International Journal of Application on Sciences, Technology and Engineering (IJASTE)
Volume 1, Issue 2, 2023.ISSN:2987-2499

E-COMMERCE WEBSITE APPLICATION WITH CUSTOMER LOYALTY AND RECOMMENDATIONS ITEMS DEPENDS ON PRICE FEATURES USING K- MEANS CLUSTERING METHOD

Marcelino Fernandez[1)], Bagus Mulyawan[2)], Manatap Dolok Lauro[3)]

Computer Science Department, Faculty of Information Technology Tarumanagara University, Jln. Letjen S. Parman No. 1, Jakarta, 11440, Indonesia

[a)] marcelino.535180111@stu.untar.ac.id, [b)]bagusm@fti.untar.ac.id, [c)]manataps@fti.untar.ac.id
b)

Submitted: January-February 2023, Revised: March 25 2023, Accepted: May 23, 2023

Abstract.

In the current era of computer technology development, the internet is a media of information that can be accessed across countries. To overcome problems in the commercial system, E-Commerce was created. E-Commerce is a commercial activity using internet media. In the midst of the pandemic Covid-19, many Indonesians have opened E- Commerce to fullfill their daily needs because they have been affected by this Covid-19 pandemic. This program is one of the forms of E-Commerce that was created during the Covid-19 pandemic, starting from December 2020 which was named Maniak Mainan at Indonesian E-Commerce, Tokopedia. Currently the products listed on Tokopedia Maniak Mainan is around 6100 products, therefore a recommendation program is made to provide users with suitable choices according to the price chosen through the bottom of the product page. This program also has a customer loyalty feature that allows users to get discounted prices by adding the nominal amount of purchases on the Maniak Mainan website. The recommendations and customer loyalty made using the K-Means Clustering method obtained using original sales and product data from Tokopedia Maniak Mainan. However, this cluster center will adjusted when the admin uploads a product with a higher or lower price, and the customer loyalty tier will also change if the user's purchase nominal exceeds the previous highest nominal. Through the results of the test, it can be concluded that 20 cluster centers are the right distance for this because it includes the price of products that are nearby. The compatibility test was carried out with the Silhouette Score with the results of 0.579 on the recommendation and 0.689 on the customer tier

Keyword: E-Commerce, K-Means Clustering, Recommendations, Customer Loyalty, Maniak Mainan.

1. INTRODUCTION

The growing use of the Internet, tablet devices, and smart phones coupled with larger consumer confidence will see that ecommerce will continue to evolve and expand. With social media growing exponentially in recent years, the conversation between businesses and consumers has become more engaging, making it easier for transactional exchanges to happen online [1]. E-Commerce is the buying, selling and marketing of products and services marketed through electronic systems, such as television, websites, radio and computer networks. [2] In the online commerce sector, the existence of internet technology allows business transactions not only to be carried out directly, but also to use this technology. Internet media itself has begun to be widely used as a media for business

https://doi.org/10.24912/ijaste.v1.i2.457-472

activities, mainly because of its contribution to efficiency. Many companies have moved from physical stores to e-commerce only for various reasons. One of them is the convenience provided by technology so as to make e- commerce an efficient and effective sales alternative. Another important thing that makes companies switch to e- commerce is that internet users increase from time to time, and sellers can create new jobs for themselves and others. In this modern era, shopping online is one of the daily activities of the Indonesian people, this is done to meet the needs of users or their hobbies. By utilizing current technology, online shopping activities can be done anywhere and anytime, such as shopping through an online shop that can be accessed via smartphones and computers. Which is a very wide open business opportunity for sellers in Indonesia. In the midst of this Covid- 19 pandemic virus or to be precise in December 2020 the author opened an online store in one of the well-known marketplaces in Indonesia, namely Tokopedia. The creation of this website so that the total purchases made by buyers will received 100% (percent), the existence of rules that are burdensome to the seller is also one of the factors where there are violations of several products that make the product must be forcibly removed by Tokopedia. And also there is a delivery limited time determined by Tokopedia. The unsupervised machine learning algorithm typically identifies insight structures of the data from unlabelled data contained in the dataset. The clustering algorithm finds and divides the data points according to the similarity of the hidden structures of the dataset [3]. Recommender Systems (RS) appear to suggest items individually for each user based on his historical or preferences. The purpose is to inform the user about the items that should be of interest in many different choices [4]. The results of this program will produce a marketing strategy by giving discounts to each cluster group called the Tier system starting from Tier 1, 2, 3, and 4. This cluster is made based on nominal purchases because a turnover of 100 million per month can be achieved within 80 buyers. The nominal purchase is the main factor in the center of this customer tier cluster.

