COMPARISON STUDY OF FUZZY C-MEANS AND FUZZY SUBTRACTIVE CLUSTERING IMPLEMENTATION IN QUALITY OF INDIHOME FIBER OPTIC NETWORK (Case Study in PT. TELKOM INDONESIA)

THESIS

Submitted to International Program
Department of Industrial Engineering
as The Partial Requirement of Acquiring Bachelor's Degree of Industrial Engineering at Universitas Islam Indonesia

Name : Delia Isti Astari
Student No. : 14 522 166

INTERNATIONAL PROGRAM
DEPARTMENT OF INDUSTRIAL ENGINEERING
UNIVERSITAS ISLAM INDONESIA
YOGYAKARTA

2018

AUTHENTICITY STATEMENT

In the name of Allah, I hereby certify that this research was conducted by me except the summaries which these are already mentioned the source and rewrite by myself. If someday the declaration letter is untrue and violates the legitimate regulations in the papers and intellectual property rights then I am willing to the diploma I received to be withdrawn by Universitas Islam Indonesia.

Yogyakarta, August 2018

Delia Isti Astari

THESIS APPROVAL OF SUPERVISOR

COMPARISON STUDY OF FUZZY C-MEANS AND FUZZY SUBTRACTIVE CLUSTERING IMPLEMENTATION IN QUALITY OF INDIHOME FIBER OPTIC NETWORK
(Case Study in PT. TELKOM INDONESIA)

THESIS

By

Name : **Delia Isti Astari**

Student Number : **14 522 166**

Yogyakarta, August 2018

Supervisor

(Muhammad Ridwan Andi Purnomo, ST., M.Sc., PhD.)

THESIS APPROVAL OF EXAMINATION COMMITTEE

COMPARISON STUDY OF FUZZY C-MEANS AND FUZZY SUBTRACTIVE CLUSTERING IMPLEMENTATION IN QUALITY OF INDIHOME FIBER OPTIC NETWORK
(Case Study in PT. TELKOM INDONESIA)

By

Name : Delia Isti Astari

Student Number : 14 522 166

Was defended before Examination Committee in Partial Fulfillment of the requirement for the bachelor degree of Industrial Engineering Department

Universitas Islam Indonesia

Examination Committee

Annisa Uswatun Khasanah, ST., M.Sc
Examination Committee Chair

Muhammad Ridwan Andi Purnomo, ST., M.Sc., PhD
Member I

Agus Mansur, ST., M.Eng.Sc
Member II

Acknowledged by,

Head of Department - Industrial Engineering

Faculty of Industrial Technology

Universitas Islam Indonesia

UNIVERSITAS ISLAM INDONESIA ★ YOGYAKARTA ★ INTERNATIONAL PROGRAM

(Muhammad Ridwan Andi Purnomo, ST., M.Sc., PhD.)

DEDICATION

This thesis is dedicated for the one and only my mom, my dad, my brother, and all my beloved family.

Thesis Supervisor,

Mr. Muhammad Ridwan Andi Purnomo, ST., M.Sc., PhD.

Best Friends and Industrial Engineering International Program UII Batch 2014 Family

MOTTO

“Allah does not charge a soul except [with that within] its capacity. It will have [the consequence of] what [good] it has gained, and it will bear [the consequence of] what [evil] it has earned. "Our Lord, do not impose blame upon us if we have forgotten or erred. Our Lord, and lay not upon us a burden like that which You laid upon those before us. Our Lord, and burden us not with that which we have no ability to bear. And pardon us; and forgive us; and have mercy upon us. You are our protector, so give us victory over the disbelieving people.”

(Qur’an Surah (QS) Al Baqarah verse 286)

PREFACE

Assalamualaikum warahmatullah wabarakatuh

Alhamdullillahirabbil'alamiin, praise the presence of Allah Almighty who has delegated all his mercy and grace, so that the writer can complete this Thesis in expected time accordance with Rasulullah SAW and his family, friends, and his followers who have fought and guided us out of the darkness to the bright way to reach the blessings of Allah SWT. Thanks to Allah SWT's grace, the thesis entitled "Comparison Study of Fuzzy C-Means and Fuzzy Subtractive Clustering Implementation for Quality of IndiHome Fiber Optic Network (Case Study in PT Telkom Indonesia)" can be solved well. This thesis is arranged as one of the requirements that must be fulfilled as The Partial Requirement of Acquiring Bachelor’s Degree of Industrial Engineering at Universitas Islam Indonesia.

In completing the preparation of this thesis can not be finished from the support, assistance, and guidance from various parties. For that the authors would like to thank and reward the parties who have provided support directly or indirectly, therefore with gratitude the authors thank to:

1. Both parents and all the big family member who always give prayers, encouragement, and love to me.
2. Mr. Muhammad Ridwan Andi Purnomo, ST., M.Sc., Ph.D. who has guided and provided solutions and suggestions in the completion of this thesis.
3. PT Telkom Indonesia branch Yogyakarta which has provided opportunities and facilities that have facilitated the author in completing the Thesis.
4. Head of Laboratory, Laboran, and entire family of Laboratorium Statistika Industri and Optimasi (SIOP) to my special bestfriend Citra, Feny, Dhaniya, Febri, Adi and Alfiqra.

5. Industrial Engineering International Program batch 2014 family and all those who have prayed for, supported and motivated during the writing of the final task that can not be mentioned one by one.
6. Mrs. Diana and Mrs. Devi that patiently help the students, especially the author.

The Authors also thanks to all of concerned parties that cannot be mentioned one by one who have helped the author in completing this report. Hopefully, the goods which are given by all parties to the Author will be replied by the kindness from Allah. Finally, the Author realizes that there are still shortcomings as well as weaknesses in this report, so the building suggestions and critics are fully expected. The author hopes this research would bring advantages for everyone who reads this.

Wassalamu'alaikum Warahmatullah Wabarakatuh

Yogyakarta, August 2018

Delia Isti Astari

ABSTRACT

This research is conducted for grouping or clustering the quality assessment rule of IndiHome fiber optic cable network using fuzzy clustering method in PT Telkom company and to understand the difference type of clustering by observing the mapping and clustering data results that presented by each algorithm method of Fuzzy Subtractive Clustering and Fuzzy C-Means Clustering results. It applied ten predictor variables that affect the quality of the system through the study of previous research literature. Several factors that affect the transmission are Tx Power, Rx Power, Temperature, Power Supply, and Bias Current. Later, cluster validation is performed by using Partition Coefficient Index (PCI) and Partition Coefficient Index (PEI) indicator. This research uses the Fuzzy Subtractive Clustering process with cluster radius is from 0.1 until 1. Each radius has each number of clusters, nevertheless, for radius 0.1 the number of clusters that formed are 4, while radius 0.2 to 1, there is only one cluster formed. In Fuzzy Subtractive Clustering, it is considering some of the parameter which are the accept ratio 0.5, the reject ratio 0.15, and squash factor 1.25. In Fuzzy C-Means result, the value of the PCI (Partition Coefficient Index) is 0.662786731. Then, the value of the PEI (Partition Entropy Index) is 0.546967522. From the results of Fuzzy Subtractive Clustering, highest value of PCI are resulted in radius 1 with the value of 0.451738. The smallest PEI is in radius 0.2 with the value of 0. 0.070139. Then, it can be stated that both methods are better within each parameter. But after considering the number of clusters that are formed, compared to fuzzy c-means method has 4 clusters and in fuzzy subtractive only two clustering numbers are formed, which are 41 and 1. In conclusion, the method that will be preferred in terms of grouping quality is Fuzzy C-Means.

Keywords: Quality, Fuzzy C-Means, Fuzzy Subtractive Clustering

TABLE OF CONTENT

AUTHENTICITY STATEMENT.. ii

THESIS APPROVAL OF SUPERVISOR iii

THESIS APPROVAL OF EXAMINATION COMMITTEE iv

DEDICATION.. v

MOTTO vi

PREFACE vii

ABSTRACT...ix

TABLE OF CONTENT x

LIST OF TABLES xii

LIST OF FIGURES xiii

CHAPTER I INTRODUCTION 1

1.1 Background 1

1.2 Problem Formulation 4

1.3 Objectives of Research 5

1.4 Scope of Problem 5

1.5 Benefits of Research 5

1.6 Systematical Writing 6

CHAPTER II LITERATURE REVIEW 8

2.1 Deductive Study 8

2.1.1 Data Mining Concept 8

2.1.2 Clustering Analysis 11

2.1.3 Fuzzy Logic 14

2.1.4 Fuzzy Clustering 17

2.2 Inductive Study 26

CHAPTER III RESEARCH METHODOLOGY 34

3.1 Research Flowchart 34

3.2 Problem Identification 35

3.3 Problem Formulation... 35

3.4 Literature Review 36

3.5 Data Collection 36
3.6 Pre-Processing Data 37
3.7 Data Processing 37
3.7.1 Fuzzy C-Means Processing 38
3.7.2 Fuzzy Subtractive Clustering Processing... 38
3.8 Clustering Validation 38
3.9 Discussion 39
3.10 Conclusion and Recommendation 39
CHAPTER IV DATA COLLECTING AND PROCESSING 40
4.1 Data Collection 40
4.2 Pre-Processing Data 45
4.3 Data Processing 46
4.3.1 Fuzzy C-Means Processing 46
4.3.2 Fuzzy C-Means Validation.. 55
4.3.3 Fuzzy Subtractive Clustering Processing... 58
4.3.4 Fuzzy Subtractive Clustering Validation... 73
CHAPTER V DISCUSSION 77
5.1 Grouping System Quality with Fuzzy C-Means Method 77
5.2 Grouping System Quality with Fuzzy Subtractive Clustering Method ... 79
5.3 Sensitivity Analysis on Fuzzy Subtractive Clustering 81
CHAPTER VI CONCLUSION AND RECOMMENDATION 84
6.1 Conclusion 84
6.2 Recommendation 85
6.2.1 For PT Telkom Indonesia branch Yogyakarta.................................... 85
6.2.2 For Further Researchers.. 85
REFFERENCES 86
APPENDICES 90

LIST OF TABLES

Table 1.1 Competition Map of Local and Foreign Telecommunication Industry in Indonesia 2
Table 2.1 Inductive Study 31
Table 4.1 Data Recapitulation IndiHome System 41
Table 4.2 Degree of Membership 49
Table 4.3 Clustering Result 50
Table 4.4 PCI Result on Fuzzy C-Means 55
Table 4.5 PEI Result on Fuzzy C-Means 57
Table 4.6 Normalization Data 60
Table 4.7 Initial Potential 65
Table 4.8 New Potential 67
Table 4.9 Normalization and Denormalization Data 70
Table 4.10 Sigma Cluster 71
Table 4.11 Degree of Membership using Radius 0.2 72
Table 4.12 PCI Result on Fuzzy Subtractive Clustering 73
Table 4.13 PEI Result on Fuzzy Subtractive Clustering 75
Table 5.1 Number of Cluster 80
Table 5.2 PCI and PEI value 82

LIST OF FIGURES

Figure 2.1 KDD Process.. 9
Figure 2.2 Fuzzy Logic.. 15
Figure 2.3 Gauss Curve Membership Function.. 24
Figure 3.1 Research Flowchart.. 34
Figure 4.1 Relationship between Objective Function with the Number of Iterations 48
Figure 4.2 Data Plot for Each Cluster in 4 Clusters.. 54
Figure 5.1 Sensitivity Analysis based on PCI and PEI... 82

CHAPTER I

INTRODUCTION

1.1 Background

It is important to consider that the processes and operations are often linked via process intra- and inter-relations to each other and thus, the variations can, even being tolerable from an individual (isolated) process perspective, lead to an unacceptable accumulation causing failure of the final product to meet the customer requirements (Wuest et al. 2013). An important things is the quality by both producers and consumers so it has a very important meaning for the survival of business activities in the field of services and manufacturing. Quality has become a demand of society in the era of global competition. Maintaining quality is important because; it can reduce costs. Companies by harvesting technological innovations can provide high quality and personalized service at reasonable costs (Rust & Miu, 2006). Companies that make quality as a strategy tool will have the advantage to compete against its competitors in the market because not all companies can achieve the superiority of quality. In this case the company is required to produce products with high quality, low price and timely delivery.

Under conditions of intense competition among telecommunication service providers, companies are required to improve the quality of services and products produced. With more choices in the market, consumers have a higher bargaining power in choosing products according to their needs. The competition in telecommunication industry in Indonesia is increasingly tight, in addition to competition among local telecommunication companies is also enlivened by the increasing number of foreign telecommunications companies entering Indonesia, where in general the area of competition is done with a variety of bonus facilities, cheap tariffs, and product differentiation offered (Wiryono & Suharto, 2008). This can be seen in Table 1.1. Nevertheless, each telecommunication service provider strives to concentrate on expanding the market.

Table 1.1 Competition Map of Local and Foreign Telecommunication Industry in Indonesia

Company	**Tech**	**License**
Telkomsel	GSM & 3G	Nation wide
Indosat		
Excel		
Natrindo	GSM	Regional
CAC		Not Operated Yet
Telkom	CDMA (Fixed Wireless)	Nation wide
Mobile-8		
Bakrie Tel		Regional
Telkom	Fixed Wireline	Nation wide
BBT		Limited Area
Sampoerna Tel	NMT-450	Regional

Source: (Wiryono & Suharto, 2008)

This research was conducted in PT. Telekomunikasi Indonesia Tbk which is one of the SOEs whose currently owned by the Government of Indonesia (52.56%), and 47.44% is owned by the Public, Bank of New York and Domestic Investors. PT. Telekomunikasi Indonesia Tbk, which is now better known as Telkom Group is the only State-Owned Enterprises telecommunication company and the largest telecommunication and network service provider in Indonesia. PT Telkom Indonesia with Speedy products that now changed to Indonesia Home (IndiHome) is the largest internet service provider in Indonesia, with relatively cheap for its monthly cost, this internet service is used by many customers all over Indonesia. For the IndiHome service using fiber optic, sometimes, issues will happen such as network break or the network becomes slow. Occasionally, customers complain about the presence of interruptions and a sudden drop in speed. The company will perform controlling and maintenance only when the customer propose the complaints. The company unaware on know the quality of the system for each customer. The company also unspecify the quality that should be understood by its employees that leads to the lack of preparedness to respond the complain on time.

Based on these problems, this research is conducted for grouping or clustering the quality system of IndiHome fiber optic cable network using fuzzy clustering method in PT Telkom company and to compare the prediction result or to evaluate the performance obtained by Fuzzy Subtractive Clustering and Fuzzy C-Means Clustering results. Clustering is one of the data mining functions that is used to group data into a class or cluster, so that objects on a cluster have a very large similarity with other objects on the same cluster, but it is very similar to other cluster objects (Tan, Steinbach, Karpatne, & Kumar, 2013). Fuzzy clustering techniques allow the automatic generation of fuzzy models and can be utilized to predict the quality. In fact, fuzzy modeling means more flexible modeling-by extending a zero-one membership in the interval (0.1), can be said to be more flexible (Takagi & Sugeno, 1985). Then using fuzzy modeling is simplifying the formulation of the problem as it reduces the cost of computing. This is due to the fact that the non-fuzzy (generally crisp) model generally produces a complete search in large space (since some key variables can only take values 0 and 1), whereas in the fuzzy model all variables are continuous, so the derivative can be calculated to find the direction for the search (Gorrostieta, Pedraza, & Carlos, 2005).

Finally fuzzy modeling can be an automated or semi-automated process using grouping techniques such as Fuzzy C-means Clustering (FCM) and Fuzzy Subtractive Clustering (FSC) (Yager & Filev, 1994). In this research, the data did not use the class label so therefore it is categorized as an unsupervised method. Where, for Fuzzy C-Means is an unsupervised method and Fuzzy Subtractive Clustering is a supervised method and each cluster center can be used as a rule base that describes system behavior. For fuzzy subtractive clustering also can be said as an unsupervised method. According to Yaqin, et al. (2018), fuzzy subtractive clustering method relatively unsupervised clustering method in which the number of cluster centers is unknown. Implementation of data mining algorithm using Fuzzy Subtractive Clustering and Fuzzy C-Means when being viewed from some previous researches can provide the best clustering data by some parameter. Both models have significant results and also have some differences in the shape and pattern of the cluster. Therefore a comparison test between two methods of data mining on both modeling is made to understand the different type of clustering. Particularly, for designated case study in analyzing the Quality of IndiHome Fiber Optic Network by considering the mapping and clustering data results from the clustering presented by each algorithm method.

At the same time, the sensitivity analysis is important to do in Fuzzy Subtractive Clustering with the range 0 – 1 because the output which produced by the different radius will have the variation in results. From the result, it can be seen that varying the cluster radius will obtain the different outputs. Analysis should be performed to examine the sensitivity due to the uncertainties result. FSC has an inconsistency problem where different way in running the FSC yields different results. Bataineh, Nadji & Saqer (2011) conducted the comparison for both methods was based on the validity measurement of their clustering method. The effects of different parameters on the performance of the algorithms are investigated. The parameter of validity measurement is Partition Coefficient Index (PCI) and Partition Entropy Index (PEI). Highly non-linear functions are modeled and a comparison is made between the two algorithms according to their capabilities of modeling. The number of clusters for the fuzzy c-mean algorithm is determined. The validity results are calculated for several cases. As for fuzzy subtractive clustering, the radius parameter is changed to obtain a different number of clusters. Generally, increasing the number of generated clusters yields an improvement in the validity index value. The optimal modeling results are obtained when the validity indices are on their optimal values.

1.2 Problem Formulation

Based on the background of research elaborated above, the problem formulation in this research are:

1. What is the result of the cluster validity performance value with the Partition Coefficient Index (PCI) and Partition Entropy Index (PEI) indicator produced by

Fuzzy C-Means and Fuzzy Subtractive Clustering in clustering the IndiHome quality system?

2. How is the sensitivity from testing the influence of a radius of 0.1 to 1 in the Fuzzy Subtractive Clustering method used in clustering the IndiHome quality system?

1.3 Objective of Research

In this section, the objectives in creating this research are revealed, as follows:

1. To identify the result of the validity performance value with the Partition Coefficient Index (PCI) and Partition Coefficient Index (PEI) indicator produced by Fuzzy C-Means and Fuzzy Subtractive Clustering in clustering the IndiHome quality system.
2. To find out how the sensitivity testing the influence of radius of 0.1 to 1 in the Fuzzy Subtractive Clustering method used in clustering the IndiHome quality system.

1.4 Scope of Problem

There are several limitations that existed in this research, as mentioned as follows:

1. The data used in quality measurement only a few variables.
2. This research did not examine the fuzzy clustering in depth until to get the final IF-THEN rule result.

1.5 Benefit of Research

The following are the benefits of this research:

1. The company can produce quality groupings so that they will be able to make different treatment from each cluster.
2. Comparative results can be used to see the performance differences of the Fuzzy Subtractive Clustering and Fuzzy C-Means Clustering methods.

1.6 Systematical Writing

The systematical writing in this study are:

CHAPTER I INTRODUCTION

This chapter explains the introduction of the research. In this chapter, there will be elaborated the problem background, problem formulation, research objective, scope of the problem, research benefit, and systematical writing.

CHAPTER II LITERATURE REVIEW

This chapter focuses to determine the current study of the related previous researches by finding the state of the art of the previous researches to make difference with other researches. In this chapter, there will be elaboration between inductive and deductive studies related to the topic. The chapter contains information about the result of related previous research and supporting literature underlying the research.

CHAPTER III RESEARCH METHODOLOGY

This chapter will describe the research methodology. In this chapter, there will be described the detailed series of research object, research flow, and method used for the research including data collecting, data processing, and analyzing method.

CHAPTER IV DATA COLLECTING AND PROCESSING

This chapter describes the data collection and processing, analysis and results, including images, graphics, and tables obtained. In addition, this chapter also explains thoroughly about the data processed using the aforesaid method. This chapter is a reference for the discussion of the results that will be written in Chapter V.

CHAPTER V DISCUSSION

This chapter contains the analysis about the result of the previous chapter. In this chapter, core discussion will be conducted in order to get a comprehensive understanding of the whole research.

CHAPTER VI CONCLUSION AND SUGGESTION

This chapter provides short and precise statements described in the previous chapter which answer the problem formulation of the research. Suggestion related to the current study in the purpose of the advancement of the future research is given based on the limitations of the current research.

REFERENCES

APPENDIX

CHAPTER II

LITERATURE REVIEW

2.1 Deductive Study

2.1.1 Data Mining Concept

Data Mining is a series of processes to explore the added value of information that has not been known manually from a database by extracting patterns from the data in order to manipulate data into more valuable information obtained by extracting and recognizing important patterns or pulling from the data contained in the database. Due to the wide variety of Data Mining techniques and many different types of information and forms of data presentation, it is necessary to define the limits of the applicability and relevance of certain methods according to the provided data and the achieved objectives. It is also necessary to understand how the problem should be solved with the Data Mining south as classification, regression, clustering and so on (Vadim, 2018). The main reason why data mining has attracted the attention of the information industry in recent years is because of the availability of large amounts of data and the increasing need for transforming the data into useful information and knowledge as it focussed on the field of science that is doing extracting or mining activities of the data size / large quantities, this information that will be very useful for development.

Data mining is also known by other names such as Knowledge discovery (mining) in databases (KDD), knowledge extraction, data analysis and business intelligence and is an important tool for manipulating data for presenting information as needed users with the aim to assist in the analysis of behavioral observation collections, in general the definition of data-mining can be interpreted as follows:

- The process of finding interesting patterns from large amounts of stored data.
- The extraction of useful or interesting information (non-trivial, implicit, as yet unknown potential use) pattern or knowledge of data stored in large sums.
- Exploration of automated or semi-automatic analysis of large amounts of data to search for meaningful patterns and rules.

Figure 2.1 below shows the process of Knowledge Discovery in Database. The phases of the process are as follows:

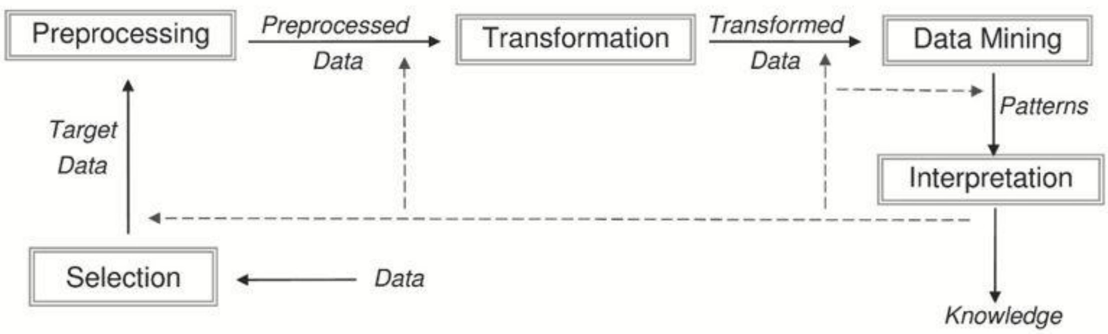

Figure 1.1 KDD Process

Source: Vannozzi, Croce, Starita, Benvenuti, & Cappozzo (2004)

1. Selection

- Creating a target data set, selecting a data set, or focusing on a subset of variables or sample data, where discovery will be performed.
- Data selection from a set of operational data needs to be done before the stage of extracting information in KDD begins. Selected data will be used for the data mining process, stored in a file, separate from the operational database.

2. Pre-processing

- Preliminary processing and data cleaning are basic operations such as noise removal.
- Before the data mining process can be implemented, it is necessary to do the cleaning process on the data that became the focus of KDD.
- The cleaning process includes removing data duplication, checking inconsistent data, and correcting data errors, such as typographical errors.
- Enrichment process is carried out, ie the process of existing data with other relevant data or information required for KDD, such as external data or information.

3. Transformation

- The search for useful features for presenting the data depends on the goal to be achieved.
- A process of transformation on the data that has been selected, so the data is appropriate for the process of data mining. This process is a creative process and depends on the type or pattern of information to be searched in the database.

4. Data mining

- Selection of data mining tasks; the selection of goals from the KDD process such as classification, regression, clustering, etc.
- Selection of data mining algorithm for searching.

- Data Mining process is the process of finding patterns or interesting information in selected data using a particular technique or method. Techniques, methods, or algorithms in data mining vary widely. The choice of the appropriate method or algorithm depends heavily on the purpose and process of KDD as a whole.

5. Interpretation / Evaluation

- Translation of patterns resulting from data mining.
- The pattern of information generated from the data mining process needs to be displayed in a form that is easily understood by interested parties.
- This stage is part of the KDD process that includes examining whether the pattern or information found is contrary to previous facts or hypotheses.

2.1.2 Clustering Analysis

A.1 Introduction

Clustering refers to the process of grouping samples so that the samples are similar within each group (Gose, Johnsonbaugh, & Jost, 2018). Clustering can be considered the most important unsupervised learning problem; so, as every other problem of this kind, it deals with finding a structure in a collection of unlabeled data. An example where this might be used is in the field of psychiatry, where the characterization of patients on the basis of clusters of symptoms can be useful in the identification of an appropriate form of therapy. In marketing, it may be useful to identify distinct groups of potential customers so that, for example, advertising can be appropriately targetted. Benefits of Clustering is a method of data segmentation that is very useful in predicting and analyzing certain business problems. For example market segmentation, marketing, and territorial mapping. Furthermore, the identification of objects in fields such as computer vision and image processing.

A good clustering will result in a high degree of commonality in one class and a low degree of commonality between classes. The similarity is a numerical measurement of two objects. The value of similarity between the two objects will be higher if the two objects are compared have a high similarity. The quality of clustering results depends on the method used. In clustering known four data types. The four data types are:

- Interval-scale variable
- Binary variables
- The nominal, ordinal, and ratio variables
- Variables with other types.

