

TESIS

**OPTIMASI STRUKTUR METODE ADAPTIVE NEURO FUZZY
INFERENCE SYSTEM MENGGUNAKAN ALGORITMA GENETIKA
DALAM MENINGKATKAN AKURASI PREDIKSI DATA TIME SERIES**



Disusun oleh:

Nama : Ma'ruf Aziz Muzani
NIM : 20.55.1415
Konsentrasi : Business Intelligence

PROGRAM STUDI S2 TEKNIK INFORMATIKA
PROGRAM PASCASARJANA UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA

2022

TESIS

**OPTIMASI STRUKTUR METODE ADAPTIVE NEURO FUZZY
INFERENCE SYSTEM MENGGUNAKAN ALGORITMA GENETIKA
DALAM MENINGKATKAN AKURASI PREDIKSI DATA TIME SERIES**

**STRUCTURE OPTIMIZATION OF ADAPTIVE NEURO FUZZY
INFERENCE SYSTEM METHOD USING GENETIC ALGORITHM TO
INCREASE TIME SERIES DATA PREDICTION ACCURACY**

Diajukan untuk memenuhi salah satu syarat memperoleh derajat Magister



Disusun oleh:

Nama : Ma'ruf Aziz Muzani
NIM : 20.55.1415
Konsentrasi : Business Intelligence

**PROGRAM STUDI S2 TEKNIK INFORMATIKA
PROGRAM PASCASARJANA UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2012**

HALAMAN PENGESAHAN

**OPTIMASI STRUKTUR METODE ADAPTIVE NEURO FUZZY INFERENCE
SYSTEM MENGGUNAKAN ALGORITMA GENETIKA DALAM
MENINGKATKAN AKURASI PREDIKSI DATA TIME SERIES**

**STRUCTURE OPTIMIZATION OF ADAPTIVE NEURO FUZZY INFERENCE
SYSTEM METHOD USING GENETIC ALGORITHM TO INCREASE TIME
SERIES DATA PREDICTION ACCURACY**

Dipersiapkan dan Disusun oleh

Ma'ruf Aziz Muzani

20.55.1415

Telah Diujikan dan Dipertahankan dalam Sidang Ujian Tesis
Program Studi S2 Teknik Informatika
Program Pascasarjana Universitas AMIKOM Yogyakarta
pada hari Senin, 5 Desember 2022

Tesis ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Magister Komputer

Yogyakarta, 5 Desember 2022

Rektor

Prof. Dr. M. Suyanto, M.M.
NIK. 190302001

HALAMAN PERSETUJUAN

OPTIMASI STRUKTUR METODE ADAPTIVE NEURO FUZZY INFERENCE SYSTEM MENGGUNAKAN ALGORITMA GENETIKA DALAM MENINGKATKAN AKURASI PREDIKSI DATA TIME SERIES

STRUCTURE OPTIMIZATION OF ADAPTIVE NEURO FUZZY INFERENCE SYSTEM METHOD USING GENETIC ALGORITHM TO INCREASE TIME SERIES DATA PREDICTION ACCURACY

Dipersiapkan dan Disusun oleh

Ma'ruf Aziz Muzani

20.55.1415

Telah Diujikan dan Dipertahankan dalam Sidang Ujian Tesis
Program Studi S2 Teknik Informatika
Program Pascasarjana Universitas AMIKOM Yogyakarta
pada hari Senin, 5 Desember 2022

Pembimbing Utama

Prof. Dr. Ema Utami, S.Si., M.Kom.
NIK. 190302037

Anggota Tim Penguji

Dr. Andi Sunyoto, M.Kom.
NIK. 190302052

Pembimbing Pendamping

Angrit Dwi Hartanto, M.Kom.
NIK. 190302163

Hanafi, S.Kom., M.Eng., Ph.D.
NIK. 190302024

Prof. Dr. Ema Utami, S.Si., M.Kom
NIK. 190302037

Tesis ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Magister Komputer

Yogyakarta, 5 Desember 2022

Direktur Program Pascasarjana

Dr. Kusrini, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN TESIS

Yang bertandatangan di bawah ini,

Nama mahasiswa : Ma'ruf Aziz Muzani
NIM : 20.55.1415
Konsentrasi : Business Intelligence

Menyatakan bahwa Tesis dengan judul berikut:
**OPTIMASI STRUKTUR METODE ADAPTIVE NEURO FUZZY INFERENCE
SYSTEM MENGGUNAKAN ALGORITMA GENETIKA DALAM
MENINGKATKAN AKURASI PREDIKSI DATA TIME SERIES**

Dosen Pembimbing Utama : Prof. Dr. Ema Utami, S.Si., M.Kom.
Dosen Pembimbing Pendamping : Anggit Dwi Hartanto, M.Kom.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Tim Dosen Pembimbing
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi

Yogyakarta, 5 Desember 2022

Yang Menyatakan,



Ma'ruf Aziz Muzani

HALAMAN PERSEMBAHAN

Tesis ini saya persembahkan untuk:

Yang paling pertama terima kasih kepada Allah SWT atas segala nikmat dan rahmatnya, sehingga tesis ini dapat terselesaikan.

Tak lupa juga terimakasih kepada kedua orang tua saya, yang tanpa lelah sudah mendukung apapun keputusan dan pilihan saya dalam menjalani kehidupan serta doa yang tidak pernah putus kepada saya.

Sahabat-sahabat terbaik saya yang telah senantiasa memberikan arahan, motivasi, serta dukungan yang sangat berguna bagi saya.

Rekan-rekan PT GIT Solution Terimakasih atas ilmu dan kesempatannya sehingga saya dapat turut andil dalam sebuah amanah.

HALAMAN MOTTO

"Your Future Is Created By What You Do Today Not Tomorrow"

"Fabiyyi 'aalaa' i Rabbikumaa Tukadzdzibaan"

"Maka nikmat Tuhan kamu yang manakah yang kamu dustakan"

Ma'ruf Aziz Muzani



KATA PENGANTAR

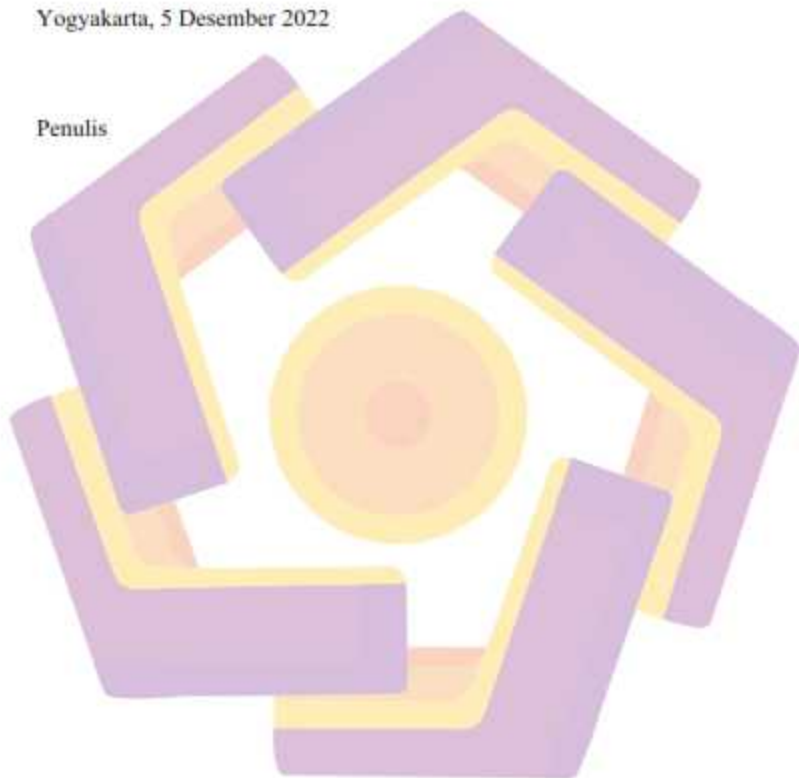
Puji syukur atas kehadiran Allah SWT yang telah memberikan rahmat-Nya sehingga penulis dapat menyelesaikan laporan penelitian tesis ini dengan baik. Penulisan laporan tesis ini dapat terselesaikan berkat bantuan dari berbagai pihak. Oleh karena itu penulis mengucapkan terimakasih kepada pihak-pihak yang terlibat dalam penelitian ini:

1. Prof. Dr. M. Suyanto, MM. selaku Rektor Universitas AMIKOM Yogyakarta.
2. Ibu Prof. Dr. Kusrini, M.Kom. selaku Direktur Program Pascasarjana Universitas AMIKOM Yogyakarta.
3. Ibu Prof. Dr. Ema Utami, S.Si., M.Kom. selaku Wakil Direktur Program Pascasarjana Universitas AMIKOM Yogyakarta sekaligus selaku Pembimbing utama.
4. Bapak Anggit Dwi Hartanto, M.Kom. selaku dosen Pembimbing Pendamping.
5. Bapak Dr. Andi Sunyoto, M.Kom., selaku penguji Seminar Hasil dan Ujian Tesis.
6. Bapak Hanafi, S.Kom., M.Eng., Ph.D., selaku penguji Seminar Hasil dan Ujian Tesis.
7. Bapak Dr. Kumara Ari Yuana, S.T, M.T., selaku penguji Seminar Proposal.
8. Bapak Hanif Al Fatta, M.Kom., selaku penguji Seminar Proposal.
9. Serta semua pihak yang tidak bisa penulis sebutkan satu persatu terima kasih atas doa dan sarannya.

Dengan diiringi doa dan ucapan terimakasih, penulis berharap semoga tesis ini dapat bermanfaat. Saran, harapan, kritik yang membangun selalu penulis untuk perbaikan di masa yang akan datang. Terimakasih

Yogyakarta, 5 Desember 2022

Penulis



DAFTAR ISI

HALAMAN JUDUL	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN	iv
HALAMAN PERNYATAAN KEASLIAN TESIS	v
HALAMAN PERSEMBAHAN	vi
HALAMAN MOTTO	vii
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xv
DAFTAR ISTILAH	xvi
INTISARI	xviii
<i>ABSTRACT</i>	xix
BAB I PENDAHULUAN	1
1.1. Latar Belakang Masalah	1
1.2. Rumusan Masalah	7
1.3. Batasan Masalah	8
1.4. Tujuan Penelitian	8
1.5. Manfaat Penelitian	9
1.5.1. Manfaat Teoritis	9
1.5.2. Manfaat Praktik	9

BAB II TINJAUAN PUSTAKA	10
2.1. Tinjauan Pustaka.....	10
2.2. Keaslian Penelitian	14
2.3. Landasan Teori	28
2.3.1. Data Runtut (<i>Time Series</i>).....	28
2.3.2. Prediksi (<i>Forecasting</i>).....	29
2.3.3. Teknik Prediksi	30
2.3.4. Tahapan Prediksi.....	30
2.3.5. Tujuan Prediksi	31
2.3.6. ANFIS	31
2.3.7. Arsitektur ANFIS.....	32
2.3.8. Algoritma Pembelajaran Hybrid	36
2.3.9. Pembelajaran Arah Maju (<i>LSE Recursive</i>).....	37
2.3.10. Pembelajaran Arah Mundur (<i>Error Backpropagation</i>)	38
2.3.11. Pengukuran Hasil Akurasi	42
2.3.12. Algoritma Genetika.....	45
BAB III METODE PENELITIAN	48
3.1. Jenis, Sifat, dan Pendekatan Penelitian	48
3.1.1. Jenis Penelitian.....	48
3.1.2. Sifat Penelitian	48
3.1.3. Pendekatan Penelitian	48

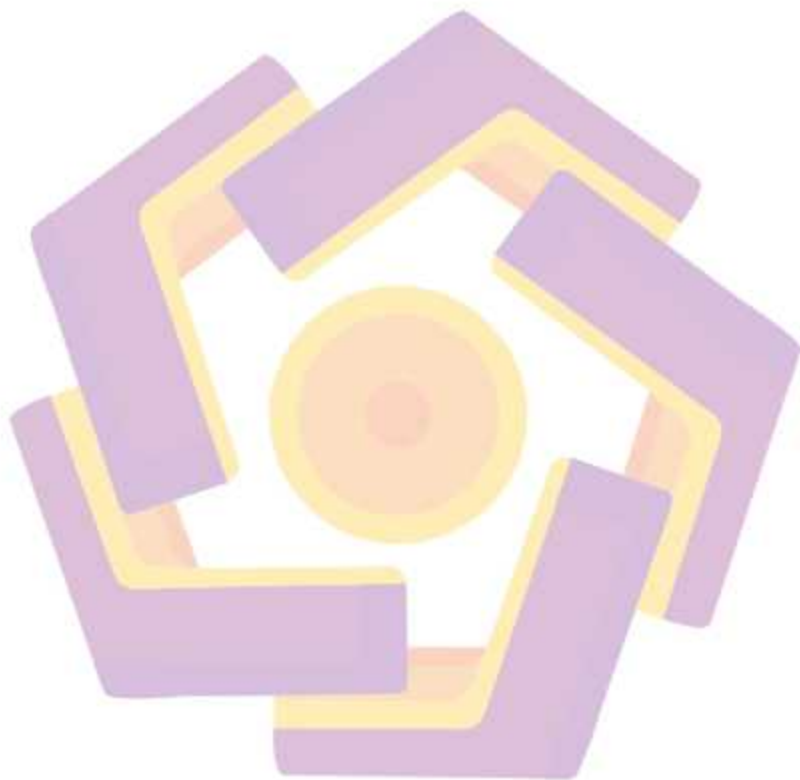
3.2. Metode Pengumpulan Data	49
3.3. Metode Analisis Data	49
3.4. Dataset	50
3.5. Alur Penelitian	53
3.6. Alur Proses Genetika	58
BAB IV HASIL PENELITIAN DAN PEMBAHASAN	61
4.1. Pengumpulan Data	61
4.2. Pengolahan Data	64
4.3. Mekanisme Training	66
4.4. Model ANFIS	69
4.4.1. Pelatihan Data	69
4.4.2. Pengujian Data	74
4.4.3. Perhitungan Akurasi	78
4.5. Model ANFIS – GA	79
4.5.1. Pemilihan Operator Genetika	80
4.5.2. Proses Pelatihan dan Pengujian ANFIS - GA	81
4.5.3. Analisa Perhitungan Akurasi	85
4.6. Pembahasan Hasil	86
BAB V PENUTUP	91
5.1. Kesimpulan	91
5.2. Saran	92
DAFTAR PUSTAKA	93

DAFTAR TABEL

Tabel 2.1. Matriks literatur review dan posisi penelitian Optimasi Struktur Metode Adaptive Neuro Fuzzy Inference System Menggunakan Algoritma Genetika Dalam Meningkatkan Akurasi Prediksi Data Time Series.....	14
Tabel 2.2. Proses Pembelajaran Hybrid.....	37
Tabel 3.1. Contoh dataset.....	51
Tabel 3.2. Contoh dataset dengan variabel yang digunakan.....	51
Tabel 4.1. Dataset.....	62
Tabel 4.2. Dataset setelah preprocessing.....	65
Tabel 4.3. Dataset normalisasi.....	66
Tabel 4.4. Variasi Skenario Training Data.....	68
Tabel 4.5. Parameter ANFIS.....	68
Tabel 4.6. Sampel hasil pelatihan.....	70
Tabel 4.7. Sampel hasil pelatihan dengan proses normalisasi.....	70
Tabel 4.8. Sampel hasil pengujian.....	74
Tabel 4.9. Sampel hasil pengujian dengan normalisasi.....	75
Tabel 4.10. Perbandingan Akurasi.....	78
Tabel 4.11. Pengujian skema 80:20.....	82
Tabel 4.12. Pengujian skema 70:30.....	82
Tabel 4.13. Pengujian skema 80:20 dengan normalisasi.....	83
Tabel 4.14. Pengujian skema 70:30 dengan normalisasi.....	83
Tabel 4.15. Hasil seluruh uji yang dilakukan.....	87

Tabel 4.16. Perbandingan model ANFIS dengan penelitian lain88

Tabel 4.17. Perbandingan optimasi ANFIS89



DAFTAR GAMBAR

Gambar 2.1. Struktur ANFIS.....	33
Gambar 2.2. Alur algoritma genetika	46
Gambar 3.1. Tahapan <i>preprocessing</i>	53
Gambar 3.2. Alur penelitian	53
Gambar 3.3. Struktur ANFIS.....	56
Gambar 3.4. Alur proses genetika	58
Gambar 4.1. Sebaran data temperatur.....	63
Gambar 4.2. Sebaran data kelembapan.....	63
Gambar 4.3. Sebaran data kecepatan angin.....	64
Gambar 4.4. Sebaran data curah hujan	64
Gambar 4.5. Grafik training data 80:20.....	71
Gambar 4.6. Grafik training data 70:30.....	72
Gambar 4.7. Grafik training dataset 80:20 dengan normalisasi	73
Gambar 4.8. Grafik training dataset 70:30 dengan normalisasi	73
Gambar 4.9. Testing dataset 80:20	76
Gambar 4.10. Testing dataset 70:30	76
Gambar 4.11. Testing dataset 80:20 dengan normalisasi	77
Gambar 4.12. Testing dataset 70:30 dengan normalisasi	77
Gambar 4.13. Pengujian jumlah generasi untuk mf 4.....	84
Gambar 4.14. Pengujian jumlah generasi dengan normalisasi untuk mf 4.....	85

DAFTAR ISTILAH

Dataset: Kumpulan data yang umumnya digunakan dalam penelitian di bidang data science, machine learning dan deep learning. Ada dataset yang bersifat public, yakni bisa diakses secara bebas dan ada pula yang bersifat private.

Time series: Merupakan serial dari kumpulan data yang teratur oleh urutan waktu bisa meliputi tahunan, bulanan, jam, atau bahkan mili-detik

Model: Istilah umum yang digunakan untuk merepresentasikan kecerdasan buatan yang sudah dilatih.

Epoch: Sebuah kecerdasan buatan dapat dilatih dengan data yang sama berkali-kali. Setiap menyelesaikan satu kali latihan (epoch), nilai weight atau bobot pada setiap neuron diperbaharui agar dapat nilai keluarannya dapat mendekati output yang seharusnya berdasarkan label data

Genotype (Gen): Sebuah nilai yang menyatakan satuan dasar yang membentuk suatu arti tertentu dalam satu kesatuan gen yang dinamakan kromosom. Gen bisa berupa nilai biner, float, integer, maupun karakter, atau kombinatorial.

Kromosom: Gabungan gen – gen yang membentuk suatu nilai tertentu.

Individu: Suatu nilai atau keadaan yang menyatakan salah satu solusi yang mungkin dari permasalahan yang diangkat.

Populasi: Sekumpulan individu yang akan diproses bersama dalam satu siklus proses evolusi.

Fitness: Menyatakan seberapa baik nilai dari suatu individu atau solusi yang didapatkan.

Generasi: Menyatakan satu – satuan siklus proses evolusi.

Crossover: Sebuah proses mengkombinasikan dua individu untuk memperoleh individu – individu baru yang diharapkan mempunyai nilai *fitness* yang lebih baik.



INTISARI

ANFIS merupakan gabungan dari 2 metode yaitu fuzzy dan jaringan syaraf tiruan. Dari penggabungan 2 metode tersebut ANFIS memiliki kelebihan dalam melakukan prediksi data time series. Namun penggunaan metode ANFIS juga terdapat kekurangan. Kekurangan tersebut yaitu jika jumlah aturan samar dalam skala besar yang didasari dengan adanya jumlah membership function dan jumlah masukan yang besar. Permasalahan tersebut disebut dengan "*curse of dimensionality*". Proses optimasi merupakan suatu usaha untuk mengatasi permasalahan yang terjadi sebelumnya pada metode ANFIS. Optimasi dilakukan dengan menggunakan Algoritma genetika yang dalam prosesnya menggantikan algoritma *steepest descent*.

Penelitian ini menggunakan dataset curah hujan kabupaten sleman dengan dataset yang didapat berjumlah 3045 baris data. Data yang diperoleh akan dilakukan preprocessing lalu kemudian akan diproses menggunakan algoritma genetika. Dalam pengujian data akan dibagi menjadi 2 skema yaitu dengan rasio 80:20 dan 70:30 serta data dengan normalisasi dan tanpa normalisasi.

Berdasarkan hasil penerapan metode ANFIS yang di optimasi dengan algoritma genetika dan normalisasi pada dataset curah hujan kabupaten sleman diterapkan pada beberapa skenario pengujian. Hasil dari pengujian dengan berbagai skenario diperoleh nilai error terendah sebesar 2,046 untuk ANFIS-GA dengan proses normalisasi data.

Kata kunci: ANFIS, Algoritma Genetika, Normalisasi, Optimasi

ABSTRACT

ANFIS is a combination of 2 methods, namely fuzzy and artificial neural networks. From the combination of these two methods, ANFIS has advantages in predicting time series data. However, the use of the ANFIS method also has drawbacks. The drawback is if the number of vague rules on a large scale is based on the number of membership functions and a large number of inputs. This problem is called the "curse of dimensionality". The optimization process is an attempt to overcome the problems that occurred previously in the ANFIS method. Optimization is done by using a genetic algorithm which in the process replaces the steepest descent algorithm.

This study uses a rainfall dataset of Sleman district with a dataset of 3045 rows of data obtained. The data obtained will be preprocessed and then will be processed using a genetic algorithm. In testing the data will be divided into 2 schemes, namely with a ratio of 80:20 and 70:30 and data with normalization and without normalization.

Based on the results of the application of the ANFIS method which was optimized with genetic algorithms and normalization on the rainfall dataset of Sleman Regency, it was applied to several test scenarios. The results of testing with various scenarios obtained the lowest error value of 2.046 for ANFIS-GA with the data normalization process.

Keyword: ANFIS, Genetic Algorithm, Normalization, Optimization

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Data Time Series atau dapat disebut dengan data rentet waktu merupakan sebuah ilmu yang digunakan dalam dunia statistika serta pemrosesan sinyal. Time Series merupakan serangkaian informasi yang berbentuk nilai dari hasil pengamatan yang diukur dalam kurun waktu tertentu, bersumber pada waktu dengan interval yang sama (Piegel L.J dkk., 2007). Dalam data time series dipergunakan data siklus hari, minggu, bulan, tahun, dan rentet waktu lainnya (Cryer & Chan, 2008).

Implementasi time series sering kali dipergunakan dalam bidang peramalan ataupun forecasting dimana model time series mengamati sikap observasi yang diambil secara berurutan (Wei, 2006). Model informasi time series sering kali diaplikasikan untuk penelitian terkait dengan peramalan atau prediksi seperti penelitian yang dilakukan oleh (Anggraini, 2018) dengan melakukan penelitian yang mangulas tentang peramalan curah hujan dengan memanfaatkan informasi time series multivariate. Analisis informasi time series membolehkan kita untuk bisa mendapat informasi mengenai sesuatu perkembangan ataupun peristiwa dan pengaruhnya terhadap peristiwa lainnya (Halimi & Kusuma, 2018). Pergerakan informasi ataupun nilai variabel bisa dikenal dengan terdapatnya informasi time series, sehingga bisa dijadikan sebagai pertimbangan untuk membuat keputusan dan juga perencanaan pada waktu yang akan datang karna pola pada informasi time

series dipercaya di masa depan pola yang pernah terjadi akan kembali terulang (Gen & Cheng, 1997).

Analisis data time series dapat diuraikan untuk memisahkan informasi masa lalu menjadi beberapa bagian dan memasukkannya ke dalam peramalan atau forecasting (M.E Render B., 2003). Dengan kata lain analisis data time series memiliki tujuan untuk mengenali faktor-faktor bagian yang dapat mempengaruhi nilai, sehingga sering kali diaplikasikan untuk mengantisipasi dalam jangka pendek ataupun dalam jangka panjang (Subanar & Suhartono, 2009).

Forecasting merupakan suatu kegiatan untuk meramalkan, melakukan, atau mengukur kemungkinan hasil yang berbeda yang akan terjadi di kemudian hari (R.K Buffa, 2007). Penerapan forecasting dapat dengan cara mengaitkan proses penggalan informasi masa lampau kemudian meramalkan masa yang akan datang. Proses tersebut dilakukan dalam perhitungan matematis (Cryer & Chan, 2008). Dalam forecasting dikategorikan menjadi 2 metode yakni kuantitatif dan kualitatif. Dalam metode kuantitatif terdapat beberapa metode dan salah satunya merupakan metode time series (Lee & Suhartono, 2012).

