



Jurnal Teknologi Informasi dan Komunikasi

Vol: 14 No 02 2023

E-ISSN: 2477-3255

Diterima Redaksi: 21-10-2023 | Revisi: 24-11-2023 | Diterbitkan: 30-11-2023

Constellation of Football Players Determination Based on Cost and Performance History Using the K-Means Clustering

Eko Prasetyo¹, Almendaris Shandy Priyatama², Fardanto Setyatama³

^{1,2,3}Fakultas Teknik, Universitas Bhayangkara Surabaya, Indonesia

^{1,2,3}Jl. Ahmad Yani, No. 114, Surabaya

e-mail: eko@ubhara.ac.id

Abstract

Determining the constellation of football players determines a team's success when competing on the field. Disassembling players is an option that must be made considering performance history and costs. This research experiments with K-Means to automate the search for groups of players based on performance and price history. Grouping can achieve a constellation of players with high-performance characteristics but at an affordable price. The dataset used in this research is 580 football players for the 2022/2023 season from Sofifa, Fbref, and SofaScore. The data is divided into four player positions: goalkeeper, defender, midfielder, and attacker. Data for each position is grouped into 3 clusters. Each cluster is analyzed to obtain dominant performance indicator values and determine the characteristics of the cluster. Experimental results using K-Means show that cluster 1 is a team with medium player prices but low performance. Cluster 2 has the cheapest price but with the best performance. Meanwhile, cluster 3 is the most expensive but perform similarity to cluster 2

Keywords: constellation, soccer player, football, k-means, performance indicator, transfer cost

Sistem Penentuan Konstelasi Pemain Sepak Bola Berdasarkan Harga dan Riwayat Performa Menggunakan Metode K-Means

Abstrak

Penentuan konstelasi pemain sepak bola sangat menentukan keberhasilan suatu tim saat bertanding di lapangan. Bongkar pasang pemain menjadi pilihan yang harus dilakukan dengan pertimbangan riwayat performa dan biaya. Penelitian ini bereksperimen dengan K-Means untuk otomatisasi pencarian kelompok pemain berdasarkan riwayat performa dan harga. Pengelompokan dapat mencapai konstelasi pemain dengan karakter performa tinggi tapi dengan harga terjangkau. Dataset yang digunakan dalam penelitian ini adalah 580 pemain sepak bola musim 2022/2023 dari Sofifa, Fbref, dan SofaScore. Data terbagi menjadi 4 posisi pemain yaitu penjaga gawang, bek, gelandang dan penyerang. Data tiap posisi dikelompokkan menjadi 3 cluster. Masing-masing cluster di analisis untuk mendapatkan nilai indikator performa yang dominan dan mengetahui karakteristik cluster. Hasil eksperimen menggunakan K-Means menunjukkan bahwa cluster 1 merupakan tim dengan harga pemain sedang namun performa rendah. Cluster 2 harga paling murah tapi dengan performa terbaik. Sedangkan cluster 3 harga paling mahal tapi dengan performa mirip dengan cluster 2.

Kata kunci: konstelasi, pemain sepak bola, k-means, indikator performa, value

<https://doi.org/10.31849/digitalzone.v14i2.17106>

Digital Zone is licensed under a Creative Commons Attribution International (CC BY-SA 4.0)

1. Pendahuluan

Konstelasi atau bongkar pasang pemain diperlukan untuk memilih pemain terbaik sebagai bagian dari tim sepak bola. Pemilihan ini sangat menentukan keberhasilan suatu tim bertanding di lapangan. Dengan banyaknya pemain bola berbakat dengan kemampuan olah bola yang tinggi, menyebabkan banyak tim atau klub sepak bola bersaing untuk merekrut pemain dengan kemampuan yang tinggi untuk meningkatkan performa tim. Performa ini dipengaruhi oleh beberapa faktor seperti pergerakan, perhatian, memori, citra motorik [1]. Namun, tidak semua klub sepak bola mampu mendatangkan pemain yang diincar karena keterbatasan anggaran. Lebih lanjut lagi, pemain baru juga dibutuhkan untuk meningkatkan kekuatan tim. Oleh karena itu, dibutuhkan tim yang terdiri dari pemain dengan harga yang sesuai anggaran namun memiliki performa yang baik. Pemilihan pemain, untuk menyusun sebuah tim, memerlukan usaha khusus agar karakteristik, kekuatan, dan harga konstelasi tim yang dihasilkan dapat diketahui dan disesuaikan dengan anggaran yang tersedia. Untuk kebutuhan penyusunan pemain yang tepat, perkiraan kinerja pemain sangat diperlukan guna mencapai konstelasi tim yang tepat [2].

Penyelesaian masalah penyusunan tim sepak bola dan transfer pemain sudah dilakukan oleh peneliti sebelumnya, seperti pemilihan kapten tim berdasarkan enam kriteria [3], analisis hubungan antara kekuatan pinggul pemain dengan control kesehatan [4], perbandingan kinerja pemain sepak bola dengan dan tanpa nyeri pinggul [5], penyelidikan preferensi pemain sepak bola anti-doping [6], analisa pelanggaran oleh pemain sepak bola di Australia [7], perekrutan pemain dengan mempertimbangkan kesehatan mental [8], efek musim latihan sepak bola jangka panjang terhadap performance [9], prediksi cedera pemain sepak bola dengan machine learning [10], analisis visual perpindahan pemain sepak bola [11], pemodelan dan analisis pemain sepak bola [12], efek minuman dingin pada pemain sepak bola [13].