The clustering method should also be able to measure its own ability in an attempt to find a hidden pattern on the data under study. There are various methods that can be used to measure the value of similarity between the objects that are compared. One of them is the weighted Euclidean Distance. Euclidean distance calculates the distance of two points by knowing the value of each attribute on both points. Here's the formula used to calculate the distance with Euclidean distance in Equation 2.1:

$$Distance(p,q) = \left(\sum_{k}^{n} \mu k \, |Pk - qk| r\right) 1/r \qquad \text{... (2.1)}$$

Where:

n = Total of record data

k = Sequence of data fields

r = 2

μk = Weight field given a user

Distance is a common approach used to determine the similarity or inequality of two feature vectors expressed by rank. If the value of the resulting rank the smaller the value the closer or higher similarity between the two vectors. Distance measurement techniques with the Euclidean method become one of the most commonly used methods. Distance measurement with the euclidean method can be written in the Equation 2.2:

$$j(v_1, v_2) = \sqrt{\sum_{k=1}^{N} \left(v_1(k) - v_2(k)\right)^2} \quad \text{... (2.2)}$$

where v_1 and v_2 are two vectors whose distance will be calculated and N denotes the length of the vector.

A.2 Clustering Procedures

1. Non-hierarchical clustering (also called k-means clustering)

In this analysis, k number of clusters is chosen. Each of these clusters is assigned a centroid (or center). These initial centroids can be taken randomly, but it is important for researchers to recognize that different locations of the centroids can cause different results. Next, we determine the distance of each object to the nearest centroid. Then, we need to recalculate new centroids, which result from the clusters of the previous step. After we have these new centroids, we have to bind the points again to their nearest centroid. We group each object by a minimum distance to the centroid and continue doing so until we find convergence and stability (i.e. centroids do not move anymore). The goal of k-means cluster analysis is to minimize the summed distance between all data points and the cluster centroids. The diagram to the right explains the procedure. This process does not always find the optimal configuration and results can be easily affected by randomly selected centroids. Outliers may also have a strong effect on results.

2. Hierarchical clustering

In hierarchical clustering, objects are organized into a hierarchical structure as part of the procedure. We start with *n* total points and clusters each containing a single point (*n* total clusters). Then we look for the closest two clusters (using one of several distance measures explained above). This leaves us with *n-1* total clusters, with all but one containing a single element. We continue this process using the distance between cluster centroids. This agglomerative process uses a "bottom-up" strategy to cluster single elements into successively larger clusters. There are several methods for agglomerative clustering, including:

a. Centroid methods mean clusters are generated that maximize the distance between the centers of clusters (a centroid is a mean value for all the objects in the cluster).
b. Variance methods mean clusters are generated that minimize the within-cluster variance.
c. Ward's Procedure means clusters are generated that minimize the squared Euclidean distance to the center mean.
d. Linkage methods mean objects are clustered based on the distance between them.

2.1.3 Fuzzy Logic

Since 1985 when the fuzzy model methodology suggested by Takagi-Sugeno (Takagi & Sugeno, Fuzzy Identification of Systems and its Application to Modeling and Control, 1985), as well known as the TSK model, has been widely applied on theoretical analysis, control applications and fuzzy modeling. The fuzzy system needs the antecedent and consequence to express the logical connection between the input-data and output-data that is used as a basis to produce the desired system behavior (Sin & De, 1993). Fuzzy Logic is a troubleshooting methodology with thousands of applications in stored controls and information processing. Suitable to be implemented on a simple, small, embedded system on a microcontroller, multi-channel PC or workstation based data acquisition and control system. Figure 2.2 shows that the fuzzy logic provides a simple way to describe the exact conclusions of information that is ambiguous, vague, or incorrect. In a sense, fuzzy logic resembles human decision making with its ability to work from interpreted data and find the right solution.

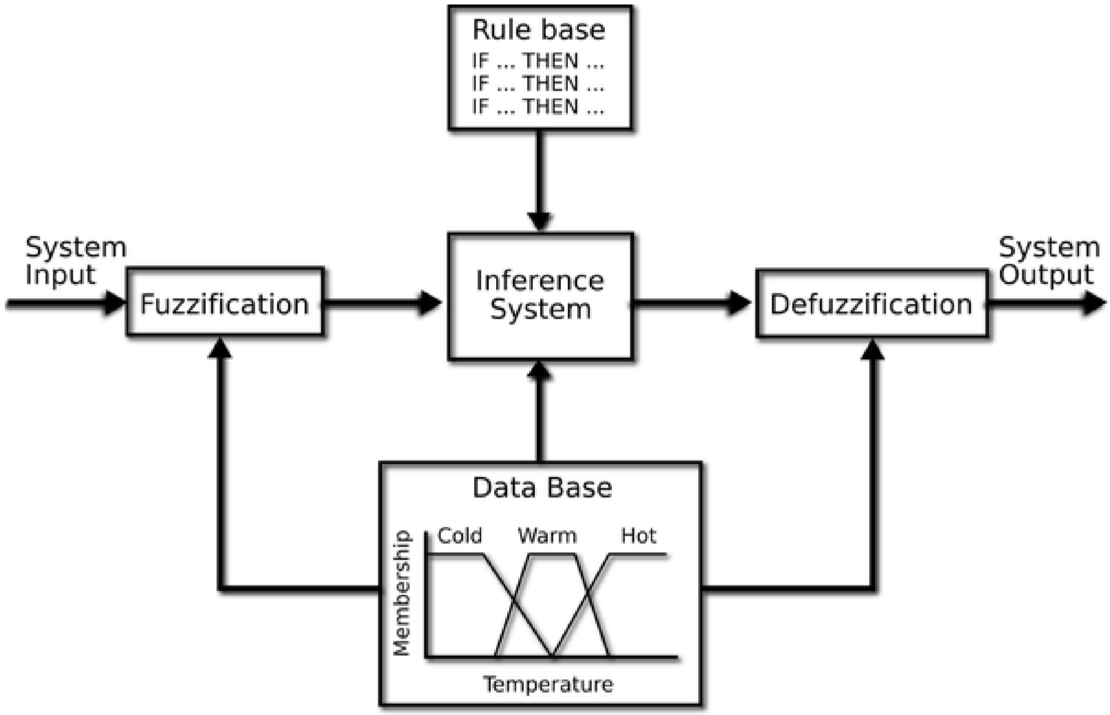

Figure 2.2 Fuzzy Logic

Source: Puspita & Yulianti (2016)

Fuzzy logic is basically a logical value that can define values between conventional states like yes or no, true or false, black or white, and so on. Fuzzy reasoning provides a way to understand the performance of the system by assessing the input and output system of observations. To do the design of a fuzzy system needs to do some of the following stages:

a. Defines model characteristics functionally and operationally.

In this section to note what characteristics of existing systems, then formulated the characteristics of operations that will be used on the fuzzy model.

b. Decomposition of model variables into fuzzy sets

From the variables that have been formulated, formed related fuzzy sets without overriding the domain.

c. Creating fuzzy rules

The rules on a fuzzy show how a system operates. The way of writing rules in general is: If (X_1 is A_1). .... (X_a is A_n) Then Y is B with (.) Is operator (OR or AND), X is scalar and A is linguistic variable.

Things to consider in creating rules are:

- Grouping all rules that have solutions on the same variable.
- Sorting rules for easy reading.
- Using an identity to show the rule structure.
- Using common naming to identify variables in different classes.
- Using comments to describe the purpose of a or a group of rules.
- Providing spaces between rules.
- Writing variable with big letters, fuzzy set with capital letters and other language elements with lowercase letters.

d. Define the defuzzy method for each solution variable

In the defuzzy stage, a value of a solution variable which is consequently selected from the fuzzy region is selected. The most commonly used method is the centroid method, this method has a high consistency, has a high and total width of a sensitive fuzzy area.

2.1.4 Fuzzy Clustering

A.1 Definition

Traditional clustering approaches generate partitions; in a partition, each instance belongs to one and only one cluster. Hence, the clusters in a hard clustering are disjointed. Fuzzy clustering, for instance, extends this notion and suggests a soft clustering schema. In this case, each pattern is associated with every cluster using some sort of membership function, namely, each cluster is a fuzzy set of all the patterns. Larger membership values indicate higher confidence in the assignment of the pattern to the cluster. The purpose of clustering is to identify natural groupings of data from a large data set to produce a concise representation of a system's behavior. A hard clustering can be obtained from a fuzzy partition by using a threshold of the membership value. The results of traditional clustering approaches are not appropriate to define clusters as modules in product design. Fuzzy clustering approaches can use fuzziness related to product design features and provide more useful solutions. Measuring result processing is proposed to be performed with a cluster analysis method enabling division of pooled data under consideration into groups of similar objects (clusters) and record distribution into different groups or segments.

Most clustering algorithms may be used under conditions of almost the whole unavailability of information on data distribution laws. Objects with quantitative (numerical), qualitative or mixed attributes are subject to clustering. Division of sampled information into groups of similar objects simplifies further data processing and decision-making as a specific analysis method may be used for each cluster. The clustering algorithm is the a function: X->Y that assigns to all x€X objects numbers of y€Y clusters. The Y range is known in advance in some cases but normally the objective is to determine an optimum cluster number in terms of the specified criterion of clustering quality. Membership grades are assigned to each of the data points. These membership grades indicate the degree to which data points belong to each cluster. Thus, points on the edge of a cluster, with lower membership grades, may be in the cluster to a lesser degree than points in the center of a cluster.

A.2 Fuzzy Clustering Algorithm Method

The different fuzzy clustering methods are described as follows.

1. Fuzzy C-Means Clustering Method

The most popular fuzzy clustering algorithm is the fuzzy c-means (FCM) algorithm. Even though it is better than the hard K-means algorithm at avoiding local minima, FCM can still converge to local minima of the squared error criterion (Elmzabi, Bellafkih, Ramdani, & Zeitouni, 2004). The fuzzy c-means algorithm attempts to partition a finite collection of elements X={ $x_1,x_2,...,x_n$} into a collection of c fuzzy clusters with respect to some given criterions. Fuzzy sets allow for degrees of membership. A single point can have partial membership in more than one class.

There can be no empty classes and no class that contains no data points. The output of such algorithms is a clustering, but not a partition sometimes (Nugraheni, 2013). This algorithm, data are a leap to every cluster by membership procedure, which represents the fuzzy performance of algorithms. The algorithm constructs a suitable matrix named U, factors are numbers between 0 and 1 also represent the level of membership among data and centers of clusters.

According to Gusti (2012), the earliest stages of Fuzzy C-Means concept were to determine the center of the cluster (centroid) that would identify the average location or space for each cluster. In the initial conditions, the center of this cluster cannot be accurately said this is caused by each data has a degree of membership for each cluster. Improvements to the central cluster (centroid) and each of the data values by repetition, it will be seen that the center of the cluster (centroid) will move closer to the correct space/location.

Based on the minimization of the rational function that represents the distance given to the centroid or cluster center of the data points by repairing the centroid and the membership value of each data repetitive or repetitive, the exact center position of the cluster (centroid) can be found. Fuzzy C-Means modeling stages from the beginning of the algorithm start as determining each cluster number, initial objective function, initial iteration, maximum iteration, rank, smallest expected error, generate random numbers, calculate the sum of each column and then calculate the center the kth cluster to produce the final data clustering. The output generated from Fuzzy C-Means (FCM) is a row of cluster centers and some degree of membership for each data point.

In this research, the development of a prediction model on Fuzzy C-Means Clustering method is done in 7 stages. The following is the development stage of the prediction model using the Fuzzy C-Means method (Prihatini, 2015):

1. Input data to be grouped, ie X is a matrix of size n x m (n = number of data samples, m = attribute of each data). X_{ij} the sample data to-i (i = 1,2, ..., N), j-attribute (j = 1,2, ..., m).
2. Determine:
   a. the number of clusters (c)
   b. ranks for the partition matrix (w)
   c. maximum iteration (maxIter)
   d. least expected error (ξ)
   e. initial objective function (Po = 0)
   f. and initial iteration (t = 1).
3. Generate random numbers μik, i = 1,2, ..., n; k = 1,2, ..., c as elements of the initial partition matrix U.

4. Calculate the center of the k-cluster: V_{kj}, with k = 1,2, ..., c; and j = 1,2, ..., m, using Equation 2.3 (Yan, Ryan, & Power, 1994):

$$V_{kj} = \frac{\sum_{i=1}^{n}((\mu_{ik})^{w}.X_{ij})}{\sum_{i=1}^{n}(\mu_{ik})^{w}} \quad \text{... (2.3)}$$

with:

V_{kj} = center of k-cluster for j-attribute

μ_{ik} = degree of membership for i-th sample data at the k-th cluster

X_{ij} = i-data, j-attribute

5. Compute the objective function on the t iteration using Equation 2.4 (Yan, Ryan, & Power, 1994):

$$P_t = \sum_{i=1}^{n}\sum_{k=1}^{c}\left(\left[\sum_{j=1}^{m}(X_{ij} - V_{kj})^2\right](\mu_{ik})^{w}\right) \quad \text{... (2.4)}$$

with:

V_{kj} = center of cluster to k for attribute to j

μ_{ik} = degree of membership for sample data to i on the k-th cluster

X_{ij} = i-data, j-attribute

P_t = objective function on the t iteration

6. Calculate the partition matrix change using Equation 2.5 (Yan, Ryan, & Power, 1994):

$$\mu_{ik} = \frac{\left[\sum_{j=1}^{m}(X_{ij} - V_{kj})^2\right]^{\frac{1}{w-1}}}{\sum_{k=1}^{c}\left[\sum_{j=1}^{m}(X_{ij} - V_{kj})^2\right]^{\frac{1}{w-1}}} \quad \text{... (2.5)}$$

with i = 1,2, ..., n; and k = 1,2, ..., c

Where :

V_{kj} = center of cluster to k for attribute to j

X_{ij} = data to i, attribute to j

μ_{ik} = degree of membership for sample data to i on cluster to k

7. Check stop condition:

If: ($|P_t - P_{t-1}| < \xi$ or ($t >$ MaxIter) then stop. If not: t = t +1, repeat step 4.

2. Fuzzy Subtractive Clustering Method

Clustering algorithms typically require the user to pre-specify the number of cluster centers and their initial locations. Estimated number and initial location of cluster centers in simple and effective algorithm is called as the mountain method. The method is based on gridding the data space and computing a potential value for each grid point based on its distances to the actual data points. A grid point with many data points nearby will have a high potential value. The grid point with the highest potential value is chosen as the first cluster center. The key idea in their method is that once the first cluster center is chosen, the potential of all grid points is reduced according to their distance from the cluster center. Grid points near the first cluster center will have greatly reduced potential. The next cluster center is then placed at the grid point with the highest remaining potential value. This procedure of acquiring new cluster center and reducing the potential of surrounding grid points repeats until the potential of all grid points falls below a threshold.

According to Chiu (1994), it uses data points as the candidates for cluster centers, instead of grid points as in mountain clustering. The computation for this technique is now proportional to the problem size instead of the problem dimension. The problem with this method is that sometimes the actual cluster centres are not necessarily located at one of the data points. However, this method provides a good approximation, especially with the reduced computation that this method offers. It also eliminates the need to specify a grid resolution, in which tradeoffs between accuracy and computational complexity must be considered. The subtractive clustering method also extends the mountain method's criterion for accepting and rejecting cluster centres. Although the Subtractive Clustering is fast, robust and accurate, the user-specified parameter (the radius of influence of cluster center) in this method, strongly affects the number of rules generated. A large generally results in fewer rules, while a small can produce immoderate number of clusters.

In the implementation, it can be used 2 fractions as a comparator factor, that is accept ratio and reject ratio. The accept ratio is the lower limit at which a point the data being candidate (candidate) cluster center is allowed to become the center of the cluster. While the reject ratio is the upper limit in which a data point becomes a candidate (candidate) cluster center is not allowed to become the center of the cluster. At an iteration, if it has been found a data point with the highest potential, then it will be continued by searching the potential ratio of that data point with the highest potential of a data point at the beginning of the iteration. For the development of prediction model on Fuzzy Subtractive Clustering method is done in 7 stages. The following is the development stage of the prediction model using the Fuzzy Subtractive method (Kusumadewi & Purnomo, 2013):

1. Input data to be clustered: X_{ij}, with i = 1,2, ... n; and j = 1,2, ... m.
2. Set value:
   a. rj (the radius of each data attribute); j = 1,2, ... m;
   b. q (squash factor);
   c. accept_ratio
   d. reject_ratio;
   e. XMin (minimum data allowed)
   f. XMax (maximum data allowed)
3. Normalization
   Calculate the normalization using Equation 2.6:
   $$X_{ij} = \frac{Xij - XMinj}{XMaxj - XMinj} \quad \text{... (2.6)}$$
   i = 1,2, ..., n; j = 1,2, ..., m
4. Determine the initial potential of each data point
   a. i = 1
   b. Do it up to i = n,
   1.) $Tj = X_{ij}$; j = 1,2, ..., m (2)
   2.) Calculate the $Dist_{kj}$ based on Equation 2.7:

$$\text{Dist}_{kj} = \left(\frac{Tj - Kkj}{ra}\right) \qquad \text{... (2.7)}$$

j = 1,2,....,m; k = 1,2,....n

3.) Initial potential

If m = 1, then follow the Equation 2.8:

$$D_i = \sum_{k=1}^{n} e^{-4(Dist^2 kj)} \qquad \text{... (2.8)}$$

If m > 1, then follow the Equation 2.9:

$$Di = \sum_{k=1}^{n} e^{-4(\sum_{j=1}^{m} Dist^2 kj)} \qquad \text{... (2.9)}$$

4.) i = i + 1

5. Find the point with the highest potential

a. M = max [In | i = 1,2, ..., n];

b. h = i, such that Di = M;

6. Determine the cluster center and reduce its potential to the surrounding points.

a. Center = []

b. $V_j = X_{hj}$; j = 1,2, ..., m;

c. C = 0 (number of clusters);

d. Condition = 1;

e. Z = M;

f. Do if (condition ≠ 0) and (Z ≠ 0):

1) Condition = 0 (there is no new center candidate yet);

2) Ratio = Z / M

3) If ratio> acceptance ratio, then condition = 1; (there is a new center candidate)

4) If not then the ratio> refusal ratio, (a new center candidate will be accepted as the center if its existence will provide balance to the data that is located far enough with the existing cluster center)

7. Return the cluster center from the normalized shape to the original shape.

$$\text{Center}_{ij} = \text{Center}_{ij} * (\text{XMax}_j - \text{Xmin}_j) + \text{XMin}_j; \qquad \text{... (2.10)}$$

8. Calculate the sigma value of the cluster using Equation 2.11:

$$\sigma_j = r_j * \left(\frac{XMaxj - XMinj}{\sqrt{8}}\right) \qquad \text{... (2.11)}$$

The result of this algorithm is the sigma value (σ) used to determine the parameter value of the fuzzy membership function. In this study used Gauss membership function as seen in Figure 2.3.

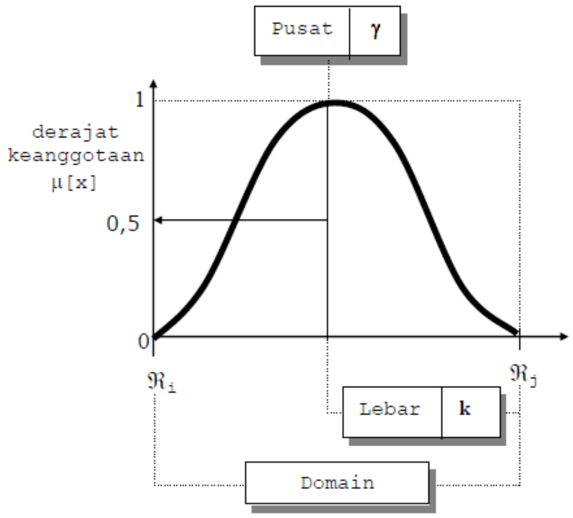

Figure 2.3 Gauss Curve Membership Function

Source: Fat (2014)

With the Gauss curve the membership degree of a data x_i in k-group is in Equation 2.12:

$$\mu_{ik} = e^{-\sum_{j=1}^{m} \frac{(x_{ij}-c_{kj})^2}{2\sigma_j^2}} \quad \text{... (2.12)}$$

From the description above can be seen that FSC has 4 parameters that are radius cluster, with upper acceptance limit and lower rejection limits and squash factor. These four parameters will affect the number of rules and error size (Kusumadewi, 2002).

a) Squash factor is used to multiply the radius value, in determining the center of the nearby cluster where its existence against the other center of the cluster will be reduced (default = 1.25).
b) Accept ratio is used to set the potential of each member to be the center of the cluster. If a member has a potential above the accept ratio then expected to be a cluster center (default = 0.5).

c) Reject ratio is used to set the potential of each member to be the center of the cluster. If any member has a potential under the reject ratio then the member will never be a cluster center (default = 0.15).

d) The cluster radius is used as the distance to be used in forming group members from each cluster. The higher the radius value then the number of clusters will be lesser, and dominant will generate a high error value.

A.3 Clustering Validation

Cluster analysis aims at identifying groups of similar objects and, therefore helps to discover distribution of patterns and interesting correlations in large data sets. However, it is a difficult problem, which combines concepts of diverse scientific fields (such as databases, machine learning, pattern recognition, statistics). Thus, the differences in assumptions and context among different research communities caused a number of clustering methodologies and algorithms to be defined (Halkidi, Batistakis, & Vazirgiannis, 2001). Validation includes efforts by the researcher to ensure that the cluster results are representative of the population in general and thus can be generalized to other objects and stable for a certain time.

Type of clustering validation are:

1. Partition Coefficient Index (PCI)

Bezdek (1981) proposes validity by calculating the partition coefficient (PC) as an evaluation of the value of data membership in each cluster. The PC Index value (PCI) only evaluates the degree of membership, regardless of the vector (data) value that usually contains geometric information (data distribution). The value of PCI is said to be able to measure the amount of overlapping between groups. The value in the range [0,1], the larger value (close to 1) means that the cluster quality is getting better. Here's the formula for calculating PC Index using Equation 2.13:

$$PCI= \frac{1}{N} (\sum_{i=1}^{N} \sum_{j=1}^{N} \mu ij2) \qquad ... (2.13)$$

Where N represents the amount of data in the data set, K represents the number of clusters, whereas u_{ij} denotes the membership value of the i data of the j-cluster.

2. Partition Entropy Index (PEI)

The partition entropy (PE) index is another fuzzy validity index that involves only the membership values. It is defined as Bezdek (1981) in Equation 2.14:

$$PEI = - \frac{1}{N} \sum_{l=1}^{c} \sum_{i=1}^{n} \mu_{li} log_{a}(\mu_{li}) \qquad ... (2.14)$$

Where a is the base of the logarithm and $U = (\mu_{li})$ is the membership matrix of a fuzzy c-partition. The values of the PE index range in [0, $\log_a$ c]. The closer the value of PEI to 0, the harder the clustering is. The values of PEI close to the upper bound indicate the absence of any clustering structure inherent in the data set of the inability of the algorithm to extract it. The PE index has the same drawbacks as the PC index.

2.2 Inductive Study

Fuzzy clustering is especially useful for fuzzy modeling especially in identifying fuzzy rules. Research on the application of Fuzzy Clustering method conducted by Ferarro & Giordani (2017) modifies fuzzy k-means clustering method for LR fuzzy data (PFkM-F). This paper focuses on robust clustering of data affected by imprecision. The clustering process is based on the fuzzy and possibilistic approaches. This has been done by a comparing the performance of PFkM-F with the ones of other related clustering methods for fuzzy data. The researchers have found that PFkM-F worked in a satisfactory way also in comparison with its competitors.

Other studies that apply the Fuzzy Clustering method are done by Zhu, Pedrycz, & Li (2017) using Particle Swarm Optimization (PSO) and Fuzzy K-Means. Two data transformation methods are proposed, Particle Swarm Optimization (PSO) is used to determine the optimal transformation realized on a basis of a certain performance index. Experimental studies completed for a synthetic data set and a number of data sets coming from the Machine Learning Repository demonstrate the performance of the FCM with transformed data. The experiments show that the proposed fuzzy clustering method achieves better performance (in terms of the clustering accuracy and the reconstruction error) in comparison with the outcomes produced by the generic version of the FCM algorithm.

Research by using Fuzzy Subtractive Clustering method by researcher Marzouk & Alaraby (2012) presented a fuzzy subtractive modelling technique to predict the weight of telecommunication towers which was used to estimate their respective costs. The towers considering four input parameters: tower height; allowed tilt or deflection; antenna subjected area loading; and wind load. Telecommunication towers were classified according to designated code (TIA-222-F and TIA-222-G standards) and structures type (Self-Supporting Tower (SST) and Roof Top (RT)). As such, four fuzzy subtractive models were developed to represent the four classes. Sensitivity analysis was carried to validate the model and observe the effect of clusters' radius on models performance.

Respati (2017) implemented forecasting optimization (STLF) using fuzzy subtractive clustering method (FSC). Characteristics of the load anomaly patterns showed inconsistency. Usually industrial activity stopped for a while and the workers took time off from work a few days. Parameter setting for short term load forecasting optimization (STLF) using fuzzy subtractive clustering method (FSC) which consisted of three input parameters, cluster radius or influence range and epoch. The optimization in this study was very influential in optimizing the value of forecasting accuracy that has not been optimized.

Another study conducted by Ramos, et al. (2017), using Noise Clustering, Density Oriented Fuzzy C-Means algorithms, Kernel Fuzzy C-Means, and Dierential Evolution algorithm. A design data driven based fault diagnosis systems using fuzzy clustering techniques was presented. As a rest part of the classifcation process, the data was pre-processed to eliminate outliers and reduce the confusion. To achieve this, the Noise Clustering and Density Oriented Fuzzy C-Means algorithms were used. Secondly, the Kernel Fuzzy C-Means algorithm was used to achieve greater separability among the classes, and reduce the classifcation errors. Finally, a third step is developed to optimize the two parameters used in the algorithms in the training stage using the Dierential Evolution algorithm.

Mittal & Suman (2014) conducted the research using k-Means Clustering, Hierarchical Clustering and Density. Data mining is covering every field of our life. In this paper, provided an overview of the comparison, classification of clustering algorithms. Under partitioning methods, applied k-means, and its variant k-medicine weka tool. Under hierarchical, discussed the two approaches which are the top-down approach and the bottom-up approach. The DBSCAN and OPTICS algorithms under the density based methods. The STING and CLIQUE algorithms under the grid based methods.

