ASSESSING THE PERFORMANCE OF ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM FOR COMPLEX NONLINEAR SYSTEMS USING MULTIPLE MEMBERSHIP FUNCTIONS

Thesis

By

MEILIANA ZALIANTI
NPM. 2117031071

FACULTY OF MATHEMATICS AND NATURAL SCIENCES
UNIVERSITY OF LAMPUNG
BANDAR LAMPUNG
2025

ABSTRACT

ASSESSING THE PERFORMANCE OF ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM FOR COMPLEX NONLINEAR SYSTEMS USING MULTIPLE MEMBERSHIP FUNCTIONS

By

Meiliana Zalianti

ANFIS is a form of artificial intelligence that integrates neural networks, fuzzy logic, and inference systems to develop intelligent decision-making frameworks. It can be applied to various tasks including classification, regression, clustering, and control. One of the benefits of ANFIS is its capability to manage complex and uncertain data while also learning from past experiences and adjusting to evolving conditions. The purpose of this research is to apply ANFIS to predict cooling load using three different membership functions namely sigmoidal, gaussian, and generalized bell. Furthermore, estimating the ANFIS model and obtaining the accuracy level of the cooling load prediction model with the ANFIS method. The predictor variables used are Relative Compactness (RC), Surface Area (SA), Wall Area (WA), Roof Area (RA), Overall Height (OH), Orientation (OR), Glazing Area (GA), and Glazing Area Distribution (GAD), while the response variable used is Cooling Load (CL). The best model is selected based on the RMSE value. The results of the analysis show that the use of the ANFIS method is effective for data prediction because the prediction results are quite close to the actual data. The ANFIS model with generalized bell membership function provides the best level of accuracy to predict the cooling load with an RMSE value of 2.058887.

Keywords: Adaptive Neuro Fuzzy Inference System, Artificial Neural Network, Fuzzy Logic, Backpropagation Algorithms, Prediction.

ASSESSING THE PERFORMANCE OF ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM FOR COMPLEX NONLINEAR SYSTEMS USING MULTIPLE MEMBERSHIP FUNCTIONS

MEILIANA ZALIANTI

Thesis

Submitted as a Partial Fulfillment of the Requirements for the Degree of
BACHELOR OF MATHEMATICS

In

Department of Mathematics

Faculty of Mathematics and Natural Science

FACULTY OF MATHEMATICS AND NATURAL SCIENCES
UNIVERSITY OF LAMPUNG
BANDAR LAMPUNG
2025

Thesis Title : **ASSESSING THE PERFORMANCE OF ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM FOR COMPLEX NONLINEAR SYSTEMS USING MULTIPLE MEMBERSHIP FUNCTIONS**

Student Name : ***Meiliana Zalianti***

Student Identification Number : **2117031071**

Study Program : **Mathematics**

Faculty : **Faculty of Mathematics and Natural Sciences**

APPROVED BY

1. Supervisory Committee

Dr. Khoirin Nisa, S.Si., M.Si.
NIP 197407262000032001

Bernadhita Herindri S. U., S.Si., M.Sc.
NIP 199206302023212034

2. Head of Mathematics Department

Dr.Aang Nuryaman, S.Si.,M.Si.
NIP. 197403162005011001

RATIFIED BY

1. Examination Committee

Head	:	**Dr. Khoirin Nisa, S.Si., M.Si.**	....
Secretary	:	**Bernadhita Herindri S. U., S.Si., M.Sc.**	....
Examiner Non-Supervisor	:	**Prof. Drs. Mustofa, M.A., Ph.D.**	....

2. Dean of the Faculty of Mathematics and Natural Science

Dr. Eng. Heri Satria, S.Si., M.Si.
NIP. 197110012005011002

Thesis Defense Date: **February 20, 2025**

STUDENT THESIS STATEMENT

The undersigned below:

Name	:	**Meiliana Zalianti**
Student Identification Number	:	**2117031071**
Department	:	**Mathematics**
Thesis Title	:	**Assessing the Performance of Adaptive Neuro-Fuzzy Inference System for Complex Nonlinear Systems Using Multiple Membership Functions**

Hereby declare that this thesis is my own work. If, at any time in the future, it is proven that this thesis is plagiarized or written by someone else, I am willing to accept any sanctions in accordance with the applicable academic regulations.

Bandar Lampung, 20 February 2025

Author,

Meiliana Zalianti

BIOGRAPHY

The author has the full name Meiliana Zalianti, born in Margototo, Metro Kibang District, East Lampung Regency on May 22, 2004, is the only child of Mr. Saipan and Mrs. Bonatin.

The author began her education at TK Margototo from 2009 to 2010, then continued to SD Negeri 3 Margototo from 2010 to 2016. She pursued her secondary education at SMP Negeri 1 Metro Kibang from 2016 to 2019 and completed her high school education at SMA Negeri 1 Metro in an accelerated class program from 2019 to 2021. Subsequently, she was admitted to the Mathematics Study Program, Faculty of Mathematics and Natural Sciences, Universitas Lampung through the SBMPTN selection pathway and pursued her undergraduate studies from 2021 to 2025.

During her university years, the author actively participated in various organizations. In 2023, she joined the Mathematics Student Association (HIMATIKA) in the Keilmuan division and also became a member of the Student Executive Board (BEM) of FMIPA Universitas Lampung in the Internal Youth Relations (HIK) division. Additionally, she undertook an internship at the National Research and Innovation Agency (BRIN) in Tanjung Bintang, South Lampung, from January to February 2024. From June to August 2024, she participated in the Community Service Program (KKN) in Sriwangi Village, Way Jepara District, East Lampung Regency.

INSPIRATIONAL QUOTE

"But they plan, and Allah plans. And Allah is the best of planners."

(QS Al-Anfal: 30)

"Indeed, Allah will not change the condition of a people until they change what is in themselves."

(QS Ar-Ra'd: 11)

"Allah will raise those who have believed among you and those who were given knowledge, by degrees."

(Q.S. Al-Mujadilah: 11)

"Whoever has never tasted the bitterness of seeking knowledge, even for a moment, will endure the humiliation of ignorance for a lifetime."

(Imam Ash-Shafi'i)

DEDICATION

Alhamdulillah, praise be upon Allah SWT, the Most Gracious and the Most Merciful, for His endless blessings, guidance, and strength throughout this journey. I dedicate my deepest thanks and appreciation to the following people:

My Beloved Parents

To my beloved parents, father (Saipan) and mother (Bonatin) who have been my greatest source of love, support, and inspiration. Thank you for your endless prayers, sacrifices, and encouragement that have shaped me into the person I am today.

Supervisors and Examiner

To my supervisors and examiner, for their invaluable guidance, patience, and constructive feedback that have helped me grow academically and personally.

My Dearest Friends

To my dearest friends, who have stood by my side through the highs and lows, offering unwavering support, laughter, and motivation. Your presence has made this journey even more meaningful.

My Beloved Almamater
University of Lampung

AKNOWLEDGEMENT

Alhamdulillah, all praise and gratitude be to Allah SWT for His abundant blessings and mercy, which have enabled me to complete this thesis, titled ”Assessing the Performance of Adaptive Neuro-Fuzzy Inference System for Complex Nonlinear Systems Using Multiple Membership Functions,” successfully and on time.

In the process of preparing this thesis, many parties have helped provide guidance, support, direction, motivation and advice so that this thesis can be completed. Therefore, on this occasion the author would like to thank:

1. Dr. Khoirin Nisa, S.Si., M.Si. as the first supervisor who has taken a lot of time to provide direction, guidance, motivation, advice and support to the author so that she can complete this thesis.

2. Bernadhita Herindri S.U, S.Si., M.Sc. as Supervisor II who has provided direction, guidance and support to the author so that she can complete this thesis.

3. Prof. Drs. Mustofa, M.A., Ph.D as an examiner who has been willing to provide criticism and suggestions and evaluation to the author so that it can be even better.

4. Dr. Aang Nuryaman, S.Si., M.Si. as the Head of Mathematics Department, Faculty of Mathematics and Natural Sciences, University of Lampung.

5. Dr. Agus Sutrisno, S.Si., M.Si. as academic advisor.

6. All lecturers, staff and employees of the Department of Mathematics, Faculty of Mathematics and Natural Sciences, University of Lampung.

7. My parents who always provide support, prayers, and motivation during the completion of this thesis. This achievement is as much theirs as it is mine.

8. My best friends Atun, Eva, Nabila, Maya, Rahma, Adinda, Sherina, Tasya, Dita, Fathan, Vedisya and Niluh who have helped, encouraged, and motivated me so that I can complete this thesis well.
9. Keilmuan 2021 friends, thank you for all the experiences and togetherness we have shared.

Hopefully this thesis can be useful for all of us. The author realizes that this thesis is far from perfect, so the author hopes for constructive criticism and suggestions to make this thesis even better.

Bandar Lampung, 20 February 2025

Meiliana Zalianti

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

Introduction

1.1 Background of the Problem

Increasing complexity and technological advances such as big data have driven rapid developments in complex systems science. Complex systems consist of many interacting components, usually in ways that are not linear or simple. Complex systems can emerge and evolve through a process of self-organization, such that the patterns in the system are not completely ordered but also not completely random. This process allows new behaviors to emerge that can be observed on a larger (macroscopic) scale than just looking at the components in isolation. In the real world, complexity can arise in a variety of contexts such as gene regulatory networks within a cell, physiological systems of an organism, brains and other neural systems, food webs, the global climate, stock markets, the internet, social media, national and international economies, and even human cultures and civilizations (Sayama, 2015).

The development of the concept of complex systems has made modeling an essential element in understanding, analyzing, and designing these systems (Buragohain and Mahanta, 2008). According to Sayama (2015), a model is a simplified representation of a system, which can be conceptual, verbal, diagrammatic, physical, or formal (mathematical). In complex systems, emergent behaviors often arise from interactions between components and modeling allows researchers to simulate these interactions. In addition, modeling helps in testing hypotheses, predicting future system behavior and designing effective optimization. With modeling, researchers can identify patterns and interrelationships in large and complex data, thereby creating new insights that are useful in decision-making. The interdisciplinary approach applied in modeling complex systems also allows for a more holistic understanding of these intricately dynamic systems (Buragohain and Mahanta, 2008).

The distinct characteristics of complex systems such as networks, nonlinearity, uncertainty, and self-organization make developing a model present significant challenges. Traditional modeling approaches using linear system tools are inadequate for these types of systems. However, with the widespread availability of computers, computational modeling and simulation have emerged as practical and effective methods for this task. We can create our models using general-purpose programming languages like Python, R, Mathematica, MATLAB, and others (Sayama, 2015).

