Cluster Analysis for Customer Segmentation Based on the Reasons for Choosing IndiHome Products Using the K-Means Method at PT Telkom Access Witel SBU

Iin Mutmainnah[1*]

[1]Department of Management, Faculty of Economics, Universitas Negeri Surabaya

Abstract

The Internet is a global communication network that interconnects computer networks, facilitating the sharing of information and communication worldwide. Therefore, Bambang Gunawan, the Director of Information and Political Communication, Legal, and Security at the Ministry of Communication and Information Technology, has stated that the number of internet users in Indonesia has reached 202 million people. Consequently, this positions Indonesia as the fourth-largest internet user in the world in the year 2020.

PT Telkom Indonesia is a State-Owned Enterprise (SOE) operating in the field of information technology and communication (ITC) services and telecommunication networks in Indonesia. IndiHome is one of the flagship programs and projects of Telkom, Indonesia Digital Network 2015. IndiHome is a bundled triple-play service product from Telkom, comprising high-speed communication and data services, including home phone (voice), internet (fiberoptic/high-speed internet), and interactive TV services (TV Cable, IPTV, & Netflix). IndiHome services are exclusively available to homes with fiber optic cable networks provided by Telkom FTTH (Fibre To The Home) and areas that still rely on optical cables.

The purpose of this research is to categorize IndiHome customers based on their reasons for using the service at PT Telkom Indonesia. The methodology employed in this study is quantitative analysis, involving data collection through survey techniques using Google Forms. The sample size includes 30 respondents who are IndiHome users within the specified region. The K-Means method is applied using SPSS for the purpose of data analysis.

Based on the research findings and analysis conducted, it can be concluded that there are three consumer groups at PT Telkom Indonesia. Within these three consumer groups, a total of 7 (seven) attributes or variables related to the reasons for using IndiHome are identified. Among these variables, 7 (seven) possess a Significance value of <0.05, and thus, they can be considered differentiating variables between clusters. Specifically, there is one respondent in cluster 1, 18 respondents in cluster 2, and 11 respondents in cluster 3.

This clustering result can subsequently serve as a recommendation for marketing managers at PT Telkom Indonesia to enhance their marketing system in order to better understand the reasons behind customer usage of IndiHome.

Keywords:

Clustering; Customers; IndiHome; K-Means Method"

JEL Code: Written by editor

Received Month Date Years; Received in revised form Month Date Years; Accepted Month Date Years; Available online Month Date Years (Written by editor)

Email:iinmutmainnah.20038@mhs.unesa.ac.id

1. Introduction

The internet constitutes a global communication network connecting computer networks that facilitate the exchange of information and communication worldwide. Consequently, Bambang Gunawan, the Director of Information and Political Communication, Law, and Security at the Ministry of Communication and Informatics, asserted that the number of internet users in Indonesia has reached 202 million individuals. This places Indonesia as the fourth-largest internet user in the world as of the year 2020.

PT Telkom Indonesia is a State-Owned Enterprise (SOE) specializing in information and communication technology (ICT) services and telecommunications networks in Indonesia. Telkom's legal status is a Limited Liability Company (PT), referred to as such due to the capital used for its establishment, which consists of tradable shares. IndiHome stands as one of Telkom's principal programs and projects, part of the Indonesia Digital Network 2015 initiative. IndiHome serves as a bundled triple-play service, encompassing high-speed communication and data packages, including home phone services (voice), internet connectivity (fiber-based/high-speed internet), and interactive television services (TV Cable, IPTV, and Netflix). IndiHome services are exclusively available in

households with access to Telkom's Fiber To The Home (FTTH) optical cable network and areas utilizing optical cables.

Customers are individuals whose activities involve the continuous purchase and use of a product, be it a commodity or service. Customers or users of a product are individuals who are directly or indirectly associated with business enterprises.

K-Means is one of the methods in data mining used to partition existing data into several clusters, ensuring that data with similar characteristics are grouped into one cluster, while data with distinct characteristics are placed in another cluster.