Next research was conducted by Tiwari & Yadav (2015) using Fuzzy Subtractive Clustering and ANFIS. Applicability and capability of Fuzzy Subtractive Clustering based approach to develop a prediction model prior to the implementation of the actual machining has been investigated. Subtractive Clustering is a fast one-pass algorithm for estimating the number of clusters and determining the cluster centres in a set of data . In all three input variable were used, consisting of Spindle Speed S, Feed rate F, and Depth of Cut DOC, and one output variable as tool vibration.

Pereira, et al. (2014) researched about Fuzzy Subtractive Clustering. This paper focused on demand response in a smart grid scope using a fuzzy subtractive clustering technique for modeling demand response. Domestic consumption was classified into profiles in order to favorable cover the adequate modeling. The fuzzy subtractive clustering technique was applied to a case study of domestic consumption demand response with three scenarios and a comparison of the results.

Radionov, Evdokimov, Sarlybaev, & Karandaeva (2015) conducted a study using Subtractive Clustering. A promising diagnostic condition control technique for the high-voltage oil-filled electrical facilities is the method of positioning partial discharges (PDs) and their intense measuring. The paper provideds outcome of experiments enabling acoustic PD positioning at the transformers of the power plant units. It considered the methods and algorithm of processing results of the periodical acoustic PD positioning based on the subtractive clustering technique.

Another study conducted by Rao, Sood, & Jarial (2015), using Subtractive Clustering. This paper helped in tuning and designing the membership functions that are best suited for the problem statement by integrating subtractive clustering method for fuzzy expert system design. The proposed integrated design of clustering based fuzzy expert system acted in improving the accuracy and leads to a précised decision making environment.

Ahmad & Dang (2015) conducted a study using some method which are Simple K-mean, DBSCAN, HCA and MDBCA. In this paper the four major clustering algorithms namely Simple K-mean, DBSCAN, HCA and MDBCA were compared to identify the performance of these four clustering algorithms. Performance of these four techniques were presented and compared using a clustering tool WEKA. The results were tested on different datasets namely Abalone, Bankdata, Router, SMS and Webtk dataset using WEKA interface and compute instances, attributes and the time taken to build the model.

Soni & Patel (2017) conducted a study using K-means and K-medoids Algorithm. In this paper, they strived to compare K-means and Kmedoids algorithms using the dataset of Iris plants from UCI Machine Learning Repository. The results obtained were in favour of K medoids algorithm owing to its ability to be better at scalability for the larger dataset and also due to it being more efficient than K-means. K-medoids showed its superiority over k means in execution time, sensitivity towards outlier data and to reduce the noise.

Another study conducted by Rani & Rohil (2013) conducted research by applying CURE, BIRCH, ROCK, CHEMELEON, Linkage, and Bisceting k-means. The quality of a pure hierarchical clustering method suffered from its inability to perform adjustment, once a merge or split decision has been executed. This paper presented an overview of improved hierarchical clustering algorithm. Hierarchical clustering is a method of cluster analysis which seeks to build a hierarchy of clusters.

Arumugadevi & Seenivasagam (2015) performed the research by implementing Fuzzy C-Means (FCM) clustering and Self Organizing Map (SOM). Image segmentation was the first step for any image processing based applications. The Conventional methods are unable to produce good segmentation results for color images. The researcher presented two soft computing approaches namely Fuzzy C-Means (FCM) clustering and Self Organizing Map (SOM) network were used to segment the color images. The segmentation results of FCM and SOM compared to the results of K-Means clustering. The results shown that the Fuzzy C-Means and SOM produced the better results than K-means for segmenting complex color images. The time required for the training of SOM was higher.

Chitra & Maheswari (2017) performed the research using partition-based algorithms, hierarchical based algorithms, and density-based algorithms. Clustering is a significant task in data analysis and data mining applications. Clustering algorithms can be classified into partition-based algorithms, hierarchical based algorithms, density-based algorithms and grid-based algorithms. This paper focused on a keen study of different clustering algorithms in data mining. In short, partitioning algorithms attempted to determine k clusters that optimize a certain, often distance-based criterion function.

After making the inductive study, the research position for the researches can be seen in Table 2.1 below:

Table 2.1 Inductive Study

No	Title	Author	Fuzzy C-Means	Hierarchical Clustering	Optimization	Fuzzy Subtractive	Noise Clustering	Density Algorithm	ANFIS	K-Medoids	Self Organizing Map
1	Possibilistic and fuzzy clustering methods for robust	Ferarro & Giordani, 2017	√	-	-	-	-	-	-	-	-
2	Fuzzy Clustering with Nonlinearly Transformed Data	Zhu, Pedrycz, & Li, 2017	√	-	√	-	-	-	-	-	-
3	Predicting Telecommunication Tower Costs Using Fuzzy Subtractive Clustering	Marzouk & Alaraby, 2012	-	-	-	√	-	-	-	-	-
4	The Impact of Influence Range Fuzzy Subtractive Clustering Modification to Accuracy Anomalous Load Forecasting	Respati, 2017	-	-	√	√	-	-	-	-	-
5	A novel fault diagnosis scheme applying fuzzy clustering	Ramos, et al., 2017	√	-	-	-	√	√	-	-	-

No	Title	Author	Fuzzy C-Means	Hierarchical Clustering	Optimization	Fuzzy Subtractive	Noise Clustering	Density Algorithm	ANFIS	K-Medoids	Self Organizing Map
6	Comparison and Analysis of Various Clustering Methods	Mittal & Suman, 2014	√	√	-	-	-	√	-	-	-
7	Fuzzy Subtractive Clustering Based Prediction Approach for Machine Tool Vibration	Tiwari & Yadav, 2015	-	-	-	√	-	-	√	-	-
8	Fuzzy subtractive clustering technique applied to demand response in a smart grid scope	Pereira, et al, 2014	-	-	-	√	-	-	-	-	-
9	Application of subtractive clustering for power transformer fault diagnostics	Radionov, et al, 2015	-	-	-	√	-	-	-	-	-
10	Subtractive clustering Fuzzy Expert System for Engineering Applications	Rao, Sood, & Jarial, 2015	-	-	-	√	-	-	-	-	-
11	Performance Evaluation of Clustering Algorithm Using Different Dataset	Ahmad & Dang, 2015	√	-	-	-	-	-	-	-	-

No	Title	Author	Fuzzy C-Means	Hierarchical Clustering	Optimization	Fuzzy Subtractive	Noise Clustering	Density Algorithm	ANFIS	K-Medoids	Self Organizing Map
12	Comparative Analysis of K-means and K-medoids	Soni & Patel, 2017	√	-	-	-	-	-	-	√	-
13	A Study of Hierarchical Clustering Algorithm	Rani & Rohil, 2013	√	√	-	-	-	-	-	-	-
14	Clustering Methods with Qualitative Data: A Mixed Methods Approach for Prevention Research	Arumugadevi & Seenivasagam, 2015	√	-	-	-	-	-	-	-	√
15	A Comparative Study of Various Clustering Algorithms in Data Mining	Chitra & Maheswari, 2017	-	√	-	-	√	√	-	-	-

Comparison to the Author Research

Title	Author	Fuzzy C-Means	Hierarchical Clustering	Optimization	Fuzzy Subtractive	Noise Clustering	Density Algorithm	ANFIS	K-Medoids	Self Organizing Map
Comparison Study of Fuzzy C-Means and Fuzzy Subtractive Clustering Implementation for Quality of IndiHome Fiber Optic Network PT Telkom Indonesia	Astari, 2018	√	-	-	√	-	-	-	-	-

CHAPTER III

RESEARCH METHODOLOGY

3.1 Research Flowchart

The research flowchart of this study is depicted in Figure 3.1 below:

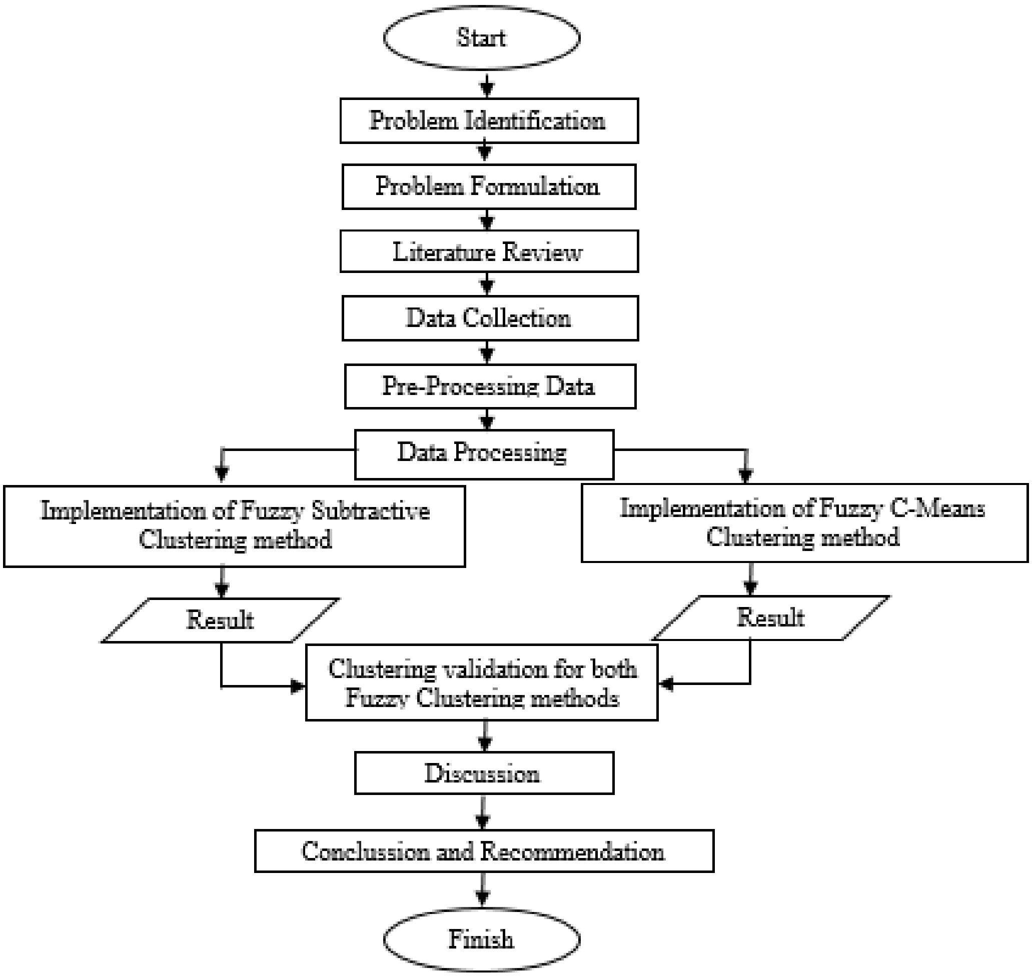

Figure 2.1 Research Flowchart

3.2 Problem Identification

Research is initiated by identifying the problems that exist within the concept of quality maintenance, especially on quality of telecommunication network. Consumer perceptions do not always result in the same judgment because not all consumers have full knowledge of the condition of the product or service, which will have an impact on IndiHome's buying interest. The company unaware on the quality of the system for each customer. The company also unclassify the quality to be understood by its employees which resulted in the lack of preparedness of the company to respond the complaints. This research is conducted for grouping or clustering the quality system of IndiHome fiber optic cable network using fuzzy clustering method in PT Telkom company and to evaluate the performance obtained on Fuzzy Subtractive Clustering and Fuzzy C-Means Clustering results.

3.3 Problem Formulation

From the problems found in the concept, furthermore, the formulation of the problem according to the problem identification is identified. From the above conditions that stated in problem identification the quality of IndiHome system PT Telkom Indonesia branch of Yogyakarta determined by using clustering method with Fuzzy C-Means and Fuzzy Subtractive Clustering based on predetermined criteria. Thus, the purpose of this study is to form the cluster group members of the IndiHome fiber optic cable network quality by implementing Fuzzy C-Means method and Fuzzy Subtractive Clustering.

3.4 Literature Review

The literature review in accordance with the discussion are gathered. Literature review consists of two types, deductive study and inductive study. Deductive studies are often known by theoretical studies derived from the theories of experts who are often used as a source of study. While the inductive study is a study derived from previous studies that can be used as a reference or comparison between previous studies with the research to avoid the existence of plagiarism.

3.5 Data Collection

Data collection methods are used to form the clustering group member of IndiHome quality by using fuzzy subtractive clustering algorithm compare with fuzzy c-means (FCM). The data were collected at PT. Telkom Indonesia (Yogyakarta) on March 2018. The data that collected from the company is the primary data. The primary data will be used for the main calculation and information for this research. The data were collected by using some variables to get the performance result of fuzzy clustering method. The type of data are:

a. Primary Data

Primary data in this study were obtained from direct observation by interviewing the manager of the PT TELKOM in Yogyakarta. Interview was conducted with several managers to identify the factors about the prediction of IndiHome network quality according to the clustering method used in this particular task as well as some other information on the operations. Then, historical data are is quantitative data obtained from the manager of PT TELKOM Yogyakarta that contain the quality variable covering some indicators or variables that affect the network performance such as Tx Power (the path through which to send data between devices), Rx Power (commonly called received which is useful to capture data transmitted by the transmitter or Tx Power), Temperature, Power Supply (component that supplies power at least one electric load), and Bias Current (establishing predetermined voltages or currents at various point of an electronic circuit for the purpose of establishing proper operating conditions in electronic components) from both ONT (Optical Network Unit) from customer system and OLT (Optical Network Termination) from company system.

b. Secondary Data

Secondary data obtained from literature-literature like journals, articles, explanation from experts to find the information about methods and problems on this particular task.

3.6 Pre-Processing Data

In the KDD (Knowledge Discovery in Database) stage, preprocessing data is the second stage in which a series of processes are used to clean up unnecessary data or if it is wrong to fit the purpose and data will be ready for processing. In order for the research process to run properly, the needs of data to be clustered need to be translated and converted into data form in accordance with the clustering method used is the algorithm with Fuzzy Subtractive Clustering modeling and Fuzzy C-Means modeling.

3.7 Data Processing

Preprocessing data are applied using tools or MATLAB and Microsoft Excel software by applying both clustering techniques using Fuzzy Subtractive Clustering modeling algorithm and using Fuzzy C-Means modeling.

3.7.1 Fuzzy C-Means Processing

The parameters required in the clustering process are the number of clusters (c), the rank (m), maximum iteration (MaxIter), the smallest expected error (x), the initial objective function (P0), and the initial iteration (t). The outputs resulting from the clustering process are the industries that fall into clusters (1 or 2 or 3 or 4) of different kinds of quality: excellent, good, bad, and very bad then made into 4 clusters according to the cluster number parameters. The first test is performed to get the minimum error value (ξ). The error value is obtained by calculating the difference of the objective function obtained on each iteration. The objective function will conclude as converged if the resulting value is constant, so the error (error) produced is worth 0. In this test, the cluster number parameter used is 4 with maximum 100 iterations.

3.7.2 Fuzzy Subtractive Clustering Processing

In the data processing, there are 4 parameters required for the formation of FIS model that is 3 parameters follow the standard provisions of squash factor of 1.25, accept ratio and reject ratio of 0.50 and 0.15 respectively and the radius value (r) used is a value from range 0 to 1 (to obtain optimal cluster number). The accept ratio is the lower bound in which a data point being a candidate cluster center is allowed to become a cluster center. While the reject ratio is the upper limit in which a data point being a candidate cluster center is not allowed to become a cluster center

3.8 Clustering Validation

After the clustering process is complete, the evaluation process proceeded by using PC index and PE index to obtain the value of cluster validity. After the whole process is complete, the output in the form of PC index and PE index values and quality groups will be acquired. The purpose of this test is to get the best value of c that has PC index and the best PE index value addressed in subsequent tests. The test of cluster number is done by comparing PC index and PE index values from different clusters for each variation of radius which have been used in range 0.1 to 1 in Fuzzy Subtractive Clustering, also in Fuzzy C-Means with the consideration of the number of clusters 4.

3.9 Discussion

After doing the data processing and obtaining the results with MATLAB software and Microsoft Excel, then analysis of result discussion will be resumed. Then, the result will be discussed by comparing the performance of Fuzzy Subtractive Clustering and Fuzzy C-Means Clustering by considering the value of the PC index and PE index value. On both artificial and real datasets, this algorithm is able, not only to determine the optimal number of clusters but also to provide better clustering partitions than standard algorithms.

3.10 Conclusion and Recommendation

The conclusion contains an explanation on the answer to the problem formulation that was set at the beginning of the study briefly. In addition, there are suggestions or recommendations that can be used by the hospitality and can also be used as further research material.

CHAPTER IV

DATA COLLECTING AND PROCESSING

4.1 Data Collection

Data are collected from IndiHome status system information contained in each of the existing ownership in the area of Yogyakarta in March 2018. The amount of data processed that have experienced pre-processing are 100 data. The data are divided into 2 categories namely:

a. OLT (Optical Network Termination)

   OLT is a device that becomes the endpoint, which is the root of an ODN. The function of OLT is to control information going to both ways and to be on a server at the head office. OLT is also called optical path termination, as a hardware endpoint device on passive optical networks. OLT will send ethernet data to ONU.

b. ONU (Optical Network Unit)

   Then, ONU (Optical Network Unit) or Optical Network Terminal (ONT) is a customer-side device that provides both data, voice, and video interfaces. The ONU's main function is to receive traffic in an optical format and convert it to the desired shape, such as data, voice, and video.

The number of variables used is as many as 10 variables with both of categories include data Tx Power, Rx Power, Temperature, Power Supply, and Bias Current as the factors related in determining the quality of the network based on interviews conducted with the manager of PT Telkom Yogyakarta. The data can be seen in Table 4.1. which are used as a reference for model building.

Table 4.1 Data Recapitulation IndiHome System

No	ONU - Optical Network Unit (from Customer)					OLT - Optical Network Termination (from Company)				
	Tx Power (dBm)	Rx Power (dBm)	Temperature (°C)	Power Supply (Volt)	Bias Current (mA)	Tx Power (dBm)	Rx Power (dBm)	Temperature (°C)	Power Supply (Volt)	Bias Current (mA)
1	2.31	-24.09	50	3.34	17	3.7	-13.882	37	3.24	15
2	2.686	-15.392	38.941	3.28	12.85	2.85	-14.136	43	3.135	30.678
3	2.1	-19.28	42	3.24	13	3.67	-13.324	44	3.2	17
4	2.35	-17.03	40	3.26	11	3.86	-12.754	42	3.18	10
5	2.24	-17.3	50	3.24	11	3.67	-12.672	51	3.21	13
6	2.108	-20.088	55.383	3.22	16.1	0	-19.146	0	0	0
7	1.98	-31.54	42	3.28	12	4.05	-15.604	35	3.19	8
8	1.93	-33.97	43	3.26	10	3.76	-15.796	37	3.26	14
9	2.02	-15.49	40	3.28	12	3.49	-12.184	48	3.2	12
10	2.05	-21.25	48	3.28	12	3.41	-13.568	48	3.19	11
11	2.3	-18.15	47	3.28	14	3.79	-12.918	32	3.32	11
12	2.28	-18.66	42	3.24	10	3.61	-12.544	47	3.17	17
13	2.08	-16.73	50	3.3	8	3.65	-12.76	22	3.2	9
14	2.18	-17.05	45	3.28	7	3.81	-12.902	22	3.2	9
15	2.27	-19.58	43	3.32	13	3.67	-13.532	51	3.19	18
16	2.02	-18.89	56	3.28	10	3.9	-13.162	33	3.2	9
17	2.31	-20.81	47	3.26	12	3.84	-13.46	40	3.23	11
18	2	-16.32	52	3.28	9	4.01	-12.614	39	3.21	9
19	2.556	-20.758	51.406	3.18	21.148	3.514	-21.621	45	3.142	29.644
20	2.37	-24.68	45	3.3	13	4.1	-14.648	35	3.17	9
21	2.28	-16.55	45	3.18	8	3.57	-12.902	42	3.21	11
22	2.12	-17.98	41	3.24	8	4.08	-13.184	48	3.17	13

No	ONU - Optical Network Unit (from Customer)					OLT - Optical Network Termination (from Company)				
	Tx Power (dBm)	Rx Power (dBm)	Temperature (^{o}C)	Power Supply (Volt)	Bias Current (mA)	Tx Power (dBm)	Rx Power (dBm)	Temperature (^{o}C)	Power Supply (Volt)	Bias Current (mA)
23	2.25	-17.54	48	3.24	11	3.67	-12.836	36	3.24	13
24	2.1	-19.46	47	3.24	10	3.93	-13.122	35	3.18	10
25	2.42	-19.47	42.77	3.24	12.2	3.46	-21.627	39.1	3.212	12.214
26	2.35	-16.23	51	3.3	8	3.68	-12.77	45	3.21	17
27	1.99	-18.79	45	3.28	11	3.76	-13.074	40	3.23	11
28	2.14	-18.07	49	3.26	8	3.7	-12.938	29	3.24	15
29	2.03	-18.41	53	3.28	9	3.63	-12.934	37	3.3	13
30	2.16	-22.84	53	3.28	18	3.75	-14.186	36	3.23	11
31	2.582	-13.8	53.609	3.22	19.3	3.248	-15.08	59	3.148	41.481
32	2.31	-19.706	53.254	3.22	17.9	3.648	-20.893	41.105	3	30.014
33	2.23	-19.79	45	3.28	6	3.8	-13.286	40	3.2	11
34	2.24	-16.14	48	3.3	7	3.72	-12.708	49	3.19	11
35	2.24	-16.14	48	3.3	7	3.72	-12.708	49	3.19	11
36	2.18	-18.32	45	3.28	12	3.31	-12.648	35	3.18	11
37	2.32	-19.83	44	3.32	10	3.48	-13.216	42	3.2	7
38	1.86	-18.86	43	3.24	9	3.76	-13.424	27	3.17	14
39	2.35	-18.09	46	3.26	13	3.35	-12.632	45	3.18	11
40	2.02	-21.3	54	3.26	10	0	-13.938	0	0	0
41	2.28	-17.3	51	3.3	8	3.66	-12.654	33	3.33	14
42	2.03	-19.43	40	3.2	9	3.73	-12.614	46	3.21	18
43	2.15	-17.98	46	3.3	8	3.64	-13.046	39	3.18	9
44	2.204	98.064	44.535	3.2	12.6	3.53	-21.434	52	3.176	32.354
45	2.35	-19.39	50	3.3	16	4.05	-12.648	39	3.21	8

No	ONU - Optical Network Unit (from Customer)					OLT - Optical Network Termination (from Company)				
	Tx Power (dBm)	Rx Power (dBm)	Temperature (°C)	Power Supply (Volt)	Bias Current (mA)	Tx Power (dBm)	Rx Power (dBm)	Temperature (°C)	Power Supply (Volt)	Bias Current (mA)
46	2.22	-19.17	52	3.28	9	3.75	-13.348	31	3.2	10
47	2.04	-21.94	50	3.3	7	3.75	-13.888	39	3.28	15
48	2.19	-18.09	63	3.28	10	3.64	-12.78	37	3.21	7
49	2.27	-26.57	44	3.32	11	3.66	-13.99	48	3.17	11
50	2.59	-24.09	52.902	3.18	14	3.789	-23.027	52.167	3	29.468
51	2.27	-20.75	53	3.28	9	3.42	-13.718	55	3.21	16
52	1.94	-18.01	47	3.26	12	3.75	-12.836	37	3.34	12
53	2.17	-18.29	52	3.28	8	3.9	-13.174	52	3.2	13
54	2.02	-14.95	53	3.28	10	3.73	-12.43	45	3.23	16
55	2.29	-18.18	50	3.22	10	4.03	-12.644	38	3.21	9
56	2.26	-19.7	43	3.28	11	3.95	-13.24	50	3.22	9
57	2.16	-17.85	54	3.3	7	3.69	-12.888	35	3.3	11
58	2.612	-23.098	46.656	3.2	16.1	0	-17.122	0	0	0
59	2.16	-21.25	36	3.28	7	3.96	-13.638	37	3.21	7
60	2.38	-13.65	47	3.32	14	4.11	-12.398	43	3.19	8
61	2.134	-18.014	48.102	3.2	9.7	0	-17.748	0	0	0
62	2.21	-20.97	46	3.3	14	3.69	-13.336	40	3.16	10
63	3.23	-23.98	54.574	3.26	16.866	0	-17.214	0	0	0
64	2.476	13.316	58.602	3.3	13.65	0	-23.962	0	0	0
65	2.31	-15.34	47	3.3	14	3.68	-12.43	47	3.21	12
66	2.25	-19.13	46	3.32	14	3.66	-13.274	40	3.16	10
67	2.394	-18.24	44.039	3.2	13.7	0	-14.922	0	0	0
68	2.34	-23.37	46	3.28	13	3.7	-13.912	38	3.2	8

No	ONU - Optical Network Unit (from Customer)					OLT - Optical Network Termination (from Company)				
	Tx Power (dBm)	Rx Power (dBm)	Temperature (^oC)	Power Supply (Volt)	Bias Current (mA)	Tx Power (dBm)	Rx Power (dBm)	Temperature (^oC)	Power Supply (Volt)	Bias Current (mA)
69	2.588	22.22	41.707	3.2	12.55	0	-16.99	0	0	0
70	2.89	-19.47	53.254	3.2	19.4	0	-14.934	0	0	0
71	2.14	-19.39	48	3.24	12	3.69	-13.054	48	3.21	12
72	2.05	-28.862	45.242	3.2	13.9	0	-17.8	0	0	0
73	2.504	-22.758	50.418	3.2	16.05	0	-20.704	0	0	0
74	2.07	-22.44	44	3.28	13	3.56	-13.99	42	3.2	7
75	2.364	-24.814	46.781	3.28	9.4	0	-22.68	0	0	0
76	2.338	-18.762	51.129	3.2	16.85	0	-26.022	0	0	0
77	2.33	-15.78	49	3.3	15	3.6	-12.714	36	3.2	8
78	3.184	-24.684	63.016	3.28	16.908	0	-17.056	0	0	0
79	2.178	98.064	44.535	3.2	12.45	3.473	-21.407	52	3.173	32.236
80	2.16	-16.34	47	3.24	11	3.67	12.814	46	3.21	9
81	2.13	-18.5	42	3.26	8	3.69	-12.992	37	3.19	11
82	2.27	-20.5	48	3.24	14	3.75	-13.228	38	3.32	12
83	3.42	-19.788	51.012	3.32	11.396	0	-23.768	0	0	0
84	2.538	-15.592	47.938	3.22	16.55	0	-23.012	0	0	0
85	2.15	-18.44	47	3.28	9	3.73	-13.46	36	3.28	14
86	1.94	-18.82	54	3.26	5	3.42	-13.44	56	3.21	17
87	2.02	-18.63	55	3.18	10	3.61	-12.976	48	3.16	12
88	2.32	-27.96	48	3.28	11	3.8	-15.31	28	3.2	6
89	2.172	-22.678	52.191	3.22	16.9	0	-18.894	0	0	0
90	3.202	-18.664	52.992	3.28	13.962	0	-24.89	0	0	0
91	2.28	-19.39	53	3.3	9	3.73	-13.52	38	3.26	14

No	ONU - Optical Network Unit (from Customer)					OLT - Optical Network Termination (from Company)				
	Tx Power (dBm)	Rx Power (dBm)	Temperature (^{o}C)	Power Supply (Volt)	Bias Current (mA)	Tx Power (dBm)	Rx Power (dBm)	Temperature (^{o}C)	Power Supply (Volt)	Bias Current (mA)
92	2.02	-17.37	52	3.24	7	3.87	-12.754	40	3.2	9
93	2.502	-16.576	47.363	3.22	16.65	0	-21.55	0	0	0
94	2.33	-23.46	45	3.2	9	3.82	-13.734	39	3.23	11
95	2.21	-24.95	45	3.3	13	3.97	-14.842	35	3.17	9
96	2.344	-18.182	44.891	3.22	13.85	0	-21.74	0	0	0
97	2.468	-20.058	55.602	3.3	10.214	0	-19.21	0	0	0
98	2.17	-18.99	50	3.24	13	3.83	-12.684	35	3.22	9
99	2.74	-20.606	49	3.18	15.85	0	-20.52	0	0	0
100	2.23	-18.69	52	3.28	10	3.64	-12.918	36	3.28	14

4.2 Pre-Processing Data

Preprocessing is the stage where the selection of data are processed and changed to be more structured. In this case, the preprocessing stages of data include:

a. Data Cleaning

Data cleaning is eliminating false data values, fixing data clutter and checking inconsistent data. The data that has been labeled as incomplete or lack of attribute values. Therefore, incomplete data is not used or discarded and replaced with new data.

b. Data Integration

Data integration are merged data from multiple sources. Combination of technical and business processes used to combine data from disparate sources into meaningful and valuable information.

