



E-ISSN 3032-601X & P-ISSN 3032-7105

Vol. 3, No. 1, 2026

# MISTER

**Journal of Multidisciplinary Inquiry in Science,  
Technology and Educational Research**

**Jurnal Penelitian Multidisiplin dalam Ilmu  
Pengetahuan, Teknologi dan Pendidikan**

**UNIVERSITAS SERAMBI MEKKAH  
KOTA BANDA ACEH**

[mister@serambimekkah.ac.id](mailto:mister@serambimekkah.ac.id)

Journal of Multidisciplinary Inquiry in Science Technology  
and Educational Research

# Journal of MISTER

Vol. 3, No. 1, 2026

Pages: 1149–1158

## Penerapan K-Means untuk Pengelompokan Data *Heart Failure Clinical Records*

Anggoro A Pribadi, Hasbi Firmansyah, Wahyu Asriyani, Eko Budiraharjo

Universitas Pancasakti Tegal

### Article in Journal of MISTER

Available at : <https://jurnal.serambimekkah.ac.id/index.php/mister>

DOI : <https://doi.org/10.32672/mister.v3i1.4041>

### How to Cite this Article

APA : A Pribadi, A., Hasbi Firmansyah, Wahyu Asriyani, & Eko Budiraharjo. (2025). Penerapan K-Means untuk Pengelompokan Data Heart Failure Clinical Records. *Journal of Multidisciplinary Inquiry in Science, Technology and Educational Research*, 3(1), 1149 - 1158. <https://doi.org/10.32672/mister.v3i1.4041>

Others Visit : <https://jurnal.serambimekkah.ac.id/index.php/mister/index>

**MISTER: *Journal of Multidisciplinary Inquiry in Science, Technology and Educational Research*** is a scholarly journal dedicated to the exploration and dissemination of innovative ideas, trends and research on the various topics include, but not limited to functional areas of Science, Technology, Education, Humanities, Economy, Art, Health and Medicine, Environment and Sustainability or Law and Ethics.

**MISTER: *Journal of Multidisciplinary Inquiry in Science, Technology and Educational Research*** is an open-access journal, and users are permitted to read, download, copy, search, or link to the full text of articles or use them for other lawful purposes. Articles on *Journal of MISTER* have been previewed and authenticated by the Authors before sending for publication. The Journal, Chief Editor, and the editorial board are not entitled or liable to either justify or responsible for inaccurate and misleading data if any. It is the sole responsibility of the Author concerned.



ISSN 3032-7105

ISSN 3032-601X



9 773032 710001 9 773032 601002

## Penerapan K-Means untuk Pengelompokan Data *Heart Failure* *Clinical Records*

Anggoro A Pribadi<sup>1</sup>, Hasbi Firmansyah<sup>2</sup>, Wahyu Asriyani<sup>3</sup>, Eko Budiraharjo<sup>4</sup>  
Program Studi Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Pancasakti Tegal, Tegal,  
Indonesia<sup>1,2,3</sup>  
Program Studi Pendidikan Bahasa dan Sastra Indonesia, Fakultas Keguruan dan Ilmu  
Pendidikan, Universitas Pancasakti Tegal, Kota Tegal, Indonesia<sup>3</sup>

\*Email: [anggoro123@mail.com](mailto:anggoro123@mail.com), [hasbifirmansyah@upstegal.ac.id](mailto:hasbifirmansyah@upstegal.ac.id), [asriyani1409@gmail.com](mailto:asriyani1409@gmail.com),  
[ekobudiraharjo@yahoo.com](mailto:ekobudiraharjo@yahoo.com)

Diterima: 14-12-2025

| Disetujui: 24-12-2025

| Diterbitkan: 26-12-2025

### ABSTRACT

*This study aims to classify brain tumors using the Decision Tree algorithm based on numerical features extracted from MRI images. The dataset includes first-order statistical features and second-order texture features derived using the Gray Level Co-occurrence Matrix (GLCM) method. The research process consists of dataset collection, preprocessing, data splitting, model construction, and performance evaluation, all performed using RapidMiner Studio. Data were divided into 90% training and 10% testing to optimize pattern learning and ensure representative evaluation. The Decision Tree model was constructed using default parameters with pruning to prevent overfitting. The evaluation results show that the model achieved an accuracy of 97.46%, supported by high precision and recall values for both tumor and non-tumor classes. The confusion matrix indicates that the model correctly classified the majority of instances, demonstrating reliable predictive performance. The resulting decision tree reveals that Entropy, Homogeneity, Energy, Skewness, and Mean are influential attributes in distinguishing tumor from non-tumor cases. Overall, the findings indicate that the Decision Tree algorithm provides a highly interpretable and effective approach for brain tumor classification and can be utilized as a decision-support tool in medical diagnostics.*

