Sorghum Entrepreneur Marketplace Using K-Medoid Clustering

Nia Saurina[a,*], Endang Noerhartati[a], Marina Revitriani[a]
[a,b,c]Universitas Wijaya Kusuma Surabaya, Dukuh Kupang Street, Surabaya, Indonesia
[*]correspondence email : niasaurina@gmail.com

Abstract— **Marketplace is a virtual market where the market becomes a meeting place for buyers and sellers to make transactions. Wijaya Kusuma University Surabaya has established an Entrepreneurship Sorghum Unit (UES) starting in 2009 and is collaborating with Sorghum Entrepreneurs in several areas in East Java. The method used is K-Medoid Clustering which aims to (1) create a marketplace for Sorghum Entrepreneurs who have joined UES, (2) Classify using the K-Medoid method to classify Sorghum Entrepreneurs based on the income that has been earned by UES, (3) Conduct mapping based on clustering results so that UES can provide assistance to sorghum entrepreneurs. From the results of the questionnaire, obtained data on the level of desire for UES for each variable from 35 sorghum entrepreneurs as respondents. Quantitative data obtained is then entered and then proceed with the clustering stage. The data is divided into several clusters based on the type of sorghum, products sold and sales results. After that, the cluster performance calculation is carried out by calculating the k-fold value. analysis of the results of testing the effect of the k-fold value is the best percentage found in the k-fold value = 6 with an accuracy value of 34.597%.**

Index Terms— **entrepreneur sorghum; clustering; k-medoid; marketplace.**

Abstrak–-**Marketplace merupakan sebuah pasar virtual dimana pasar tersebut menjadi tempat bertemunya pembeli dan penjual untuk melakukan transaksi. Universitas Wijaya Kusuma Surabaya telah mendirikan Unit Entrepreneurship Sorghum (UES) mulai tahun 2009 dan berkolaborasi dengan Entrepeneur Sorgum di beberapa daerah di Jawa Timur. Metode yang digunakan adalah Klastering K-Medoid yang bertujuan untuk (1) pembuatan marketplace untuk Entrepeneur Sorgum yang telah bergabung di UES, (2) Melakukan klasifikasi menggunakan metode K-Medoid untuk mengelompokkan Entrepeneur Sorgum berdasarkan pendapatan yang telah didapatkan oleh UES, (3) Melakukan pemetaaan berdasarkan hasil klastering agar UES dapat memberikan bantuan kepada entrepreneur sorgum. Dari hasil kuesioner, didapatkan data tingkat keinginan UES untuk tiap-tiap variabel dari 35 entrepreneur sorgum sebagai responden. Data kuantitatif yang diperoleh kemudian dimasukan dan selanjutnya dilanjutkan dengan tahap pengklasteran. Data dibagi menjadi beberapa klaster yaitu berdasarkan jenis sorgum, produk yang dijual dan hasil penjualan. Setelah itu dilakukan perhitungan performansi klaster dengan menghitung nilai k-fold. analisis hasil pengujian pengaruh nilai k-fold adalah persentase terbaik terdapat pada nilai k-fold=6 dengan nilai akurasi 34,597%.**

Kata Kunci— **entrepreneur sorgum; klastering; k-medoid; marketplace.**

I. INTRODUCTION

Marketplace is an electronic product marketing online that brings many sellers and buyers to transact with each other [1]. Marketplace is a virtual market where the market is a meeting place for buyers and sellers to make transactions. Marketplaces have the same function as a traditional market, the difference is that marketplaces are more computerized by using the help of a network to support a market so that it can be carried out efficiently in providing updated information and services for different sellers and buyers [2].

The problem that happens is that there are still many entrepreneurs who are reluctant to enter digital platforms, because they find it difficult to access and partner with marketplaces [3]. Only 13% of culinary entrepreneurs have gone online, the remaining 87% have not yet gone online [4]. Entrepreneurs are still faced with several challenges, including: (1) Digital knowledge, where sellers must understand new online features to be able to use them properly, (2) From an operational standpoint, namely suitable facilities for the production and storage of goods in a larger size in order to maintain product quality, and (3) Prepare competent human resources [5]. Digitalization is one of the solutions to the problems faced by Entrepreneurs today. Therefore, strong collaboration between the government and digital platforms is needed to revive Entrepreneurs in Indonesia. The government must prepare a national economic recovery scheme, including for Entrepreneurs [6].

UES has several activities such as seminars, talk shows, training, workshops, Entrepreneurship Expo. UES also helps Entrepreneurs to grow and develop in support of training and community service activities. The number of Entrepreneurs at UES is increasing day by day as well as the types of sorghum processed products which are increasingly diverse. If they don't have electronic version, the following styles and sizes must be used:

Clustering is a tool in data mining which has the goal of grouping objects into clusters. A cluster is a collection of data objects that are similar to one another in the same cluster and are dissimilar to objects that are in different clusters. Objects can be grouped into clusters so that objects in one cluster have high similarity to one another [7]. Clustering is the process of grouping records, observations, or those that have the same object. Clustering and classification have differences, namely the absence of a target variable in the clustering process [8].

