

## Analisis Rekam Medis untuk Prediksi Keberlangsungan Hidup Pasien Gagal Jantung dengan Metode Data Mining

Ade Irma Rizmayanti<sup>1\*</sup>, Nurril Cholifatul Izza<sup>2</sup>

<sup>1,2</sup> Politeknik Karya Husada, Jl. Margonda Raya No. 28 Kel. Pondok Cina, Kec. Beji  
Kota Depok, Jawa Barat 16424

<sup>1</sup>ade.irma.r@gmail.com\*; <sup>2</sup>nurril.cholifatul@gmail.com

### Abstrak

Penelitian tentang analisis rekam medis pasien gagal jantung merupakan sebuah isu kesehatan yang kompleks dengan implikasi luas di berbagai sektor. Fokus utama adalah untuk menggali pengetahuan dari data yang besar dan kompleks, menggunakan algoritma ensemble seperti *Random forest*, *Extreme Gradient Boosting*, *Extra Tree*, dan *AdaBoost*. Hasilnya *Random forest* menunjukkan performa terbaik dengan akurasi 0,84% dan nilai AUC 0,89 dari 299 data rekam medis yang dijadikan sample pada penelitian ini, hal ini mengindikasikan efektivitas tinggi dalam klasifikasi pasien berdasarkan potensi keberlangsungan hidup, data mining dapat secara signifikan mendukung pengambilan keputusan medis berbasis bukti dan meningkatkan manajemen penyakit gagal jantung, dengan memberikan wawasan yang lebih dalam melalui identifikasi pola dan korelasi dalam data kesehatan. Pendekatan ini memungkinkan peningkatan kualitas perawatan pasien dan memberikan rekomendasi metodologi untuk praktik klinis di masa depan.

**Kata kunci:** Gagal Jantung; Rekam Medis; Data Mining; *Random forest*

### Abstract

*Research on medical record analysis of heart failure patients is a complex health issue with broad implications in various sectors. The main focus is to extract knowledge from large and complex data, using ensemble algorithms such as Random forest, Extreme Gradient Boosting, Extra Tree, and AdaBoost. The results of Random forest show the best performance with an accuracy of 0.84% and an AUC value of 0.89 from 299 medical record data sampled in this study, this indicates high effectiveness in classifying patients based on potential survival, data mining can significantly support the decision making. evidence-based medical decisions and improve heart failure disease management, by providing deeper insights through identifying patterns and correlations in health data. This approach allows improving the quality of patient care and provides methodological recommendations for future clinical practice.*

**Keywords:** Heart failure; Medical records; Data Mining; *Random forest*

### PENDAHULUAN

Isu kesehatan adalah masalah yang sangat rumit yang memiliki hubungan timbal balik dengan berbagai sektor di luar kesehatan, termasuk sosial, ekonomi, pendidikan, demografi, teknologi, dan lainnya (Salim & Sugeng, 2018). Oleh karena itu, penanganan masalah kesehatan publik harus mempertimbangkan aspek-aspek kesehatan secara menyeluruh dan

dampaknya (Ishaq et al., 2021). Saat ini, kita menghadapi masalah double burden disease, di mana terjadi peningkatan kasus penyakit tidak menular atau degeneratif, sementara penyakit menular juga masih bertambah (Salim & Sugeng, 2018).

Gagal jantung merupakan salah satu penyakit kardiovaskular yang paling serius dan menjadi penyebab utama morbiditas dan mortalitas di seluruh dunia (Moreno-Sanchez, 2020). Dengan prevalensi yang terus meningkat seiring dengan penuaan populasi, terdapat kebutuhan mendesak untuk mengembangkan metode yang lebih efektif dalam diagnosis, pengelolaan, dan pencegahan penyakit ini (Pratama et al., 2022). Data rekam medis pasien gagal jantung menyimpan informasi klinis yang berharga, namun kompleksitas dan volume data ini seringkali menjadi tantangan dalam analisis manual (Kim et al., 2020).

Dalam dekade terakhir, data mining telah muncul sebagai alat yang sangat bermanfaat dalam menggali pengetahuan dari dataset besar di berbagai bidang, termasuk kesehatan (Putri et al., 2023). Perkembangan dalam teknologi sistem informasi telah menyelesaikan banyak masalah di berbagai sektor, salah satunya dalam bidang kesehatan, melalui penggunaan data mining. Data mining merupakan metode ekstraksi pengetahuan atau pola penting dan relevan dari dataset yang luas dan rumit. Tujuannya adalah untuk mendeteksi hubungan, pola, atau tren yang tidak jelas pada data, yang dapat menyediakan pemahaman mendalam dan bermanfaat untuk keperluan pengambilan keputusan (Akbarollah et al., 2023). Namun, penerapan teknik ini pada data rekam medis gagal jantung masih terbatas, terutama di negara berkembang (Novaldy & Herliana, 2021).