Penggunaan bobot kriteria untuk mendapatkan peringkat rekomendasi pemain yang akan dibeli oleh klub yang sudah dilakukan adalah menggunakan metode *Analytical Hierarchy Process* [14], klasifikasi posisi pemain gelandang menggunakan Naïve Bayes [15], dan pemilihan posisi pemain berdasarkan variabel yang lebih dominan menggunakan metode K-Means [16]. Penggunaan K-Means dalam menyelesaikan masalah pengelompokan juga terbukti dibanyak penelitian, seperti pengelompokan prestasi belajar siswa dalam menentukan strategi pembelajaran yang tepat [17], pengelompokan data peminjaman buku perpustakaan [18], pengelompokan penjualan suku cadang mobil [19], pemetaan kelompok mahasiswa berdasarkan aktivitas kuliah [20], pengelompokan produk [21], pengelompokan mahasiswa penerima beasiswa berdasarkan penghasilan orang tua, keadaan rumah, jumlah tanggungan orang tua, dan prestasi [22], pemetaan pemanfaatan teknologi di daerah [23], pengelompokan merk laptop [24]. Dengan popularnya K-Means, metode ini dapat menjadi solusi yang dapat dimanfaatkan pada masalah konstelasi pemain sepak bola.

Penelitian sebelumnya yang terkait pemain sepak bola umumnya adalah pemilihan pemain yang tepat saja. Kriteria yang digunakan seperti kecepatan, arah, dan putaran tendangan. Kriteria tersebut penting namun bersifat umum dan tidak spesifik terhadap posisi didalam tim sepak bola [25]. Satu hasil baik yang dicapai dalam penelitian tersebut adalah penggunaan machine learning dalam pemilihan pemain sepak bola adalah sangat mungkin dilakukan, namun belum mempertimbangkan kombinasi harga dan performa antar pemain. Misalnya, penjaga gawang yang baik jika tidak didukung bek dan gelandang yang baik juga maka kinerja tim di lapangan juga tidak dapat optimal. Oleh karena itu, penelitian ini melakukan eksperimen penentuan konstelasi pemain sepak bola berdasarkan harga dan riwayat performa menggunakan metode K-Means. Penelitian ini menggunakan data pemain sepak bola berdasarkan acuan indikator performa sepak bola dari [26]. Dalam penelitian ini data diolah menggunakan metode K-Means untuk mendapatkan pemain terbaik berdasarkan rata-rata dari keseluruhan indikator dari setiap posisi pemain dan masing-masing cluster. Selanjutnya data tersebut di pisahkan lagi kedalam masing-masing label clusternya. Setelah tim terbentuk, maka dilakukan perbandingan susunan tim dari cluster. Perbandingan dilakukan antara harga dan nilai kekuatan keseluruhan yang

menarik atau seimbang. Sebagai contoh, dataset yang digunakan sebagai sampel data adalah data untuk posisi attacking midfielder dengan jumlah sampel data 70 dari keseluruhan data yang digunakan yaitu 580. Variabel yang digunakan berjumlah 14, diantaranya yaitu tinggi badan, kekuatan fisik, stamina, kecepatan, dan masih banyak lagi.

Penelitian sebelumnya menggunakan machine learning untuk prediksi kualitas pemain berdasarkan kecepatan, arah, dan putaran tendangan, sehingga model yang dibentuk bekerja hanya pada satu pemain saja. Kualitas pemainpun juga didasarkan pada variabel kinerja individu yang berlaku secara umum. Tidak ada penggunaan biaya atau harga pemain sebagai parameter ketika berminat merekrut pemain tersebut [25]. Penelitian ini melakukan pengembangan dengan menggunakan 14 indikator kinerja dan harga transfer pemain untuk mendapatkan pola susunan pemain bersama dengan harga transfer pemain tersebut berdasarkan hasil clustering. Sehingga dihasilkan pengetahuan tentang karakter dominan dari masing-masing cluster yang terbentuk. Pemain dengan nilai rata-rata indikator performa terbaik dari setiap posisi dan masing-masing cluster menjadi karakter dominan. Tim disusun berdasarkan jumlah cluster yang berisi pemain di posisinya masing-masing dan memiliki nilai rata-rata indikator terbaik. Tim yang terbentuk juga dilakukan perbandingan antara harga tim terhadap nilai keseluruhan tim. Hal ini berguna untuk mendapatkan tim manakah yang lebih baik dan lebih seimbang.

2. Metode Penelitian

2.1. Dataset

Eksperimen yang dilakukan dalam penelitian ini menggunakan dataset yang berasal dari data pemain sepak bola di musim 2022/2023. Data ini diambil dan disaring agar sesuai dengan acuan dari indikator performa sepakbola yang terdiri atas pemain dari berbagai posisi, mulai dari penjaga gawang hingga penyerang. Sampel data yang digunakan disajikan pada Tabel 1. Variabel yang digunakan dalam eksperimen penelitian konstelasi pemain sepakbola adalah dari Indikator Performa Sepakbola atau *Football Performance Indicator* (FPI) sebagai ukuran kinerja pemain bola internasional. Ada 14 indikator kinerja sebagai berikut *height, strength, speed, power, stamina, agility, tackle, interc-anticip, heade, shooting, heading, passing, dribbling, dan crossing*. Variabel ini ditambah lagi dengan harga atau biaya transfer pemain. Penggunaan semua variabel tersebut sangat rasional karena menjadi acuan sepak bola internasional sehingga hasil yang dicapai dalam penelitian ini juga sudah menggunakan data dan variabel berskala internasional.