Pengolahan data time series dapat dicoba dengan memakai metode soft computing. Metode tersebut merupakan pendekatan yang mengaplikasikan komputasi. Metode tersebut meniru pola pikir yang dimiliki manusia atau individu serta mempunyai keahlian berfikir serta mampu mempelajari suatu ketidakpastian yang terjadi di lingkungan sekitar (Jang dkk., 1997). Sistem fuzzy merupakan salah satu bagian utama dalam soft computing. Selain itu terdapat juga komponen lain

yang juga merupakan bagian utama yakni algoritma evolusioner serta jaringan syaraf tiruan.

Salah satu metode untuk peramalan data rentet waktu (time series) yang bisa digunakan guna melakukan prediksi atau forecasting yakni meliputi metode ANFIS yang merupakan metode soft computing dengan penerapan penggabungan logika fuzzy dengan jaringan syaraf tiruan (JST) (Jang dkk., 1997). Selain ANFIS, terdapat beberapa metode yang populer untuk mengelola data time series yaitu meliputi Support Vector Machine (SVM) (Sang dkk., 2008), Neural Network (NN) (Chang dkk., 2007), dan Autoregressive Integrated Moving Average (ARIMA) (Diaz-Robles dkk., 2008).

ANFIS memiliki keahlian untuk menanggulangi sistem yang kompleks, nonlinear serta dapat menyesuaikan waktu dengan menggunakan algoritma pembelajaran terhadap informasi yang dihasilkan dari sistem. ANFIS dapat digunakan guna mengaplikasikan proses peramalan informasi time series. Informasi yang mampu diramalkan dapat berupa data stasioner, random, non stasioner serta data yang terkait waktu atau data musim. NN mempunyai kelebihan dalam melakukan prediksi nonlinear, kemampuan parallel processing yang kuat serta mempunyai kemampuan untuk menerima suatu kesalahan, tetapi NN mempunyai kelemahan yaitu diperlukannya data training dengan skala yang besar, terdapat suatu proses yang membuat model sangat akurat, proses konvergensi yang lama (Sanay dkk., 2007). ARIMA merupakan suatu model tanpa teori dengan variabel yang digunakan merupakan nilai lampau dan kesalahan yang mengikutinya. SVM mempunyai kelebihan untuk memecahkan over fitting, proses

konvergensi yang lama, dan data training yang sedikit (Vapnik, 2000). Tetapi metode SVM mempunyai kekurangan yaitu sulitnya menentukan parameter terbaik dan optimal (Coussement & van den Poel, 2006).

Menurut referensi yang didapat, pengolahan data time series untuk forecasting sudah banyak dilakukan penelitian sebelumnya dengan menerapkan berbagai metode seperti yang dilakukan oleh (Angraini, 2018) Penelitian bertujuan untuk memprediksi curah hujan dengan data rentet waktu (time series) multivariate dengan membandingkan setiap membership function yang terdapat pada ANFIS. membership function yang dibandingkan meliputi membership function trapmf, trimf, gaussmf serta bellmf. Dari hasil perbandingan tersebut didapatkan hasil bahwa membersip function trimf (segitiga) mendapatkan nilai error terkecil dibandingkan dengan membership function yang lain. Pada penelitian tersebut didapatkan nilai error terkecil dengan RMSE (root mean score error) sebesar 0,73436. Penelitian lain yang dilakukan oleh (Hadiansyah, 2017) menerapkan pemodelan time series ARIMA untuk memprediksi harga cabai. Data yang digunakan berjumlah 236 data yang diperoleh dari periode 1 mei 2016 hingga 14 januari 2017. Hasil yang didapatkan dengan data training menggunakan model ARIMA (1 2 1) mendapatkan nilai RMSE sebesar 964,005 dan nilai MAPE sebesar 1,479%. Dalam pengujian lain yang dilakukan pada data testing dengan berdasarkan 3 pengukuran nilai error yaitu dengan RMSE, MAPE, serta R-Squared, model ARIMA (1) merupakan model yang efektif dalam penelitian ini dengan nilai RMSE sebesar 305,348, nilai MAPE sebesar 0,651% dan nilai R-Squared 0,635. Pada penelitian lain yang dilakukan (Nikentari dkk., 2018) menerapkan Adaptive

Neuro Fuzzy (ANFIS) serta dalam penelitian tersebut juga diterapkan metode Radial Basis Function Neural Network (RBFNN). Penelitian tersebut bertujuan meramalkan kecepatan angin. Informasi rata-rata kecepatan angin pada tahun 2017 dengan jumlah data sebanyak 296 data dari bulan Januari hingga Oktober digunakan sebagai dataset dalam penelitian ini. Dari jumlah data tersebut kemudian untuk menguji model maka dataset akan dikelompokkan menjadi data latih (train) dan data uji (test). Pembagian data yang digunakan yaitu dengan membagi 80% data, serta menempatkan 20% sisa dataset untuk dialokasikan sebagai data uji (test). Pada penelitian tersebut didapatkan hasil error terendah pada metode RBFNN sebesar 0,177 dan untuk RMSE pada metode ANFIS sebesar 1,146.

Pemodelan ANFIS dibagi menjadi 2 bagian, yakni identifikasi struktur serta optimisasi parameter (Lutfy dkk., 2011). Identifikasi struktur meliputi penentuan banyaknya MF serta aturan samar dengan menyusun segala gabungan ketentuan sesuai banyaknya masukan serta fungsi keanggotaan (MF) pada tiap masukan (Jang dkk., 1997). Sehingga apabila dalam suatu sistem terdapat (n) yang merupakan banyaknya variabel input kemudian (p) yang merupakan jumlah membership function dari masing-masing masukan, maka akan dihasilkan hukum samar sebesar (pn). Terdapat akibat yang ditimbulkan dari penggunaan metode ini. Akibat yang ditimbulkan yakni terbentuknya jumlah aturan samar dalam skala besar yang didasari dengan adanya jumlah membership function dan jumlah masukan yang besar. Selain itu dengan jumlah membership yang besar maka akan membutuhkan waktu yang relatif lebih lama untuk pemrosesan data. Untuk mengatasi akibat yang ditimbulkan tersebut atau yang dapat disebut sebagai permasalahan "curse of

dimensionality", banyak riset yang dicoba guna menghasilkan jumlah yang sedikit atau minimum pada aturan samar. Untuk mengurangi jumlah aturan samar, dapat dengan cara membuat aturan samar sebesar membership function atau membuat aturan samar menjadi sebesar (p) (Liu dkk., 2002).

Proses optimasi merupakan suatu usaha untuk mengatasi permasalahan yang terjadi sebelumnya pada metode ANFIS. Optimasi dapat dilakukan pada parameter yang digunakan dalam proses pengolahan data menggunakan ANFIS dengan memanfaatkan keahlian pada sistem neural yakni keahlian untuk mempelajari hal baru yang bertujuan untuk memperbarui nilai pada parameter premis dan juga memperbarui nilai parameter konsekuen yang sebelumnya dihasilkan dari struktur ANFIS menggunakan algoritma LSE. Dalam metode ANFIS algoritma hybrid merupakan suatu algoritma yang digunakan untuk melakukan pembelajaran arah maju dan mundur (Jang dkk., 1997). Algoritma steep descent (SD) merupakan salah satu algoritma yang digunakan untuk mengenali parameter liner, selain itu terdapat juga algoritma least square estimator (LSE) yang berfungsi sama. Kedua algoritma tersebut dikombinasikan sehingga membentuk proses pelatihan data untuk mengenali parameter liner pada metode ANFIS. Dalam riset ini, algoritma yang digunakan untuk membentuk proses pelatihan data bukan lagi menerapkan kombinasi antara algoritma SD dan juga LSE, namun terdapat pergantian di salah satu algoritma dimana algoritma SD akan digantikan dengan algoritma genetika (GA) yang bertujuan sesuai tujuan utama yaitu untuk mendapat parameter terbaik yang hendak digunakan pada metode ANFIS. Penggunaan algoritma genetika dipergunakan untuk menggantikan proses pembelajaran arah mundur dengan

algoritma steepest descent. Algoritma genetika merupakan suatu teknik yang mengambil prinsip – prinsip proses genetika yang terdapat dalam makhluk hidup, teknik tersebut merupakan pertumbuhan generasi yang berlangsung dalam suatu populasi dimana proses tersebut mengikuti dan menerapkan prinsip seleksi alam. Penggunaan prinsip tersebut dengan maksud untuk mendapatkan jalur pemecahan kasus yang berlangsung di dalam dunia nyata (Gen & Cheng, 1997).

Berdasarkan masalah yang sudah dijelaskan diatas, penelitian ini menerapkan optimasi yang bertujuan untuk memperbaiki struktur metode ANFIS sehingga dapat meningkatkan nilai hasil akurasi dengan menggunakan Algoritma Genetika. Penerapan algoritma genetika dimaksudkan untuk mengoptimasi proses pelatihan struktur ANFIS yang hasil dari proses tersebut berupa parameter parameter utama yang nantinya akan digunakan dalam proses ANFIS.

1.2. Rumusan Masalah

Berdasarkan pemaparan pada latar belakang, dalam penelitian ini terdapat rumusan masalah sebagai berikut:

1. Apakah penggunaan Algoritma Genetika sebagai algoritma optimasi memberikan dampak terhadap performa metode ANFIS?
2. Parameter Algoritma Genetika apa saja yang memberikan hasil terbaik dalam optimasi ANFIS?
3. Bagaimana penerapan Algoritma Genetika dalam proses optimasi ANFIS pada dataset *Time Series Multivariate*?

1.3. Batasan Masalah

Batasan – batasan pada penelitian ini agar pembahasan masalah dapat teratur dan tidak menyimpang ditetapkan sebagai berikut:

- a. Penentuan prediksi dilakukan menggunakan metode ANFIS
- b. *Membership function* yang digunakan adalah *membership function Gause*
- c. Algoritma yang digunakan untuk optimasi yaitu menggunakan Algoritma Genetika
- d. Jenis *Crossover* yang digunakan yaitu *Uniform Crossover*
- e. Model seleksi yang digunakan meliputi *Tournament Selection*
- f. Data yang digunakan acuan berdasarkan data dari website BMKG
- g. Dataset yang digunakan yaitu data curah hujan pada periode januari 2014 sampai dengan September 2022
- h. Parameter yang digunakan meliputi temperatur, kelembapan, kecepatan angin, dan curah hujan
- i. Optimasi yang dilakukan berfokus untuk mendapat parameter premis dan parameter konsekuen yang optimal
- j. Algoritma genetika digunakan untuk menggantikan Algoritma *steepest descent* pada proses pembentukan parameter premis dan konsekuen
- k. Pengukuran akurasi error menggunakan *Root Mean Squared Error (RMSE)*.
- l. Pembagian data pelatihan dan data pengujian yaitu sebesar 80:20 dan 70:30

1.4. Tujuan Penelitian

Tujuan dari penelitian ini adalah:

- a. Mengetahui faktor apa saja yang dapat mengurangi nilai akurasi pada metode ANFIS.
- b. Mengetahui tingkat akurasi yang didapatkan dengan menerapkan algoritma genetika untuk mengoptimasi struktur proses pelatihan metode ANFIS.
- c. Mengembangkan metode ANFIS dengan menggunakan Algoritma Genetika.

1.5. Manfaat Penelitian

1.5.1. Manfaat Teoritis

- a. Membuat model untuk memprediksi data *time series*
- b. Dapat menjadi rekomendasi untuk penelitian mendatang tentang pengolahan data *time series*

1.5.2. Manfaat Praktik

- a. Mampu menguasai dan mengimplementasikan metode prediksi data rentet waktu (*time series*) pada kasus nyata.
- b. Mendapat informasi ramalan curah hujan kabupaten sleman di masa yang akan datang.
- c. Penelitian ini diharapkan menambah wawasan ilmu pengetahuan dalam peramalan curah hujan dan mampu dijadikan rujukan bagi penelitian yang akan datang.

BAB II

TINJAUAN PUSTAKA

2.1. Tinjauan Pustaka

(Wahyuni dkk., 2017) mengimplementasikan algoritma genetika untuk mengoptimasi FIS Sugeno dalam batasan fungsi keanggotaan sebelum masuk ke dalam proses pelatihan ANFIS untuk memprediksi curah hujan. Dalam penelitian tersebut dengan menggunakan 2 parameter masukan, ANFIS-GA mendapat nilai RMSE rendah dengan rincian metode G-Star (4.9), Tsukamoto FIS (8.95), Tsukamoto FIS & ANFIS-GA (7.5), dan GA (5.87) untuk daerah puspo, Tengger.

(Aziz dkk., 2017) melakukan penelitian untuk meningkatkan prediksi energi biochar. Dalam penelitian tersebut menerapkan metode ANFIS. Dalam penelitian ini dilakukan proses optimasi metode pada ANFIS menggunakan algoritma optimasi partical swarm optimization (PSO). Pendekatan ANFIS PSO digunakan untuk memaksimalkan produksi biochar. Pendekatan ANFIS PSO memiliki RMSE dan AAPRE yang jauh lebih kecil yaitu masing-masing 0,2673 dan 3,4529%. Juga, pendekatan RMSE untuk ANFIS berdasarkan GA adalah yang terbaik daripada NN dan LS-SVM. Namun, perbandingan antara regresi nonlinier dan ANFIS, regresi terbaik dalam semua hal (AAPRE, R2, dan RMSE) Dari penelitian tersebut dengan penerapan ANFIS-PSO menghasilkan nilai RMSE sebesar 0.2373 dan menunjukkan nilai dengan akurasi yang lebih baik dari metode lain seperti metode ANFIS, ANFIS-GA, LS-SVM, dan pendekatan ANN.

Dalam penelitian lain yang dilakukan oleh (Damayanti & Cahyadi, 2016) menerapkan metode ANFIS dengan mengaplikasikan optimasi pada struktur dengan memanfaatkan algoritma genetika yang bertujuan guna mengetahui secara dini terdapatnya defleksi yang ada pada proses pertumbuhan balita. Penelitian tersebut, dalam proses pembelajaran untuk menghasilkan parameter yang akan digunakan dalam metode ANFIS dengan menggunakan algoritma steepest descent akan digantikan dengan menggunakan algoritma genetika. Penggantian metode disini bertujuan guna mengoptimasi proses pelatihan pada struktur ANFIS. Dengan penerapan algoritma genetika untuk optimasi ANFIS dapat disimpulkan bahwa struktur ANFIS sama seperti jaringan neural dengan fungsi radial yang terdapat sedikit batas tertentu, ataupun bisa dikatakan ANFIS mampu melakukan proses pengolahan data menggunakan algoritma pelatihan tertentu terhadap kumpulan informasi. Pada penelitian ini dihasilkan nilai error sebesar 0,002899.

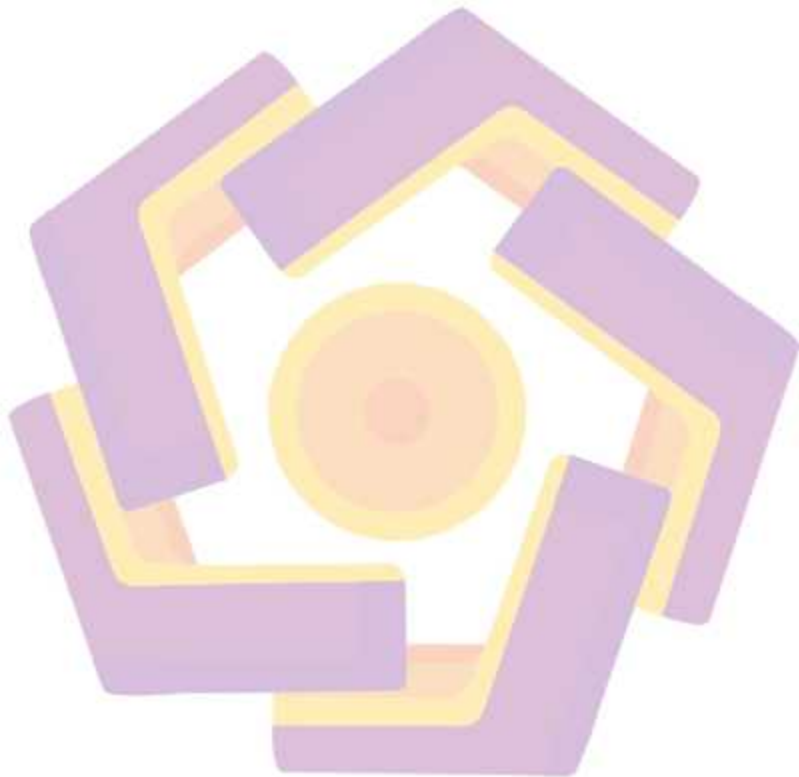
Penelitian yang dilakukan oleh (Anggraini, 2018) menerapkan metode ANFIS untuk prediksi data curah hujan. Hasil perhitungan prediksi yang dilakukan, mendekati data yang sebenarnya dengan nilai RMSE 0,73436. Nilai tersebut didapat dari fungsi keanggotaan *bellmf* yang mempunyai akurasi nilai tertinggi dari pada fungsi keanggotaan *trapmf*, *trimf* dan *gaussmf*. Dengan memanfaatkan informasi pendukung seperti kelembapan 58.4% serta temperatur 47.5o C dapat dihasilkan prediksi curah hujan yang akan terjadi dengan ukuran sebesar 32.2 mm yang artinya hujan berlangsung dengan intensitas rendah serta dari hasil akurasi error menggunakan RMSE mendapat nilai sebesar 0,734 pada fungsi segitiga yang memiliki error terkecil dari fungsi keanggotaan yang lainnya.

(Azhar & Mahmudy, 2018) juga menerapkan metode ANFIS guna memprediksi curah hujan yang terjadi di kabupaten malang. Parameter input yang digunakan dalam penelitian tersebut memanfaatkan data harian. Pada penelitian tersebut data dibagi menjadi 3 cluster yakni hujan dengan intensitas rendah, hujan dengan intensitas menengah, dan hujan dengan intensitas tinggi menggunakan algoritma K-Means. Berdasarkan hasil dari penelitian dengan menggunakan 150 data latih yang merupakan pengujian dengan nilai RMSE terbaik kemudian dengan jumlah iterasi sebanyak 400 dapat mampu menghasilkan error yang rendah dengan pengukuran menggunakan RMSE sebesar 1,88 menunjukkan data hasil prediksi mendekati data asli.

Hal yang sama dilakukan oleh (Sinaga dkk., 2019) menggunakan ANFIS sebagai metode untuk meramalan besaran curah hujan. Penelitian tersebut menerapkan parameter seperti fungsi keangotaan (MF), learning rate, tipe fungsi keanggotaan dan rasio data. Dari hasil perhitungan menghasilkan tingkatan error dengan akurasi yang baik dengan nilai MSE sebesar 0,063290962 dan 0,020498283 untuk rasio 3:2 dan 4:1 dengan menggunakan parameter MF, tipe MF, Learning Rate dengan rasio data yang masing masing menggunakan rasio 3:2 dan 4:1. Dari hal tersebut menunjukkan bahwa metode ANFIS mampu menjadi alternative metode untuk proses peramalan tingkat curah hujan dengan jenis data time series pada Kota Samarinda, Kalimantan.

Sehingga dalam penelitian ini akan melakukan optimasi terhadap struktur metode ANFIS dengan menggunakan algoritma genetika untuk meningkatkan akurasi prediksi pada data time series menggunakan pendekatan yang sama dengan

(Damayanti & Cahyadi, 2016). Perbedaan yang akan distruktur pengolahan data serta penerapan beberapa MF yang terdapat dalam metode anfis mengadaptasi dari (Sinaga dkk., 2019) dengan parameter yang berbeda. Serta dalam penelitian ini juga akan menggunakan beberapa skenario pengujian data.



2.2. Keaslian Penelitian

Tabel 2.1. Matriks literatur review dan posisi penelitian Optimasi Struktur Metode Adaptive Neuro Fuzzy Inference System Menggunakan Algoritma Genetika Dalam Meningkatkan Akurasi Prediksi Data Time Series

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
1	<i>Rainfall Prediction Using Hybrid Adaptive Neuro-Fuzzy Inference System (ANFIS) and Genetic Algorithm</i>	Ida Wahyuni, Wayan Firdaus Mahmady & Atiek Iriany, <i>Journal of Telecommunication, Electronic and Computer Engineering</i> , 2017	Peningkatan hasil dengan mengoptimalkan FIS Sugeno dalam batasan fungsi keanggotaan dan aturan konsekuensi sebelum masuk ke dalam proses pelatihan ANFIS dengan menggunakan algoritma genetika agar mendapat hasil prediksi yang lebih baik dari rediksi curah hujan sebelumnya	Dengan penerapan 2 proses optimasi pada fungsi keanggotaan serta 2 parameter masukan ANFIS GA didapatkan nilai RMSE rendah dengan rincian metode G-STAR mendapat nilai RMSE 4,9, Tsukamoto FIS mendapat nilai RMSE 8,95, Tsukamoto FIS & GA ANFIS mendapat nilai RMSE 7,5, dan GA mendapat nilai RMSE 5,87 untuk daerah Puspo.	Saran menambahkan jumlah skenario pengujian dan pelatihan agar bisa didapatkan skenario terbaik yang mampu menghasilkan akurasi kesalahan minimum	Penulis menggunakan penelitian sebagai dasar penerapan optimasi metode, penelitian berbeda dalam penggunaan dataset dan proses optimasi yang diterapkan pada metode ANFIS

Tabel 2.1. Matriks literatur review dan posisi penelitian
 Optimasi Struktur Metode Adaptive Neuro Fuzzy Inference System Menggunakan Algoritma Genetika Dalam Meningkatkan Akurasi Prediksi Data Time Series (Lanjutan)

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
2	<i>A hybrid forecasting system based on fuzzy time series and multi-objective optimization for wind speed forecasting</i>	Ping Jiang, Hufang Yang & Junji Heng, Applied Energy – ELSEVIER, 2019	Penerapan sistem peramalan hybrid dengan 3 modul (pemrosesan data, optimasi, dan peramalan) untuk meningkatkan akurasi dan stabilitas peramalan pada data deret waktu kecepatan angin	Hasil pengujian yang didapat untuk pencapaian akurasi dan stabilitas menunjukkan persentase kesalahan absolut rata-rata di bawah 4%.	Saran menambahkan preprocessing data untuk mengatasi data yang memiliki keacakan dan non stasioner yang cukup besar agar didapat data yang lebih normal.	Penulis menggunakan penelitian sebagai dasar terhadap penerapan time series, penelitian berbeda dalam penggunaan algoritma dan metode optimasi
3	<i>Optimization of EPB Shield Performance with Adaptive Neuro-Fuzzy Inference System and Genetic Algorithm</i>	Khalid Elbaz, Shui-Long Shen, Annan Zhou, Da-Jun Yuan and Ye-Shuang Xu, applied sciences – MDPI, 2019	Penerapan optimasi multi fungsi untuk memprediksi kinerja keamanan proses tunneling dengan menerapkan metode gabungan antara metode ANFIS dengan Algoritma Genetika	Hasil pada penelitian ini, disimpulkan bahwa metode hibrida ANFIS-GA memberikan akurasi yang cukup baik dengan nilai RMSE 0.11 dengan perbandingan metode ANFIS non GA mendapat nilai RMSE 0.15 pada data pengujian.	Saran menambahkan jumlah skenario pengujian dan pelatihan agar bisa didapatkan skenario terbaik yang mampu menghasilkan akurasi kesalahan minimum	penelitian berbeda dalam penggunaan dataset dan proses optimasi yang diterapkan pada metode ANFIS

Tabel 2.1. Matriks literatur review dan posisi penelitian
Optimasi Struktur Metode Adaptive Neuro Fuzzy Inference System Menggunakan Algoritma Genetika Dalam Meningkatkan Akurasi Prediksi Data Time Series (Lanjutan)

