

# **IMPLEMENTASI MODEL ANFIS DENGAN SELEKSI FITUR PCA UNTUK KLASIFIKASI BINER KEBERADAAN PENYAKIT JANTUNG**

**Rizky Akhmad Fahreza<sup>1</sup>**

<sup>1</sup>Program Studi Informatika, Universitas Diponegoro;  
[rafahreza@students.undip.ac.id](mailto:rafahreza@students.undip.ac.id)



## **Abstrak**

Penelitian ini mengimplementasikan model Adaptive Neuro-Fuzzy Inference System (ANFIS) dengan seleksi fitur Principal Component Analysis (PCA) untuk klasifikasi biner keberadaan penyakit jantung. Tujuan utama dari penelitian ini adalah untuk membedakan pasien yang memiliki dan tidak memiliki penyakit jantung. PCA diterapkan untuk mereduksi dimensi dan memilih fitur yang paling relevan, sehingga meningkatkan efisiensi model. Model ANFIS dipilih karena kemampuannya untuk menggabungkan keunggulan jaringan saraf tiruan dan logika fuzzy, yang memungkinkan model ini menangani ketidakpastian dan ketidaktepatan dalam data medis. Hasil penelitian menunjukkan bahwa model ANFIS dengan seleksi fitur mencapai akurasi 83% dalam klasifikasi keberadaan penyakit jantung. Keterbatasan jumlah data menjadi salah satu faktor yang memengaruhi hasil ini. Penelitian selanjutnya diharapkan menggunakan data yang lebih besar untuk meningkatkan akurasi dan keandalan deteksi dini.

Kata Kunci: ANFIS, PCA, klasifikasi biner, penyakit jantung, seleksi fitur, jaringan saraf tiruan, logika fuzzy, deteksi dini.

# **1. Pendahuluan**

## **1.1. Latar belakang**

Teknologi Neuro-Fuzzy merupakan kombinasi dari jaringan saraf tiruan dan logika fuzzy, yang dirancang untuk menangani ketidakpastian dan kompleksitas dalam data. Pendekatan ini sangat berguna dalam bidang medis, terutama dalam diagnosis penyakit, di mana data sering kali tidak lengkap atau tidak pasti. Dengan memanfaatkan kemampuan adaptif dari jaringan saraf dan fleksibilitas dari logika fuzzy, model ini dapat memberikan hasil yang lebih akurat dalam klasifikasi dan prediksi.

## **1.2. Permasalahan**

Meskipun terdapat berbagai metode untuk mendiagnosis penyakit jantung, banyak dari metode tersebut masih menghadapi tantangan dalam hal akurasi dan efisiensi. Data medis sering kali memiliki dimensi tinggi dan fitur yang saling berkorelasi, sehingga menyulitkan model untuk belajar dengan baik. Oleh karena itu, penting untuk mengembangkan pendekatan yang dapat mereduksi dimensi data sekaligus meningkatkan akurasi klasifikasi.

## **1.3. Tujuan Penelitian**

Penelitian ini bertujuan untuk mengimplementasikan model ANFIS dengan seleksi fitur PCA untuk klasifikasi biner keberadaan penyakit jantung. Dengan cara ini, diharapkan dapat meningkatkan akurasi deteksi dini penyakit jantung serta memberikan wawasan tentang fitur-fitur penting yang berkontribusi pada diagnosis.

## **1.4. Signifikansi Penelitian**

Penelitian ini penting karena dapat memberikan kontribusi signifikan bagi pengembangan sistem diagnosis penyakit jantung yang lebih efisien. Dengan menggunakan teknologi Neuro-fuzzy yang adaptif dan metode seleksi fitur, hasil penelitian ini diharapkan dapat memberi opsi alternatif yang dapat membantu dokter atau praktisi bidang Kesehatan dalam pengambilan keputusan klinis dan meningkatkan kualitas perawatan pasien.

# **2. Tinjauan Pustaka**

## **2.1. Konsep Neuro-Fuzzy**

Neuro-fuzzy adalah sebuah sistem yang menggabungkan dua sistem logika fuzzy dengan jaringan saraf tiruan. Ouifak & Idri (2023) menjelaskan bahwa Neuro-Fuzzy adalah hasil dari penggabungan antara jaringan saraf tiruan dengan sistem inferensi fuzzy (FIS) yang berdasar pada logika fuzzy dan teori himpunan fuzzy. Ouifak & Idri (2023) juga menjelaskan bahwa sistem inferensi fuzzy dapat dengan mudah diinterpretasikan karena dapat merepresentasikan pengetahuan yang diperoleh

dalam bentuk aturan linguistik. Namun, sistem ini sangat bergantung pada ahli dan tidak memiliki kemampuan untuk belajar dari data historis seperti yang dilakukan oleh jaringan saraf tiruan. Oleh karena itu, penggabungan kedua sistem ini diharapkan bisa menjadi sebuah kesatuan sistem yang mampu menutupi kekurangan yang dimiliki oleh sistem inferensi fuzzy.