Tabel 1 Sampel Data (*sofifa.com*)

| nama | asal_klub | tfr | hgt | str | spd | pwr | stm | agl | tck | iac | hdr | sht | hdg | psg | drb | crx |
|--------------|-------------|-----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| L. Messi | PSG | 54000000 | 169 | 68 | 81 | 74 | 70 | 91 | 30 | 40 | 70 | 89 | 70 | 90 | 94 | 84 |
| B. Silva | Man City | 97500000 | 173 | 54 | 77 | 68 | 93 | 94 | 62 | 60 | 51 | 78 | 51 | 84 | 92 | 85 |
| T. Muller | FC Bayern | 63000000 | 185 | 66 | 69 | 77 | 86 | 72 | 52 | 58 | 81 | 84 | 81 | 83 | 80 | 85 |
| P. Dybala | Roma | 69500000 | 177 | 60 | 80 | 70 | 68 | 88 | 40 | 42 | 58 | 85 | 58 | 85 | 90 | 82 |
| B. Fernandes | Man United | 72500000 | 179 | 66 | 72 | 78 | 92 | 78 | 69 | 64 | 64 | 86 | 64 | 88 | 81 | 85 |
| C. Nkunku | RB Leipzig | 94000000 | 175 | 64 | 88 | 69 | 78 | 88 | 64 | 67 | 61 | 81 | 61 | 83 | 88 | 81 |
| N. Fekir | Real Betis | 54500000 | 173 | 81 | 83 | 74 | 80 | 85 | 36 | 34 | 49 | 82 | 49 | 82 | 87 | 81 |
| P. Foden | Man City | 109500000 | 171 | 47 | 82 | 64 | 83 | 91 | 53 | 59 | 47 | 78 | 47 | 81 | 88 | 76 |
| M. Reus | B. Dortmund | 33500000 | 180 | 66 | 70 | 69 | 72 | 83 | 53 | 56 | 44 | 84 | 44 | 84 | 85 | 84 |
| A. Gomez | Sevilla FC | 20000000 | 167 | 44 | 86 | 62 | 71 | 91 | 32 | 40 | 38 | 78 | 38 | 83 | 85 | 83 |

Variabel performa yang digunakan dalam penelitian ini meliputi harga transfer pemain dan performa pemain. Untuk performa pemain terdiri dari: *height, strength, speed, power, stamina, agility, tackle, interc-anticip, heade, shooting, heading, passing, dribbling, dan crossing*. Rentang nilai masing-masing variabel disajikan pada Tabel 2. Penjelasan variabel tersebut merujuk pada [27], [28] dan [29] sebagai berikut:

<https://doi.org/10.31849/digitalzone.v14i2.17106>

- *Height*: tinggi badan pemain
- *Strength*: kekuatan fisik pemain
- *Speed*: kecepatan pemain
- *Power*: daya ledak pemain
- *Stamina*: daya tahan tubuh pemain
- *Agility*: kelenturan dan kecepatan gerakan pemain
- *Tackle*: teknik untuk merebut bola dari lawan
- *Interc-Anticip*: kemampuan membaca pergerakan lawan dan mengantisipasi pergerakan bola
- *Header*: teknik memainkan bola dengan kepala
- *Shooting*: teknik menembak bola ke arah gawang
- *Heading*: teknik memainkan bola dengan kepala
- *Passing*: teknik melempar bola ke rekan satu tim
- *Dribbling*: teknik menggiring bola
- *Crossing*: teknik melempar bola ke dalam kotak penalti lawan dari sisi lapangan
- *Transfer*: Nilai atau harga untuk suatu pemain

Tabel 2 Indikator performa pemain sepak bola

| Variable | Description | Value range |
|----------|-----------------------|------------------------------|
| HGT | <i>Height</i> | 163 - 190 |
| STR | <i>Strength</i> | 35 - 86 |
| SPD | <i>Speed</i> | 51 - 92 |
| PWR | <i>Power</i> | 54 - 81 |
| STM | <i>Stamina</i> | 50 - 93 |
| AGL | <i>Agility</i> | 70 - 94 |
| TCK | <i>Tackle</i> | 18 - 77 |
| IAC | <i>Interc-Anticip</i> | 20 - 83 |
| HDR | <i>Heade</i> | 34 - 83 |
| SHT | <i>Shooting</i> | 68 - 89 |
| HDG | <i>Heading</i> | 34 - 83 |
| PSG | <i>Passing</i> | 72 - 90 |
| DRB | <i>Dribbling</i> | 79 - 94 |
| CRX | <i>Crossing</i> | 63 - 86 |
| TFR | <i>Transfer</i> | 10,5 - 109, 5 (Million Euro) |

Setiap variabel memiliki rentang nilai performa beragam, misalnya *height* dengan rentang 163-190 cm, nilai ini didasarkan pada jangkauan nilai pada data yang kami kumpulkan. Demikian pula pada variabel lain yaitu *strength*, *speed*, *power*, *stamina*, *agility*, *tackle*, *interc-anticip*, *heade*, *shooting*, *heading*, *passing*, *dribbling*, dan *crossing*. Semua rentang nilai tersebut didasarkan pada nilai minimal dan maksimal dari dataset yang digunakan.