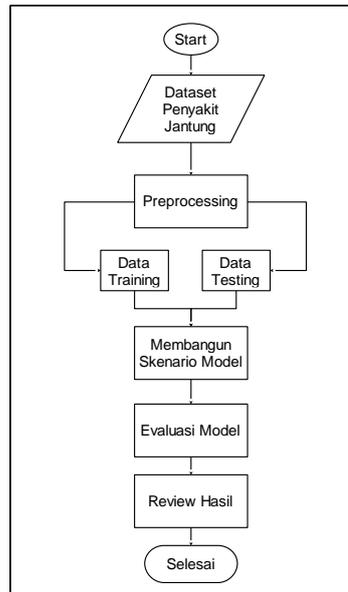
Penelitian ini bertujuan untuk menganalisis rekam medis pasien gagal jantung dengan menggunakan metode data mining, sehingga diharapkan dapat memberikan kontribusi signifikan dalam pengelolaan dan pencegahan gagal jantung. Melalui pendekatan ini, diharapkan dapat diidentifikasi pola-pola tersembunyi dalam data yang bisa membantu dalam pengambilan keputusan klinis dan peningkatan kualitas perawatan pasien. Selain itu, penelitian ini juga akan mengevaluasi efektivitas dan efisiensi berbagai algoritma data mining dalam konteks data rekam medis, sehingga dapat memberikan rekomendasi metodologi terbaik untuk praktik klinis masa depan.

Dengan meningkatnya digitalisasi data kesehatan dan kemajuan dalam teknologi analitik, penerapan data mining dalam rekam medis pasien gagal jantung menawarkan peluang yang belum dimanfaatkan untuk memajukan penelitian dan praktek medis dalam menghadapi tantangan kesehatan global ini.

Penelitian ini melakukan pengujian terhadap dataset rekam medis pasien gagal jantung dengan mengidentifikasi korelasi antar variabel dan pengujian dengan menggunakan model data mining, hasil pengujian dengan model data mining yang diusulkan seperti *Random forest*, *Extreme Gradient Boosting*, *Extra Tree* dan *Adaboost* didapatkan model paling bagus yaitu *Random forest* berdasarkan nilai akurasi paling tinggi diangka 0,84% dalam memprediksi keakuratan model pada dataset dengan 13 label dan 1 variabel target sebagai prediksi keberlangsungan hidup pasien gagal jantung. Selain mendapatkan akurasi yang tinggi, model *random forest* mendapatkan nilai AUC/ROC untuk kelas klasifikasi di angka 0,89 dimana ini dikategorikan sebagai klasifikasi sangat baik (Novaldy & Herliana, 2021).

## BAHAN DAN METODE

Dalam penelitian ini, metode yang digunakan adalah eksperimental dengan pendekatan kuantitatif. Penelitian eksperimental digunakan untuk mengidentifikasi efek dari sebuah tindakan yang secara khusus dilakukan oleh peneliti (Edric & Tamba, 2022) sedangkan pendekatan kuantitatif dipilih sebab data yang dikumpulkan berbentuk numerik (Novaldy & Herliana, 2021). Adapun tahapan pada penelitian ini dijelaskan pada gambar 1.



Gambar 1. Alur Penelitian

Pada gambar 1 dijelaskan proses penelitian dimulai dengan pengambilan dataset penyakit jantung yang kemudian melalui tahap *preprocessing*, di mana data dibersihkan dan disiapkan. Selanjutnya data tersebut dibagi menjadi dua bagian yaitu data training dan data testing, data training digunakan untuk membangun model, dan data testing dipakai untuk menguji model. Setelah model dibangun berdasarkan skenario yang ditentukan, dilakukan evaluasi terhadap kinerja model tersebut. Penelitian ini berakhir dengan tahap review hasil, di mana penemuan dan performa model dikaji untuk menarik kesimpulan penelitian.