https://doi.org/10.24912/ijaste.v1.i2.457-472

2. METHOD AND MATERIALS

2.1 Data Used

Maniak Mainan is active in Tokopedia, all data for this program is originally taken from Tokopedia. Data will be used for creating and testing recommendation system and loyalty customer systems.

2.2 Method

K-means is one of the relaxed unsupervised erudition algorithms that illuminate the well-known clustering issue. The k-means clustering algorithm is one of the widely used data clustering methods where the datasets having "n" data points are partitioned into "k" groups or clusters. The k-means grouping algorithm was initially proposed by MacQueen in 1967 [5]. K-Means Clustering is a non-hierarchical data cluster method that groups data in the form of one or more clusters with similar data. implemented a document clustering technique using singular vector decomposition to find out the number of clusters required (i.e., the value of k). The k-means algorithm is used to create clusters [6]. Data that has the same characteristics will be collected in one cluster, while data that has different characteristics will be collected into another cluster. The characteristics of the largest data in one cluster have a small degree of variation The stages in performing clustering with the K-Means method are as follows: [7]

1. Select the number of clusters k.
2. The value of k cluster center can be done in various ways and one of them is random. The cluster center is assigned an initial value with a random number.
3. Allocate all data/objects to the nearest cluster. The proximity of two objects is determined based on the distance between the two objects. Likewise, the proximity of a data to a particular cluster is determined by the distance between the data and the center of the cluster. In this stage, it is necessary to calculate the distance of each data to each cluster center. The distance between one data and a certain cluster will determine which data will fall into which cluster. The formula for calculating the data distance to each cluster center point can use the Euclidian Distance, which is written in:

$$d_{ik=\sqrt{\sum^{m}(X_{ij}-C_{kj})2}}$$

https://doi.org/10.24912/ijaste.v1.i2.457-472

m=1

Where:
1. dik is the distance between data to i with the centroid of each cluster
2. m is the number of attributes
3. xij is the original price of the item
4. ckj is the center of the cluster

A recommender system is used to achieve useful information of the products, information and services to the user by combining preference from the other users, from different authorities and from user attributes [8]. Recommender system is defined by accommodating and supplementing for prediction of social procedure for others to make some relevant preferences when there is lack of information. Recommender system is also used as an information filtering process in the situation where there is lots of information exists. [9]

The following data is an example of a cluster calculation with K-Means Clustering. The data obtained is the original data in accordance with the products in the Maniak Mainan online store. The following are 50 product data consisting of product names and prices which can be seen in the table

Table 1 Products Data

No	Product Name	Price
1	Figure Vocaloid Hatsune Miku - Miku With You 2019 Ver.	5300000
2	Figure Vocaloid Hatsune Miku - Miku Yukata Harvest Moon Ver.	900000
3	Figure Vocaloid Hatsune Miku - Nendoroid 303 Ichigo Shiromuku Ver.	550000
4	Figure Vocaloid Hatsune Miku - Nendoroid 493 Snow Miku: Snow Bell Ver.	400000
5	Figure Vocaloid Hatsune Miku - Nendoroid 539 Miku Harvest Moon Ver.	380000
6	Figure Vocaloid Hatsune Miku - Nendoroid 831 Miku 10th Anniversary Ver	250000
7	Figure Vocaloid Hatsune Miku - Nendoroid 93 Megurine Luka	350000
8	Figure Vocaloid Hatsune Miku - Taito Miku Shifuku Casual Roomwear Ver.	380000
9	Figure Walkure Romance - Celia Cumani Aintree Apron Ver.	580000
10	Figure Walkure Romanze - Celia Cumani Aintree Swimsuit Suntanned Ver.	2250000