Tabel 3 Posisi pemain sepak bola

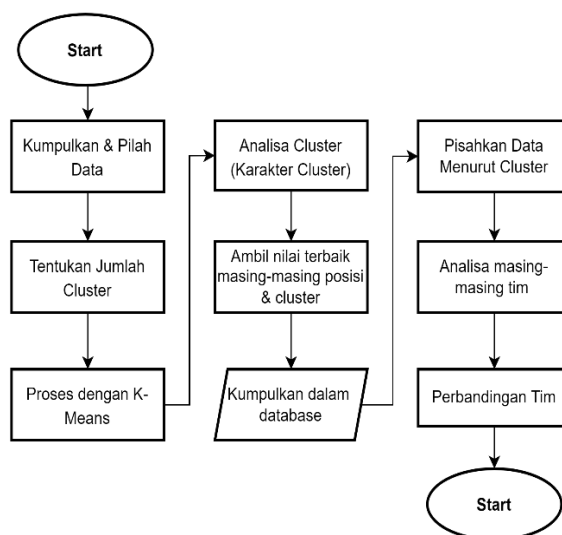
| No. | FPI position | Position | Number of data |
|-----|---------------------------|----------|----------------|
| 1 | <i>GoalKeeper</i> | GK | 100 |
| 2 | | LB | 25 |
| 3 | <i>Full Back</i> | LWB | 25 |
| 4 | | RB | 25 |
| 5 | | RWB | 25 |
| 6 | <i>Center Back</i> | CB | 70 |
| 7 | <i>Holding Midfielder</i> | CDM | 35 |
| 8 | | CM | 35 |

| | | | |
|----|-----------------------------|-----|----|
| 9 | <i>Attacking Midfielder</i> | CAM | 70 |
| 10 | <i>Winger Midfielder</i> | LM | 35 |
| 11 | | RM | 35 |
| 12 | <i>Strikers</i> | CF | 25 |
| 13 | | LW | 25 |
| 14 | | RW | 25 |
| 15 | | ST | 25 |

Tabel 3 menyajikan posisi masing-masing pemain sepak bola, yang terdiri dari penjaga gawang (*goal keeper*), bek (*back*), gelandang (*midfielder*), dan penyerang (*striker*). Masing-masing juga memiliki sub posisi sesuai dengan denah pemain.

2.2 Metodologi Penelitian

Penelitian ini dilakukan dalam beberapa tahap, yaitu: pengumpulan dan pemisahan data pemain berdasarkan posisinya, menentukan jumlah cluster, pemrosesan dengan algoritma K-Means, dan analisis. Metodologi riset pada penelitian ini digambarkan sesuai flowchart Gambar 1 yang akan disertai dengan penjelasan mengenai tahapan-tahapan pada penelitian ini.



Gambar 1. Flowchart Alur Penelitian

Berikut adalah penjelasan mengenai flowchart dari alur penelitian:

- Kumpulkan dan pilah data agar sesuai dengan acuan yang telah ditentukan.
- Tentukan jumlah cluster yang akan dibentuk lalu proses data yang telah di pilah dengan algoritma K-Means hingga data tidak berpindah cluster.
- Analisa masing-masing cluster untuk mengetahui karakter dominan cluster yang telah dibentuk dan ambil data dengan nilai terbaik dari cluster dan juga posisi yang berbeda, setelah itu simpan dan kumpulkan data terbaik kedalam database.
- Setelah seluruh data terkumpul, pisahkan data berdasarkan label cluster lalu analisa kembali masing-masing cluster tersebut untuk mengetahui nilai rata-rata tim serta total harga untuk membentuk tim tersebut.
- Bandingkan tim yang telah dibuat sehingga terlihat tim mana yang memiliki kombinasi antara total harga dan nilai rata-rata tim yang seimbang. Hal ini berarti tim mana yang memiliki nilai rata-rata yang baik namun dengan total harga yang lebih rendah.

2.3. K-Means

K-Means merupakan suatu algoritma yang digunakan dalam pengelompokan secara pertisi yang memisahkan data ke dalam kelompok yang berbeda-beda. Algoritma ini mampu meminimalkan jarak antara data ke cluster nya [30] menggunakan rumus jarak Euclidean (Euclidean Distance) hingga ditemukan jarak yang paling dekat dari setiap data dengan centroid [31].

$$De = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2} \quad (1)$$

Dimana:

De is Euclidean Distance

i is the number of object

(*x*, *y*) is object coordinat

(*s*, *t*) is centroid coordinat

Karena titik centroid akan selalu diperbarui untuk hingga tidak adanya objek data yang berpindah cluster, maka kita menggunakan rumus dibawah ini untuk menentukan centroid baru.

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

Dimana:

v_{ij} is *i*-th centroid or cluster means to for *j*-th variable,

N_i is number of data that member of *i*-th cluster,

i, *k* is index of cluster,

j is index of variable

X_{kj} is value of *k*-th data in cluster for *j*-th variable [32].

2.4. Hasil Pengelompokan

Tahap analisa yang dibahas adalah perbandingan karakteristik antar cluster yang dibentuk berdasarkan nilai rata-rata masing-masing variabel kemudian mengambil nilai terbaik dari masing-masing cluster. Tabel 4 berikut adalah hasil pengelompokan posisi Attacking Midfielder.

| Variabel | Posisi Cluster | | |
|--|----------------|--------|--------|
| | C1 | C2 | C3 |
| HGT (<i>Height</i>) | 181.6 | 176.58 | 175.96 |
| STR (<i>Strength</i>) | 72.35 | 63.08 | 57.75 |
| SPD (<i>Speed</i>) | 77.2 | 79.65 | 72.25 |
| PWR (<i>Power</i>) | 74.8 | 67.15 | 65.54 |
| STM (<i>Stamina</i>) | 81.65 | 72.15 | 77.92 |
| AGL (<i>Agility</i>) | 79.55 | 83.23 | 85.21 |
| TCK (<i>Tackle</i>) | 62.2 | 32.69 | 57.875 |
| IAC (<i>Interception-Anticipation</i>) | 63.45 | 35.38 | 62.75 |
| HDR (<i>Header</i>) | 67.35 | 53.46 | 48.17 |
| SHT (<i>Shoot</i>) | 78.05 | 77.65 | 76.42 |
| HDG (<i>Heading</i>) | 67.35 | 53.46 | 48.17 |
| PSG (<i>Passing</i>) | 79.9 | 79.54 | 81.875 |
| DRB (<i>Dribbling</i>) | 82.7 | 84.15 | 85.125 |
| CRX (<i>Crossing</i>) | 75.7 | 77.73 | 78.83 |

