Grouping 2 Hour Load Data Substation Using Hierarchical Clustering

Yusmartato\(^1\), Indra Gunawan\(^2\), Murni Sari Rahayu\(^2\), Dedi Kusbiantoro\(^2\), Surya Dharma\(^2\), Anuar Sadat\(^1\), Rachmat Rizaldi\(^1\), Supriadi\(^1\), Zenni Riana\(^1\), Dediyarwan Nst\(^6\), Mayang Sari Ayu\(^6\), Nanda Novziansyah\(^6\), Aulia\(^6\), Ismurrizal\(^6\), Dewi Pangestuti\(^6\), M. Joesoef S\(^6\), Tri Makmun\(^6\), Zain Anshari\(^6\), Nondang Purnama Siregar\(^6\), Hardy Hasibuan\(^6\), Dian Afriandi\(^6\), Budi Kurniawan\(^6\), Isnaini Halimah Rambe\(^7\), Evi Susilawati\(^7\) and Adilah Wirdhani Lubis\(^7\)

\(^1\)Faculty of Engineering, Universitas Islam Sumatera Utara Medan, Indonesia
\(^2\)Faculty of Agriculture, Universitas Islam Sumatera Utara, Medan, Indonesia
\(^3\)Faculty of Social and Political Science, Universitas Islam Sumatera Utara Medan, Indonesia
\(^4\)Faculty of Education and Teaching Training, Universitas Islam Sumatera Utara, Medan, Indonesia
\(^5\)Faculty of Economic, Universitas Islam Sumatera Utara Medan, Indonesia
\(^6\)Faculty of Medicine, Universitas Islam Sumatera Utara Medan, Indonesia
\(^7\)Faculty of Teacher Training and Education, Universitas Islam Sumatera Utara, Medan, Indonesia

\(*\)yusmartato@ft UISU.ac.id

Abstract. In this modern era, information technology is very influential in many ways, especially in the business and business world. Computer is a device that is needed for data processing, so the data processed can be effectively and efficiently to provide information needed by individuals, companies, and agencies. Clustering is a method for finding and grouping data that has similarity between one data and others. In the clustering algorithm, data will be grouped into clusters based on the similarity of one data to another. The principle of clustering is to maximize the similarity between members of one cluster and minimize similarities between different cluster members. The most well-known category of clustering algorithms is Hierarchical Clustering. As well as PLN (State Electricity Company), where PLN is a state-owned enterprise (BUMN) that provides electricity services to all communities and all circles, both private and business entities whose job it is to serve users in electricity. Services will be maximized and can be accessed by all parties if in the service there is a role for the media as a assistant in conveying information to all circles. This information will be right on the target if the information is processed from a reliable source. PT PLN (Persero) Krueng Geukuh did not yet have a load data collection system that makes staff still did data collection by manual methods. The result of this study shows the system of 2 hour load cluster data grouping using hierarchical clustering of PT PLN (Persero) Krueng Geukuh is an effective solution to help staff record computerized loads and can increasing the work performance in data collection. Designing systems of Grouping Data Load 2 Hours Substation Using Hierarchical Clustering PT PLN (Persero) Krueng Geukuh was built by using PHP so it has a more attractive appearance and uses a MySQL database so that load managers become more effective.
1. Introduction
The development of information technology today in the world is not only of industry, trade, services and production, but these advances have expanded into the world of service [1][2]. By utilizing computers as a means, all service information can be accessed and enjoyed by anyone for certain purposes.

In this modern era, information technology is very influential in many ways, especially in the business and business world. Computer is a device that is needed for data processing, so the data processed can be effectively and efficiently to provide information needed by individuals, companies, and agencies [3][4]. Data processing was done by manually will require a lot of time and energy and is not efficient. The optimal utilization of computers and information systems in an agency or company is the right choice to processing data efficiently to obtain the information needed [5].

Service in the business world can be interpreted as a center in serving all the information needed both for organization purposes or for purposes outside the organization. Services will be maximized and can be accessed by all parties if in the service there is a role for the media as a assistant in conveying information to all circles. This information will be right on the target if the information is processed from a reliable source. As well as PLN (State Electricity Company), where PLN is a state-owned enterprise (BUMN) that provides electricity services to all communities and all circles, both private and business entities whose job it is to serve users in electricity.