Soft computing is one of the paradigms in computing that is very effective in handling complex and nonlinear problems. The idea of soft computing (SC), which combines fuzzy logic, neuro-computing, evolutionary computing, and probabilistic computing, originated from this. Fuzzy logic, first introduced by Prof. Lotfi A. Zadeh in 1965, works well under conditions of uncertainty. Meanwhile, artificial neural networks are widely known as methods that can learn from data and form robust models to predict system behavior. Therefore, a hybrid approach emerged that combines the advantages of artificial neural networks (ANN) and fuzzy inference systems (FIS) (Siddique and Adeli, 2013).

First proposed by Jang in 1993 and 1997, ANFIS (Adaptive Neuro-Fuzzy Inference System) is a class of adaptive networks functionally equivalent to fuzzy inference systems. Sugeno and Tsukamoto's fuzzy models are represented by the ANFIS architecture. The benefits of neural networks and fuzzy identification are combined in ANFIS. For extremely real nonlinear systems, ANFIS requires less computing power than neural networks. ANFIS is also capable of delving deeply into the issue while operating with less structural knowledge. Furthermore, rather than using real physical systems, ANFIS can be used to regulate the output of online forecasting systems (Hou et al., 2003).

In this research, the Anfis method will be used to predict the cooling load in buildings. The building thermal system is an example of a complex and uncertain real system. Buildings contribute as one of the largest sources of carbon dioxide (CO_2) emissions with a percentage reaching 40% (Yudelson, 2007). CO_2 is a major factor causing global warming, an issue that is currently a major concern around the world. According to Adiwoso and Prasetyoadi (2010), of the total energy available in the world, buildings consume 48%. This trend has been increasing in recent years as buildings are major energy consumers globally and living standards are improving. Understanding heating and cooling loads is essential during building construction to lower overall energy consumption.

Access to indoor cooling is critical to ensure health and safety in a rapidly warming world with more frequent and intense extreme heat events. Air conditioning allows us to work and learn effectively, keeps food and medical supplies safe, and reduces the risk of heat-related illnesses. But increased energy use needed to meet accelerating cooling demand can strain electricity grids, drive up emissions, and worsen urban heat islands (Adiwoso and Prasetyoadi, 2010).

The energy used to cool buildings currently relies heavily on fossil fuels. Without decarbonizing energy supply, more cooling therefore means higher carbon pollution and more warming. The growing demand for air conditioning can strain power grids. This can become dangerous, especially if a power grid fails during a heatwave. Chemicals used in air conditioning, such as hydrofluorocarbons (HFCs), are extremely powerful greenhouse gases with a global warming potential far greater than carbon dioxide. Global agreements in place aim to phase down HFCs in the decades ahead. Air conditioners emit waste heat back outside. In a city with millions of air conditioners running and releasing waste heat into the outside air, it can increase the urban heat island effect (Central, 2023).

To address these challenges, the Adaptive Neuro-Fuzzy Inference System (ANFIS) model is employed to simulate and predict energy efficiency in building cooling systems. By integrating both neural network learning capabilities and fuzzy logic reasoning, ANFIS provides an effective tool for modeling the nonlinear relationships between various factors influencing energy consumption and efficiency. This approach enables better optimization of cooling strategies, potentially reducing reliance on fossil fuels and mitigating environmental impacts such as greenhouse gas emissions and the urban heat island effect.

ANFIS has been successfully applied in a variety of application areas. Researchers from various fields have recently shown a strong preference for a combined neuro-fuzzy approach. Ziasabounchi and Askerzade (2014) studied the use of ANFIS in the medical field to categorize heart disease. With an accuracy of 92.30%, the model was able to accurately predict the patient’s heart disease. Furthermore, Savić et al. (2014) used ANFIS in the field of environmental science to model tropospheric ozone concentrations. The results of this study demonstrated that ANFIS modeling was highly accurate, with a prediction potential of over 80%, as indicated by the coefficient of determination (R^2) of 0.802. Another study also conducted by Liu et al. (2018) is predicting the unfrozen water content in saline soil using BPNN and ANFIS and shows that the ANFIS method has outperformed the BPNN method.

Against the background of these studies, the use of ANFIS to predict cooling load in buildings is highly relevant. The building energy load system includes several variables like Relative Compactness (RC), Surface Area (SA), Wall Area (WA), Roof Area (RA), Overall Height (OH), Orientation (OR), Glazing Area (GA), and Glazing Area Distribution (GAD) has complex and non-linear characteristics. This study aims to evaluate how well the ANFIS model comprehends and predicts the behavior of complex non-linear systems, especially in the context of energy efficiency or building thermal systems. The results of this modeling are also expected that the outcomes of this modeling will offer the best model for assisting in the design of more energy-efficient and ecologically friendly building structures.

1.2 Problem Statement

In conducting this research, the problem formulations put forward are as follows:

1. How does the performance of three different membership functions (Sigmoidal, Gaussian, and Generalized Bell) in the ANFIS model compare in predicting energy efficiency, especially in terms of RMSE value?
2. Which membership function gives the most accurate prediction results with the smallest RMSE in the ANFIS model for energy efficiency data?

1.3 Research Objectives

Based on this background, the objectives of this study are as follows:

1. Build Adaptive Neuro-Fuzzy Inference System (ANFIS) model to predict cooling load based on input variables such as Relative Compactness (RC), Surface Area (SA), Wall Area (WA), Roof Area (RA), Overall Height (OH), Orientation (OR), Glazing Area (GA), and Glazing Area Distribution (GAD).
2. Comparing three different membership functions (Sigmoidal, Gaussian, and Generalized Bell) in the ANFIS model to determine the membership function that produces the most accurate prediction based on the RMSE value.
3. Determining the best membership function in the ANFIS model that provides prediction results with the lowest RMSE.

1.4 Research Benefits

This research is expected to provide the following benefits:

1. Produce an accurate model in predicting cooling load in buildings, which can help design more energy efficient buildings.
2. Provide a deeper understanding of how different types of membership functions affect ANFIS performance, thus providing a richer reference for research in the field of soft computing and non-linear system modeling.
3. Provide insights to building designers and engineers to integrate energy efficiency principles in future building construction, in line with global efforts to reduce environmental impact and excessive energy use.

CHAPTER II

LITERATURE REVIEW

2.1 Related Research

This research uses a literature review related to the analysis reference to increase understanding related to the research to be carried out. Table 1 below shows a summary of previous research related to this study.

Table 1. ANFIS Applications in Various Fields

No	Research	Method	Performance
1	Analisis Kinerja Metode ANFIS untuk Peramalan Kasus Demam Berdarah di Kabupaten Malang (Anggraeni et al., 2018)	ANFIS vs ARIMA, Decomposition, Holt-Winters; time series data of dengue cases in Malang	ANFIS has better accuracy (MAPE 15%-24%)
2	Bank Soundness Level Prediction: ANFIS vs Deep Learning (Anggraeni et al., 2018)	ANFIS vs LSTM and CNN; bank health prediction with 5 input variables and triangular membership function	ANFIS is superior (MAPE 0.14) to LSTM and CNN.

No	Research	Method	Performance
3	Penerapan ANFIS untuk Peramalan Gangguan pada Transformator Daya (Zakri and Firdaus, 2019)	ANFIS with Gbell and triangular MF; short circuit current data for 25 MVA transformer.	ANFIS effectively predicts transformer faults with Gbell function and low RMSE.
4	Application of an ANFIS model to Optimize the Liquid Flow Rate of a Process Control System (Dutta and Kumar, 2018)	Develop an ANFIS model for controlling liquid flow rates	RMSE: 2.143, MAE: 0.504.
5	AI different approaches and ANFIS data mining: A novel approach to predicting early employment readiness in middle eastern nations (Alkashami et al., 2023)	ANFIS vs Decision Tree, SVM, Naïve Bayes, MLP; 3-7 attributes for job readiness prediction	ANFIS is effective in prediction with high accuracy and Kappa with RMSE 0.3025.
6	Prediksi Curah Hujan Menggunakan Metode Adaptive Neuro Fuzzy Inference System (ANFIS) (Azhar and Mahmudy, 2018)	Predict rainfall in Kabupaten Malang, Indonesia, using the ANFIS.	RMSE: 1.88 .
7	Comparative study of ANN and ANFIS models for predicting temperature in machining (Masoudi et al., 2018)	Model and predict temperature during machining processes using ANN and ANFIS models.	ANFIS model is superior to the ANN model.

The findings from these studies indicate that ANFIS performs effectively across multiple domains, including health, environmental science, finance, education and human resources, industry, and energy. ANFIS is proven to be able to provide accurate predictions and cope with high data complexity. Moreover, these studies show that ANFIS is often superior to traditional methods and other machine learning techniques, especially in situations where data is non-linear and complex. This confirms the importance of ANFIS as an effective tool in data processing and decision-making in various sectors.

2.2 Artificial Neural Network

An artificial neural network (ANN) is a type of computer model inspired by the workings of the human brain. In an artificial neural network, the basic units are called artificial neurons or nodes. The first and simplest ANN model is the perceptron, specifically the Single Layer Perceptron (SLP). The SLP has only one layer of neurons directly connected to the inputs and outputs with no hidden layers. This model is typically used for binary classification with a simple activation function, such as a step or sigmoid function (Siddique and Adeli, 2013).

In neural network terminology, the term 'weight' refers to the strength of the connection between two neurons, representing the influence of information transferring from one neuron to another within the neural network. Each neuron model includes a processing unit that has synaptic input connections and one output. The initial stage involves a calculation where the inputs $x_1, x_2, \ldots, x_n$ are multiplied by their corresponding weights $w_1, w_2, \ldots, w_n$ and then summed by the neuron. A threshold value b, known as the bias, is crucial for certain neuron models and should be explicitly identified as a distinct parameter of the neuron model (Siddique and Adeli, 2013).

In each layer, the input values are transformed into the layer nonlinearly by the process elements and then processed forward to the next layer. Finally, the output values $\hat{y}$, which can be scalar or vector values, are calculated at the output layer with the following equation.

$$\hat{y}(k) = f_o\left(\sum_{j=1}^{a} v_j^o f_j^h\left(\sum_{j=1}^{p} w_{j,i}^h x_i(k) + b_j^h\right) + b_o\right) \tag{2. 1}$$

where:

$x_i(k)$: input variable change for the i-th input, $i = 1, 2, \ldots, p$
p : total number of inputs
$\hat{y}(k)$: estimated output variable
k : index of the input-target data pair $(x_i(k), \hat{y}(k))$, with $k = 1, 2, \ldots, n$
n : number of patterns
$w^h_{j,i}$: weight from the i-th input to the j-th neuron in the hidden layer
b^h_j : bias for the j-th neuron in the hidden layer, with $j = 1, 2, \ldots, q$
q : number of nodes in the hidden layer
f^h_j : activation function at the j-th neuron in the hidden layer
v^o_j : weight from the j-th neuron in the hidden layer to the output neuron
b^o : bias at the output neuron
f^o : activation function at the output neuron.