Informasi yang didapatkan dari perbandingan nilai rata-rata pada tabel 4 sebagai berikut:

- Cluster 1 unggul paling banyak pada kategori, yaitu **HGT (Height)**, **PWR (Power)**, **STM (Stamina)**, **TCK (Tackle)**, **IAC (Interception-Anticipation)**, **HDR (Header)**, **SHT (Shoot)**, dan **HDG (Heading)**. Yang berarti unggul dalam kondisi fisik, dan juga pertahanan serta tembakan.
- Cluster 2 unggul hanya dalam kategori **SPD (Speed)**
- Cluster 3 unggul dalam kategori **STR (Strength)**, **AGL (Agility)**, **PSG (Passing)**, **DRB (Dribbling)**, dan **CRX (Crossing)**. Yang berarti unggul dalam kekuatan fisik, kelincahan, dan juga umpan.

Tabel 5 Nilai Terbaik dari Masing-masing Cluster

| Nama | Posisi | Asal Klub | TFR | Rata-Rata | Cluster |
|------------------------|--------|---------------------|----------|-----------|-------------|
| Marcos Llorente | CAM | Atlético de Madrid | 48000000 | 86.21429 | Sangat Baik |
| L. Messi | CAM | Paris Saint-Germain | 54000000 | 80 | Cukup Baik |
| T. Lemar | CAM | Atlético de Madrid | 47000000 | 81.21429 | Baik |

Pada Tabel 5 terlihat data pertama dan data ketiga adalah unggul karena memiliki nilai rata-rata yang tinggi dengan nilai transfer yang lebih rendah. Kondisi sebaliknya terjadi pada data kedua dimana nilai rata-rata paling rendah namun memiliki nilai transfer paling tinggi.

3. Hasil dan Pembahasan

Hasil dari penelitian ini adalah terbentuknya tim dengan jumlah yang sama sesuai jumlah cluster yang ditentukan. Konstelasi tim yang terbentuk adalah seluruh posisi pemain sepak bola yang ada, terdiri dari penjaga gawang, bek, gelandang dan penyerang. Susunan tim disajikan pada Tabel 6, 7, dan 8 sebagai berikut.

Tabel 6 Tim Cluster 1

| nama | posisi | asal_klub | tfr | rata-rata |
|--------------------------|--------|--------------------|-----------|-----------|
| M. Flekken | GK | S.C Freiburg | 16000000 | 75 |
| Nuno Mendes | LB | PSG | 44000000 | 81.3571 |
| Reguilon | LWB | Atlético de Madrid | 25500000 | 80.7143 |
| Rubén Peña | RB | CA Osasuna | 12000000 | 80.4286 |
| M. Darmian | RWB | Inter | 14500000 | 79.7857 |
| J. Matip | CB | Liverpool | 34500000 | 79.2143 |
| A. Zambo Anguissa | CDM | Napoli | 28500000 | 86.2143 |
| F. de Jong | CM | FC Barcelona | 116500000 | 86.4286 |
| Marcos Llorente | CAM | Atlético de Madrid | 48000000 | 86.2143 |
| S. Mane | LM | FC Bayern | 99500000 | 83.3571 |
| Ronaldo Cabrais | RM | Palmeiras | 35000000 | 79.2857 |
| Nuno Santos | LW | Sporting CP | 15000000 | 78.1429 |
| L. Ocampos | RW | Ajax | 27500000 | 84.4286 |
| Diogo Jota | ST | Liverpool | 76000000 | 83.7857 |
| R. Malinovskyi | CF | Atalanta | 30500000 | 83.5714 |

Tabel 7 Tim Cluster 2

| nama | posisi | asal_klub | tfr | rata-rata |
|---------------------|--------|----------------|----------|-----------|
| M. Neuer | GK | FC Bayern | 13500000 | 81.8333 |
| F. Mendy | LB | Real Madrid CF | 38000000 | 86 |
| T. Hernandez | LWB | AC Milan | 76000000 | 87.5 |

<https://doi.org/10.31849/digitalzone.v14i2.17106>

| | | | | |
|------------------------|-----|-------------------|-----------|---------|
| K. Walker | RB | Man City | 37500000 | 85.2143 |
| D. Dumfries | RWB | Inter | 34500000 | 86.3571 |
| E. Can | CB | B. Dortmund | 30500000 | 87.3571 |
| Sergio Busquets | CDM | FC Barcelona | 26500000 | 82.1429 |
| K. De Bruyne | CM | Man City | 107500000 | 83.3571 |
| L. Messi | CAM | PSG | 54000000 | 80 |
| I. Perisic | LM | Tottenham Hotspur | 26500000 | 86.7143 |
| N. Mazraoui | RM | FC Bayern | 40500000 | 83.9286 |
| Neymar Jr. | LW | PSG | 99500000 | 78.0714 |
| Á. Di María | RW | Juventus | 20000000 | 75.0714 |
| Romarinho | CF | Al Ittihad | 12000000 | 75.3571 |