https://doi.org/10.24912/ijaste.v1.i2.457-472

11	Figure Warcraft Dota - Pandaren Brewmaster Battle Ultimate Ver.	860000
12	Figure Warship Girls - Saratoga Flower of Isha Ver. - Merah	1100000
13	Figure Warship Girls - Saratoga Flower of Isha Ver. - Hitam	2530000
14	Figure Warship Girls - Saratoga Flower of Isha Ver. - Putih	1430000
15	Figure Warship Girls - Saratoga Flower of Isha Ver. (Bonus) - Merah	1100000
16	Figure Warship Girls - Saratoga Flower of Isha Ver. (Bonus) - Hitam	2650000
17	Figure Warship Girls - Saratoga Flower of Isha Ver. (Bonus) - Putih	1550000
18	Figure Warship Girls R - Lexington Swimsuit Ver.	650000
19	Figure Warship Girls Senkan Shoujo - Light Cruiser Yi Xian Battle Ver.	850000
20	Figure Wataten! - Nendoroid 1195 Hinata Hoshino (Bonus)	1150000
21	Figure Wataten! - Nendoroid 1730 Hana Shirosaki - Merah	200000
22	Figure Wataten! - Nendoroid 1730 Hana Shirosaki - Hitam	580000
23	Figure Wataten! - Nendoroid 1730 Hana Shirosaki - Putih	380000
24	Figure Wataten! - Nendoroid 1731 Noa Himesaki - Merah	200000
25	Figure Wataten! - Nendoroid 1731 Noa Himesaki - Hitam	580000
26	Figure Wataten! - Nendoroid 1731 Noa Himesaki - Putih	380000

https://doi.org/10.24912/ijaste.v1.i2.457-472

27	Figure White Album - Kotobukiya Morikawa Yuki 4-Leaves Legend Girls	1200000
28	Figure White Album 2 - Myethos Kazusa Touma Wedding Ver.	2450000
29	Figure White Album 2 - Myethos Setsuna Ogiso Wedding Ver.	2450000
30	Figure White Album 2 Shiawase no Mukougawa - Ogiso Setsuna	2650000
31	Figure White Album 2 Shiawase no Mukougawa - Touma Kazusa	2650000
32	Figure Wisteria - Witch of the Night Hag Lilith Reaper Battle Ver.	880000
33	Figure Wixoss - Kokuten no Miko Tamayorihime Battle Ver.	720000
34	Figure Wonder Woman - Nendoroid 818 Wonder Woman: Hero's Edition	360000
35	Figure World of Warcraft - Demon Hunter Illidan Stormrage Battle Ver.	1350000
36	Figure World of Warcraft - Durotan Frostwolf Battle Ver.	2700000
37	Figure World of Warcraft - Grommash Hellscream Warlord of Draenor Ver.	1600000
38	Figure World of Warcraft - Lich King Arthas Menethil Battle Ver.	420000
39	Figure World of Warcraft Dota - Sylvanas Traxex Markmanship Battle Ver	680000
40	Figure Wreck-It Ralph - Nendoroid 1492-DX Vanellope DX Ver.	1150000
41	Figure X-Men - Cable 1/10 Marvel Event Exclusive by Iron Studios	3500000
42	Figure Xenoblade Chronicles - Kos Mos Sexy Swimwear Alter Ver.	2200000
43	Figure Xenoblade Chronicles 2 - Hikari Mythra (Good Smile Company)	4600000
44	Figure Xenoblade Chronicles 2 - Homura / Pyra (Good Smile Company)	4800000
45	Figure Xenoblade Chronicles 2 - KOS-MOS Kosmos	4400000
46	Figure Xenoblade Chronicles 2 - Mythra Hikari Aegis Sword Battle Ver.	780000
47	Figure Xenoblade Chronicles 2 - Nia - Merah	1500000
48	Figure Xenoblade Chronicles 2 - Nia - Hitam	3380000
49	Figure Xenoblade Chronicles 2 - Nia - Putih	1880000
50	Figure Xenoblade Chronicles 2 - Pyra Homura Aegis Sword Battle Ver.	780000

Calculation is done by dividing the center of the cluster into 7. Then search for the maximum and minimum. Then the maximum and minimum are subtracted to find the average according to the 7 number of clusters. In the results of clustering this section, the results are obtained, center of the cluster as a recommendation is fairly small in nominal, but that small number only applies if it is calculated with 50 training data where the highest price of the data is 5.300.000. the lowest is 200,000. The following are the results of the clustering of the 50 training data.

