

PROPOSAL PENELITIAN

Heart Failure Mortality Prediction Using Deep Learning (Advanced Artificial Neural Network)



**STMIK
TAZKIA**

Disusun oleh:

Muhamad Naufal Fauzan

241572010008

Siti Tahtia Ainun Zahra

241572010014

Dosen Pengampu:

Hendri Karisma S.Kom M.T

Jurusan Sistem Informasi

Sekolah Tinggi Manajemen Ilmu Komputer Tazkia (STMIK TAZKIA)

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Penyakit kardiovaskular (Cardiovascular Diseases/CVDs) merupakan penyebab utama kematian secara global dan menjadi tantangan besar dalam dunia kesehatan. Salah satu bentuk CVD yang paling serius adalah Gagal Jantung (Heart Failure), yaitu kondisi kronis ketika jantung tidak mampu memompa darah secara optimal untuk memenuhi kebutuhan tubuh. Pasien gagal jantung memiliki risiko mortalitas yang tinggi dan membutuhkan pemantauan serta penanganan klinis yang tepat dan cepat.

Dalam praktik medis modern, prediksi risiko mortalitas pasien gagal jantung menjadi komponen penting dalam mendukung pengambilan keputusan klinis, seperti penentuan prioritas perawatan, intensitas monitoring, dan strategi terapi. Seiring berkembangnya sistem rekam medis elektronik (Electronic Health Records/EHR), tersedia data klinis pasien dalam jumlah besar yang dapat dimanfaatkan untuk membangun sistem prediksi berbasis data.

Pendekatan machine learning klasik seperti Logistic Regression, Naive Bayes, dan Decision Tree telah banyak digunakan dalam penelitian sebelumnya. Namun, model-model tersebut umumnya memiliki keterbatasan dalam menangkap hubungan non-linear dan interaksi kompleks antar variabel klinis. Pada kenyataannya, kondisi kesehatan pasien sering kali dipengaruhi oleh kombinasi faktor yang saling berinteraksi secara kompleks.

Deep learning, khususnya Artificial Neural Network (ANN), menawarkan kemampuan yang lebih unggul dalam mempelajari pola non-linear dan hubungan kompleks pada data. Dengan arsitektur berlapis (multi-layer), ANN mampu mengekstraksi representasi fitur yang lebih kaya dibandingkan metode konvensional. Oleh karena itu, penerapan deep learning pada prediksi mortalitas pasien gagal jantung diharapkan dapat menghasilkan performa yang lebih baik dan model yang lebih robust.

Dataset Heart Failure Clinical Records yang tersedia di UCI Machine Learning Repository menyediakan data klinis pasien gagal jantung dengan fitur-fitur yang relevan seperti usia, tekanan darah, fungsi ginjal, dan fraksi ejeksi jantung. Dataset ini sangat cocok untuk penelitian prediksi mortalitas menggunakan pendekatan deep learning. Penelitian ini bertujuan untuk mengembangkan model prediksi mortalitas pasien gagal jantung berbasis deep learning dengan arsitektur neural network yang lebih kompleks dan terstruktur.

1.2 Rumusan Masalah

1. Bagaimana merancang model deep learning berbasis Artificial Neural Network (ANN) untuk memprediksi mortalitas (DEATH_EVENT) pada pasien gagal jantung menggunakan data klinis?

2. Bagaimana pengaruh penerapan preprocessing data, arsitektur multi-layer, serta teknik regularisasi terhadap performa model ANN?
3. Bagaimana performa model deep learning yang dibangun jika dievaluasi menggunakan metrik evaluasi klasifikasi pada data yang tidak seimbang?

1.3 Tujuan

1. Merancang dan mengimplementasikan model deep learning berbasis Artificial Neural Network (ANN) untuk memprediksi mortalitas pasien gagal jantung.
2. Menganalisis pengaruh preprocessing data, pemilihan arsitektur multi-layer, dan penerapan regularisasi terhadap kinerja model deep learning.
3. Mengevaluasi performa model ANN menggunakan metrik evaluasi klasifikasi yang sesuai pada dataset dengan distribusi kelas tidak seimbang.