## HASIL DAN PEMBAHASAN

Berdasarkan sumber dari dataset yang dijadikan object penelitian terhadap hasil rekam medis pasien gagal jantung yang besumber dari dataset publik kaggle.com beberapa informasi terkait fitur-fitur yang digunakan untuk mendeteksi penyakit jantung dimana terdapat 13 fitur yang terbagi menjadi 2 bagian yaitu 12 atribut label dan 1 atribut target. Pada penelitian ini jumlah sample yang digunakan sebanyak 299 data. Berikut adalah penjelasan dari masing-masing atribut

1. Age/Umur: Menunjukkan umur pasien.
2. Anemia: Mengindikasikan penurunan sel darah merah atau hemoglobin.

3. *Creatine Phosphokinase (CPK)*: Menunjukkan tingkat enzim CPK dalam darah, yang bisa meningkat dengan adanya kerusakan otot, termasuk otot jantung.
4. *Diabetes*: Menyatakan apakah pasien memiliki diabetes.
5. *Ejection Fraction*: Mengukur persentase darah yang dikeluarkan dari jantung pada setiap kontraksi, yang merupakan indikator fungsi jantung.
6. *High Blood Pressure*: Menunjukkan apakah pasien menderita hipertensi.
7. *Platelets*: Jumlah trombosit dalam darah, yang penting dalam proses pembekuan darah.
8. *Serum Creatine*: Mengukur tingkat kreatinin dalam darah, yang merupakan indikator fungsi ginjal.
9. *Serum Sodium*: Tingkat natrium dalam darah, yang berperan penting dalam keseimbangan air dan elektrolit.
10. *Sex*: Jenis kelamin pasien.
11. *Smoking*: Menandakan apakah pasien merokok.
12. *Time*: Durasi waktu *follow-up* pasien.
13. *Dead Event*, merupakan variabel target apakah pasien meninggal atau tidak selama masa *follow up*

Sample dari dataset penelitian di tampilkan pada table 1.

Tabel 1. Data Rekam Medis Pasien Penyakit Jantung

Age	Anaemia	Creatinine Phosphokinase	Diabetes	Ejection Fraction	High Blood Pressure	Platelets	Serum Creatinine	Serum Sodium	Sex	Smoking	Time	Death Event
75	0	582	0	20	1	265000	1.9	130	1	0	4	1
55	0	7861	0	38	0	263358.03	1.1	136	1	0	6	1
65	0	146	0	20	0	162000	1.3	129	1	1	7	1
50	1	111	0	20	0	210000	1.9	137	1	0	7	1
65	1	160	1	20	0	327000	2.7	116	0	0	8	1
...	...	...	...	...	...	...	...	...	..	...	...	...
...	...	...	...	...	...	...	...	...	..	...	...	...
90	1	47	0	40	1	204000	2.1	132	1	1	8	1
75	1	246	0	15	0	127000	1.2	137	1	0	10	1
60	0	2261	0	35	1	228000	0.9	136	1	0	115	0
62	0	30	1	60	1	244000	0.9	139	1	0	117	0
50	0	115	0	45	1	184000	0.9	134	1	1	118	0

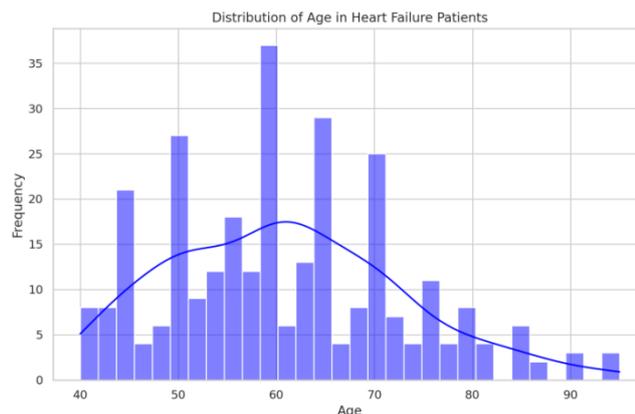
Pada tabel 1 diatas menampilkan data-data rekam medis dari pasien selama masa *follow up* analisis berdasarkan data hasil rekam medis dan analisis dengan data mining akan digunakan untuk mendapatkan informasi secara detail dari hasil rekam medis dari tabel 1 diatas.