https://doi.org/10.24912/ijaste.v1.i2.457-472

Table 2 Price Cluster

CLUSTER	AVG	AVG RESULT
1	0.125	637,500.00
2	0.25	1,275,000.00
3	0.375	1,912,500.00
4	0.5	2,550,000.00
5	0.625	3,187,500.00
6	0.75	3,825,000.00
7	0.875	4,462,500.00

For the calculation of the customer loyalty feature using data taken from Tokopedia Maniak Mainan in February 2022 which has been combined based on the buyer's name and nominal. Customer name is hidden due to privacy policy and will be codenamed as customer 01-73. The following data is order created in February 2022 :

https://doi.org/10.24912/ijaste.v1.i2.457-472

Table 3 February 2022 Customers Data

Customer Name	Total Purchase
Customer 01	1800000
Customer 02	360000
Customer 03	1500000
Customer 04	750000
Customer 05	350000
Customer 06	750000
Customer 07	600000
Customer 08	1100000
Customer 09	350000
Customer 10	350000
Customer 11	720000
Customer 12	7330000
Customer 13	400000
Customer 14	2570000
Customer 15	400000
Customer 16	360000
Customer 17	620000
Customer 18	2350000
Customer 19	600000

https://doi.org/10.24912/ijaste.v1.i2.457-472

Customer 20	3000000
Customer 21	670000
Customer 22	1600000
Customer 23	360000
Customer 24	2680000
Customer 25	380000
Customer 26	440000
Customer 27	1220000
Customer 28	320000
Customer 29	280000
Customer 30	1400000
Customer 31	3000000
Customer 32	360000
Customer 33	1000000
Customer 34	280000
Customer 35	320000
Customer 36	215000
Customer 37	200000
Customer 38	300000
Customer 39	360000
Customer 40	1540000
Customer 41	350000
Customer 42	390000
Customer 43	350000
Customer 44	2000000
Customer 45	100000
Customer 46	200000
Customer 47	3000000
Customer 48	1800000
Customer 49	550000

https://doi.org/10.24912/ijaste.v1.i2.457-472

Customer 50	1650000
Customer 51	2100000
Customer 52	400000
Customer 53	300000
Customer 54	100000
Customer 55	480000
Customer 56	780000
Customer 57	1910000
Customer 58	360000
Customer 59	7900000
Customer 60	700000
Customer 61	2800000
Customer 62	500000
Customer 63	840000
Customer 64	900000
Customer 65	200000
Customer 66	200000
Customer 67	200000
Customer 68	440000
Customer 69	350000
Customer 70	3650000
Customer 71	5600000
Customer 72	380000
Customer 73	780000