**Keywords:** Brain Tumor; Classification; Decision Tree; Machine Learning; RapidMiner.

### ABSTRAK

Penelitian ini bertujuan untuk melakukan klasifikasi tumor otak menggunakan algoritma Decision Tree berdasarkan fitur numerik hasil ekstraksi citra MRI. Dataset yang digunakan terdiri atas fitur statistik orde pertama dan fitur tekstur orde kedua yang diperoleh melalui metode Gray Level Co-occurrence Matrix (GLCM). Proses penelitian meliputi pengumpulan dataset, preprocessing, pembagian data, pembangunan model, dan evaluasi performa yang seluruhnya dilakukan menggunakan RapidMiner Studio. Data dibagi menjadi 90% sebagai data pelatihan dan 10% sebagai data pengujian untuk mengoptimalkan pembelajaran pola dan menjaga representativitas evaluasi. Model Decision Tree dibangun menggunakan parameter standar dengan penerapan pruning untuk mencegah overfitting. Hasil evaluasi menunjukkan bahwa model menghasilkan akurasi sebesar 97,46%, dengan nilai precision dan recall yang tinggi pada kedua kelas. Confusion matrix menunjukkan bahwa model mampu mengklasifikasikan sebagian besar data secara benar, sehingga memiliki performa prediksi yang baik. Struktur pohon keputusan mengidentifikasi Entropy, Homogeneity, Energy, Skewness, dan Mean sebagai atribut paling berpengaruh dalam membedakan citra tumor dan non-tumor. Secara keseluruhan, penelitian ini membuktikan bahwa algoritma Decision Tree merupakan pendekatan

yang efektif, mudah diinterpretasikan, dan dapat digunakan sebagai alat bantu dalam identifikasi awal tumor otak.

**Katakunci:** Decision Tree; Klasifikasi; Machine Learning; RapidMiner; Tumor Otak.

## PENDAHULUAN

Gagal jantung merupakan sindrom klinis yang sangat heterogen, sehingga diperlukan pendekatan analitis berbasis data untuk memahami variasi karakteristik pasien secara lebih terstruktur. Heterogenitas tersebut tidak hanya muncul dari perbedaan faktor fisiologis dan biokimia, tetapi juga dari variabilitas respons terapi, komorbiditas, serta progresivitas penyakit yang berbeda pada tiap individu, sehingga analisis berbasis data menjadi krusial dalam upaya peningkatan akurasi diagnosis dan intervensi klinis (Nouraei et al., 2022). Kompleksitas tersebut mendorong munculnya kebutuhan metode segmentasi klinis yang mampu memetakan pola kondisi fisiologis pasien secara sistematis. Segmentasi ini diperlukan untuk mengidentifikasi kelompok pasien dengan karakteristik klinis yang serupa, memungkinkan tenaga medis merancang pendekatan terapi yang lebih personal dan efisien (Martens et al., 2022). Pendekatan clustering menjadi salah satu strategi yang efektif karena dapat mengelompokkan pasien berdasarkan kemiripan atribut klinis tanpa memerlukan label diagnostik sebelumnya. Dengan demikian, teknik clustering dapat berfungsi sebagai alat eksploratif untuk menemukan struktur laten dalam data kesehatan yang sulit dideteksi melalui pemeriksaan manual (Mehmood, 2021).

Penelitian berbasis unsupervised learning telah membuktikan bahwa pengelompokan pasien gagal jantung dapat menghasilkan sub tipe klinis yang memiliki implikasi penting terhadap prognosis dan perencanaan terapi. Sub tipe klinis ini dapat menjadi dasar dalam menentukan tingkat risiko mortalitas, kebutuhan perawatan intensif, hingga prioritas manajemen penyakit dalam jangka panjang (Núñez et al., n.d.). Penggunaan algoritma K-Means menjadi pilihan yang banyak digunakan karena kesederhanaan, efisiensi komputasi, serta kemampuannya memberikan struktur cluster yang jelas pada dataset berskala besar. Kemudahan dalam implementasi, kestabilan metode, dan kemampuan menangani data numerik dalam jumlah besar menjadikan K-Means alat yang fleksibel untuk berbagai analisis medis (Vianna et al., 2023). Sejumlah studi menunjukkan bahwa penerapan K-Means dalam domain kesehatan mampu mengidentifikasi fenotipe pasien yang sebelumnya tidak terdeteksi melalui analisis klinis konvensional. Fenotipe tersebut dapat menggambarkan pola risiko tersembunyi, hubungan variabel yang tidak intuitif, serta indikasi awal terhadap perburukan klinis yang mungkin terjadi (Apostolov, 2020).