2.2.1 Activation Functions

Activation functions are utilized in neural networks to calculate the weighted sum of inputs and biases, determining whether a neuron should be activated. The activation function can be categorized as either linear or non-linear based on the function it represents, and it plays a role in regulating the outputs of our neural networks (Nwankpa et al., 2018).

The sigmoid is a non-linear activation function primarily used in feedforward neural networks. It is a bounded and differentiable real function that is defined for real input values, having positive derivatives throughout and exhibiting a certain level of smoothness (Han and Moraga, 1995). The Sigmoid function can be expressed through the relationship

$$f(x) = \frac{1}{1 + \exp^{-x}}. \tag{2. 2}$$

The ReLU is known for promoting faster learning (LeCun et al., 2015) and has shown to be the most effective and widely adopted function (Ramachandran et al., 2017). In comparison to the Sigmoid and tanh activation functions, it provides superior performance and generalization in deep learning. The ReLU behaves almost like a linear function, thus maintaining the favorable attributes of linear models that facilitate optimization through gradient descent methods. The ReLU activation function applies a threshold operation to each input value, setting values below zero to zero; hence, the ReLU is defined as

$$f(x) = \max(0, x) = \begin{cases} x_i, & \text{if } x_i \geq 0 \\ 0, & \text{if } x_i < 0. \end{cases} \tag{2. 3}$$

This function corrects any input values that are below zero, setting them to zero, which helps address the vanishing gradient issue seen in earlier activation functions. The ReLU function is applied within the hidden units of deep neural networks, while a different activation function is typically used in the output layers, with common applications seen in object classification.

2.2.2 Artificial Neural Network Architecture

The architecture of an ANN is defined by the number of layers of processing elements or nodes, which includes input, output, and any hidden layers, as well as the number of nodes present in each layer (Walczak, 2019). In feed-forward networks, data flows only from the input layer to the output layer without any feedback or reconnection, meaning the data only moves forward through the network. In contrast, recurrent networks have feedback connections, which means information can flow back to the previous layer (Rajasekaran and Pai, 2003).

Single-layer feedforward networks are the simplest type of artificial neural network with only two layers, the input layer and the output layer. Unlike single-layer networks, multi-layer networks have one or more hidden layers between the input and output layers. Recurrent networks differ from feedforward types due to the presence of at least one feedback loop. The general architecture of ANN is shown in Figure 1.

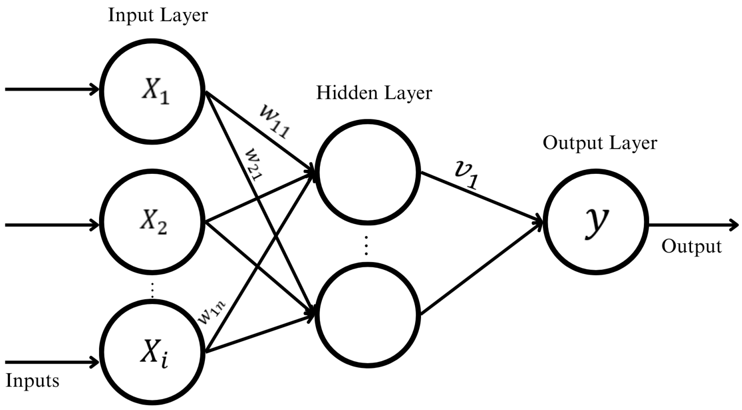

Figure 1. Artificial Neural Network Architecture

2.3 Fuzzy Logic

Fuzzy Logic is a concept developed by Lotfi Zadeh in 1965 to handle uncertainty. Unlike the traditional binary set theory that describes clear-cut events, i.e. events that occur or do not occur, fuzzy logic uses probabilities where there is no definite answer. Therefore, fuzzy logic is used to model uncertain or ambiguous data, as often encountered in real life (Sivanandam et al., 2007).

2.3.1 Fuzzy Sets

A fuzzy set A in X is defined by a membership function $\mu_A(x)$ that assigns a real number within the interval $[0, 1]$ to each point in X, where the values of $\mu_A(x)$ indicate the degree of membership of x in A. This means that the class of objects A contains elements from X with a range of membership grades represented by $\mu_A(x)$. For instance, a fuzzy set $A = \{x_1, x_2, x_3, x_4\}$ in X can be defined by the membership function $\mu_A(x)$, which maps each x in X to the real values $0.5, 1, 0.75$, and 0.5. Here, $\mu_A(x)$ indicates the degree to which x belongs to A, and the mapping is constrained only by $\mu_A(x) \in [0, 1]$. In conventional set theory, the membership function can result in only two possible values: 0 and 1, meaning either $\mu_A(x) = 1$ or $\mu_A(x) = 0$. This is expressed in set-theoretic terms as $\mu_A(x) \in \{0, 1\}$. A fuzzy set serves as a generalization of a classical set. If X represents the universe of discourse with its elements denoted by x, then a fuzzy set A within X is defined as a collection of ordered pairs (Siddique and Adeli, 2013).

$$A = \{(x, \mu_A(x)) \mid x \in X\}$$

2.3.2 Membership Functions

Membership function (MF) is the function that characterizes the degree of fuzziness of a fuzzy set A within X by assigning a real number between 0 and 1 to each point in X. There are no rigid rules for defining a membership function. The membership function in ANFIS can be learned by updating the tunable parameters. The tunable parameters are membership function parameters and consequent parameters. The most commonly used membership functions in fuzzy logic include trapezoidal, gaussian, and generalized bell functions (Siddique and Adeli, 2013).

a. Sigmoidal

The sigmoid membership function is a smooth, "S"-shaped function that defines fuzzy set membership based on a gradual transition from 0 to 1. The parameter a controls the slope of the MF at the cross-point $x = c$. Sigmoidal MFs are shown in Figure 2. A parameterized sigmoidal MF is defined by the Equation 2.4.

$$\text{Sigmoidal}(x, a, c) = \frac{1}{1 + \exp[-a(x - c)]} \tag{2.4}$$

where:

x : input value
a : slopes
c : mean.

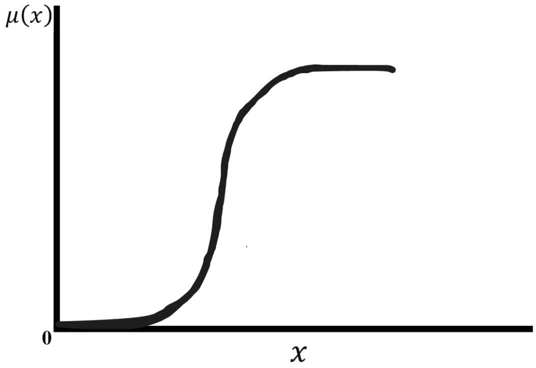

Figure 2. Sigmoidal Membership Function

b. Gaussian

A Gaussian membership function (MF) is specified by two parameters $\{a, c\}$, which are defined by the Equation 2. 5.

$$\text{Gauss}(x, a, c) = \exp\left(-\frac{1}{2}\left(\frac{x-c}{a}\right)^2\right) \tag{2. 5}$$

where:

x : input value
a : standard deviation
c : mean.

The parameters a and c represent the center and width of the Gaussian MF, respectively. Gaussian MFs are shown in Figure 3..

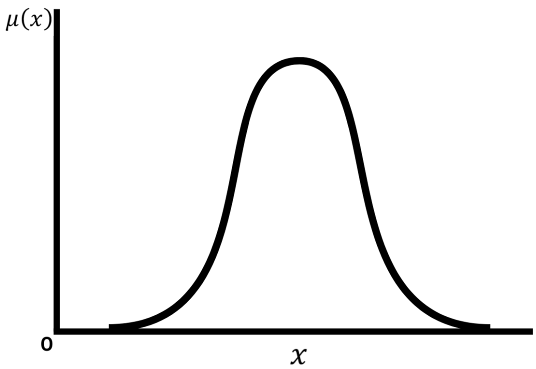

Figure 3. Gaussian Membership Function

c. Generalized Bell

The generalized bell shaped membership function (Gbellmf) has a symmetrical form resembling a bell. According to Equation 2. 6, this function utilizes three parameters: the parameter a defines the breadth of the bell-shaped curve, b is a positive integer, and c determines the center of the curve within the universe of discourse. Generalized Bell MFs are shown in Figure 4.

$$\text{Gbell}(x, a, b, c) = \frac{1}{1 + \left(\frac{|x-c|}{a}\right)^{2b}} \tag{2. 6}$$

where:

x : input value
a : standard deviation
b : slopes
c : mean.

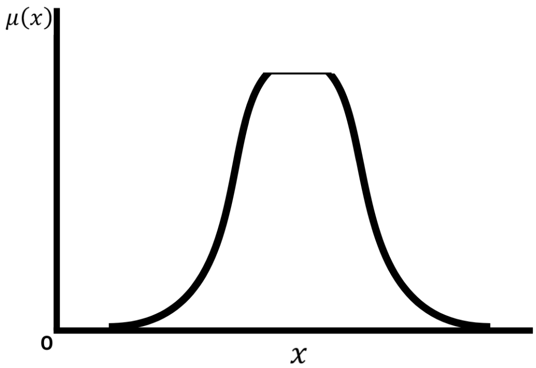

Figure 4. Generalized Bell Membership Function.

2.4 Fuzzy Inference System

Fuzzy inference systems (FISs) are frequently referred to as fuzzy rule-based systems, fuzzy models, fuzzy expert systems, or fuzzy associative memories. This component is a fundamental part of a fuzzy logic system. Decision-making plays a crucial role within the entire framework. The FIS establishes appropriate rules, and decisions are made based on those rules. This process is primarily grounded in the principles of fuzzy set theory, fuzzy IF–THEN statements, and fuzzy reasoning. FIS employs “IF... THEN...” expressions, and the connectors in the rule statements are “OR” or “AND” to construct the necessary decision rules. A basic FIS can accept either fuzzy or crisp inputs, but it typically generates fuzzy set outputs. When the FIS functions as a controller, producing a crisp output is essential. Consequently, a defuzzification method is used to derive a crisp value that effectively represents a fuzzy set (Sivanandam et al., 2007).