https://doi.org/10.24912/ijaste.v2.i2.457-472

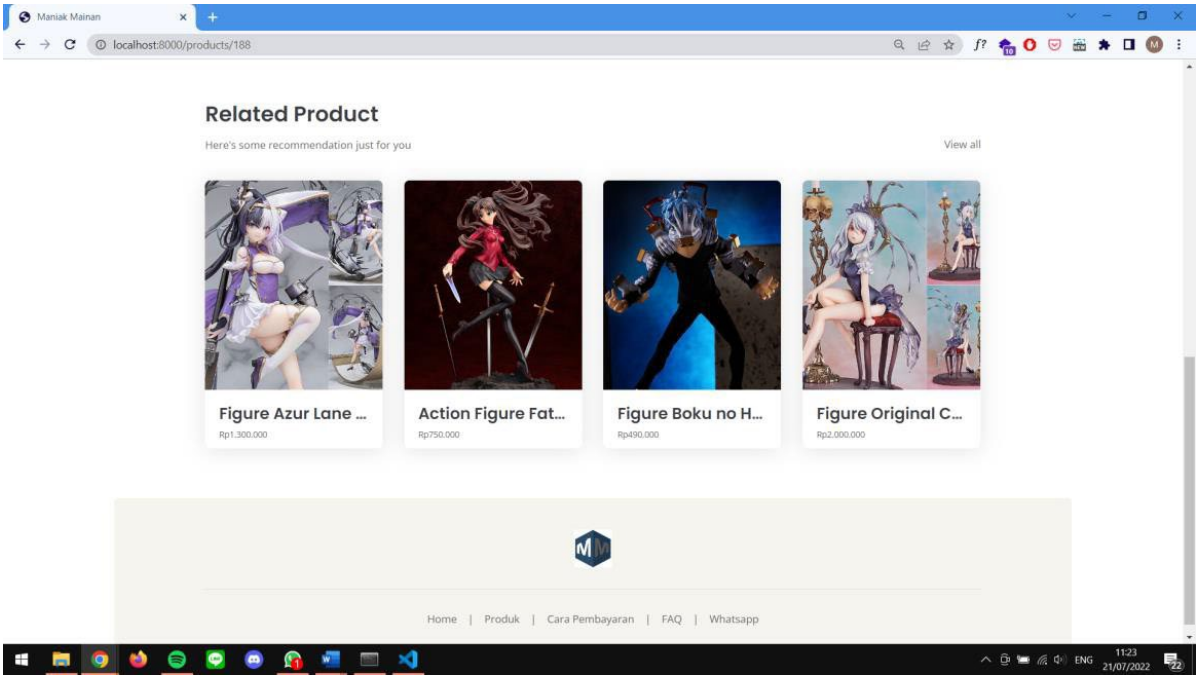
Determination of cluster centers based on total purchases is divided into 4, namely: Tier 1,2,3 and 4 where tier 1 does not get a loyalty discount. Starting from tier 2 getting 2%, tier 3 getting 3% and tier 4 getting 5% loyalty discount, the cluster is centered on the highest and lowest total transactions where the system will look for the average of the total transactions and then divided into clusters in the data on the number of purchases in February it can be concluded that the center of the cluster is as follows:

Table 4 Loyalty Cluster

C1	1.560.000
C2	3.120.000
C3	4.680.000
C4	6.240.000

RESULT AND DISCUSSION

The results of the calculation test are made by system by using average price of the items, recommendation will be different everytime admin input new highest or lowest product. This cluster test will be carried out several times on the tier or product recommendation level to get the desired results.



https://doi.org/10.24912/ijaste.v2.i2.457-472

Figure 1 Cluster Division 4

Cluster division 4 is showing recommendation range with Rp. 850.000 Product. But recommendation range are from Rp. 490.000 up to Rp.2.000.000. This means cluster division 4 is not suitable. Because clusters are too far.