## **2.2. Literatur Terkait**

Terdapat beberapa literatur yang sudah melakukan studi terkait topik yang akan dibahas. Ufumaka (2020) melakukan analisis komparatif terhadap beberapa algoritma *machine learning* seperti Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Logistic Regression (LR), Naïve Bayes (NB), Random Forest (RF), dan Gradient Boosting (GD) untuk melakukan prediksi adanya penyakit jantung menggunakan dataset serupa yang digunakan pada penelitian ini. Ufumaka (2020) melampirkan bahwa algoritma SVM memberikan hasil akurasi paling tinggi (83,8%) dibandingkan algoritma *machine learning* lainnya yang memberikan akurasi  $\pm 82\%$ .

Ziasabounchi & Askerzade (2014) melakukan studi untuk memprediksi penyakit jantung menggunakan model klasifikasi berbasis ANFIS yang diimplementasikan di MATLAB. (Opeyemi & Justice, 2012) memberikan kesimpulan bahwa ANFIS telah digunakan untuk klasifikasi yang menggabungkan keunggulan jaringan saraf tiruan dan logika fuzzy. Pada model yang diusulkan, nilai error pelatihan dan error rata-rata pengujian masing-masing adalah 0,01 dan 0,15. Hasil eksperimen yang dilakukan menunjukkan bahwa teknik yang diusulkan memiliki akurasi yang tinggi, terutama jika dibandingkan dengan studi lain yang menggunakan basis data yang sama.

Berdasarkan studi-studi tersebut, penulis tertarik melakukan studi dengan menggunakan data serupa untuk diimplementasikan kepada model ANFIS yang dikombinasikan dengan Principal Component Analysis untuk menseleksi fitur-fitur yang mempengaruhi performa model. Berbeda dari studi sebelumnya yang mengimplementasikan ANFIS di MATLAB, penulis akan menggunakan TensorFlow sebagai *framework* dan Google Colab sebagai lingkungan pengembangan.

## **2.3. Kerangka Teori**

### **2.3.1. Adaptive Neuro-Fuzzy Inference System**

Adaptive Neuro-Fuzzy Inference System (ANFIS) adalah salah satu sistem pakar inferensi neuro-fuzzy hibrid yang bekerja dalam sistem inferensi fuzzy tipe Takagi-Sugeno, yang dikembangkan oleh Jyh-Shing dan Roger Jang pada tahun 1993 (Opeyemi & Justice, 2012). ANFIS sangat mudah diimplementasikan untuk penugasan yang *input/output*-nya sudah diberikan (*Supervised Learning*), sehingga sistem ini sangat sukses diaplikasikan di berbagai macam area (Loganathan & Girija, 2013). Sistem ini menawarkan pendekatan untuk pemodelan fuzzy yang bertujuan

mempelajari informasi dari suatu kumpulan data. Proses ini dirancang untuk menghitung parameter fungsi keanggotaan, sehingga sistem inferensi fuzzy dapat mengikuti data *input* dan *output* yang diberikan. Metode pembelajaran yang dilakukan memiliki cara kerja serupa dengan jaringan saraf tiruan (Opeyemi & Justice, 2012).