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
4	<i>Prediction of Biochar Yield Using Adaptive Neuro-Fuzzy Inference System with Particle Swarm Optimization</i>	Mohamed Abd El Aziz, Ahmed Monem Hemdan, dkk, IEEE PES-IAS PowerAfrica, 2017	Penggunaan sistem <i>adaptive neuro fuzzy inference system</i> (ANFIS) dengan <i>Particle Swarm Optimization</i> (PSO) untuk meningkatkan prediksi energi biochar	Hasil penelitian yang dilakukan oleh peneliti dengan menerapkan metode ANFIS – PSO menghasilkan nilai RMSE sebesar 0.2373 dan menunjukkan nilai yang lebih optimal dari metode lain yang serupa seperti metode ANFIS, GA-ANFIS LS – SVM dan pendekatan ANN	Jumlah dataset yang digunakan terlalu sedikit Saran menambah jumlah dataset dan menambahkan jumlah skenario pengujian dan pelatihan agar bisa didapatkan skenario terbaik yang mampu menghasilkan akurasi kesalahan minimum serta dapat menguji pola yang didapat	penelitian berbeda dalam penggunaan dataset dan metode optimasi dan proses optimasi yang diterapkan pada metode ANFIS
5	<i>Improved adaptive neuro-fuzzy inference system based on modified glowworm swarm and differential evolution optimization algorithm for medical diagnosis</i>	Kishore Balasubramanian, N. P. Ananthamoorthy, <i>Neural Computing and Applications</i> – Springer, 2021	Peningkatan efisiensi metode ANFIS dengan menggunakan algoritma optimasi <i>glowworm swarm</i> yang dimodifikasi (M-GSO) serta algoritma <i>differential evolution</i> (DE) untuk memprediksi	Hasil penelitian dengan menerapkan metode ANFIS yang dioptimasi dengan M-GSO dan DE mendapatkan nilai RMSE sebesar 0.1104 dan hasil tersebut lebih rendah dibandingkan dengan	Skenario pelatihan dan pengujian hanya terdapat 1 skenario. Saran menambahkan jumlah skenario pengujian dan pelatihan agar bisa didapatkan skenario	Penulis menggunakan penelitian sebagai dasar terhadap optimasi ANFIS, penelitian berbeda dalam penggunaan dataset dan algoritma optimasi

Tabel 2.1. Matriks literatur review dan posisi penelitian
 Optimasi Struktur Metode Adaptive Neuro Fuzzy Inference System Menggunakan Algoritma Genetika Dalam Meningkatkan Akurasi Prediksi Data Time Series (Lanjutan)

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
			gangguan medis dan membantu dalam skrining dan pengelolaan penyakit	metode ANFIS (0.2037), PSO-ANFIS (0.1341), GA-ANFIS (0,1469), LOA-ANFIS (0.1685), dan GSO-ANFIS (0.1267) untuk nilai RMSE	terbaik yang mampu menghasilkan akurasi kesalahan minimum	
6	<i>Performance evaluation of GIS-based new ensemble data mining techniques of adaptive neuro-fuzzy inference system (ANFIS) with genetic algorithm (GA), differential evolution (DE), and particle swarm optimization (PSO) for landslide spatial modelling</i>	Wei Chena, Mahdi Panahb, Hamid Reza Pourghasemi. <i>Catena</i> – ELSEVIER, 2017	Penerapan 3 teknik data mining yang dikombinasikan dengan model ANFIS. Model yang diterapkan meliputi ANFIS-GA, ANFIS-DE, dan ANFIS-PSO untuk permodelan spasial longsor dan zonasinya	Hasil penelitian yang dilakukan peneliti dalam menerapkan 3 teknik data mining yang dikombinasikan dengan model ANFIS menghasilkan nilai AUC tertinggi untuk model ANFIS-DE (0.844), diikuti oleh ANFIS-GA (0.821), dan ANFIS-PSO (0.780).	Saran menambahkan jumlah skenario pengujian dan pelatihan agar bisa didapatkan skenario terbaik yang mampu menghasilkan akurasi kesalahan minimum	Penulis menggunakan penelitian sebagai dasar terhadap optimasi ANFIS dengan Algoritma Genetika, penelitian berbeda dalam penggunaan dataset dan model optimasi yang diterapkan

Tabel 2.1. Matriks literatur review dan posisi penelitian
Optimasi Struktur Metode Adaptive Neuro Fuzzy Inference System Menggunakan Algoritma Genetika Dalam Meningkatkan Akurasi Prediksi Data Time Series (Lanjutan)

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
7	<i>Time-series prediction of wind speed using machine learning algorithms: A case study Osorio wind farm, Brazil</i>	A. Khosravi, L. Machado, R.O. Nunes, <i>Applied Energy</i> – ELSEVIER, 2018	Penerapan metode <i>multilayer feed-forward neural network</i> (MLFFNN), <i>support vector regression</i> (SVR), <i>fuzzy inference system</i> (FIS), <i>group method of data handling</i> (GMDH), dan <i>adaptif neuro-fuzzy inference system</i> (ANFIS) dengan model ANFIS dilakukan optimasi dengan menerapkan algoritma optimasi <i>particle swarm optimization</i> (ANFIS-PSO) dan algoritma genetika (ANFIS-GA) untuk memprediksi kecepatan angin	Hasil penelitian yang dilakukan peneliti dengan metode – metode <i>machine learning algorithms</i> dapat disimpulkan bahwa secara umum, performa maksimal diperoleh untuk teknik MLFFNN, SVR, GMDH, ANFIS-PSO dan ANFIS-GA. Untuk empat kumpulan data, metode ANFIS menunjukkan koefisien korelasi kira-kira di atas 0,9 untuk data uji	Tidak dijelaskan dengan detail tentang data yang diolah pada penelitian di masing-masing algoritma	Penulis menggunakan penelitian sebagai dasar terhadap optimasi ANFIS dengan Algoritma Genetika, penelitian berbeda dalam penggunaan dataset dan hanya menggunakan satu algoritma untuk optimasi metode ANFIS
8	<i>Optimization of biodiesel production from</i>	Benjamin Ogaga Ighose, Ibrahim A. Adeleke, Mueji	Penerapan ANFIS digabungkan dengan algoritma genetika dan	Dari hasil penelitian disimpulkan penerapan ANFIS	Jumlah dataset yang digunakan tidak djeskan secara	Penulis menggunakan penelitian sebagai dasar terhadap optimasi ANFIS

Tabel 2.1. Matriks literatur review dan posisi penelitian
Optimasi Struktur Metode Adaptive Neuro Fuzzy Inference System Menggunakan Algoritma Genetika Dalam Meningkatkan Akurasi Prediksi Data Time Series (Lanjutan)

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
	<i>Thevetia peruviana seed oil by adaptive neuro-fuzzy inference system coupled with genetic algorithm and response surface methodology</i>	Damos, Hamidah, Adeola Jumaid, Kelechi Ernest Okpalaek, Eriola Betiku, <i>Energy Conversion and Management</i> – ELSEVIER, 2017	<i>response surface methodology (RSM)</i> sebagai alat prediksi untuk produksi metil ester asam lemak (FAME) dari oleander kuning (<i>Thevetia peruviana</i>) minyak biji	digabungkan dengan GA dan RSM mendapatkan hasil yang cukup memuaskan dengan GA mampu memprediksi hasil TPME yang lebih tinggi (99,8 wt.%) dan RSM (98,8 wt.%) lebih rendah dibandingkan dengan GA	mendetail serta tidak dijelaskan skenario pembagian pelatihan dan pengujian	dengan Algoritma Genetika, penelitian berbeda dalam penggunaan dataset dan model optimasi yang diterapkan serta penulis menerapkan ke dalam peramalan
9	<i>Adaptive neuro-fuzzy inference system combined with genetic algorithm to improve power extraction capability in fuel cell applications</i>	Murat Mustafa Savrın, Mustafa İnci, <i>Journal of Cleaner Production</i> – ELSEVIER, 2021	Penerapan metode MPPT berbasis ANFIS yang ditingkatkan untuk memaksimalkan ekstraksi daya dari sistem. Metode ANFIS kemudian dioptimasi dengan algoritma agenetika (ANFIS-GA)	Hasil penelitian yang dilakukan peneliti dapat disimpulkan ANFIS-GA menghasilkan nilai efisiensi antara 98,58% dan 98,99%, ANFIS Konvensional dan PI antara 97,45%-98,62% dan 76,57%	Saran untuk lebih spesifik dalam menunjukkan data dan pengolahannya	Penulis menggunakan penelitian sebagai dasar terhadap optimasi ANFIS dengan Algoritma Genetika, penelitian berbeda dalam penggunaan dataset dan model optimasi yang diterapkan serta penulis menerapkan ke dalam peramalan

Tabel 2.1. Matriks literatur review dan posisi penelitian
 Optimasi Struktur Metode Adaptive Neuro Fuzzy Inference System Menggunakan Algoritma Genetika Dalam Meningkatkan Akurasi Prediksi Data Time Series (Lanjutan)

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
10	<i>A Nonlinear Fuzzy Neural Network Modeling Approach Using Improved Genetic Algorithm</i>	Ridong Zhang and Jili Tao, <i>iecc transactions on industrial electronics</i> , 2017	Penerapan algoritma genetika untuk mengoptimalkan struktur dan parameter jaringan saraf fuzzy	Dari hasil penelitian dapat disimpulkan bahwa metode yang diusulkan lebih unggul dalam penyederhanaan struktur, presisi pemodelan dan kemampuan generalisasi serta mampu menghasilkan RMSE pelatihan sebesar 0,3187 dan RMSE pengujian sebesar 0,6083. pada aturan fuzzy cukup kecil yang terdistribusi antara 1,4 dengan nilai rata-rata RMSE dari 0,4425 hingga 0,7148.	Dataset yang digunakan tidak dijelaskan dengan rinci	Penulis menggunakan penelitian sebagai dasar terhadap penerapan Algoritma Genetika, penelitian berbeda dalam penggunaan dataset dan metode yang digunakan serta penulis menerapkan ke dalam peramalan
11	Prediksi Besaran Curah Hujan Menggunakan	Diera Desmonda, Tursina, M Azhar Irwansyah, Jurnal	Pemanfaatan teknik FST untuk membangun aplikasi yang mampu	Hasil penelitian yang dilakukan oleh peneliti disimpulkan bahwa dengan model	Saran menambahkan variabel input untuk pengolahan data	Penulis menggunakan penelitian sebagai dasar terhadap penerapan <i>time series</i> , penelitian berbeda

Tabel 2.1. Matriks literatur review dan posisi penelitian
Optimasi Struktur Metode Adaptive Neuro Fuzzy Inference System Menggunakan Algoritma Genetika Dalam Meningkatkan Akurasi Prediksi Data Time Series (Lanjutan)

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
	Metode Fuzzy Time Series	Sistem dan Teknologi Informasi, 2018	berinteraksi serta meramalkan informasi curah hujan yang dimungkinkan terjadi di masa depan	yang digunakan dalam pengujian menghasilkan nilai MAPE pada pengujian data sebesar 0.1509 % yang artinya mendapat akurasi ketepatan 99.8491 %. Sedangkan untuk data pengujian didapat nilai MAPE terbaik dengan 0.6460 % yang didapat dengan menggunakan data pada tahun 2014 hingga 2017 dan memiliki nilai akurasi ketepatan sebesar 99.354%.		dalam penggunaan algoritma dan metode optimasi
12	Peramalan Curah Hujan Dengan Pendekatan Adaptive Neuro	Chrisman Bonor Sinaga, Havaluddin, dkk, Sains, Aplikasi, Komputasi dan	Meramalkan tingkat curah hujan yang terjadi pada kota samarinda dengan memanfaatkan	Hasil dari penelitian yang dilakukan disimpulkan bahwa model yang digunakan mampu	Saran menambahkan metode perhitungan akurasi	Penulis menggunakan penelitian sebagai dasar terhadap penerapan ANFIS, penelitian berbeda di dataset dan dalam

Tabel 2.1. Matriks literatur review dan posisi penelitian
Optimasi Struktur Metode Adaptive Neuro Fuzzy Inference System Menggunakan Algoritma Genetika Dalam Meningkatkan Akurasi Prediksi Data Time Series (Lanjutan)

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
	Fuzzy Inference System	Teknologi Informasi, 2019	pengimplementasian pendekatan ANFIS	mendapat nilai <i>error</i> yang diukur dengan menggunakan RMSE dan MSE masing masing sebesar 0,734 dan 0,0633 dengan menggunakan rasio 4:1		menambahkan metode optimasi
13	Prediksi curah hujan menggunakan metode anfis (studi kasus: kabupaten hulu sungai utara)	Muhammad Fajrian Noor, Ema Utami, Eko Pramono, Jurnal INFORMA Politeknik Indonusa Surakarta, 2019	Mengimplementasikan model untuk meramalkan curah hujan dengan menerapkan metode ANFIS. penerapan tersebut dengan tujuan untuk mencari tahu faktor yang saya dapat menyebabkan banjir.	Hasil dari penelitian disimpulkan bahwa ANFIS mampu menghasilkan akurasi prediksi yang baik dengan pola rata-rata curah hujan dari 3 kecamatan menunjukkan variasi yang berbeda beda dengan mampu memprediksi bahwa pada tiap-tiap tahun terdapat 2 bulan yang mempunyai tingkat curah hujan diatas batas normal.	Parameter input yang digunakan tidak dijelaskan dengan rinci Saran menambahkan jumlah skenario pengujian dan pelatihan agar bisa didapatkan skenario terbaik yang mampu menghasilkan akurasi kesalahan minimum	Penulis menggunakan penelitian sebagai dasar terhadap penerapan ANFIS, penelitian berbeda di dataset dan dalam menambahkan metode optimasi

Tabel 2.1. Matriks literatur review dan posisi penelitian
Optimasi Struktur Metode Adaptive Neuro Fuzzy Inference System Menggunakan Algoritma Genetika Dalam Meningkatkan Akurasi Prediksi Data Time Series (Lanjutan)

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
14	Prediksi Harga Cabai dengan Pemodelan Time Series ARIMA	F.N Hadiansyah, Indonesian Journal of Computing, 2017	Membangun sistem yang mampu meramalkan atau memprediksi harga cabai. Sistem tersebut juga harus mampu meramalkan kemungkinan yang akan terjadi di masa depan dengan menggunakan penerapan model ARIMA	Hasil dari penelitian dapat disimpulkan bahwa model ARIMA (1 2 1), IMA (2 1) mampu menghasilkan akurasi yang cukup baik dengan jika diukur dengan menggunakan data latih. Namun untuk pengujian maka model AR (1) yang menjadi model dengan nilai terbaik dengan nilai <i>error</i>) mendapat nilai RMSE sebesar 305.348, MAPE sebesar 0.651%, dan R-square sebesar 0.635	Saran untuk lebih spesifik dalam menunjukkan data dan pengolahannya	Penulis menggunakan penelitian sebagai dasar terhadap penerapan <i>time series</i> , penelitian berbeda dalam penggunaan metode dan adanya proses optimasi
15	Algoritma Genetika Untuk Optimasi Struktur Anfis	Ariesta Damayanti, Rudy Cahyadi, Seminar Riset	Penggantian algoritma pelatihan pada struktur ANFIS untuk mendapatkan nilai	Berdasarkan hasil penelitian yang dilakukan, dapat disimpulkan	Saran untuk lebih spesifik dalam menunjukkan data dan pengolahannya dan	Penulis menggunakan penelitian sebagai dasar terhadap optimasi ANFIS dengan Algoritma

Tabel 2.1. Matriks literatur review dan posisi penelitian
 Optimasi Struktur Metode Adaptive Neuro Fuzzy Inference System Menggunakan Algoritma Genetika Dalam Meningkatkan Akurasi Prediksi Data Time Series (Lanjutan)

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
		Teknologi Informasi (SRITI), 2016	parameter yang lebih maksimal dengan menerapkan algoritma genetika. Penerapan tersebut dimaksudkan untuk mengatasi permasalahan yang terjadi akibat jumlah aturan samar terlalu besar atau "curse of dimensionality"	penerapan algoritma genetika untuk optimasi struktur anfis mendapatkan nilai RMSE terkecil sebesar 0,002899	menambahkan jumlah skenario pengujian dan pelatihan agar bisa didapatkan skenario terbaik yang mampu menghasilkan akurasi kesalahan minimum	Genetika, penelitian berbeda dalam penggunaan dataset dan model optimasi yang diterapkan
16	Prediksi Kecepatan Angin Menggunakan Adaptive Neuro Fuzzy (ANFIS) dan Radial Basis Function Neural Network (RBFNN)	Nerfita Nikentari, Martalefi Bettiza, Helen Sasty Pratiwi, Jurnal Edukasi dan Penelitian Informatika (JEPIN), 2018	Mengimplementasikan model untuk memprediksi kecepatan angin dengan menggunakan penerapan metode ANFIS serta RBFNN	Dari hasil yang didapat, disimpulkan bahwa penelitian tersebut didapatkan hasil error terendah pada metode RBFNN sebesar 0,1766 dan untuk RMSE pada ANFIS sebesar 1,1456.	Saran menambahkan jumlah skenario pengujian dan pelatihan agar bisa didapatkan skenario terbaik yang mampu menghasilkan akurasi kesalahan minimum	Penulis menggunakan penelitian sebagai dasar terhadap penerapan ANFIS, penelitian berbeda di dataset dan dalam menambahkan metode optimasi
17	Optimasi Fuzzy Artificial Neural Network dengan Algoritma Genetika untuk	Anwar Rifa'i, Jurnal Teknik Informatika dan Sistem Informasi, 2020	Penerapan algoritma genetika untuk optimasi metode FRBFNN dalam peramalan harga CPO	Berdasarkan hasil penelitian yang dilakukan, metode FRBFNN mendapatkan hasil	Saran untuk lebih spesifik dalam menunjukkan data dan pengolahannya dan menambahkan jumlah	Penulis menggunakan penelitian sebagai dasar terhadap penerapan Algoritma Genetika, penelitian berbeda dalam

Tabel 2.1. Matriks literatur review dan posisi penelitian
 Optimasi Struktur Metode Adaptive Neuro Fuzzy Inference System Menggunakan Algoritma Genetika Dalam Meningkatkan Akurasi Prediksi Data Time Series (Lanjutan)

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
	Prediksi Harga <i>Crude Palm Oil</i>			prediksi yang cukup baik dengan tingkat error sebesar 11,7% untuk data latih dan 9,4% untuk data uji. Sedangkan pada metode FRBFNN yang dioptimasi dengan GA mengalami peningkatan dengan menghasilkan nilai error sebesar 10,2% untuk data latih dan 8,3% untuk data uji	skenario pengujian dan pelatihan agar bisa didapatkan skenario terbaik yang mampu menghasilkan akurasi kesalahan minimum	penggunaan dataset dan metode yang digunakan
18	Algoritma Genetika Untuk Optimasi <i>Fuzzy Time Series</i> Dalam Memprediksi Debit Air (Studi Kasus: PDAM Indramayu)	Mohamad Alfi Fauzan, Budi Darma Setiawan, Indriati, Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, 2019	Membuat peramalan untuk meramalkan tingkat distribusi debit air dengan penerapan algoritma genetika yang bertujuan memperbaiki <i>fuzzy time series</i>	Berdasarkan hasil penelitian dapat diambil kesimpulan bahwa metode <i>Average Forecasting Error Rate (AFER)</i> mampu mendapat nilai <i>error</i> sebesar 15,33%, dari nilai tersebut peneliti memasukkan hasil	saran untuk menambah banyak data agar dapat di ketahui berapa akurasi jika menggunakan data dalam jumlah banyak serta menambah metode <i>fuzzy time series</i> lainnya agar dapat dibandingkan	Penulis menggunakan penelitian sebagai dasar terhadap penerapan Algoritma Genetika, penelitian berbeda dalam penggunaan dataset dan metode yang digunakan

Tabel 2.1. Matriks literatur review dan posisi penelitian
Optimasi Struktur Metode Adaptive Neuro Fuzzy Inference System Menggunakan Algoritma Genetika Dalam Meningkatkan Akurasi Prediksi Data Time Series (Lanjutan)

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
				tersebut kedalam klasifikasi akurasi yang baik	dan didapat model yang terbaik	
19	ANFIS dengan membership function untuk prediksi curah hujan pada data rentet waktu multivariate	Lilis Anggraini, Technologia, 2018	Pemanfaatan metode ANFIS yang bertujuan untuk meramalkan atau memprediksi tingkat curah hujan dengan memanfaatkan <i>membership function trapmf, trimf, gaussmf</i> serta <i>bellmf</i> dan menggunakan data rentet waktu	Hasil perhitungan prediksi yang dilakukan, mendekati data yang sebenarnya dengan nilai RMSE 0,73436. Nilai tersebut didapat dari fungsi keanggotaan <i>bellmf</i> yang mempunyai akurasi nilai tertinggi. Dengan menggunakan data pendukung seperti kelembapan 58,4% dan suhu 47,5 celcius maka curah hujan sebesar 32,2 mm dengan hasil akurasi / RMSE 0,73436	Saran untuk lebih spesifik dalam menunjukkan data dan pengolahannya dan menambahkan jumlah skenario pengujian dan pelatihan agar bisa didapatkan skenario terbaik yang mampu menghasilkan skenario terbaik	Penulis menggunakan penelitian sebagai dasar terhadap penerapan ANFIS, penelitian berbeda di dataset dan dalam menambahkan metode optimasi
20	Prediksi curah hujan menggunakan	Muhammad Isradi Azhar, Wayan Firdaus Mahmudy, Jurnal	Pemanfaatam metode ANFIS yang bertujuan untuk membuat	Hasil dari penelitian dengan menggunakan 150	Saran jumlah dataset diperbanyak dan ditambahkan skenario	Penulis menggunakan penelitian sebagai dasar terhadap penerapan ANFIS,

Tabel 2.1. Matriks literatur review dan posisi penelitian
 Optimasi Struktur Metode Adaptive Neuro Fuzzy Inference System Menggunakan Algoritma Genetika Dalam Meningkatkan Akurasi Prediksi Data Time Series (Lanjutan)

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
	metode <i>adaptive neuro fuzzy inference system</i> (ANFIS)	Pengembangan Teknologi Informasi dan Ilmu Komputer, 2018	peramalan curuh hujan dengan menggunakan 4 parameter masukan yang berbeda dimana parameter tersebut dapat saling mempengaruhi	data latih yang merupakan pengujian dengan nilai RMSE terbaik kemudian dengan jumlah iterasi sebanyak 400 dapat menghasilkan nilai <i>error</i> yang diukur menggunakan RMSE sebesar 1,88 menunjukkan untuk model yang digunakan mampu menghasilkan data hasil prediksi mendekati data asli	pengujian yang berbeda.	penelitian berbeda di dataset dan dalam menambahkan metode optimasi

2.3. Landasan Teori

Terdapat beberapa landasan teori yang dibutuhkan dalam penelitian ini, mulai dari landasan teori untuk data berkala (time series), forecasting, ANFIS, serta Algoritma Genetika.

2.3.1. Data Runtut (*Time Series*)

Data runtut (*Time Series*) berlangsung dan bersumber pada waktu- waktu tertentu secara berentetan. *Time Series* merupakan serangkaian informasi yang berbentuk nilai dari hasil pengamatan yang diukur dalam kurun waktu tertentu, bersumber pada waktu dengan interval yang sama (Piegel L.J dkk., 2007) Waktu yang dipergunakan bisa berbentuk setiap hari, mingguan, bulanan, tahunan, serta sebagainya (Cryer & Chan, 2008). Pola gerakan informasi ataupun nilai- nilai variabel bisa diiringi ataupun diketahui dengan terdapatnya informasi berkala, sehingga mampu dijadikan sebagai bahan pertimbangan untuk membantu dalam pembuatan keputusan, peramalan pada masa yang akan datang, serta perencanaan aktivitas pada waktu yang akan datang (Chia-Hoang Lee dkk., 1997).

Implementasi *time series* sering kali dipergunakan dalam bidang peramalan ataupun forecasting dimana model *time series* mengamati sikap observasi yang diambil secara berurutan (Wei, 2006). Model data *time series* telah banyak dimanfaatkan guna memprediksi atau meramalkan sesuatu yang terkait dengan waktu, seperti (Anggraini, 2018) yang membahas tentang prediksi curah hujan serta dalam penerapannya digunakan data *time series*. Analisis *time series* memungkinkan kita untuk mendapat informasi mengenai suatu perkembangan atau peristiwa serta pengaruh terhadap peristiwa lain (Halimi & Kusuma, 2018).