## A. Analisis Data Rekam Medis

Dalam upaya menemukan korelasi antar variable label terhadap variable target maka diperlukan analisis untuk melihat korelasi dari masing-masing variable yang dianggap penting terhadap keberlangsungan hidup pasien gagal jantung selama masa *follow-up*

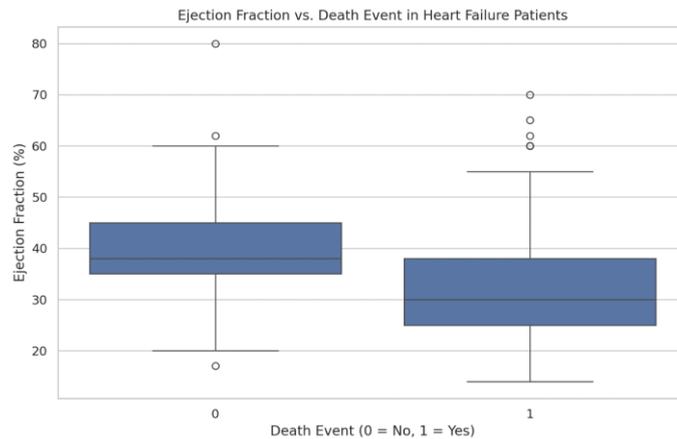
### 1. Distribusi Usia dan Hubungan Antara Fraksi Ejeksi pada Kejadian Kematian

Pada analisis ini, kita akan menyelidiki dua aspek kunci yang berpengaruh dalam pengelolaan pasien gagal jantung, distribusi usia pasien dan hubungan antara fraksi ejeksi dengan kejadian kematian. Pertama, kita akan melihat bagaimana usia pasien terdistribusi, mengingat usia adalah faktor risiko utama dalam gagal jantung (Moreno-Sanchez, 2020). Kedua, kita akan mengeksplorasi bagaimana fraksi ejeksi, yang mengukur efektivitas jantung dalam memompa darah, mempengaruhi kemungkinan kematian di antara pasien. Analisis ini bertujuan untuk memberikan wawasan yang lebih dalam tentang bagaimana faktor-faktor ini berinteraksi dan berkontribusi terhadap hasil klinis pasien dengan gagal jantung. Grafik dari distribusi usia dan fraksi ejeksi pada gambar 2 dan gambar 3.



Gambar 2. Distribusi Usia Pasien Gagal Jantung

Pada Gambar 2 grafik ini menggambarkan rentang usia pasien, dengan mayoritas berada di kelompok usia yang lebih tua.

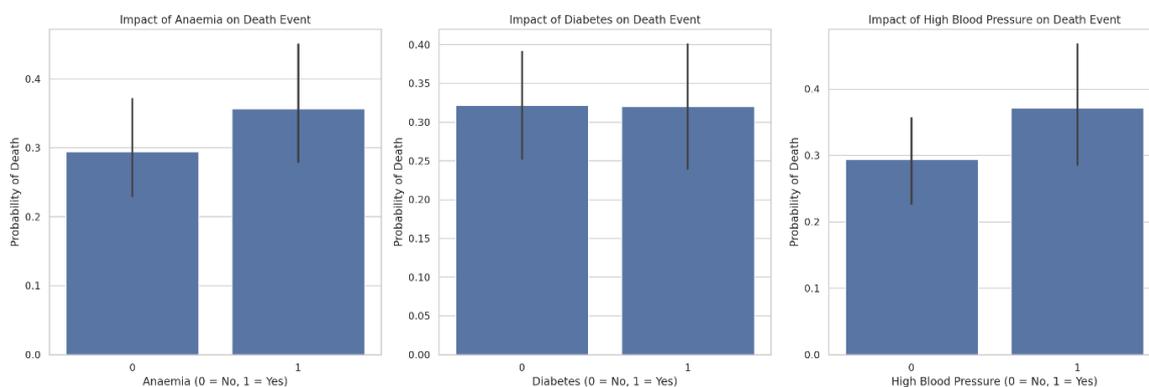


Gambar 3. Distribusi Fraksi Ejection Gagal Jantung

Pada Gambar 3 grafik ini mengeksplorasi hubungan antara fraksi ejeksi yaitu proporsi darah yang dikeluarkan dari ventrikel pada setiap denyutan jantung dan kejadian kematian. Terlihat bahwa fraksi ejeksi yang rendah cenderung berhubungan dengan risiko kematian yang lebih tinggi, sementara pasien yang masih hidup biasanya memiliki fraksi ejeksi yang lebih tinggi.

## 2. Analisis Dampak Anemia, Diabetes dan Tekanan Darah Tinggi Terhadap Kejadian Kematian

Dalam bagian kedua dari analisis ini, kita akan fokus pada pengaruh anemia, diabetes, dan tekanan darah tinggi terhadap kejadian kematian di kalangan pasien gagal jantung. Ketiga kondisi ini sering kali ditemukan bersamaan dengan gagal jantung dan dapat memperburuk prognosis pasien. Dengan memahami seberapa signifikan dampak dari masing-masing kondisi ini terhadap risiko kematian (Ishaq et al., 2021). Analisis ini bertujuan untuk mengidentifikasi hubungan yang mungkin antara kondisi-kondisi kesehatan ini dan outcome klinis, sehingga memberikan panduan yang lebih jelas bagi pengelolaan pasien yang kompleks ini.