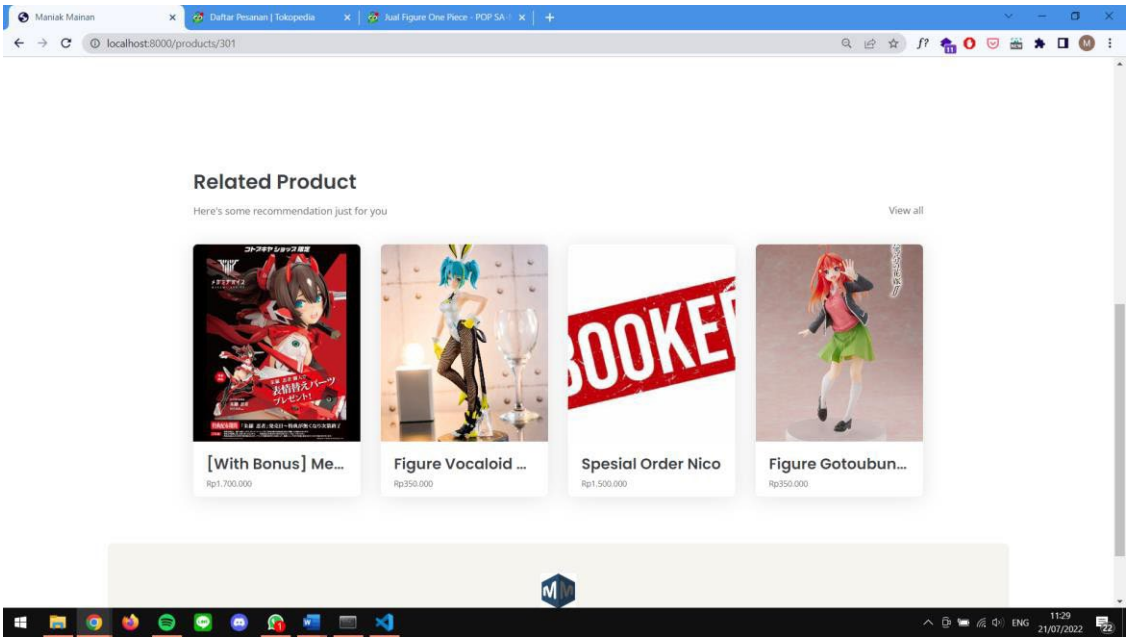


Figure 2 Cluster Division 8

In this cluster 8 division, the recommendations are starting to getting approached where the product clicked on the display above is a product worth Rp. 850,000, getting a recommendation of Rp. 350,000 to Rp. 1,700,000.

https://doi.org/10.24912/ijaste.v2.i2.457-472

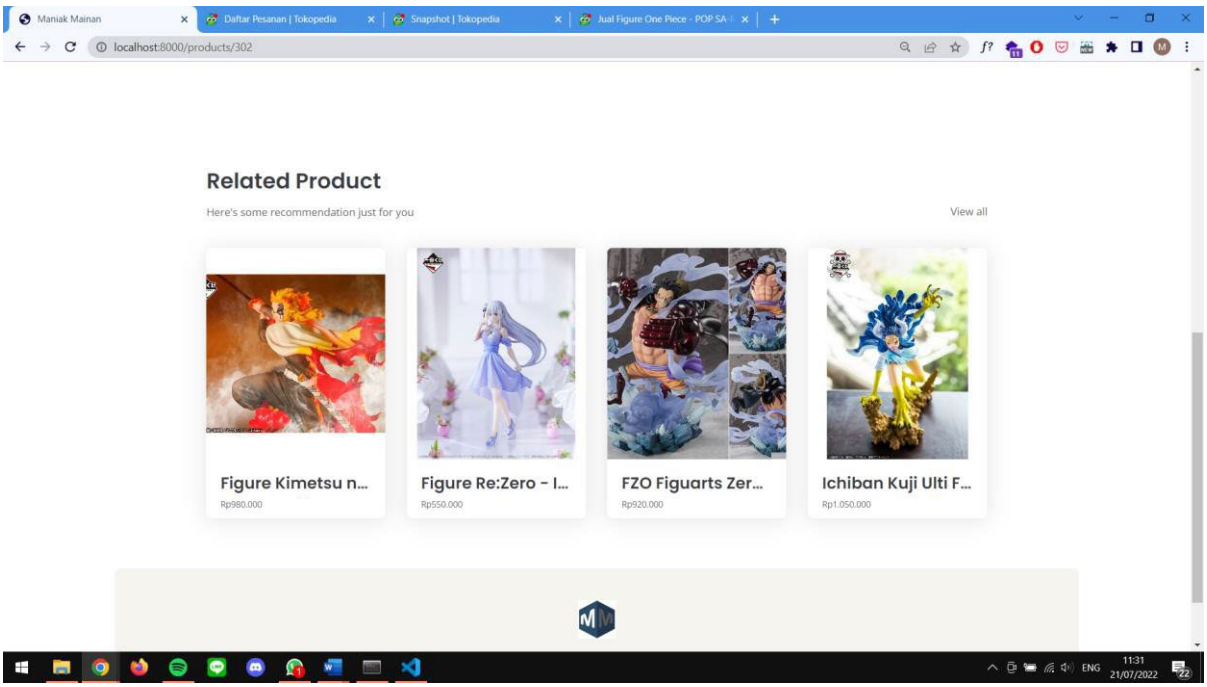


Figure 3 Cluster Division 12

In the cluster division 12, the products used are also worth Rp. 850,000 where the recommendations are narrowed so that the recommendations obtained start from Rp. 550,000 to Rp. 1,050,000

