



ANALISIS K-MEANS *CLUSTERING* DENGAN *SEQUENTIAL FORWARD SELECTION* DAN *MULTIPLE VALIDATION METHODS* PADA DATA PERFORMA PEMAIN NBA

Skripsi

**Untuk Memenuhi Persyaratan
Dalam Menyelesaikan Strata-1 Ilmu Komputer**

Oleh

THAHA WAFIQ ADLY

NIM 2111016310016

**PROGRAM STUDI S-1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU**

NOVEMBER 2025



**ANALISIS K-MEANS *CLUSTERING* DENGAN *SEQUENTIAL FORWARD SELECTION* DAN *MULTIPLE VALIDATION METHODS* PADA DATA PERFORMA
PEMAIN NBA**

Skripsi

**Untuk Memenuhi Persyaratan
Dalam Menyelesaikan Strata-1 Ilmu Komputer**

Oleh

THAHA WAFIQ ADLY

NIM 2111016310016

**PROGRAM STUDI S-1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU**

NOVEMBER 2025

SKRIPSI

**ANALISIS K-MEANS CLUSTERING DENGAN SEQUENTIAL FORWARD SELECTION DAN
MULTIPLE VALIDATION METHODS PADA DATA PERFORMA PEMAIN NBA**

Oleh:


THAHA WAFIQ ADLY

NIM. 2111016310016


Telah dipertahankan di depan Dosen Penguji pada tanggal 3 November 2025.

Susunan Dosen Penguji:


Pembimbing I


Irwan Budiman, S.T., M.Kom
NIP. 197703252008121001

Dosen Penguji I


Rudy Herterio, S.Kom, M.Kom
NIP. 198809252022031003

Pembimbing II


Muhammad Itqan Mazdadi, S.Kom., M.Kom.
NIP. 199006122019031013

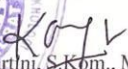
Dosen Penguji II


Triando Hamonangan Saragih, S.Kom, M.Kom
NIP. 199308242019031012

Banjarbaru, 28 November 2025

Koordinator PS Ilmu Komputer




Dwi Kartini, S.Kom., M.Kom.
NIP. 198704212012122003

PERNYATAAN

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Banjarbaru, 2 November 2025

Yang Menyatakan,



Thaha Wafiq Adly

NIM. 2111016310016

ABSTRAK

ANALISIS K-MEANS CLUSTERING DENGAN SEQUENTIAL FORWARD SELECTION DAN MULTIPLE VALIDATION METHODS PADA DATA PERFORMA PEMAIN NBA

(Oleh : Thaha Wafiq Adly; Pembimbing: Irwan Budiman., S.T., M.Kom. dan Muhammad Itqan Mazdadi., S.Kom., M.Kom.; 2025; 91 halaman)

Analisis performa pemain dalam liga profesional seperti NBA menuntut pendekatan objektif berbasis data untuk memahami kontribusi yang bervariasi antar individu. Penelitian sebelumnya dalam *clustering* pemain seringkali memiliki keterbatasan, seperti penentuan jumlah cluster (K) yang subjektif, ketiadaan seleksi fitur yang sistematis, dan kurangnya evaluasi stabilitas hasil. Penelitian ini mengusulkan kerangka kerja analisis *clustering* yang komprehensif menggunakan algoritma K-Means pada data statistik 213 pemain NBA musim 2023-2024. Pendekatan ini mengintegrasikan *Sequential Forward Selection* (SFS) untuk seleksi fitur, *Elbow Method* untuk menentukan jumlah cluster, serta validasi statistik menggunakan ANOVA, uji Tukey HSD, dan evaluasi stabilitas dengan *Adjusted Rand Index* (ARI). Hasil penelitian menetapkan K=5 sebagai jumlah cluster optimal dengan empat fitur terpilih (DRtg, ORtg, TO%, dan USG%). Uji ANOVA mengonfirmasi adanya perbedaan yang sangat signifikan secara statistik antar kelima *cluster* ($p < 0.05$), dan hasil *clustering* terbukti sangat stabil dengan skor rata-rata ARI sebesar 0.9573. Penelitian ini berhasil menyajikan sebuah metode analisis performa pemain NBA yang lebih objektif dan konsisten, yang mampu mengungkap lima arketipe pemain yang berbeda secara bermakna.