Cluster analysis is a statistical technique aimed at segregating cases/objects into multiple groups with differing attributes among the groups. In this analysis, each group exhibits homogeneity among its members, or it can be said that the variation among objects/individuals within one formed group is kept as minimal as possible."

2. Literature review

2.1 Customers behavior

Consumer behavior reflects the entirety of consumer decisions related to the acquisition, consumption, and disposition of goods, services, activities, experiences, individuals, and ideas by decision-making units (i.e., individuals) [over time] (Hoyer *et al*., 2021:4).

Consumer behavior is the study of consumer actions when seeking, purchasing, using, evaluating, and disposing of products and services they anticipate will fulfill their needs (Schiffman & Kanuk, 2015:30).

Consumer behavior studies elucidate which products and brands consumers purchase, why they make these purchases, when they make them, where they make them, how frequently they make such purchases, how

often they utilize these products, how they evaluate them post-purchase, and whether they engage in repeat purchases (Schiffman & Kanuk, 2015:30).

Contemporary consumer behavior is widely comprehended as a process involving an individual's engagement in information search, decisionmaking, purchase activities, usage, evaluation, and management (disposition) of products or services to satisfy their needs and (Rumondang *et al*., 2020:14)

2.2 Purchase Interest Process

As defined by Kotler and Keller (2009:15), the concept of purchase interest represents behavior that emerges in response to objects indicating a consumer's desire to make a purchase.

According to Tjiptono (2019:152), there are various roles in the purchasing decision process:

a) Users, referring to individuals who will use the purchased goods or services.

b) Influencers, these are individuals who exert influence on the purchasing decision. They often assist in determining product specifications and provide information for evaluating available alternatives.

c) Deciders, individuals who make determinations regarding product requirements and/or suppliers.

d) Approvers, individuals who authorize actions proposed by decisionmakers or buyers.

e) Buyers, those with the official authority to select suppliers and establish purchasing terms. Buyers may assist in specifying product requirements, but they play a crucial role in supplier selection and negotiation.

f) Gatekeepers, individuals who control the flow of information to other parties. For example, purchasing agents are authorized to prevent salespeople from directly contacting users or decision-makers.

As defined by Engel, Blackwell, and Miniard (2004:294), they present a model of the purchase decision process that involves several stages, including problem recognition, information search, alternative evaluation, purchase decision, and post-purchase behavior.

2.2 Purchase Decision

Consumer purchase decision encompasses two influencing factors: the company's marketing efforts (i.e., product, price, promotion, and place) and social influences (such as family, friends, neighbors, social class, and cultural and subcultural entities). This stage also includes the methods by which information from the company and social sources is channeled to consumers (Schiffman & Kanuk, 2015:47).

Consumer purchase decision can be entirely based on information retained in memory. Thus, consumers engage in external search from external sources, such as trusted dealers, friends or relatives, published sources (magazines, brochures, or books), advertisements, the Internet, or product packaging. Consumers employ external search to gather additional information about available brands, as well as attributes and benefits associated with brands in their consideration (Hoyer *et al*., 2021:169).

The cognitive model of purchase decision depicts how consumers systematically utilize attribute information to make decisions. Marketers should also recognize that consumers can make decisions based on their feelings or emotions by using the affective decision-making model. Therefore, marketers need to understand how consumers make choices when the decision is more cognitive or emotional in nature (Hoyer *et al*., 2021:195).

Different purchase decision models can explain how consumers make distinct decisions effectively. Due to their opportunistic and adaptable nature, consumers do not always follow a consistent process when making decisions. Instead, they may choose one model or utilize elements from multiple models, depending on the situation, and they may use one or more purchase decision rules, sometimes merely for the sake of changing things. Furthermore, consumer choices may be related to other choices. For example, one decision (buying a computer) may lead to another decision (Hoyer *et al*., 2021:195).

Purchase decision is where consumers can evaluate one brand at a time. Therefore, consumers interested in purchasing a laptop are likely to gather information about Apple models and evaluate that model before moving on to the next brand (Hoyer *et al*., 2021:196).