### **2.3.2. Principal Component Analysis**

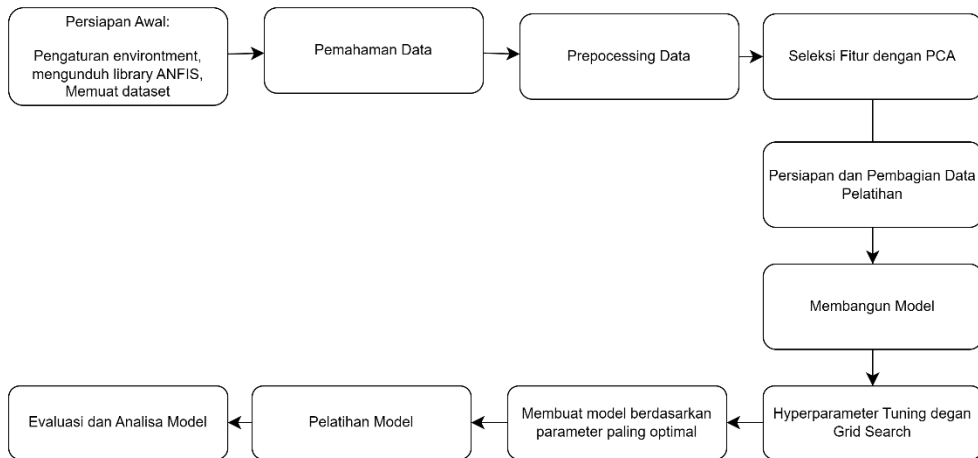
Principal Component Analysis (PCA), diperkenalkan oleh Karl Pearson pada tahun 1901, adalah metode statistik yang dikenal sebagai transformasi linier ortogonal. Metode ini berfokus pada variasi data dan mengidentifikasi pola-pola dominan dalam sebuah dataset. PCA berfungsi untuk mereduksi dimensi *dataset* yang sebelumnya besar menjadi lebih kecil, sambil tetap mempertahankan sebagian besar informasi penting dari dataset awal dengan melibatkan perhitungan rata-rata dan identifikasi komponen utama. Sebagai metode statistik yang efisien, PCA sering digunakan untuk memaksimalkan variasi data dan menyoroti pola-pola kuat dalam fitur dataset (Ejaz et al., 2019).

### **2.3.3. Grid Search**

Grid search adalah metode pencarian secara menyeluruh pada subset hiperparameter yang ditentukan secara manual untuk algoritma target. Pendekatan ini melibatkan menjalankan eksperimen pada sejumlah kondisi yang telah ditentukan sebelumnya. Meskipun praktis untuk masalah dengan jumlah parameter yang sedikit dan solusi yang diketahui berada dalam kisaran tertentu, grid search memiliki keterbatasan, seperti jumlah eksperimen yang bisa sangat banyak, kesalahan eksperimen yang signifikan, dan kemungkinan menemukan titik optimum palsu atau lokal. Selain itu, jika grid yang digunakan terlalu kecil, metode ini bisa kehilangan karakteristik penting atau tidak menemukan solusi optimal (Hossain & Timmer, 2021).

### 3. Metodologi Penelitian

#### 3.1. Desain Penelitian



Gambar 1. Desain penelitian

#### 3.2. Data

Data yang digunakan pada penelitian ini berasal dari koleksi basis data University of California. Data ini bertujuan untuk mendiagnosis penyakit jantung berdasarkan hasil tes medis. Dari 76 atribut yang tersedia, sebagian besar eksperimen menggunakan data yang sudah diproses dimana hanya 14 atribut saja yang digunakan dengan banyak baris data berjumlah 303 baris. Berikut deskripsi data yang digunakan:

Tabel 1. Tabel atribut *dataset Heart Disease Presence*

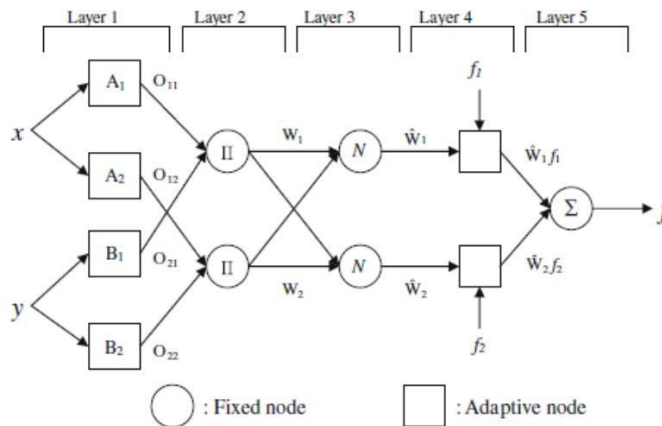
Heart Disease Presence Attributes				
No	Nama		Label	Deskripsi
1	Age		age	Diukur dengan satuan tahun
2	Sex		sex	1 = pria 0 = female
3	Chest Pain		cp	1: typical angina 2: atypical angina 3: non-anginal pain 4: asymptomatic
4	Resting Pressure	Blood	trestbps	Diukur dengan satuan mm Hg
5	Cholesterol		chol	Diukur dengan satuan mg/dl
6	Fasting Sugar	Blood	fbs	Jika lebih dari 120 mg/dl (normal): 1 = true