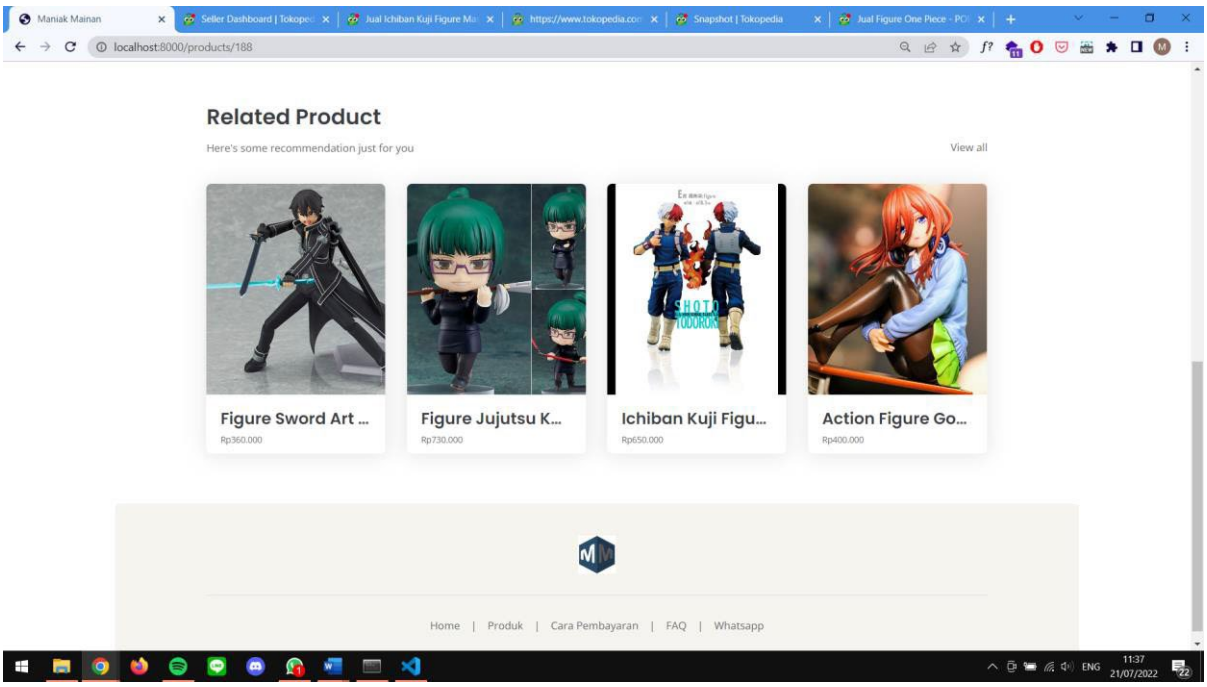


Figure 4 Cluster Division 16

https://doi.org/10.24912/ijaste.v2.i2.457-472

The picture above shows the cluster 16 division using a product worth Rp. 850.000 which is getting more realistic recommendations with prices, buyers get the lowest recommendation at a price of Rp. 360.000 and the highest is Rp. 730.000