Tabel 8 Tim Cluster 3

| nama | posisi | asal_klub | tfr | rata-rata |
|-----------------------|--------|--------------------|-----------|-----------|
| De Gea | GK | Man United | 42000000 | 76 |
| João Cancelo | LB | Man City | 82500000 | 86.7143 |
| Renan Lodi | LWB | Nottingham Forest | 29500000 | 83.1429 |
| A. Hakimi | RB | PSG | 53500000 | 86 |
| Pedro Porro | RWB | Sporting CP | 38500000 | 83.5 |
| V. van Dijk | CB | Liverpool | 98000000 | 85.3571 |
| G. Kondogbia | CDM | Atlético de Madrid | 29500000 | 88.2857 |
| L. Goretzka | CM | FC Bayern | 91000000 | 90.2143 |
| T. Lemar | CAM | Atlético de Madrid | 47000000 | 81.2143 |
| Raphinha | LM | FC Barcelona | 48000000 | 77.7857 |
| F. Chiesa | RM | Juventus | 66500000 | 76.6429 |
| H. Son | LW | Tottenham Hotspur | 101000000 | 80.8571 |
| A. Carrillo | RW | Al Hilal | 11500000 | 77.3571 |
| R. Lewandowski | ST | FC Barcelona | 84000000 | 83.8571 |
| K. Benzema | CF | Real Madrid CF | 64000000 | 83.0714 |

Susunan tim pada Tabel 6,7, 8 merupakan hasil dari perhitungan data menurut posisi yang sudah ditentukan oleh *Football Performance Indicator* (FPI). Hasil analisa kelompok yang dibentuk oleh algoritma K-Means digunakan untuk menyusun tim, masing-masing posisi dari *GK* (Penjaga Gawang) hingga *Strikers* (Pemain Depan, Penyerang). Susunan tim terdiri atas variabel nama, posisi, lalu nama klub, harga pemain, dan nilai rata-rata yang merupakan hasil dari pembagian seluruh variabel indikator performa. Hal ini untuk menggambarkan bagaimana nilai performa keseluruhan dari masing-masing pemain dalam tim.

Nilai transfer pada susunan tim dijumlahkan untuk mengetahui dana yang dibutuhkan untuk membangun tim dengan daftar pemain yang terbentuk. Rata-rata dari nilai semua pemain juga dihitung untuk mendapatkan performa keseluruhan anggota tim berdasarkan daftar pemain terpilih. Total dana yang dibutuhkan untuk membentuk tim juga dihitung sehingga kita dapat mengetahui alokasi biaya susunan pemain. Rata-rata kinerja semua pemain juga disajikan sehingga kita dapat membandingkan kinerja keseluruhan antar tim. Dengan menyajikan 3 data tersebut, analisa dapat dilakukan untuk memberikan pertimbangan kinerja dan biaya transfer pemain.

Berdasarkan data yang disajikan pada Tabel 9, tim cluster 1 membutuhkan dana 623 juta Euro untuk membangun tim dengan nilai performa tim sebesar 81.8619. Tim cluster 2 membutuhkan dana 616.5 juta Euro untuk membangun tim dengan nilai performa tim sebesar

82.7789. Lalu yang terakhir yaitu tim cluster 3 membutuhkan dana 886.5 juta Euro untuk membangun tim dengan nilai performa tim 82.6667

Table 9 Kinerja keseluruhan dan biaya tim

| Cluster | Total TFR (Euro) | Mean TFR | Performance average |
|---------|------------------|-------------|---------------------|
| 1 | 623000000 | 41533333.33 | 81.8619 |
| 2 | 616500000 | 44035714.29 | 82.7789 |
| 3 | 886500000 | 59100000 | 82.6667 |

Tim cluster 1 memiliki nilai transfer atau harga yang lebih tinggi dari cluster 2 namun lebih rendah dari cluster 3 dan nilai rata-ratanya adalah yang terendah diantara ketiganya. Tim cluster 2 memiliki nilai transfer yang paling rendah namun memiliki nilai rata-rata paling besar jika dibandingkan dengan dua tim lainnya. Tim cluster 3 memiliki nilai transfer tertinggi, sedangkan nilai rata-ratanya berada di antara tim cluster 1 dan juga cluster 2.

Analisa keseluruhan yang didapatkan adalah bahwa tim cluster 2 unggul daripada tim cluster lainnya dalam perbandingan antara harga transfer dan nilai performa keseluruhan. Hal ini karena tim cluster 2 memiliki harga transfer paling rendah, yang berarti tim tersebut adalah yang paling murah daripada dua tim lainnya, namun memiliki nilai performa yang paling tinggi. Tim terbaik kedua adalah tim cluster 1 dengan nilai transfer yang sedikit lebih tinggi daripada tim cluster 2 namun memiliki nilai performa yang lebih rendah. Akan tetapi, tim ini juga menjadi tim alternatif karena selisih performa kedua tim adaah sedikit. Tim yang terakhir cluster 3 dengan nilai transfer paling tinggi dengan nilai performa yang sangat dekat dengan tim cluster 2. Tim ini memiliki harga transfer yang jauh lebih mahal tetapi performa tidak lebih unggul.