				0 = false
7	Resting Electrocardiogram	restecg		0: normal 1: abnormal 2: hipertrofi ventrikel kiri yang mungkin atau pasti berdasarkan kriteria Estes
8	Maximum Heart	thalach		Denyut jantung dicapai selama Stress Testing pasien
9	Exercise Induced Angina	exang		1 = yes 0 = no
10	Stress Test Depression	oldpeak		Nilai numerik yang menandakan penanda kejadian buruk pada jantung
11	Slope for Peak Exercise	slope		1: up sloping 2: flat 3: down sloping
12	Number of major vessels	ca		Jumlah pembuluh darah besar (0-3) diwarnai dengan fluoroskopi
13	Thallium Heart Rate	thal		3: normal 6: fixed defect 7: reverseable defect
14	Heart disease predicition	num		0: absence 1, 2, 3, 4: presence

### 3.3. Implementasi Model Neuro Fuzzy

#### 3.3.1 Arsitektur ANFIS

Adaptive Neuro-Fuzzy Inference System memiliki lima layer, yaitu *layer* fuzzifikasi, *layer rule*, *layer* normalisasi, *layer* defuzzifikasi, dan hasil *neuro* tunggal. Untuk arsitektur jelas dari ANFIS dapat dilihat pada gambar.



Gambar 2. Struktur ANFIS (Ciptaningrum et al., 2022)

### 3.3.2. Algoritma Pelatihan dan Parameter

Pada studi kasus penelitian ini, algoritma pelatihan yang digunakan adalah sebagai berikut:

#### 3.3.2.1. Backpropagation

*Backpropagation* digunakan untuk memperbarui parameter *neural network*, seperti bobot dan bias, berdasarkan perhitungan gradien dari fungsi loss terhadap parameter-parameter tersebut. Ini adalah bagian dari proses pembelajaran berbasis gradien.

#### 3.3.2.1. Optimizer

Adam digunakan sebagai optimizer untuk mempercepat dan menstabilkan proses pelatihan. Adam merupakan algoritma optimasi berbasis gradien dengan kemampuan adaptif yang menggabungkan keunggulan Momentum dan RMSProp.

#### 3.3.2.1. Loss Function

Digunakan untuk menghitung error atau selisih antara prediksi model dan label sebenarnya. Dalam kasus ini, Binary Focal Crossentropy dipilih karena kemampuannya menangani ketidakseimbangan kelas.

#### 3.3.2.1. Epoch dan Batch

Pelatihan dilakukan dalam sejumlah *epoch* (iterasi penuh melalui dataset), dan data dipecah menjadi *batch* (subset kecil) untuk memperbarui parameter model secara bertahap.

### 3.4. Pengujian dan Validasi

Pengujian dan validasi yang dilakukan pada penelitian studi kasus ini menggunakan *confusion matrix*. Hal ini dilakukan untuk mendapatkan nilai akurasi dan F1-Score yang didapat dari model yang dibuat.