Pergerakan informasi ataupun nilai variabel bisa dikenal dengan terdapatnya informasi time series, sehingga bisa dijadikan sebagai pertimbangan untuk membuat keputusan dan juga perencanaan pada waktu yang akan datang karna pola pada informasi time series dipercaya di masa yang akan datang akan kembali terulang (Gen & Cheng, 1997).

2.3.2. Prediksi (*Forecasting*)

Prediksi (*forecasting*) merupakan suatu kegiatan yang bertujuan salah satunya untuk memprekirakan atau meramalkan data. Peramalan dilakukan secara sistematis mengenai suatu hal peristiwa atau kejadian yang bisa jadi akan berlangsung pada masa mendatang. Data yang digunakan untuk melakukan prediksi bersumber pada data masa lampau yang saat ini masih berlangsung dengan tujuan untuk memperkecil kesalahan. Data masa lampau berguna juga untuk menghitung selisih data hasil peramalan dengan data aktual yang informasi tersebut sangat diperlukan agar mampu menghasilkan nilai kesalahan yang kecil. Sebutan prediksi sama dengan ramalan ataupun perkiraan (*forecasting*) (Stevenson & Chee Chuong, 2014) Penerapan *forecasting* dapat dengan cara mengaitkan proses penggalian informasi masa lampau kemudian meramalkan masa yang akan datang. Proses tersebut dilakukan dalam perhitungan matematis (Cryer & Chan, 2008). Prediksi tidak harus memberikan informasi secara detail atau pasti tentang kejadian yang akan terjadi di masa yang akan datang, melainkan prediksi berusaha untuk mencari tau informasi sedekat mungkin mengenai hal yang dapat terjadi di kedepannya. Prediksi menunjukkan gambaran yang akan terjadi pada masa mendatang pada

suatu keadaan tertentu serta merupakan bahan dasar untuk perencanaan dan pengambilan keputusan

2.3.3. Teknik Prediksi

Prediksi dikategorikan kedalam 2 kategori, antara lain:

1. Prediksi Kuantitatif

Prediksi kuantitatif merupakan suatu teknik peramalan berdasarkan data kuantitatif yang telah berlangsung sebelumnya di masa lampau. Teknik ini sangat bergantung pada jenis metode peramalan yang nantinya akan diterapkan. Peramalan ini dapat diterapkan sebagai teknik peramalan apabila terdapat kondisi berikut (S. Makridakis, 1995):

1. Tersedianya data yang terkait dengan masa lalu.
2. Informasi yang diperoleh dapat berupa bilangan numerik

Pola data sebelumnya yang terjadi pada masa lampau diasumsikan dapat kembali terjadi di masa mendatang.

2. Prediksi Kualitatif

Prediksi kualitatif merupakan suatu metode berdasarkan data kualitatif yang terjadi pada masa lalu. Prediksi dengan teknik ini bergantung terhadap peneliti. Perihal tersebut sangat mempengaruhi hampir disetiap aspek dikarenakan hasil dari peramalan ini akan ditentukan bersumber pada pemikiran yang memiliki naluri, pendapatan pengetahuan dan pengalaman dalam proses penyusunan peramalan.

2.3.4. Tahapan Prediksi

Prediksi yang baik merupakan prediksi yang dilakukan dengan mengaplikasikan langkah penyusunan yang baik. Ada 9 langkah yang wajib

dicermati agar mampu mendapat prediksi yang efektifitas dan efisien (V. Gaspresz, 2005).

1. Menentukan tujuan
2. Memilih data yang akan dilakukan peramalan
3. Menetapkan barasan interval dari peramalan (pendek, menengah, ataupun, panjang)
4. Memilah model yang berkaitan dengan *forecasting* atau prediksi
5. Mendapatkan data yang akan digunakan untuk bahan dasar melakukan prediksi
6. Menetapkan tingkat keakuratan suatu metode
7. Membuat prediksi
8. Melakukan implementasi dari hasil yang didapat dari proses peramalan
9. Mengevaluasi hasil

2.3.5. Tujuan Prediksi

Tujuan prediksi dibuat agar senantiasa dapat:

1. Untuk memperoleh informasi dimasa yang akan datang.
2. Mendapatkan prediksi yang mampu mendapat kesalahan minimum yang dapat diukur dengan pengukuran akurasi error menggunakan *Root Mean Squared Error (RMSE)*.

2.3.6. ANFIS

Adaptive Neuro Fuzzy Inference System atau dapat disebut dengan ANFIS adalah suatu permodelan dengan struktur yang menerapkan base model sugeno, *fuzzy rule base*, dan jaringan syaraf tiruan (JST). Dalam ANFIS, terdapat kesamaan dengan JST dimana terdapat pembeda atau batasan. Pada ANFIS memperbolehkan

ketentuan atau aturan untuk menyesuaikan diri (s. kusumadewi, 2010).
Supaya jaringan fungsi basis sebanding dengan *fuzzy* Sugeno orde 1, dibutuhkan batas batas sebagai berikut (s. kusumadewi, 2010):

1. Setiap aturan wajib mempunyai metode penghimpunan guna menciptakan seluruh keluaran.
2. fungsi aktivasi sama dengan jumlah aturan
3. Apabila terdapat beberapa masukan pada basis aturannya, maka tiap – tiap fungsi aktivasi harus sama dengan fungsi keanggotannya tiap – tiap masukannya.
4. Fungsi aktivasi serta aturan *fuzzy* harus mempunyai fungsi yang sama untuk tiap – tiap *neuron – neuron* dan aturan – aturan yang ada di sisi keluarannya.

2.3.7. Arsitektur ANFIS

Dimisalkan terdapat dua nilai masukan x_1, x_2 serta satu nilai keluaran y . terdapat 2 aturan pada basis aturan model Sugeno (Damayanti & Cahyadi, 2016):

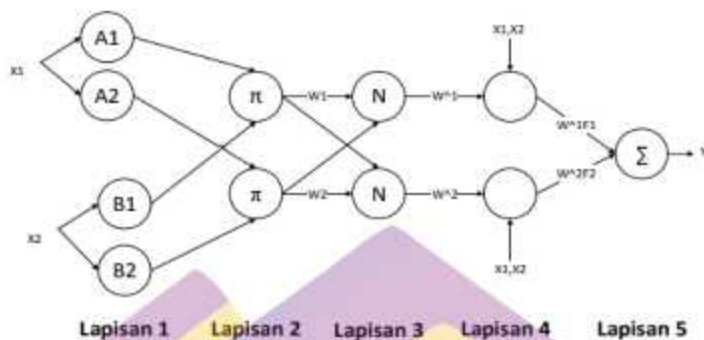
$$\text{if } x_1 \text{ is } A_1 \text{ and } x_2 \text{ is } B_1 \text{ Then } y_1 = c_{11}x_1 + c_{12}x_2 + c_{10}$$

$$\text{if } x_1 \text{ is } A_2 \text{ and } x_2 \text{ is } B_2 \text{ Then } y_2 = c_{21}x_1 + c_{22}x_2 + c_{20}$$

Jika α predikat untuk aturan ke-2 aturan adalah w_1 dan w_2 , maka dapat dihitung rata-rata terbobot:

$$y^i = \frac{w_1 y_1 + w_2 y_2}{w_1 w_2} = \bar{w}_1 y_1 + \bar{w}_2 y_2$$

(2.1)



Gambar 2.1. Struktur ANFIS

Lapisan 1 (fuzzyfikasi): pada lapisan 1 masing - masing parameter neuron (i) pada lapisan ini adaptif dengan fungsi aktivasi. Keluaran yang dihasilkan dipresentasikan sebagai derajat keanggotaan. Nilai pada derajat keanggotaan didapat dari pemberian dari *membership function* $\alpha A1(X1)$, $\alpha B1(X2)$, $\alpha A2(X1)$ atau $\alpha B2(X2)$. Pada lapisan ini, *membership function* yang digunakan merupakan *generalized bell (gbell)*. Fungsi umumnya adalah sebagai berikut:

$$gbell(x, a, b, c) = \frac{1}{1 + \left| \frac{x-c}{a} \right|^{2b}}$$

(2.2)

Nilai koefisien b merupakan bilangan positif. Dalam penerapannya koefisien b didapat dari pendefinisian nilai secara manual dengan nilai yang digunakan adalah 1. Untuk nilai koefisien a didedinisikan sebagai lebar antara titik pusat sampai dengan batas nilai keanggotaan. Untuk koefisien c mempresentasikan suatu pusat dari fungsi keanggotaan yang digunakan. Deklarasi koefisien ini terjadi pada saat

instalasi awal pembelajaran kemudian nilai koefisien ini dapat mengalami perubahan seiring dengan pembelajaran yang ada lalu selama berada didalam proses hingga dihasilkan sebuah parameter yang berniali adaptif.

Lapisan 2 (Operasi Logika *Fuzzy*): masing masing sinyal dari lapisan ini merupakan sinyal tetap yang berlabel z_i yang berarti nilai keluaran yang dihasilkan adalah produk dari semua sinyal yang datang. Pada lapis ini akan dikalikan derajat keanggotaan dengan fungsi umumnya sebagai berikut:

$$O_{z,i} = w_i = \mu_{A_i}(x_1) \cdot \mu_{B_i}(x_2), \text{ untuk } i = 1,2 \quad (2.3)$$

Keterangan:

- x_1 – Nilai masukan
- x_2 – Nilai masukan
- w_i – Bobot ke-i
- μ_i – Nilai keanggotaan kelas-i
- $O_{z,i}$ – Nilai keluaran pada lapisan 2i

Masing – masing keluaran menerangkan tentang nilai derajat pengaktifan pada masing masing aturan yang ada. Jika himpunan fuzzy diatas dua, maka nilai fungsi ini bisa dijabarkan kembali terutama pada bagian parameter premis. Jumlah aturan yang dibentuk didasarkan dari banyaknya simpul yang. Sehingga jika semakin banyak simpul, maka akan semakin banyak aturan yang terbentuk. Operator *t-norm* dalam lapisan ini berfungsi untuk melakukan fungsi perkalian.

Lapisan 3 (normalisasi derajat pengaktifan): masing masing simpul yang terbentuk pada lapisan ini merupakan simpul tetap. Simpul tetap ini dilambangkan dengan N . Lapisan ini merupakan lapisan penampung hasil dari pengolahan data menggunakan nilai rasio data α predikat (w), dari aturan ke-1 terhadap jumlah dari keseluruhan α predikat. Fungsi umumnya sebagai berikut:

$$O_{3,i} = \bar{w}_i = \frac{w_i}{w_1 + w_2}, \text{ dengan } i = 1,2 \quad (2.4)$$

Keterangan:

- w_i – Bobot ke- i
- \bar{w}_i – Derajat pengaktifan ternormalisasi ke- i
- $O_{3,i}$ – Nilai output dari lapisan 3 ke- i

Lapisan 4 (*Defuzzification*): Tiap simpul yang dihasilkan pada lapisan ini merupakan simpul yang bersifat adaptif. Fungsi *node* dinotasikan sebagai berikut:

$$O_{4,i} = \bar{w}_i f_i = \bar{w}_i (c_{i1} x_1 + c_{i2} x_2 + c_{i0}), \text{ dengan } i = 1,2 \quad (2.5)$$

Dengan

- \bar{w}_i – derajat pengaktifan ternormalisasi ke- i
- f_i – sinyal kontrol pada data ke- i
- $O_{3,i}$ – output lapisan 4 ke- i
- (c_{i1}, c_{i2}, c_{i0}) – parameter pada lapisan *neuron*.

Lapisan 5 (Perhitungan nilai keluaran): simpul dari lapisan ini bernilai tunggal dan merupakan simpul tetap dengan Σ yang mampu memproses semua sinyal yang masuk menjadi sebuah keluaran. Fungsi umumnya sebagai berikut:

$$O_{5,i} = \sum_i \bar{w}_i f_i = \bar{w}_1 f_1 + \bar{w}_2 f_2 \quad (2.6)$$

Keterangan:

- \bar{w}_i – nilai derajat pengaktifan ternormalisasi ke-i
- f_i – sinyal kontrol ke-i
- $O_{5,i}$ – keluaran pada lapisan 5 ke-i

Parameter metode ANFIS bisa dipisahkan menjadi 2 bagian, yakni premis dan juga konsekuen. Kedua parameter tersebut mampu beradaptasi dalam proses pembelajaran yang terdapat pada pemrosesan ANFIS, baik pembelajaran maju dan juga mundur yang di lakukan oleh algoritma pembelajaran hybrid (T. S. Widodo, 2005).

2.3.8. Algoritma Pembelajaran Hybrid

Pembelajaran *hybird* merupakan suatu algoritma yang dipergunakan untuk mengatur parameter-parameter yang terdapat dalam metode ANFIS dalam pembelajaran yang berlangsung secara maju maupun secara mundur (S. Kusumadewi and H. Purnomo, 2010).

Tabel 2.2. Proses Pembelajaran Hybrid

	Arah Maju	Arah Mundur
Parameter Premis	Tetap	<i>steepest descent</i>
Parameter Konsekuen	RLSE	Tetap
Sinyal	Keluaran Simpul	Laju Kesalahan

2.3.9. Pembelajaran Arah Maju (LSE *Recursive*)

Berdasarkan struktur ANFIS yang di presentasikan pada Gambar 1, diketahui bahwa apabila pada parameter premis didapatkan nilai yang tetap maka untuk semua keluaran dapat dinyatakan sebagai kombinasi linear (T. S. Widodo, 2005). Pada pembelajaran ini akan digunakan algoritma *LSE Recursive*.

Persamaannya seperti berikut:

$$\begin{aligned}
 y &= \frac{w_1}{w_1 + w_2} y_1 + \frac{w_2}{w_1 + w_2} y_2 \\
 &= \hat{w}_1(p_1x_1 + q_1x_2 + r_1) + \hat{w}_2(p_2x_1 + q_2x_2 + r_2) \\
 &= (\hat{w}_1x_1)p_1 + (\hat{w}_1x_2)q_1 + \hat{w}_1r_1 + (\hat{w}_2x_1)p_2 + (\hat{w}_2x_2)q_2 + \hat{w}_2r_2
 \end{aligned}
 \tag{2.7}$$

Untuk p pasangan data pelatihan X_1, X_2 serta keluaran y , maka diperbolehkan. Persamaannya sebagai berikut:

$$\begin{aligned}
 y_1 &= (\hat{w}_1x_1)p_1 + (\hat{w}_1x_2)q_1 + \hat{w}_1r_1 + (\hat{w}_2x_1)p_2 + (\hat{w}_2x_2)q_2 + \hat{w}_2r_2 \\
 y_2 &= (\hat{w}_1x_1)p_1 + (\hat{w}_1x_2)q_1 + \hat{w}_1r_1 + (\hat{w}_2x_1)p_2 + (\hat{w}_2x_2)q_2 + \hat{w}_2r_2 \\
 &\dots = \dots + \dots + \dots + \dots + \dots \\
 y_n &= (\hat{w}_1x_1)p_1 + (\hat{w}_1x_2)q_1 + \hat{w}_1r_1 + (\hat{w}_2x_1)p_2 + (\hat{w}_2x_2)q_2 + \hat{w}_2r_2
 \end{aligned}
 \tag{2.8}$$

Jika dimisalkan

$$A = [px6] = \begin{bmatrix} \hat{w}_1 x_1 & \hat{w}_1 x_2 & \hat{w}_1 & \hat{w}_2 x_1 & \hat{w}_2 x_2 & \hat{w}_2 \\ \hat{w}_1 x_1 & \hat{w}_1 x_2 & \hat{w}_1 & \hat{w}_2 x_1 & \hat{w}_2 x_2 & \hat{w}_2 \\ \hat{w}_1 x_1 & \hat{w}_1 x_2 & \hat{w}_1 & \hat{w}_2 x_1 & \hat{w}_2 x_2 & \hat{w}_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \hat{w}_1 x_1 & \hat{w}_1 x_2 & \hat{w}_1 & \hat{w}_2 x_1 & \hat{w}_2 x_2 & \hat{w}_2 \end{bmatrix} \rightarrow \text{data ke } i$$

(2.9)

$$\theta = [6 \times 1] = \begin{pmatrix} r_1 \\ p_2 \\ q_2 \\ r_2 \end{pmatrix}, y = [p \times 1] = \begin{pmatrix} \dots \\ y_p \end{pmatrix}$$

(2.10)

Maka pada baris ke p dari matriks data $[A: y]$ dirumuskan sebagai $(a_p^T: y_p)$ sebab hampir seluruh perhitungan bersumber pada matriks A dan juga vektor y , maka $(a_p^T: y_p)$ disebut sebagai pasangan data ke p dari himpunan data pelatihan. Apabila nilai $m = n$ (A persegi) dan A yaitu *nonsingular*, maka nilai persamaan parameter q dinotasikan berikut:

$$\theta = A^{-1}y$$

(2.11)

2.3.10. Pembelajaran Arah Mundur (*Error Backpropagation*)

Error Backpropagation (EBP) diaplikasikan untuk mendapatkan nilai kesalahan pada masing - masing lapisan. Nilai kesalahan tersebut didapatkan dari nilai keluaran jaringan adaptif. Selain itu nilai target akan dipropagasi mundur dengan menerapkan algoritma *Gradient Descent* guna melaukan proses yang bertujuan untuk mendapatkan parameter premis yang baru.

Perhitungan nilai *error* pada masing-masing lapisan dapat dirumuskan sebagai berikut (s. kusumadewi, 2010):

1. Lapisan 5

Pada lapisan ini rumus untuk menghitung propagasi *error* diformulasikan sebagai berikut:

$$\varepsilon_{5j} = \frac{\partial E_p}{\partial f} = -2(d_p - O_{5j}) \quad (2.15)$$

2. Lapisan 4

Pada lapisan ini rumus untuk menghitung propagasi *error* diformulasikan sebagai berikut:

$$\varepsilon_{4j} = \left(\frac{\partial E_p}{\partial O_{5i}} \right) \left(\frac{\partial f_{5i}}{\partial O_{4j}} \right) = \varepsilon_{5i} \left(\frac{\partial f_{5i}}{\partial O_{4j}} \right) = \varepsilon_{5i} (1) \quad (2.16)$$

3. Lapisan 3

Pada lapisan ini rumus untuk menghitung propagasi *error* diformulasikan sebagai berikut:

$$\varepsilon_{3j} = \left(\frac{\partial E_p}{\partial O_{5i}} \right) \left(\frac{\partial f_{5i}}{\partial O_{4j}} \right) \left(\frac{\partial f_{4j}}{\partial O_{3j}} \right) = \varepsilon_{4i} \left(\frac{\partial f_{4j}}{\partial O_{3j}} \right) = \varepsilon_{4i} f_i \quad (2.17)$$

4. Lapisan 2

Pada lapisan ini rumus untuk menghitung propagasi *error* diformulasikan sebagai berikut:

$$\varepsilon_{2j} = \frac{w_2}{w_1 + w_2} (\varepsilon_{3,1} - \varepsilon_{3,2})$$

(2.18)

5. Lapisan 1

Pada lapisan ini rumus untuk menghitung propagasi *error* diformulaikan sebagai berikut:

$$\varepsilon_{1,j} = \varepsilon_{2,j} \mu B_1(X_2) \quad (2.19)$$

$$\varepsilon_{1,j} = \varepsilon_{2,j} \mu B_2(X_2) \quad (2.20)$$

$$\varepsilon_{1,j} = \varepsilon_{2,j} \mu A_1(X_1) \quad (2.21)$$

$$\varepsilon_{1,j} = \varepsilon_{2,j} \mu A_2(X_1) \quad (2.22)$$

6. Lapisan 1 dan parameter input *Standar Deviasi*

Rumus persamaan *error* yang terdapat diantara lapisan 1 dan parameter masukan *stdev* diformulaikan sebagai berikut:

$$\varepsilon_{a11} = \varepsilon_{1,j} * \frac{2(x_1 - c_{11})^2}{a_{11}^3 \left[1 + \left(\frac{x_1 - c_{11}}{a_{11}} \right)^2 \right]^2} \quad (2.23)$$

$$\varepsilon_{a12} = \varepsilon_{1,j} * \frac{2(x_1 - c_{12})^2}{a_{12}^3 \left[1 + \left(\frac{x_1 - c_{12}}{a_{12}} \right)^2 \right]^2} \quad (2.24)$$

$$\varepsilon_{a21} = \varepsilon_{1,i} * \frac{2(x_2 - c_{21})^2}{a_{21}^2 \left[1 + \left(\frac{x_2 - c_{21}}{a_{21}} \right)^2 \right]^2}$$
(2.25)

$$\varepsilon_{a22} = \varepsilon_{1,i} * \frac{2(x_2 - c_{22})^2}{a_{22}^2 \left[1 + \left(\frac{x_2 - c_{22}}{a_{22}} \right)^2 \right]^2}$$
(2.26)

7. Lapisan 1 dengan Parameter Masukan *Mean*

Rumus persamaan *error* yang terdapat diantara lapisan 1 dan parameter masukan *mean* diformulasikan sebagai berikut:

$$\varepsilon_{c11} = \varepsilon_{1,i} * \frac{2(x_1 - c_{11})}{a_{11}^2 \left[1 + \left(\frac{x_1 - c_{11}}{a_{11}} \right)^2 \right]^2}$$
(2.27)

$$\varepsilon_{c12} = \varepsilon_{1,i} * \frac{2(x_1 - c_{12})}{a_{12}^2 \left[1 + \left(\frac{x_1 - c_{12}}{a_{12}} \right)^2 \right]^2}$$
(2.28)

$$\varepsilon_{c21} = \varepsilon_{1,i} * \frac{2(x_2 - c_{21})}{a_{21}^2 \left[1 + \left(\frac{x_2 - c_{21}}{a_{21}} \right)^2 \right]^2}$$
(2.29)

$$\varepsilon_{c22} = \varepsilon_{1,i} * \frac{2(x_2 - c_{22})}{a_{22}^2 \left[1 + \left(\frac{x_2 - c_{22}}{a_{22}} \right)^2 \right]^2}$$
(2.30)

8. Perubahan Nilai Parameter Masukan

Menentukan perubahan atau pembaruan nilai pada parameter a_{ij} dan c_{ij}

(Δa_{ij} dan Δc_{ij}) sebagai berikut:

$$\Delta a_{ij} = \eta \varepsilon_{a_{ij}} x_1, \text{ dan} \quad (2.31)$$

$$\Delta c_{ij} = \eta \varepsilon_{c_{ij}} x_1 \quad (2.32)$$

η merupakan *learning rate* dan terdapat pada interval $[0, 1]$ sehingga nilai a_{ij} dan c_{ij} yang baru dapat diformulaikan sebagai berikut:

$$a_{ij} = a_{ij} (\text{lama}) + \Delta a_{ij} \quad (2.33)$$

$$c_{ij} = c_{ij} (\text{lama}) + \Delta c_{ij} \quad (2.34)$$

2.3.11. Pengukuran Hasil Akurasi

Dalam permodelan *time series*, sebagian dari data dapat digunakan untuk meramalkan data yang akan terjadi di kemudian hari. Selisih antara kesalahan dengan data yang sesungguhnya dan yang sedang terjadi merupakan informasi yang sangat berharga. Informasi tersebut dapat dipergunakan untuk meninjau kembali atau melakukan evaluasi kembali mengenai akurasi teknik pengukuran yang digunakan. Suatu kesalahan terkecil bisa didapatkan dengan cara melakukan perbandingan akurasi dari beberapa metode yang diterapkan. Dengan begitu akan akan didapatkan strategi atau metode yang tepat untuk jenis data yang digunakan sehingga dapat dipergunakan sebagai prespektif yang sangat menentukan kegiatan ataupun kebutuhan masa mendatang.