Kata Kunci: Analisis Clustering, K-Means, Pemain NBA, Seleksi Fitur, Validasi Statistik

ABSTRACT

K-MEANS CLUSTERING ANALYSIS WITH SEQUENTIAL FORWARD SELECTION AND MULTIPLE VALIDATION METHODS ON NBA PLAYER PERFORMANCE DATA

(By : Thaha Wafiq Adly; Supervisors: Irwan Budiman., S.T., M.Kom. and Muhammad Itqan Mazdadi., S.Kom., M.Kom.; 2025; 91 pages)

Player performance analysis in professional leagues such as the NBA demands an objective, data-driven approach to understand the varying contributions among individuals. Previous research on player clustering often has limitations, such as the subjective determination of the number of clusters (K), the absence of systematic feature selection, and a lack of result stability evaluation. This study proposes a comprehensive clustering analysis framework using the K-Means algorithm on the statistical data of 213 NBA players from the 2023–2024 season. This approach integrates Sequential Forward Selection (SFS) for feature selection, the Elbow Method to determine the number of clusters, as well as statistical validation using ANOVA, the Tukey HSD test, and stability evaluation with the Adjusted Rand Index (ARI). The results established K=5 as the optimal number of clusters with four selected features (DRtg, ORtg, TO%, and USG%). The ANOVA test confirmed a statistically significant difference among the five clusters ($p < 0.05$), and the clustering result proved to be highly stable with a mean ARI score of 0.9573. This study successfully presents a more objective and consistent method for NBA player performance analysis, capable of uncovering five meaningfully different player archetypes.

Keywords: *Clustering Analysis, K-Means, NBA Players, Feature Selection, Statistical Validation*

PRAKATA

Puji syukur penulis panjatkan ke hadirat Allah Subhanahu Wa Ta'ala atas berkat rahmat dan karunia-Nya penulis dapat menyelesaikan skripsi yang berjudul “Analisis K-means *clustering* dengan *sequential forward selection* dan *multiple validation* pada data performa NBA” untuk memenuhi syarat dalam menyelesaikan pendidikan S1 Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat.

Pada lembar ini penulis ingin menyampaikan ucapan terima kasih kepada pihak-pihak yang sangat mendukung penulis dalam pembuatan dan penyusunan skripsi ini, adapun yang dimaksud adalah sebagai berikut:

1. Terima kasih yang sebesar-besarnya untuk Ayah, Bapak Riduan Masykur atas segala doa dan nasihat yang telah membentuk penulis hingga hari ini. Dukungan Ayah adalah pilar utama yang menguatkan penulis dalam menyelesaikan studi ini.
2. Ucapan terima kasih yang sebesar-besarnya penulis sampaikan kepada Kak ulwan, Penulis sangat berterima kasih atas segala dukungan moril dan materiel yang telah diberikan selama masa perkuliahan. Bantuan finansial yang diberikan, terutama dalam pembiayaan perkuliahan, sangatlah berarti dan menjadi penunjang utama bagi penulis untuk dapat fokus dan menyelesaikan studi ini.
3. Bapak Irwan Budiman., S.T., M.Kom dan Bapak Muhammad Itqan Mazdadi., S.Kom., M.Kom selaku dosen pembimbing atas kesabaran, pengarahan, saran dan meluangkan waktunya untuk penulis dalam penyelesaian tugas akhir.
4. Terima kasih sebesar-besarnya kepada Ibu Dwi Kartini, S.Kom., M.Kom., selaku Koordinator Program Studi Ilmu Komputer FMIPA ULM, atas kebijakan dan solusi krusial yang memungkinkan penulis untuk dapat melanjutkan dan menyelesaikan sidang skripsi ini.
5. Seluruh Dosen dan Staff Prodi Ilmu Komputer atas ilmu dan bantuan yang diberikan selama perkuliahan hingga penyelesaian studi.
6. Teman-teman di RyzenGen'21, dan sahabat seperjuangan atas semangat dan

dukungan.

7. Semua pihak yang tidak dapat disebutkan satu persatu yang telah turut membantu dalam penyelesaian skripsi ini.

Akhir kata penulis menyadari sepenuhnya bahwa penulisan ini jauh lebih dari sempurna dan penulis mengharapkan bantuan berupa kritik dan saran yang membangun dari semua pihak demi kesempurnaan dan mutu penulisan skripsi ini. Semoga tulisan ini dapat bermanfaat bagi ilmu pengetahuan dan pembaca khususnya serta mendapat keridhaan Allah SWT.