## 4. Hasil dan Pembahasan

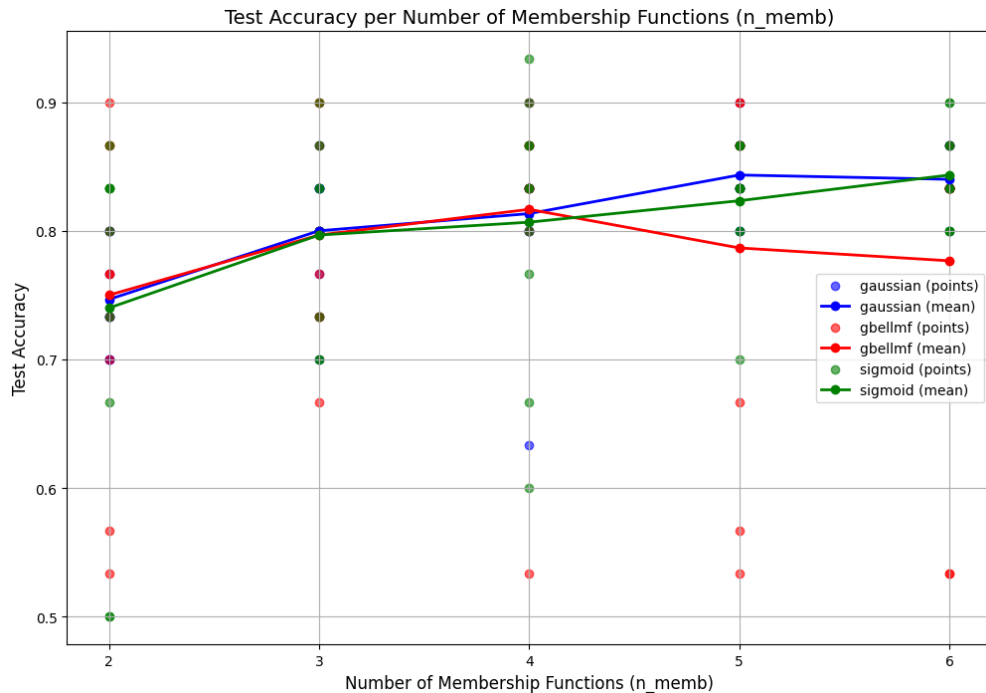
### 4.1. Hasil Eksperimen

Pada tahap awal, dilakukan pemrosesan terhadap *dataset* yang digunakan. Proses ini meliputi perubahan nama kolom “num” menjadi “target” dan menjadikan nilai 1,2,3,4 hanya menjadi satu nilai yaitu 1. Setelah itu, dilakukan pengisian nilai yang hilang dengan imputasi dan dilakukan standarisasi fitur untuk normalisasi nilai pada data.

Data yang sudah melalui tahap pemrosesan kemudian direduksi dimensinya menggunakan metode PCA. Pada eksperimen ini, jumlah komponen yang dipilih adalah 6. Hal ini dikarenakan model ANFIS pada *library* yang digunakan saat ini masih terbatas pada 2 – 6 *input* saja. Demi memaksimalkan proporsi varians pada fitur data, maka jumlah komponen disamakan dengan jumlah input maksimal yang bisa diproses oleh model.

Fitur yang sudah terseleksi kemudian dijadikan sebagai data yang diuji coba dalam *hyperparameter tuning* dengan menerapkan metode Grid Search untuk menemukan parameter utama yang paling mempengaruhi performa model, yaitu parameter *n\_memb* (banyaknya derajat fungsi keanggotaan) dan parameter *memb\_func* (jenis fungsi keanggotaan yang digunakan). Kondisi yang diberikan untuk metode Grid Search pada penelitian ini adalah berupa opsi untuk parameter *n\_member* berupa 1,2,3,4,5, dan 6. Sedangkan untuk parameter *memb\_func* berupa Gaussian, Generalized Bell, dan Sigmoid. Hasil dari Grid Search yang dilakukan sebanyak 10 iterasi adalah sebagai berikut:





Gambar 3. Grafik akurasi pengujian Grid Search sebanyak 10 iterasi

Berdasarkan grafik akurasi pengujian Grid Search di atas, maka nilai optimal untuk parameter `n_memb` adalah 5 dan parameter `memb_func` adalah Gaussian. Sehingga model yang dibangun adalah seperti berikut:

Model: "heart\_disease\_presence"

Layer (type)	Output Shape	Param #	Connected to
inputLayer (InputLayer)	[(10, 6)]	0	[]
fuzzyLayer (FuzzyLayer)	(10, 5, 6)	60	['inputLayer[0][0]']
ruleLayer (RuleLayer)	(10, 15625)	0	['fuzzyLayer[0][0]']
normLayer (NormLayer)	(10, 15625)	0	['ruleLayer[0][0]']
defuzzLayer (DefuzzLayer)	(10, 15625)	109375	['normLayer[0][0]', 'inputLayer[0][0]']
sumLayer (SummationLayer)	(10, 1)	0	['defuzzLayer[0][0]']

