This chapter present a review of the related literature of this study. The discussion of relevant concepts in this study is presented, including Customer Segmentation Based on Electricity Consumption Data, Customer Segmentation Based on Customer Lifetime Value. The related literature is classified as to clarify the knowledge gap and this research’s position. Finally, the proposed research is presented at the end of this chapter.

**2.1 Previous Segmentation Studies Based on Elctricity Consumption Data**

Table 1 presents an overview of previous studies that focuses on customer segmentation using transaction/ customer credentials data. As shown, we categorize related articles based on its business context, dataset, segmentation features, and the segmentation method.

Previous studies in customer segmentation in electricity consumption have explored various dimensions of the customer clustering problem[4], [7], [8], [10],[12]. They use the context of electricity consumption as a case study to find out patterns of electricity use in predicting future electricity consumption. Several clustering models, one of which is often used, namely K-Means Clustering, has explored customer grouping by considering patterns of electricity use and electricity demand to meet electricity consumption based on what has been prepared by company[7], [9]– [11].

A context study of load profile electricity [4] using experimental data by installing 4000 intelligent meters in several homes in Ireland with existing methods used to classify household electricity use, in general, can be divided into four categories, statistics, manipulation, time series, and clustering. Statistical methods have been widely used in the unregulated power market to form a standard load Personal Classes (PC). A typical load PC is used for settlement purposes and estimates the amount and Time of Use of electricity used. A series of PCs are manufactured for different market segments (e.g., residential, commercial, industrial) and derived on an average for all customers within a customer class.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Article | Business Context | Dataset | Segmentation  Features | Segmentation  Method |
| [4] | Electricity Load Profile in Ireland | Experimental data period January 1, 2009, to December 31, 2010, | Dwelling type, No. of bedrooms, Age, Social Class, Electronic Type | K-means, k-medoid and Self Organizing Maps (SOM) |
| [6] | Electricity Consumption in South Africa | South Africa Electric Load Profile Data from 1994 to 2014 | X=Hour (load profile multiple 1 day)  Y= X multiple All household | K-Means  And Self Organizing Maps (SOM) |
| [5] | Electricity  Demand Signature in Andalusian | The load data of 64 buildings located in Andalusia, Spain | Identity, Industrial Division, Industrial Categories, Mean Power Consumption, Power Consumption | Variable selection (Feature Selection), Model (K-Means, Hierarchical Clustering, K-Medoid Clustering), Validation (Connectivity, Dunn and  Silhouette indexes) |
| [10] | Electricity Load Profile | Smart Metering Data in 2009 | Identity, Social Status, age, gender, Demand kWh, Income | Regression Ordinary Least Square (OLS), Evaluation (Root Mean Square Error (RMSE)) |
| [9] | Electricity Load Profile | Residential Demand Data during November,2017 until February, 2018 | Identity, Daily Consumption, Load Profile, Peak Hour, Demand | K-means, Fuzzy C-Means (FCM) and Self Organizing Maps (SOM) |
| [8] | Electricity Consumption Forecasting | Electricity Consumption Data from 46 homes in Texas | Identity, Time, Total kWh | Model (Artificial neural networks, regression  trees, random forest regression, 𝑘-nearest neighbors’ regression,  and support vector regression), Evaluation (Naive forecast, random  forecast, the ARIMA model, and stepwise regression) |
| [15] | Electricity Demand with Renewable Technologies | Half -hourly energy use for 1 year data | Average energy use,  energy–temperature correlation, entropy of the load-shape representative vector, and distance to  wind generation patterns. | Model (K-Medoids), Validities (average silhouette) |
| [11] | Electricity Consumption in Indonesia | Customer Transaction in September 2021 | Rate, Power, Total kWh, Total Cost, Flash Time | Variable selection with correlation  Model (K-Means)  Validity (Silhouette Method)  Explores (Customer Relationship Management (CRM)) |

A context study of load profile electricity [4] using experimental data by installing 4000 intelligent meters in several homes in Ireland with existing methods used to classify household electricity use, in general, can be divided into four categories, statistics, manipulation, time series, and clustering. Statistical methods have been widely used in the unregulated power market to form a standard load Personal Classes (PC). A typical load PC is used for settlement purposes and estimates the amount and Time of Use of electricity used. A series of PCs are manufactured for different market segments (e.g., residential, commercial, industrial) and derived on an average for all customers within a customer class.

Research on electricity consumption in South Africa [6] focus on household customers, which aims to classify customers based on patterns and types using electricity using the K-Means clustering model and Self Organizing Maps (SOM). They used internal and external validation to evaluate the clustering structure based on the expected behavior of South African households' daily electricity consumption. Another study used electrical load data also in Andalusia, Spain [5], but the research context was about electricity demand. They determine interrelated variables to predict customer segmentation models using a combination model between K-Means clustering and K-medoid clustering. This study aims to provide an alternative customer segmentation that can manage several types of customers. It then presents the segmentation results based on the characteristics of the load curve. Finally, they compare the two marks and provide solutions to the effects of classification and segmentation.