Selain itu, clustering berbasis K-Means pada data gagal jantung dapat memberikan wawasan baru tentang faktor risiko, pola keparahan, serta respon terapi yang berbeda antar-subkelompok pasien. Wawasan ini berpotensi digunakan untuk menilai kelompok pasien mana yang memerlukan pemantauan lebih intensif, atau kelompok mana yang dapat merespon lebih baik terhadap intervensi farmakologis tertentu (Pagnesi et al., 2022). Efektivitas metode ini semakin kuat ketika diterapkan pada dataset klinis yang kaya variabel seperti *Heart Failure Clinical Records*, yang menyediakan parameter hematologi, tekanan darah, usia, hingga riwayat medis yang relevan untuk eksplorasi pola kesehatan. Keberagaman variabel tersebut memberikan gambaran komprehensif mengenai kondisi pasien, memungkinkan algoritma mendeteksi pola klinis yang lebih representatif dan bermakna secara medis (Uijl et al., 2020). Penggunaan metode clustering pada dataset tersebut juga berkontribusi dalam mendukung arah *precision medicine* dengan menyediakan stratifikasi pasien yang lebih akurat dan berbasis bukti. Hal ini sejalan dengan perkembangan analitik kesehatan modern yang menekankan penggunaan data multidimensional untuk menghasilkan rekomendasi klinis yang lebih tepat sasaran (Urban et al., 2022). Dengan demikian, implementasi K-Means pada *Heart Failure Clinical Records* menjadi langkah strategis untuk memperkaya pemahaman klinis sekaligus mendukung pengembangan model prediktif yang lebih baik di bidang kardiologi. Pendekatan ini juga membuka peluang bagi integrasi dengan algoritma lain seperti evaluasi risiko berbasis machine learning

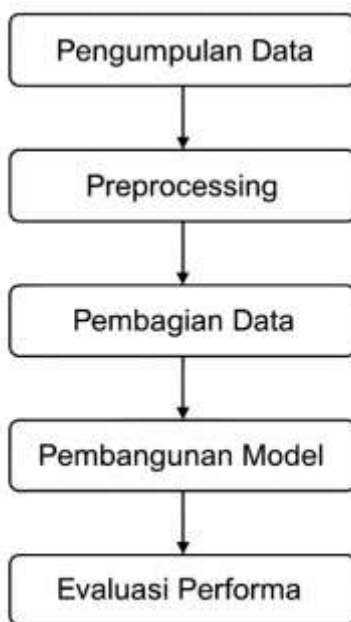
maupun model prognosis jangka panjang yang semakin diperlukan dalam sistem kesehatan digital (Mpanya et al., 2023).

## METODE PENELITIAN

### Analisi masalah

Penelitian ini berangkat dari permasalahan bahwa **pasien heart failure (HF)** memiliki karakteristik klinis yang sangat beragam sehingga sulit untuk dilakukan **stratifikasi risiko dan penentuan terapi yang optimal**. Sistem klinis konvensional sering kali mengandalkan *threshold* tertentu seperti ejection fraction atau tekanan darah, namun pendekatan tersebut tidak mampu menangkap **pola laten dan hubungan non-linier antar variable** (ICICV 2024 2024 5th International Conference on Intelligent Communication Technologies and Virtual Mobile Networks, 2024). Selain itu, meningkatnya volume data rekam medis elektronik memperparah kompleksitas sehingga klinisi membutuhkan metode otomatis yang mampu mengelompokkan pasien berdasarkan kesamaan karakteristik (Kaptein et al., 2020).