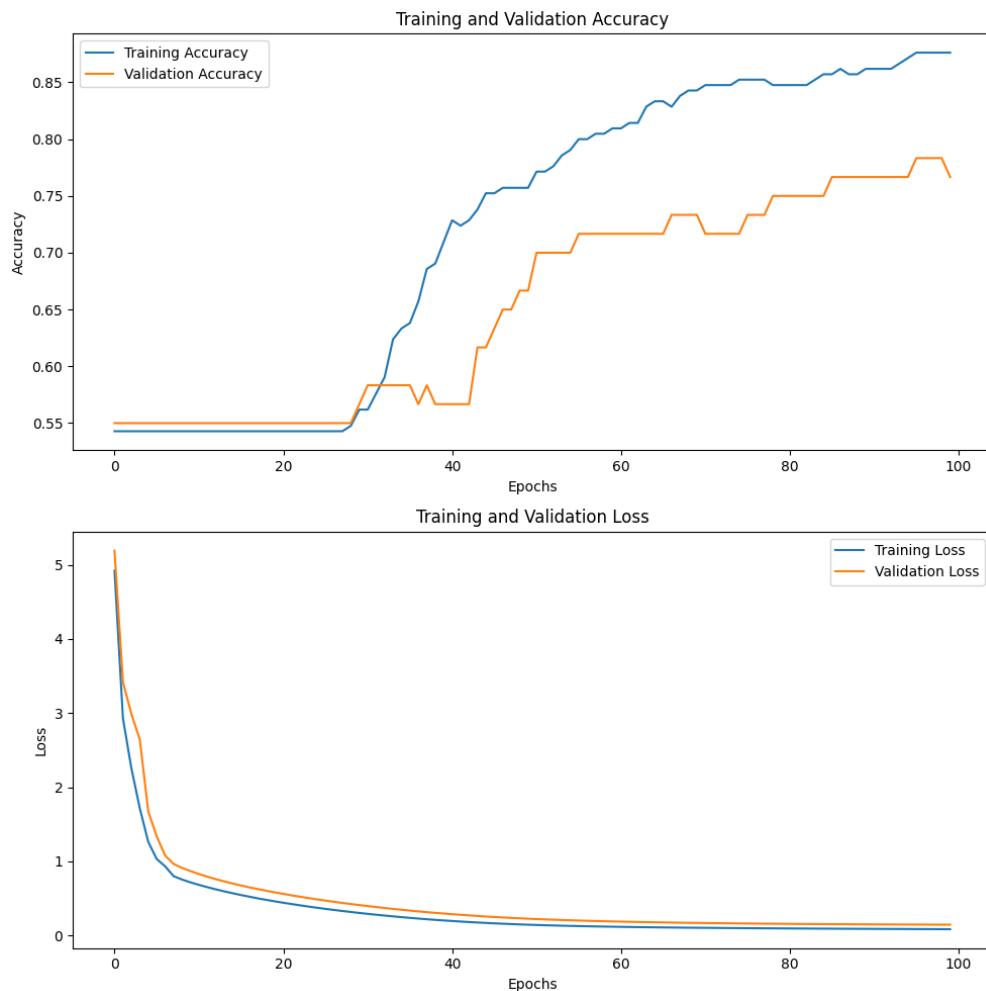
=====

Total params: 109,435  
Trainable params: 109,435  
Non-trainable params: 0

Gambar 4. Summary model ANFIS yang dibangun

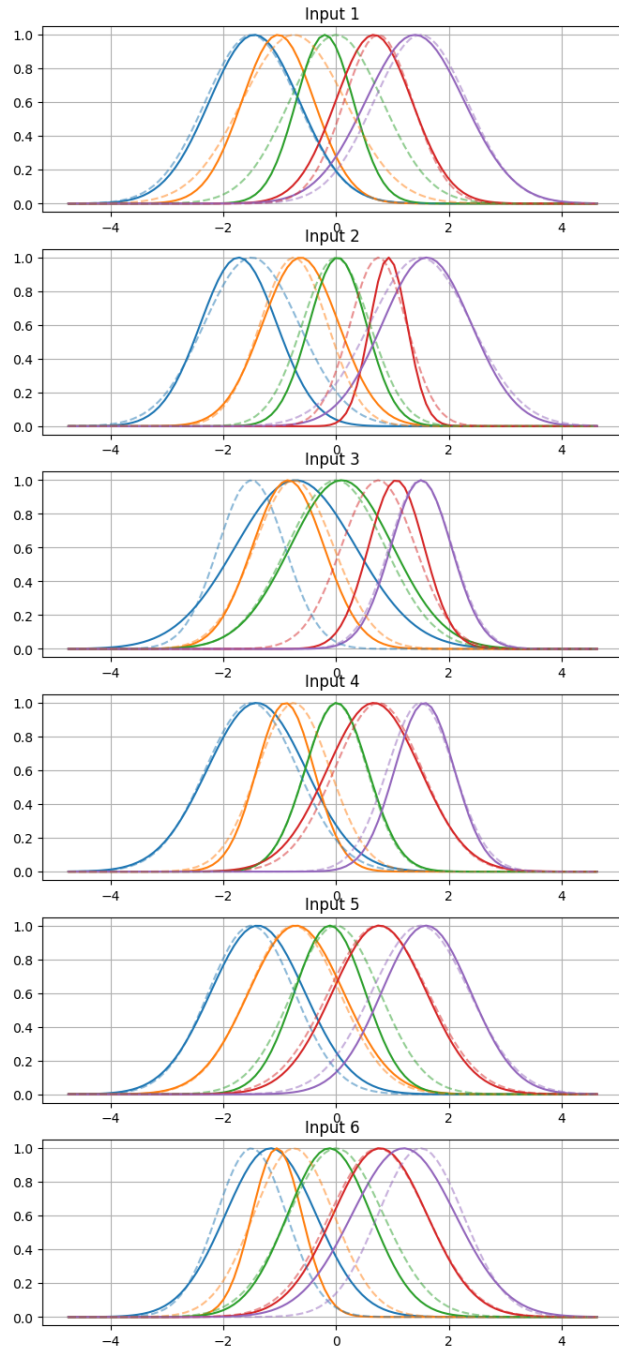
## 4.2. Analisis Hasil

Dilakukan pelatihan terhadap model yang dibangun dengan memanfaatkan Adam dengan *learning rate* sebesar 0.001 sebagai algoritma optimasi dan Binary Focal Crossentropy sebagai *loss function*. Pelatihan dilakukan sebanyak 100 epoch dengan ukuran batch 10. Berikut ditampilkan beberapa metrik hasil pelatihan dan pengujian model:

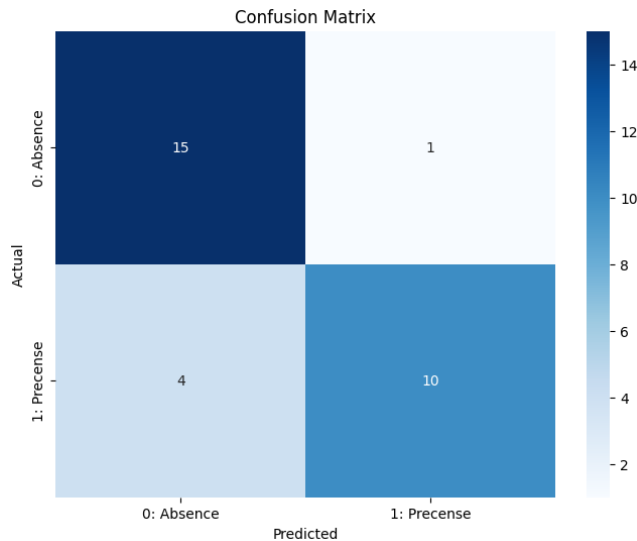


Gambar 5. Grafik Akurasi dan Loss untuk training dan validation

## Membership functions



Gambar 6. Fungsi Keanggotaan untuk masing-masing *input*



Gambar 7. Confusion Matrix dari pengujian model

Classification Report:				
	precision	recall	f1-score	support
0: Absence	0.79	0.94	0.86	16
1: Precense	0.91	0.71	0.80	14
accuracy			0.83	30
macro avg	0.85	0.83	0.83	30
weighted avg	0.85	0.83	0.83	30
Accuracy Score: 0.8333333333333334				
Precision: 0.9090909090909091				
Recall: 0.7142857142857143				
F1-Score: 0.8				

Gambar 8. Metrik dari *classification report*

Hasil pelatihan menunjukkan bahwa model ANFIS dengan parameter `n_memb: 5` dan `memb_func: gaussian` dapat memberikan akurasi sebesar 83% dengan metrik F1-Score 80%.