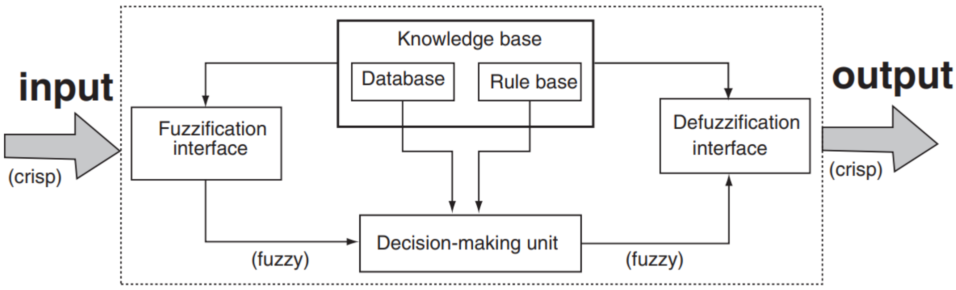

Figure 5. Fuzzy Inference System
(Source : Sivanandam et al., 2007)

A fuzzy inference system comprises a fuzzification interface, a rule base, a database, a decision-making unit, and a defuzzification interface. The FIS includes five functional components as illustrated in Figure 5. The roles of each component are as follows:

- a rule base that holds several fuzzy IF–THEN rules;
- a database that specifies the membership functions for the fuzzy sets utilized in the fuzzy rules;
- a decision-making unit that executes the inference operations on the established rules;
- a fuzzification interface is a process that converts system inputs with crisp values into linguistic variables using membership functions stored in the knowledge base and

- a defuzzification interface that translates the fuzzy outcomes of the inference back into a definite output.

2.4.1 Takagi–Sugeno Fuzzy Method (TS Method)

The Sugeno fuzzy model was introduced by Takagi, Sugeno, and Kang to establish a systematic method for creating fuzzy rules based on an input-output dataset. This model is also referred to as the Sugeno-Takagi model. A standard fuzzy rule within a Sugeno fuzzy model follows this structure.

$$\text{IF } x \text{ is } A \text{ and } y \text{ is } B \text{ THEN } z = f(x, y)$$

where AB represent fuzzy sets in the antecedent; $z = f(x, y)$ denotes a crisp function in the consequent. Typically, $f(x, y)$ is a polynomial involving the input variables x and y, though it can also be any other functions that suitably capture the output of the system within the fuzzy region defined by the antecedent of the rule. In cases where $f(x, y)$ is a first-order polynomial, we refer to this as the first-order Sugeno fuzzy model. If f is a constant, we have the zero-order Sugeno fuzzy model. A standard rule in a Sugeno fuzzy model is structured as follows.

$$\text{IF Input } x = 1 \text{ AND Input } y = 2, \text{ THEN Output is } z = ax + by + c$$

For a zero-order Sugeno model, the output z_i is a constant, which implies that the parameters a and b are both zero (Sivanandam et al., 2007).

2.5 Adaptive Neuro Fuzzy Inference System (ANFIS)

The synergy between neural and fuzzy systems arises from adaptive networks. An adaptive network features a structure defined by a set of parameters that can be adjusted. This type of network integrates both neural networks and fuzzy logic frameworks. An adaptive network is a multilayer feedforward structure where every node executes a specific operation (node function) on input signals along with a set of parameters relevant to that node. The equations for the node functions can differ from one node to another, and the selection of each node function is influenced by the overall input-output relationship that the adaptive network needs to perform. It's important to note that the connections in an adaptive network merely represent the direction of signal flow between nodes, without any weights assigned to the connections (Jang, 1993).

The ANFIS model is developed using pairs of input and output vectors, which modifies the parameters of the membership functions associated with the input and output variables (Jang, 1996). The training process utilizes a hybrid algorithm that merges the gradient descent technique with the Least Squares Estimation (LSE). Fuzzy inference is structured on the Takagi-Sugeno system, where a typical rule can be expressed as: IF A THEN B, with A and B being fuzzy sets defined by their respective membership functions. ANFIS features a five-layer architecture and operates as a feed-forward network, where the outputs of neurons are passed to the inputs of neurons in the subsequent layer, continuing in a linear manner without any cycles. The primary uses of this neuro-fuzzy model include modeling nonlinear systems, predicting chaotic time series, and performing clustering tasks (Stojčić, 2018).

2.5.1 ANFIS Architecture

The ANFIS structure that describes the Takagi-Sugeno-Kang (TSK) fuzzy inference system is in Figure 6. Suppose there are 2 inputs and 1 output. There are 2 rules on the Sugeno rule base as in Equation 2. 7.

$$\begin{aligned} &\text{IF } x_1 \text{ is } A_1 \text{ AND } x_2 \text{ is } B_1 \text{ THEN} \quad y_1 = c_{11}x_1 + c_{12}x_2 + c_{10} \\ &\text{IF } x_1 \text{ is } A_2 \text{ AND } x_2 \text{ is } B_2 \text{ THEN} \quad y_2 = c_{21}x_1 + c_{22}x_2 + c_{20} \end{aligned} \tag{2. 7}$$

where:

y_i : output
c_i, j : consequent parameter

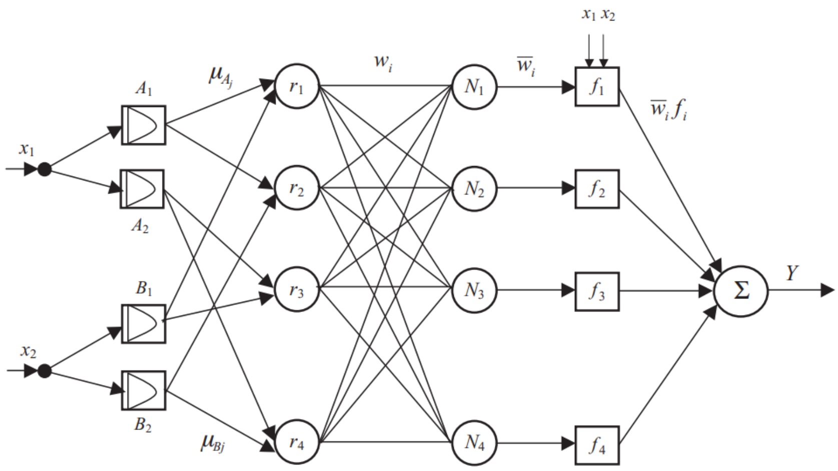

Figure 6. ANFIS Architecture

The following is an explanation for each layer in the ANFIS network architecture:

Layer 1: All nodes in this layer are adaptive (parameters can change) with the node function as follows:

$$\mu_{A_j}(x_1), \mu_{B_j}(x_2), \quad \text{for } j = 1, 2 \tag{2. 8}$$

where x_1 and x_2 are inputs to node i, A_i or B_{i-2} represent the membership functions of each node. The node $O_{1,i}$ indicates the degree of membership of each input with respect to fuzzy sets A and B.

Layer 2: All nodes in this layer are non-adaptive (parameters are fixed). Each node multiplies all incoming signals. The function for this layer is:

$$w_i = \mu_{A_j}(x_1) \cdot \mu_{B_j}(x_2), \quad \text{for } i = 1, 2, 3, 4 \tag{2. 9}$$

Each node output represents the firing strength of each fuzzy rule. This function can be expanded if the premise part has more than two fuzzy sets. The number of nodes in this layer corresponds to the number of rules formed.

Layer 3: Each node in this layer is non-adaptive and displays the normalized firing strength, which is the ratio of the output from the i-th node in the previous layer to the total output from the previous layer:

$$\hat{w}_i = \frac{w_i}{\sum_{i=1}^{4} w_i}, \quad \text{for } i = 1, 2, 3, 4 \tag{2. 10}$$

If more than two rules are formed, this function can be extended by dividing w_i by the total sum of w for all rules.

Layer 4: Each node in this layer is adaptive with the following node function:

$$\hat{w}_i \cdot f_i(a_i x_1 + b_i x_2 + c_i), \quad \text{for } i = 1, 2, 3, 4 \tag{2. 11}$$

where $\bar{w}_i$ is the normalized firing strength from Layer 3, and the parameters a_i, b_i, c_i are adaptive consequent parameters.

Layer 5: In this layer, there is only one fixed node whose function is to sum all inputs. The function is as follows:

$$Y = \sum_{i=1} \bar{w}_i f_i \quad \text{for } i = 1, 2, 3, 4 \tag{2. 12}$$

The adaptive network with five layers is equivalent to the TSK fuzzy inference system (Jang, 1993).

2.5.2 Least Square Estimator (LSE)

If the output of the linear model y is known, expressed through Equation 2. 13.

$$y = k_1 x_1 + k_2 x_2 + \cdots + k_n x_n \tag{2. 13}$$

where:

y : output value
x_n : input variable
k_n : estimated parameter.

By using matrix notation, we get:

$$\mathbf{Ak} = \mathbf{y} \tag{2. 14}$$

In this context, k denotes an unknown vector, while the dimensions of $A, k,$ and y are $P \times M, M \times 1,$ and $P \times 1$, respectively. Given that P (the number of training data pairs) typically exceeds M (the number of linear parameters), the goal is to find a least squares estimate (LSE) of k, denoted as k^*, that minimizes the squared error $||y - Ak||^2$. This represents a common problem foundational to linear regression, adaptive filtering, and signal processing. The most recognized formula for k^* is shown in Equation 2. 15.

$$\mathbf{k}^* = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{y} \tag{2. 15}$$

where $\mathbf{A}^T$ represents the transpose of $\mathbf{A}$, and $(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ is the standard least squares solution when $\mathbf{A}^T\mathbf{A}$ is non-singular. If $\mathbf{A}^T\mathbf{A}$ is singular or does not have an inverse, the pseudo-inverse A^+ can be used instead, leading to $k^* = A^+y$ (Jang, 1993).

2.6 Error Propagation Model

Error propagation is one of the learning methods used to optimize parameters in the Adaptive Neuro-Fuzzy Inference System (ANFIS) model, specifically parameters a and c. Optimizing a and c through error propagation helps ANFIS model non-linear relationships more accurately. The error propagation model in ANFIS includes five main layers as per the architecture of ANFIS described. At each iteration, error propagation optimizes the parameters of these layers to reduce the difference between the prediction output and the target.