Gambar 4. Analisis Dampak Anemia, Diabetes dan Tekanan Darah Tinggi

Hasil analisis dari Gambar 4 adalah analisis lebih lanjut tentang rekam medis pasien gagal jantung yang mengevaluasi dampak anemia, diabetes, dan tekanan darah tinggi terhadap kejadian kematian. Grafik terkait anemia menunjukkan bahwa pasien dengan anemia memiliki probabilitas kematian yang sedikit lebih tinggi dibandingkan dengan pasien tanpa anemia. Dari analisis terhadap pasien dengan diabetes, tidak terlihat perbedaan yang signifikan dalam probabilitas kematian antara pasien dengan dan tanpa diabetes. Sementara itu, pasien dengan tekanan darah tinggi cenderung memiliki probabilitas kematian yang lebih tinggi dibandingkan dengan mereka yang tidak memiliki tekanan darah tinggi. Analisis ini membantu dalam mengidentifikasi faktor risiko yang dapat meningkatkan probabilitas kematian di kalangan pasien gagal jantung.

### 3. Analisis Waktu Hingga Kematian

Analisis ini bertujuan untuk memahami hubungan antara durasi pemantauan (waktu) dan kematian. Kita akan mengevaluasi distribusi waktu follow-up antara pasien yang telah meninggal dibandingkan dengan mereka yang masih hidup. Hasil analisis pada Gambar 5.



Gambar 5. Analisis Waktu Hingga Kematian

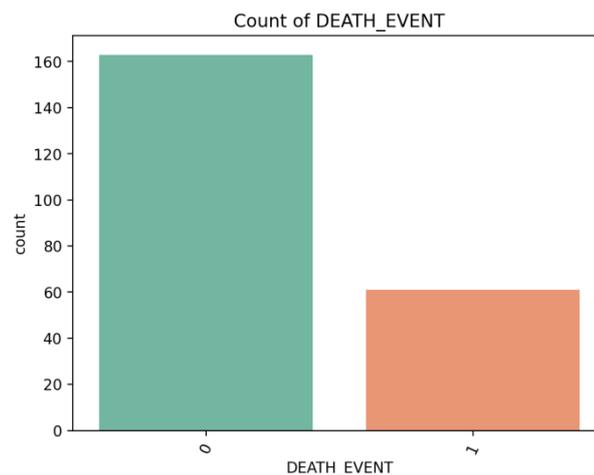
Pada Gambar 5 Grafik yang disajikan menggambarkan distribusi durasi *follow-up* (dalam hari) antara pasien yang meninggal dengan yang selamat, temuan ini mengidentifikasi beberapa temuan penting.: Pasien yang meninggal cenderung memiliki periode *follow-up* yang lebih singkat, menandakan bahwa kematian mereka terjadi lebih awal selama masa pengamatan. Sebaliknya, pasien yang selamat memiliki distribusi waktu *follow-up* yang lebih panjang dan lebih merata, yang menunjukkan bahwa mereka mampu bertahan hidup lebih lama pasca-diagnosis awal atau intervensi. Temuan ini menekankan pentingnya pertimbangan untuk intervensi lebih awal atau pemantauan yang lebih intensif bagi pasien dengan risiko tinggi kematian dini, sebagai upaya meningkatkan kesempatan bertahan hidup.

## B. Analisis Data Mining Terhadap Penyakit Gagal Jantung

Dalam upaya untuk meningkatkan pemahaman dan pengelolaan penyakit gagal jantung, membangun model prediktif menjadi langkah penting. Model ini dirancang untuk mengidentifikasi pola dan prediktor yang dapat mempengaruhi hasil pasien, termasuk kematian dan kelangsungan hidup. Dengan menerapkan teknik data mining dan pembelajaran mesin pada dataset yang telah dikumpulkan, kita dapat mengembangkan algoritma yang mampu memprediksi risiko kematian atau kejadian klinis lainnya berdasarkan berbagai variabel klinis dan demografis.