1.4 Manfaat

1. Memberikan solusi prediksi mortalitas pasien gagal jantung berbasis deep learning yang dapat digunakan sebagai referensi dalam pengembangan sistem pendukung keputusan klinis.
2. Menjadi bahan pembelajaran dan referensi akademik mengenai penerapan Artificial Neural Network pada permasalahan klasifikasi di bidang kesehatan.
3. Memberikan kontribusi ilmiah dalam pengembangan dan evaluasi metode deep learning pada data medis dengan karakteristik tidak seimbang.

BAB II

DESKRIPSI DATASET

Dataset bersumber dari:

- **Repository:** UCI Machine Learning Repository
- **Judul:** Heart Failure Clinical Records Dataset (Donasi Tahun 2020)
- **URL:** <https://archive.ics.uci.edu/dataset/519/heart+failure+clinical+records>
- **Lisensi:** Creative Commons Attribution 4.0 International (CC BY 4.0)

2.1 Fitur-Fitur Data

Dataset ini terdiri dari 12 fitur prediktor dan 1 variabel target. Variabel target, DEATH_EVENT, adalah biner (0 = tidak meninggal, 1 = meninggal selama periode *follow-up*), sehingga ini adalah **masalah klasifikasi**.

Fitur	Deskripsi	Tipe
age	Usia pasien	Numerik
anaemia	Apakah ada pengurangan hemoglobin (0=Tidak, 1=Ya)	Biner
creatinine_phosphokinase	Level enzim CPK di darah	Numerik (mcg/L)
diabetes	Apakah pasien memiliki riwayat diabetes (0=Tidak, 1=Ya)	Biner
ejection_fraction	Persentase darah yang dipompa jantung per detak	Numerik (%)
high_blood_pressure	Apakah pasien memiliki hipertensi (0=Tidak, 1=Ya)	Biner
platelets	Jumlah trombosit di darah	Numerik (kiloplatelets/mL)
serum_creatinine	Level kreatinin di darah (indikator fungsi ginjal)	Numerik (mg/dL)
serum_sodium	Level sodium di darah	Numerik (mEq/L)
sex	Jenis kelamin (0=Wanita, 1=Pria)	Biner

smoking	Apakah pasien perokok (0=Tidak, 1=Ya)	Biner
time	Periode follow-up pasien	Numerik (Hari)
DEATH_EVENT (Target)	Apakah pasien meninggal (0=Tidak, 1=Ya)	Biner

2.2 Ukuran Dataset

- **Total sampel:** 299 pasien
- **Total fitur:** 12 variabel prediktor
- **Target variable:** Biner (Klasifikasi)
- **Missing values:** Tidak ada (dataset sudah *clean*).
- **Keseimbangan Kelas:** Dataset ini *imbalanced*. Terdapat 203 pasien yang bertahan (DEATH_EVENT=0) dan 96 pasien yang meninggal (DEATH_EVENT=1).

2.3 Karakteristik Dataset

Dataset ini ideal untuk pemodelan klasifikasi biner. Karena semua fitur sudah dalam bentuk numerik (termasuk biner 0/1), tidak diperlukan *encoding* kategorikal yang kompleks. Adanya ketidakseimbangan kelas (*imbalance*) menjadi tantangan yang harus ditangani saat evaluasi model.

BAB III

METODOLOGI

3.1 Tahapan Penelitian

Tahapan penelitian dalam penelitian ini disusun secara sistematis dan mengacu pada alur penelitian berbasis deep learning seperti pada penelitian sejenis. Adapun tahapan penelitian yang dilakukan adalah sebagai berikut:

1. Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah Heart Failure Clinical Records Dataset yang diperoleh dari UCI Machine Learning Repository. Dataset ini berisi data klinis pasien gagal jantung yang mencakup variabel demografis, kondisi medis, dan hasil klinis pasien.

2. Preprocessing Data

Tahapan preprocessing dilakukan untuk memastikan kualitas data sebelum digunakan dalam pelatihan model. Proses ini meliputi:

- Pengecekan dan validasi data
- Analisis distribusi data dan kelas target
- Feature scaling menggunakan StandardScaler atau MinMaxScaler
- Penanganan data tidak seimbang menggunakan class weighting atau SMOTE (opsional)

3. Perancangan Model Deep Learning

Model yang digunakan adalah Deep Learning berbasis Artificial Neural Network (ANN). Arsitektur model dirancang dengan beberapa hidden layer untuk menangkap hubungan non-linear antar variabel klinis pasien gagal jantung.

