International Journal of Society Reviews (INJOSER) e-ISSN: 3030-802X
Vol. 1 No. 8 2024, pages: 912~922

GROUPING PRODUCTS IN SUPERMARKETS USING THE K-MEANS ALGORITHM

Mohammad Aldinugroho Abdullah
Universitas Nasional, Jakarta, Indonesia
Email : nugrohoaldi48@gmail.com

Rima Tamara Aldisa*
Universitas Nasional, Jakarta, Indonesia
Correspondence author Email: rimatamaraa@gmail.com

Abstract

Supermarket, a shop that provides various products for use, especially for daily life, including food products, drinks, kitchen utensils, clothing, electronic equipment and others. It is not surprising that many mothers now choose to shop for daily necessities at supermarkets rather than the nearest stall. With self-service, it can make it easier for us consumers to buy different products in one place. So there is no need to change shops to buy other items. Of course, products have different levels of popularity, not only because of taste but also because of price. The number of products provided by supermarkets is relatively large and if you look at the level of popularity, it is difficult to determine, so data mining is needed. The data mining used is clustering. After implementing and using the K-Means algorithm in clustering (grouping) supermarket products, there are two centroids used (C1 for Not Selling Products and C2 for Best Selling Products). The initial centroid value is determined randomly, while the subsequent centroids are adjusted according to the results of calculating the closest distance (maximum distance). The final result obtained is that the best-selling group consists of 12 products, namely products with serial numbers 1, 4, 5, 6, 7, 8, 9, 11, 14, 15, 16 and 17. Meanwhile, the product group does not There are 6 best-selling products, namely products with serial numbers 2, 3, 10, 12, 13 and 18.

Keywords: Data mining, Clustering, K-Means, Self-Service Products

INTRODUCTION

Supermarkets are shops that provide various products for use, especially for daily life, including food products, drinks, kitchen utensils, clothing, electronic equipment and others. It is not surprising that many mothers now choose to shop for daily necessities at supermarkets rather than the nearest stall. With self-service, it can make it easier for us consumers to buy different products in one place. So there is no need to change shops to buy other items. Of course, products have different levels of popularity, not only because of taste but also because of price. The number of products provided by supermarkets is relatively large and if you look at the level of popularity, it is difficult to determine, so data mining is needed.

Data mining is a process of artificial intelligence, machine learning and statistics to analyze and identify large amounts of information (data) [1]. The data mining groupings are association, description, estimation, prediction, classification and clustering. This research focuses on the clustering method because it is in accordance with the aim of this research, namely grouping.

Clustering is a data mining method, used to group data, grouping is done based on the similarity of the data [2][3][4]. The k-Means algorithm was used for this research.

The initial step is to determine how many clusters are formed, then determine the initial