## 1. *Preprocessing*

Pada tahapan ini sebanyak 299 data dikumpulkan dan dianalisis untuk mendapatkan data-data yang bersih sehingga hasil yang diharapkan bisa lebih optimal ketika melakukan pemodelan. Pada tahapan ini data dibagi menjadi dua bagian, yakni sebagai data training dan data testing dengan perbandingan 80:20 (Nawawi et al., 2019) . Pada tahap ini melakukan analisis terhadap jumlah sebaran data yang diteliti apakah terdapat kesenjangan atau ketidakseimbangan yang signifikan, yang bisa mempengaruhi keakuratan dan efektivitas model yang akan dikembangkan. Gambaran dataset pada penelitian ini pada gambar 6.



Gambar 6. Kelas Dataset Pasien Gagal Jantung

Dari Gambar 6 perbandingan jumlah pasien gagal jantung berdasarkan status kematian mereka, dengan warna hijau mewakili kelas pasien yang meninggal dan warna merah mewakili kelas pasien yang selamat. Dari grafik, dapat dilihat bahwa jumlah pasien yang selamat (ditunjukkan dengan warna merah) lebih rendah dibandingkan dengan pasien yang meninggal (ditunjukkan dengan warna hijau). Ini menunjukkan bahwa dalam dataset ini, lebih banyak pasien gagal jantung yang mengalami kematian daripada yang berhasil selamat. Data ini dapat digunakan untuk menganalisis faktor-faktor yang mungkin mempengaruhi kelangsungan hidup pasien setelah didiagnosis dengan gagal jantung.

## 2. *Membangun Model*

Dalam tahap pembangunan model dari penelitian ini, berbagai model prediktif diuji pada dataset penelitian, termasuk model-model ensemble prediktif seperti *Random forest*, *Extreme Gradient Boosting (XGBoost)*, *Extra Trees*, dan *AdaBoost*. Model *Random forest* dipilih berdasarkan ketahanannya terhadap dataset dengan berbagai ukuran (Moreno-Sanchez, 2020), *XGBoost* dikenal sebagai implementasi yang efisien dan skalabel dari peningkatan gradien yang dirancang untuk kecepatan dan performa, *Extra Trees* adalah variasi dari pengklasifikasi bagging yang menggunakan teknik berbasis pohon untuk meningkatkan akurasi dan efisiensi komputasi (Ishaq et al., 2021) sementara *AdaBoost* mengombinasikan sejumlah pengklasifikasi yang lebih sederhana untuk membentuk pengklasifikasi yang lebih kuat.

### 3. Evaluasi Hasil

Dalam proses pengembangan model prediktif untuk penyakit gagal jantung, evaluasi model adalah langkah kritis yang menentukan seberapa efektif model tersebut dalam memprediksi kejadian yang diinginkan, seperti kematian. Untuk memperkirakan kinerja klasifikasi, beberapa metrik digunakan dalam penelitian ini yaitu akurasi, presisi, recall/sensitivitas, presisi, f1-score dan roc-auc (Moreno-Sanchez, 2020). Akurasi menggambarkan tingkat prediksi yang sebenarnya dan cocok untuk data yang seimbang antar kelas (Cahya & Buani, 2021). Namun, karena kumpulan data yang digunakan tidak seimbang dalam fitur targetnya, metrik lain akan digunakan untuk menilai pengklasifikasi model. Hasil dari analisis terhadap data dengan menggunakan model data mining yang diuji, pada table 2.

Tabel 2. Hasil Pemodelan

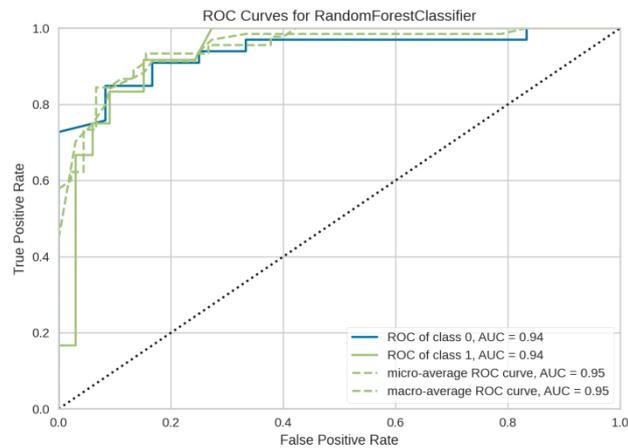
Model	Accuracy	AUC	Recall	Precision	F1	TT (Sec)
rf ( <i>Random forest Classifier</i> )	0.8493	0.8958	0.6500	0.7833	0.6925	9.230
xgboost ( <i>Extreme Gradient Boosting</i> )	0.8320	0.8773	0.6933	0.6403	0.5363	0.0730
et ( <i>Extra Trees Classifier</i> )	0.8042	0.8485	0.6917	0.5302	0.4219	0.1690
ada ( <i>Ada Boost Classifier</i> )	0.7931	0.8219	0.6648	0.5926	0.4575	0.1250