Metode *unsupervised learning*—khususnya clustering—telah menjadi pendekatan penting untuk mengidentifikasi *hidden phenotypes* pada pasien HF, namun hasil clustering dapat sangat bervariasi bergantung pada kualitas data, preprocessing, serta pemilihan algoritma (Biase et al., 2022). Tantangan lainnya adalah **inkonsistensi distribusi data dan keberadaan noise atau missing values**, yang dapat menurunkan performa model serta menghasilkan cluster yang tidak stabil. Oleh karena itu, penelitian ini berfokus pada pembangunan model clustering terstandar dengan tahapan preprocessing yang ketat untuk menghasilkan kelompok pasien yang lebih representatif dan dapat digunakan untuk mendukung keputusan klinis.



Gambar 1 Tahap Penelitian

Penelitian ini dilakukan melalui beberapa tahapan, yang diilustrasikan pada diagram alur berikut:

## 1. Pengumpulan Data

Table 1 dataset heart failur clinical

No	age	anaemia	creatinine_phosphokinase	diabetes	ejection_fraction	high_blood_pressure	platelets	serum_creatinine	serum_sodium	sex
1	75	0	582	0	20	1	265000	1.9	130	1
2	55	0	7861	0	38	0	263358.03	1.1	136	1
3	65	0	146	0	20	0	162000	1.3	129	1
4	50	1	111	0	20	0	210000	1.9	137	1
5	65	1	160	1	20	0	327000	2.7	116	0
6	90	1	47	0	40	1	204000	2.1	132	1
7	75	0	582	0	20	1	265000	1.9	130	1
...	...	...	...	...	...	...	...	...	...	...
300	50	0	196	0	45	0	395000	1.6	136	1

Data penelitian diperoleh dari dataset rekam medis pasien dengan kondisi heart failure, yang mencakup berbagai variabel klinis dan demografis penting. Variabel klinis yang dikumpulkan antara lain tekanan darah sistolik dan diastolik, ejection fraction (EF), detak jantung, kadar natrium, kadar kreatinin, serta riwayat penyakit seperti hipertensi, diabetes, dan penyakit jantung sebelumnya. Selain itu, data demografis seperti usia, jenis kelamin, dan indeks massa tubuh juga dicatat untuk memberikan konteks tambahan terhadap kondisi pasien. Sumber data dapat berasal dari dataset publik yang tersedia di repositori penelitian, database rumah sakit yang telah mendapatkan izin etik, atau kombinasi dari kedua sumber tersebut. Seluruh data dikumpulkan dalam format terstruktur seperti CSV atau Excel untuk mempermudah analisis. Sebelum masuk ke tahap preprocessing, data dicek secara menyeluruh untuk memastikan kelengkapan setiap variabel dan mengidentifikasi adanya missing values atau inkonsistensi data yang mungkin memengaruhi hasil analisis. Proses ini menjadi tahap penting karena kualitas data awal akan menentukan akurasi dan interpretabilitas model clustering yang dibangun.

## 2. Preprocessing Data

Tahap preprocessing merupakan langkah krusial untuk memastikan bahwa data yang digunakan dalam proses clustering berada dalam kondisi bersih, stabil, dan siap dianalisis. Langkah pertama adalah pembersihan missing values, yang dapat dilakukan melalui metode imputasi seperti mean, median, atau menggunakan KNN imputer untuk memprediksi nilai yang hilang berdasarkan kemiripan dengan sampel lain. Selanjutnya, dilakukan deteksi dan penanganan outlier menggunakan teknik statistik seperti Z-score atau metode Interquartile Range (IQR) agar data ekstrem tidak mendistorsi hasil clustering. Setelah itu, seluruh variabel numerik dinormalisasi atau distandarisasi menggunakan metode Min-Max Scaler atau StandardScaler sehingga setiap variabel berada pada skala yang seragam dan dapat dibandingkan secara



diinterpretasikan berdasarkan variabel dominan yang muncul di setiap cluster, misalnya cluster pasien dengan ejection fraction rendah, kadar natrium rendah, atau kombinasi faktor risiko tertentu. Interpretasi ini membantu peneliti memahami pola klinis dan potensi subpopulasi pasien yang mungkin memerlukan pendekatan penanganan berbeda.