4. Kesimpulan

Berdasarkan eksperimen dan hasil yang dicapai dari penelitian ini, dapat disimpulkan bahwa metode K-Means clustering dapat diimplementasikan kedalam sistem untuk melakukan pengelompokan pemain sepak bola berdasarkan harga dan riwayat performa. Pengelompokan dilakukan pada empat posisi pemain yaitu penjaga gawang, bek, gelandang, dan penyerang. Data pemain pada tiap posisi dipisahkan menjadi tiga kelompok untuk mendapatkan tiga alternatif tim. Hasil pengelompokan dilakukan analisis untuk menyusun tim sesuai anggota pada tiap cluster. Hasil analisis menunjukkan bahwa terbantu tim dengan performa yang tinggi tapi dengan harga transfer yang tidak mahal. Hasil yang dicapai ini dapat digunakan sebagai alat bantu untuk menyusun tim yang berisi rekomendasi pemain dari masing-masing posisi sesuai indikator performa sepak bola. Saran untuk peneliti selanjutnya adalah untuk membuktikan manfaat mengenai penelitian ini lebih jauh diperlukan pihak dari tim atau organisasi yang berkaitan dengan sepak bola untuk pendalaman hasil dan pertimbangan lainnya dalam pembentukan sebuah tim sepak bola.

Acknowledgement

Penelitian dilaksanakan dengan pendanaan penelitian mandiri dari Universitas Bhayangkara Surabaya dengan nomor kontrak 06/VIII/2023/LPPM/UBHARA tanggal 1 Agustus 2023.

Daftar Pustaka

- [1] C. Rominger *et al.*, "Female and male soccer players recruited different cognitive processes when generating creative soccer moves," *Psychology of Sport and Exercise*, vol. 50, p. 101748, Sep. 2020, <https://doi.org/10.1016/J.PSYCHSPORT.2020.101748>.
- [2] E. Vallance, N. Sutton-Charani, P. Guyot, and S. Perrey, "Predictive modeling of the ratings of perceived exertion during training and competition in professional soccer players," *Journal of Science and Medicine in Sport*, vol. 26, no. 6, pp. 322–327, Jun. 2023, <https://doi.org/10.1016/J.JSAMS.2023.05.001>.

- [3] I. Coker, S. T. Cotterill, and J. Griffin, "Player perceptions of athlete leadership and leadership development in an English Premier League football academy," *Asian Journal of Sport and Exercise Psychology*, vol. 2, no. 3, pp. 182–189, Dec. 2022, <https://doi.org/10.1016/J.AJSEP.2021.12.001>.
- [4] C. Bonello *et al.*, "The association between hip/groin pain and hip strength in football players: An exploratory analysis of the FORCe cohort," *Journal of Science and Medicine in Sport*, Jul. 2023, <https://doi.org/10.1016/J.JSAMS.2023.07.008>.
- [5] E. A. Roughead *et al.*, "Football players with long standing hip and groin pain display deficits in functional task performance," *Physical Therapy in Sport*, vol. 55, pp. 46–54, May 2022, <https://doi.org/10.1016/J.PTSP.2022.02.023>.
- [6] L. Skoufa, G. Daroglou, A. Loukovitis, H. Lunde, G. Guižauskaitė, and V. Barkoukis, "Football players' preferences for anti-doping education: A cross-country study," *Performance Enhancement & Health*, vol. 10, no. 1, p. 100217, Feb. 2022, <https://doi.org/10.1016/J.PEH.2021.100217>.
- [7] J. B. Farley, J. W. L. Keogh, C. T. Woods, and N. Milne, "Injury profiles of Australian football players across five, women's and girls' competition levels," *Journal of Science and Medicine in Sport*, vol. 25, no. 1, pp. 58–63, Jan. 2022, <https://doi.org/10.1016/J.JSAMS.2021.08.016>.
- [8] T. Ito, M. D. Fetters, C. Kipps, and B. Kumar, "Depressive symptoms among male professional soccer players in Japan," *Asian Journal of Sport and Exercise Psychology*, Mar. 2023, <https://doi.org/10.1016/J.AJSEP.2023.02.002>.
- [9] M. A. Hammami *et al.*, "The effects of a soccer season on anthropometric characteristics, physical fitness, and soccer skills in North African elite female youth soccer players," *Science & Sports*, vol. 38, no. 4, pp. 401–410, Jun. 2023, <https://doi.org/10.1016/J.SCISPO.2022.08.002>.
- [10] F. J. Robles-Palazón *et al.*, "Predicting injury risk using machine learning in male youth soccer players," *Chaos, Solitons & Fractals*, vol. 167, p. 113079, Feb. 2023, <https://doi.org/10.1016/J.CHAOS.2022.113079>.
- [11] A. Cao *et al.*, "MIG-Viewer: Visual analytics of soccer player migration," *Visual Informatics*, vol. 5, no. 3, pp. 102–113, Sep. 2021, <https://doi.org/10.1016/J.VISINF.2021.09.002>.
- [12] A. Beheshtian-Ardakani, M. Salehi, and R. Sharma, "CMPN: Modeling and analysis of soccer teams using Complex Multiplex Passing Network," *Chaos, Solitons & Fractals*, vol. 174, p. 113778, Sep. 2023, <https://doi.org/10.1016/J.CHAOS.2023.113778>.
- [13] P. Schons *et al.*, "Effects of cold water immersion on the physical performance of soccer players: A systematic review," *Science & Sports*, vol. 37, no. 3, pp. 159–166, May 2022, <https://doi.org/10.1016/J.SCISPO.2021.03.008>.
- [14] H. P. K. Negara, E. Santoso, and H. Nurwasito, "Sistem Pendukung Keputusan Menentukan Transfer Pemain Sepak Bola Menggunakan Metode AHP (Analytical Hierarchy Process)," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer (J-PTIHK) Universitas Brawijaya*, vol. 2, no. 7, pp. 2670–2678, 2018.
- [15] N. W. Wisswani, T. R. Tulili, M. F. Andrijasa, M. R. Satria, E. Wahyudi, and S. Jumaini, "Klasifikasi Tipe Gelandang Sepak Bola Berdasarkan Data Kemampuan Menggunakan Metode Naive Bayes," *Just TI (Jurnal Sains Terapan Teknologi Informasi)*, vol. 12, no. 2, pp. 46–50, Jul. 2020, <https://doi.org/10.46964/JUSTTI.V12I2.370>.
- [16] M. A. Akbar, F. Fatimah, and J. Jaenudin, "Penerapan Data Mining Untuk Pengelompokan Posisi Pemain Sepak Bola Menggunakan Algoritma K-Means Clustering," *Semnati*, vol. x, No.x, no. x, pp. 278–282, 2019.
- [17] S. Dewi, S. Defit, and Y. Yuhandri, "Akurasi Pemetaan Kelompok Belajar Siswa Menuju Prestasi Menggunakan Metode K-Means," *Jurnal Sistim Informasi dan Teknologi*, vol. 3, pp. 28–33, 2021, <https://doi.org/10.37034/jsisfotek.v3i1.40>.
- [18] D. Siburian, S. R. Andani, and I. P. Sari, "Implementasi Algoritma K-Means untuk
-