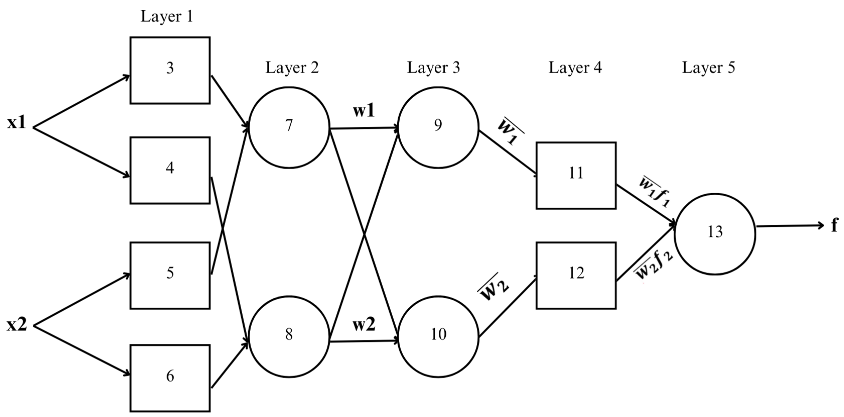

Figure 7. Error Propagation Model in ANFIS

Error on layer 5

If we choose an adaptive network as in Figure 7, which only has 1 neuron in the output layer (neuron 13), then the error propagation to the 5th layer is as in Equation 2. 16.

$$\varepsilon_{13} = \frac{\partial E_p}{\partial x_{13}} = -2(d_{13} - x_{13}) = -2(y_p - y_p'). \tag{2. 16}$$

where y_p is the target output of the p_{th} training data, and y'_p is the network output at the p_{th} training data.

Error on layer 4

The adaptive network in Figure 7 the error propagation that goes to the 4th layer, namely neuron 11 and neuron 12 is as in Equation 2. 17 and Equation 2. 18.

$$\varepsilon_{11} = \left(\frac{\partial E_p}{\partial x_{13}}\right)\left(\frac{\partial f_{13}}{\partial x_{11}}\right) = \varepsilon_{13}\left(\frac{\partial f_{13}}{\partial x_{11}}\right) = \varepsilon_{13}(1) = \varepsilon_{13}. \tag{2. 17}$$

Since $f_{13} = \overline{w_1} f_1 + \overline{w_2} f_2,$ then $\frac{\partial f_{13}}{\partial(\overline{w_1} f_1)} = 1.$

$$\varepsilon_{12} = \left(\frac{\partial E_p}{\partial x_{13}}\right)\left(\frac{\partial f_{13}}{\partial x_{12}}\right) = \varepsilon_{13}\left(\frac{\partial f_{13}}{\partial x_{12}}\right) = \varepsilon_{13}(1) = \varepsilon_{13}. \tag{2. 18}$$

Since $f_{13} = \overline{w_1} f_1 + \overline{w_2} f_2,$ then $\frac{\partial f_{13}}{\partial(\overline{w_2} f_2)} = 1.$

Error on layer 3

The adaptive network in Figure 7 propagation of errors that go to the 3rd layer, namely neuron 9 and neuron 10 as in Equation 2. 19 and Equation 2. 20.

$$\varepsilon_9 = \left(\frac{\partial E_p}{\partial x_{13}}\right)\left(\frac{\partial f_{13}}{\partial x_{11}}\right)\left(\frac{\partial f_{11}}{\partial x_9}\right) = \varepsilon_{11}\left(\frac{\partial f_{11}}{\partial x_9}\right) = \varepsilon_{11} f_1 \tag{2. 19}$$

and

$$\varepsilon_{10} = \left(\frac{\partial E_p}{\partial x_{13}}\right)\left(\frac{\partial f_{13}}{\partial x_{12}}\right)\left(\frac{\partial f_{12}}{\partial x_{10}}\right) = \varepsilon_{12}\left(\frac{\partial f_{12}}{\partial x_{10}}\right) = \varepsilon_{12} f_2. \tag{2. 20}$$

Error on layer 2

The adaptive network in Figure 7 propagation of errors that go to the 2nd layer, namely neuron 7 and neuron 8 as in Equation 2. 21 and Equation 2. 22.

$$\begin{aligned}
\varepsilon_7 &= \left(\frac{\partial E_p}{\partial x_{13}}\right)\left(\frac{\partial f_{13}}{\partial x_{11}}\right)\left(\frac{\partial f_{11}}{\partial x_9}\right)\left(\frac{\partial f_9}{\partial x_7}\right) + \left(\frac{\partial E_p}{\partial x_{13}}\right)\left(\frac{\partial f_{13}}{\partial x_{12}}\right)\left(\frac{\partial f_{12}}{\partial x_{10}}\right)\left(\frac{\partial f_{10}}{\partial x_7}\right) \\
&= \varepsilon_9\left(\frac{\partial f_9}{\partial x_7}\right) + \varepsilon_{10}\left(\frac{\partial f_{10}}{\partial x_7}\right) \\
&= \varepsilon_9\left(\frac{w_2}{(w_1+w_2)^2}\right) + \varepsilon_{10}\left(-\frac{w_2}{(w_1+w_2)^2}\right) \\
&= \frac{w_2}{(w_1+w_2)^2}\left(\varepsilon_9 - \varepsilon_{10}\right).
\end{aligned} \tag{2. 21}$$

$$\begin{aligned}
\varepsilon_8 &= \left(\frac{\partial E_p}{\partial x_{13}}\right)\left(\frac{\partial f_{13}}{\partial x_{12}}\right)\left(\frac{\partial f_{12}}{\partial x_{10}}\right)\left(\frac{\partial f_{10}}{\partial x_8}\right) + \left(\frac{\partial E_p}{\partial x_{13}}\right)\left(\frac{\partial f_{13}}{\partial x_{11}}\right)\left(\frac{\partial f_{11}}{\partial x_9}\right)\left(\frac{\partial f_9}{\partial x_8}\right) \\
&= \varepsilon_{10}\left(\frac{\partial f_{10}}{\partial x_8}\right) + \varepsilon_9\left(\frac{\partial f_9}{\partial x_8}\right) \\
&= \varepsilon_{10}\left(\frac{w_1}{(w_1+w_2)^2}\right) + \varepsilon_9\left(-\frac{w_1}{(w_1+w_2)^2}\right) \\
&= \frac{w_1}{(w_1+w_2)^2}\left(\varepsilon_{10}-\varepsilon_9\right). \qquad (2.\ 22)
\end{aligned}$$

Error on layer 1

The adaptive network in Figure 7 error propagation that goes to the 1st layer, namely neurons 3, 4, 5, and 6 as in Equation 2. 23, Equation 2. 24, Equation 2. 25 and Equation 2. 26.

$$\varepsilon_3 = \varepsilon_7\left(\frac{\partial f_7}{\partial x_3}\right) = \varepsilon_7\mu_{B1}(x_2)\mu_{C1}(x_3)\mu_{D1}(x_4)\mu_{E1}(x_5)\mu_{F1}(x_6). \qquad (2.\ 23)$$

$$\varepsilon_4 = \varepsilon_8\left(\frac{\partial f_8}{\partial x_4}\right) = \varepsilon_8\mu_{B2}(x_2)\mu_{C2}(x_3)\mu_{D2}(x_4)\mu_{E2}(x_5)\mu_{F2}(x_6). \qquad (2.\ 24)$$

$$\varepsilon_5 = \varepsilon_7\left(\frac{\partial f_7}{\partial x_5}\right) = \varepsilon_7\mu_{A1}(x_1)\mu_{C1}(x_3)\mu_{D1}(x_4)\mu_{E1}(x_5)\mu_{F1}(x_6). \qquad (2.\ 25)$$

$$\varepsilon_6 = \varepsilon_8\left(\frac{\partial f_8}{\partial x_6}\right) = \varepsilon_8\mu_{A2}(x_1)\mu_{C2}(x_3)\mu_{D2}(x_4)\mu_{E2}(x_5)\mu_{F2}(x_6). \qquad (2.\ 26)$$

Furthermore, the error can be used to find error information for the parameters a and c, as in Equation 2. 27 and Equation 2. 28.

$$\varepsilon_{aik} = \frac{2(x_i - c_{ik})^2}{a_{ik}^2\left(1+\left(\frac{x_i-c_{ik}}{a_{ik}}\right)^2\right)^2} * \varepsilon_i \qquad (2.\ 27)$$

and

$$\varepsilon_{cik} = \frac{2(x_i - c_{ik})^2}{a_{ik}^2\left(1+\left(\frac{x_i-c_{ik}}{a_{ik}}\right)^2\right)^2} * \varepsilon_i. \qquad (2.\ 28)$$

From the above equation, the change in parameter values a_{ik} and c_{ik} ($\Delta\alpha_{aik}$ and $\Delta\alpha_{cik}$) can be determined as in Equation 2. 29.

$$\Delta a_{ik} = \eta\varepsilon_{aik}, \text{ dan } \Delta c_{ik} = \eta\varepsilon_{cik}x_i. \tag{2. 29}$$

where η is the learning rate which lies in the interval $[0, 1]$ so that the new a_{ik} and c_{ik} values can be calculated as in Equation 2. 30 and Equation 2. 31 (Suyanto and Sc, 2008).

$$a_{ik} = a_{ik(old)} - \Delta a_{ik}, \tag{2. 30}$$

$$c_{ik} = c_{ik(old)} - \Delta c_{ik}. \tag{2. 31}$$

2.7 Data Preprocessing

Data preprocessing is crucial for ensuring valid data analysis. It plays an essential role in the development of operational data analysis due to the inherent complexity of building operations and the shortcomings in data quality. Data preprocessing is the initial stage in data analysis that aims to clean, organize, and transform raw data into a more suitable format for further analysis. This includes several steps such as data cleaning, data normalization, and feature selection (Fan et al., 2021).

2.7.1 Data Cleaning

Many methods in data mining depend on a dataset that is assumed to be complete or devoid of noise. Nevertheless, actual data is often far from clean or exhaustive. In the data preprocessing stage, it is typical to apply techniques aimed at either eliminating noisy data or imputing (replacing) the missing values. Missing values frequently occur during data acquisition processes. A missing value refers to a piece of data that has not been recorded or captured due to issues in the sampling process, financial constraints, or limitations in the data gathering method. Missing values is unavoidable in data analysis and often presents significant challenges for practitioners. Addressing missing values is complex and mishandling them can easily result in poor insights and incorrect conclusions. Missing values has been known to reduce efficiency in the knowledge extraction process, introduce strong biases if the mechanism of missingness is not correctly managed, and create substantial difficulties in data management (Luengo et al., 2020).

Numerous methods exist to address the challenges posed by missing values during data preprocessing. The most common initial option is to remove instances that may contain an missing values. However, this strategy rarely proves advantageous, as discarding instances can introduce bias into the learning process and lead to the loss of critical information. Foundational studies on data imputation originated from statistics. These studies model the probability distributions of the data and consider the mechanisms that contribute to missingness. By employing maximum likelihood techniques, they simulate the approximate probabilistic models to fill in the missing values. Since the actual probability model for a specific dataset is typically unknown, the application of machine learning techniques has gained significant traction lately, as they can be used without requiring any prior information (Luengo et al., 2020).