#### 4.3. Komparasi dengan Metode Lain

Terdapat beberapa metode lain yang sudah diterapkan sebelumnya untuk mengklasifikasikan *dataset Heart Disease* ini. Dilansir dari *repository* University of California, berikut data dan perbandingan dengan model ANFIS pada eksperimen ini:

Tabel 2. Tabel perbandingan akurasi dengan metode lain

Model	Metrik Akurasi		
	Batas bawah	Mean	Batas Atas
<b>Xgboost Classification</b>	72%	81%	89%
<b>Support Vector Classification</b>	55%	65%	76%
<b>Random Forest Classification</b>	71%	80%	88%
<b>Neural Network Classification</b>	69%	78%	88%
<b>Logistic Regression</b>	72%	81%	89%
<b>ANFIS + PCA</b>	80%	82%	83%

## 5. Kesimpulan

### 5.1. Ringkasan Temuan

Eksperimen ini menggunakan ANFIS yang dipadukan dengan PCA untuk mendeteksi penyakit jantung. Dua parameter utama, yaitu *n\_memb* (jumlah fungsi keanggotaan) dan *memb\_func* (jenis fungsi keanggotaan), dievaluasi melalui metode Grid Search sebanyak 10 iterasi. Parameter *n\_memb* diuji dengan nilai 2, 3, 4, 5, dan 6, sementara *memb\_func* mencakup Gaussian, Generalized Bell, dan Sigmoid. Hasil Grid Search menunjukkan bahwa kombinasi parameter terbaik adalah *n\_memb* = 5 dan *memb\_func* = Gaussian. Dengan kombinasi ini, model mencapai akurasi pengujian sebesar 83% dan F1-Score sebesar 80%.