Untuk melakukan pengukuran kesalahan dari hasil peramalan yang telah dilakukan, terdapat beberapa metode pengukuran yang disarankan yakni meliputi (S. Makridakis, 1995):

1. RMSE

Root mean squared error (RMSE) merupakan suatu pengukuran tingkat akurasi dengan mengukur rata-rata nilai kuadrat. Nilai tersebut didapat dari selisih nilai hasil peramalan dengan nilai target. Tingkat keakuratan metode ini berdasarkan nilai yang lebih kecil atau yang lebih rendah. Semakin kecil RMSE yang didapat maka semakin akurat dan semakin valid hasil dari peramalan yang dilakukan. Rumus RMSE dinotasikan sebagai berikut:

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (Y_t - w_i f_t)^2}{n}} \quad (2.35)$$

Keterangan:

- Y – nilai target
- $w_i f_t$ – output peramalan
- n – banyak data

2. MAD

Mean Absolute Deviation (MAD) merupakan suatu pengukuran tingkat kesalahan dengan mengukur ketepatan angka hasil dari peramalan dengan cara mencari rata-rata kesalahan. MAD dapat sangat bermanfaat ketika penganalisis perlu mengukur kesalahan nilai hasil peramalan dengan data yang digunakan data deret asli. Rumus MAD dinotasikan sebagai berikut:

$$MAD = \frac{\sum |A_t - F_t|}{n}$$
(2.36)

Keterangan:

- A_t – Data aktual
- F_t – nilai keluaran peramalan
- n – jumlah data

3. MAPE

Mean absolute percentage error (MAPE) merupakan suatu pengukuran tingkat *mean* dari kesalahan yang didapat dengan menggunakan nilai mutlak (absolut) yang terdapat dalam masing masing periode. Nilai mutlak tersebut kemudian dibagi dengan data asli pada periode yang sama, kemudian akan ditentukan nilai *mean* untuk seluruh kesalahan. Semakin rendah nilai kesalahan yang didapat, maka semakin tinggi nilai akurasi yang di dapat dari peramalan yang di lakukan. Rumus MAPE dinotasikan sebagai berikut:

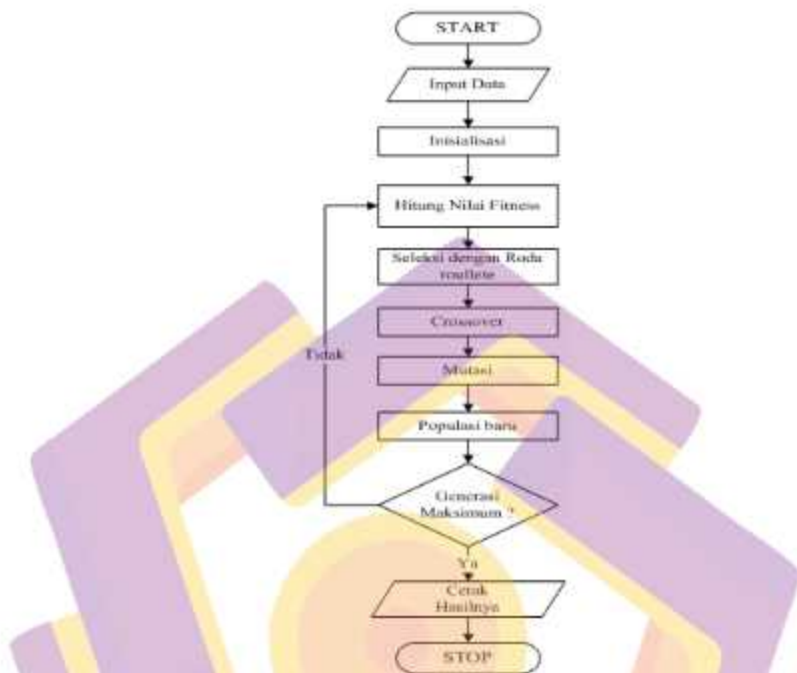
$$MAPE = \frac{\sum_{t=1}^n \left| \frac{e_t}{X_t} \right| 100\%}{n}$$
(2.37)

Keterangan:

- e_t – *error* pada periode t
- n – jumlah data
- X_t – selisih hasil peramalan dengan nilai sebenarnya
- MAPE – rata-rata persen *error*

2.3.12. Algoritma Genetika

Genetic Algorithm atau Algoritma genetika adalah suatu algoritma yang mengaplikasikan prinsip genetika serta seleksi alam. Algoritma Genetika merupakan metode yang bisa digunakan untuk memecahkan kasus peramalan. Algoritma genetika merupakan metode pencarian yang di dalam ilmu komputer digunakan untuk mendapatkan penyelesaian perkiraan guna melakukan optimisasi serta permasalahan pencarian. Algoritma genetik merupakan kelas spesial dari algoritma evolusioner⁶ dengan memanfaatkan metode yang termotivasi oleh biologi evolusioner seperti peninggalan, mutasi, seleksi alam serta rekombinasi (ataupun crossover). Dalam riset yang dilakukan oleh (Rahmi & Mahmudy, 2016) yang menerapkan Algoritma Genetika untuk memprediksi harga saham, disimpulkan bahwa penggunaan algoritma genetika mampu menjadi solusi dalam penentuan koefisien terbaik yang mampu menghasilkan data prediksi yang mendekati data asli.



Gambar 2.2. Alur algoritma genetika

Algoritma genetik memiliki karakteristik yang perlu diperhatikan sehingga bisa terbedakan dari mekanisme pencarian ataupun optimasi yang lain, yakni: (Goldberg, 1989):

1. Algoritma genetika memanfaatkan pengkodean berasal himpunan pemecahan kasus bersumber pada parameter yang telah ditetapkan dan bukan hanya parameter itu sendiri
2. Algoritma genetika pencarian pada suatu penyelesaian dari beberapa individu yang menggambarkan pemecahan kasus bukan hanya dari suatu individu

3. Algoritma genetika informasi fungsi objektif (*fitness*), menjadi solusi untuk mengevaluasi individu yang memiliki solusi optimal, bukan turunan dari suatu fungsi
4. Algoritma genetika mengenakan aturan hukum transisi peluang, bukan aturan-aturan deterministik

Parameter yang digunakan pada algoritma genetika meliputi:

1. *Fitness* yang dimiliki oleh tiap individu guna memilah tingkatan kesesuaian individu tersebut menggunakan kriteria yang ingin dicapai
2. Populasi jumlah individu yang dilibatkan pada tiap generasi.
3. Probabilitas terbentuknya persilangan (*crossover*) pada sesuatu generasi.
4. Probabilitas terbentuknya mutasi pada tiap individu.
5. Jumlah generasi yang hendak dibangun tersebut yang akan memutuskan berapa lama pelaksanaan algoritma genetik.

BAB III METODE PENELITIAN

3.1. Jenis, Sifat, dan Pendekatan Penelitian

3.1.1. Jenis Penelitian

Pada penelitian ini, menggunakan Jenis penelitian berupa penelitian kuantitatif, dimana peneliti melakukan perhitungan matematis untuk menemukan hasil yang diinginkan. Pada penelitian ini bertujuan untuk memprediksi melalui pengumpulan data sedalam dalamnya yang dapat digunakan sebagai bahan analisis.

3.1.2. Sifat Penelitian

Sifat dari penelitian yang akan dilakukan adalah *Deskriptif*. Penelitian *deskriptif* adalah penelitian yang dilakukan untuk mengetahui nilai variable mandiri, baik satu variable atau lebih tanpa membuat perbandingan, atau menghubungkan antara variabel satu dengan variabel yang lain. Penelitian *deskriptif*, artinya penelitian yang menggambarkan objek tertentu dan menjelaskan hal hal yang terkait dengan atau melukiskan secara sistematis fakta-fakta atau karakteristik populasi atau pola tertentu dalam bidang tertentu secara faktual dan cermat.

3.1.3. Pendekatan Penelitian

Pada penelitian ini, peneliti menggunakan pendekatan kuantitatif dimana peneliti akan melakukan penelitian sesuai alur yang telah peneliti buat serta dalam penelitian ini untuk proses pengumpulan data akan dilakukan dengan metode observasi.

3.2. Metode Pengumpulan Data

Metode pengumpulan data pada penelitian ini menggunakan metode observasi berstruktur. Pada proses pengumpulan dataset, peneliti sudah mengetahui secara terstruktur proses pencarian dataset dan mengetahui variabel yang akan digunakan dalam penelitian. Data yang digunakan pada analisis ini berasal dari data time series yang diperoleh dari database BMKG. Jumlah data yang diambil pada rentang waktu Januari 2017 - Desember 2021. Selanjutnya akan dilakukan pengolahan data untuk memastikan bahwa data yang digunakan tidak terdapat data dengan nilai kosong (tanpa nilai) atau *missing value*.

3.3. Metode Analisis Data

Analisis data dilakukan menggunakan *preprocessing* untuk memastikan bahwa tidak terdapat data dengan nilai kosong (tanpa nilai). Selanjutnya data akan dilakukan normalisasi kedalam nilai *fuzzy* dimana nilai data berkisar diantara 0 – 1. Untuk analisa performa akan dibandingkan hasil pengolahan data menggunakan metode ANFIS dengan metode ANFIS-GA serta dengan skenario yang sudah disiapkan sebelumnya. Hasil tersebut dinilai dari metode mana yang menghasilkan nilai *error* terkecil dan selisih nilai hasil akurasi dalam pemrosesan data dan juga skenario mana yang menghasilkan nilai *error* terkecil. Metode pada penelitian ini menggunakan metode deskriptif analisis. Dari metode ini data yang diperoleh disajikan dalam berbentuk grafik atau tabel dan bisa juga dalam bentuk angka seperti hasil perhitungan.

3.4. Dataset

Dataset yang digunakan dalam penelitian ini didapatkan dari database online BMKG kabupaten Sleman. Pengumpulan data diambil dari bulan Januari 2014 hingga bulan September 2022 dan didapatkan data sejumlah 3045 data. Selanjutnya data tersebut perlu dilakukan filtering data dimana variabel – variabel yang tidak dibutuhkan dihapus. Pada penelitian ini akan menggunakan variabel yang meliputi data temperatur (T_{avg}), kelembapan (RH_{avg}), kecepatan angin (ff_{avg}), dan curah hujan (RR), sehingga variabel selain itu akan dihapus atau tidak digunakan.

Dataset yang digunakan dalam penelitian ini sama dengan penelitian sebelumnya yang dilakukan oleh (Anggraini, 2018), namun variabel yang digunakan berbeda, dimana pada penelitian tersebut menggunakan 3 (tiga) parameter yang meliputi temperatur, kelembapan, dan curah hujan serta range waktu yang diambil mulai dari 2010 hingga 2015. Kemudian jenis dataset yang digunakan berbeda dengan penelitian sebelumnya yang dilakukan oleh (Damayanti & Cahyadi, 2016) dimana pada penelitian tersebut menggunakan dataset *time series* tumbuh kembang bayi. Tabel 3.1. merupakan sampel hasil dari pengumpulan data. Data tersebut perlu dilakukan filtering dan pemrosesan terlebih dahulu agar data dapat digunakan pada tahap selanjutnya. Tabel 3.2. merupakan sampel hasil dataset dengan variabel yang sudah dilakukan filtering sesuai dengan batasan dalam penelitian ini.

Tabel 3.1. Contoh dataset

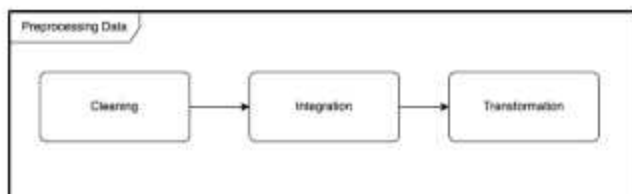
Tanggal	Tn	Tx	Tavg	RH_avg	RR	SS	ff_x	fdd_x	ff_avg	ddd_car
01-03-2016	24	32,2	27,3	89	5,5	1,6	2	90	0	N
02-03-2016	25	32,6	27	88		0,9	2	200	1	N
03-03-2016	24	32,6	26,9	92	3,9	2,7	1	160	0	N
04-03-2016	24	30,6	26,5	89	11,5	4,9	2	180	0	N
05-03-2016	24	30,2	25,7	96	8888	2,4	2	200	0	N
06-03-2016	23	32,4	25,9	93		2	2	200	0	N
07-03-2016	23	33	26,6	89	2,5	5,6	2	200	0	N
08-03-2016	24	32,8	26,5	92	24	6,4	3	180	1	S
09-03-2016	23	33	26,6	88	123	6,8	3	180	1	N
10-03-2016	25	32,2	27,2	92		5,8	2	160	0	N
11-03-2016	26	30,4	27	93	2	0,6	2	180	0	N
12-03-2016	24	32,8	27,2	87	2,2	0	2	180	1	N
13-03-2016	24	33,2	27,1	87	0		2	180	0	N
14-03-2016	24	33,4	27,8	83	8888	6,9	2	200	0	N
15-03-2016	23	32,8	27,7	83		8,4	2	180	1	S

Tabel 3.2. Contoh dataset dengan variabel yang digunakan.

Tanggal	Tavg	RH_avg	RR	FF_avg
01-03-2016	27,3	89	5,5	0
02-03-2016	27	88		1
03-03-2016	26,9	92	3,9	0
04-03-2016	26,5	89	11,5	0
05-03-2016	25,7	96	8888	0
06-03-2016	25,9	93		0
07-03-2016	26,6	89	2,5	0
08-03-2016	26,5	92	24	1
09-03-2016	26,6	88	123	1
10-03-2016	27,2	92		0
11-03-2016	27	93	2	0
12-03-2016	27,2	87	2,2	1
13-03-2016	27,1	87	0	0
14-03-2016	27,8	83	8888	0
15-03-2016	27,7	83		1

Tahap selanjutnya dari pengolahan dataset ini adalah melakukan *pre-processing*. Ada beberapa tahapan *preprocessing* yang dilakukan, dimulai dari *cleaning*, *integration*, dan *transformation*. Secara singkat, alur *preprocessing* data dapat dilihat pada Gambar 3.1.

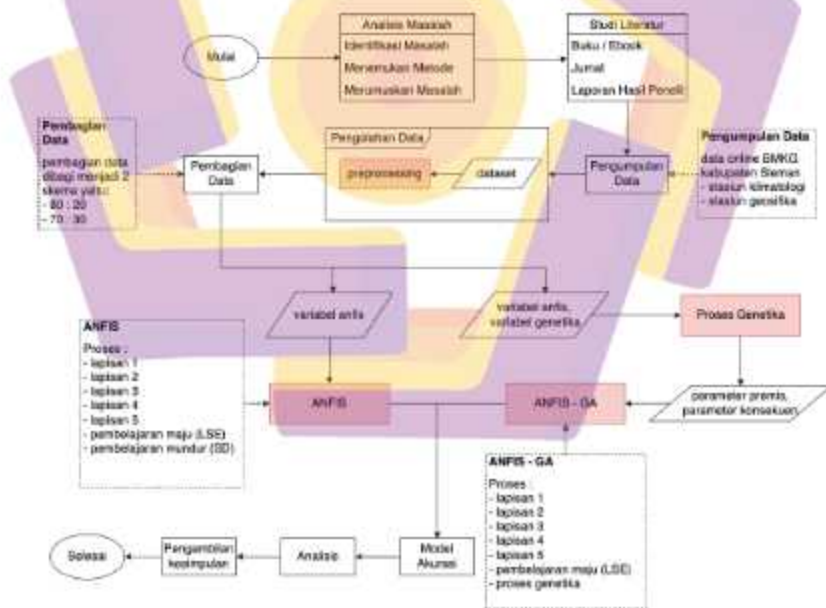
- a. *Cleaning* merupakan tahapan dimana data akan dibersihkan melalui proses seperti mengisi nilai yang hilang, menghaluskan noisy data, dan menyelesaikan inkonsistensi yang ditemukan. Tujuan dari data cleaning ini adalah untuk menyeleksi data dan membuang data yang berpotensi mengurangi akurasi.
- b. *Integration* merupakan tahapan untuk menggabungkan data dari berbagai sumber menjadi satu kesatuan data yang lebih besar. Tahap kerja ini merupakan proses lanjutan dari data cleansing dengan tujuan untuk membuat data lebih halus. Dalam penelitian ini menggabungkan data dari 2 sumber berbeda yaitu stasiun geofisika dan stasiun klimatologi.
- c. *Transformation* merupakan tahapan dimana dalam proses tersebut dataset akan dinormalisasi dan digeneralisasikan. Normalisasi sendiri merupakan suatu proses memastikan bahwa tidak ada data yang berlebihan. Pada penelitian ini, metode normalisasi menggunakan teknik min – max. sama halnya dengan penelitian (Sinaga dkk., 2019), pada dataset yang digunakan dalam penelitian ini akan menerapkan teknik normalisasi min – max dengan scala nilai fuzzy atau dengan *range* nilai 0 – 1. Penerapan normalisasi data menurut (Sinaga dkk., 2019) dapat mempengaruhi performa yang dihasilkan.



Gambar 3.1. Tahapan *preprocessing*.

3.5. Alur Penelitian

Alur penelitian diilustrasikan pada gambar 3.2. dimulai dari analisis masalah, studi literatur, pengumpulan data, *filtering data*, *preprocessing data*, pembuatan model genetika, pelatihan dan pengujian model ANFIS hingga mendapatkan hasil prediksi serta akurasi.



Gambar 3.2. Alur penelitian

Alur penelitian dimulai dengan melakukan analisis masalah. Dalam analisis masalah dilakukan identifikasi masalah kemudian menentukan metode yang akan digunakan hingga merumuskan masalah. Selanjutnya tahapan akan masuk ke tahap studi literatur dimana dalam tahapan ini akan dilakukan pencarian referensi baik dari buku, *ebook*, jurnal, serta laporan hasil penelitian. Hasil dari tahapan studi literatur akan berupa tabel *literatur review*.

Selanjutnya yaitu masuk kedalam tahapan pengumpulan data. Pengumpulan data bersumber dari data online BMKG kabupaten Sleman yang diambil dari stasiun geofisika dan stasiun klimatologi. Dari hasil pengumpulan data tersebut, didapatkan data mentah sebanyak 3.045 baris data dengan 10 variabel yang selanjutnya dilakukan *filtering* sehingga variabel yang tersisa sejumlah 4 variabel. Dataset yang digunakan dalam penelitian ini berjenis *time series multivariate*.

Dari data tersebut, akan dilakukan proses *preprocessing* untuk membersihkan data dan mengatasi jika terdapat data yang *missing value* dan kemudian dilakukan normalisasi data. Proses normalisasi data akan menggunakan teknik min – max dengan scala nilai fuzzy atau dengan *range* nilai 0 – 1. Penerapan normalisasi data menurut (Sinaga dkk., 2019) dapat mempengaruhi performa yang dihasilkan.

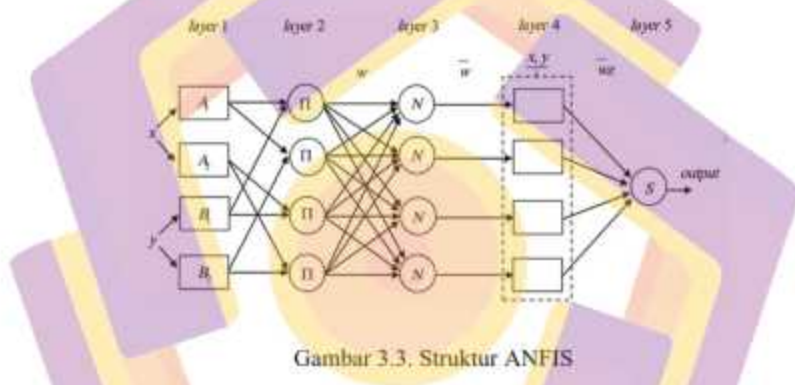
Dataset yang tersedia, selanjutnya dilakukan pembagian data. Tujuan dari pembagian data yaitu untuk memisahkan data latih dan data uji dengan proporsi 80:20 dan 70:30. Pembagian dengan proporsi ini dapat merepresentasikan data, evaluasi model yang cukup cepat, dapat menggambarkan distribusi yang sama pada

data latih dan data uji, dan mendapatkan akurasi terbaik (Pratama & Tjahyanto, 2021).

Proses *training* data dilakukan menggunakan dua jenis data, yaitu data original dan data yang telah dinormalisasi. Penggunaan dua jenis data tersebut bertujuan untuk mengetahui seberapa efektif proses normalisasi data pada dataset *time series multivariate* curah hujan yang *inconsisten*. Beberapa penelitian sebelumnya melatih model menggunakan jumlah iterasi yang berbeda-beda. Sebagai contoh, penelitian yang dilakukan oleh (Anggraini, 2018) melatih model dengan iterasi 100. Penelitian yang dilakukan oleh (Wahyuni dkk., 2017) melatih model dalam 300 iterasi. Pada penelitian yang dilakukan (Azhar & Mahmudy, 2018) melakukan uji coba pada jumlah iterasi yang digunakan dimana dalam penelitian tersebut iterasi yang digunakan yaitu 100 hingga 1000 dengan kelipatan 100. Uji coba tersebut ditujukan untuk mencari iterasi mana yang menghasilkan akurasi terbaik. Dalam penelitian ini model dicoba untuk dilatih hingga mencapai 100 iterasi.

Dalam proses *training*, dataset akan dicoba menggunakan dua model, dimana dalam penelitian ini, akan dilakukan dua permodelan yaitu dengan ANFIS dan juga ANFIS-GA. Dua model tersebut bertujuan untuk membandingkan hasil dari metode ANFIS yang akan dioptimasi dengan GA dengan ANFIS yang tanpa optimasi. Selain untuk membandingkan hasil, juga akan dilakukan monitoring apakah algoritma genetika memberikan pengaruh kepada metode ANFIS dengan dataset yang *inconsisten*.

Pada proses dengan model ANFIS, tentunya dataset akan melalui tahapan – tahapan sebagaimana model ANFIS bekerja, dimana akan terdapat 5 lapisan dengan algoritma pembelajaran maju dan mundur. Algoritma pembelajaran maju dalam ANFIS akan digunakan metode *LSE Recursive*, serta untuk algoritma pembelajaran mundur akan digunakan algoritma *steepest descent*. Algoritma *steepest descent* ditujukan untuk memperbarui parameter premis yang akan digunakan dalam iterasi berikutnya. Struktur ANFIS dapat dilihat pada gambar 3.3.



Berbeda pada model ANFIS-GA, dimana model ANFIS akan dioptimasi menggunakan algoritma genetika. Tahapan pada model ANFIS-GA dimulai dengan melakukan pemilihan operator genetika. Pemilihan operator genetika merupakan suatu tahapan yang penting sebelum melakukan proses genetika. Pemilihan operator genetika ditahapan ini meliputi pemilihan jenis *crossover*, mutasi, dan seleksi. Selain itu variabel yang akan digunakan juga ditentukan dalam tahapan ini yang meliputi *membership function*, *max generation*, *mutation rate*, *learning rate*, dan *population*.

Berikutnya tahapan akan masuk dalam proses genetika. **Proses genetika merupakan suatu tahapan dimana proses ini menggantikan peran dari algoritma pembelajaran arah mundur (*steepest descent*). Proses ini ditujukan sebagai optimasi pada metode ANFIS.** Proses ini akan menghasilkan nilai parameter premis dan parameter konsekuen yang nantinya akan diteruskan menggunakan metode ANFIS. Nilai parameter premis dan konsekuen tersebut dihasilkan dari pencarian nilai terbaik dari generasi ke generasi. Pada tahapan awal, populasi awal akan dibangkitkan secara acak dengan rentan nilai yang telah ditetapkan. Kemudian dilakukan proses *crossover* dan mutasi serta seleksi untuk menghasilkan kromosom – kromosom yang mempunyai nilai *fitness* yang lebih tinggi daripada generasi sebelumnya. Selanjutnya untuk menjaga agar generasi yang baik tidak tereliminasi pada saat proses seleksi, maka dilakukan proses *elitism*. Proses *elitism* yaitu mengcopykan kromosom terbaik ke generasi berikutnya. Proses genetika ini akan terus dilakukan sesuai dengan jumlah generasi yang ditentukan.