2.7.2 Normalization

Normalization methods play a crucial role in speeding up the training process and enhancing the generalization capabilities of neural networks, and they have been effectively applied in numerous applications (Huang et al., 2023). Among various normalization techniques, Min-Max normalization is widely used due to its simplicity and effectiveness in scaling data to a fixed range, typically $[0, 1]$ (Han et al., 2022).

Min-Max normalization adjusts the values of a dataset linearly by subtracting the minimum value of the dataset and dividing by the range, which is the difference between the maximum and minimum values. The formula for this transformation can be expressed as shown in Equation 2. 32.

$$x' = \frac{x - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}}. \tag{2. 32}$$

The advantages of Min-Max normalization include its interpretability and ability to maintain the original distribution of the data. It is particularly beneficial in scenarios where the features have different units or scales, as it standardizes the data without distorting its inherent patterns. However, it is sensitive to outliers, extreme values can disproportionately affect the scaling process, making it essential to handle outliers before applying this method (Patro, 2015).

2.7.3 Feature Selection

Real-world datasets often have a significant amount of irrelevant, redundant, and noisy features. By utilizing feature selection to eliminate these features, we can decrease storage and computational expenses while maintaining crucial information and preserving learning performance. Feature selection serves as an effective and efficient data preprocessing technique, particularly beneficial for preparing high-dimensional data for a range of data mining and machine learning challenges. When there are a substantial number of features, learning models are prone to overfitting, which can lead to a decline in performance when dealing with new data. High-dimensional data can considerably escalate the memory storage needs and computational expenses associated with data analysis (Li et al., 2017).

Dimensionality reduction serves as a highly effective method to tackle the previously mentioned challenges. It can primarily be divided into two key components: feature extraction and feature selection. Feature extraction transforms the original high-dimensional features into a new feature space with reduced dimensions. The newly created feature space typically consists of either linear or nonlinear combinations of the original features. In contrast, feature selection involves directly choosing a subset of relevant features for constructing a model (Guyon et al., 2008).

Both feature extraction and feature selection offer benefits like enhancing learning outcomes, boosting computational efficiency, minimizing memory usage, and developing more effective generalization models. Consequently, both methods are considered valuable dimensionality reduction strategies. In scenarios where the raw input data lacks understandable features for a specific learning algorithm, feature extraction is generally favored. Conversely, since feature extraction generates a new set of features, further analysis can become challenging as the original physical meanings of these features are lost. In contrast, by preserving some original features, feature selection retains their physical interpretations, thus providing models with improved readability and interpretability. As a result, feature selection is frequently favored in various applications such as text mining and genetic research. It is important to note that even when feature dimensionality is relatively low, feature extraction or selection can still be crucial for enhancing learning performance, mitigating overfitting, and lowering computational expenses (Li et al., 2017).

One effective approach for feature selection is to evaluate the strength of the relationship between each feature and the target variable. In this context, the Spearman correlation coefficient serves as a valuable tool, as it measures the monotonic relationship between variables and helps identify features that are most relevant to the target. It is a metric that indicates both the intensity and direction of the association between two phenomena only. The relationship can be classified as either negative or positive, as well as weak or strong. Equation 2. 33 is a formula for calculating the Spearman rank correlation coefficient (r_s) (Ali Abd Al-Hameed, 2022).

$$r_s = 1 - \frac{6 \sum d_i^2}{n^3 - n} \tag{2. 33}$$

where:

r_s	: strength of the rank correlation between variables
d_i	: the difference between the x-variable rank and the y-variable rank for each pair of data
$\sum d_i^2$	: sum of the squared differences between x- and y-variable ranks
n	: sample size.

The Spearman correlation coefficient (r) is between -1 and 1, where:

- $r = 1$: Perfect positive relationship.
- $r = -1$: Perfect negative relationship.
- $r = 0$: No linear relationship.

2.8 K-Means Algorithm

K-Means algorithm is a non-hierarchical algorithm derived from data clustering method. K-Means is among the most straightforward unsupervised learning algorithms designed to address the common clustering issue. The method uses a clear and uncomplicated approach to group a specific dataset into a predetermined number of clusters (let's assume k clusters) specified beforehand. The following are the steps in clustering data using the k-means method:

1. Set the value of k as the number of clusters to be formed.
2. Initialize k cluster centroids. This can be done in various ways, with the most common method being random selection from the data points.
3. Calculate the distance of each data point from each centroid using the Euclidean Distance formula to find the closest centroid for each data point. The Euclidean distance between a data point $\mathbf{X} = (x_1, x_2, \ldots, x_j)$ and a centroid $\mathbf{Y}_k = (y_{k1}, y_{k2}, \ldots, y_{kj})$ is defined as:

$$d(\mathbf{X}, \mathbf{Y}_k) = \sqrt{\sum_{i=1}^{j}(x_i - y_{ki})^2} \tag{2. 34}$$

   where:

   $d(\mathbf{X}, \mathbf{Y}_k)$: Euclidean distance between data point $\mathbf{X}$ and centroid $\mathbf{Y}_k$
   x_i : i-th component of the data point $\mathbf{X}$
   y_{ki} : i-th component of the k-th centroid
   j : total number of features or dimensions

4. Classify each data point based on its proximity to the nearest centroid (smallest distance).
5. Update the centroid values. This is done by finding the mean (average) of all points in the cluster.
6. Repeat steps 2 through 5 until there is no change in the cluster members.

2.9 Learning Rate

Learning rate is a parameter that control how many steps are taken to update the model weights during training. In the context of optimization algorithms, the learning rate determines how fast or slow the model learns from the training data. It is a measure of the weight change at each iteration based on the gradient of the loss function (Jacobs, 1988).

Typically, a deep neural network is updated via stochastic gradient descent, where the parameters θ (weights) are adjusted according to

$$\theta_t = \theta_{t-1} - \eta_t \frac{\partial L}{\partial \theta}$$

with ∂L representing the loss function and η_t as the learning rate. It is well-established that a learning rate that is too small will result in a slow convergence of the training algorithm, while a learning rate that is excessively large will cause the training algorithm to diverge. Therefore, it is essential to test various learning rates and schedules (Smith, 2017).

2.10 Accuracy of Testing Results

Root mean square error (RMSE) quantifies the discrepancies between the values forecasted by a model and those that are truly observed. RMSE is particularly useful for evaluating continuous prediction models as it offers a straightforward measure of accuracy. A lower RMSE signifies superior model effectiveness in reflecting the variability of the data (Hodson, 2022).

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \bar{y}_i)^2}{N}} \tag{2. 35}$$

where:

N : number of data points
y_i : actual output (target of the network)
$\bar{y}_i$: predicted output (output of the network)

2.11 Energy Efficiency

According to statistics from design and construction to demolition, buildings consume about 40% of the earth's energy and then produce 40% of the earth's total waste. In Indonesia, the use of electricity contributes CO_2 from the greenhouse effect and causes global warming so climate change often occurs even though Indonesia is located on the equator with solar potential with high intensity (Indarto et al., 2015). The release of CO_2 emissions resulting from energy combustion in power plants, the transportation sector, the industrial sector, commercial, households, and other sectors into the atmosphere in a certain amount where this will have an impact on increasing global warming. Efforts to reduce the causes of increased global warming can be made through increasing the efficiency of energy technology and utilizing energy sources with low carbon content (Nasional et al., 2019).

The swift economic development globally has driven a surge in energy usage. Fossil fuels continue to dominate the global energy landscape, accounting for approximately 81% of the market. The extraction process releases harmful gases into the atmosphere, raising societal environmental concerns (Memon, 2014). Presently, the rapid expansion of urban areas and heightened comfort levels have led to increased energy consumption, marking it as a major issue for contemporary society. This dilemma arises from excessive reliance on non-renewable energy sources, which inflict serious environmental consequences. A significant portion of electricity consumption in residential areas is linked to heating and cooling needs. Therefore, there is an urgent need to implement strategies that enhance the energy efficiency of buildings (da Cunha and de Aguiar, 2020).

Annually, the sun provides around 5×10^{24} Joule of energy that reaches the Earth's surface. This figure is roughly 10000 times greater than the total energy consumed worldwide each year. Thus, the urgency to harness this natural resource, coupled with the demand for improvements in the quality of the built environment, is critical. In this context, the scientific community is now focused on integrating solar energy with innovative building materials that can help reduce energy consumption in structures (Diamanti et al., 2008).

CHAPTER III

RESEARCH METHODS

3.1 Time and Place of Research

This research was conducted in the odd semester of the 2024/2025 academic year at the Department of Mathematics, Faculty of Mathematics and Natural Sciences, University of Lampung, which is located at Jalan Prof. Dr. Ir. Soemantri Brojonegoro, Gedong Meneng, Rajabasa District, Bandar Lampung City, Lampung.

3.2 Research Data

The data used in this research is the Energy Efficiency dataset which was donated on November 29, 2012. The data was obtained through the UCI Machine Learning Repository website. This dataset includes 768 samples with 8 main features, namely Relative Compactness (RC), Surface Area (SA), Wall Area (WA), Roof Area (RA), Overall Height (OH), Orientation (OR), Glazing Area (GA), and Glazing Area Distribution (GAD). The targets in this dataset are the heating load and cooling load of the building.

3.3 Research Methods

This chapter discusses the research methods used to predict heating and cooling loads using the Adaptive Neuro Fuzzy Inference System (ANFIS) method. The methodology in this research is carried out in several stages, namely as follows:

1. Descriptive analysis of the variables Relative Compactness (x_1), Surface Area

(x_2), Wall Area (x_3), Roof Area (x_4), Overall Height (x_5), Orientation (x_6), Glazing Area (x_7), and Glazing Area Distribution (x_8), and cooling load (Y).

2. Perform data pre-processing to clean and prepare the data, including the following steps:
   (a) Identifying missing values in the dataset.
   (b) Identifying the most relevant features for the model using the Spearman correlation test.
   (c) Perform data normalization to balance the scale of each different input variable.
3. Splitting data into training and testing sets with proportions of $50 : 50, 60 : 40, 70 : 30, 80 : 20$, and $90 : 10$.
4. Build ANFIS model using training data with 3 membership functions namely simoidal, gaussian, and generalized bell. ANFIS model training is done with the following steps:
   (a) Clustering using K-Means. The process of grouping data using the K-Means clustering method aims to group data into 2 clusters. The clustering steps using the K-Means algorithm are as follows:
      i. Determine the initial cluster center (centroid) randomly.
      ii. Determine the members of each cluster using the Euclidean distance formula according to Equation 2. 34.
      iii. Calculating the new cluster center.
      iv. Perform the iteration process until there is no change in the members of each cluster formed.
      v. Obtained output in the form of members of each cluster.
   (b) Calculate the mean and standard deviation of each cluster formed which will be used as premise parameters to calculate the degree of membership using sigmoidal, gaussian, and generalized bell membership functions used as ANFIS input.
   (c) Calculate the degree of membership based on input data using Equation 2. 8.
   (d) Performing the firing strength process. The fire strength value is obtained from the multiplication process between the values of each membership degree obtained using Equation 2. 9.