### 5.2. Implikasi Praktis

Hasil penelitian ini dapat diterapkan dalam sistem pendukung keputusan klinis untuk mendukung diagnosis dini penyakit jantung. Model berbasis ANFIS dengan PCA memungkinkan analisis data pasien secara lebih akurat dan efisien, menghemat waktu dan biaya, serta cocok untuk digunakan di fasilitas kesehatan dengan sumber daya terbatas. Sistem ini juga dapat dikembangkan lebih lanjut untuk personalisasi diagnosis dan integrasi ke aplikasi berbasis AI untuk mendukung keputusan real-time.

### 5.3. Rekomendasi Untuk Penelitian Selanjutnya

Diharapkan pada penelitian selanjutnya untuk menggunakan *dataset* yang memiliki baris data lebih banyak untuk meningkatkan akurasi pada model dan mengurangi kemungkinan *overfitting* yang bisa berakibat kepada performa model.

### Daftar Pustaka

- Ciptaningrum, A., Arliawan, A., Nova Wima Sari, F., Bayu Muchlas Aditya, M., Zulkarnain, M., Indah Kusumawati, N., Budisatriya, R., Fajar Sari, S., Perkeretaapian, T., Negeri Madiun Jl Serayu No, P., Taman, K., Madiun, K., & Timur, J. (2022). Adaptive Neuro Fuzzy Inference System (ANFIS) terhadap Klasifikasi Kendaraan Roda Dua Mahasiswa Teknik Perkeretaapian Angkatan 2019 di Politeknik Negeri Madiun. In *JTE UNIBA* (Vol. 7, Issue 1).
- Ejaz, M. S., Islam, M. R., Sifatullah, M., & Sarker, A. (2019). Implementation of Principal Component Analysis on Masked and Non-masked Face Recognition. *1st International Conference on Advances in Science, Engineering and Robotics Technology 2019, ICASERT 2019*, 1.  
<https://doi.org/10.1109/ICASERT.2019.8934543>
- Hossain, R., & Timmer, D. (2021). *Machine Learning Model Optimization with Hyper Parameter Tuning Approach*.
- Janosi, A., Steinbrunn, W., Pfisterer, M., & Detrano, R. (1989). Heart Disease [Dataset]. UCI Machine Learning Repository.  
<https://doi.org/10.24432/C52P4X>.
- Lenhard, G. (2020). Adaptive-Network-Based Fuzzy Inference System (ANFIS) based on Keras on top of Tensorflow 2.0.
- Loganathan, C., & Girija, K. V. (2013). Hybrid Learning For Adaptive Neuro Fuzzy Inference System. In *Research Inventy: International Journal Of Engineering And Science* (Vol. 2, Issue 11).
- Opeyemi, O., & Justice, E. O. (2012). Development of Neuro-fuzzy System for Early Prediction of Heart Attack. *International Journal of Information Technology and Computer Science*, 4(9), 22–28.  
<https://doi.org/10.5815/ijitcs.2012.09.03>
- Ouifak, H., & Idri, A. (2023). Application of neuro-fuzzy ensembles across domains: A systematic review of the two last decades (2000–2022). In *Engineering Applications of Artificial Intelligence* (Vol. 124). Elsevier Ltd.  
<https://doi.org/10.1016/j.engappai.2023.106582>

Ufumaka, I. (2020). Comparative Analysis of Machine Learning Algorithms for Heart Disease Prediction. *International Journal of Scientific and Research Publications (IJSRP)*, 11(1), 339–346.

<https://doi.org/10.29322/ijsrp.11.01.2021.p10936>

Ziasabounchi, N., & Askerzade, I. (2014). ANFIS Based Classification Model for Heart Disease Prediction. In *International Journal of Electrical & Computer Sciences IJECS-IJENS* (Vol. 14).

## Lampiran

- Kodingan Google Collab :  
<https://colab.research.google.com/drive/1GHSvBpK0qPvvSQWBtzNsdS8wE7XnPPxi?usp=sharing>
- Dataset Heart Disease : <https://archive.ics.uci.edu/dataset/45/heart+disease>
- GitHub Library Adaptive-Network-Based Fuzzy Inference System (ANFIS) based on Keras on top of Tensorflow 2.0 :  
<https://github.com/gregorLen/AnfisTensorflow2.0>