4.3 Data Processing

At this stage, we will present the steps of applying clustering techniques to both Fuzzy C-Means and Fuzzy Subtractive Clustering methods using Matlab and Microsoft Excel software.

4.3.1 Fuzzy C-Means Processing

Clustering process is done using the input data, which are:

n	= 100 (there are 100 historical data quality system performances)
m	= 10 (there are ten criteria that are Tx Power, Rx Power, Temperature, Power Supply, and Bias Current in ONU and Tx Power, Rx Power, Temperature, Power Supply, and Bias Current in OLT)

According to Prihatini, P. M. (2015), the values used for parameter initialization are:

C	= 4 (4 clusters)
m	= 2 (ranks for the partition matrix)
MaxIter	= 100
ξ	= 0,000016 (least expected error)
P_0	= 0 (initial objective function is 0)
t	= 1 (initial iteration is 1)

By using MATLAB R2013a Software, the calculation result is center cluster or center, the degree of membership or matrix U and value of an objective function or ObjFcn. The first result is the result of functional value calculation, as follows:

X = load('data.dat');

[Center, U, ObjFcn]=fcm(X,4);

Iteration count = 1, obj. fcn = 24348.721886

Iteration count = 2, obj. fcn = 18293.255652

Iteration count = 3, obj. fcn = 17336.118489

Iteration count = 4, obj. fcn = 15924.139002

Iteration count = 5, obj. fcn = 14008.353018

Iteration count = 6, obj. fcn = 13275.303860

Iteration count = 7, obj. fcn = 13094.172823

Iteration count = 8, obj. fcn = 12922.304448

Iteration count = 9, obj. fcn = 12445.800347

Iteration count = 10, obj. fcn = 10861.606226

Iteration count = 11, obj. fcn = 7876.017889

Iteration count = 12, obj. fcn = 7263.455978

Iteration count = 13, obj. fcn = 7262.034144

Iteration count = 14, obj. fcn = 7261.993259

Iteration count = 15, obj. fcn = 7261.981248

Iteration count = 16, obj. fcn = 7261.977304

Iteration count = 17, obj. fcn = 7261.975987

Iteration count = 18, obj. fcn = 7261.975544

Iteration count = 19, obj. fcn = 7261.975395

Iteration count = 20, obj. fcn = 7261.975345

Iteration count = 21, obj. fcn = 7261.975328

Iteration count = 22, obj. fcn = 7261.975322

Interpretation, software MATLAB R2013a. It requires 22 iterations before obtaining the optimal solution for the functional value of 7261.975322. The iteration process stops at the 22nd iteration where the value $| Pt - Pt - 1 | < \xi$. To further illustrate it can be seen in Figure 4.1 graph relationship between objective function with the number of iterations in the figure below:

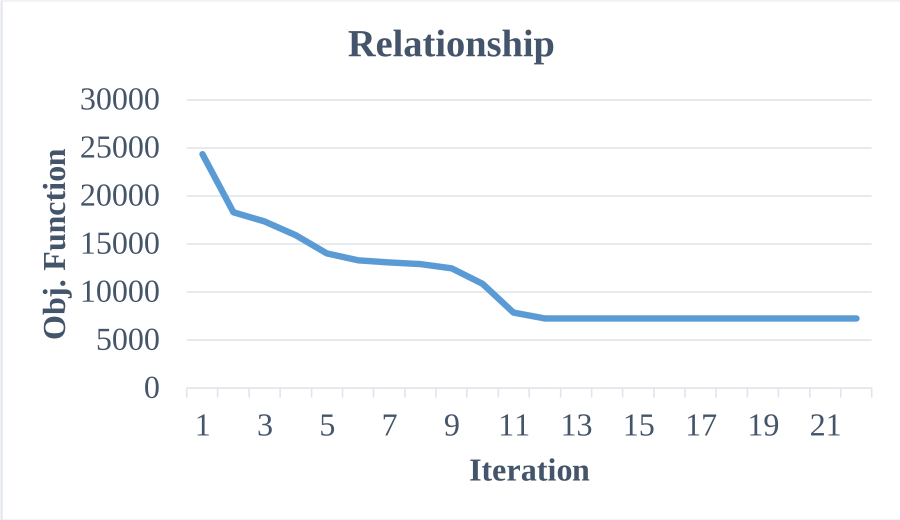

Figure 4.1 Relationship between Objective Function with the Number of Iterations

From the picture above can be seen that the value of the minimum objective function achieved by the iteration process as much as 22 times or function is converged with the iteration process as much as 22 times. The second result is the result of the calculation of v_{ij} values as follows:

Center =

1.0e+04 *

2.217	-18.603	47.078	3.264	10.857	3.646	-13.323	46.073	3.176	14.077
2.193	97.369	44.561	3.200	12.533	3.481	-21.388	51.699	3.156	32.102
2.554	-19.688	50.592	3.233	14.455	0.038	-19.695	0.335	0.033	0.139
2.196	-19.559	48.047	3.271	11.048	3.729	-13.387	36.675	3.207	11.343

At the last iteration (the 22nd iteration), the vkj cluster center produced by the software Matlab with k = 1,2,3,4; and j = 1,2,3,4,5,6,7,8,9,10 are:

$$\begin{bmatrix} 2.217 & -18.603 & 47.078 & 3.264 & 10.857 & 3.646 & -13.323 & 46.073 & 3.176 & 14.077 \\ 2.193 & 97.369 & 44.561 & 3.200 & 12.533 & 3.481 & -21.388 & 51.699 & 3.156 & 32.102 \\ 2.554 & -19.688 & 50.592 & 3.233 & 14.455 & 0.038 & -19.695 & 0.335 & 0.033 & 0.139 \\ 2.196 & -19.559 & 48.047 & 3.271 & 11.048 & 3.729 & -13.387 & 36.675 & 3.207 & 11.343 \end{bmatrix}$$

The second result is the degree of membership values. Fuzzy c-means has a membership degree that is useful for grouping data into appropriate clusters. The result shows as follows in Table 4.2:

Table 4.2 Degree of Membership

Data	µi1	µi2	µi3	µi4	Data	µi1	µi2	µi3	µi4
1	0.3046	0.0032	0.0295	0.6627	**38**	0.2313	0.0061	0.0842	0.6784
2	0.5377	0.0153	0.0667	0.3803	**39**	0.8173	0.001	0.0064	0.1753
3	0.7308	0.0023	0.0137	0.2532	**40**	0.0262	0.0035	0.9294	0.0409
4	0.5275	0.0033	0.0223	0.4469	**41**	0.1786	0.0025	0.0254	0.7935
5	0.8463	0.0022	0.0108	0.1406	**42**	0.724	0.0036	0.0194	0.2529
6	0.0104	0.0015	0.9722	0.016	**43**	0.2366	0.0014	0.0119	0.7501
7	0.3276	0.0067	0.0771	0.5886	**44**	0	0.9999	0	0
8	0.3866	0.0074	0.0712	0.5348	**45**	0.2642	0.0022	0.0199	0.7137
9	0.7331	0.0039	0.0197	0.2433	**46**	0.1589	0.003	0.0381	0.8001
10	0.847	0.0013	0.0077	0.144	**47**	0.3376	0.002	0.0158	0.6447
11	0.1316	0.002	0.0238	0.8426	**48**	0.3526	0.0094	0.0833	0.5547
12	0.8183	0.0022	0.0114	0.1682	**49**	0.6733	0.0037	0.0229	0.3001
13	0.2193	0.0094	0.2049	0.5664	**50**	0.5645	0.0168	0.0692	0.3494
14	0.2272	0.0098	0.2	0.563	**51**	0.728	0.0064	0.0277	0.2379
15	0.8019	0.0036	0.0165	0.178	**52**	0.0565	0.0004	0.0032	0.94
16	0.2201	0.0042	0.0483	0.7274	**53**	0.7758	0.0038	0.0181	0.2022
17	0.2176	0.0008	0.0064	0.7752	**54**	0.703	0.0029	0.0163	0.2778
18	0.2716	0.0022	0.0175	0.7087	**55**	0.126	0.0009	0.008	0.8651
19	0.517	0.0162	0.0777	0.3891	**56**	0.7635	0.0032	0.0168	0.2165
20	0.195	0.0024	0.0275	0.775	**57**	0.221	0.003	0.0296	0.7464
21	0.5598	0.0018	0.012	0.4265	**58**	0.0146	0.0019	0.9608	0.0227
22	0.7779	0.0029	0.0149	0.2043	**59**	0.3717	0.0068	0.0598	0.5617
23	0.0672	0.0005	0.0046	0.9277	**60**	0.5289	0.0033	0.0215	0.4463
24	0.0467	0.0005	0.0047	0.9481	**61**	0.0145	0.002	0.9609	0.0226
25	0.4062	0.0041	0.0332	0.5565	**62**	0.2799	0.0013	0.0112	0.7076
26	0.7582	0.0022	0.0124	0.2272	**63**	0.018	0.0024	0.9519	0.0277
27	0.2914	0.001	0.0082	0.6993	**64**	0.1754	0.0577	0.5411	0.2258
28	0.2052	0.0044	0.054	0.7364	**65**	0.8274	0.0017	0.009	0.1619
29	0.2095	0.0018	0.0158	0.7728	**66**	0.2826	0.0013	0.0104	0.7056
30	0.27	0.0038	0.0395	0.6867	**67**	0.0269	0.0037	0.9277	0.0417
31	0.5073	0.0424	0.1021	0.3482	**68**	0.2103	0.0018	0.0175	0.7704
32	0.4887	0.0149	0.08	0.4164	**69**	0.2059	0.0888	0.4522	0.2531
33	0.3716	0.002	0.0151	0.6113	**70**	0.0209	0.0029	0.944	0.0322
34	0.8058	0.0024	0.0123	0.1795	**71**	0.9182	0.0007	0.004	0.077
35	0.8058	0.0024	0.0123	0.1795	**72**	0.0419	0.0053	0.8886	0.0642
36	0.099	0.001	0.0096	0.8905	**73**	0.0053	0.0007	0.9859	0.0081
37	0.4433	0.0024	0.0182	0.5361	**74**	0.4315	0.0027	0.0219	0.5439

Data	μi1	μi2	μi3	μi4	Data	μi1	μi2	μi3	μi4
75	0.0282	0.0037	0.9249	0.0432	**88**	0.2343	0.0067	0.122	0.6371
76	0.0176	0.0026	0.9535	0.0264	**89**	0.0073	0.001	0.9803	0.0113
77	0.1986	0.0024	0.0234	0.7756	**90**	0.0132	0.0019	0.9648	0.02
78	0.0603	0.0085	0.8409	0.0902	**91**	0.2601	0.0019	0.0159	0.7221
79	0	0.9999	0	0	**92**	0.3341	0.0025	0.0194	0.644
80	0.4609	0.0224	0.1003	0.4164	**93**	0.0113	0.0017	0.9698	0.0172
81	0.2672	0.0024	0.0204	0.71	**94**	0.2697	0.0016	0.0141	0.7146
82	0.1227	0.0007	0.0063	0.8703	**95**	0.2017	0.0026	0.029	0.7667
83	0.0106	0.0015	0.9716	0.0162	**96**	0.0156	0.0022	0.9583	0.0239
84	0.0152	0.0023	0.9594	0.0231	**97**	0.017	0.0024	0.9545	0.0262
85	0.1172	0.0009	0.0078	0.8741	**98**	0.0925	0.001	0.0111	0.8954
86	0.6834	0.0093	0.0365	0.2708	**99**	0.0025	0.0003	0.9932	0.0039
87	0.6976	0.0036	0.0198	0.279	**100**	0.1637	0.0015	0.0133	0.8216

The output of Fuzzy C-Means is a central cluster and some degree of membership for each data point. This will provide information on the similarity of each object. One of fuzzy clustering algorithms used is the fuzzy clustering c means algorithm. The vector of fuzzy clustering, V = {v_1, v_2, v_3, ..., vc}, is an objective function that is defined by the degree of membership of the data X_j and the center of cluster V_j. Fuzzy clustering is the process of determining the degree of membership. This information can be used to build a fuzzy inference system. The degree of membership refers to how likely a data can be a member of a cluster. The position and value of the matrix are constructed randomly. Where the value of the membership lies on the interval 0 to 1. An IndiHome system has a certain degree of membership to become a member of a cluster. Certainly, the greatest degree of membership shows the highest tendency of a cluster system to enter a cluster member. The degree of membership of each system in each cluster is shown in the following Table 4.3.

Table 4.3 Clustering Result

Data	Degree of membership on cluster				Data tends to get into the cluster			
	1	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1	0.3046	0.0032	0.0295	0.6627				*
2	0.5377	0.0153	0.0667	0.3803	x			
3	0.7308	0.0023	0.0137	0.2532	x			
4	0.5275	0.0033	0.0223	0.4469	x			
5	0.8463	0.0022	0.0108	0.1406	x			
6	0.0104	0.0015	0.9722	0.016			+	

Data	Degree of membership on cluster				Data tends to get into the cluster			
	1	2	3	4	1	2	3	4
7	0.3276	0.0067	0.0771	0.5886				*
8	0.3866	0.0074	0.0712	0.5348				*
9	0.7331	0.0039	0.0197	0.2433	x			
10	0.847	0.0013	0.0077	0.144	x			
11	0.1316	0.002	0.0238	0.8426				*
12	0.8183	0.0022	0.0114	0.1682	x			
13	0.2193	0.0094	0.2049	0.5664				*
14	0.2272	0.0098	0.2	0.563				*
15	0.8019	0.0036	0.0165	0.178	x			
16	0.2201	0.0042	0.0483	0.7274				*
17	0.2176	0.0008	0.0064	0.7752				*
18	0.2716	0.0022	0.0175	0.7087				*
19	0.517	0.0162	0.0777	0.3891	x			
20	0.195	0.0024	0.0275	0.775				*
21	0.5598	0.0018	0.012	0.4265	x			
22	0.7779	0.0029	0.0149	0.2043	x			
23	0.0672	0.0005	0.0046	0.9277				*
24	0.0467	0.0005	0.0047	0.9481				*
25	0.4062	0.0041	0.0332	0.5565				*
26	0.7582	0.0022	0.0124	0.2272	x			
27	0.2914	0.001	0.0082	0.6993				*
28	0.2052	0.0044	0.054	0.7364				*
29	0.2095	0.0018	0.0158	0.7728				*
30	0.27	0.0038	0.0395	0.6867				*
31	0.5073	0.0424	0.1021	0.3482	x			
32	0.4887	0.0149	0.08	0.4164	x			
33	0.3716	0.002	0.0151	0.6113				*
34	0.8058	0.0024	0.0123	0.1795	x			
35	0.8058	0.0024	0.0123	0.1795	x			
36	0.099	0.001	0.0096	0.8905				*
37	0.4433	0.0024	0.0182	0.5361				*
38	0.2313	0.0061	0.0842	0.6784				*
39	0.8173	0.001	0.0064	0.1753	x			
40	0.0262	0.0035	0.9294	0.0409			+	
41	0.1786	0.0025	0.0254	0.7935				*
42	0.724	0.0036	0.0194	0.2529	x			
43	0.2366	0.0014	0.0119	0.7501				*
44	0	0.9999	0	0		o		
45	0.2642	0.0022	0.0199	0.7137				*
46	0.1589	0.003	0.0381	0.8001				*

Data	Degree of membership on cluster				Data tends to get into the cluster			
	1	2	3	4	1	2	3	4
47	0.3376	0.002	0.0158	0.6447				*
48	0.3526	0.0094	0.0833	0.5547				*
49	0.6733	0.0037	0.0229	0.3001	x			
50	0.5645	0.0168	0.0692	0.3494	x			
51	0.728	0.0064	0.0277	0.2379	x			
52	0.0565	0.0004	0.0032	0.94				*
53	0.7758	0.0038	0.0181	0.2022	x			
54	0.703	0.0029	0.0163	0.2778	x			
55	0.126	0.0009	0.008	0.8651				*
56	0.7635	0.0032	0.0168	0.2165	x			
57	0.221	0.003	0.0296	0.7464				*
58	0.0146	0.0019	0.9608	0.0227			+	
59	0.3717	0.0068	0.0598	0.5617				*
60	0.5289	0.0033	0.0215	0.4463	x			
61	0.0145	0.002	0.9609	0.0226			+	
62	0.2799	0.0013	0.0112	0.7076				*
63	0.018	0.0024	0.9519	0.0277			+	
64	0.1754	0.0577	0.5411	0.2258			+	
65	0.8274	0.0017	0.009	0.1619	x			
66	0.2826	0.0013	0.0104	0.7056				*
67	0.0269	0.0037	0.9277	0.0417			+	
68	0.2103	0.0018	0.0175	0.7704				*
69	0.2059	0.0888	0.4522	0.2531			+	
70	0.0209	0.0029	0.944	0.0322				*
71	0.9182	0.0007	0.004	0.077	x			
72	0.0419	0.0053	0.8886	0.0642			+	
73	0.0053	0.0007	0.9859	0.0081			+	
74	0.4315	0.0027	0.0219	0.5439				*
75	0.0282	0.0037	0.9249	0.0432			+	
76	0.0176	0.0026	0.9535	0.0264			+	
77	0.1986	0.0024	0.0234	0.7756				*
78	0.0603	0.0085	0.8409	0.0902			+	
79	0	0.9999	0	0		o		
80	0.4609	0.0224	0.1003	0.4164	x			
81	0.2672	0.0024	0.0204	0.71				*
82	0.1227	0.0007	0.0063	0.8703				*
83	0.0106	0.0015	0.9716	0.0162			+	
84	0.0152	0.0023	0.9594	0.0231			+	
85	0.1172	0.0009	0.0078	0.8741				*
86	0.6834	0.0093	0.0365	0.2708	x			

Data	Degree of membership on cluster				Data tends to get into the cluster			
	1	2	3	4	1	2	3	4
87	0.6976	0.0036	0.0198	0.279	x			
88	0.2343	0.0067	0.122	0.6371				*
89	0.0073	0.001	0.9803	0.0113			+	
90	0.0132	0.0019	0.9648	0.02				*
91	0.2601	0.0019	0.0159	0.7221				*
92	0.3341	0.0025	0.0194	0.644				*
93	0.0113	0.0017	0.9698	0.0172			+	
94	0.2697	0.0016	0.0141	0.7146				*
95	0.2017	0.0026	0.029	0.7667				*
96	0.0156	0.0022	0.9583	0.0239			+	
97	0.017	0.0024	0.9545	0.0262			+	
98	0.0925	0.001	0.0111	0.8954				*
99	0.0025	0.0003	0.9932	0.0039			+	
100	0.1637	0.0015	0.0133	0.8216				*

From the degree of membership in the last iteration can be obtained information on the tendency for each observation goes into which cluster. The greatest degree of membership shows that the highest tendency of observation to enter into a particular cluster member. At the first observation, the membership degree value for the first cluster is 0.3046 while the membership degree value for the second cluster is 0.0032. Then, the membership degree value for the third cluster is 0.0295. After that, the membership degree value for the fourth cluster is 0.6627. From that value, the first observation entered in the fourth cluster. That's because the first observation has the highest degree of membership in the fourth cluster rather than the other cluster.

Furthermore, on the second observation the value of membership degrees for the first cluster is 0.5377. While the degree of membership for the second cluster is 0.0153. Then, the membership degree value for the third cluster is 0.0667. After that, the membership degree value for the fourth cluster is 0.3803. From that value, the second observation entered in the first cluster.

The determination continues until the 100th observation, with a membership degree value for the first cluster of 0.1637 while the membership degree value for the second cluster is 0.0015. Then, the membership degree value for the third cluster is

0.0133. After that, the membership degree value for the fourth cluster is 0.8216. From that value, the 100th observation entered in the fourth cluster. Figure 4.2 shows the plot of data for each cluster in 4 clusters.

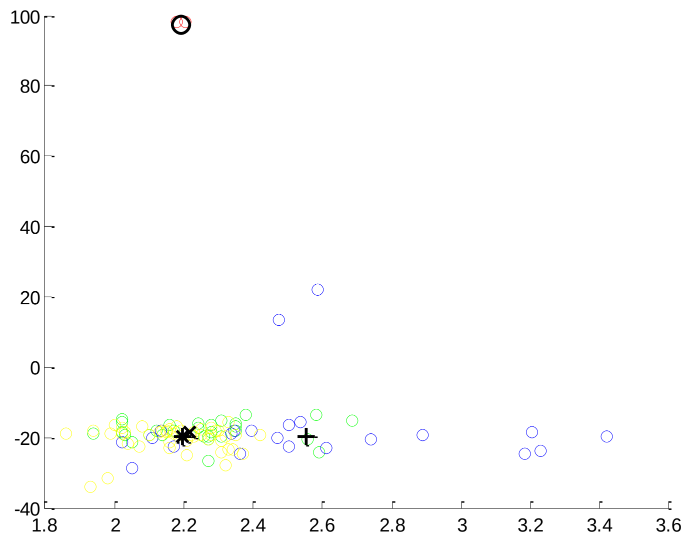

Figure 4.2 Data Plot for Each Cluster in 4 Clusters

Information:

x = cluster 1

o = cluster 2

+ = cluster 3

* = cluster 4

The end result of the clustering of 100 IndiHome system quality data with ten criteria are generate into 4 clusters as follows:

a. Group 1 (cluster 1), contains some of the data with the number 6, 40, 58, 61, 63, 64, 67, 69, 70, 72, 73, 75, 76, 78, 83, 84, 89, 90, 93, 96, 97, and 99
b. Group 2 (cluster 2), contains some of the data with the number 2, 3, 5, 9, 10, 12, 15, 19, 22, 26, 31, 32, 34, 35, 39, 42, 44, 49, 50, 51, 53, 54, 56, 65, 71, 79, 80, 86, and 87
c. Group 3 (cluster 3), contains some of the data with the number 1, 11, 13, 14, 16, 18, 23, 24, 28, 29, 30, 36, 38, 41, 46, 47, 48, 55, 57, 77, 85, 88, 91, 92, 98, and 100

d. Group 4 (cluster 4), contains some of the data with the number 4, 7, 8, 17, 20, 21, 25, 27, 33, 37, 43, 45, 52, 59, 60, 62, 66, 68, 74, 81, 82, 94, and 95

4.3.2 Fuzzy C-Means Validation

Next, the clustering validity on Fuzzy C-Means method is carried out with 2 indicators as follows.

a. Partition Coefficient Index (PCI)

To calculate the value of PCI, the Equation 2.13 is employed. The result of PCI validation on Fuzzy C-Means method is shown in Table 4.4.