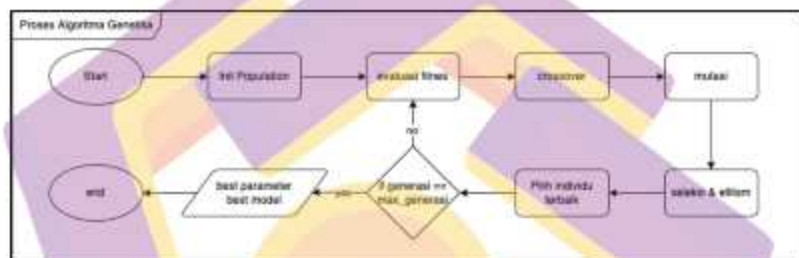
Dari proses genetika, tahapan akan dilanjutkan dengan pembentukan struktur ANFIS. Struktur ANFIS yang dibangun seperti ditunjuk pada gambar 3.5. Pada struktur tersebut metode inferensi samar yang digunakan ialah inferensi samar Sugeno orde I dengan bentuk umum yaitu:

$$\text{if } x \text{ is } A \text{ and } y \text{ is } B \text{ Then } F = px + qy + r$$

Dimana A dan B merupakan parameter premis dan F adalah parameter konsekuen. Dalam penelitian ini, nilai parameter premis dan parameter konsekuen pada lapisan kedua dan keempat yang merupakan parameter adaptif diperoleh dari hasil optimasi pelatihan menggunakan algoritma genetika yang mana pada

umumnya diperoleh dari hasil pelatihan menggunakan algoritma *hybird* maju dan mundur. Dalam penelitian ini parameter – parameter tersebut dioptimasi menggunakan algoritma genetika dan *least square estimator* (LSE), sehingga peran algoritma genetika yaitu untuk menggantikan algoritma *steepest descent*.

3.6. Alur Proses Genetika



Gambar 3.4. Alur proses genetika

Alur proses genetika dapat dijelaskan sebagai berikut:

- Langkah 1 yaitu inialisasi pada data latih. Inialisasi dilakukan untuk mendefinisikan parameter parameter yang nantinya akan digunakan dalam proses optimasi. parameter yang didefinisikan meliputi nilai `populasi_awal`, `jumlah_generasi`, nilai `max_error`, `crossover` serta `probabilitas_mutasi`.
- Langkah 2 yaitu instalasi nilai `populasi_awal`, dilakukan dengan pengambilan data random serta bilangan yang digunakan merupakan bilangan asli dengan batasan untuk nilai sudah ditentukan sebelumnya. nilai `populasi_awal` merupakan penggambaran dari nilai premis yang nantinya akan diolah dan dievaluasi menggunakan sistem. Langkah berikutnya setelah didapat nilai premis yakni melanjutkan dengan mencari nilai konsekuen. pencarian nilai tersebut

menggunakan algoritma yang terdapat dalam metode ANFIS yaitu dengan menggunakan algoritma RLSE.

c. Langkah 3 yaitu setelah didapat nilai premis dan konsekuen, dilanjutkan dengan melakukan proses *crossover* dimana proses ini merupakan proses genetika. Proses pada tahapan ini bertujuan memperoleh nilai baru dengan menciptakan individu baru. Individu yang baru diperoleh berdasarkan sifat asalnya.

d. Langkah 4 selanjutnya dilakukan proses mutasi dengan tujuan mencegah agar tidak terjadi konvergensi. Konvergensi terjadi dikarenakan adanya homogenisasi populasi.

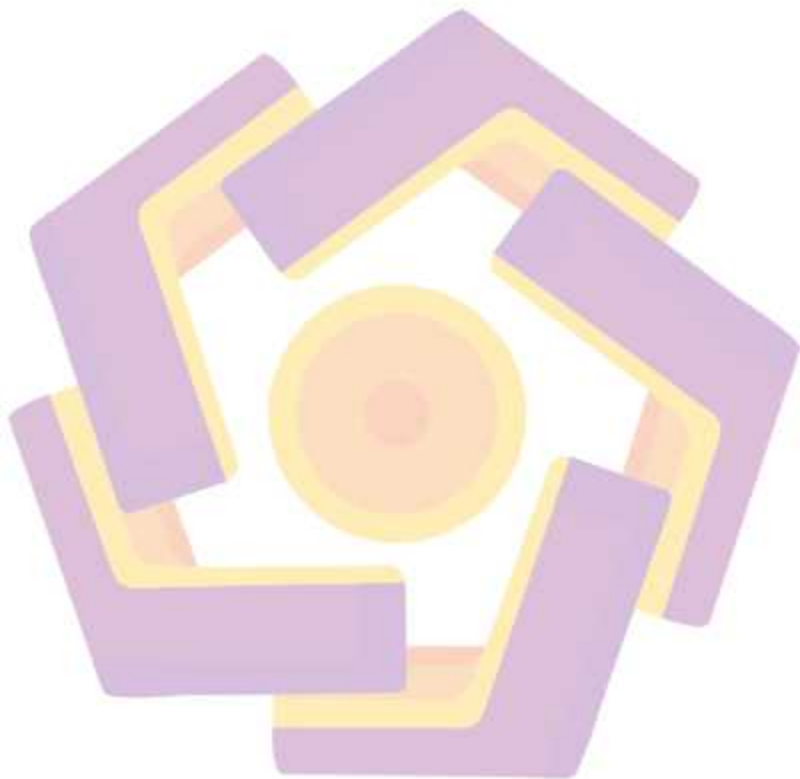
e. Langkah 5 akan dilakukan proses *elitism* yang bertujuan untuk mencegah data individu terbaik hilang yang diakibatkan dari proses pengacakan yang terjadi. Pada proses *elitism* terjadi proses dimana akan dipilihnya dua nilai individu yang memiliki nilai *fitness* terbaik yang akan ditempatkan pada generasi berikutnya.

f. Langkah 6 Setelah proses *elitism*, dilakukan pemilihan terhadap salah satu premis yang mempunyai nilai error paling kecil, atau dapat juga dengan mencari satu premis yang memiliki nilai *fitness* paling besar.

g. Langkah 7 setelah mendapat nilai premis, dilakukan pengecekan iterasi, apakah iterasi sudah mencapai batas maksimal seperti yang dimasukkan pada saat instalasi atau belum.

h. langkah 8 setelah mendapat nilai premis, dilakukan pengecekan iterasi, apakah iterasi sudah mencapai batas maksimal seperti yang dimasukkan pada saat instalasi atau belum. Jika iterasi belum mencapai batas maksimal, maka proses akan

kembali diulang dengan genetika berikutnya. Proses yang diulang meliputi proses *crossover*, mutasi, dan elitism.



BAB IV

HASIL PENELITIAN DAN PEMBAHASAN

4.1. Pengumpulan Data

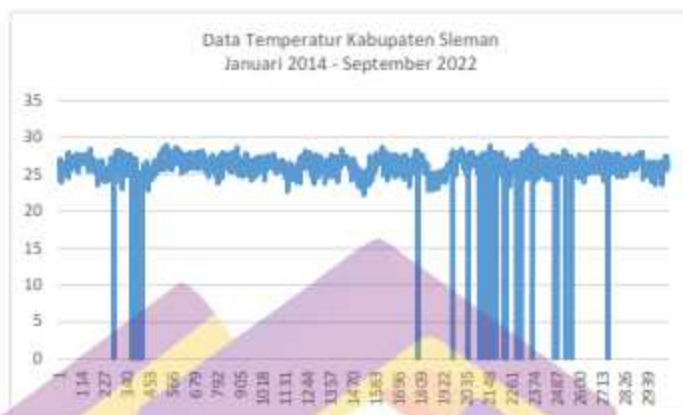
Pada penelitian ini, dataset yang digunakan berjenis *time series multivariate* dimana dataset yang digunakan terdapat 4 (empat) variabel. Berdasarkan pengumpulan data dari Januari 2014 hingga September 2022 didapatkan data sejumlah 3045 data yang meliputi data temperatur, kelembapan, kecepatan angin, dan curah hujan. Parameter yang digunakan berbeda dengan penelitian sebelumnya yang dilakukan oleh (Angraini, 2018) dimana pada penelitian tersebut menggunakan 3 (tiga) parameter yang meliputi temperatur, kelembapan, dan curah hujan serta range waktu yang diambil mulai dari 2010 hingga 2015.

Data *time series* yang digunakan berupa data-harian agar dataset yang didapat untuk pelatihan dan pengujian berjumlah banyak, berbeda dengan penelitian sebelumnya yang dilakukan oleh (Sinaga dkk., 2019) yang menggunakan data *time series* bulanan dengan rentang waktu 2014 – 2018. Data tersebut diperoleh dari dataset online BMKG kabupaten Sleman pada badan stasiun klimatologi dan geofisika. Dataset yang di peroleh merupakan data mentah yang belum siap digunakan. Dataset tersebut harus melewati tahapan – tahapan berikutnya agar dataset dapat digunakan untuk pelatihan dan pengujian metode. Dataset yang diperoleh divisualisasikan dalam bentuk tabel serta grafik yang ditunjukkan pada tabel 4.1.

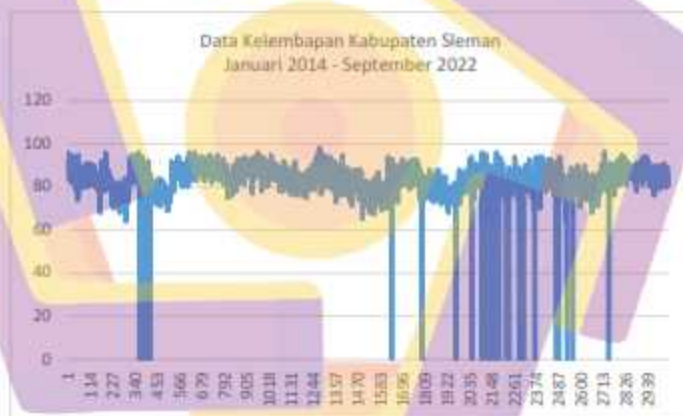
Tabel 4.1. Dataset

No	Tanggal	T	RH	FF	RR
1	1-Jan-2015	24,3	96	1	8,1
2	2-Jan-2015	26,2	87		21,7
3	3-Jan-2015	26,7	84	1	0
4	4-Jan-2015	26,6	86	1	0
5	5-Jan-2015	25,1	93	2	8
...					
...					
3041	26-Sep-22	26,3	85	2	
3042	27-Sep-22	26,2	86	1	0,5
3043	28-Sep-22	26,3	81	2	0
3044	29-Sep-22	26,2	82	2	0
3045	30-Sep-22	25,7	89	1	0,2

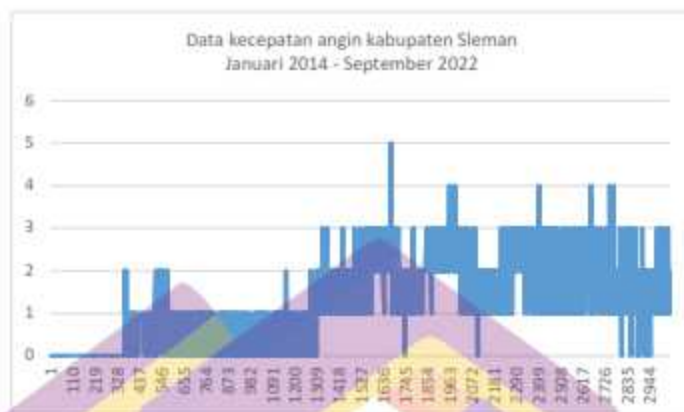
Selama rentang waktu periode januari 2014 – september 2022 diperoleh data cuaca yang cukup beragam sebanyak 3045 data. Data tersebut merupakan data yang tersedia pada database online BMKG kabupaten sleman. Gambar 4.1 menunjukkan grafik data temperatur. Pada grafik tersebut terdapat beberapa kali data kosong dikarenakan tidak adanya pendataan pada tanggal tersebut. Sama halnya pada gambar 4.2 yang menunjukkan grafik data kelembapan yang tidak dilakukan pendataan pada tanggal - tanggal tertentu hingga pada grafik terlihat turun secara signifikan. Pada gambar 4.3 dan gambar 4.4 yang merupakan grafik data kecepatan angin dan curah hujan terlihat stabil kenaikan dan penurunan data sesuai dengan musim yang sedang berlangsung.



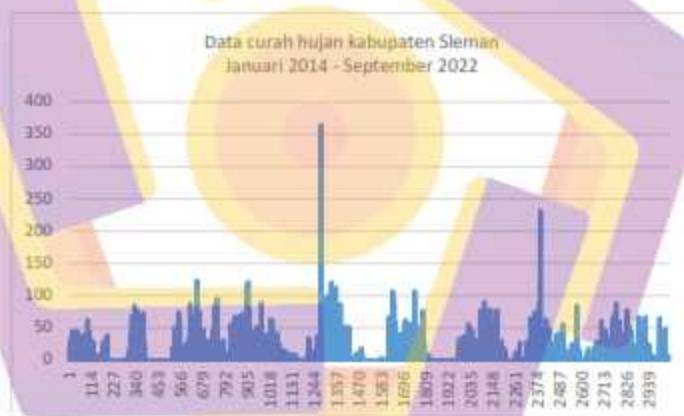
Gambar 4.1. Sebaran data temperatur



Gambar 4.2. Sebaran data kelembapan



Gambar 4.3. Sebaran data kecepatan angin



Gambar 4.4. Sebaran data curah hujan

4.2. Pengolahan Data

Berdasarkan data yang didapat, data tersebut selanjutnya akan melalui proses pengolahan data yang terdiri dari proses *cleaning*, *integration*, *transformation*, dan *reduction*. Proses tersebut bertujuan agar data yang digunakan dalam proses pelatihan dan pengujian bersih dari nilai error atau nilai kosong.

Proses preprocessing juga dilakukan dalam penelitian yang dilakukan oleh (Anggraini, 2018) dimana dalam penelitian tersebut dilakukan proses preprocessing sebelum dilakukan proses pengolahan menggunakan algoritma yang di tentukan. Dalam penelitian tersebut dijelaskan bahwa proses preprocessing ditujukan untuk menskalakan nilai – nilai yang telah disesuaikan dengan beberapa parameter.

Dalam implementasinya, data mining tidak dapat memproses data mentah, sehingga proses ini sangat penting dilakukan untuk mempermudah proses berikutnya, yakni analisis data. Proses ini utamanya dilakukan untuk memastikan kualitas data baik sebelum digunakan saat analisis data. Teknik preprocessing yang digunakan ialah teknik data *cleaning* dan *imputation* atau pengisian pada nilai *missing value* dengan nilai *average*. Hasil preprocessing data dapat dilihat pada tabel 4.2. Dari data awal sejumlah 3045 setelah melalui proses *preprocessing*, data yang dapat digunakan menjadi sejumlah 3045 karena menggunakan teknik *imputation*.

Tabel 4.2. Dataset setelah preprocessing

	Tanggal	T	RH	FF	RR
1	1-Jan-2015	24,3	96	1	8,1
2	2-Jan-2015	26,2	87	1	21,7
3	3-Jan-2015	26,7	84	1	0
4	4-Jan-2015	26,6	86	1	0
5	5-Jan-2015	25,1	93	2	8
...					
...					
3041	26-Sep-22	26,3	85	2	2,2
3042	27-Sep-22	26,2	86	1	0,5
3043	28-Sep-22	26,3	81	2	0
3044	29-Sep-22	26,2	82	2	0
3045	30-Sep-22	25,7	89	1	0,2

Pada tabel 4.2. merupakan dataset yang telah dilakukan proses *preprocessing*. Setelah melalui proses *preprocessing*, dataset akan melalui proses normalisasi data. Normalisasi data bertujuan untuk menskalakan data suatu atribut sehingga berada dalam rentang yang lebih kecil. Pada dataset yang digunakan dalam penelitian (Sinaga dkk., 2019) dilakukan proses normalisasi menggunakan teknik min – max. sama halnya dengan penelitian (Sinaga dkk., 2019), pada dataset yang digunakan dalam penelitian ini akan menerapkan teknik normalisasi min – max dengan *scala* nilai fuzzy atau dengan *range* nilai 0 – 1. Data hasil normalisasi dapat dilihat pada tabel 4.3.

Tabel 4.3. Dataset normalisasi.

No.	Tanggal	T	RH	FF	RR
1	1-Jan-2015	0,3	0,9	0,2	8,1
2	2-Jan-2015	0,6	0,7	0,3	21,7
3	3-Jan-2015	0,6	0,6	0,2	0
4	4-Jan-2015	0,6	0,6	0,2	0
5	5-Jan-2015	0,4	0,9	0,4	8
...					
...					
3041	26-Sep-22	0,6	0,6	0,4	7,7
3042	27-Sep-22	0,6	0,6	0,2	0,5
3043	28-Sep-22	0,6	0,5	0,4	0
3044	29-Sep-22	0,6	0,5	0,4	0
3045	30-Sep-22	0,5	0,7	0,2	0,2

4.3. Mekanisme Training

Secara umum proses eksperimen dalam penelitian ini dimulai dari proses persiapan data. Data yang sudah siap dan telah diproses melalui teknik *preprocessing* data, selanjutnya akan dilakukan pengolahan menggunakan 2 (dua) jenis pengujian yaitu menggunakan metode ANFIS dan pengujian dengan

menerapkan algoritma genetika untuk mengoptimasi metode ANFIS. Penggunaan atau pemilihan algoritma genetika untuk mengoptimasi metode ANFIS dikarenakan menurut (Wahyuni dkk., 2017) dan (Damayanti & Cahyadi, 2016) algoritma genetika mampu memberikan kinerja yang bagus pada ANFIS serta dapat mengatasi permasalahan yang terjadi pada metode ANFIS.

Pada jenis pengujian menggunakan ANFIS, dataset yang telah dilakukan proses preprocessing akan diolah menggunakan metode ANFIS dengan 5 lapisan, algoritma pembelajaran maju (*LSE Recursive*) serta algoritma pembelajaran mundur (*steepest descent*). Untuk jenis pengujian ANFIS – GA, dataset akan diolah terlebih dahulu menggunakan algoritma genetika untuk mendapatkan parameter premis dan konsekuen terbaik sebelum dilanjutkan dengan metode ANFIS dengan menghilangkan proses pada algoritma *steepest descent*.

Dalam penelitian ini digunakan *tools* matlab serta python sebagai *environment* untuk melakukan eksperimen. Parameter yang digunakan untuk melakukan training berupa *epoch / generations*, *learning rate*, *input function*, *membership function*, *offspring*, *mutation rate*. Pada penelitian ini, *membership function* yang digunakan yaitu jenis *gaussmf*, berbeda dengan penelitian sebelumnya oleh (Anggraini, 2018) yang menggunakan jenis *tripfmf* dan (Azhar & Mahmudy, 2018) yang menggunakan jenis *gbellmf*.

Dataset yang digunakan dalam penelitian ini merupakan dataset *time series* hairan yang berjumlah 3045 baris data yang sudah dilakukan proses *preprocessing* data dan normalisasi data dengan 4 parameter yang akan digunakan. Pada penelitian ini dataset akan dibagi menjadi 80:20 dan 70:30 masing – masing untuk data

training dan data testing. Penggunaan 2 (dua) skema yang berbeda ditujukan untuk membandingkan antara kedua skema agar dapat diketahui skema mana yang menghasilkan akurasi hasil yang lebih baik. Pemilihan skema tersebut didasarkan pada *machine learning mastery* dan oleh (Rosebrock, 2018). Dalam pengujiannya akan digunakan 2 (dua) jenis dataset yaitu dataset dengan normalisasi dan tanpa normalisasi.

Tabel 4.4. Variasi Skenario Training Data

Metode	Training Data	Skema Data
ANFIS	Normalisasi	80:20
		70:30
	Non Normalisasi	80:20
		70:30
ANFIS - GA	Normalisasi	80:20
		70:30
	Non Normalisasi	80:20
		70:30

Tabel 4.5. Parameter ANFIS

Parameter	Value
Epoch	100
Learning Rate	0.2
Membership Function	<i>Gaussmf</i>

Proses training dalam penelitian ini dilakukan dengan beberapa skenario dengan beberapa variasi variabel yang berbeda seperti jenis metode yang digunakan, jenis data training yang digunakan, serta skema pengujian dataset yang digunakan. Secara singkat, tabel 4.4. dan tabel 4.5. menunjukkan variasi skenario yang akan digunakan dalam penelitian ini.

4.4. Model ANFIS

Pada model ANFIS dalam pengujian ini, semua tahapan dalam metode ANFIS akan dijalankan. Mengingat ANFIS merupakan gabungan antara JST dengan *Fuzzy*, seperti yang dituliskan oleh (Kusumadewi S & Hartati S, 2010) maka dalam pengujian ini akan melalui 5 lapisan dan algoritma *hybird* yang meliputi algoritma pembelajaran maju dan algoritma pembelajaran mundur. Dalam pengujian ini akan diterapkan fungsi fungsi sesuai dari metode ANFIS atau dalam artian proses pada pengujian ini dilakukan tanpa adanya perubahan atau optimasi dalam struktur ANFIS.

Dalam pembuatan model, *library* yang digunakan adalah *anfisedit* yang telah disediakan oleh matlab. Data yang digunakan akan terlebih dahulu di proses diluar aplikasi dan akan di import kedalam aplikasi matlab dalam keadaan data siap digunakan untuk pelatihan dan pengujian. Seperti yang dijabarkan pada mekanisme pelatihan, pada proses pengujian dengan menggunakan ANFIS ini akan digunakan parameter *epoch* sebesar 100, *membership function* jenis *gaussmf*, serta *learning rate* sebesar 0,2 dengan skema data 80:20 dan 70:30.

4.4.1. Pelatihan Data

Sesuai dengan pembagian dataset untuk pelatihan, dalam pelatihan ANFIS ini akan digunakan data sebanyak 2436 dan 2132. Hasil pelatihan dengan 2 skema dapat dilihat pada tabel 4.6 dan tabel 4.7.

Table 4.6. Sampel hasil pelatihan

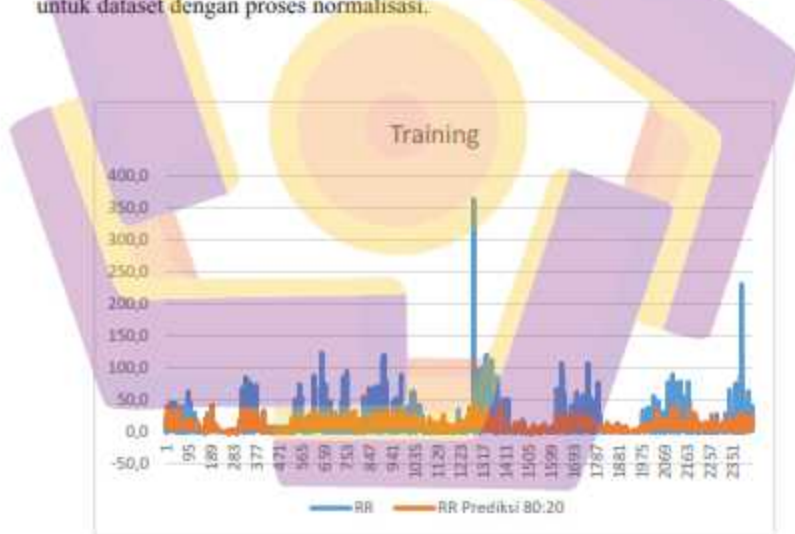
RR	RR (80:20)	RR (70:30)
8,1	39,0	40,4
21,7	10,9	10,9
0,0	6,6	6,4
0,0	8,6	8,4
8,0	25,3	23,0
0,2	16,3	16,1
15,1	17,5	17,4
0,0	6,4	6,5
0,2	6,7	7,0
0,0	5,7	5,5
0,0	21,6	20,2
34,2	25,4	23,1
28,0	21,8	21,7
25,9	13,0	12,8

Tabel 4.7. Sampel hasil pelatihan dengan proses normalisasi

RR	RR (80:20)	RR (70:30)
8,1	27,6	30,7
21,7	13,8	13,4
0,0	6,6	7,6
0,0	6,6	7,6
8,0	22,5	22,0
0,2	17,8	17,2
15,1	20,9	21,1
0,0	6,2	7,3
0,2	6,6	8,1
0,0	6,3	7,3
0,0	19,2	17,8
34,2	22,5	22,0
28,0	20,9	21,1
25,9	17,8	17,2

Dari tabel 4.6 dan 4.7 dapat ditarik kesimpulan bahwa penggunaan rasio data 80:20 dan 70:30 perbedaannya tidak begitu signifikan dalam pelatihan data. Terlihat bahwa dari masing – masing rasio data mampu menghasilkan nilai prediksi

yang mendekati nilai asli. Namun pada beberapa data terutama pada data yang bernilai atau mendekati nilai 0, dari kedua rasio kurang baik menghasilkan nilai yang jauh dari nilai asli. Hal tersebut bukan karena pembagian data, namun hal tersebut terjadi dikarenakan nilai 0 pada dataset yang terlampaui banyak. Mengingat dalam dataset tercantum data pada musim kemarau dan pada saat tidak terjadi hujan yang nilai parameter curah hujan sama dengan 0. Sehingga dari hal tersebut yang membuat nilai prediksi yang dihasilkan kurang optimal. Untuk visualisasi data prediksi pada masing – masing rasio data, dapat dilihat pada gambar grafik 4.5. dan gambar grafik 4.6. untuk dataset tanpa normalisasi dan gambar 4.7. dan gambar 4.8. untuk dataset dengan proses normalisasi.