Sumber (Hasil Penelitian, 2024)

Berdasarkan hasil pemodelan dengan menggunakan data mining pada tabel 2 kesimpulan yang dapat diambil adalah model *random forest classifier* menunjukkan kinerja yang paling konsisten dan tinggi di antara model-model yang diuji, dan menjadikan model yang baik untuk aplikasi lebih lanjut dalam klasifikasi pasien gagal jantung berdasarkan data. Namun, *Gradient Boosting Classifier* memiliki keunggulan dalam mengidentifikasi kasus positif, yang bisa sangat penting tergantung pada konteks aplikasi klinis. Perlu pertimbangan lebih lanjut terkait trade-off antara precision dan recall, terutama dalam konteks medis di mana setiap aspek memiliki konsekuensi yang berbeda.

### 4. Review Hasil Pemodelan

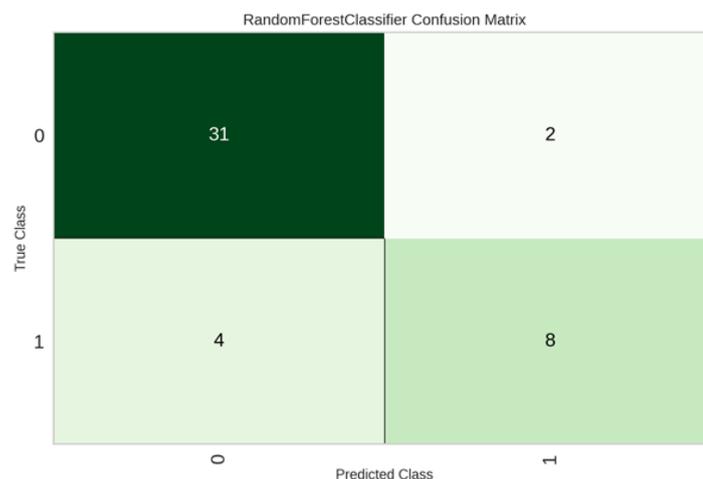
Berdasarkan hasil pemodelan dengan menggunakan beberapa model prediktif hasil yang diperoleh adalah model *Random forest Classifier* yang paling baik dibandingkan dengan model lainnya, tahap selanjutnya adalah melihat grafik AUC pada model *Random forest*, hasilnya pada Gambar 7.



Gambar 7. Curva AUC ROC Model *Random forest*

Pada Gambar 7 Grafik ROC untuk *Random forest Classifier* menunjukkan bahwa model ini memiliki kinerja yang sangat baik dalam mengklasifikasikan kedua kelas pasien gagal jantung, dengan nilai AUC sebesar 0.94 untuk kedua kelas. Hal ini menunjukkan model sangat efektif dalam membedakan pasien yang tidak mengalami event kematian (kelas 0) serta mengidentifikasi mereka yang mengalami event kematian (kelas 1), dengan tingkat prediksi yang tinggi dan kesalahan yang minimal.

Selanjutnya menampilkan confusion matriks untuk menguji ke efektifan model *Random forest* dalam membaca hasil prediksi, secara detail digambarkan pada gambar 8.



Gambar 8. Confussion Matriks

Dari gambar 8 dapat disimpulkan sebagai berikut:

1. Prediksi benar dan hasilnya benar, model berhasil memprediksi bahwa 31 pasien tidak akan mengalami event kematian, dan prediksi ini benar.
2. Prediksi benar hasilnya salah, Model salah memprediksi bahwa 2 pasien tidak akan mengalami event kematian, padahal mereka sebenarnya mengalami event kematian.
3. Prediksi salah hasilnya benar, Model salah memprediksi bahwa 4 pasien akan mengalami event kematian, padahal mereka sebenarnya tidak mengalami event kematian.
4. Prediksi salah dan hasilnya salah, Model berhasil memprediksi bahwa 8 pasien akan mengalami event kematian, dan prediksi ini benar.