Table 4.4 PCI Result on Fuzzy C-Means

Data	$\mu i1^2$	$\mu i2^2$	$\mu i3^2$	$\mu i4^2$	Data	$\mu i1^2$	$\mu i2^2$	$\mu i3^2$	$\mu i4^2$
1	0.0928	0.0000	0.0009	0.4392	**22**	0.6051	0.0000	0.0002	0.0417
2	0.2891	0.0002	0.0044	0.1446	**23**	0.0045	0.0000	0.0000	0.8606
3	0.5341	0.0000	0.0002	0.0641	**24**	0.0022	0.0000	0.0000	0.8989
4	0.2783	0.0000	0.0005	0.1997	**25**	0.1650	0.0000	0.0011	0.3097
5	0.7162	0.0000	0.0001	0.0198	**26**	0.5749	0.0000	0.0002	0.0516
6	0.0001	0.0000	0.9452	0.0003	**27**	0.0849	0.0000	0.0001	0.4890
7	0.1073	0.0000	0.0059	0.3464	**28**	0.0421	0.0000	0.0029	0.5423
8	0.1495	0.0001	0.0051	0.2860	**29**	0.0439	0.0000	0.0002	0.5972
9	0.5374	0.0000	0.0004	0.0592	**30**	0.0729	0.0000	0.0016	0.4716
10	0.7174	0.0000	0.0001	0.0207	**31**	0.2574	0.0018	0.0104	0.1212
11	0.0173	0.0000	0.0006	0.7100	**32**	0.2388	0.0002	0.0064	0.1734
12	0.6696	0.0000	0.0001	0.0283	**33**	0.1381	0.0000	0.0002	0.3737
13	0.0481	0.0001	0.0420	0.3208	**34**	0.6493	0.0000	0.0002	0.0322
14	0.0516	0.0001	0.0400	0.3170	**35**	0.6493	0.0000	0.0002	0.0322
15	0.6430	0.0000	0.0003	0.0317	**36**	0.0098	0.0000	0.0001	0.7930
16	0.0484	0.0000	0.0023	0.5291	**37**	0.1965	0.0000	0.0003	0.2874
17	0.0473	0.0000	0.0000	0.6009	**38**	0.0535	0.0000	0.0071	0.4602
18	0.0738	0.0000	0.0003	0.5023	**39**	0.6680	0.0000	0.0000	0.0307
19	0.2673	0.0003	0.0060	0.1514	**40**	0.0007	0.0000	0.8638	0.0017
20	0.0380	0.0000	0.0008	0.6006	**41**	0.0319	0.0000	0.0006	0.6296
21	0.3134	0.0000	0.0001	0.1819	**42**	0.5242	0.0000	0.0004	0.0640

Data	$\mu i1^2$	$\mu i2^2$	$\mu i3^2$	$\mu i4^2$	Data	$\mu i1^2$	$\mu i2^2$	$\mu i3^2$	$\mu i4^2$
43	0.0560	0.0000	0.0001	0.5627	**72**	0.0018	0.0000	0.7896	0.0041
44	0.0000	0.9998	0.0000	0.0000	**73**	0.0000	0.0000	0.9720	0.0001
45	0.0698	0.0000	0.0004	0.5094	**74**	0.1862	0.0000	0.0005	0.2958
46	0.0252	0.0000	0.0015	0.6402	**75**	0.0008	0.0000	0.8554	0.0019
47	0.1140	0.0000	0.0002	0.4156	**76**	0.0003	0.0000	0.9092	0.0007
48	0.1243	0.0001	0.0069	0.3077	**77**	0.0394	0.0000	0.0005	0.6016
49	0.4533	0.0000	0.0005	0.0901	**78**	0.0036	0.0001	0.7071	0.0081
50	0.3187	0.0003	0.0048	0.1221	**79**	0.0000	0.9998	0.0000	0.0000
51	0.5300	0.0000	0.0008	0.0566	**80**	0.2124	0.0005	0.0101	0.1734
52	0.0032	0.0000	0.0000	0.8836	**81**	0.0714	0.0000	0.0004	0.5041
53	0.6019	0.0000	0.0003	0.0409	**82**	0.0151	0.0000	0.0000	0.7574
54	0.4942	0.0000	0.0003	0.0772	**83**	0.0001	0.0000	0.9440	0.0003
55	0.0159	0.0000	0.0001	0.7484	**84**	0.0002	0.0000	0.9204	0.0005
56	0.5829	0.0000	0.0003	0.0469	**85**	0.0137	0.0000	0.0001	0.7641
57	0.0488	0.0000	0.0009	0.5571	**86**	0.4670	0.0001	0.0013	0.0733
58	0.0002	0.0000	0.9231	0.0005	**87**	0.4866	0.0000	0.0004	0.0778
59	0.1382	0.0000	0.0036	0.3155	**88**	0.0549	0.0000	0.0149	0.4059
60	0.2797	0.0000	0.0005	0.1992	**89**	0.0001	0.0000	0.9610	0.0001
61	0.0002	0.0000	0.9233	0.0005	**90**	0.0002	0.0000	0.9308	0.0004
62	0.0783	0.0000	0.0001	0.5007	**91**	0.0677	0.0000	0.0003	0.5214
63	0.0003	0.0000	0.9061	0.0008	**92**	0.1116	0.0000	0.0004	0.4147
64	0.0308	0.0033	0.2928	0.0510	**93**	0.0001	0.0000	0.9405	0.0003
65	0.6846	0.0000	0.0001	0.0262	**94**	0.0727	0.0000	0.0002	0.5107
66	0.0799	0.0000	0.0001	0.4979	**95**	0.0407	0.0000	0.0008	0.5878
67	0.0007	0.0000	0.8606	0.0017	**96**	0.0002	0.0000	0.9183	0.0006
68	0.0442	0.0000	0.0003	0.5935	**97**	0.0003	0.0000	0.9111	0.0007
69	0.0424	0.0079	0.2045	0.0641	**98**	0.0086	0.0000	0.0001	0.8017
70	0.0004	0.0000	0.8911	0.0010	**99**	0.0000	0.0000	0.9864	0.0000
71	0.8431	0.0000	0.0000	0.0059	**100**	0.0268	0.0000	0.0002	0.6750

Here is the calculation of the PCI value:

$$PCI = \frac{1}{100}(66.27867309)$$
$$= 0.662786731$$

The value of the PCI (Partition Coefficient Index) is 0.662786731.

b. Partition Entropy Index (PEI)

To calculate the value of PEI, the Equation 2.14 is employed. The result of PEI validation on Fuzzy C-Means method is shown in Table 4.5.

Table 4.5 PEI Result on Fuzzy C-Means

Data	µi1	µi2	µi3	µi4	Data	µi1	µi2	µi3	µi4
1	-0.169	-0.058	-0.104	-0.273	**35**	-0.119	-0.035	-0.054	-0.308
2	-0.268	-0.123	-0.181	-0.368	**36**	-0.119	-0.035	-0.045	-0.103
3	-0.102	-0.028	-0.059	-0.348	**37**	-0.101	-0.028	-0.073	-0.334
4	-0.119	-0.035	-0.085	-0.360	**38**	-0.318	-0.174	-0.208	-0.263
5	-0.129	-0.039	-0.049	-0.276	**39**	-0.057	-0.013	-0.032	-0.305
6	-0.024	-0.091	-0.027	-0.066	**40**	-0.046	-0.144	-0.068	-0.131
7	-0.290	-0.144	-0.198	-0.312	**41**	-0.182	-0.065	-0.093	-0.184
8	-0.275	-0.130	-0.188	-0.335	**42**	-0.142	-0.045	-0.076	-0.348
9	-0.143	-0.046	-0.077	-0.344	**43**	-0.076	-0.019	-0.053	-0.216
10	-0.097	-0.027	-0.037	-0.279	**44**	-0.357	-0.333	0.000	0.000
11	-0.204	-0.078	-0.089	-0.144	**45**	-0.126	-0.038	-0.078	-0.241
12	-0.117	-0.034	-0.051	-0.300	**46**	-0.251	-0.110	-0.124	-0.178
13	-0.365	-0.314	-0.325	-0.322	**47**	-0.095	-0.026	-0.066	-0.283
14	-0.366	-0.307	-0.322	-0.323	**48**	-0.296	-0.150	-0.207	-0.327
15	-0.155	-0.051	-0.068	-0.307	**49**	-0.156	-0.052	-0.086	-0.361
16	-0.253	-0.111	-0.146	-0.232	**50**	-0.279	-0.133	-0.185	-0.367
17	-0.036	-0.007	-0.032	-0.197	**51**	-0.193	-0.071	-0.099	-0.342
18	-0.105	-0.030	-0.071	-0.244	**52**	-0.056	-0.013	-0.018	-0.058
19	-0.287	-0.141	-0.199	-0.367	**53**	-0.155	-0.051	-0.073	-0.323
20	-0.194	-0.072	-0.099	-0.198	**54**	-0.109	-0.031	-0.067	-0.356
21	-0.055	-0.013	-0.053	-0.363	**55**	-0.082	-0.021	-0.039	-0.125
22	-0.130	-0.040	-0.063	-0.324	**56**	-0.141	-0.045	-0.069	-0.331
23	-0.074	-0.019	-0.025	-0.070	**57**	-0.186	-0.067	-0.104	-0.218
24	-0.119	-0.035	-0.025	-0.051	**58**	-0.037	-0.124	-0.038	-0.086
25	-0.165	-0.056	-0.113	-0.326	**59**	-0.250	-0.109	-0.168	-0.324
26	-0.097	-0.026	-0.054	-0.337	**60**	-0.109	-0.031	-0.083	-0.360
27	-0.034	-0.007	-0.039	-0.250	**61**	-0.020	-0.079	-0.038	-0.086
28	-0.272	-0.127	-0.158	-0.225	**62**	-0.069	-0.017	-0.050	-0.245
29	-0.114	-0.033	-0.066	-0.199	**63**	-0.045	-0.141	-0.047	-0.099
30	-0.213	-0.083	-0.128	-0.258	**64**	-0.208	-0.342	-0.332	-0.336
31	-0.326	-0.184	-0.233	-0.367	**65**	-0.092	-0.024	-0.042	-0.295
32	-0.288	-0.142	-0.202	-0.365	**66**	-0.056	-0.013	-0.047	-0.246
33	-0.084	-0.022	-0.063	-0.301	**67**	-0.036	-0.122	-0.070	-0.132
34	-0.119	-0.035	-0.054	-0.308	**68**	-0.134	-0.042	-0.071	-0.201

Data	μi1	μi2	μi3	μi4	Data	μi1	μi2	μi3	μi4
69	-0.254	-0.363	-0.359	-0.348	**85**	-0.082	-0.021	-0.038	-0.118
70	-0.035	-0.119	-0.054	-0.111	**86**	-0.216	-0.085	-0.121	-0.354
71	-0.086	-0.022	-0.022	-0.197	**87**	-0.138	-0.043	-0.078	-0.356
72	-0.084	-0.214	-0.105	-0.176	**88**	-0.351	-0.225	-0.257	-0.287
73	-0.026	-0.096	-0.014	-0.039	**89**	-0.028	-0.102	-0.020	-0.051
74	-0.124	-0.037	-0.084	-0.331	**90**	-0.023	-0.088	-0.035	-0.078
75	-0.058	-0.168	-0.072	-0.136	**91**	-0.102	-0.028	-0.066	-0.235
76	-0.029	-0.104	-0.045	-0.096	**92**	-0.109	-0.031	-0.076	-0.283
77	-0.159	-0.053	-0.088	-0.197	**93**	-0.013	-0.057	-0.030	-0.070
78	-0.097	-0.233	-0.146	-0.217	**94**	-0.098	-0.027	-0.060	-0.240
79	-0.357	-0.333	0.000	0.000	**95**	-0.198	-0.074	-0.103	-0.204
80	-0.318	-0.174	-0.231	-0.365	**96**	-0.023	-0.087	-0.041	-0.089
81	-0.125	-0.038	-0.079	-0.243	**97**	-0.031	-0.109	-0.044	-0.095
82	-0.066	-0.016	-0.032	-0.121	**98**	-0.146	-0.047	-0.050	-0.099
83	-0.023	-0.086	-0.028	-0.067	**99**	-0.016	-0.065	-0.007	-0.022
84	-0.015	-0.064	-0.040	-0.087	**100**	-0.113	-0.033	-0.057	-0.161

Below is the calculation of the PEI value:

$$\text{PEI} = -\frac{1}{100}(-54.69675219)$$
$$= 0.546967522$$

The value of the PEI (Partition Entropy Index) is 0.546967522.

4.3.3 Fuzzy Subtractive Clustering Processing

The determination of the number of clusters is still the same as the Fuzzy C-Means method based on 10 (ten) main variables, both of categories include data Tx Power, Rx Power, Temperature, Power Supply, and Bias Current. The data can be seen in Table 4.1. The parameters used in the process of clustering using the Fuzzy Subtractive Clustering algorithm are:

Influence range (r) = 0.2;

Accept ratio = 0.5;

Reject ratio = 0.15;

Squash factor (q) = 1.25;

Bottom line (Xmin) = [0;0;0;0;0;0;0;0;0;0]

Upper limit (Xmax) = [5;100;70;5;25;5;5;70;5;50]

The first step in the grouping process with Fuzzy Subtractive Clustering is the normalization of data is to equalize the range between the 10 variables. Normalization of data can be calculated using the following Equation 2.6. Example on the first data:

- First Variable

  $X_{11} = \frac{2.31-0}{5-0} = \frac{2.31}{5} = 0.462$

- Second Variable

  $X_{12} = \frac{-24.09-0}{100-0} = \frac{-24.09}{100} = -0.2409$

- Third Variable

  $X_{13} = \frac{50-0}{70-0} = \frac{50}{70} = 0.714285714$

- Fourth Variable

  $X_{14} = \frac{3.34-0}{5-0} = \frac{3.34}{5} = 0.668$

- Fifth Variable

  $X_{15} = \frac{17-0}{25-0} = \frac{17}{25} = 0.68$

- Sixth Variable

  $X_{16} = \frac{3.7-0}{5-0} = \frac{3.7}{5} = 0.74$

- Seventh Variable

  $X_{17} = \frac{-13.882-0}{100-0} = \frac{-13.882}{5} = -0.13882$

- Eight Variable

  $X_{18} = \frac{37-0}{70-0} = \frac{37}{70} = 0.528571429$

- Ninth Variable

  $X_{19} = \frac{3.24-0}{5-0} = \frac{3.24}{5} = 0.648$

- Tenth Variable

  $X_{110} = \frac{15-0}{50-0} = \frac{15}{50} = 0.68$

The above step is also done to the 2nd data up to the 100th data. So the final result of normalization as in Table 4.6.

Tabel 4.6 Normalization Data

No	ONU - Optical Network Unit (from Customer)					OLT - Optical Network Termination (from Company)				
	Tx Power (dBm)	**Rx Power (dBm)**	**Temperature (°C)**	**Power Supply (Volt)**	**Bias Current (mA)**	**Tx Power (dBm)**	**Rx Power (dBm)**	**Temperature (°C)**	**Power Supply (Volt)**	**Bias Current (mA)**
1	0.462	-0.2409	0.714286	0.668	0.68	0.74	-2.7764	0.528571	0.648	0.3
2	0.5372	-0.15392	0.5563	0.656	0.514	0.57	-2.8272	0.614286	0.627	0.61356
3	0.42	-0.1928	0.6	0.648	0.52	0.734	-2.6648	0.628571	0.64	0.34
4	0.47	-0.1703	0.571429	0.652	0.44	0.772	-2.5508	0.6	0.636	0.2
5	0.448	-0.173	0.714286	0.648	0.44	0.734	-2.5344	0.728571	0.642	0.26
6	0.4216	-0.20088	0.791186	0.644	0.644	0	-3.8292	0	0	0
7	0.396	-0.3154	0.6	0.656	0.48	0.81	-3.1208	0.5	0.638	0.16
8	0.386	-0.3397	0.614286	0.652	0.4	0.752	-3.1592	0.528571	0.652	0.28
9	0.404	-0.1549	0.571429	0.656	0.48	0.698	-2.4368	0.685714	0.64	0.24
10	0.41	-0.2125	0.685714	0.656	0.48	0.682	-2.7136	0.685714	0.638	0.22
11	0.46	-0.1815	0.671429	0.656	0.56	0.758	-2.5836	0.457143	0.664	0.22
12	0.456	-0.1866	0.6	0.648	0.4	0.722	-2.5088	0.671429	0.634	0.34
13	0.416	-0.1673	0.714286	0.66	0.32	0.73	-2.552	0.314286	0.64	0.18
14	0.436	-0.1705	0.642857	0.656	0.28	0.762	-2.5804	0.314286	0.64	0.18
15	0.454	-0.1958	0.614286	0.664	0.52	0.734	-2.7064	0.728571	0.638	0.36
16	0.404	-0.1889	0.8	0.656	0.4	0.78	-2.6324	0.471429	0.64	0.18
17	0.462	-0.2081	0.671429	0.652	0.48	0.768	-2.692	0.571429	0.646	0.22
18	0.4	-0.1632	0.742857	0.656	0.36	0.802	-2.5228	0.557143	0.642	0.18
19	0.5112	-0.20758	0.734371	0.636	0.84592	0.7028	-4.3242	0.642857	0.6284	0.59288

No	ONU - Optical Network Unit (from Customer)					OLT - Optical Network Termination (from Company)				
	Tx Power (dBm)	Rx Power (dBm)	Temperature (°C)	Power Supply (Volt)	Bias Current (mA)	Tx Power (dBm)	Rx Power (dBm)	Temperature (°C)	Power Supply (Volt)	Bias Current (mA)
20	0.474	-0.2468	0.642857	0.66	0.52	0.82	-2.9296	0.5	0.634	0.18
21	0.456	-0.1655	0.642857	0.636	0.32	0.714	-2.5804	0.6	0.642	0.22
22	0.424	-0.1798	0.585714	0.648	0.32	0.816	-2.6368	0.685714	0.634	0.26
23	0.45	-0.1754	0.685714	0.648	0.44	0.734	-2.5672	0.514286	0.648	0.26
24	0.42	-0.1946	0.671429	0.648	0.4	0.786	-2.6244	0.5	0.636	0.2
25	0.484	-0.1947	0.611	0.648	0.488	0.692	-4.3254	0.558571	0.6424	0.24428
26	0.47	-0.1623	0.728571	0.66	0.32	0.736	-2.554	0.642857	0.642	0.34
27	0.398	-0.1879	0.642857	0.656	0.44	0.752	-2.6148	0.571429	0.646	0.22
28	0.428	-0.1807	0.7	0.652	0.32	0.74	-2.5876	0.414286	0.648	0.3
29	0.406	-0.1841	0.757143	0.656	0.36	0.726	-2.5868	0.528571	0.66	0.26
30	0.432	-0.2284	0.757143	0.656	0.72	0.75	-2.8372	0.514286	0.646	0.22
31	0.5164	-0.138	0.765843	0.644	0.772	0.6496	-3.016	0.842857	0.6296	0.82962
32	0.462	-0.19706	0.760771	0.644	0.716	0.7296	-4.1786	0.587214	0.6	0.60028
33	0.446	-0.1979	0.642857	0.656	0.24	0.76	-2.6572	0.571429	0.64	0.22
34	0.448	-0.1614	0.685714	0.66	0.28	0.744	-2.5416	0.7	0.638	0.22
35	0.448	-0.1614	0.685714	0.66	0.28	0.744	-2.5416	0.7	0.638	0.22
36	0.436	-0.1832	0.642857	0.656	0.48	0.662	-2.5296	0.5	0.636	0.22
37	0.464	-0.1983	0.628571	0.664	0.4	0.696	-2.6432	0.6	0.64	0.14
38	0.372	-0.1886	0.614286	0.648	0.36	0.752	-2.6848	0.385714	0.634	0.28
39	0.47	-0.1809	0.657143	0.652	0.52	0.67	-2.5264	0.642857	0.636	0.22
40	0.404	-0.213	0.771429	0.652	0.4	0	-2.7876	0	0	0
41	0.456	-0.173	0.728571	0.66	0.32	0.732	-2.5308	0.471429	0.666	0.28
42	0.406	-0.1943	0.571429	0.64	0.36	0.746	-2.5228	0.657143	0.642	0.36

No	ONU - Optical Network Unit (from Customer)					OLT - Optical Network Termination (from Company)				
	Tx Power (dBm)	Rx Power (dBm)	Temperature (^oC)	Power Supply (Volt)	Bias Current (mA)	Tx Power (dBm)	Rx Power (dBm)	Temperature (^oC)	Power Supply (Volt)	Bias Current (mA)
43	0.43	-0.1798	0.657143	0.66	0.32	0.728	-2.6092	0.557143	0.636	0.18
44	0.4408	0.98064	0.636214	0.64	0.504	0.706	-4.2868	0.742857	0.6352	0.64708
45	0.47	-0.1939	0.714286	0.66	0.64	0.81	-2.5296	0.557143	0.642	0.16
46	0.444	-0.1917	0.742857	0.656	0.36	0.75	-2.6696	0.442857	0.64	0.2
47	0.408	-0.2194	0.714286	0.66	0.28	0.75	-2.7776	0.557143	0.656	0.3
48	0.438	-0.1809	0.9	0.656	0.4	0.728	-2.556	0.528571	0.642	0.14
49	0.454	-0.2657	0.628571	0.664	0.44	0.732	-2.798	0.685714	0.634	0.22
50	0.518	-0.2409	0.755743	0.636	0.56	0.7578	-4.6054	0.745243	0.6	0.58936
51	0.454	-0.2075	0.757143	0.656	0.36	0.684	-2.7436	0.785714	0.642	0.32
52	0.388	-0.1801	0.671429	0.652	0.48	0.75	-2.5672	0.528571	0.668	0.24
53	0.434	-0.1829	0.742857	0.656	0.32	0.78	-2.6348	0.742857	0.64	0.26
54	0.404	-0.1495	0.757143	0.656	0.4	0.746	-2.486	0.642857	0.646	0.32
55	0.458	-0.1818	0.714286	0.644	0.4	0.806	-2.5288	0.542857	0.642	0.18
56	0.452	-0.197	0.614286	0.656	0.44	0.79	-2.648	0.714286	0.644	0.18
57	0.432	-0.1785	0.771429	0.66	0.28	0.738	-2.5776	0.5	0.66	0.22
58	0.5224	-0.23098	0.666514	0.64	0.644	0	-3.4244	0	0	0
59	0.432	-0.2125	0.514286	0.656	0.28	0.792	-2.7276	0.528571	0.642	0.14
60	0.476	-0.1365	0.671429	0.664	0.56	0.822	-2.4796	0.614286	0.638	0.16
61	0.4268	-0.18014	0.687171	0.64	0.388	0	-3.5496	0	0	0
62	0.442	-0.2097	0.657143	0.66	0.56	0.738	-2.6672	0.571429	0.632	0.2
63	0.646	-0.2398	0.779629	0.652	0.67464	0	-3.4428	0	0	0
64	0.4952	0.13316	0.837171	0.66	0.546	0	-4.7924	0	0	0
65	0.462	-0.1534	0.671429	0.66	0.56	0.736	-2.486	0.671429	0.642	0.24

No	ONU - Optical Network Unit (from Customer)					OLT - Optical Network Termination (from Company)				
	Tx Power (dBm)	Rx Power (dBm)	Temperature (^oC)	Power Supply (Volt)	Bias Current (mA)	Tx Power (dBm)	Rx Power (dBm)	Temperature (^oC)	Power Supply (Volt)	Bias Current (mA)
66	0.45	-0.1913	0.657143	0.664	0.56	0.732	-2.6548	0.571429	0.632	0.2
67	0.4788	-0.1824	0.629129	0.64	0.548	0	-2.9844	0	0	0
68	0.468	-0.2337	0.657143	0.656	0.52	0.74	-2.7824	0.542857	0.64	0.16
69	0.5176	0.2222	0.595814	0.64	0.502	0	-3.398	0	0	0
70	0.578	-0.1947	0.760771	0.64	0.776	0	-2.9868	0	0	0
71	0.428	-0.1939	0.685714	0.648	0.48	0.738	-2.6108	0.685714	0.642	0.24
72	0.41	-0.28862	0.646314	0.64	0.556	0	-3.56	0	0	0
73	0.5008	-0.22758	0.720257	0.64	0.642	0	-4.1408	0	0	0
74	0.414	-0.2244	0.628571	0.656	0.52	0.712	-2.798	0.6	0.64	0.14
75	0.4728	-0.24814	0.6683	0.656	0.376	0	-4.536	0	0	0
76	0.4676	-0.18762	0.730414	0.64	0.674	0	-5.2044	0	0	0
77	0.466	-0.1578	0.7	0.66	0.6	0.72	-2.5428	0.514286	0.64	0.16
78	0.6368	-0.24684	0.900229	0.656	0.67632	0	-3.4112	0	0	0
79	0.4356	0.98064	0.636214	0.64	0.498	0.6946	-4.2814	0.742857	0.6346	0.64472
80	0.432	-0.1634	0.671429	0.648	0.44	0.734	2.5628	0.657143	0.642	0.18
81	0.426	-0.185	0.6	0.652	0.32	0.738	-2.5984	0.528571	0.638	0.22
82	0.454	-0.205	0.685714	0.648	0.56	0.75	-2.6456	0.542857	0.664	0.24
83	0.684	-0.19788	0.728743	0.664	0.45584	0	-4.7536	0	0	0
84	0.5076	-0.15592	0.684829	0.644	0.662	0	-4.6024	0	0	0
85	0.43	-0.1844	0.671429	0.656	0.36	0.746	-2.692	0.514286	0.656	0.28
86	0.388	-0.1882	0.771429	0.652	0.2	0.684	-2.688	0.8	0.642	0.34
87	0.404	-0.1863	0.785714	0.636	0.4	0.722	-2.5952	0.685714	0.632	0.24
88	0.464	-0.2796	0.685714	0.656	0.44	0.76	-3.062	0.4	0.64	0.12

No	ONU - Optical Network Unit (from Customer)					OLT - Optical Network Termination (from Company)				
	Tx Power (dBm)	Rx Power (dBm)	Temperature (^oC)	Power Supply (Volt)	Bias Current (mA)	Tx Power (dBm)	Rx Power (dBm)	Temperature (^oC)	Power Supply (Volt)	Bias Current (mA)
89	0.4344	-0.22678	0.745586	0.644	0.676	0	-3.7788	0	0	0
90	0.6404	-0.18664	0.757029	0.656	0.55848	0	-4.978	0	0	0
91	0.456	-0.1939	0.757143	0.66	0.36	0.746	-2.704	0.542857	0.652	0.28
92	0.404	-0.1737	0.742857	0.648	0.28	0.774	-2.5508	0.571429	0.64	0.18
93	0.5004	-0.16576	0.676614	0.644	0.666	0	-4.31	0	0	0
94	0.466	-0.2346	0.642857	0.64	0.36	0.764	-2.7468	0.557143	0.646	0.22
95	0.442	-0.2495	0.642857	0.66	0.52	0.794	-2.9684	0.5	0.634	0.18
96	0.4688	-0.18182	0.6413	0.644	0.554	0	-4.348	0	0	0
97	0.4936	-0.20058	0.794314	0.66	0.40856	0	-3.842	0	0	0
98	0.434	-0.1899	0.714286	0.648	0.52	0.766	-2.5368	0.5	0.644	0.18
99	0.548	-0.20606	0.7	0.636	0.634	0	-4.104	0	0	0
100	0.446	-0.1869	0.742857	0.656	0.4	0.728	-2.5836	0.514286	0.656	0.28

After obtaining the normalized data, the next step is to determine the initial potential value of the 1st data until the 100th data then search for data with the largest potential value selected as the first group center. The initial potential of each data is calculated using the Equation 2.8 and the calculation results are presented in Table 4.7.