According to Kotler and Keller (2016: 200), the purchase decision is the stage where consumers actually make a purchase of a product.

Purchase Decision, as defined by Kotler and Keller (2016:198), involves the evaluation stage, during which consumers form preferences among available brands and may also form an intention to purchase their preferred brand.

2.4 Customer Expectations

In accordance with Samuel (2006) cited in (Agustin *et al*., 2018), customer expectations refer to what customers desire or anticipate to find in a retail environment, particularly when they are engaged in shopping activities.

According to Zeithaml (2013:87), customer expectations encompass all the desires related to a product that lead to the formation of customer

beliefs, which serve as guidelines for assessing the outcomes of product usage.

In accordance with the perspective of Fandy Tjiptono (2015:146), customer satisfaction is an evaluation or judgment made after a purchase, wherein, from among various product alternatives selected or purchased, at least an outcome that is equivalent to or exceeds customer expectations is anticipated. Conversely, if the perceived outcome falls short of expectations, customer dissatisfaction is likely to ensue.

As Schifman and Kanuk (2007) elaborated in Rangga Pratama (2009:25), individuals typically base their expectations on what they anticipate seeing, and these expectations are often rooted in past habits and experiences.

2.5 Customer Satisfaction

According to Kotler and Keller (2007), customer satisfaction is the emotional response of an individual upon comparing the performance (outcomes) of a product under consideration with the anticipated performance (outcomes). If the performance falls below expectations, customers express dissatisfaction.

According to Fandy Tjiptono (2015:146), customer satisfaction is an assessment or judgment carried out after a purchase, wherein from various product alternatives selected or purchased, at least an outcome that is equivalent to or exceeds the average customer expectation should be obtained. However, if the perceived outcome falls short of the expectations, customer dissatisfaction may ensue.

Kotler (2009) posits that customer satisfaction is the emotional response of an individual, characterized by feelings of contentment or disappointment, which arises after comparing their perception or impression

of performance to be below expectations, leading to customer dissatisfaction.

Hansemark and Albinsson (2004), assert that overall customer satisfaction reflects one's attitude towards service providers or their emotional reactions to the disparities between what customers anticipate and what they receive.

According to Bitner and Zeithaml (2003), customer satisfaction entails the evaluation by customers of a product or service in terms of whether it has fulfilled their needs and expectations.

2.6 Cluster Analysis

Clustering is a technique that can be employed to categorize a set of data into several data clusters based on their shared characteristic similarities (Anisa Rahmawati, *et al*., 2023).

Clustering is a part of Data Mining which has no direction. Clustering is a process for dividing data into clusters based on the level of similarity. Clustering is a job that can separate data or vectors into a number of groups or clusters according to their respective characteristics (Egi Affandi1 *et al*., 2021).

Clustering is the process of categorizing similar entities into distinct groups or, more precisely, partitioning a data set into subsets, such that the data within each subset carries meaningful significance. Each cluster comprises a collection of objects that share similarities with each other and differ from objects found in other clusters. Clustering algorithms can be categorized into two main approaches: hierarchical and partitional. Hierarchical algorithms determine clusters sequentially, where clusters are predefined, while partitional algorithms establish all clusters at a given point in time (Madhulatha, 2012).

2.7 K-Means Clustering

As per Suyanto (2017:262), k-means represents a clustering algorithm founded on a straightforward principle, seeking to minimize the Sum of Squared Error (SSE) between data objects and a set of k centroids.

As defined by Anisa Rahmawati and Endah Setyowati (2023), the KMeans algorithm is one of the non-hierarchical data clustering methods designed to partition a set of data points into one or more clusters. This process aims to group data with similar characteristics into the same cluster, while data exhibiting distinct characteristics are grouped into separate clusters.