Clustering is a method for finding and grouping data that has similarity between one data and others [6]. Clustering is one of the data mining methods that is non-directive (unsupervised). It is means by unsupervised method was applied without training and teacher did not require an output target [7].

2. Related Works
Objects in the cluster have similar characteristics between each other and are different from other clusters. Partitions are not done manually but with a clustering algorithm. Therefore, clustering is very useful and can find unknown groups or groups in the data. Clustering is widely used in various applications such as business intelligence, image pattern recognition, web search, the field of biology, and for security [9]. In business intelligence, clustering can manage a lot of customers into various groups. For example, grouping customers into clusters with strong common characteristics. Clustering is also known as data segmentation because clustering partitioned many data sets into many groups based on their similarity. In addition clustering can also be an outlier detection [8][10].

3. Research Methodology

3.1. Clustering of Hierarchy
The clustering builds a cluster hierarchy or in other words a cluster tree, also known as a dendrogram. Each cluster node contains several cluster; Clusters that share the points covered by previous cluster above. The clustering hierarchical methods are categorized into agglomerative (bottom-up) and divisive (top-down) [11][12]. Agglomerative clustering begin with cluster one point (singleton) and repeatedly combines two or more of the most appropriate clusters.

In the clustering algorithm, data will be grouped into clusters based on the similarity of one data to another. The principle of clustering is to maximize the similarity between members of one cluster and minimize similarities between different cluster members. The most well-known category of clustering algorithms is Hierarchical Clustering [13][14]. Hierarchical Clustering is one of the clustering algorithms that can be used to cluster documents (document clustering). The overall results of the hierarchical clustering algorithm graphically can be described as a tree, called a dendogram. This tree graphically describes the process of combining existing clusters, resulting in higher levels of clusters.
3.2. Database
Database is a collection of information stored in a computer systematically and structured so that it can be checked using a computer program to obtain information from the database.

Database consists of base words and data. Base can be interpreted as a base or warehouse. While data is a record of a collection of real-world facts that represent objects such as humans, goods, animals, concepts, events and so on that are manifested in the form of letters, numbers, symbols, images, text, sounds or combinations thereof. Database or database is a collection of data that forms a file was interconnected (relations) with certain procedures to form new data or information [15]. Or a database is a collection of interconnected data (relations) between one another which is organized based on a particular scheme or structure. On computers, databases were stored in storage hardware devices, and with certain software manipulated for certain interests or uses. Relationships or data relations are usually indicated by a key from each file that exists. Data is a fact or value that is recorded or represents a description of an object.

3.3. Programming Tools
In designing and developing an information system certainly requires several tools in the form of a programming language or a tool. In this information system uses several tools, namely: Personal Home Page (PHP) and XAMPP as a web server, using a database from MySQL, Sublime Text for application coding, as well as HyperText Markup Language (HTML) and CSS to create application display.

4. Result and Discussion
4.1. Former System Analysis
After the author conducted a research on the former system (using manual methods) was running, the problems were obtained in the former system have to be evaluated and designed a new system in order to overcome the problems that occur in the old system. From the author's observations while undergoing practical work at PT. PLN (Persero) Krueng Geukueh District, almost all activities in various fields of the company had used a computerized system. In the Engineering staff section, the information system used to manage the data, especially regarding data recording the load of 2 hours of relay stations, still uses Excel and the agenda book as a data storage media in special filing cabinets. Because managing data still uses manual methods in data collection, so that often the occurrence of duplicate data is excessive and this problem can also make it difficult for the staff in searching for previously stored data manually. This system can be an effective solution to help staff record computerized loads, can speed up work in data collection and facilitate the process of finding data and information in the agency.

4.2. New System Analysis
The system to be built is a search system that will be used by staff or PT PLN (Persero) Krueng Geukuh. This system aims to store incoming load data. In this latest system, staff will be able to input load data in/out, perform KH searches, view loads and print load reports quickly and efficiently. With the existence of this system, the staff will be very helpful to make and store load data quickly and efficiently.