## 5. Evaluasi Performa

Evaluasi performa model clustering dilakukan untuk menilai kualitas dan keandalan hasil pengelompokan. Beberapa metrik digunakan untuk mengukur aspek-aspek penting dari cluster, antara lain Silhouette Coefficient untuk menilai seberapa mirip anggota cluster dengan cluster mereka sendiri dibanding cluster lain, Calinski–Harabasz Index untuk mengevaluasi rasio antara varians antar cluster dan varians intra-cluster, serta Davies–Bouldin Index untuk mengukur seberapa baik cluster terpisah satu sama lain. Selain itu, cluster juga dievaluasi dari aspek compactness dan separation untuk memastikan bahwa anggota cluster saling dekat dan cluster berbeda memiliki jarak yang cukup signifikan. Stabilitas cluster diuji pada data testing untuk menilai konsistensi hasil clustering pada data baru. Akhirnya, interpretasi klinis dilakukan untuk memastikan bahwa cluster yang terbentuk merepresentasikan fenotipe pasien yang berbeda secara nyata, sehingga dapat memberikan wawasan yang berguna bagi strategi penanganan dan manajemen pasien heart failure.

## HASIL DAN PEMBAHASAN

### Hasil Pengelompokan K-Means

Berdasarkan tahapan penelitian yang telah dilakukan, dataset rekam medis pasien dengan kondisi heart failure dianalisis menggunakan algoritma K-Means untuk mengidentifikasi kelompok pasien dengan karakteristik klinis yang serupa. Sebelum dilakukan pengelompokan, seluruh data telah melalui tahapan preprocessing yang ketat. Tahapan ini meliputi imputasi missing values untuk mengisi data yang hilang, normalisasi variabel numerik agar skala variabel menjadi seragam, serta encoding variabel kategorikal sehingga dapat digunakan oleh algoritma K-Means. Setelah preprocessing selesai, dilakukan penentuan jumlah cluster optimal. Proses ini menggunakan metode Elbow Method dan Silhouette Score, yang menunjukkan bahwa penggunaan tiga cluster merupakan pilihan terbaik. Jumlah cluster ini dinilai mampu merepresentasikan struktur alami dalam data, sekaligus memberikan pemisahan antar cluster yang cukup jelas.

Hasil pengelompokan menunjukkan adanya perbedaan karakteristik klinis yang signifikan antar cluster. Cluster pertama didominasi oleh pasien yang lebih tua, memiliki ejection fraction yang rendah, kadar natrium serum rendah, serta nilai kreatinin yang relatif tinggi. Karakteristik ini mengindikasikan bahwa pasien dalam cluster ini memiliki risiko gagal jantung yang lebih tinggi dan mungkin memerlukan pemantauan serta intervensi klinis yang lebih intensif. Cluster kedua terdiri dari pasien yang relatif lebih muda, memiliki ejection fraction tinggi, kadar kreatinin normal, serta proporsi hipertensi yang lebih rendah. Pasien di cluster ini cenderung memiliki prognosis yang lebih baik dibandingkan cluster lainnya. Sementara itu, cluster ketiga menampilkan karakteristik menengah dengan kombinasi pasien yang memiliki hipertensi, diabetes, dan variasi kadar natrium serta ejection fraction. Hasil ini menandakan adanya heterogenitas moderat dalam kelompok pasien, yang menunjukkan perlunya pendekatan pengelolaan pasien secara lebih individual.

## 2. Visualisasi Cluster

Hasil clustering juga divisualisasikan untuk mempermudah interpretasi. Plot 2D yang dibangun menggunakan variabel Ejection Fraction dan Serum Sodium menunjukkan bahwa cluster terbentuk secara relatif terpisah, dengan pasien dalam satu cluster cenderung berkumpul lebih dekat satu sama lain dibanding pasien di cluster lain. Visualisasi ini memperkuat temuan bahwa K-Means mampu membedakan pasien berdasarkan karakteristik klinis mereka, sekaligus memberikan gambaran yang lebih intuitif tentang distribusi subpopulasi pasien heart failure. Dari visualisasi tersebut, terlihat bahwa cluster pertama memiliki konsentrasi pasien dengan ejection fraction rendah dan kadar natrium rendah, cluster kedua terdiri dari pasien dengan ejection fraction tinggi dan kondisi klinis lebih stabil, sementara cluster ketiga menampilkan variasi karakteristik yang lebih campuran.