3. Research method

3.1 Data Collection Technique

The research methodology employed in the application of the KMeans algorithm involves the determination of clustering analysis for the purpose of categorizing customers based on their reasons for choosing Indihome products. The composition of this analytical report employs a data collection method in the form of a questionnHoyer, W. D., MacInnis, D. J., Pieters, R., Chan, E., & Northey, G. (2021). *Behaviour Consumer 2nd Asia-Pacific Edition*.

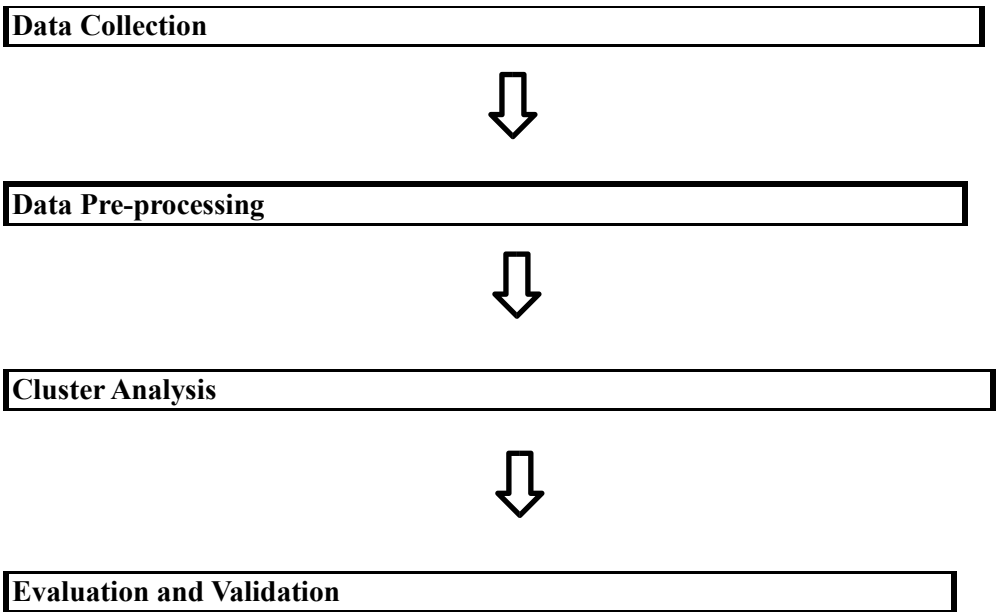
Rumondang, A., Sudirman, A., Sitorus, S., Kusuma, aditya H., Manuhutu, M., Sudarso, A., Simarmata, J., Hasdiana, D., Tasnim, T., & Arif, nina F. (2020). *Pemasaran Digital dan Perilaku Konsumen*.

aire, which was distributed to Indihome customers using the simple random sampling technique with a sample size of 30 respondents. The analysis is carried out using the method of cluster analysis, specifically the K-Means Cluster method, in accordance with the designed research framework.

This research methodology leverages the K-Means Clustering algorithm to conduct an analysis of customer grouping with respect to the

rationales behind Indihome product usage. The process for achieving the clustering results using the K-Means Clustering algorithm is presented in the following flowchart.

Figure 1. *Research Flowchart*



3.1.1 Questionnaire Data

In light of the aforementioned issues, the author proceeded to distribute a questionnaire simultaneously to Indihome users utilizing Google Forms. The questionnaire data employed by the author followed the semantic differential scale methodology, gauging satisfaction levels on a scale ranging from 1 to 5. The measurement scale utilized for this study was the Rating Scale. The resultant outcomes are manifested in the form of preference rankings, complemented by categorical descriptions as follows:

a) 1: Highly Dissatisfied

b) 2: Dissatisfied

c) 3: Neutral

d) 4: Satisfied

e) 5: Highly Satisfied

This scale serves as an instrument for measuring preferences concerning attribute combinations. From the questionnaire, 30 respondents with recorded responses were obtained as follows:

CASE NO.	NAMA	X_1	X_2	X_3	X_4	X_5	X_6	X_7
1	Ita	3,00	3,00	4,00	4,00	5,00	5,00	3,00
2	Laila M	5,00	4,00	4,00	4,00	5,00	5,00	3,00
3	Karina	3,00	3,00	2,00	4,00	5,00	5,00	5,00
4	Firdausi	4,00	5,00	5,00	3,00	4,00	2,00	1,00
5	Miftahul	4,00	4,00	5,00	5,00	1,00	4,00	2,00
6	Agam	5,00	4,00	5,00	5,00	2,00	4,00	5,00
7	Suci W	5,00	4,00	5,00	4,00	4,00	1,00	1,00
8	Iin M	2,00	1,00	4,00	1,00	2,00	1,00	1,00
9	Itsnani	3,00	4,00	3,00	2,00	4,00	5,00	5,00
10	M.Kurniawan	2,00	5,00	5,00	1,00	5,00	4,00	1,00
11	Listina	4,00	4,00	3,00	4,00	3,00	4,00	5,00
12	Nurul	5,00	4,00	1,00	4,00	3,00	5,00	5,00
13	Warsini	2,00	5,00	5,00	5,00	5,00	5,00	5,00
14	Nisa	2,00	3,00	1,00	3,00	3,00	5,00	3,00
15	Heri	5,00	4,00	2,00	5,00	2,00	4,00	1,00
16	Nisa M	3,00	3,00	3,00	3,00	3,00	3,00	3,00
17	Edi	5,00	5,00	2,00	4,00	2,00	4,00	5,00
18	Doni	5,00	5,00	5,00	3,00	4,00	2,00	1,00
19	Fuad	1,00	1,00	4,00	5,00	5,00	4,00	5,00

20	**Sasa**	5,00	2,00	1,00	5,00	4,00	1,00	1,00
21	**Imdadu**	4,00	4,00	4,00	4,00	4,00	4,00	4,00
22	**Alfiansya**	2,00	3,00	3,00	4,00	1,00	4,00	3,00
23	**Amelia**	3,00	3,00	4,00	2,00	4,00	3,00	4,00
24	**Elgie**	4,00	4,00	4,00	4,00	4,00	4,00	5,00
25	**Agustin**	5,00	1,00	3,00	2,00	3,00	4,00	4,00
26	**Arin**	4,00	4,00	3,00	4,00	4,00	3,00	2,00
27	**Maya**	5,00	5,00	4,00	5,00	1,00	4,00	4,00
28	**Hapit**	3,00	3,00	4,00	3,00	4,00	3,00	1,00
29	**M zidmi**	5,00	5,00	4,00	5,00	4,00	3,00	4,00
30	**Qurrotul**	5,00	4,00	4,00	4,00	3,00	4,00	1,00

Description:

X1 = Network speed quality

X2 = Clear image quality

X3 = Affordable product price

X4 = Comprehensive menu variants

X5 = Highly adequate service

X6 = Extensive geographical coverage X7 = Well-known brand

3.1.2 Final Cluster Center

Final Cluster Centers			
	Cluster		
	1	2	3
X1	**4,40**	2,63	3,71
X2	**3,87**	3,25	3,57
X3	3,33	3,38	**4,14**
X4	**4,13**	3,63	2,86
X5	2,73	**4,38**	3,86
X6	3,93	**4,50**	2,00
X7	3,40	**4,38**	1,00

Based on the Final Cluster Centers table, three clusters have been formed, with the highest-value members as follows:

Cluster 1 = X1 (Network speed quality), X2 (Clarity of image quality), and X4 (Comprehensive menu variety).

Cluster 2 = X5 (Highly adequate service), X6 (Extensive geographical coverage), and X7 (Well-known brand). **Cluster 3** = X3 (Affordability of the product price).