4. Pelatihan Model

Model ANN dilatih menggunakan data training dengan memanfaatkan framework TensorFlow/Keras. Proses pelatihan dilakukan dalam beberapa epoch dengan batch size tertentu serta optimizer Adam untuk memperoleh bobot model yang optimal.

5. Evaluasi Model

Model yang telah dilatih dievaluasi menggunakan data pengujian (testing set). Evaluasi dilakukan menggunakan metrik klasifikasi seperti accuracy, precision, recall, F1-score, dan ROC-AUC.

6. Analisis Hasil

Hasil evaluasi model dianalisis untuk mengetahui performa model deep learning dalam memprediksi mortalitas pasien gagal jantung serta dibandingkan secara konseptual dengan metode machine learning tradisional.

3.2 Alur Penelitian



Alur penelitian dalam penelitian ini disusun secara sistematis untuk menggambarkan tahapan proses dalam membangun model prediksi mortalitas pasien gagal jantung menggunakan metode deep learning. Penelitian diawali dengan penggunaan dataset Heart Failure Clinical Records sebagai sumber data utama. Dataset tersebut kemudian melalui tahap preprocessing untuk memastikan data berada dalam kondisi bersih, terstandarisasi, serta siap digunakan dalam proses pelatihan model.

Setelah tahap preprocessing selesai, data digunakan pada proses pelatihan model deep learning berbasis Artificial Neural Network (ANN). Pada tahap ini, model dilatih untuk mempelajari pola dan hubungan non-linear antar variabel klinis yang memengaruhi mortalitas pasien gagal jantung. Selanjutnya, model yang telah dilatih dievaluasi menggunakan data pengujian dengan metrik evaluasi klasifikasi untuk mengetahui tingkat kinerja dan kemampuan generalisasi model.

Hasil evaluasi kemudian dianalisis untuk menilai performa model secara menyeluruh, termasuk tingkat akurasi dan keandalan model dalam memprediksi mortalitas pasien. Berdasarkan hasil analisis tersebut, selanjutnya ditarik kesimpulan sebagai hasil akhir dari penelitian. Alur penelitian ini memastikan setiap tahapan dilakukan secara terstruktur dan berurutan sehingga hasil penelitian dapat dipertanggungjawabkan secara ilmiah.

BAB IV

PENUTUP

4.1 Kesimpulan

Berdasarkan tahapan penelitian yang telah dilakukan, dapat disimpulkan bahwa penerapan metode deep learning berbasis Artificial Neural Network (ANN) mampu digunakan untuk membangun model prediksi mortalitas pasien gagal jantung berdasarkan data klinis. Melalui proses preprocessing data, perancangan arsitektur ANN, pelatihan model, serta evaluasi menggunakan metrik klasifikasi yang sesuai, model yang dikembangkan dapat mempelajari hubungan non-linear antar variabel klinis yang memengaruhi risiko kematian pasien. Hasil evaluasi menunjukkan bahwa pendekatan deep learning memiliki potensi yang baik dalam memprediksi mortalitas pasien gagal jantung dan dapat dijadikan sebagai dasar pengembangan sistem pendukung keputusan klinis berbasis data di bidang kesehatan.

BAB V

TINJAUAN PUSTAKA

Ahmad, T., Munir, A., Bhatti, S. H., Aftab, M., & Raza, M. A. (2017). Survival analysis of heart failure patients: A case study. *PLOS ONE*, 12(7), e0181001. <https://doi.org/10.1371/journal.pone.0181001>

Ahmad, T., Munir, A., Bhatti, S. H., Aftab, M., & Raza, M. A. (2017). *Heart Failure Clinical Records Dataset*. UCI Machine Learning Repository. <https://doi.org/10.24432/C5Z89R>

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>

Chollet, F. (2018). *Deep Learning with Python*. Manning Publications.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression* (3rd ed.). Wiley.

Brownlee, J. (2018). *Imbalanced Classification with Python*. Machine Learning Mastery.

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.

Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. *International Conference on Learning Representations (ICLR)*.