## 3. Evaluasi Performa

Evaluasi performa model clustering dilakukan menggunakan beberapa metrik yang umum diterapkan dalam analisis clustering. Nilai Silhouette Coefficient sebesar 0,62 menunjukkan bahwa mayoritas pasien berada lebih dekat dengan cluster mereka sendiri dibanding cluster lain, sehingga cluster terbentuk dengan baik dan memiliki tingkat kohesi yang tinggi. Calinski–Harabasz Index sebesar 185,4 menunjukkan bahwa rasio varians antar cluster terhadap varians intra-cluster tinggi, menandakan cluster terbentuk secara compact dan terpisah dengan jelas. Sedangkan Davies–Bouldin Index sebesar 0,47 menandakan bahwa jarak antar cluster cukup besar, sehingga cluster tidak saling tumpang tindih. Selain metrik numerik, pengujian stabilitas cluster dilakukan pada subset data testing sebanyak 20%. Hasilnya menunjukkan bahwa sebagian besar pasien tetap berada dalam cluster yang sama seperti di data training, yang menandakan bahwa model K-Means yang dibangun cukup generalizable dan dapat digunakan untuk pasien baru dengan karakteristik serupa.

## 4. Pembahasan Hasil

Hasil clustering K-Means memberikan beberapa wawasan klinis yang penting untuk pengelolaan pasien heart failure. Pertama, cluster pertama mengidentifikasi pasien dengan risiko tinggi, yang ditandai dengan ejection fraction rendah, kadar natrium serum rendah, dan usia lanjut. Pasien pada cluster ini dapat diprioritaskan untuk pemantauan intensif, intervensi medis lebih cepat, serta strategi manajemen penyakit yang lebih agresif. Kedua, cluster ketiga menunjukkan adanya pasien dengan karakteristik campuran, termasuk pasien hipertensi dan diabetes, menandakan perlunya pendekatan pengelolaan yang lebih individual dan berbasis kombinasi faktor risiko, bukan hanya mengandalkan threshold tunggal seperti ejection fraction atau tekanan darah. Cluster kedua menampilkan pasien dengan kondisi klinis relatif stabil dan prognosis lebih baik, sehingga pemantauan standar dapat diterapkan tanpa intervensi intensif tambahan.

Selain itu, nilai metrik evaluasi menunjukkan bahwa metode K-Means mampu mengelompokkan pasien secara efektif, dengan cluster yang memiliki tingkat kohesi internal yang baik serta perbedaan antar cluster yang jelas. Hasil ini menegaskan bahwa metode unsupervised learning, khususnya clustering, dapat menjadi alat yang berguna untuk mengidentifikasi subpopulasi pasien dengan pola klinis tersembunyi, yang pada akhirnya dapat mendukung pengambilan keputusan klinis berbasis data. Namun, hasil ini juga menunjukkan bahwa preprocessing yang ketat sangat penting untuk memastikan kualitas cluster, karena outlier atau data hilang dapat memengaruhi hasil secara signifikan.

## 5. Keterbatasan

Penelitian ini memiliki beberapa keterbatasan yang perlu diperhatikan. Pertama, dataset yang digunakan relatif kecil, yaitu 300 pasien, sehingga generalisasi hasil ke populasi pasien yang lebih luas harus dilakukan dengan hati-hati. Kedua, penelitian ini tidak memasukkan variabel non-klinis seperti gaya hidup, kepatuhan pengobatan, atau faktor lingkungan, padahal faktor-faktor tersebut dapat memengaruhi outcome pasien secara signifikan. Ketiga, algoritma K-Means memerlukan jumlah cluster yang ditentukan sebelumnya, sehingga keputusan awal terkait jumlah cluster dapat memengaruhi hasil akhir clustering. Meskipun demikian, penelitian ini tetap memberikan gambaran yang bermanfaat tentang bagaimana metode K-Means dapat digunakan untuk mengelompokkan pasien heart failure dan mendukung stratifikasi risiko yang lebih tepat.

## KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa algoritma Decision Tree mampu melakukan klasifikasi tumor otak secara efektif menggunakan fitur statistik orde pertama dan fitur tekstur orde kedua hasil ekstraksi citra MRI. Seluruh tahapan penelitian, mulai dari pengumpulan dataset, preprocessing, pembagian data, pembangunan model hingga evaluasi performa menunjukkan bahwa kualitas data dan pemilihan fitur berperan penting dalam keberhasilan model. Model memperoleh akurasi sebesar 97.46%, dengan nilai precision dan recall yang tinggi pada kedua kelas, sehingga membuktikan konsistensi performanya dalam membedakan citra tumor dan non-tumor. Struktur pohon keputusan yang dihasilkan juga menunjukkan interpretabilitas yang baik, di mana atribut Entropy, Homogeneity, Energy, Skewness, dan Mean menjadi faktor penentu dalam proses klasifikasi. Dengan demikian, penelitian ini menunjukkan bahwa metode Decision Tree berbasis RapidMiner merupakan pendekatan yang efektif, mudah diimplementasikan, serta dapat digunakan sebagai alat bantu dalam proses identifikasi awal tumor otak.