### SIMPULAN

Penelitian tentang prediksi keberlangsungan hidup pasien gagal jantung dengan penerapan metode data mining menunjukkan efektivitas berbagai metode data mining, termasuk *Random forest Classifier*, *Gradient Boosting Classifier*, dan *Light Gradient Boosting*, dalam menganalisis dan memprediksi keberlangsungan hidup pasien gagal jantung. Hasilnya *Random forest Classifier*, menunjukkan kinerja yang paling baik dalam memprediksi prediksi keberlangsungan hidup pasien gagal jantung, terbukti melalui nilai AUC dan akurasi yang tinggi. Meski masih terdapat beberapa kesalahan prediksi tetapi dengan hasil *Random forest* mendapatkan nilai akurasi 0,84% menunjukkan keberhasilan pemodelan yang baik. Penerapan metode data mining dalam rekam medis pasien gagal jantung ini mendukung pengambilan keputusan medis yang berbasis bukti, meningkatkan efektivitas manajemen penyakit.

### DAFTAR PUSTAKA

- Akbarollah, M. F., Wiyanto, W., Ardiatma, D., & Zy, A. T. (2023). Penerapan Algoritma K-Nearest Neighbor Dalam Klasifikasi Penyakit Jantung. *Journal of Computer System and Informatics (JoSYC)*, 4(4), 850–860. <https://doi.org/10.47065/josyc.v4i4.4071>
- Cahya, D., & Buani, P. (2021). Penerapan Algoritma Naïve Bayes dengan Seleksi Fitur Algoritma Genetika Untuk Prediksi Gagal Jantung. *Evolusi: Jurnal Sains Dan Manajemen*, 9(2), 43–48.
- Edric, E., & Tamba, S. P. (2022). Prediksi Penyakit Gagal Jantung Dengan Menggunakan *Random forest*. *Jurnal Sistem Informasi Dan Ilmu Komputer Prima (JUSIKOM PRIMA)*, 5(2), 176–181. <https://doi.org/10.34012/jurnalsisteminformasidanilmukomputer.v5i2.2445>
- Ishaq, A., Sadiq, S., Umer, M., Ullah, S., Mirjalili, S., Rupapara, V., & Nappi, M. (2021). Improving the Prediction of Heart Failure Patients' Survival Using SMOTE and Effective Data Mining Techniques. *IEEE Access*, 9, 39707–39716. <https://doi.org/10.1109/ACCESS.2021.3064084>

- Kim, Y. T., Kim, D. K., Kim, H., & Kim, D. J. (2020). A Comparison of Oversampling Methods for Constructing a Prognostic Model in the Patient with Heart Failure. *International Conference on ICT Convergence, 2020-October*, 379–383. <https://doi.org/10.1109/ICTC49870.2020.9289522>
- Moreno-Sanchez, P. A. (2020). Development of an Explainable Prediction Model of Heart Failure Survival by Using Ensemble Trees. *Proceedings - 2020 IEEE International Conference on Big Data, Big Data 2020*, 4902–4910. <https://doi.org/10.1109/BigData50022.2020.9378460>
- Nawawi, H. M., Purnama, J. J., & Hikmah, A. B. (2019). Komparasi Algoritma Neural Network Dan Naïve Bayes Untuk Memprediksi Penyakit Jantung. *Jurnal Pilar Nusa Mandiri, 15*(2), 189–194. <https://doi.org/10.33480/pilar.v15i2.669>
- Novaldy, F., & Herliana, A. (2021). Penerapan Pso Pada Naïve Bayes Untuk Prediksi Harapan Hidup Pasien Gagal Jantung. *Jurnal Responsif: Riset Sains Dan Informatika, 3*(1), 37–43. <https://doi.org/10.51977/jti.v3i1.396>
- Pratama, Y., Prayitno, A., Azrian, D., Aini, N., Rizki, Y., & Rasywir, E. (2022). Klasifikasi Penyakit Gagal Jantung Menggunakan Algoritma K-Nearest Neighbor. *Bulletin of Computer Science Research, 3*(1), 52–56. <https://doi.org/10.47065/bulletincsr.v3i1.203>
- Putri, T. A. E., Widiari, T., & Santoso, R. (2023). Penerapan Tuning Hyperparameter Randomsearchcv Pada Adaptive Boosting Untuk Prediksi Kelangsungan Hidup Pasien Gagal Jantung. *Jurnal Gaussian, 11*(3), 397–406. <https://doi.org/10.14710/j.gauss.11.3.397-406>
- Salim, M. F., & Sugeng, S. (2018). Analisis Rekam Medis Pasien Diabetes Mellitus Melalui Implementasi Teknik Data Mining di RSUP Dr. Sardjito Yogyakarta. *Jurnal Kesehatan Vokasional, 2*(2), 167. <https://doi.org/10.22146/jkesvo.30331>