Banjarbaru, 2 November 2025



Thaha Wafiq Adly

DAFTAR ISI

	Halaman
HALAMAN PENGESAHAN	i
PERNYATAAN	iii
ABSTRAK	iv
ABSTRACT	v
PRAKATA	vi
DAFTAR ISI	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	4
1.3. Tujuan Penelitian	4
1.4. Manfaat Penelitian.....	4
1.5. Batasan Masalah	5
BAB II TINJAUAN PUSTAKA	6
2.1 Kajian Terdahulu.....	6
2.2 Landasan Teori	15
2.2.1 Data Mining.....	15
2.2.2 Clustering	15
2.2.3 Algoritma K-Means	16
2.2.4 <i>Sequential Forward Selection (SFS)</i>	18
2.2.5 MinMax Scaler	19
2.2.6 Elbow Method	19
2.2.7 Silhoutte Score.....	20
2.2.8 ANOVA (Analysis of Variance).....	21
2.2.9 Tuskey HSD.....	21
2.2.10 <i>Adjusted Rand Index (ARI)</i>	22

BAB III METODE PENELITIAN	23
3.1. Alat Penelitian	23
3.2. Bahan Penelitian	23
3.3. Variabel Penelitian	24
3.4. Prosedur Penelitian	25
BAB IV HASIL DAN PEMBAHASAN	30
4.1 Hasil	30
4.1.1 Dataset dan Fitur yang Digunakan	30
4.1.2 Seleksi Fitur	31
4.1.3 Preprocessing Data	36
4.1.4 Pemilihan Jumlah K	37
4.1.5 Clustering	38
4.1.6 Silhoutte Score	39
4.1.7 ANOVA	41
4.1.8 Tukey HSD	42
4.1.9 Adjusted Rand Index (ARI)	43
4.2 Pembahasan	45
BAB V PENUTUP	64
5.1 Kesimpulan	64
5.2 Saran	64
DAFTAR PUSTAKA	66
LAMPIRAN	70
RIWAYAT PENULIS	92

DAFTAR TABEL

Tabel	Halaman
Table 1.Keaslian Penelitian.....	10
Table 2.Perancangan Penelitian.....	14
Table 3. Fitur Dataset NBA 2023-2024	30
Table 4. Seleksi Fitur Iterasi 1	31
Table 5. Seleksi Fitur Iterasi 2.....	32
Table 6. Seleksi Fitur Iterasi 3.....	33
Table 7. Seleksi Fitur Iterasi 4.....	34
Table 8. 10 Pemain Dataset dengan Fitur yang terpilih.....	36
Table 9. Statistik Deskriptif Fitur terpilih.....	36
Table 10. Data pemain Sesudah MinMax	37
Table 11. Cluster Pemain.....	39
Table 12. Silhoutte Score per K.....	40
Table 13. Silhoutte Score Per Cluster	41
Table 14. Fitur DRtg	42
Table 15. Fitur ORtg	42
Table 16. Fitur TO%	43
Table 17. Fitur USG%	43
Table 18. Iterasi Adjusted Rand Index	44
Table 19. Rata-rata Adjusted Rand Index.....	45
Table 20. Tukey HSD Fitur DRtg	53
Table 21. Tukey HSD ORtg.....	53
Table 22. Tukey HSD fitur TO%	54
Table 23. Tukey HSD fitur USG%	54
Table 24. ORtg dan TO% per Cluster	56

DAFTAR GAMBAR

Gambar	Halaman
Gambar 1 Flowchart Algoritma K-Means (Hutagalung & Sonata, 2021).....	17
Gambar 2. Alur Penelitian Menggunakan K-Means.....	25
Gambar 4. K optimal Elbow Method.....	38
Gambar 5. K optimal Berdasarkan Silhoutte Score	40
Gambar 6. Adjusted Rand Index.....	44
Gambar 7. Grafik SFS 10 Iterasi.....	46
Gambar 8. Cluster tanpa MinMax.....	48
Gambar 9. Cluster Dengan MinMax.....	49
Gambar 10. Perbedaan Elbow Method dan Silhoutte Score.....	51
Gambar 11. Visualisasi Perbandingan Rata-rata Statistik Per Cluster.....	62

DAFTAR LAMPIRAN

Lampiran	Halaman
Lampiran 1. Pengaturan Awal dan Impor Pustaka.....	71
Lampiran 2. Konfigurasi Parameter Utama	71
Lampiran 3. Fungsi Visualisasi - Bar Char	71
Lampiran 4. Fungsi Visualisasi - SFS Progress	72
Lampiran 5. Fungsi Sequential Feature Selection (SFS)	74
Lampiran 6. Fungsi Visualisasi - Elbow & Silhouette.....	78
Lampiran 7. Fungsi Visualisasi - Radar Chart	79
Lampiran 8. Fungsi Analisis Stabilitas (ARI)	80
Lampiran 9. Fungsi Evaluasi Cluster	82
Lampiran 10. Fungsi Analisis Clustering Utama	85
Lampiran 11. Fungsi Main - Load Data	90