(e) Performing the fire strength value standardization process to get the normalized fire strength value using Equation 2. 10.

(f) Calculating the value of the consequent parameter using the Least Square Estimator (LSE) algorithm according to Equation 2. 15.

(g) Calculate the output value at layer 4 using Equation 2. 11.

(h) Summing up all values at layer 4 to get the output at layer 5 (network output).

(i) Calculate the propagation error of the network. If the resulting error in the network is smaller or equal to the maximum error value set and the epoch value is greater than the maximum iteration then proceed to the next process. If not, then perform the improvement process on the consequent parameters and premise parameters using the steepest descent algorithm until the appropriate value is obtained.

(j) Obtain the values of the premise parameters and consequent parameters in the form of mean (c) and standard deviation (a).

5. The model that has been formed through the training process is tested using the testing data obtained from splitting the data with a predetermined proportion. The training process is carried out with specific parameters, such as the number of iterations (epochs) and the predetermined learning rate.

6. Evaluate the performance of the model using test data with evaluation metrics represented by Root Mean Squared Error (RMSE) for each scenario.

7. Comparing the performance results of several models to determine the best model based on predefined evaluation metrics.

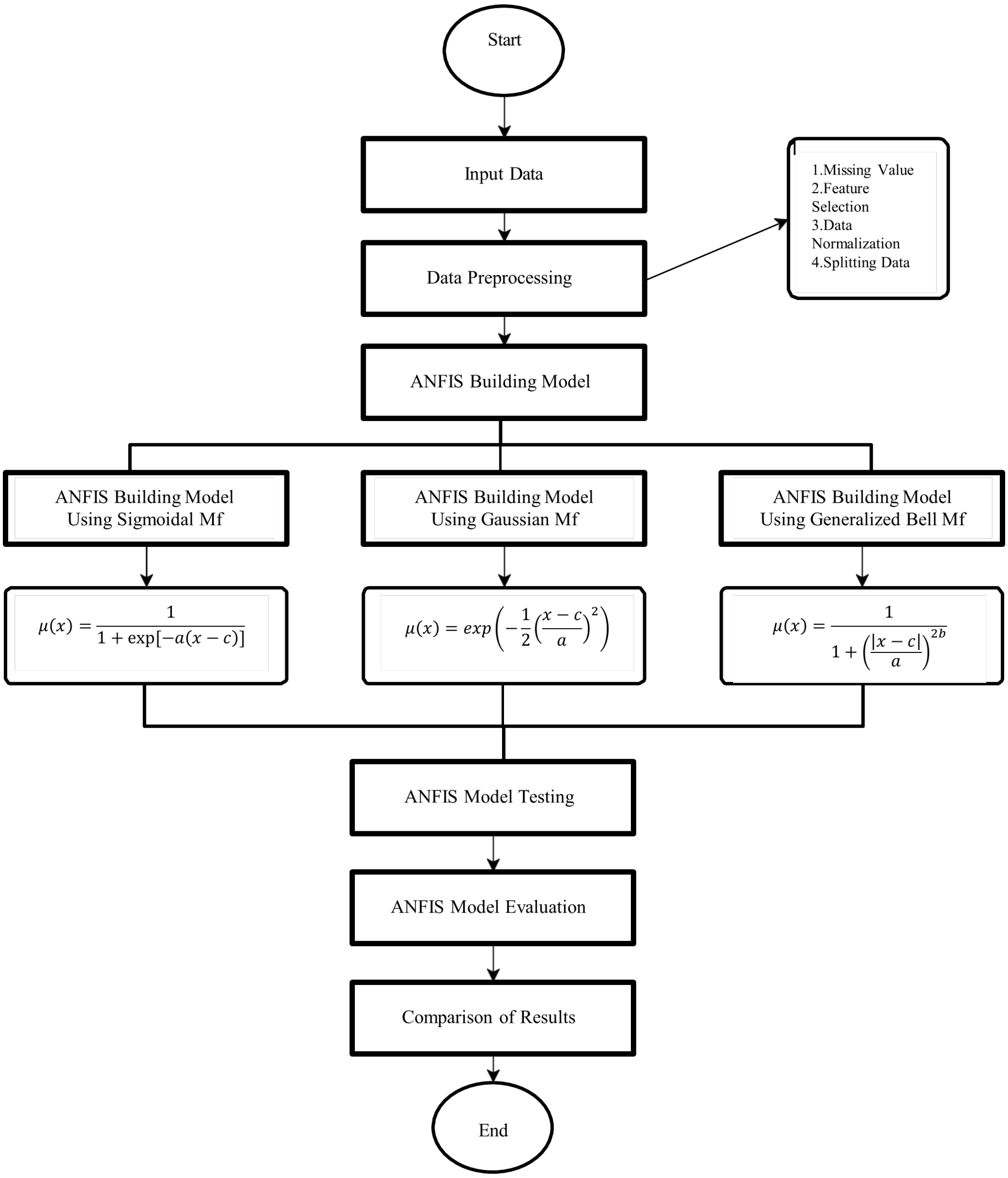

Figure 8. Flowchart of Research Methods Comparison of ANFIS Models with Multiple Membership Functions