Table 4.7 Inital Potential

Data	Initial Potential	Data	Initial Potential
1	3.276159146	**42**	3.05581125
2	1.000310815	**43**	10.00685708
3	4.1270792	**44**	1.98294683
4	6.190572893	**45**	5.179353148
5	5.851388426	**46**	6.667127884
6	3.313221293	**47**	0
7	2.659123335	**48**	1.428574518
8	1.936308772	**49**	4.913351358
9	3.762197928	**50**	1.049349778
10	5.831236185	**51**	2.26617063
11	7.018811926	**52**	10.26778953
12	4.415836643	**53**	3.938108254
13	1.965670626	**54**	4.191252417
14	1.853361045	**55**	9.190336649
15	2.762023334	**56**	3.155739715
16	4.503353971	**57**	6.809755242
17	13.2673036	**58**	5.110787063
18	6.875952963	**59**	1.852572376
19	1.320020299	**60**	4.103295809
20	6.231248337	**61**	3.077395944
21	8.846517591	**62**	11.64060376
22	3.32240608	**63**	2.348530758
23	12.04167891	**64**	1.000761166
24	11.02598771	**65**	6.347572338
25	4.324908138	**66**	11.54631924
26	3.892154155	**67**	3.907034509
27	0.00000000	**68**	9.766880298
28	4.219938598	**69**	1.000715577
29	8.640826755	**70**	2.584350232
30	3.297782113	**71**	8.094687041
31	1.000037395	**72**	2.439496374
32	1.317543263	**73**	6.171326863
33	7.662633712	**74**	6.371475372
34	5.696934499	**75**	2.684113847
35	5.696934499	**76**	4.5960941
36	6.329872041	**77**	0
37	0	**78**	1.349290593
38	2.15731114	**79**	1.98294683
39	6.593156423	**80**	1.010341706
40	2.250267828	**81**	7.886066853
41	7.186045346	**82**	11.08180887

Data	Initial Potential	Data	Initial Potential
85	10.37810748	**93**	5.854686624
86	1.521320815	**94**	10.14108712
87	0	**95**	7.886706952
88	2.055603647	**96**	4.988301099
89	4.442336353	**97**	0
90	2.500674274	**98**	10.1518002
91	8.361222521	**99**	5.703634866
92	6.400144661	**100**	9.355512058

Based on Table 4.7 it can be seen that the data with the greatest initial potential value is in the 23rd data selected as the first group center with 12.04167891.

Next, look for the point with the highest initial potential.

- M = max [At | i = 1,2, ..., n] = 12.04167891 the potential in the 23rd data;
- h = i = 3, such that D23 = M = 12.04167891;

After that, determine the center of the cluster and reduce its potential to the surrounding points:

a) Center = []
b) C = 0
c) Conditions = 1
d) Z = M = 12.04167891

Due to Condition≠0 and Z≠0, then do the calculation for #^{1st}Iteration

a. Determine the center of the cluster:
   The highest potential lies in the 23rd data, then the result of normalization of the 23rd data becomes the center of the cluster, namely: V1 = 0.45; V2 = -0.1754; V3 = 0.685714286; V4 = 0.648; V5 = 0.44; V6 = 0.734; V7 = -2.5672; V8 = 0.514285714; V9 = 0.648; V10 = 0.26
b. Determining whether a cluster center is accepted or not as a cluster center
   Ratio = Z / M = 12.04167891 / 12.04167891 = 1
   Ratio > accept_ratio (0.5), then Condition = 1.

c. Condition = 1 means cluster center candidate accepted as a cluster center, then:

1) C = C + 1 = 1

2) Center1 = 0.45; -0.1754; 0.685714286; 0.648; 0.44; 0.734; -2.5672; 0.514285714; 0.648; 0.26

3) Reduce the potential of the points near the center of the cluster from the first data up to the hundredth data.

Example on the first data:

The potential abatement value for the first data is:

$$Dc_1 = M * e^{-4|ST1|}$$
$$= 12.04167891 * 2.7187^{-4|1.408004|}$$
$$= 3.54869605$$

The new potential value is calculated by subtracting the potential of the first data with the first potential degradation value:

$$D_1 = D_1 - Dc_1$$
$$= 3.276159146 - 3.54869605$$
$$= -0.272536904$$

The calculation results for the 2nd data until the hundredth data are presented in Table 4.8.

Table 4.8 New Potential

Data	New Potential	Data	New Potential
D44	1.977449198	**D63**	1.260131
D79	1.977449198	**D93**	1.231021
D73	1.668210576	**D78**	1.209095
D99	1.644165223	**D43**	1.205291
D95	1.561545146	**D96**	1.196209
D6	1.525501475	**D62**	1.146511
D89	1.523932431	**D7**	1.125091
D20	1.412718284	**D34**	1.098794

Data	New Potential	Data	New Potential
D35	1.098794	**D85**	0.185429
D58	1.067294	**D15**	0.14628
D8	1.055217	**D59**	0.13181
D66	1.054353	**D10**	0.11719
D88	1.039186	**D100**	0.069958
D72	1.029713	**D5**	0.053118
D61	1.015042	**D14**	0.010086
D19	1.010005	**D94**	0.002433
D32	1.010005	**D23**	0
D90	1.00177	**D97**	-2.2E-20
D83	1.00177	**D18**	-0.06117
D75	1.000818	**D22**	-0.09818
D84	1.000618	**D45**	-0.13345
D67	1.000517	**D46**	-0.16963
D70	1.000359	**D91**	-0.19627
D40	1.000158	**D52**	-0.21235
D76	1.000072	**D13**	-0.24679
D64	1	**D65**	-0.33734
D50	1	**D42**	-0.34198
D69	1	**D9**	-0.41075
D25	1	**D12**	-0.42353
D80	1	**D55**	-0.48643
D31	0.999975	**D41**	-0.49802
D2	0.921618	**D16**	-0.53104
D30	0.848634	**D60**	-0.53183
D86	0.746032	**D98**	-0.5365
D29	0.652979	**D26**	-0.55793
D74	0.642866	**D39**	-0.63623
D17	0.631461	**D48**	-0.64122
D21	0.619483	**D3**	-0.69201
D24	0.546035	**D28**	-0.70052
D68	0.485935	**D11**	-0.81884
D92	0.472328	**D38**	-0.90043
D51	0.434219	**D56**	-0.93916
D82	0.417202	**D4**	-0.96054
D49	0.401231	**D54**	-0.97844
D1	0.397254	**D47**	-1.44198
D53	0.395333	**D36**	-1.65212
D57	0.385249	**D87**	-2.45806
D33	0.355057	**D77**	-2.77507
D71	0.306278	**D37**	-2.88816
D81	0.219735	**D27**	-4.1498

Based on Table 4.8 it can be seen that the data with the greatest potential value is in the 44th data with the potential value of 1.977449198. To determine the center of the group so it will be used the value of the ratio.

To determine the next group center, done the same way. If the ratio ≤ reject ratio, then there is no longer data that is considered to be a candidate for group center and iteration stop. After that, seeking the point with the highest new potential:

a) Z = max [At | i = 1,2, ..., n] = 1.977449198 the potential on the 44th data
b) h = i = 44, such that D44 = Z = 1.977449198

Due to Condition ≠ 0 and Z ≠ 0, then do the calculation for #2nd Iteration with the same step with the previous point c, so obtained:

Ratio = Z / M
= 1.977449198 / 12.04167891
= 0.39189028

Ratio ≤ reject_ratio (0,5), then Condition = 0

Condition = 0 means cluster center candidate is not accepted as cluster center, then process is stopped with cluster number 44 and center of cluster as in Table 4.9. Returns the cluster center of the normalized shape to its original shape which is denormalization data. By using the Equation 2.10 the example of calculation for the first data is:

- 1st Variable
  $Center_{11}$ = 0.462 * (5 – 0) + 0
  = 2.31
- 2nd Variable
  $Center_{12}$ = -0.2409 * (100 – 0) + 0
  = -24.09
- 3rd Variable

Center$_{13}$ = 0.714285714 * (70 – 0) + 0

= 50

- 4th Variable

Center$_{14}$ = 0.668 * (50 – 0) + 0

= 3.34

- 5th Variable

Center$_{15}$ = 0.34 * (50 – 0) + 0

= 17

- 6th Variable

Center$_{16}$ = 0.74 * (5 – 0) + 0

= 3.72.31

- 7th Variable

Center$_{17}$ = -0.13882 * (100 – 0) + 0

= -13.882

- 8th Variable

Center$_{18}$ = 0.528571429 * (70 – 0) + 0

= 37

- 9th Variable

Center$_{19}$ = 0.648 * (50 – 0) + 0

= 3.24

- 10th Variable

Center$_{110}$ = 0.3 * (50 – 0) + 0

= 15

Table 4.9 Normalization and Denormalization Data

23th Data	**NORMALIZATION**	Tx Power (dBm)	0.45
		Rx Power (dBm)	-0.1754
		Temperature (^{o}C)	0.685714
		Power Supply (Volt)	0.648
		Bias Current (mA)	0.44
		Tx Power (dBm)	0.734
		Rx Power (dBm)	-2.5672
		Temperature (^{o}C)	0.514286
		Power Supply (Volt)	0.648
		Bias Current (mA)	0.26

DENORMALIZATION	Tx Power (dBm)	2.25
	Rx Power (dBm)	-17.54
	Temperature (^{o}C)	48
	Power Supply (Volt)	3.24
	Bias Current (mA)	11
	Tx Power (dBm)	3.67
	Rx Power (dBm)	-12.836
	Temperature (^{o}C)	36
	Power Supply (Volt)	3.24
	Bias Current (mA)	13

At the last iteration obtained the sigma value. By using the Equation 2.11 the example of calculation for the first variable is:

$$\sigma_1 = 0.2 * \frac{5-0}{\sqrt{8}}$$

$$= 0.353553391$$

Do the calculations also on the second variable to the tenth variable. More data is shown in Table 4.10.

Table 4.10 Sigma Cluster

Variable	**Sigma Value**
v1	0.353553391
v2	7.071067812
v3	4.949747468
v4	0.353553391
v5	1.767766953
v6	0.353553391
v7	0.353553391
v8	4.949747468
v9	0.353553391
v10	3.535533906

By using the Gauss function in Equation 2.12, it can be found the degree of membership of each data in each group. The degree of membership of the 1st data (i = 1,2, ..., 100) in cluster 1:

$$\mu_{11} = e^{-|\left(\frac{(2.31-2.25)^2}{2*0.35355}\right)+\left(\frac{(-24.09+17.54)^2}{2*7.07106}\right)+\cdots+\left(\frac{(15+13)^2}{2*3.5355}\right)}$$

$= 0.110802385$

The above step is also done to the second until the hundredth data so that the final result of degree of membership for all data such as in Table 4.11.

Table 4.11 Degree of Membership using Radius 0.2

Data	**Degree of Membership On the cluster 1**	**Data**	**Degree of Membership On the cluster 1**
1	0.110802385	**32**	1.81308E-07
2	6.34438E-09	**33**	0.164498458
3	0.051309787	**34**	0.013520565
4	0.073985617	**35**	0.013520565
5	0.009294668	**36**	0.379341254
6	5.91087E-58	**37**	0.061353105
7	0.008939101	**38**	0.04761789
8	0.021431848	**39**	0.080159648
9	0.008869171	**40**	2.80375E-57
10	0.024027056	**41**	0.440848421
11	0.378959092	**42**	0.008555742
12	0.019544169	**43**	0.261823904
13	0.005373631	**44**	4.61368E-68
14	0.00378421	**45**	0.05310498
15	0.001754273	**46**	0.239805972
16	0.072433696	**47**	0.224748526
17	0.453953034	**48**	0.00219121
18	0.128995092	**49**	0.013551849
19	9.3301E-09	**50**	4.30385E-09
20	0.092744111	**51**	0.000155328
21	0.220951534	**52**	0.562915739
22	0.006335905	**53**	0.002095867
23	1	**54**	0.056994833
24	0.426007517	**55**	0.252150299
25	0.143437117	**56**	0.004003877
26	0.054348446	**57**	0.198118473
27	0.36935556	**58**	1.07457E-57
28	0.202519074	**59**	0.003840869
29	0.39888693	**60**	0.033000004
30	0.050080224	**61**	6.05455E-57
31	1.39807E-21	**62**	0.273538673

Data	Degree of Membership On the cluster	Data	Degree of Membership On the cluster
	1		1
63	1.07202E-59	**82**	0.536019202
64	1.1422E-62	**83**	9E-60
65	0.051167829	**84**	6.27105E-58
66	0.299104369	**85**	0.741113182
67	4.1488E-57	**86**	8.82792E-06
68	0.17843479	**87**	0.013611333
69	2.56422E-64	**88**	0.010964492
70	5.28947E-59	**89**	7.90248E-58
71	0.044731571	**90**	2.2723E-59
72	9.86142E-58	**91**	0.421160916
73	8.84402E-58	**92**	0.099120652
74	0.044804726	**93**	8.7192E-58
75	1.6211E-57	**94**	0.310902793
76	3.05028E-58	**95**	0.136176561
77	0.172469822	**96**	2.26218E-57
78	3.32217E-61	**97**	1.3419E-57
79	5.17287E-68	**98**	0.348925497
80	0	**99**	6.3879E-58
81	0.257854089	**100**	0.645440111

4.3.4 Fuzzy Subtractive Clustering Validation

Next, do the clustering validity on Fuzzy Subtractive Clustering method with 2 indicators as follows.

a. Partition Coefficient Index (PCI)

To calculate the value of PCI, the Equation 2.13 is employed. Below is the result of PCI validation on Fuzzy Subtractive Clustering method as shown in Table 4.12.

Table 4.12 PCI Result on Fuzzy Subtractive Clustering

Data	$\mu i1^2$	Data	$\mu i1^2$
1	3.50711E-10	**3**	0.000150712
2	2.46075E-23	**4**	0.005187898

Data	$\mu i1^2$	Data	$\mu i1^2$
5	6.97036E-05	**48**	3.68357E-06
6	6.9631E-256	**49**	4.45358E-09
7	1.75794E-31	**50**	0
8	1.55982E-34	**51**	1.86016E-11
9	2.08097E-06	**52**	0.249262018
10	6.31221E-06	**53**	2.03573E-07
11	0.015696594	**54**	0.000685762
12	0.000152162	**55**	0.037267464
13	3.18011E-06	**56**	4.35821E-06
14	2.97278E-07	**57**	0.000825635
15	2.46857E-08	**58**	4.8728E-181
16	0.001767393	**59**	1.87083E-09
17	0.007250074	**60**	2.71696E-05
18	0.004299536	**61**	5.8778E-197
19	5.4774E-295	**62**	0.001173664
20	1.37614E-14	**63**	1.1162E-188
21	0.005437778	**64**	0
22	1.76131E-06	**65**	8.10309E-05
23	1	**66**	0.002232037
24	0.074323458	**67**	2.486E-129
25	2.1401E-270	**68**	1.18459E-06
26	0.000329004	**69**	5.8194E-188
27	0.086812022	**70**	6.8834E-140
28	0.004353103	**71**	0.001077197
29	0.056427735	**72**	5.1606E-201
30	9.45901E-15	**73**	0
31	4.56146E-67	**74**	1.86396E-08
32	3.7983E-244	**75**	0
33	1.33278E-05	**76**	0
34	3.4476E-06	**77**	0.000567744
35	3.4476E-06	**78**	4.8606E-187
36	0.085376596	**79**	0
37	0.000935405	**80**	0
38	5.50026E-05	**81**	0.00631439
39	0.001765069	**82**	0.009721559
40	3.824E-118	**83**	0
41	0.017207019	**84**	0
42	1.89143E-05	**85**	0.009405842
43	0.005560405	**86**	7.50033E-16
44	0	**87**	0.000124636
45	5.27246E-06	**88**	7.37453E-26
46	0.002718271	**89**	9.5362E-246
47	1.58574E-07	**90**	0

Data	$\mu i1^2$	Data	$\mu i1^2$
91	0.001623824	**96**	0
92	0.00020014	**97**	2.4448E-255
93	0	**98**	0.038767864
94	5.9374E-05	**99**	0
95	8.04084E-17	**100**	0.310583273

Below is the calculation of the PCI value:

$$\text{PCI} = \frac{1}{100}(2.045943245)$$
$$= 0.020459432$$

The value of the PCI (Partition Coefficient Index) is 0.020459432.

b. Partition Entropy Index (PEI)

To calculate the value of PEI, the Equation 2.14 is employed. Below is the result of PEI validation on Fuzzy Subtractive Clustering method as shown in Table 4.13.

Table 4.13 PEI Result on Fuzzy Subtractive Clustering

Data	$\mu i1$	Data	$\mu i1$
1	-0.000203856	**18**	-0.17865591
2	-1.29122E-10	**19**	-2.5073E-145
3	-0.054017443	**20**	-1.87207E-06
4	-0.18948256	**21**	-0.192257736
5	-0.039954585	**22**	-0.008791954
6	-7.7517E-126	**23**	0
7	-1.48458E-14	**24**	-0.354318765
8	-4.86103E-16	**25**	-4.5418E-133
9	-0.009436246	**26**	-0.072730216
10	-0.015040579	**27**	-0.36005011
11	-0.260238637	**28**	-0.17935692
12	-0.054217582	**29**	-0.341446855
13	-0.011286947	**30**	-1.57031E-06
14	-0.004097039	**31**	-5.15845E-32
15	-0.001376111	**32**	-5.4618E-120
16	-0.133231186	**33**	-0.020490904
17	-0.209749625	**34**	-0.011677085

Data	μi1	Data	μi1
35	-0.011677085	**68**	-0.007426129
36	-0.35949689	**69**	-5.20012E-92
37	-0.106655892	**70**	-4.20347E-68
38	-0.036370372	**71**	-0.112138299
39	-0.133171197	**72**	-1.65649E-98
40	-2.64351E-57	**73**	-8.455E-164
41	-0.266446218	**74**	-0.001214952
42	-0.023649295	**75**	-7.4348E-223
43	-0.193581997	**76**	0
44	-1.2114E-193	**77**	-0.089040997
45	-0.013952772	**78**	-1.49546E-91
46	-0.154006536	**79**	-9.0809E-193
47	-0.003117424	**80**	0
48	-0.012006554	**81**	-0.20123738
49	-0.000641643	**82**	-0.228422365
50	-6.8691E-187	**83**	-4.1752E-264
51	-5.32817E-05	**84**	-3.1228E-236
52	-0.346799664	**85**	-0.226283592
53	-0.003475799	**86**	-4.76891E-07
54	-0.095386072	**87**	-0.050183034
55	-0.317528409	**88**	-7.8575E-12
56	-0.012884299	**89**	-8.7111E-121
57	-0.101995983	**90**	0
58	-1.44911E-88	**91**	-0.129412273
59	-0.000434627	**92**	-0.060241799
60	-0.027400316	**93**	-2.1765E-188
61	-1.73205E-96	**94**	-0.037493404
62	-0.115582707	**95**	-1.66157E-07
63	-2.28615E-92	**96**	-1.5885E-192
64	-9.8221E-275	**97**	-1.4494E-125
65	-0.042401148	**98**	-0.319971403
66	-0.144209892	**99**	0
67	-7.38231E-63	**100**	-0.325826334

Below is the calculation of the PEI value:

$$\text{PEI} = -\frac{1}{100}(-7.013930968)$$

$$= 0.07013931$$

The value of the PEI (Partition Entropy Index) is 0.07013931.

CHAPTER V

DISCUSSION

5.1 Grouping System Quality with Fuzzy C-Means Method

Conceptually, there are two algorithms in grouping techniques which are supervised and unsupervised. The difference between the two lies in determining the number of clusters that are formed. For supervised algorithms, it is necessary to know in advance the number of clusters to be formed, where the method usually used is Fuzzy C-means (FCM). The output of FCM is basically not a fuzzy inference system (IF-THEN) but is a collection of cluster centers as well as some degree of membership for each data point, then that information can be used to build a fuzzy inference system. FCM uses a fuzzy grouping model so that data can be a member of all classes or clusters formed with different degrees or membership levels between 0 and 1. The level of data presented in a class or cluster is determined by the degree of membership. The FCM method requires a large number of pre-defined group and group membership matrices. Typically, initial group membership matrices are randomly initialized which causes the FCM method to have inconsistency issues. The Fuzzy C-Means algorithm is one of the easiest and often used algorithms in data grouping techniques because it makes efficient estimates and does not require any parameters.

Based on the results of this study, it can be concluded for the grouping by Fuzzy C-Means method there are several things related to the results of the system quality classification based on 10 variables and 100 data. In accordance with the results obtained, the researchers concluded by grouping into 4 groups of quality based on the level of excellent quality, good quality, poor quality, very poor quality. The distribution of its cluster can be seen in Table 4.3. Several factors or strong variables that affect the quality of the system on IndiHome still needs further investigation. Because in this study, the final results obtained only to get the distribution of data to each cluster and look for the degree of membership to be done on the next validity research. And not yet known exactly which variables that affect grouping group. Thus, after getting the results of clustering it should be identified the solution to treat the good quality system always be in control to maintain the customer loyalty.

Cluster 1 consists of 5 factors in the ONU namely Tx Power with the data in the range of 2.02 to 3.42 dBm, the data range in Rx Power is -28.862 to 22.22 dBm, temperature from 41.707 to 63.016 °C, power supply from 3.18 to 3.32 volt, bias current from 9.4 to 19.4 mA. Then the 5 factors in the ONT are Tx Power with the data on the entire cluster member is 0 dBm, the data range in Rx Power is -26,022 to -13,938 dBm, the temperature with the data on the whole cluster member is 0, the power supply with the data on the entire cluster member is 0 volt, the bias current with the data on all cluster members is 0 mA. A significant difference in cluster 1 with the other 3 clusters is that the characteristics of this cluster lie in the Tx Power, Temperature, Power Supply, and Bias Current factors in the ONT. The data contained in these factors only have the data that is 0 dBm. Then in Tx Power in ONU, cluster 1 has higher data up to 3.42 dBm.

In cluster 2 consists of 5 factors in the ONU namely Tx Power with the data in the range of 1.94 to 2.686 dBm, the data range in Rx Power is -26.57 to 98.064 dBm, temperature from 38.941 to 55 °C, power supply from 3.18 to 3.32 volt, bias current from 5 to 21.148 mA. Then the 5 factors in the ONT are Tx Power with the data in the range of 2.85 to 4.08 dBm, the data range in Rx Power is -23.027 to 12.814 dBm, temperature from 41.105 to 59 °C, power supply from 3 to 3.23 volt, bias current from 9 to 41.481 mA. In the 2nd cluster, the characteristics that show significant differences from this cluster are the Temperature and Bias Current factors in the ONT, the data obtained has a higher yield than the other 3 clusters. In the Temperature factor, the data goes to 59°C and the data on the current bias has data up to 41,481 mA. Then in the Power Rx factor at ONU, it has very high data which is 98,064 dBm.

In cluster 3 consists of 5 factors in the ONU namely Tx Power with the data in the range of 1.86 to 2.33 dBm, the data range in Rx Power is -27.96 to -15.78 dBm, temperature from 43 to 63 °C, power supply from 3.22 to 3.34 volt, bias current from 7 to 18 mA. Then the 5 factors in the ONT are Tx Power with the data in the range of 3.31 to 4.03 dBm, the data range in Rx Power is -15.31 to -12.614 dBm, temperature from 22 to 40 °C, power supply from 3.17 to 3.33 volt, bias current from 6 to 15 mA. In clusters 3 and 4 have almost identical data on their characteristics, but the significant difference between the two clusters is that cluster 3 has a larger than the cluster 4 on the 5 factors in the ONU.

In cluster 4 consists of 5 factors in the ONU namely Tx Power with the data in the range of 1.93 to 2.42 dBm, the data range in Rx Power is -33.97 to -13.65 dBm, temperature from 36 to 50 °C, power supply from 3.18 to 3.32 volt, bias current from 6 to 16 mA. Then the 5 factors in the ONT are Tx Power with the data in the range of 3.46 to 4.11 dBm, the data range in Rx Power is -21.627 to -12.398 dBm, temperature from 35 to 43 °C, power supply from 3.16 to 3.34 volt, bias current from 7 to 14 mA. However, on the ONT factor, cluster 4 has greater data from cluster 3.

In cluster 1, the available data that have unstable data marked by the number 0. For clusters 2, 3 and 4, there are differences that are not too significant and none with the problem of unstable data. However, with the limitations of this research, the researcher did not carry out the analysis until it goes to the IF-THEN rule stage, so there was no consideration of the final rule towards what type of quality existed in each cluster. On the other hand, according to the parameter of validity clustering, Fuzzy C-Means are better than the Fuzzy Subtractive Clustering in Partition Coefficient Index (PCI) because the result is higher. Based on this case study, this method is better in terms of measuring the amount of overlap among groups and evaluating the degree of membership without considering the vector data.