3.1.3 ANOVA

ANOVA						
	Clust er		Erro r			
	Mean Square	df	Mean Square	df	F	Sig.
X1	8.232	2	1.071	27	7.689	.002
X2	1.010	2	1.443	27	.700	.505
X3	1.701	2	1.558	27	1.092	.350
X4	3.917	2	1.202	27	3.258	.054
X5	7.851	2	1.099	27	7.145	.003
X6	13.017	2	.627	27	20.755	.000
X7	22.612	2	1.240	27	18.239	.000

The analysis of variance (ANOVA) reveals that the seven variables under consideration impart distinct characteristics to the three formed clusters. Certain attributes, namely Network Speed Quality, Highly Satisfactory Service, Extensive Geographic Coverage, and a Renowned Brand, exhibit a Significance (Sig.) value of <0.05. This suggests that all instruments can be employed to differentiate each cluster effectively. In contrast, some other attributes, such as Image Clarity Quality, Affordable Product Pricing, and Comprehensive Menu Variety, exhibit a Significance (Sig.) value of >0.05. This implies that these attributes or variables do not provide effective discrimination between the clusters.

3.1.4 Number of Cases in each Cluster

Number of Cases in each Cluster		
Cluster	1	15.000
	2	8.000
	3	7.000
Valid		30.000
Missing		.000

Description:

Cluster 1 = 15 respondents

Cluster 2 = 8 respondents

Cluster 3 = 7 respondents

Total respondents = 30 respondents

4. Conclusion

The clustering analysis employed in this research aims to categorize customers who utilize IndiHome services with respect to the reasons for their usage at PT Telkom Indonesia. Based on the research findings and analysis conducted, it can be inferred that three consumer groups are discernible within PT Telkom Indonesia. Within these three consumer groups, a total of seven attributes or variables associated with the reasons for using IndiHome are identified. Among these variables, seven possess a Significance value of <0.05, indicating their suitability as differentiating variables among clusters. Specifically, one respondent is found within cluster 1, while 18 respondents belong to cluster 2, and 11 respondents are allocated to cluster 3.

The results of this clustering can subsequently serve as recommendations for marketing managers at PT Telkom Indonesia to enhance their marketing systems, thus allowing for a more comprehensive understanding of the reasons behind customers' adoption of IndiHome.

References

Hoyer, W. D., MacInnis, D. J., Pieters, R., Chan, E., & Northey, G. (2021). *Behaviour Consumer 2nd Asia-Pacific Edition*.

Rumondang, A., Sudirman, A., Sitorus, S., Kusuma, aditya H., Manuhutu, M., Sudarso, A., Simarmata, J., Hasdiana, D., Tasnim, T., & Arif, nina F. (2020). *Pemasaran Digital dan Perilaku Konsumen*.

Rahmawati, A. Setyowati, E. (2023). Analisis Cluster K-Means untuk Pengelompokan Kabupaten/Kota di Provinsi Bengkulu berdasarkan Jumlah Base Transceiver Station dan Kekuatan Sinyal Telepon Seluler. Journal of Computing Engineering, System and Science. Vol. 8, No. 1, Hal 184-196.

Egi, A. Syahrin, S. Yohanni, S. Amrullah. (2021) Data Mining in Grouping Indihome Customer Data Using the K–Means Clustering Method at PT.Telkom Akses. Journal mantik. Vol. 4, No. 4, Hal 2604-26124

Fatihah, R. Azahro1, S. Handojo, D. (2022). Pengaruh Harga dan Kualitas Produk Terhadap Kepuasan Pelanggan Produk Indihome PT Telkom Kabupaten Pati. Jurnal Ilmu Administrasi Bisnis. Vol. 11, No. 4, Hal 2746-1297.

Aliya, A. Endang, M. Bida, S. (2020). Pengaruh Bauran Promosi, Harga, dan Kualitas Layanan Terhadap Kepuasan Pelanggan Indihome PT. Telkom Indonesia, Jakarta Timur. Jurnal IKRA-ITH Ekonomika. Vol.
3, No. 1, Hal. 2654 – 4946.

(BOOK)

Hoyer, Macinnis, Pieters, Chan and Northey. 2021.
Consumer Behaviour. 2nd Asia–Pacific Edition. Max McMaster. Australia.