4.3. System Design
The design of a system will be done after an analysis of a system has been passed. Designing can be defined as drawing, planning and making a pattern or sketch of a number of separate elements into one whole unit or often called an interface. The design of this information system was built aimed at facilitating the part of PT PLN Krueng Geukuh in terms of recording the burden, especially for staff at the company.
Region of PT PLN Kreueng Geukuh can do archiving just by opening the search system and then doing the next step. The design of the archiving search system built is object oriented by using the Unified Modeling Language (UML) as a modeling language.

### Table 1. User Table

| No | Field Name   | Data Type | Size | Information     |
|----|--------------|-----------|------|-----------------|
| 1  | User_ID      | Integer   | 11   | Primary Key     |
| 2  | User_Name    | Vanchar   | 30   |                 |
| 3  | Address      | Text      | 50   |                 |
| 4  | Picture      | Vanchar   | 50   |                 |
| 5  | Username     | Vanchar   | 15   |                 |
| 6  | Password     | Vanchar   | 8    |                 |
| 7  | Level        | Vanchar   | 30   |                 |

### Table 2. Load Table

| No | Field Name   | Data Type | Size | Information     |
|----|--------------|-----------|------|-----------------|
| 1  | Code         | Vanchar   | 10   | Primary Key     |
| 2  | Feeder       | Vanchar   | 10   |                 |
| 3  | Current      | Vanchar   | 10   |                 |
| 4  | Cos_Q        | Vanchar   | 15   |                 |
| 5  | Voltage_TM   | Vanchar   | 10   |                 |
| 6  | Power        | Vanchar   | 10   |                 |
| 7  | Information  | Vanchar   | 100  |                 |

### 4.4. System Implementation

Implementation is an application of a system that has been analyzed. Implementation steps cannot be carried out before the system analysis phase has been completed. The implementation phase starts with the interface process or the design of a system interface. This interface design will be a container for inputting, changing and deleting a data that will be managed. After the interface has been created, the system will also be used. Here, the author will explain the implementation of a 2-hour load clustering data grouping using hierarchical clustering.

![Login Form](image)

#### Figure 1. Login Form

In the picture above shows the beginning form or the start page of a system. Where if a user wants to use the system, the user must first enter or be called login. This form functions for system user validation. If the user wants to enter the system, at first the user does is enter the username and password on the login form. After the user enters a username and password, the next is the user does to press the login button. After pressing the login button, the system will automatically read the validity and data entered. There are two choices, if the data entered is invalid, the system will display
5. Conclusion

From the research that has been done conclusions can be drawn:

1. PT PLN (Persero) Krueng Geukuh did not yet have a load data collection system that makes staff still did data collection by manual methods. Therefore, the author attempt to design a system of 2 hour load cluster data grouping using hierarchical clustering of PT PLN (Persero) Krueng Geukuh which is an effective solution to help staff record computerized loads, can increasing the work performance in data collection.

2. Designing systems of Grouping Data Load 2 Hours Substation Using Hierarchical Clustering PT PLN (persero) Krueng Geukuh was built by using PHP so it has a more attractive appearance and uses a MySQL database so that load managers become more effective.

References

[1] H. Nuryanto, Sejarah Perkembangan Teknologi Informasi dan Komunikasi. PT Balai Pustaka (Persero), 2012.
[2] M. E. P. M. E. Porter, Competition in global industries. Harvard Business Press, 1986.
[3] K. C. Laudon and J. P. Laudon, Management information system. Pearson Education India, 2016.
[4] D. Abdullah, “Merancang Aplikasi Perpustakaan Menggunakan SDLC,” 2017.
[5] E. Irwansyah and M. Faisal, Advanced Clustering: Teori dan Aplikasi. DeePublish, 2015.
[6] A. K. Jain, M. N. Murty, and P. J. Flynn, “Data clustering: a review,” ACM Comput. Surv., vol. 31, no. 3, pp. 264–323, 1999.
[7] L. Kaufman and P. J. Rousseeuw, “Finding groups in data, 1990,” New York, pp. 22–52, 1990.