5.2 Grouping System Quality with Fuzzy Subtractive Clustering Method

Fuzzy subtractive clustering is an unsupervised clustering algorithm that can form the number and center of the cluster corresponding to the data conditions. Fuzzy subtractive clustering is based on density size or potential data points in a space or variable. The basic concept of fuzzy subtractive clustering is to determine the regions in a variable that has high density to the points around it. In contrast to the supervised algorithm, grouping on an unsupervised algorithm cannot determine the number of clusters first, where the method usually used is the subtractive clustering method (Kusumadewi & Purnomo, 2004). FSC method has advantages in learning abilities that can solve complex problems without formulating. Data processing in forming estimation models based on subtractive grouping has special parameters in forming the number of clusters which is radius parameters (influence range). Used radius is usually in the range of 0 to 1. The higher the radius used will produce a small number of clusters, and vice versa.

In this research using the subtractive clustering process, the cluster radius is from 0.1 until 1. For each radius have each number of clusters but when in radius 0.1 the number of the cluster formed is 41 and then for radius 0.2 to 1, clusters formed only one. It can be seen in Table 5.1. The ratio of acceptance and rejectance is a constant value between 0 and 1 which is used as a measure for accepting and rejecting a cluster central candidate data point into a cluster center. It is considering the accept ratio 0.5, reject ratio 0.15, and squash factor 1.25. Similarly, at the center of different clusters, although with the same number of clusters as this will certainly affect the degree of membership. The grouping method that uses the fuzzy concept is a concept that a data can be a member in all clusters with the degree of membership value it has. The higher the degree of membership in a cluster, the greater the tendency to be a cluster member. The value of its influence range close to 0 will affect the accuracy of the predicted data and the number of clusters. There is a cluster central equilibrium of radius 0.2 to 0.8 because the cluster center of the radius is determined by the iteration process to locate the points with the largest number of neighbors and on the radius having the same number of iterations.

Table 5.1 Number of Cluster

Radius	Number of Clusters Formed	Cluster Center
0.1	41	34
0.2	1	23
0.3	1	23
0.4	1	23
0.5	1	23
0.6	1	23
0.7	1	23
0.8	1	23
0.9	1	17
1	1	17

According to the parameter of validity clustering, Fuzzy Subtractive Clustering is better than the Fuzzy C-Means in Partition Entropy Index (PEI) because the result is smaller. This method is better to evaluate the randomness of the data in a cluster. But in this case, after getting the number of clusters in each radius. It can be said that the Fuzzy Subtractive Clustering method cannot be used by the company because the results of the number of clusters obtained cannot represent the concept of the fuzzy clustering. There are very far results when the number of clusters is 41 in the radius of 0.1 but after that it drops dramatically at a radius of 0.2 and so on with the number of clusters only 1. Then, the number of clusters is 1 cannot enter the concept of clustering itself, because clustering can be said to reach its destination if there is more than 1 cluster.

5.3 Sensitivity Analysis on Fuzzy Subtractive Clustering

The clustering process in the fuzzy clustering algorithm always finds the best solution for defined parameters. However, this best solution may not necessarily determine the best description of the data structure. To determine the most optimal number of clusters and can validate whether fuzzy partitions applied in the clustering process are in accordance with the data, the validity measurement index is used. The cluster validity means the procedure to evaluate the results of the cluster analysis quantitatively so that the optimum group is produced. An optimum group is a group that has a solid range between individuals in the group and is isolated from other groups well. In fuzzy subtractive clustering, if the higher or lower radius is used it cannot promise good estimation results, which is calculated based on several parameters of clustering validity testing such as PCI and PEI. On testing the influence of the radius value is used finger value 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1 with the value reject ratio is taken from the best value on the previous test is 0.15 and the value accept ratio is equal to 0.5. test result the influence of the radius values is shown in Table 5.2 and Figure 5.1

Table 5.2 PCI and PEI value

Radius	**PCI**	**PEI**
0.1	0.436662	0.074157
0.2	0.020459	0.070139
0.3	0.059397	0.155595
0.4	0.124276	0.192753
0.5	0.197615	0.193938
0.6	0.266852	0.180695
0.7	0.327196	0.163564
0.8	0.378117	0.14654
0.9	0.413066	0.139047
1	0.451738	0.123053

From the results obtained in Table 5.2, PCI results are used to measure the number of overlap clusters and the highest results are in radius 1 with the value of 0.451738. It shows that with the largest PCI value, the cluster is the most optimal. Then in PEI, it is necessary to see the degree of fuzziness of the resulting cluster partition. The smallest PEI results show that the radius has an optimal number of clusters. The smallest is in radius 0.2 with the value of 0.070139.

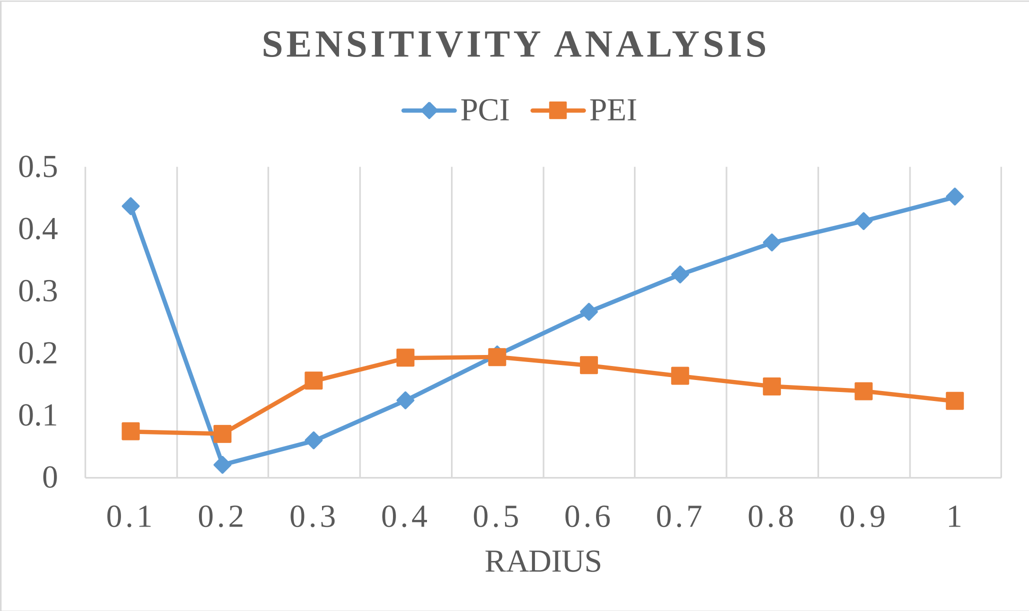

Figure 5.1 Sensitivity Analysis based on PCI and PEI

In the PCI results there is a decrease in the radius of 0.1 to radius 0.2 but after that it will increase again. This happens because the formula on the PCI is a quadratic result of the degree of membership. In the result, in radius 0.1 there are 41 clusters, therefore there is 41 degrees membership, and the radius of 0.2 to 1 has only 1 cluster, so there is only 1 degree of membership, making the graph decrease from radius 0.1 to radius 0.2. After that increase again because the higher the radius then the higher the PCI results due to the results of squaring.

In PEI results there is a decrease as it goes to radius 0.2 and again rises at a radius of 0.3 but an increase is only in a radius of 0.5 because after that the graph decreases until heading to radius 1. The decrease in the graph from radius 0.1 to radius 0.2 due to the differences number of clusters or the number of membership degrees significantly, and then increased again because of the increasing radius that affects the calculation on the PEI formula. However, when the graph is at a radius of 0.5 it decreases again as it goes to a radius of 0.6 and so on, due to the result of different degrees of membership. Then, when at a radius of 0.5 the result of membership degree is greater than the radius of 0.6. Then at radius 0.9 and 1, having a different cluster center from the previous radius and with the center of cluster 17 having fewer PEI results with that radius. Then the graph will continue to decrease when the radius is at the point of radius 0.6 to radius 1.

CHAPTER VI

CONCLUSION AND RECOMMENDATION

6.1 CONCLUSION

After conducting the study, it can be concluded that:

1. For PCI, the better method is Fuzzy C-Means in measuring the amount of overlapping among groups because it has the higher result. In Fuzzy C-Means result, the value of the PCI (Partition Coefficient Index) is 0.662786731. It shows that with the largest PCI value, the cluster is the most optimal. But in the PEI result, Fuzzy Subtractive clustering is having the smallest value and it becomes the better method especially to evaluate the randomness of the data in a cluster. The smallest PEI is in radius 0.2 with the value of 0.070139. The smallest PEI results show that the radius has an optimal number of clusters. Then both methods can be said to be better with each parameter. But after considering the number of clusters that are formed, compared to fuzzy c-means method has 4 clusters and in fuzzy subtractive only two clustering numbers are formed, which are 41 and 1. Then in the number of clusters parameter, the method that can be used in terms of grouping quality is Fuzzy C-Means.
2. There is a change in the PCI indicator graph that experiencing the increases and decreases. It occurs due to the formula that squaring the degree of membership results. Similarly, sensitivity to the validity of PEI, the graphs are not experiencing constant results, it can be said from the results of cluster validity suffered significant sensitive changes.

6.2 RECOMMENDATION

The advice given by the author to the PT. Telkom Indonesia branch Yogyakarta and researchers for further research development as follows:

6.2.1 For PT. Telkom Indonesia branch Yogyakarta

1. The company can consider to cluster the quality to differentiate the treatment carried out to the customer to reduce and overcome the complaint. Then in the parameters of the number of the cluster formed, the company can use the Fuzzy C-Means method. Besides to identify how many clusters are formed, they can also see what kind of data will become the cluster members in each cluster and find out the characteristics of each cluster to analyze the cluster types.
2. The company must focus and perform special treatment on cluster 1 because the available data in the cluster has unstable data that marked 0. This also shows the instability of the data, the system is categorized as bad quality.

6.2.2 For Further Researchers

1. To get the optimal result, it needs several times processes as comparisons. It needs several times processes with different parameters.
2. A formal method in Fuzzy Subtractive Clustering is required to determine the most optimal radius value in constructing estimation models.

REFERENCES

Ahmad, P. H., & Dang, S. (2015). Performance Evaluation of Clustering Algorithm Using DIfferent Dataset. *International Journal of Advance Research in Computer Science and Management Studies*, 167-173.

Arumugadevi, S., & Seenivasagam, V. (2015). Comparison of Clustering Methods for Segmenting Color Images. *Indian Journal of Science and Technology*, 670-677.

Bataineh, K. M., Naji, M., & Saqer, M. (2011). A Comparison Study between Various Fuzzy Clustering Algorithms. *Jordan Journal of Mechanical and Industrial Engineering*, 335 - 343.

Bezdek, J. C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms.* New York:: Plenum.

Chitra, K., & Maheswari, D. (2017). A Comparative Study of Various Clustering Algorithms in Data Mining. *International Journal of Computer Science and Mobile Computing*, 109-115.

Chiu, S. L. (1994). Fuzzy model identification based on cluster estimation. *Fuzzy Systems*, 267-268.

Elmzabi, A., Bellafkih, M., Ramdani, M., & Zeitouni, K. (2004). A Conditional Fuzzy Clustering with Adaptive Method. *IPMU*, 1-7.

Fat, J. (2014). Simulasi Sistem Prediksi Arah Gerak Indeks Harga Saham Gabungan Berbasis Adaptive Neuro-Fuzzy Interference System. *Jurnal Universitas Tarumanagara*, 33 - 44.

Ferarro, M. B., & Giordani, P. (2017). Possibilistic and fuzzy clustering methods for robust. *International Journal of Approximate Reasoning*.

Gorrostieta, E., Pedraza, J. C., & Carlos, R. J. (2005). Fuzzy Modelling of System. *Proceedings of 11 th IEEE International Conference on Methods and Models in Automation and Robotics MMAR 2005*. Miedzyzdroje, Poland.

Gose, E., Johnsonbaugh, R., & Jost, S. (2018). *Pattern Recognition and Image*. Rusia: Pleiades.

Gusti, S. K. (2012). Analisis Sebaran Puskesmas Untuk Peningkatan Pelayanan Kesehatan Dengan Metode Fuzzy C-Means. Seminar Nasional Teknologi Informasi Komunikasi dan Industri (SNTIKI) (hal. 78-84). Pekanbaru: Fakultas Sains dan Teknologi UIN Sultan Syarif Kasim Riau.

Halkidi, M., Batistakis, Y., & Vazirgiannis, M. (2001). On Clustering Validation Techniques. *Journal of Intelligent Information Systems*, 107-145.

Kahimpong, I. R., & Tielung, M. V. (2016). The Influence of Consumer Perception on Purchase Intention of Using IndiHome Product in Manado City. *Jurnal EMBA*, 1381-1391.

Kusumadewi, S. (2002). *Analisis dan Desain Sistem Fuzzy Menggunakan Toolbox Matlab.* Yogyakarta: Graha Ilmu.

Kusumadewi, S., & Purnomo, H. (2004). *Aplikasi Logika Fuzzy untuk Pendukung Keputusan.* Yogyakarta: Graha Ilmu.

Kusumadewi, S., & Purnomo, H. (2013). *Aplikasi Logika Fuzzy*. Yogyakarta: Graha Ilmu.

Marzouk, M., & Alaraby, M. (2012). Predicting Telecommunication Tower Costs Using Fuzzy Subtractive Clustering. *Journal of CIvil Engineering and Management*.

Mittal, P., & Suman. (2014). Comparison and Analysis of Various Clustering Methods in Data mining On Education data set Using the weak tool. *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), 3*(2), 240-244.

Nugraheni, Y. (2013). Data Mining Using Fuzzy Method for Customer Relationship Management in Retail Industry. *Lontar Komputer*.

Pereira, R., Fagundes, A., Melício, R., Mendes, V., Figueiredo, J., & Quadrado, J. (2014). Fuzzy subtractive clustering technique applied to demand response. *Procedia Technology*, 478 – 486.

Prihatini, P. M. (2015). Klustering Berbsis Prototipe dengan Metode Fuzzy C-Means. *Jurnal Logic*, 51-57.

Puspita, E. S., & Yulianti, L. (2016). Perancangan Sistem Peramalan Cuaca Berbasis Logika Fuzzy. *Jurnal Media Infotama* , 1 - 10.

Radionov, A., Evdokimov, S. A., Sarlybaev, A., & Karandaeva, O. (2015). Application of subtractive clustering for power transformer fault diagnostics. *Procedia Engineering*, 22 – 28.

Ramos, A. R., Santiago, O. L., Lazaro, B. D., Corona, C. C., Neto, A. S., & Galdean, J. V. (2017). A novel fault diagnosis scheme applying fuzzy clustering. *Applied Soft Computing*.

Rani, Y., & Rohil, H. (2013). A Study of Hierarchical Clustering Algorithm. *International Journal of Information and Computation Technology*, 1225-1232.

Rao, M., Sood, R., & Jarial, R. (2015). Subtractive clustering Fuzzy Expert System for Engineering Applications. *International Conference on Intelligent Computing, Communication & Convergence* , 77-83.

Respati, F. A. (2017). The Impact of Influence Range Fuzzy Subtractive Clustering Modification to Accuracy Anomalous Load Forecasting. *IOP Conference Series: Materials Science and Engineering*.

Rust, R. T., & Miu, C. (2006). What academic research tells us about service. *Communications of the ACM*, 49–54.

Sin, S. K., & De, F. (1993). Fuzzy System Designing Through Fuzzy Clustering and Optimal preDefuzzification. *IEEE International Conference on Fuzzy Systems*, 190-195.

Soni, K. G., & Patel, A. (2017). Comparative Analysis of K-means and K-medoids Algorithm on IRIS Data. *International Journal of Computational Intelligence Research*, 899-906.

Takagi, T., & Sugeno, M. (1985). Fuzzy Identification of Systems and its Application to Modeling and Control. *IEEE Trans. Systems Man and Cybernetics*, 116-132.

Tan, P. N., Steinbach, M., Karpatne, A., & Kumar, V. (2013). Introduction to Data Mining (Second Edition). Boston: Pearson Education.

Tiwari, S. K., & Yadav, D. N. (2015). Fuzzy Subtractive Clustering Based Prediction Approach for Machine Tool Vibration. *International Journal of Research and Development in Applied Science and Engineering (IJRDASE), 8*(2).

Vadim, K. (2018). Overview of different approaches to solving problems of Data Mining. *Procedia Computer Science*, 234-239.

Vannozzi, G., Croce, U. D., Starita, A., Benvenuti, F., & Cappozzo, A. (2004). Knowledge discovery in databases of biomechanical variables: application to the sit to stand motor task. *Journal of NeuroEngineering and Rehabilitation*, 1 - 7.

Wiryono, S. K., & Suharto. (2008). Analisis Risiko Operasional Di PT TELKOM Dengan Pendekatan Metode ERM. *Jurnal Manajemen Teknologi*, 58-90.

Wuest, T., Irgens, C., & Thoben, K. D. (2013). An approach to monitoring quality in manufacturing using supervised machine learning on product state data. *Journal of Intelligent Manufacturing*, 1167–1180.

Yager, R., & Filev, D. (1994). Generation of Fuzzy Rules by Mountain Clustering. *Journal of Intelligent & Fuzzy Systems*, 209-219.

Yan, J., Ryan, M., & Power, J. (1994). *Algoritma dari fuzzy c-means adalah sebagai berikut.* New York: PrenticeHall.

Yaqin, E. N., Abdullah, A. G., Chandra, D., Pratiwi, T. A., Adhigunarto, S., Shidiq, A. M., . . . Nandiyanto, A. B. (2018). Short Term Load Forecasting for Weekends in Indonesia:. *The 2nd Annual Applied Science and Engineering Conference (AASEC 2017)* (pp. 1-9). IOP Publishing.

Zhu, X., Pedrycz, W., & Li, Z. (2017). Fuzzy Clustering with Nonlinearly Transformed Data. *Applied Soft Computing*.

APPENDICES

Table Appendix 1. Fuzzy Subtractive Clustering Degree of Membership Radius 0.1

Data	Degree of Membership (in cluster)													
	1	**2**	**3**	**4**	**5**	**6**	**...**	**...**	**...**	**37**	**38**	**39**	**40**	**41**
1	1.2E-45	3.2E-08	6.6E-60	0	0	5.73E-14	...	...	...	0	0	0	0	2.04E-12
2	3.5E-62	2.3E-44	4.2E-35	0	0	7.69E-52	...	...	...	0	0	0	0	4.82E-26
3	2.2E-18	1.0E-05	2.9E-63	0	0	7.11E-23	...	...	...	0	0	0	0	1
4	1.6E-09	1.2E-07	5.1E-61	0	0	6.87E-31	...	...	...	0	0	0	0	1.46E-08
5	8.2E-06	4.6E-12	1.1E-57	0	0	1.99E-42	...	...	...	0	0	0	0	4.09E-10
6	0	0	7.2E-28	0	0	0	...	...	...	7.16E-36	9.1E-42	1.53E-44	3.04E-53	0
7	3.7E-80	1.5E-42	6.7E-22	0	0	1.22E-09	...	...	...	2.4E-275	3.5E-28	0	0	2.36E-49
8	2.3E-82	1.4E-52	1.1E-12	0	0	4.02E-18	...	...	...	2.9E-280	6.8E-28	0	0	4.02E-52
9	1.1E-12	6.0E-16	9.5E-56	0	0	1.31E-55	...	...	...	0	0	0	0	5.19E-13
10	3.1E-14	2E-05	1.5E-65	0	0	6.29E-20	...	...	...	0	0	0	0	2.49E-06
11	3.5E-25	9.2E-05	5.3E-64	0	0	3.84E-24	...	...	...	0	0	0	0	3.73E-11
12	1.4E-07	1.9E-15	8.2E-55	0	0	3.07E-46	...	...	...	0	0	0	0	4.88E-08
13	1.6E-27	1.3E-25	3.6E-44	0	0	2.04E-42	...	...	...	0	0	0	0	6.64E-34
14	1.6E-27	1.2E-27	4.9E-41	0	0	9.19E-40	...	...	...	0	0	0	0	3.57E-34
15	4.0E-20	2.1E-10	1.6E-58	0	0	4.87E-26	...	...	...	0	0	0	0	0.004071
16	5.6E-17	2.2E-11	1.9E-56	0	0	1.01E-24	...	...	...	0	0	0	0	1.6E-19
17	4.7E-15	0.0254	1.5E-65	0	0	6.58E-13	...	...	...	0	0	0	0	1.21E-05
18	2.8E-07	4.3E-14	5.4E-54	0	0	1.84E-38	...	...	...	0	0	0	0	1.36E-18
19	0	0	0	2.4E-27	3.1E-27	0	...	...	...	0	0	1.5E-294	7.9E-305	0
20	8.0E-47	1.5E-15	2.7E-50	0	0	1	...	...	...	9.8E-295	4.5E-29	0	0	7.11E-23
21	0.0013	7.0E-13	1.6E-57	0	0	3.85E-34	...	...	...	0	0	0	0	2.52E-12
22	1.0E-05	5.7E-16	1.9E-51	0	0	1.76E-31	...	...	...	0	0	0	0	9.87E-11

Data	Degree of Membership (in cluster)													
	1	**2**	**3**	**4**	**5**	**6**	**...**	**...**	**...**	**37**	**38**	**39**	**40**	**41**
23	1.1E-11	1.3E-06	5.2E-63	0	0	1.89E-28	...	...	...	0	0	0	0	2.27E-08
24	3.9E-12	5.7E-07	9.1E-61	0	0	7.93E-21	...	...	...	0	0	0	0	3.79E-11
25	0	0	0	1.3E-27	2.3E-27	0	...	...	...	0	0	1.1E-226	3.0E-221	0
26	0.00016	1.1E-18	4.7E-51	0	0	6.24E-44	...	...	...	0	0	0	0	2.45E-13
27	7.0E-10	0.00026	2.7E-64	0	0	1.18E-22	...	...	...	0	0	0	0	8.02E-06
28	7.2E-17	2.0E-18	8.6E-51	0	0	1.27E-34	...	...	...	0	0	0	0	9.27E-19
29	5.3E-09	5.5E-12	1.1E-57	0	0	2.99E-32	...	...	...	0	0	0	0	1.65E-13
30	2.9E-57	1.0E-12	8.2E-55	0	0	5.51E-13	...	...	...	2E-297	9.2E-29	0	0	3.32E-22
31	1.3E-15	1.2E-11	9.8E-46	0	0	7.1E-117	...	...	...	0	0	0	0	5.58E-91
32	0	0	0	3.5E-25	1.5E-25	0	...	...	...	0	0	2.5E-283	2.6E-290	0
33	8.5E-07	9.2E-19	1.7E-49	0	0	1.38E-29	...	...	...	0	0	0	0	5.4E-18
34	1	1.3E-20	7.2E-49	0	0	8.0E-47	...	...	...	0	0	0	0	2.2E-18
35	1	1.3E-20	7.2E-49	0	0	8.04E-47	...	...	...	0	0	0	0	2.2E-18
36	3.0E-16	2.9E-07	6.3E-66	0	0	2.27E-34	...	...	...	0	0	0	0	7.58E-11
37	4.0E-09	1.3E-06	3.4E-64	0	0	1.24E-22	...	...	...	0	0	0	0	3.53E-11
38	3.0E-25	3.8E-16	4.7E-52	0	0	5.89E-23	...	...	...	0	0	0	0	1.29E-16
39	1.3E-12	1.6E-06	2.5E-66	0	0	1.67E-37	...	...	...	0	0	0	0	2.47E-08
40	3.2E-27	6.7E-23	5.8E-51	0	0	2.3E-246	...	...	...	1.03E-98	4.8E-94	0	0	1.6E-264
41	3.0E-11	3.5E-18	1.0E-51	0	0	5.52E-41	...	...	...	0	0	0	0	5.11E-19
42	1.8E-08	1.1E-18	1.8E-50	0	0	4.86E-47	...	...	...	0	0	0	0	4.22E-09
43	6.3E-06	1.2E-11	6.0E-58	0	0	1.07E-28	...	...	...	0	0	0	0	3.41E-14
44	0	0	0	1	0.9128	0	...	...	...	0	0	0	0	0
45	1.3E-28	4.1E-07	7.3E-61	0	0	4.6E-33	...	...	...	0	0	0	0	1.19E-16
46	5.5E-17	6.4E-12	1.3E-56	0	0	7.79E-21	...	...	...	0	0	0	0	2.83E-18
47	3.9E-16	3.8E-19	1.2E-48	0	0	1.59E-20	...	...	...	0	0	0	0	1.23E-16
48	1.3E-17	1.0E-18	4.1E-51	0	0	7.61E-42	...	...	...	0	0	0	0	1.05E-29