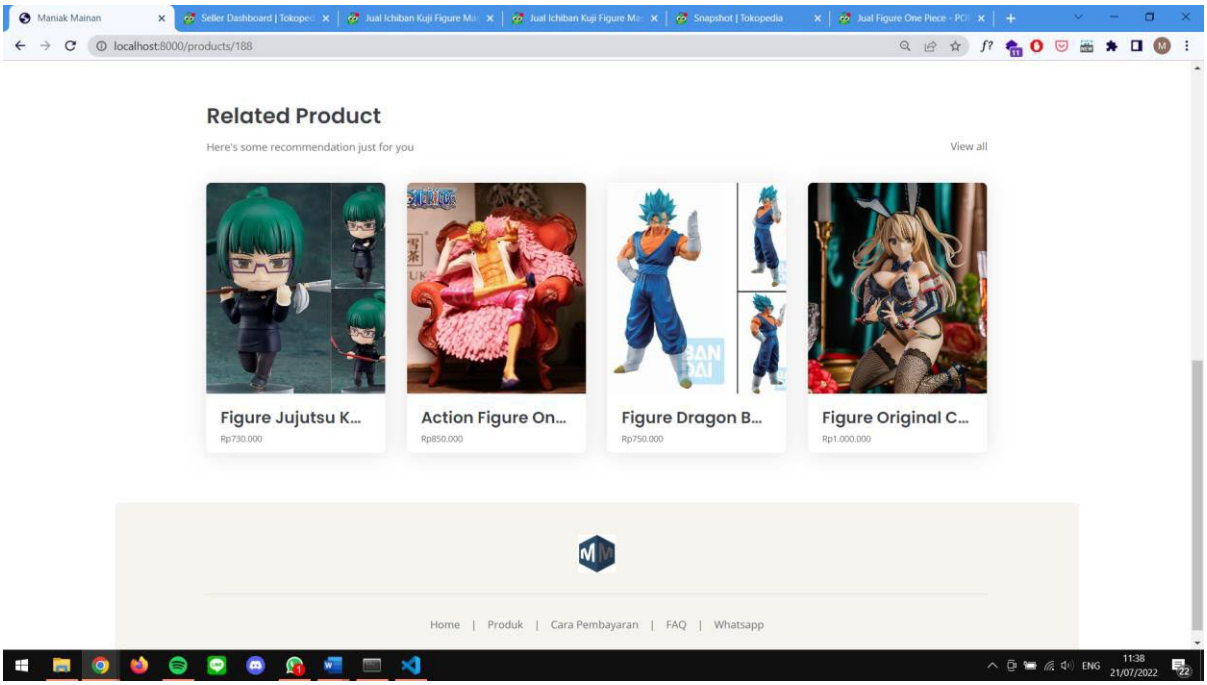


Figure 5 Cluster Division 20

The picture above shows the cluster 20 division with the same product as other clusters, but gives a balanced result with the clicked product where the recommendation is worth Rp.730,000 to Rp.1.000,000. Based on these results, this cluster division is used as a recommendation.

https://doi.org/10.24912/ijaste.v2.i2.457-472

CONCLUSION

The conclusion that can be obtained from the recommendation and customer loyalty program using K-Means Clustering is that users can get product recommendations according to the price of the product that was last clicked, and also users can get discounts through the customer loyalty feature by increasing the total purchases in the form of nominal purchases. which will be calculated by the system to get a discount given to a certain tier. And also from the results of the test data, manual calculations with the program made successful or suitable. Which means that K-Means Clustering can be used as a customer recommendation and loyalty system. User Acceptance Testing (UAT) was created using Google Forms, the question is all about this program. The forms are spreaded by chat and got 20 responses with 95% for usefulness recommendation and 95% for loyalty customer function. This program is also tested using Silhouette Score to know range from cluster to cluster, which the scores for product recommendation is 0.579 and scores for loyalty customer is 0.689.

REFERENCES

[1] Khan, Abdul Gaffar, "Electronic Commerce: A Study on Benefits and Challenges in an Emerging Economy", Global Journal Inc. (USA), Vol. 16, 2016

[2] Wong, Jony, Internet Marketing For Beginners, Elex Media Komputindo, Jakarta

[3] Han, J., Pei, J., Kamber, M, 3rd Edition, "Data Mining Concepts and Techniques, Elsevier

[4] Meira, D., Viterbo, J., and Bernardini, F., "An Experimental Analysis on Scalable Implementations of the Alternating Least Squares Algorithm", Federated Conference on Computer Science and Information Systems (FedCSIS), Vol. 15, 2018

[5] MacQueen, J.B., "Some methods for classification and analysis of multivariate observations", Vol. 5.1, 1967 [6] Gupta, H., "k-means Based Document Clustering with Automatic "k" Selection and Cluster Refinement",
International Journal of Computer Science and Mobile Applications, Vol. 2, 2014

[7] Santoso, Yessica Putri; Marlina, dan Agung Halim, "Implementasi Metode K-Means Clustering pada Sistem Rekomendasi Dosen Tetap Berdasarkan Penilaian Dosen", Jurnal Informatika Universitas Pamulang, Vol. 3, No. 4, 2018

[8] Trivedi, A., Rai, P., DuVall, S. L., "Exploiting tag and word correlations for improved webpage clustering.",

https://doi.org/10.24912/ijaste.v2.i2.457-472

Proceedings of the 2nd international workshop on Search and mining user-generated contents, 2010

[9] Schafer JB, Konstan J, Riedl J., "Recommender system in e-commerce", Proceedings of the 1st ACM conference on electronic commerce, 1999

https://doi.org/10.24912/ijaste.v2.i2.457-472