Gambar 4.5. Grafik training data 80:20

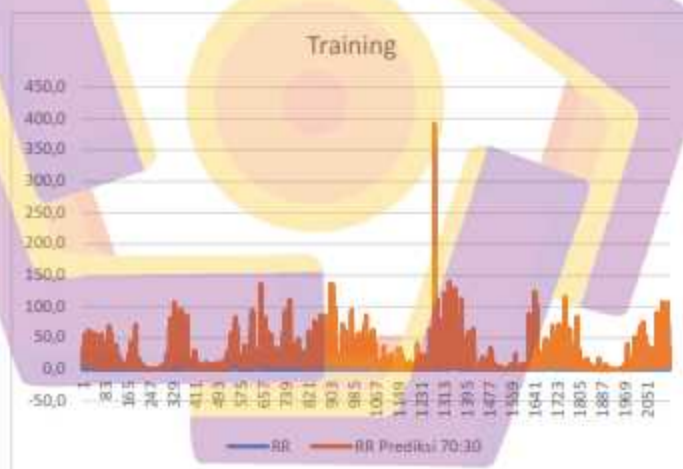


Gambar 4.6. Grafik training data 70:30

Pada gambar 4.5, yang merupakan hasil dari pelatihan data dengan rasio 80:20, prediksi yang dihasilkan memiliki rata-rata dibawah nilai asli dan terlihat stabil mengikuti pola pada nilai asli. Berbeda dengan gambar 4.6, dengan rasio data 70:30 dimana nilai prediksi yang dihasilkan relatif sama dengan nilai asli. Sehingga pada grafik terlihat parameter nilai prediksi menutupi parameter nilai asli. Pada gambar 4.6 terjadi lonjakan beberapa kali dan hal tersebut dikarenakan pada dataset memang terdapat lonjakan nilai curah hujan yang cukup signifikan. Hal tersebut merupakan hal yang wajar terjadi mengingat pada musim hujan terkadang terdapat salah satu atau beberapa hari dengan intensitas curah hujan lebat.



Gambar 4.7. Grafik training dataset 80:20 dengan normalisasi



Gambar 4.8. Grafik training dataset 70:30 dengan normalisasi

Pada gambar 4.7. dan gambar 4.8. merupakan grafik pelatihan menggunakan rasio data 80:20 dan 70:30 dengan proses normalisasi dataset terbelah dahulu. Pada gambar 4.7. dan 4.8. mampu menghasilkan nilai prediksi yang mendekati nilai asli

seperti yang tertera pada sampel hasil prediksi pada tabel 4.7, dalam hal tersebut dapat disimpulkan bahwa proses normalisasi memberikan efek pada hasil dan akurasi nilai prediksi walau dalam penelitian ini efek yang diberikan tidak begitu signifikan. Ada beberapa faktor yang mempengaruhi kinerja normalisasi salah satunya yaitu nilai 0 yang terdapat dalam dataset terlampaui banyak.

4.4.2. Pengujian Data

Pada tahapan pengujian data ini akan digunakan parameter dan model terbaik dari proses pelatihan data sebelumnya. Parameter terbaik dan model terbaik yang didapat akan digunakan dalam skema pengujian data 80:20 dan 70:30. Pengujian ini akan menggunakan data sebesar 609 dan 913. Hasil pengujian dapat dilihat pada tabel 4.8 dan tabel 4.9.

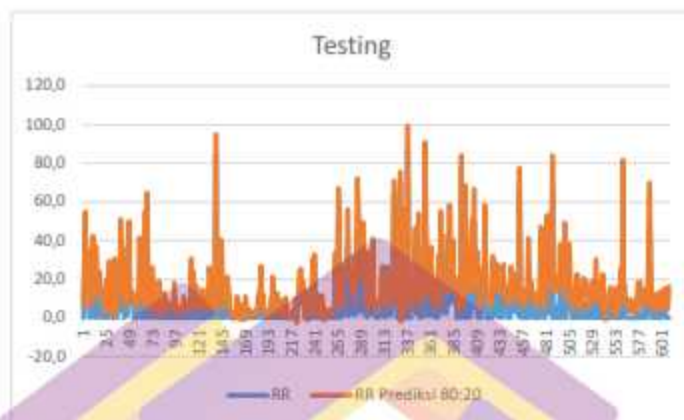
Tabel 4.8. Sampel hasil pengujian

RR	RR (80:20)	RR (70:30)
14,9	9,9	9,7
0,0	6,9	6,7
20,8	11,1	11,1
47,2	7,8	8,1
14,3	8,6	10,4
16,7	7,4	8,8
27,8	7,5	7,7
0,9	4,4	4,1
0,0	9,6	9,0
8,2	24,2	24,3
9,2	12,9	12,7
14,7	27,5	27,7
6,7	4,8	5,0
0,0	11,1	14,1
23,8	13,3	16,2

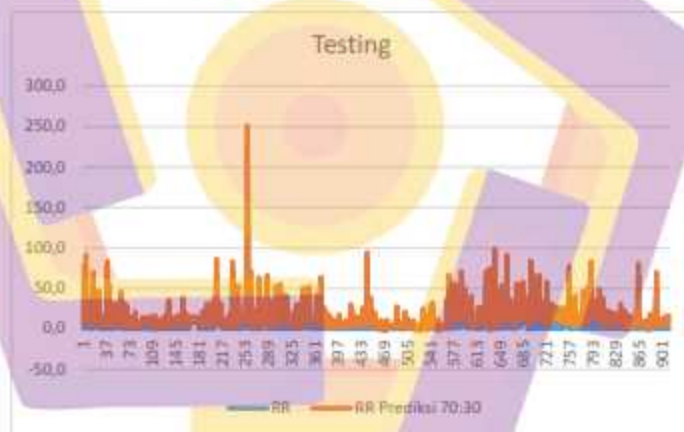
Tabel 4.9. Sampel hasil pengujian dengan normalisasi

RR	RR (80:20)	RR (70:30)
14,9	7,0	6,9
0,0	6,6	7,6
20,8	13,4	13,0
47,2	6,6	8,1
14,3	6,1	8,6
16,7	6,1	8,6
27,8	6,6	8,1
0,9	4,4	3,6
0,0	6,8	7,1
8,2	24,6	25,9
9,2	13,4	13,0
14,7	25,0	27,2
6,7	3,3	3,4
0,0	12,0	13,7
23,8	15,3	13,1

Dari tabel 4.8 dan 4.9 dapat ditarik kesimpulan bahwa penggunaan rasio data 80:20 dan 70:30 perbedaannya tidak begitu signifikan dalam pengujian data. Terlihat bahwa dari masing – masing rasio data mampu menghasilkan nilai prediksi yang mendekati nilai asli. Hal tersebut tak lepas dari model yang dihasilkan dari pelatihan data sebelumnya. Namun pada kedua rasio data tersebut, masih belum begitu bagus dalam hal memprediksi nilai asli yang bernilai atau mendekati nilai 0. Untuk visualisasi data prediksi pada masing – masing rasio data, dapat dilihat pada gambar grafik 4.9. dan gambar grafik 4.10. untuk dataset tanpa normalisasi dan gambar 4.11. dan gambar 4.12. untuk dataset dengan proses normalisasi.

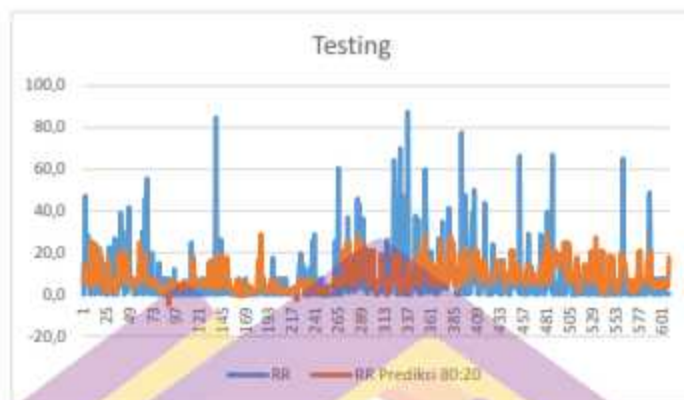


Gambar 4.9. Testing dataset 80:20

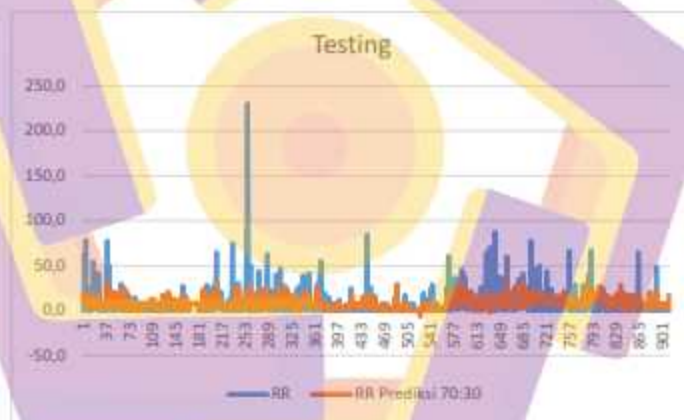


Gambar 4.10. Testing dataset 70:30

Pada gambar 4.9. yang merupakan hasil dari pengujian data dengan rasio 80:20, prediksi yang dihasilkan terlihat stabil mengikuti pola pada nilai asli. Sama halnya dengan gambar 4.10. dengan rasio data 70:30 dimana nilai prediksi yang dihasilkan juga relatif sama dengan nilai asli.



Gambar 4.11. Testing dataset 80:20 dengan normalisasi



Gambar 4.12. Testing dataset 70:30 dengan normalisasi

Pada gambar 4.11. dan gambar 4.12. merupakan grafik pengujian menggunakan rasio data 80:20 dan 70:30 dengan proses normalisasi dataset terlebih dahulu. Pada gambar 4.11. dan gambar 4.12. mampu menghasilkan nilai prediksi yang mendekati nilai asli seperti yang tertera pada sampel hasil prediksi pada tabel

4.9. dalam hal tersebut dapat disimpulkan bahwa proses normalisasi memberikan efek pada hasil dan akurasi nilai prediksi walau dalam penelitian ini efek yang diberikan tidak begitu signifikan.

4.4.3. Perhitungan Akurasi

Perhitungan Akurasi dan nilai error dilakukan dengan menggunakan perhitungan RMSE. Pada proses pengujian dan pelatihan, jumlah *epoch* yang digunakan yaitu 100. Penggunaan *epoch* 100 didasari karena selama proses percobaan, setelah iterasi ke 100 nilai RMSE sudah menunjukkan nilai yang datar atau sudah tidak terjadi perubahan sehingga pada iterasi ke 100 sudah didapatkan nilai RMSE yang terbaik pada proses yang berlangsung. Berbeda halnya ketika jumlah *epoch* kurang dari 100, hal tersebut berakibat pada nilai RMSE yang didapat belum maksimal atau belum hasil terbaik. Hal tersebut dikarenakan jumlah dataset yang digunakan cukup banyak sehingga jika *epoch* terlalu kecil, maka model yang dihasilkan belum maksimal. Dari hasil pelatihan dan pengujian menggunakan parameter berupa epoch 100 serta *membership function gaussian* dan 2 (dua) skema dapat dilihat pada tabel 4.10.

Tabel 4.10. Perbandingan Akurasi

Skema	normalisasi	Epoch	Data Lath	Data Test	RMSE
80:20	-	100	2436	609	15,2780
70:30	-	100	2132	913	15,4345
80:20	ya	100	2436	609	12,4732
70:30	ya	100	2132	913	13,0982

Dari tabel 4.10 dapat disimpulkan bahwa pembagian skema dataset dan dilakukannya proses normalisasi data terlebih dahulu sangat berpengaruh pada nilai

akurasi yang didapat. Nilai tersebut dapat dibandingkan pada tabel 4.6, tabel 4.7, tabel 4.8, dan tabel 4.9. Dari tabel tersebut tersaji sampel data hasil prediksi untuk dengan menggunakan model yang sudah dibentuk pada pelatihan sebelumnya dengan skema data yang sama. Pada pengujian tanpa normalisasi didapat nilai RMSE sebesar 15,7789 dan 15,4345 untuk masing masing rasio data, hal tersebut dipengaruhi karena dataset yang *inconsisten* serta nilai 0 pada musim penghujan terlampau banyak dan berjarak dengan nilai lain, sehingga mempengaruhi nilai prediksi. Pada nilai RMSE, semakin kecil nilai yang didapat maka semakin akurat data yang dihasilkan atau data yang di prediksi (S. Makridakis, 1995).

Hasil terbaik dalam permodelan ANFIS didapatkan nilai RMSE sebesar 12,4732 dengan rasio data 80:20 dengan normalisasi. Hasil tersebut masih kurang baik jika dibandingkan dengan penelitian sebelumnya yang dilakukan oleh (Anggraini, 2018) dengan nilai RMSE sebesar 0,734, penelitian oleh (Azhar & Mahmudy, 2018) dengan nilai RMSE sebesar 1,88 dan penelitian oleh (Sinaga dkk., 2019) dengan nilai RMSE terbaik sebesar 0,020. Tentunya pada penelitian lain menggunakan jumlah dataset dan teknik pengolahan data yang berbeda serta dengan skenario yang berberda pula. Mengingat dataset yang digunakan merupakan dataset yang *inconsisten* dengan terdapat banyak nilai 0 yang dikarenakan data pada musim kemarau, hasil akurasi yang didapat kurang maksimal atau kurang baik jika dibandingkan dengan penelitian sebelumnya.

4.5. Model ANFIS – GA

Sedikit berbeda dengan model ANFIS pada pengujian sebelumnya, pada tahapan ini metode ANFIS akan dilakukan proses optimasi terlebih dahulu

menggunakan Algoritma Genetika. Proses optimasi disini yang dimaksud ialah algoritma genetika akan menggantikan proses pada algoritma *steepest descent* sehingga dalam tahapan ini menggabungkan algoritma genetika dengan *LSE Recursive*. Dalam penelitian (Damayanti & Cahyadi, 2016) dilakukan hal yang serupa dengan tujuan algoritma genetika dapat menjadi solusi dalam mengurangi jumlah aturan samar menjadi sebanyak fungsi keanggotaan pada masukan atau lebih sering disebut permasalahan *curse of dimensionality*.

Secara teknis, tahapan ini akan dimulai dengan pengolahan dataset menggunakan algoritma genetika. Pemrosesan dengan algoritma genetika akan menghasilkan parameter premis dan konsekuen untuk selanjutnya dapat dilanjutkan pengolahan dengan metode ANFIS.

4.5.1. Pemilihan Operator Genetika

Operator genetika merupakan suatu hal penting dalam algoritma genetika. Operator genetika meliputi *crossover*, mutasi, model seleksi, maksimum generasi dan *offspring*. Jenis *crossover* yang digunakan dalam pengujian ini yaitu *crossover* menengah. Kemudian untuk jenis mutasi yang digunakan yaitu jenis mutasi seragam (*uniform*). Pemilihan jenis *crossover* dan jenis mutasi disesuaikan dengan representasi kromosom yang berupa bilangan real. Hal serupa dilakukan oleh (Damayanti & Cahyadi, 2016) yang menggunakan jenis *crossover* dan jenis mutasi yang sama dikarenakan dataset yang digunakan merupakan bilangan real. Kemudian pada model seleksi dalam pengujian ini digunakan model seleksi *Tournament Selection*.

4.5.2. Proses Pelatihan dan Pengujian ANFIS - GA

Algoritma genetika mencari solusi dengan cara berevolusi dari generasi ke generasi. Pada tahap awal dibangkitkan populasi awal secara acak dalam rentang nilai yang telah ditetapkan. Kemudian dilakukan proses crossover dan mutasi serta seleksi untuk menghasilkan kromosom-kromosom yang mempunyai nilai fitness yang lebih tinggi dibandingkan generasi sebelumnya. Agar generasi yang baik tidak tereliminasi pada saat proses seleksi yang dilakukan secara acak, maka dilakukan proses elitism yaitu dengan mengkopikan kromosom terbaik ke generasi berikutnya. Proses genetika ini dilakukan terus sampai kriteria penghentian iterasi terpenuhi, yang biasanya berupa jumlah maksimum generasi yang telah ditetapkan sebelumnya. Pada tahapan ini pengujian dilakukan terhadap kombinasi parameter-parameter yang akan digunakan dalam proses pelatihan, yakni data pelatihan dan data uji serta banyaknya fungsi keanggotaan dan parameter genetik.

- Jumlah fungsi keanggotaan - 3,4,5
- Probabilitas crossover - 0.2
- Probabilitas mutasi - 0.1
- generasi - 100, 150, 200, 250
- Offspring - 10
- Skema dataset - 80:20, 70:30

Hasil dari pelatihan dan pengujian data menggunakan parameter yang telah ditentukan dapat dilihat pada tabel 4.11 hingga tabel 4.14.

Tabel 4.11. Pengujian skema 80:20

Jumlah Data	MF	Probabilitas Crossover	Mutation Rate	Maksimum Generasi	RMSE
3045	3	0,2	0,1	100	5,074
3045	4	0,2	0,1	100	6,144
3045	5	0,2	0,1	100	7,549

Hasil pengujian pada tabel 4.11. menunjukkan bahwa untuk skema data 80:20 nilai RMSE yang paling baik ditunjukkan pada *membership function* 3 dimana mendapatkan nilai RMSE 5,074. Nilai tersebut lebih rendah daripada 2 *membership function* lainnya dimana masing – masing menghasilkan nilai RMSE 6,144 dan 7,549.

Tabel 4.12. Pengujian skema 70:30

Jumlah Data	MF	Probabilitas Crossover	Mutation Rate	Maksimum Generasi	RMSE
3045	3	0,2	0,1	100	4,801
3045	4	0,2	0,1	100	5,720
3045	5	0,2	0,1	100	7,097

Hasil pengujian pada tabel 4.12. menunjukkan bahwa untuk skema data 70:30 nilai RMSE yang paling baik ditunjukkan pada *membership function* 3 dimana mendapatkan nilai RMSE 4,801. Sama halnya dengan skema pengujian 80:20, pada skema 70:30 *membership function* 3 mendapatkan nilai RMSE terbaik dibanding dengan yang lain. Namun pada skema 70:30, nilai RMSE yang dihasilkan lebih rendah dibanding dengan skema 80:20. Hal ini menunjukkan bahwa skema data yang digunakan berpengaruh terhadap proses genetika dalam menghasilkan parameter premis dan parameter konsekuen yang nantinya akan digunakan dalam model ANFIS.

Tabel 4.13. Pengujian skema 80:20 dengan normalisasi

Jumlah Data	MF	Probabilitas Crossover	Mutation Rate	Maksimum Generasi	RMSE
3045	3	0,2	0,1	100	2,046
3045	4	0,2	0,1	100	3,591
3045	5	0,2	0,1	100	4,131

Tabel 4.13 merupakan hasil pengujian dengan skema 80:20 dan dengan proses normalisasi pada dataset. Dari hasil tersebut, nilai akurasi terbaik dihasilkan pada *membership function* 4 dengan nilai RMSE 2,046. Hasil tersebut lebih baik dari skema yang sama namun tanpa proses normalisasi pada dataset yang terdapat pada tabel 4.11. sehingga dari kedua hasil ini menunjukkan bahwa proses normalisasi data memberikan pengaruh terhadap hasil akurasi seperti yang terjadi pada model ANFIS tanpa algoritma genetika. Dari hasil ini semakin memperkuat pendapat pada penelitian oleh (Sinaga dkk., 2019) yang mengatakan bahwa proses normalisasi memberikan efek pada akurasi yang dihasilkan.

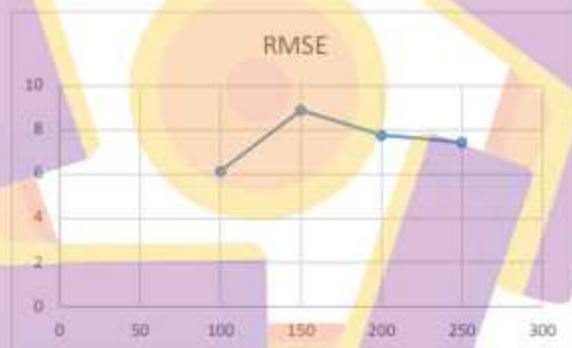
Tabel 4.14. Pengujian skema 70:30 dengan normalisasi

Jumlah Data	MF	Probabilitas Crossover	Mutation Rate	Maksimum Generasi	RMSE
3045	3	0,2	0,1	100	4,509
3045	4	0,2	0,1	100	3,350
3045	5	0,2	0,1	100	4,751

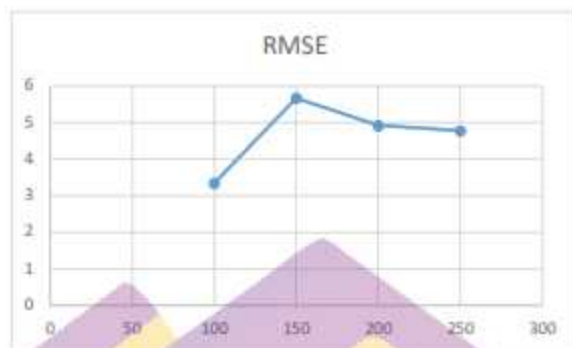
Tabel 4.14 merupakan hasil pengujian dengan skema dataset 70:30 dengan proses normalisasi data. Sama halnya dengan hasil pengujian lain dalam model ANFIS-GA ini, nilai RMSE terbaik didapatkan oleh *membership function* 4 dengan nilai capaian 3,350. Capaian ini lebih baik dari skema yang sama namun tanpa normalisasi data pada tabel 4.12. dengan nilai RMSE terbaik yang dihasilkan

sebesar 4,801. Sama halnya dengan capaian pada tabel 4.13, capaian yang didapat ini sangat dipengaruhi oleh proses normalisasi data.

Dari tabel 4.11 hingga 4.14 ditampilkan hasil dari pengujian dataset dengan dan tanpa skema normalisasi data. Pemilihan parameter – parameter genetik sangat berpengaruh terhadap jalannya program. Pemilihan parameter-parameter tersebut dilakukan dengan metode *trial and error*. Pada gambar 4.13, dan 4.14, memperlihatkan hasil efek pengaturan efek pemilihan generasi terhadap akurasi yang didapat. Tentunya efek dari pemilihan generasi juga di pengaruhi oleh parameter parameter lainnya.



Gambar 4.13. Pengujian jumlah generasi untuk mf 4



Gambar 4.14. Pengujian jumlah generasi dengan normalisasi untuk mf 4

4.5.3. Analisa Perhitungan Akurasi

Berdasarkan pengujian data pada bab sebelumnya, dan dapat dilihat dari tabel 4.11. hingga tabel 4.14, dapat disimpulkan selain skema dataset, pemilihan parameter – parameter genetika sangat berpengaruh terhadap akurasi yang di peroleh. Parameter seperti jumlah generasi, jumlah MF, dan *mutation rate* sangat berpengaruh dalam pengujian ANFIS – GA. Dari pengujian yang di lakukan nilai RMSE pada banyak fungsi keanggotaan tiga sampai lima, untuk set parameter sama, error untuk fungsi keanggotaan berjumlah 4 terlihat lebih rendah, baik pada skema data 80:30 atau 70:30 dan skema dataset dengan dan tanpa normalisasi data. Hal tersebut disebabkan oleh karakteristik data dan setting parameter yang dilakukan dalam proses pelatihan. Dalam pengujian dengan metode ANFIS yang di optimasi dengan algoritma genetika, parameter – parameter genetika yang menghasilkan nilai RMSE terendah dihasilkan dari jumlah MF – 3, *mutation rate*

- 0,1, jumlah generasi - 100, dan skema data 80:20 dengan proses normalisasi. Dari parameter - parameter tersebut, menghasilkan nilai RMSE sebesar 2,046.