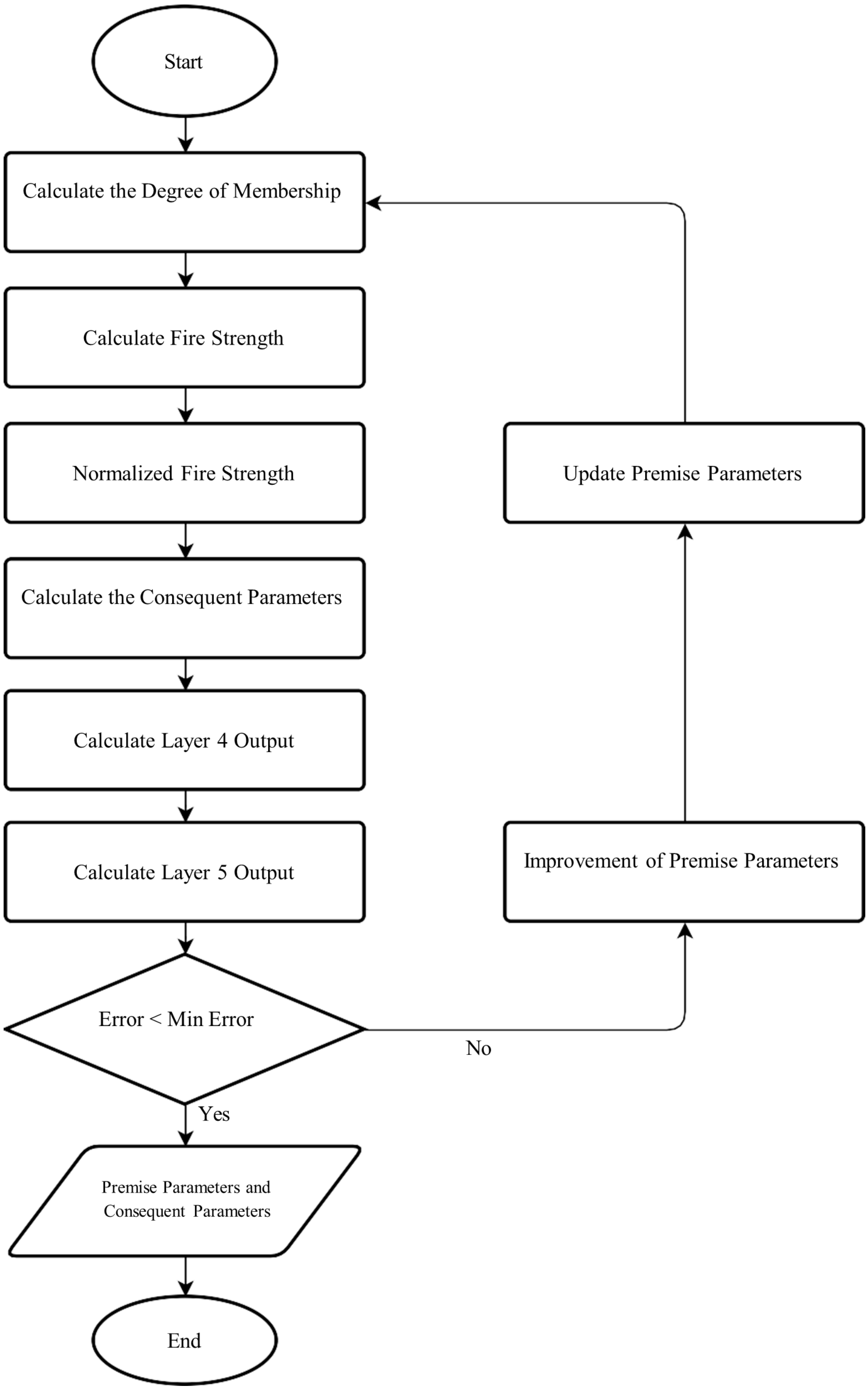

Figure 9. ANFIS Flowchart

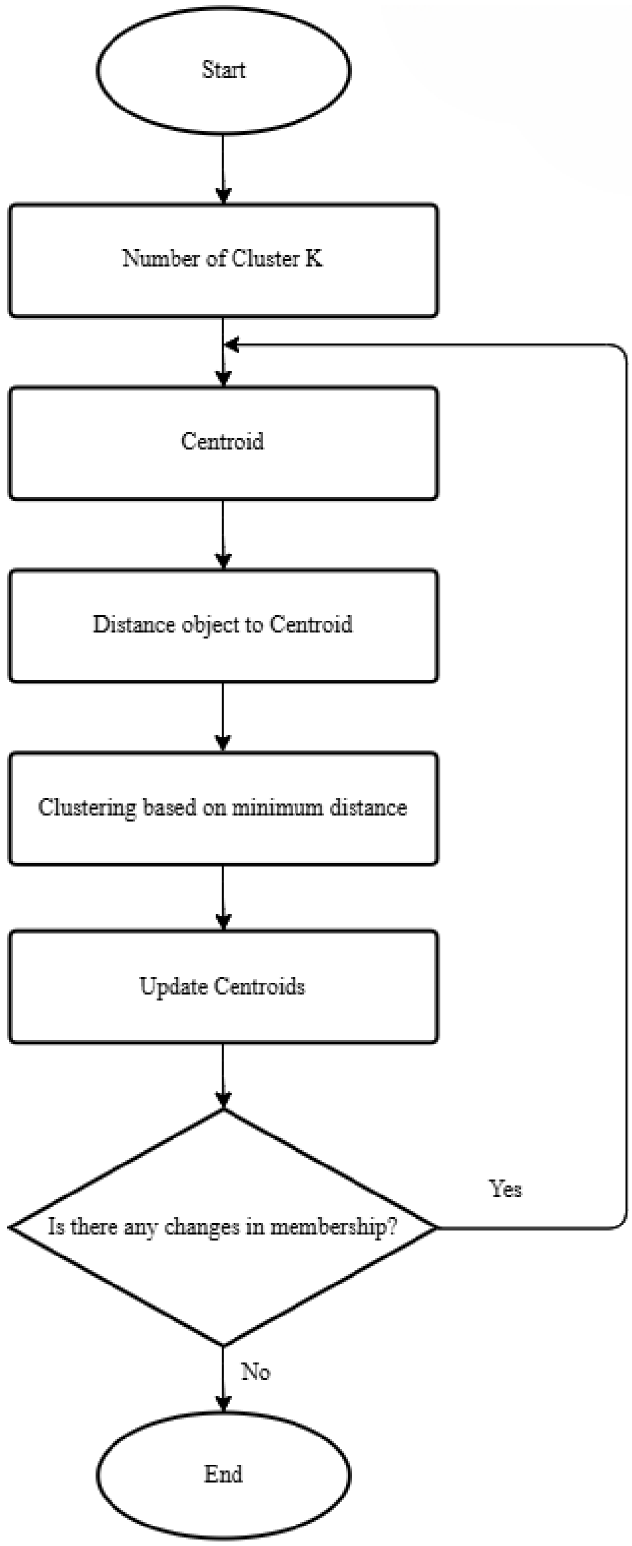

Figure 10. K-Means Flowchart

CHAPTER V

CONCLUSIONS

Based on the results of the analysis and discussion, the following conclusions are obtained:

1. This research successfully built an Adaptive Neuro-Fuzzy Inference System (ANFIS) model that can accurately predict cooling load using input variables such as Relative Compactness (RC), Surface Area (SA), Wall Area (WA), Roof Area (RA), Overall Height (OH), Orientation (OR), Glazing Area (GA), and Glazing Area Distribution (GAD).

2. The analysis of predicting the cooling load based on energy efficiency data through the ANFIS approach with three different membership functions—sigmoidal, Gaussian, and generalized bell—yields an RMSE value of 4.070644 for the sigmoidal membership function, using an $80:20$ data split, 5 iterations, and a learning rate of 0.1. For the Gaussian function, applying an $80:20$ data split, 200 iterations, and a learning rate of 0.4 results in an RMSE of 2.60803. In contrast, the generalized bell function with an $80:20$ data division, 1000 iterations, and a learning rate of 0.3 achieves an RMSE of 2.058887.

3. The best ANFIS model is the model with generalized bell membership function because it produces the lowest RMSE. The generalized bell membership function used in this model has the following form:

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

REFERENCES

Adiwoso, N. S. and Prasetyoadi, S. P. (2010). Towards indonesia's sustainable future through sustainable building and construction. In *Conference on Sustainable Buildings South East Asia (SB10SEA)"New Green opportunities and Challenges". Kuala Lumpur Malaysia*, volume 4.

Ali Abd Al-Hameed, K. (2022). Spearman's correlation coefficient in statistical analysis. *International Journal of Nonlinear Analysis and Applications*, 13(1):3249–3255.

Alkashami, M., Taamneh, A., Khadragy, S., Shwedeh, F., Aburayya, A., and Salloum, S. (2023). Ai different approaches and anfis data mining: A novel approach to predicting early employment readiness in middle eastern nations. *International Journal of Data and Network Science*, 7(3):1267–1282.

Anggraeni, W., Kurniaji, G. N., Riksakomara, E., Samopa, F., Wibowo, R. P., et al. (2018). Analisis kinerja metode anfis untuk peramalan kasus demam berdarah di kabupaten malang. *SISFO Vol 7 No 3*, 7.

Azhar, M. I. and 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.

Buragohain, M. and Mahanta, C. (2008). A novel approach for anfis modelling based on full factorial design. *Applied Soft Computing*, 8(1):609–625.

Central, C. (2023). Hotter climate, higher cooling demand — climate central. Accessed: 2024-12-05.

da Cunha, S. R. L. and de Aguiar, J. L. B. (2020). Phase change materials and energy efficiency of buildings: A review of knowledge. *Journal of Energy Storage*, 27:101083.

Diamanti, M. V., Ormellese, M., and Pedeferri, M. (2008). Characterization of photocatalytic and superhydrophilic properties of mortars containing titanium dioxide. *Cement and Concrete Research*, 38(11):1349–1353.

Dutta, P. and Kumar, A. (2018). Application of an anfis model to optimize the liquid flow rate of a process control system. *Chemical Engineering Transactions*, 71:991–996.

Fan, C., Chen, M., Wang, X., Wang, J., and Huang, B. (2021). A review on data preprocessing techniques toward efficient and reliable knowledge discovery from building operational data. *Frontiers in energy research*, 9:652801.

Guyon, I., Gunn, S., Nikravesh, M., and Zadeh, L. A. (2008). *Feature extraction: foundations and applications*, volume 207. Springer.

Han, J. and Moraga, C. (1995). The influence of the sigmoid function parameters on the speed of backpropagation learning. In *International workshop on artificial neural networks*, pages 195–201. Springer.

Han, J., Pei, J., and Tong, H. (2022). *Data mining: concepts and techniques*. Morgan kaufmann.

Hodson, T. O. (2022). Root mean square error (rmse) or mean absolute error (mae): When to use them or not. *Geoscientific Model Development Discussions*, 2022:1–10.

Hou, Z., Shen, Q., and Li, H. (2003). Nonlinear system identification based on anfis. In *International Conference on Neural Networks and Signal Processing, 2003. Proceedings of the 2003*, volume 1, pages 510–512. IEEE.

Huang, L., Qin, J., Zhou, Y., Zhu, F., Liu, L., and Shao, L. (2023). Normalization techniques in training dnns: Methodology, analysis and application. *IEEE transactions on pattern analysis and machine intelligence*, 45(8):10173–10196.

Indarto, E., Hardiman, G., and Murtomo, B. A. (2015). Kesesuaian posisi orientasi dan kemiringan solar sel pada bidang selimut bangunan dalam manifestasi arsitektur aktif desain. *Modul*, 15(1):13–22.

Jacobs, R. A. (1988). Increased rates of convergence through learning rate adaptation. *Neural networks*, 1(4):295–307.

Jang, J.-S. (1993). Anfis: adaptive-network-based fuzzy inference system. *IEEE transactions on systems, man, and cybernetics*, 23(3):665–685.

Jang, J.-S. (1996). Input selection for anfis learning. In *Proceedings of IEEE 5th international fuzzy systems*, volume 2, pages 1493–1499. IEEE.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *nature*, 521(7553):436–444.

Li, J., Cheng, K., Wang, S., Morstatter, F., Trevino, R. P., Tang, J., and Liu, H. (2017). Feature selection: A data perspective. *ACM computing surveys (CSUR)*, 50(6):1–45.

Liu, Y., Wang, Q., Zhang, X., Song, S., Niu, C., and Shangguan, Y. (2018). Using anfis and bpnn methods to predict the unfrozen water content of saline soil in western jilin, china. *Symmetry*, 11(1):16.

Luengo, J., García-Gil, D., Ramírez-Gallego, S., García, S., and Herrera, F. (2020). Big data preprocessing. *Cham: Springer*.

Masoudi, S., Sima, M., and Tolouei-Rad, M. (2018). Comparative study of ann and anfis models for predicting temperature in machining. *Journal of Engineering Science and Technology*, 13(1):211–225.

Memon, S. A. (2014). Phase change materials integrated in building walls: A state of the art review. *Renewable and sustainable energy reviews*, 31:870–906.

Nasional, T. S. J. D. E. et al. (2019). Indonesia energy out look 2019. *J. Chem. Inf. Model*, 53(9):1689–1699.

Nwankpa, C., Ijomah, W., Gachagan, A., and Marshall, S. (2018). Activation functions: Comparison of trends in practice and research for deep learning. *arXiv preprint arXiv:1811.03378*.

Patro, S. (2015). Normalization: A preprocessing stage. *arXiv preprint arXiv:1503.06462*.

Rajasekaran, S. and Pai, G. V. (2003). *Neural networks, fuzzy logic and genetic algorithm: synthesis and applications (with cd)*. PHI Learning Pvt. Ltd.

Ramachandran, P., Zoph, B., and Le, Q. V. (2017). Searching for activation functions. *arXiv preprint arXiv:1710.05941*.

Savić, M., Mihajlović, I., Arsić, M., and Živković, Ž. (2014). Adaptive-network-based fuzzy inference system (anfis) modelbased prediction of the surface ozone concentration. *Journal of the Serbian Chemical Society*, 79(10):1323–1334.

Sayama, H. (2015). *Introduction to the modeling and analysis of complex systems*. Open SUNY Textbooks.

Siddique, N. and Adeli, H. (2013). *Computational intelligence: synergies of fuzzy logic, neural networks and evolutionary computing*. John Wiley & Sons.

Sivanandam, S., Sumathi, S., and Deepa, S. (2007). *Introduction to fuzzy logic using MATLAB*. Springer.

Smith, L. N. (2017). Cyclical learning rates for training neural networks. In *2017 IEEE winter conference on applications of computer vision (WACV)*, pages 464–472. IEEE.

Stojčić, M. (2018). Application of anfis model in road traffic and transportation: a literature review from 1993 to 2018. *Operational Research in Engineering Sciences: Theory and Applications*, 1(1):40–61.

Suyanto, S. and Sc, M. (2008). Soft computing. *Bandung: Informatika*.

Walczak, S. (2019). Artificial neural networks. In *Advanced methodologies and technologies in artificial intelligence, computer simulation, and human-computer interaction*, pages 40–53. IGI global.

Yudelson, J. (2007). *Green building A to Z: Understanding the language of green building*. New Society Publishers.

Zakri, A. A. and Firdaus, H. (2019). Penerapan anfis untuk peramalan gangguan pada transformator daya. In *Prosiding Seminar Nasional Pakar*, pages 1–3.

Ziasabounchi, N. and Askerzade, I. (2014). Anfis based classification model for heart disease prediction. *International Journal of Electrical & Computer Sciences IJECS-IJENS*, 14(02):7–12.