Research on the context of electricity load data [9] uses electricity demand data to predict electricity loads per day based on the heterogeneity of electricity demand behavior by customers, then processed using a combination of K-Means clustering models and Self Organizing Maps (SOM) and Fuzzy C-Means. The segmentation results provide the proper group identification for electricity demand per day. The result shows a tremendous impact because it can save on utility costs based on electricity reduction by customers. Another study with the same context as [9], but this study uses data from smart meters in 2009 [10], they use a regression model with an evaluation of the root mean square error for customer segmentation based on electricity demand used, age, and income from the customer. The aim is to find new customer electricity usage behavior patterns based on predetermined variables. Another study uses six regression models to predict daily electricity consumption based on the total electricity consumption used by customers [8]. They compared the models to find new patterns of customers' daily electricity usage.

Research on the context of looking for energy reserves based on the number of customer electricity requests [15] uses data on customers' half-day electricity usage by selecting variables based on the average amount processed by adding wind variables as alternative electrical energy. This study uses the K-Medoid model and the Silhouette method to validate the number of clusters to apply an efficient time series clustering methodology that explicitly considers the pattern of renewable energy generation. Other research on the context of electricity consumption in Indonesia [11]. They used data on customer electricity bills in September 2021 with predictors of power, rate, total kwh, flash sale, total cost, which were tested for variable correlation. This research uses the K-Means Clustering model and the Silhouette Method as the number of clusters to get customer segmentation based on the characteristics of customers paying for electricity according to the power used. The clustering results will be explored using the CRM model to gain insight to act to customers in the future according to the wisdom that has been carried out.

**2.2. Previous Studies on Segmentation Based on Customer Lifetime Value**

Previous studies in customer segmentation have explored various dimensions of customer clustering problems [16]–[18]. Many of them use the marketing context as a case study. The K-Means clustering model and Customer Lifetime Value explores customer grouping by considering the specified product preferences and predicting customer behavior in buying products offered by the company [13]

A context study in marketing combines the Customer Lifetime Value (CLV) and K-Means models in each customer segment [16]. The grouping uses the K-Means Clustering method based on the LRFM (Length, Recency, Frequency, Monetary) model. The cluster formation process uses the Elbow method. The CLV value is generated from the multiplication of the LRFM normalization results, and then the LFRM weight value uses the Analytical Hierarchy Process (AHP). Based on the LRFM matrix, this cluster has a high loyalty value, with the symbol LRFM being a loyal customer (the best segment with a high customer loyalty value). Based on the LRFM symbol, companies can create strategies to retain customers and acquire loyal customers with high profitability.

Another study with a supermarket marketing context with the same objective and predictor variables used historical customer data processed with a combination of LRFM models to determine data selection on potential customer purchases [13]. The K-means clustering model to map customers based on the same characteristics is then classified to distinguish potential customers for repurchase and then validated using the elbow method. This study uses data from all AR-Pulsabiz ​​pulse server operators in Malang, Indonesia, to predict the future of Small and Medium Enterprises. The number of potential customers who will become operators by using a combination of the K-Means Clustering model and the LRFM model to group customers to provide services according to priority.

Research in pharmaceutical marketing [18] also has the same objective [14], but they use eight validation methods in determining the correct number of groupings. Another transportation survey uses the K-Means Clustering model and the CLV model to group customers [19] with the same research objective [3]. It also has similar goals and models [3] to marketing research in Telecommunication Companies [20]. However, they do not use the CLV model but use the Neural Network to classify priority customers after getting the results from clustering.

**2.3. Marketing Strategy in Customer Relationship Management**

Previous research, marketing strategy in customer segmentation was determined based on the CLV result, then we can evaluate target that aims to develop customer service improvement strategies based on the concept of customer relationship CRM [21] .

There are two programs from the customer relationship strategy. If the company chooses the right approach, it will increase profits and retain customers [21], [22] as follows.

1. Sustainable Marketing

This program is a program to maintain and increase customer loyalty through special long-term services and increase value by studying the characteristics of customers [28]– [31]. Implementing a sustainable marketing program from this concept will be explained as follows.

A. Continuous Replenishment Program

This program is used for less profitable customers[27]. Approaches to programs such as partnership programs to encourage increased use of the company's services to customers [27], [28].

B. Business to Business

This program is used for profitable customers[29], [30]. The approach to this program is like providing special executive services to customers to improve service, so that customer trust will increase and become more loyal [36]– [39].

2. One to One Marketing

This program is an individual program aimed at satisfying customers' unique needs [35], [36]. This program uses customer information from online news and databases, followed by personal interactions to meet customers' unique needs [37], [38]. Build interactive marketing and post-marketing programs in developing customers using individual customer information [42]– [44]. The application of the one-to-one marketing program from this concept will be explained as follows.

A. Customer Business Development

This program is used for profitable customers [40], [41]. The approach to this program is to assess the benefits of marketing, finance, and management business processes [42], [43]. This program aims to explore the customer's business development by providing the best solutions and consulting regarding customers' services [40]– [42], [44].

B. Retail Account Marketing

This program is used for less profitable customers [39], [45]. The approach to this program sees the customer as a partner to develop business opportunities. This program performs customer profiling further by using CRM, which is more integrated into the application [46], [47].

2.4 Research Position