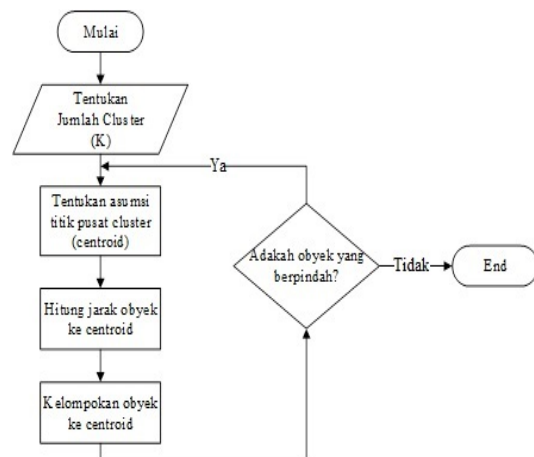
Sekolah sepak bola (SSB) salah satu wadah untuk menampung kegiatan pembelajaran mengenai sepak bola. Pada umumnya sekolah sepak bola menampung peserta didik anak-anak sampai ke usia dewasa. Hal ini dikarenakan tingkatan umur dari anak-anak menuju dewasa merupakan tingkatan umur yang mampu dibentuk untuk menjadi pemain sepak bola. SSB Jenaya FC Bogor salah satu dari sekolah sepak bola yang berpotensi untuk mengembangkan bakat muda dari peserta didiknya.

SSB Jenaya FC pada proses seleksi tim kepelatihan tidak selalu menyeleksi pemain berdasarkan kriteria yang ditetapkan, sehingga terkadang proses seleksi tidak berjalan sesuai kemampuan yang diukur dengan tepat untuk setiap pemain. Oleh karena itu diadakan penelitian yang bertujuan untuk menganalisis potensi data personal disetiap pemain yang akan dikelompokkan ke dalam *cluster* menggunakan algoritma *K-Means*.

**METODE PENELITIAN**

**a. Tahapan Review**

K-Means merupakan metode klasterisasi yang sering digunakan diberbagai bidang karena penggunaannya sederhana, mudah untuk diimplementasikan, mampu untuk mengklaster data yang besar. [3] [4] Adapun tahapan-tahapan algoritma K-Means yang lain, yaitu:



**Gambar 1. Tahapan Algoritma K-Means**

Berikut penjelasan dari Gambar 1

1. Pada saat menentukan data *centroid* pertama menggunakan inisial k sebagai data pertama dari data-data yang nantinya akan dilakukan pengclusteran dan pemilihan jumlah *cluster* sesuai dengan data yang ingin diolah.
2. Dalam penentuan titik pusat *cluster* (*Centroid*) dilakukan secara acak (*random*).
3. Menghitung jarak antara *centroid* dengan masing-masing data.
4. Mengelompokan data berdasarkan jarak minimum.
5. Jika terdapat kesamaan data dengan data sebelumnya, maka proses dihentikan. Jika tidak terjadi kesamaan, lakukan dengan cara yang ke-3.

Menghitung jarak menggunakan rumus *Euclidian Distance*. [5] [6] Berikut ini adalah rumus *Euclidian Distance*:

$$D_{ik} = \sqrt{\sum_j^M = 1(x_{ij} - C_{kj})^2} \dots\dots\dots(1)$$

Keterangan:

- $D_{ik}$  = titik, dokumen/jarak data ke-i
- $M$  = jumlah Variabel
- $x_{ij}$  = data yang akan dilakukan pengklasteran
- $C_{kj}$  = pusat dari *cluster*

Berikut merupakan rumus penentuan *centroid* baru:

$$V_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj} \dots\dots\dots(2)$$

Keterangan:

- $V_{ij}$  = *centroid*/ rata-rata *cluster* ke-i untuk variable ke-j
- $N_i$  = jumlah data yang menjadi anggota *cluster* ke-i
- $i, k$  = indeks dari *cluster*
- $j$  = indeks dari variabel
- $X_{kj}$  = nilai data ke-k yang ada di dalam *cluster* tersebut untuk variable ke- j

Setelah didapat hasil dari proses *clustering* dengan algoritma *K-Means* perlu diadakan suatu pengukuran *cluster* pengukuran. Pengujian tersebut dengan menggunakan metode *sum squared error* (SSE). Selain menggunakan SSE, pengukuran bisa digunakan juga untuk perhitungan perbandingan antara *Between-Class Variation* (BCV) dan *Within-Class Variation* (WCV) pada iterasi terakhir yang sering disebut dengan rasio. [7] [8] [9]

BCV merupakan rata-rata dari jarak antar class, sedangkan WCV adalah SSE itu sendiri. Rumus perhitungannya adalah sebagai berikut:

$$BCV = \frac{1}{n_k} - \sum_{i=1}^k d(m_i, m_i) \dots \dots \dots (3)$$

Keterangan:

*k* = jumlah *cluster*

*i* = nama yang mewakili *cluster* yang dibentuk

*m<sub>i</sub>* = jumlah anggota dari *cluster* ke-*i*

$$WCV = \sum_{j=i}^n \sum_{p \in c_i} d(p, m_i)^2 \dots \dots \dots (4)$$

Keterangan:

*p<sub>eci</sub>* = jumlah semua data

*k* = jumlah *cluster*

*p* = *cluster* jarak terdekat

*m<sub>i</sub>* = jumlah anggota dari *cluster* ke-*i*

$$Rasio = \frac{BCV}{WCV} \dots \dots \dots (5)$$

### b. Pengumpulan Data

Data yang diambil dari sekolah sepak bola Jenaya FC yaitu usia 11 tahun yg berjumlah 28 orang. Adapun kriteria posisi pemain yaitu, pemain depan (*forward*), tengah (*midfilder*) dan belakang (*defender*). Masing-masing kriteria memiliki berbagai aspek atau karekteristik yang berbeda, bisa dilihat pada Tabel 1.

Tabel 1. Tabel Kriteria Pemain

Kriteria pemain Depan	Kriteria pemain tengah	Kriteria pemain Belakang
Aspek	Aspek	Aspek
Pengukuran bentuk tubuh	Pengukuran bentuk tubuh	Pengukuran bentuk tubuh
Menendang	Menendang	Menandai lawan
Memimpin bola	Sampaikan bola	Menendang
Gerakan	Penyaluran bola	Menguasai bola
Kecepatan	Kecerdasan	Perjuangan yang khas
Sentuhan	Membaca	Kecerdasan

akhir	permainan	otak/mental
Kecerdasan		Membaca
otak		permainan
Membaca		Kepemimpinan
permainan		
Perjuangan yang seimbang		

## HASIL DAN PEMBAHASAN

### a. Analisis Dengan Metode K-Means

Dalam kasus ini perhitungan metode *K-Means* menggunakan 3 *cluster*. Kriteria pemain yang diambil untuk diklasterisasikan yaitu pengukuran bentuk tubuh (depan, tengah dan belakang), menendang, memimpin bola, gerakan, kecepatan, sentuhan akhir, kecerdasan, membaca permainan, perjuangan yang seimbang, passing, penyaluran bola, menandai lawan, menguasai bola, perjuangan yang khas, kepemimpinan. Kemudian inialisasi pusat *centroid* dan inialisasi kriteria pemain.

Proses *clustering* yang dilakukan dengan menggunakan 3 *cluster* didapatkan jumlah data tiap *cluster*, data dapat dilihat pada Tabel 2.

Tabel 2. Jumlah Data Tiap Cluster

Cluster	Jumlah
Cluster 1	9 data
Cluster 2	3 data
Cluster 3	16 data
Total	28 data

Dari pembagian *cluster* diatas didapatkan kecocokan terhadap *cluster* yang telah terbentuk dari hasil *clustering*. Dengan menghitung nilai rata-rata setiap atribut yang terdapat pada setiap *cluster* maka dapat ditentukan kelompok *cluster* mana yang masuk kedalam kriteria pemain depan, tengah dan belakang. Berikut adalah klasifikasi hasil *clustering* dapat dilihat pada Tabel 3.

Tabel 3 Klasifikasi Hasil Clustering

Cluster	Keterangan	Rata-rata keseluruhan
Cluster 1	Cluster yang masuk kelompok pemain belakang	2,22
Cluster 2	Cluster yang masuk kelompok	3

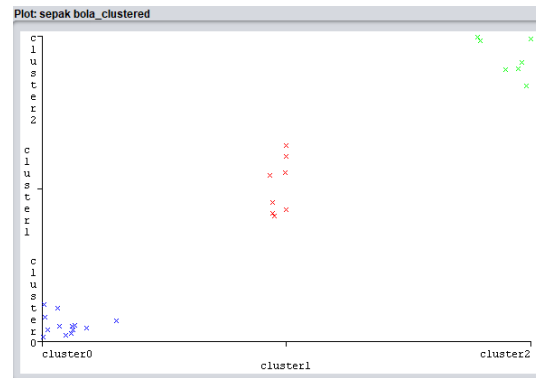
Cluster 3	pemain tengah Cluster yang masuk kelompok pemain depan	1,875
-----------	---	-------

Dilihat dari hasil yang terbentuk dari proses *clustering* menggunakan metode *K-Means*, interpretasi pengelompokan karakteristik dari tiap *cluster* yang terbentuk terhadap tiap atribut adalah sebagai berikut:

1. *Cluster* yang masuk kelompok pemain belakang. *Cluster* 1 atau yang sudah ditentukan sebagai “*cluster* yang masuk kelompok pemain belakang” memiliki karekteristik berbeda dari *cluster* yang lain. Dari data karakteristik pemain bola, nilai rata-rata dari atribut yang paling tinggi diambil untuk menentukan *cluster* mana yang masuk kedalam pemain depan, tengah atau belakang. Dari data tersebut *cluster* 1 memiliki nilai rata-rata menguasai bola sebesar 2,00 perjuangan yang khas sebesar 2,11 dan kepemimpinan sebesar 2,11 dimana kriteria tersebut adalah karakteristik yang khas dalam kriteria pemain belakang.
2. *Cluster* yang masuk kelompok pemain tengah. Dalam klasifikasi ini *cluster* 2 masuk kedalam kelompok pemain tengah, dapat dilihat dari karakteristik pemain tengah yang memiliki nilai rata-rata tertinggi dalam *cluster* 2 yaitu penyampaian bola (*passing*) sebesar 3,0 dan penyaluran bola sebesar 2,66 yang mana *passing* dan penyaluran bola adalah karakteristik yang khas dalam kriteria pemain tengah.
3. *Cluster* yang masuk kelompok pemain depan Hasil *cluster* yang masuk kelompok pemain depan yaitu *cluster* 3. Karakteristik yang menonjol yaitu kecepatan sebesar 1,875 yang mana lebih tinggi dari atribut yang lainnya dalam *cluster* 3.

#### b. Analisis Dengan Metode K-Means Menggunakan Aplikasi Weka 3.9

Didapatkan hasil dari *Clustering* data kriteria pemain sekolah sepak bola (SSB) Jenaya FC dapat dilihat pada grafik *clustering*.



Gambar 2. Hasil Grafik *Clustering* dengan Weka

Dari hasil clastering dengan aplikasi Weka pengelompokan *cluster* telah terbentuk. data tiap *cluster* dapat dilihat pada Gambar 3.

Clustered Instances	
0	13 ( 46%)
1	8 ( 29%)
2	7 ( 25%)

Gambar 3. Jumlah data tiap *cluster*

#### KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan, dapat menarik beberapa kesimpulan, antara lain:

1. Analisis Metode *K-Means clustering* pengelompokan posisi pemain sepak bola bisa dilakukan menggunakan aplikasi Weka maupun perhitungan manual.
2. Hasil *cluster* dapat dipengaruhi dari nilai *centroid* awal yang dipakai dan jumlah data yang dipakai, perbedaan pengambilan data pusat *centroid* awal yang dipakai akan mempengaruhi hasil *centroid* akhir.

#### DAFTAR PUSTAKA

- [1] Joko Risanto, "Proses *Data Mining* Dalam Meningkatkan Sistem Pembelajaran Pada Pendidikan Sekolah Menengah Pertama" Universitas Lampung, Lampung, 2013.
- [2] Achmad Fauzan, dkk "Sistem Klasterisasi Menggunakan Metode K-Means dalam Menentukan Posisi Access Point Berdasarkan Posisi Pengguna Hotspot di Universitas Muhammadiyah Purwokerto" Universitas Muhammadiyah Purwokerto, JUITA ISSN: 2086-9398 Vol. III Nomor 1,2014.
- [3] Dewi Kartika Pane, "Implementasi Data

- Mining Pada Penjualan Produk Elektronik Dengan Algoritma Apriori” Pelita Informatika Budi Darma, Volume : IV, no 3, ISSN : 2301-9425, 2013.
- [4] Kusrini, dkk “Definisi Data Mining” STMIK AMIKOM Yogyakarta, 2014.
- [5] Deka Dwinavinta Candra Nugraha, “Klasterisasi Judul Buku dengan Menggunakan Metode K-Means” Seminar Nasional Aplikasi Teknologi Informasi (SNATI), ISSN: 1907 – 5022, 2014.
- [6] Mikael Aditya Wahyu KM, “Penerapan Metode K-Means Clustering Untuk Mengelompokkan Potensi Produksi Buah-Buahan Di Provinsi Daerah Istimewa Yogyakarta”, Universitas Sanata Dharma, 2007.
- [7] Wardani Anindya Khrisna. “Implementasi Algoritma K-Means Untuk Pengelompokan Penyakit Pasien Pada Puskesmas Kajen Pekalongan,” Universitas Diponegoro, Vols. 14, N0 1, 2016.
- [8] Alfina B, dkk, “*Analisa Perbandingan Metode Hierarchical Clustering , Kmeans Dan Gabungan keduanya Dalam Membentuk Cluster Data (Studi kasus: Problem Kerja Praktek Jurusan Teknik Industri Its)*”, TEKNIK ITS, Vols. 1, pp, pp. 1-5, 2012.
- [9] Asroni, dkk, “Penerapan Metode K-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interface Studi Kasus Pada Jurusan Teknik Informatika UMM Magelang” Semesta Teknika, Vol. 18 No. 1, 76-82, 2015