Data	Degree of Membership (in cluster)													
	1	**2**	**3**	**4**	**5**	**6**	**...**	**...**	**...**	**37**	**38**	**39**	**40**	**41**
49	4.0E-19	2.9E-09	7.2E-59	0	0	1.25E-12	...	...	...	0	0	0	0	2.63E-09
50	0	0	0	6.7E-28	6.1E-28	0	...	...	...	0	0	0	0	0
51	7.8E-14	1.8E-21	1.9E-49	0	0	1.33E-34	...	...	...	0	0	0	0	1.38E-15
52	7.3E-14	3.9E-05	9.8E-64	0	0	7.54E-28	...	...	...	0	0	0	0	2.08E-07
53	0.00050	2.2E-18	1.5E-49	0	0	3.85E-37	...	...	...	0	0	0	0	3.46E-15
54	2.4E-07	6.2E-17	1.1E-52	0	0	4.17E-50	...	...	...	0	0	0	0	1.37E-13
55	9.0E-09	2.5E-10	1.0E-57	0	0	3.3E-33	...	...	...	0	0	0	0	1.31E-15
56	6.1E-09	9.0E-08	7.1E-60	0	0	2.01E-24	...	...	...	0	0	0	0	2.19E-08
57	2.3E-09	2.7E-19	4.5E-50	0	0	8.27E-38	...	...	...	0	0	0	0	1.09E-22
58	0	0	4E-146	0	0	3.1E-282	...	...	...	7.29E-06	8.6E-13	6.7E-138	2.4E-151	0
59	5.8E-19	5.1E-20	4.2E-47	0	0	8.98E-22	...	...	...	0	0	0	0	4.55E-22
60	2.9E-18	7.7E-10	1.2E-58	0	0	8.78E-41	...	...	...	0	0	0	0	5.34E-16
61	0	0	2.6E-18	0	0	8.6E-308	...	...	...	1.86E-27	7.3E-35	1.3E-115	6.3E-117	0
62	1.3E-20	1	4.5E-68	0	0	1.53E-15	...	...	...	0	0	0	0	1.02E-05
63	0	0	1.4E-16	0	0	9.8E-295	...	...	...	1	0.01876	7.2E-138	5.7E-155	0
64	0	0	0	0	0	0	...	...	...	0	0	1.05E-63	4.13E-59	0
65	3.4E-15	4.0E-09	6.0E-61	0	0	7.73E-44	...	...	...	0	0	0	0	4.2E-10
66	7.6E-20	0.784014	2.2E-68	0	0	4.25E-17	...	...	...	0	0	0	0	8.95E-06
67	0	1.8E-24	5.3E-60	0	0	6.6E-238	...	...	...	2.1E-49	4.2E-53	0	0	7.5E-273
68	7.5E-27	0.00058	3.5E-64	0	0	4.52E-06	...	...	...	0	0	0	0	2.61E-11
69	0	0	1.7E-16	0	0	0	...	...	...	4.33E-52	7.2E-63	3.2E-177	4.1E-187	0
70	0	1.9E-25	9.0E-73	0	0	3.3E-253	...	...	...	7.07E-40	3.0E-38	0	0	2E-292
71	6.7E-09	3.3E-05	6.9E-64	0	0	1.18E-27	...	...	...	0	0	0	0	4.17E-05
72	0	0	4.3E-18	0	0	9.1E-307	...	...	...	9.26E-19	1.3E-27	9.5E-105	3.5E-111	0
73	0	0	0	0	0	0	...	...	...	7.13E-90	2.4E-10	8.49E-07	3.71E-11	0
74	1.2E-26	3.4E-05	6.9E-64	0	0	1.56E-08	...	...	...	0	0	0	0	2.57E-11

Data	Degree of Membership (in cluster)													
	1	**2**	**3**	**4**	**5**	**6**	**...**	**...**	**...**	**37**	**38**	**39**	**40**	**41**
75	0	0	0	0	0	0	...	...	...	3.3E-231	3.4E-25	1.48E-25	2.74E-13	0
76	0	0	0	0	0	0	...	...	...	0	0	1.8E-140	5E-132	0
77	2.5E-25	1.7E-05	6.4E-65	0	0	1.24E-31	...	...	...	0	0	0	0	1.07E-14
78	0	0	2.6E-15	0	0	4.5E-297	...	...	...	0.001876	1	3.6E-154	4.3E-173	0
79	0	0	0	0.91280	1	0	...	...	...	0	0	0	0	0
80	0	0	0	0	0	0	...	...	...	0	0	0	0	0
81	3.8E-08	1.1E-12	1.6E-56	0	0	1.03E-29	...	...	...	0	0	0	0	1.06E-12
83	0	0	0	0	0	0	...	...	...	0	0	3.68E-49	1.82E-40	0
84	0	0	0	0	0	0	...	...	...	1.8E-240	1.1E-25	1.28E-15	1.06E-14	0
85	1.2E-12	9.5E-10	8.4E-59	0	0	6.03E-19	...	...	...	0	0	0	0	3.42E-09
86	1.8E-12	3.4E-39	7.0E-32	0	0	1E-56	...	...	...	0	0	0	0	1.94E-29
87	3.3E-06	5.5E-12	9.8E-58	0	0	4.51E-36	...	...	...	0	0	0	0	1.77E-12
88	4.1E-72	1.0E-37	2.5E-28	0	0	2.1E-08	...	...	...	7.8E-245	9.1E-24	0	0	8.86E-50
89	0	0	3.3E-26	0	0	0	...	...	...	2.33E-28	1.4E-35	5.42E-52	4.85E-62	0
90	0	0	0	0	0	0	...	...	...	0	0	7.01E-85	3.77E-77	0
91	1.8E-12	4.6E-11	1.8E-57	0	0	6.86E-20	...	...	...	0	0	0	0	6.22E-12
92	5.2E-05	8.7E-19	1.3E-49	0	0	1.84E-40	...	...	...	0	0	0	0	6.18E-22
93	0	0	0	0	0	0	...	...	...	7.2E-138	3.6E-15	1	0.001365	0
94	3.7E-14	2.4E-09	3.3E-58	0	0	1.68E-12	...	...	...	0	0	0	0	6.81E-11
95	1.0E-51	1.4E-18	2.8E-47	0	0	0.276625	...	...	...	6.9E-283	1.5E-28	0	0	1.12E-25
96	0	0	0	0	0	0	...	...	...	5.7E-155	4.3E-17	0.001365	1	0
97	0	0	3.8E-293	0	0	0	...	...	...	4.64E-45	2.6E-51	5.87E-53	3.69E-53	0
98	2.1E-18	9.5E-06	7.3E-63	0	0	8.34E-30	...	...	...	0	0	0	0	2.24E-13
99	0	0	0	0	0	0	...	...	...	5.68E-80	4.1E-93	4.66E-09	5.55E-14	0
100	6.1E-11	1.4E-09	4.0E-60	0	0	7.5E-30	...	...	...	0	0	0	0	5.35E-11

Table Appendix 2. Fuzzy Subtractive Clustering Degree of Membership Radius 0.3

Data	Degree of Membership (in cluster 1)	Data	Degree of Membership (in cluster 1)
1	0.007922779	**42**	0.007922779
2	9.45785E-06	**43**	9.45785E-06
3	0.141481555	**44**	0.141481555
4	0.310613494	**45**	0.310613494
5	0.119200731	**46**	0.119200731
6	1.98793E-57	**47**	1.98793E-57
7	1.46405E-07	**48**	1.46405E-07
8	3.0715E-08	**49**	3.0715E-08
9	0.054625102	**50**	0.054625102
10	0.069901216	**51**	0.069901216
11	0.397253626	**52**	0.397253626
12	0.141782866	**53**	0.141782866
13	0.060023377	**54**	0.060023377
14	0.035448008	**55**	0.035448008
15	0.020390709	**56**	0.020390709
16	0.244509782	**57**	0.244509782
17	0.334595848	**58**	0.334595848
18	0.297915903	**59**	0.297915903
19	4.06043E-66	**60**	4.06043E-66
20	0.000831195	**61**	0.000831195
21	0.313877607	**62**	0.313877607
22	0.052637666	**63**	0.052637666
23	1	**64**	1
24	0.561227691	**65**	0.561227691
25	1.18421E-60	**66**	1.18421E-60
26	0.168284776	**67**	0.168284776
27	0.580936763	**68**	0.580936763
28	0.298736753	**69**	0.298736753
29	0.527902738	**70**	0.527902738
30	0.000764753	**71**	0.000764753
31	1.80958E-15	**72**	1.80958E-15
32	8.06444E-55	**73**	8.06444E-55
33	0.082530224	**74**	0.082530224
34	0.061110377	**75**	0.061110377
35	0.061110377	**76**	0.061110377
36	0.578788297	**77**	0.578788297
37	0.212270097	**78**	0.212270097
38	0.1130885	**79**	0.1130885
39	0.244438291	**80**	0.244438291
40	8.07658E-27	**81**	8.07658E-27
41	0.405447464	**82**	0.405447464

Data	Degree of Membership (in cluster 1)	Data	Degree of Membership (in cluster 1)
83	5.0547E-119	**92**	0.150686482
84	1.2986E-106	**93**	2.6332E-85
85	0.3545226	**94**	0.115026844
86	0.000435418	**95**	0.000265094
87	0.135632956	**96**	3.78232E-87
88	2.60047E-06	**97**	2.62789E-57
89	3.55608E-55	**98**	0.485654125
90	6.7795E-139	**99**	3.65395E-72
91	0.239949452	**100**	0.771170984

Table Appendix 3. Fuzzy Subtractive Clustering Degree of Membership Radius 0.4

Data	Degree of Membership (in cluster 1)	Data	Degree of Membership (in cluster 1)
1	0.065783749	**31**	5.09786E-09
2	0.001492394	**32**	3.73635E-31
3	0.332865398	**33**	0.245807328
4	0.518052686	**34**	0.207581987
5	0.302278244	**35**	0.207581987
6	1.27453E-32	**36**	0.735220071
7	0.000143095	**37**	0.418191268
8	5.94476E-05	**38**	0.293459298
9	0.194887124	**39**	0.452736175
10	0.223883796	**40**	2.10288E-15
11	0.594943436	**41**	0.601815274
12	0.333263967	**42**	0.25680225
13	0.205496883	**43**	0.522562581
14	0.15280776	**44**	1.28068E-49
15	0.111958143	**45**	0.218902954
16	0.452810651	**46**	0.477844495
17	0.540185357	**47**	0.14126319
18	0.506031671	**48**	0.209306944
19	1.64938E-37	**49**	0.090383389
20	0.018506867	**50**	6.30477E-48
21	0.521107937	**51**	0.045571534
22	0.190866342	**52**	0.840585731
23	1	**53**	0.145743748
24	0.72258799	**54**	0.4022737
25	1.95571E-34	**55**	0.66285148
26	0.366986607	**56**	0.213753666
27	0.736753969	**57**	0.411716711
28	0.506815478	**58**	2.8905E-23
29	0.698130444	**59**	0.08109693
30	0.017659595	**60**	0.268695559

Data	Degree of Membership (in cluster 1)	Data	Degree of Membership (in cluster 1)
61	2.95905E-25	**81**	0.530935088
62	0.430222369	**82**	0.560359821
63	3.20603E-24	**83**	3E-67
64	6.26542E-70	**84**	2.74679E-60
65	0.30802172	**85**	0.55805205
66	0.466216544	**86**	0.012864267
67	8.40306E-17	**87**	0.325053944
68	0.18163346	**88**	0.000721883
69	3.94103E-24	**89**	2.35734E-31
70	4.02462E-18	**90**	1.90567E-78
71	0.425634566	**91**	0.448040596
72	9.20635E-26	**92**	0.344879032
73	3.85854E-42	**93**	2.65474E-48
74	0.10809475	**94**	0.296278101
75	6.15592E-57	**95**	0.009731116
76	7.12759E-91	**96**	2.44056E-49
77	0.392888257	**97**	1.49118E-32
78	5.1385E-24	**98**	0.666129985
79	2.12149E-49	**99**	6.55472E-41
80	0	**100**	0.864016941

Table Appendix 4. Fuzzy Subtractive Clustering Degree of Membership Radius 0.5

Data	Degree of Membership (in cluster 1)	Data	Degree of Membership (in cluster 1)
1	0.065783749	**20**	0.018506867
2	0.001492394	**21**	0.521107937
3	0.332865398	**22**	0.190866342
4	0.518052686	**23**	1
5	0.302278244	**24**	0.72258799
6	1.27453E-32	**25**	1.95571E-34
7	0.000143095	**26**	0.366986607
8	5.94476E-05	**27**	0.736753969
9	0.194887124	**28**	0.506815478
10	0.223883796	**29**	0.698130444
11	0.594943436	**30**	0.017659595
12	0.333263967	**31**	5.09786E-09
13	0.205496883	**32**	3.73635E-31
14	0.15280776	**33**	0.245807328
15	0.111958143	**34**	0.207581987
16	0.452810651	**35**	0.207581987
17	0.540185357	**36**	0.735220071
18	0.506031671	**37**	0.418191268
19	1.64938E-37	**38**	0.293459298

Data	Degree of Membership (in cluster 1)	Data	Degree of Membership (in cluster 1)
39	0.452736175	**70**	4.02462E-18
40	2.10288E-15	**71**	0.425634566
41	0.601815274	**72**	9.20635E-26
42	0.25680225	**73**	3.85854E-42
43	0.522562581	**74**	0.10809475
44	1.28068E-49	**75**	6.15592E-57
45	0.218902954	**76**	7.12759E-91
46	0.477844495	**77**	0.392888257
47	0.14126319	**78**	5.1385E-24
48	0.209306944	**79**	2.12149E-49
49	0.090383389	**80**	0
50	6.30477E-48	**81**	0.530935088
51	0.045571534	**82**	0.560359821
52	0.840585731	**83**	3E-67
53	0.145743748	**84**	2.74679E-60
54	0.4022737	**85**	0.55805205
55	0.66285148	**86**	0.012864267
56	0.213753666	**87**	0.325053944
57	0.411716711	**88**	0.000721883
58	2.8905E-23	**89**	2.35734E-31
59	0.08109693	**90**	1.90567E-78
60	0.268695559	**91**	0.448040596
61	2.95905E-25	**92**	0.344879032
62	0.430222369	**93**	2.65474E-48
63	3.20603E-24	**94**	0.296278101
64	6.26542E-70	**95**	0.009731116
65	0.30802172	**96**	2.44056E-49
66	0.466216544	**97**	1.49118E-32
67	8.40306E-17	**98**	0.666129985
68	0.18163346	**99**	6.55472E-41
69	3.94103E-24	**100**	0.864016941

Table Appendix 5. Fuzzy Subtractive Clustering Degree of Membership Radius 0.6

Data	Degree of Membership (in cluster 1)	Data	Degree of Membership (in cluster 1)
1	0.065783749	**9**	0.018506867
2	0.001492394	**10**	0.521107937
3	0.332865398	**11**	0.190866342
4	0.518052686	**12**	1
5	0.302278244	**13**	0.72258799
6	1.27453E-32	**14**	1.95571E-34
7	0.000143095	**15**	0.366986607
8	5.94476E-05	**16**	0.736753969

Data	Degree of Membership (in cluster 1)	Data	Degree of Membership (in cluster 1)
17	0.194887124	**59**	0.506815478
18	0.223883796	**60**	0.698130444
19	0.594943436	**61**	0.017659595
20	0.333263967	**62**	5.09786E-09
21	0.205496883	**63**	3.73635E-31
22	0.15280776	**64**	0.245807328
23	0.111958143	**65**	0.207581987
24	0.452810651	**66**	0.207581987
25	0.540185357	**67**	0.735220071
26	0.506031671	**68**	0.418191268
27	1.64938E-37	**69**	0.293459298
28	0.452736175	**70**	4.02462E-18
29	2.10288E-15	**71**	0.425634566
30	0.601815274	**72**	9.20635E-26
31	0.25680225	**73**	3.85854E-42
32	0.522562581	**74**	0.10809475
33	1.28068E-49	**75**	6.15592E-57
34	0.218902954	**76**	7.12759E-91
35	0.477844495	**77**	0.392888257
36	0.14126319	**78**	5.1385E-24
37	0.209306944	**79**	2.12149E-49
38	0.090383389	**80**	0
39	6.30477E-48	**81**	0.530935088
40	0.045571534	**82**	0.560359821
41	0.840585731	**83**	3E-67
42	0.145743748	**84**	2.74679E-60
43	0.4022737	**85**	0.55805205
44	0.66285148	**86**	0.012864267
45	0.213753666	**87**	0.325053944
46	0.411716711	**88**	0.000721883
47	2.8905E-23	**89**	2.35734E-31
48	0.08109693	**90**	1.90567E-78
49	0.268695559	**91**	0.448040596
50	2.95905E-25	**92**	0.344879032
51	0.430222369	**93**	2.65474E-48
52	3.20603E-24	**94**	0.296278101
53	6.26542E-70	**95**	0.009731116
54	0.30802172	**96**	2.44056E-49
55	0.466216544	**97**	1.49118E-32
56	8.40306E-17	**98**	0.666129985
57	0.18163346	**99**	6.55472E-41
58	3.94103E-24	**100**	0.864016941

Table Appendix 6. Fuzzy Subtractive Clustering Degree of Membership Radius 0.7

Data	Degree of Membership (in cluster 1)	Data	Degree of Membership (in cluster 1)
1	0.411225037	**42**	0.641528483
2	0.119450069	**43**	0.809027956
3	0.698242196	**44**	1.08413E-16
4	0.8067414	**45**	0.608936432
5	0.676607713	**46**	0.785737088
6	3.84964E-11	**47**	0.527786713
7	0.055551275	**48**	0.600088169
8	0.041698986	**49**	0.456174472
9	0.586262959	**50**	3.86943E-16
10	0.613426456	**51**	0.364772775
11	0.844032957	**52**	0.944873608
12	0.698515088	**53**	0.533195541
13	0.596499211	**54**	0.742786854
14	0.54150005	**55**	0.874353257
15	0.489201264	**56**	0.60422163
16	0.772051459	**57**	0.748435899
17	0.817837536	**58**	4.36835E-08
18	0.800580418	**59**	0.440307967
19	9.7573E-13	**60**	0.651082583
20	0.271788549	**61**	9.78544E-09
21	0.808291896	**62**	0.759258298
22	0.582285686	**63**	2.13049E-08
23	1	**64**	2.5298E-23
24	0.899339164	**65**	0.680779003
25	9.8419E-12	**66**	0.779441872
26	0.720850109	**67**	5.63426E-06
27	0.905058663	**68**	0.57293423
28	0.800985118	**69**	2.27903E-08
29	0.889284075	**70**	2.08886E-06
30	0.267661278	**71**	0.756604968
31	0.001959784	**72**	6.68356E-09
32	1.16004E-10	**73**	3.00054E-14
33	0.632427222	**74**	0.483623755
34	0.598468813	**75**	4.42062E-19
35	0.598468813	**76**	3.66628E-30
36	0.904442948	**77**	0.73708305
37	0.752258899	**78**	2.48528E-08
38	0.670097629	**79**	1.27838E-16
39	0.772009992	**80**	4.00098E-94
40	1.61232E-05	**81**	0.813237898
41	0.847203977	**82**	0.827688215

Data	Degree of Membership (in cluster 1)	Data	Degree of Membership (in cluster 1)
83	1.87114E-22	**92**	0.70637294
84	3.56003E-20	**93**	2.91734E-16
85	0.826573614	**94**	0.672192607
86	0.241355024	**95**	0.220329944
87	0.692848867	**96**	1.33822E-16
88	0.094230876	**97**	4.05213E-11
89	9.98063E-11	**98**	0.875763028
90	4.18848E-26	**99**	7.56615E-14
91	0.769386295	**100**	0.95339435

Table Appendix 7. Fuzzy Subtractive Clustering Degree of Membership Radius 0.8

Data	Degree of Membership (in cluster 1)	Data	Degree of Membership (in cluster 1)
1	0.506441929	**31**	0.008449811
2	0.196549001	**32**	2.47236E-08
3	0.75956888	**33**	0.704123285
4	0.848385979	**34**	0.674990121
5	0.741483895	**35**	0.674990121
6	1.06252E-08	**36**	0.925985825
7	0.109372053	**37**	0.804162357
8	0.087807898	**38**	0.736015514
9	0.664424727	**39**	0.820278445
10	0.687868886	**40**	0.000214143
11	0.878251539	**41**	0.880776668
12	0.759796153	**42**	0.711868386
13	0.673288675	**43**	0.850226382
14	0.625225025	**44**	5.98219E-13
15	0.578447464	**45**	0.684010716
16	0.820312177	**46**	0.831422258
17	0.857305682	**47**	0.61306603
18	0.843421028	**48**	0.676388023
19	6.3728E-10	**49**	0.548304936
20	0.368835936	**50**	1.58459E-12
21	0.849634075	**51**	0.462033444
22	0.660970897	**52**	0.957514817
23	1	**53**	0.617870532
24	0.921982514	**54**	0.796398457
25	3.73961E-09	**55**	0.902306441
26	0.77832782	**56**	0.679952215
27	0.926468421	**57**	0.801031554
28	0.843747438	**58**	2.31869E-06
29	0.914079869	**59**	0.533643353
30	0.364539996	**60**	0.719971196

Data	Degree of Membership (in cluster 1)	Data	Degree of Membership (in cluster 1)
61	7.37544E-07	**81**	0.853611696
62	0.809884798	**82**	0.865200471
63	1.33811E-06	**83**	2.31517E-17
64	5.00308E-18	**84**	1.28738E-15
65	0.744981238	**85**	0.864308287
66	0.826317442	**86**	0.336779977
67	9.57435E-05	**87**	0.755072859
68	0.652828287	**88**	0.163914242
69	1.40897E-06	**89**	2.20346E-08
70	4.479E-05	**90**	3.71545E-20
71	0.807716998	**91**	0.818143231
72	5.50835E-07	**92**	0.766331535
73	4.43206E-11	**93**	1.27645E-12
74	0.573391367	**94**	0.73777662
75	8.85775E-15	**95**	0.314080278
76	2.9056E-23	**96**	7.02866E-13
77	0.791712063	**97**	1.10505E-08
78	1.5056E-06	**98**	0.903420094
79	6.78673E-13	**99**	8.99784E-11
80	3.10659E-72	**100**	0.964118829

Table Appendix 8. Fuzzy Subtractive Clustering Degree of Membership Radius 0.9

Data	Degree of Membership (in cluster 1)	Data	Degree of Membership (in cluster 1)
1	0.746080464	**20**	0.70730003
2	0.309353434	**21**	0.80182171
3	0.869515906	**22**	0.760786183
4	0.843924803	**23**	0.885460005
5	0.753978059	**24**	0.911078536
6	1.47979E-06	**25**	1.79715E-06
7	0.345197495	**26**	0.705377501
8	0.281401572	**27**	0.937115098
9	0.61108692	**28**	0.703890507
10	0.888537084	**29**	0.812934484
11	0.852298054	**30**	0.6381877
12	0.699959022	**31**	0.037168592
13	0.552663182	**32**	6.38246E-06
14	0.544301301	**33**	0.743187474
15	0.777567856	**34**	0.665584223
16	0.812642688	**35**	0.665584223
17	1	**36**	0.80097415
18	0.754729096	**37**	0.891317544
19	4.63586E-07	**38**	0.726496266

Data	Degree of Membership (in cluster 1)	Data	Degree of Membership (in cluster 1)
39	0.801380931	70	0.000412058
40	0.000941636	71	0.894625526
41	0.702590183	72	2.43404E-05
42	0.660638663	73	2.94088E-08
43	0.82868471	74	0.87238262
44	1.11375E-09	75	5.2048E-11
45	0.744580429	76	2.5646E-17
46	0.829296813	77	0.783942914
47	0.746142375	78	4.60134E-05
48	0.646504466	79	1.2152E-09
49	0.850883468	80	5.59979E-60
50	5.59079E-09	81	0.804330587
51	0.649208652	82	0.948480812
52	0.88438474	83	6.39964E-13
53	0.719805884	84	1.33917E-11
54	0.678368995	85	0.890646756
55	0.822706707	86	0.436200244
56	0.864278568	87	0.786239775
57	0.703816692	88	0.40441887
58	6.58436E-05	89	2.54732E-06
59	0.688009758	90	5.37208E-15
60	0.724011199	91	0.874616388
61	2.72423E-05	92	0.70329804
62	0.955670933	93	2.18754E-09
63	4.42437E-05	94	0.909084684
64	1.79402E-13	95	0.645905124
65	0.731431971	96	1.3776E-09
66	0.949703226	97	1.42519E-06
67	0.000686236	98	0.839757391
68	0.924393965	99	4.90883E-08
69	3.5449E-05	100	0.851557602

Table Appendix 9. Fuzzy Subtractive Clustering Degree of Membership Radius 1

Data	Degree of Membership (in cluster 1)	Data	Degree of Membership (in cluster 1)
1	0.78878091	9	0.671032292
2	0.386605922	10	0.908713965
3	0.892924625	11	0.878575548
4	0.871577556	12	0.749046568
5	0.795537312	13	0.618575952
6	1.89611E-05	14	0.610984076
7	0.422507853	15	0.815639073
8	0.358058906	16	0.845315246

Data	Degree of Membership (in cluster 1)	Data	Degree of Membership (in cluster 1)
17	1	**59**	0.738671978
18	0.796179123	**60**	0.769827959
19	7.40569E-06	**61**	0.0002007
20	0.755403474	**62**	0.963939518
21	0.83618626	**63**	0.000297262
22	0.801350882	**64**	4.73806E-11
23	0.906164093	**65**	0.776212963
24	0.927342599	**66**	0.959060954
25	2.21929E-05	**67**	0.002738695
26	0.753739887	**68**	0.938305479
27	0.948751067	**69**	0.000248416
28	0.752452583	**70**	0.001811824
29	0.845561095	**71**	0.913754316
30	0.695037485	**72**	0.0001832
31	0.069477255	**73**	7.93354E-07
32	6.19499E-05	**74**	0.895308431
33	0.786302559	**75**	4.68044E-09
34	0.719108538	**76**	3.64171E-14
35	0.719108538	**77**	0.821051494
36	0.835470241	**78**	0.000306857
37	0.911016596	**79**	6.00571E-08
38	0.771967546	**80**	1.01396E-48
39	0.835813907	**81**	0.838304917
40	0.003538714	**82**	0.95806092
41	0.751326458	**83**	1.32735E-10
42	0.714777427	**84**	1.55861E-09
43	0.858806561	**85**	0.910461209
44	5.59627E-08	**86**	0.510676158
45	0.787496096	**87**	0.822999475
46	0.859320351	**88**	0.480323066
47	0.788833927	**89**	2.94393E-05
48	0.702365124	**90**	2.76322E-12
49	0.877394221	**91**	0.89716488
50	2.06754E-07	**92**	0.751939535
51	0.70474383	**93**	9.66862E-08
52	0.90527266	**94**	0.925698406
53	0.766204094	**95**	0.701837664
54	0.730276713	**96**	6.648E-08
55	0.853784922	**97**	1.83923E-05
56	0.888565675	**98**	0.868089706
57	0.752388668	**99**	1.20141E-06
58	0.0004102	**100**	0.87795724