- Pengelompokkan Peminjaman Buku Pada Perpustakaan Sekolah Implementation of K-Means Algorithm for Clustering Books Borrowing in School Libraries,” *JOMLAI: Journal of Machine Learning and Artificial Intelligence*, vol. 1, no. 2, 2022, <https://doi.org/10.55123/jomlai.v1i2.725>.
- [19] Feryanto, F. T. Kesuma, and S. P. Tamba, “Penerapan Data Mining Untuk Menentukan Penjualan Sparepart Toyota Dengan Metode K-Means Clustering,” *Jurnal Sistem Informasi dan Ilmu Komputer Prima (JUSIKOM PRIMA)*, vol. 2, no. 2, pp. 67–72, 2019, <https://doi.org/10.37034/jsisfotek.v3i1.40>.
- [20] R. Rosmini, A. Fadlil, and S. Sunardi, “Implementasi Metode K-Means Dalam Pemetaan Kelompok Mahasiswa Melalui Data Aktivitas Kuliah,” *It Journal Research and Development*, vol. 3, no. 1, pp. 22–31, 2018, [https://doi.org/10.25299/itjrd.2018.vol3\(1\).1773](https://doi.org/10.25299/itjrd.2018.vol3(1).1773).
- [21] M. Dahria, R. Gunawan, and Z. Lubis, “Implementasi K-Means Untuk Pengelompokan Produk Terbaik PT . Koko Pelli,” *Seminar Nasional Sains & Teknologi Informasi (SENSASI)*, pp. 495–498, 2019.
- [22] A. E. Rahayu, K. Hikmah, N. Yustia, and A. C. Fauzan, “Penerapan K-Means Clustering Untuk Penentuan Klasterisasi Beasiswa Bidikmisi Mahasiswa,” *ILKOMNIKA: Journal of Computer Science and Applied Informatics*, vol. 1, no. 2, pp. 82–86, 2019, <https://doi.org/10.28926/ilkomnika.v1i2.23>.
- [23] S. F. Mandang and B. N. Sari, “Penerapan K-Means Cluster Pada Daerah Penggunaan Teknologi di Indonesia,” *JOINS (Journal of Information System)*, vol. 6, no. 1, pp. 131–138, 2021, <https://doi.org/10.33633/joins.v6i1.4545>.
- [24] P. Ulil, F. Aulia, and S. Saepudin, “Penerapan Data Mining K-Means Clustering Untuk Mengelompokkan Berbagai Jenis Merk Laptop,” *SISMATIK (Seminar Nasional Sistem Informasi dan Manajemen Informatika)*, pp. 241–249, 2021.
- [25] T. I. Zohdi, “Machine-learning a perfect bending soccer goal shot,” *Computer Methods in Applied Mechanics and Engineering*, vol. 415, p. 116261, Oct. 2023, <https://doi.org/10.1016/J.CMA.2023.116261>.
- [26] M. Hughes *et al.*, “Moneyball and soccer - An analysis of the key performance indicators of elite male soccer players by position,” *Journal of Human Sport and Exercise*, vol. 7, no. SPECIALISSUE.2, pp. 402–412, 2012, <https://doi.org/10.4100/jhse.2012.72.06>.
- [27] K. T. Kristianto, “Istilah-istilah dalam Sepak Bola,” *Kompas.com*, 2021. .
- [28] K. T. Kristianto, “Istilah-istilah Sepak Bola saat Duel Maupun Bertahan,” *Kompas.com*, 2021. .
- [29] A. Istiqomah, “Istilah-Istilah dalam Transfer Pemain Sepak Bola yang Sering Digunakan,” *sportstars.id*, 2023. .
- [30] W. Pulungan, P. Poningsih, and H. Satria, “Pengelompokkan Pada Kendaraan Bermotor Menurut Kegunaannya Menggunakan Metode Data Mining K-Means,” *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, vol. 3, no. 1, pp. 746–752, 2019, <https://doi.org/10.30865/komik.v3i1.1687>.
- [31] A. Sulistiyawati and E. Supriyanto, “Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan,” *Jurnal Tekno Kompak*, vol. 15, no. 2, p. 25, 2021, <https://doi.org/10.33365/jtk.v15i2.1162>.
- [32] F. Indriyani and E. Irfiani, “Clustering Data Penjualan pada Toko Perlengkapan Outdoor Menggunakan Metode K-Means,” *JUITA : Jurnal Informatika*, vol. 7, no. 2, p. 109, 2019, <https://doi.org/10.30595/juita.v7i2.5529>.