Pada model ANFIS-GA nilai RMSE terbaik didapatkan sebesar 2,046 dengan skema dataset 80:20 dan dengan proses normalisasi. Performa yang didapat masih kurang bagus dibandingkan dengan penelitian yang dilakukan oleh (Chen dkk., 2017) dimana dalam penelitian tersebut model ANFIS-GA mendapat nilai RMSE sebesar 0,036 dan penelitian oleh (Damayanti & Cahyadi, 2016) ANFIS-GA mampu mendapat performa yang lebih baik dengan nilai RMSE yang didapat sebesar 0,00298. Tentunya dalam perhitungan suatu model dipengaruhi banyak hal seperti parameter - parameter genetika yang digunakan.

4.6. Pembahasan Hasil

Seperti yang sudah diuraikan pada masing masing pengujian, seluruh hasil pengujian menggunakan metode ANFIS dan ANFIS-GA dijabarkan secara detail pada tabel 4.15. Dari tabel tersebut dapat diketahui bahwa performa ANFIS dengan optimasi menggunakan Algoritma Genetika mampu menghasilkan nilai akurasi yang lebih baik dari metode ANFIS tanpa optimasi. Selain itu, proses normalisasi data pada dataset juga sangat berpengaruh. Dapat dilihat pada skema data 80:20 untuk metode anfis, dataset dengan normalisasi mendapatkan nilai RMSE sebesar 12,4732 sedangkan untuk dataset tanpa normalisasi mendapatkan nilai RMSE 15,7789. Pada metode ANFIS-GA hal serupa terjadi dimana dataset dengan normalisasi mendapat nilai RMSE sebesar 2,046 dan dataset tanpa normalisasi mendapat nilai RMSE sebesar 5,074 untuk *membership function* 3.

Penggunaan algoritma genetika untuk menggantikan algoritma *steepest descent* dinilai berhasil. Hal tersebut ditunjukkan dengan nilai akurasi yang didapat metode ANFIS-GA mampu lebih baik dari metode ANFIS. Tentunya pemilihan parameter – parameter juga memiliki efek yang signifikan dalam proses pengujian algoritma genetika. Kriteria optimasi yang bagus yaitu bagaimana fungsi – fungsi objektif pada algoritma yang digunakan menggambarkan bagus tidaknya suatu konfigurasi yang diberikan.

Tabel 4.15. Hasil seluruh uji yang dilakukan

metode	Skema	normalisasi	Epoch / generasi	RMSE
ANFIS	80:20	-	100	15,7789
	70:30	-	100	15,4345
	80:20	ya	100	12,4732
	70:30	ya	100	13,0982
ANFIS - GA	80:20	-	100	5,074
	70:30	-	100	4,801
	80:20	Ya	100	2,046
	70:30	ya	100	3,350

Dari seluruh hasil uji yang dilakukan peran dari normalisasi data cukup memberikan efek, dimana normalisasi data mampu memberikan efek yang baik bagi hasil prediksi pada semua model baik dalam model ANFIS dan ANFIS – GA. Sebagai perbandingan pada model ANFIS dengan rasio data 80:20 tanpa menggunakan normalisasi data, nilai RMSE yang didapat sebesar 15,7789, namun ketika menggunakan proses normalisasi data, nilai RMSE yang didapat menjadi lebih baik sebesar 12,4732.

Sama halnya dengan model ANFIS – GA dimana pada rasio data 80:20 tanpa normalisasi data, nilai RMSE yang didapat sebesar 5,074, namun ketika

menggunakan proses normalisasi data, nilai RMSE menjadi 2,046. Namun pada model ANFIS – GA untuk rasio data 70:30 penggunaan normalisasi data memang berpengaruh, tetapi tidak begitu signifikan seperti pada rasio data 80:20. Hal tersebut terjadi karena pada proses genetika terdapat suatu tahapan dimana nilai awal dibangkitkan secara acak, sehingga pada tiap – tiap percobaan akan menghasilkan nilai yang bervariasi.

Penggunaan algoritma genetika untuk mendapatkan parameter premis dan konsekuen terbaik berdasarkan penjabaran hasil pada tabel 4.15. mempengaruhi performa dari model ANFIS. Pada skema pengujian 80:20 dengan normalisasi, peningkatan performa antara model ANFIS dan ANFIS-GA cukup signifikan, dimana model ANFIS menghasilkan performa akurasi dengan nilai RMSE 12,4732 sedangkan ANFIS – GA mendapatkan nilai RMSE 2,046 hal tersebut dipengaruhi peranan dari normalisasi data. Lain halnya ketika tanpa menggunakan normalisasi data, dimana perbedaan performa cukup signifikan dimana model ANFIS mendapat nilai RMSE 15,7789 dan untuk ANFIS – GA mendapat nilai RMSE 5,074.

Tabel 4.16. Perbandingan model ANFIS dengan penelitian lain

Peneliti	Jumlah dataset	Epoch	RMSE
(Anggraini, 2018)	1827	100	0,73
(Sinaga dkk., 2019)	60	1000	0,02
(Azhar & Mahmudy, 2018)	180	400	1,88
**	3045	100	12,4732

Tabel 4.16. merupakan tabel perbandingan model ANFIS, pada baris berwarna hijau merupakan hasil model ANFIS dalam penelitian ini. Dari

perbandingan pada tabel diatas, akurasi yang didapat kurang baik jika dibandingkan dengan beberapa penelitian sebelumnya. Hasil terbaik dalam penelitian ini pada model ANFIS masih kurang baik jika dibandingkan dengan penelitian sebelumnya yang dilakukan oleh (Angraini, 2018) dengan nilai RMSE sebesar 0,734, penelitian oleh (Azhar & Mahmudy, 2018) dengan nilai RMSE sebesar 1,88 dan penelitian oleh (Sinaga dkk., 2019) dengan nilai RMSE terbaik sebesar 0,020. Tentunya pada penelitian lain menggunakan jumlah dataset yang berbeda serta dengan skenario yang berbeda seperti yang tertera pada tabel 4.16.

Tabel 4.17. Perbandingan optimasi ANFIS

Peneliti	metode	RMSE
(Damayanti & Cahyadi, 2016)	ANFIS-GA	0,0028
(Chen dkk., 2017)	ANFIS-GA	0,036
	ANFIS-DE	0,035
	ANFIS-PSO	0,040
(Aziz dkk., 2017)	ANFIS-PSO	0,267
**	ANFIS-GA	2,046

Tabel 4.17. merupakan tabel perbandingan optimasi model ANFIS dengan penelitian lain. Dalam perbandingan tersebut terdapat beberapa macam algoritma optimasi yang digunakan. Metode yang digunakan meliputi algoritma genetika (GA), *particle swarm optimization* (PSO), dan *differential evolution* (DE). Dari beberapa penelitian diatas hasil yang dicapai oleh algoritma genetika pada penelitian ini mendapatkan performa hampir mendekati penelitian yang dilakukan oleh (Chen dkk., 2017) dimana dalam penelitian tersebut model ANFIS-GA mendapat nilai RMSE sebesar 0,036. Namun dalam penelitian oleh (Damayanti &

Cahyadi, 2016) ANFIS-GA mampu mendapat performa yang lebih baik dengan nilai RMSE yang didapat sebesar 0,00298.

Dalam hal algoritma untuk optimasi, algoritma genetika yang digunakan dalam penelitian ini mempunyai performa yang hampir mendekati dengan algoritma DE dan PSO dalam penelitian (Chen dkk., 2017) dimana masing – masing algoritma tersebut mendapatkan nilai RMSE sebesar 0,036 dan 0,040 serta pada penelitian (Aziz dkk., 2017) dimana dalam penelitian tersebut ANFIS – PSO mendapat nilai RMSE sebesar 0,267.

Berdasarkan penjabaran diatas, penggunaan algoritma genetika dalam penelitian ini untuk menggantikan proses dari pembelajaran arah mundur yang sebelumnya diperankan oleh algoritma *steepest descent*, mampu mempengaruhi performa dari model ANFIS, hal tersebut dibuktikan dengan hasil akurasi yang didapatkan seperti yang tercantum pada tabel 4.15. Selain itu penerapan normalisasi data sebelum proses pengujian juga mampu memberikan hasil yang baik dan mampu meningkatkan akurasi seperti halnya penelitian yang dilakukan oleh (Angraini, 2018) dan (Sinaga dkk., 2019).

BAB V

PENUTUP

5.1. Kesimpulan

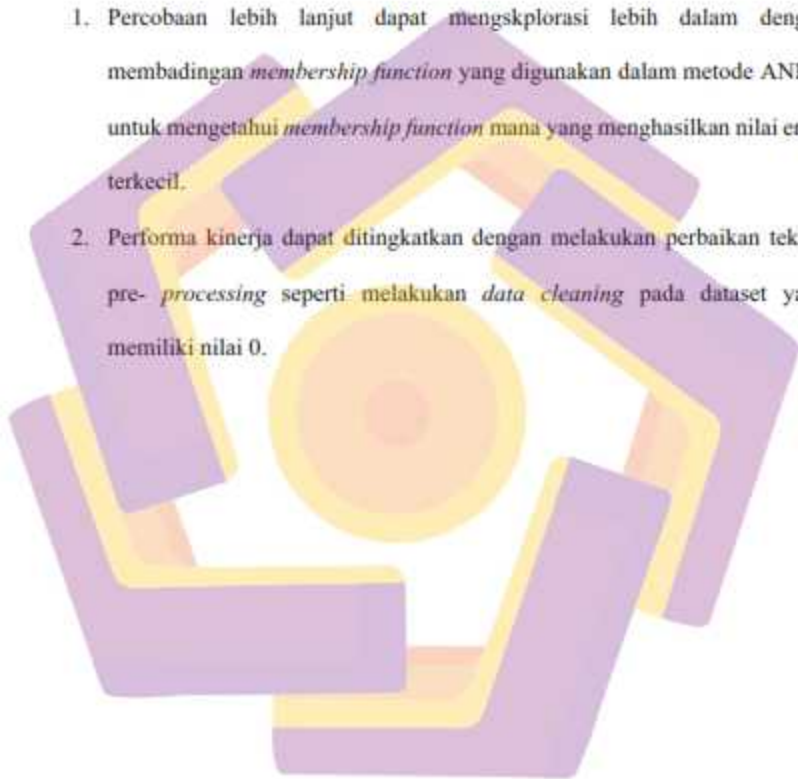
Berdasarkan hasil yang telah dicapai dalam penelitian mengenai optimasi struktur ANFIS dengan Algoritma Genetika dapat disimpulkan beberapa hal sebagai berikut:

1. Algoritma Genetika mampu menggantikan peran yang sebelumnya dijalankan oleh algoritma *steepest descent*. Hal tersebut dibuktikan bahwa penggunaan algoritma genetika mampu memberikan efek berupa penurunan nilai error yang didapat. Nilai error terkecil yang didapat yaitu sebesar 2,046 dengan dataset yang telah dilakukan normalisasi data.
2. Performa algoritma genetika yang baik tidak lepas dari pemilihan parameter – parameter genetik sebelum proses dilakukan. Parameter – parameter genetik yang mampu menghasilkan nilai error 2,046 yang diukur dengan menggunakan RMSE yaitu dengan parameter jumlah MF = 3, mutation rate = 0,1, jumlah generasi = 100.
3. Proses normalisasi data sangat berpengaruh pada hasil pengujian. Hal tersebut dibuktikan dengan pengujian yang dilakukan menggunakan metode ANFIS dengan dataset yang telah dinormalisasi mendapat nilai RMSE sebesar 12,4732 dan 13,0982 untuk rasio data 80:20 dan 70:30. Sedangkan nilai RMSE untuk dataset yang tanpa normalisasi sebesar 15,778 dan 15,434 untuk rasio data 80:20 dan 70:30

5.2. Saran

Ada beberapa saran yang direkomendasikan untuk penelitian selanjutnya pada topik penelitian yang sama, antara lain:

1. Percobaan lebih lanjut dapat mengskplorasi lebih dalam dengan membandingkan *membership function* yang digunakan dalam metode ANFIS untuk mengetahui *membership function* mana yang menghasilkan nilai error terkecil.
2. Performa kinerja dapat ditingkatkan dengan melakukan perbaikan teknik *pre-processing* seperti melakukan *data cleaning* pada dataset yang memiliki nilai 0.



DAFTAR PUSTAKA

PUSTAKA BUKU

- Cryer, J. D., & Chan, K.-S. (2008). *Time Series Analysis*. Springer New York. <https://doi.org/10.1007/978-0-387-75959-3>
- Gen, M., & Cheng. (1997). *Genetics Algorithm and Engineering Optimization* (J. Willey & Son, Ed.).
- Goldberg, D. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley.
- Jang, S. R., Sun, C. T., & Mizutani, E. (1997). *Neuro-Fuzzy And Soft Computing* (P. Hall, Ed.).
- Kusumadewi S, & Hartati S. (2010). *"Neuro-Fuzzy "Integrasi Sistem Fuzzy & Jaringan Syaraf* (2 ed.). Graha Ilmu.
- M.E Render B., S. Jr. R. M. dan H. (2003). *Quantitative Analysis for Management* (P. Education, Ed.; 8th ed.). Inc.
- Piegel L.J, R Murray, & Stephens. (2007). *STATISTIK Schaum's Outlines* (3 ed.). Erlangga.
- R.K Buffa, E. S. & S. (2007). *Modern Production / Operation Management* (J. Wiley & Sons, Ed.; 8th Ed). Inc.
- Rosebrock, A. (2018). Face recognition with OpenCV, Python, and deep learning. Dalam *Pyimagesearch*.
- s. kusumadewi, s. hartati. (2010). *"Neuro-Fuzzy "Integrasi Sistem Fuzzy & Jaringan Syaraf* (2 ed.). Graha Ilmu.
- S. Kusumadewi and H. Purnomo. (2010). *Aplikasi Logika Fuzzy Untuk Pendukung Keputusan* (2 ed.). Graha Ilmu.
- S. Makridakis, S. C. W. and V. E. M. (1995). *Metode dan Aplikasi Peramalan* (2 ed.). Erlangga.
- Stevenson, W. J., & Chee Chuong, S. (2014). *Manajemen Operasi Perspektif Asia* (9 ed., Vol. 2). Salemba Empat.
- T. S. Widodo. (2005). *Istem Neuro Fuzzy Untuk Pengolahan Informasi, Pemodelan, dan Kendali*. Graha Ilmu.

V. Gaspresz. (2005). *Sistem Manajemen Kinerja Terintegrasi Balanced Scorecard dengan Six Sigma untuk Organisasi Bisnis dan Pemerintah*. Gramedia pustaka Utama.

Vapnik, V. N. (2000). The Nature of Statistical Learning Theory. Dalam *The Nature of Statistical Learning Theory*. <https://doi.org/10.1007/978-1-4757-3264-1>

Wei, W. S. (2006). *Time Analysis Univariate and Multivariate Methods*. Addison Wesley Publishing Company, Inc.

PUSTAKA JURNAL ILMIAH ATAU PROSIDING

Angraini, L. (2018). ANFIS DENGAN MEMBERSHIP FUNCTION UNTUK PREDIKSI CURAH HUJAN PADA DATA RENTET WAKTU MULTIVARIATE Lilis Angraini Fakultas Teknologi Informasi Universitas Islam Kalimantan Muhammad Arsyad Al Banjari Banjarmasin Email : lilis.angraini@gmail.com. *Technologia*, 9(1), 18–25.

Azhar, M. I., & Mahmudy, W. F. (2018). Prediksi Curah Hujan Menggunakan Metode Adaptive Neuro Fuzzy Inference System (ANFIS). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2(11), 4932–4939.

Aziz, M. A. el, Hemdan, A. M., Ewees, A. A., Elhoseny, M., Shehab, A., Hassanien, A. E., & Xiong, S. (2017). Prediction of biochar yield using adaptive neuro-fuzzy inference system with particle swarm optimization. *Proceedings - 2017 IEEE PES-IAS PowerAfrica Conference: Harnessing Energy, Information and Communications Technology (ICT) for Affordable Electrification of Africa, PowerAfrica 2017*, 115–120. <https://doi.org/10.1109/PowerAfrica.2017.7991209>

Balasubramanian, K., & Ananthamoorthy, N. P. (2021). Improved adaptive neuro-fuzzy inference system based on modified glowworm swarm and differential evolution optimization algorithm for medical diagnosis. *Neural Computing and Applications*, 33(13), 7649–7660. <https://doi.org/10.1007/s00521-020-05507-0>

Chang, F. J., Tseng, K. Y., & Chaves, P. (2007). Shared near neighbours neural network model: A debris flow warning system. *Hydrological Processes*, 21(14). <https://doi.org/10.1002/hyp.6489>

Chen, W., Panahi, M., & Pourghasemi, H. R. (2017). Performance evaluation of GIS-based new ensemble data mining techniques of adaptive neuro-fuzzy inference system (ANFIS) with genetic algorithm (GA), differential evolution (DE), and particle swarm optimization (PSO) for landslide spatial modelling. *Catena*, 157(February), 310–324. <https://doi.org/10.1016/j.catena.2017.05.034>

- Chia-Hoang Lee, Jeng-Ren Hwang, & Shyi-Ming Chen. (1997). *handling forecasting problems using fuzzy time series* (hlm. 217–228). Elsevier B.V.
- Coussement, K., & van den Poel, D. (2006). Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques. *Expert Systems with Applications*, *34*, 313–327. <https://doi.org/10.1016/j.eswa.2006.09.038>
- Damayanti, A., & Cahyadi, R. (2016). Algoritma Genetika Untuk Optimasi Struktur Anfis. *Seminar Riset Teknologi Informasi (SRITI) Tahun*, 1–7.
- Desmonda, D., Tursina, T., & Irwansyah, M. A. (2018). Prediksi Besaran Curah Hujan Menggunakan Metode Fuzzy Time Series. *Jurnal Sistem Dan Teknologi Informast (JUSTIN)*, *6*(4), 141. <https://doi.org/10.26418/justin.v6i4.27036>
- Díaz-Robles, L. A., Ortega, J. C., Fu, J. S., Reed, G. D., Chow, J. C., Watson, J. G., & Moncada-Herrera, J. A. (2008). A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: The case of Temuco, Chile. *Atmospheric Environment*, *42*(35). <https://doi.org/10.1016/j.atmosenv.2008.07.020>
- Elbaz, K., Shen, S. L., Zhou, A., Yuan, D. J., & Xu, Y. S. (2019). Optimization of EPB shield performance with adaptive neuro-fuzzy inference system and genetic algorithm. *Applied Sciences (Switzerland)*, *9*(4). <https://doi.org/10.3390/app9040780>
- Fauzan, M. A., & Setiawan, B. D. (2019). Algoritma Genetika Untuk Optimasi Fuzzy Time Series Dalam Memprediksi Debit Air (Studi Kasus: PDAM Indramayu). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, *3*(1), 831–835. <http://j-ptiik.ub.ac.id>
- Hadiansyah, F. N. (2017). Prediksi Harga Cabai dengan Menggunakan pemodelan Time Series ARIMA. *Indonesian Journal on Computing (Indo-JC)*, *2*(1), 71. <https://doi.org/10.21108/indojc.2017.2.1.144>
- Halimi, I., & Kusuma, W. A. (2018). Prediksi Indeks Harga Saham Gabungan (IHSG) Menggunakan Algoritma Neural Network. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, *4*(1), 24. <https://doi.org/10.26418/jp.v4i1.25384>
- Hillmer, S. C., & Wei, W. W. S. (1991). Time Series Analysis: Univariate and Multivariate Methods. *Journal of the American Statistical Association*, *86*(413), 245. <https://doi.org/10.2307/2289741>

- Ighose, B. O., Adeleke, I. A., Damos, M., Junaid, H. A., Okpalaek, K. E., & Betiku, E. (2017). Optimization of biodiesel production from *Thevetia peruviana* seed oil by adaptive neuro-fuzzy inference system coupled with genetic algorithm and response surface methodology. *Energy Conversion and Management*, 132, 231–240. <https://doi.org/10.1016/j.enconman.2016.11.030>
- Jiang, P., Yang, H., & Heng, J. (2019). A hybrid forecasting system based on fuzzy time series and multi-objective optimization for wind speed forecasting. *Applied Energy*, 235(November 2018), 786–801. <https://doi.org/10.1016/j.apenergy.2018.11.012>
- Lee, M. H., & Suhartono. (2012). A weighted fuzzy time series model for forecasting seasonal data. *Journal of Quality Measurement and Analysis (JQMA)*, 8(1), 85–95.
- Liu, F., Pei, R., & Guan, X. (2002). *new algorithm of fuzzy on-line identification for nonlinear systems* . . . 103–106.
- Lutfy, O. F., Noor, S. B. M., & Marhaban, M. H. (2011). A simplified adaptive neuro-fuzzy inference system (ANFIS) controller trained by genetic algorithm to control nonlinear multi-input multi-output systems. *Scientific Research and Essays*, 6(31), 6475–6486. <https://doi.org/10.5897/sre11.1059>
- Nikentari, N., Bettiza, M., & Sastypratiwi, H. (2018). Prediksi Kecepatan Angin Menggunakan Adaptive Neuro Fuzzy (ANFIS) dan Radial Basis Function Neural Network (RBFNN). *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 4(1), 70. <https://doi.org/10.26418/jp.v4i1.25558>
- Nilakant, K., & Mitrovic, A. (2005). Applications of Data Mining in Constraint-Based Intelligent Tutoring Systems. *Proceedings of the Artificial Intelligence in Education, AIED, December*, 896–898.
- Noor, M. F., Wardhana, A. K., Utami, E., & Pramono, E. (2019). Prediksi curah hujan menggunakan metode anfis (studi kasus: kabupaten hulu sungai utara). *Jurnal INFORMA Politeknik Indonusa Surakarta*, 5(2), 24–29. <http://www.poltekindonusa.ac.id/SUB-DOMAIN/informa/index.php/informa/article/view/77>
- Pratama, R. P., & Tjahyanto, A. (2021). The influence of fake accounts on sentiment analysis related to COVID-19 in Indonesia. *Procedia Computer Science*, 197. <https://doi.org/10.1016/j.procs.2021.12.128>
- Rahmi, A., & Mahmudy, W. F. (2016). Pembentukan Model Regresi Harga Saham Menggunakan Algoritma Genetika. *Seminar Nasional Teknologi Informasi Dan Komunikasi 2016 (SENTIKA 2016) Yogyakarta, 2016*(Sentika).

- Rifa'i, A. (2020). Optimasi Fuzzy Artificial Neural Network dengan Algoritma Genetika untuk Prediksi Harga Crude Palm Oil. *Jurnal Teknik Informatika Dan Sistem Informasi*, 6(2), 234–241. <https://doi.org/10.28932/jutisi.v6i2.2617>
- Sanay, R., Voulgaris, G., & Warner, J. C. (2007). Tidal asymmetry and residual circulation over linear sandbanks and their implication on sediment transport: A process-oriented numerical study. *Journal of Geophysical Research: Oceans*, 112(12). <https://doi.org/10.1029/2007JC004101>
- Savrun, M. M., & İnci, M. (2021). Adaptive neuro-fuzzy inference system combined with genetic algorithm to improve power extraction capability in fuel cell applications. *Journal of Cleaner Production*, 299. <https://doi.org/10.1016/j.jclepro.2021.126944>
- Sinaga, C. B., Havaluddin, H., Pakpahan, H. S., Prafanto, A., & Setyadi, H. J. (2019). Peramalan Curah Hujan Dengan Pendekatan Adaptive Neuro Fuzzy Inference System. *Sains, Aplikasi, Komputasi Dan Teknologi Informasi*, 1(2), 1. <https://doi.org/10.30872/jsakti.v1i2.2599>
- Subanar & Suhartono. (2009). Wavelet Neural Networks untuk Peramalan Data Time Series Finansial. *Program Penelitian Ilmu Dasar Perguruan Tinggi, FMIPA UGM, Yogyakarta*.
- Wahyuni, I., Mahmudy, W. F., & Iriany, A. (2017). Rainfall prediction using hybrid adaptive neuro-fuzzy inference system (ANFIS) and genetic algorithm. *Journal of Telecommunication, Electronic and Computer Engineering*, 9(2–8), 51–56.
- Zhang, R., & Tao, J. (2018). A Nonlinear Fuzzy Neural Network Modeling Approach Using an Improved Genetic Algorithm. *IEEE Transactions on Industrial Electronics*, 65(7), 5882–5892. <https://doi.org/10.1109/TIE.2017.2777415>