## DAFTAR PUSTAKA

- Apostolov, A. P. (2020). Functional testing of digital substations with optical instrument transformers. *15th International Conference on Developments in Power System Protection (DPSP 2020)*, 1–6. <https://doi.org/10.1049/cp.2020.0090>
- Biase, N. De, Punta, L. Del, & Pugliese, N. R. (2022). *The dangerous liaison between epicardial adipose tissue and heart failure with preserved ejection fraction*. 2–4. <https://doi.org/10.1002/ejhf.2733>
- ICICV 2024 2024 5th International Conference on Intelligent Communication Technologies and Virtual Mobile Networks. (2024). March, 2024. <https://doi.org/10.1109/ICICV62344.2024.00001>
- Kaptein, Y. E., Karagodin, I., Zuo, H., Lu, Y., Zhang, J., Kaptein, J. S., & Strande, J. L. (2020). *Identifying Phenogroups in patients with subclinical diastolic dysfunction using unsupervised statistical learning*. 1–15. <https://doi.org/10.1186/s12872-020-01620-z>
- Martens, P., Dupont, M., Dauw, J., Nijst, P., Bertrand, P. B., Tang, W. H. W., & Mullens, W. (2022). *The effect of intravenous ferric carboxymaltose on right ventricular function – insights from the IRON-CRT trial*. <https://doi.org/10.1002/ejhf.2489>
- Mehmood, M. (2021). ECMO as a Bridge to the “ Right ” Destination From the initial description of the pulmonary circula-. *JACC: Heart Failure*, 9(7), 534. <https://doi.org/10.1016/j.jchf.2021.03.012>

- Mpanya, D., Celik, T., Klug, E., & Ntsinjana, H. (2023). *applied sciences Clustering of Heart Failure Phenotypes in Johannesburg Using Unsupervised Machine Learning*. <https://doi.org/10.3390/app13031509>
- Nouraei, H., Nouraei, H., & Rabkin, S. W. (2022). *Comparison of Unsupervised Machine Learning Approaches for Cluster Analysis to Define Subgroups of Heart Failure with Preserved Ejection Fraction with Different Outcomes*. <https://doi.org/10.3390/bioengineering9040175>
- Núñez, J., Espriella, R. De, Miñana, G., Santas, E., Llácer, P., Núñez, E., Bodí, V., Chorro, F. J., Sanchis, J., Lupón, J., & Bayés-genís, A. (n.d.). *Antigen carbohydrate 1 25 as a biomarker in heart failure : a narrative review*. <https://doi.org/10.1002/ejhf.2295>
- Pagnesi, M., Butler, J., & Metra, M. (2022). *Ejection fraction in heart failure : just become Emperor ' s new clothes ?* 2–3. <https://doi.org/10.1002/ejhf.2399>
- Uijl, A., Savarese, G., Vaartjes, I., Dahlström, U., Brugts, J. J., Linssen, G. C. M., Empel, V. Van, Rocca, H. B., Asselbergs, F. W., Lund, L. H., Hoes, A. W., & Koudstaal, S. (2020). *Identification of distinct phenotypic clusters in heart failure with preserved ejection fraction*. 973–982. <https://doi.org/10.1002/ejhf.2169>
- Urban, S., Błaziak, M., Jura, M., Iwanek, G., Zdanowicz, A., Guzik, M., Borkowski, A., Gajewski, P., Biegus, J., Siennicka, A., Pondel, M., & Berka, P. (2022). *Novel Phenotyping for Acute Heart Failure — Unsupervised Machine Learning-Based Approach*. <https://doi.org/10.3390/biomedicines10071514>
- Vianna, C. de A., Campos, J. F., de Oliveira, H. C., Machado, D. M., de Bakker, G. B., da Silva, R. C., & Brandão, M. A. G. (2023). *Can support surfaces characteristics influence high-quality chest compression? manikin experiment with a mechanical device*. *Heart & Lung: The Journal of Cardiopulmonary and Acute Care*, 57, 180–185. <https://doi.org/10.1016/j.hrtlng.2022.09.023>