In this research, sorghum entrepreneurs were collected into 5 groups, namely (1) sorghum entrepreneurs who had the smallest business capital, (2) sorghum entrepreneurs who had a large number of regular customers, (3) sorghum entrepreneurs who had few employees, (4) sorghum entrepreneurs who carry out promotions on social media, (5) Entrepreneur sorghum that relieves the least. With the help of K-Medois clustering, sorghum entrepreneurs can automatically grouped into 5 groups. This is done so that UES can provide assistance in the form of business development capital in accordance with the achievements which has been done by entrepreneur sorghum.

II. METHOD

Universitas Wijaya Kusuma Surabaya has established an Unit Entrepreneurship Sorghum (UES) starting in 2009 and is collaborating with Sorghum Entrepreneurs in several areas in East Java. Table 1 describes the number of UES from 2020 – 2022.

Table 1 Amount of UES From 2020 – 2022

No	Processed Sorghum	2020	2021	2022
1	Red Sorghum Seed	1	3	4
2	Sorghum Flour	16	18	19
3	White Sorghum Seed	-	2	6
4	Sorghum Souvenir	2	4	4
5	Sorghum Noodles	1	1	1
6	Sorghum Soy Sauce	-	1	1

This study used data from Sorghum Entrepreneurs from UES 2022. The sorghum Entrepreneurs who have joined are from the cities of Sumenep, Kediri, Malang, Sidoarjo, Nganjuk, Tulungagung, Blitar, Lamongan, Tuban, Jombang. The data collected was 35 sorghum entrepreneurs.

Figure 1 explains the research method where the first step is to enter 35 sorghum entrepreneurs according to table 1 in 2022. Then enter representative data in the form of entrepreneur id, store name, store address, store description, store location, sorghum product images, business capital owned entrepreneur, the number of employees owned, the type of social media used for promotion along with the name of the social media account used. Then enter 5 groups as a limit for the number of clusters, so that the system can perform K-Medoids clustering analysis and display the results of K-Medoid clustering. After that the system can provide advice to UES sorghum entrepreneurs who meet the criteria of 5 groups automatically so that UES can provide capital assistance for business development, the results of the K-Medoids clustering are implemented in the creation of a sorghum entrepreneur marketplace.
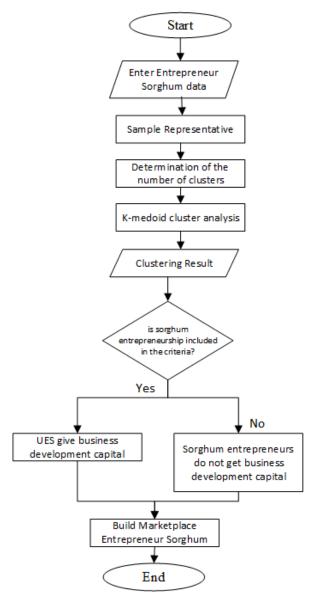
Fig. 1. Research Method

K-Medoids Clustering is a variation of the K-Means method. The difference between this method compared to k-means is that the use of medoids is not based on the average of each cluster object [9]. K-Medoids were used to reduce the sensitivity of the partitions to outlier values in the data set. The purpose of K-Medoids Clustering is to overcome the drawbacks of K-Means Clustering from its sensitivity to outlier data which can affect data distribution [10]. K-Medoids Clustering is a partitioning method of clustering that combines n objects into k clusters. Objects in a group of objects that represent a cluster are called medoids [11]. This research used a clustering algorithm because the number of datasets in this study amounted to 35 sorghum entrepreneurs whose aim was to group business assistance recipients by UES to sorghum entrepreneurs into 5 groups.

The stages use the K-Medoids algorithm as follows [12]:
1. Determine the number of clusters as much as k with theoretical and conceptual considerations.

2. Finding the distance of each object to the nearest cluster using the Euclidean distance measure.

$$d(x,y) = |x - y| = \sqrt{\sum_{i=1}^{n}(xi - yi)^2} \quad (1)$$

with:
p : The number of variables
dab : The distance between object a and object b
xak : The value of object a from variable k
xbk : The value of object b from variable k

3. Then randomly determine the new cluster centre to be used as a non-medoids candidate for each object.

4. Calculate the distance between each object in each cluster to the non-medoids candidates.

5. Calculating between the number of deviations (S) with the number of new distances – the number of old distances. If S < 0, then replace the object with non-medoids cluster data to form a new collection of k objects called medoids.

6. Next, repeat steps 3 − 5 until the medoids don't change. Then obtained the cluster and each member of the cluster.

III. RESULTS AND DISCUSSION

The marketplace in this study is intended for sorghum entrepreneurs who have joined the Sorghum Entrepreneur Unit (UES). Marketplace features are adapted to marketplace features in general where sorghum entrepreneurs as sellers can display products that can be purchased, as seen in Figure 2.

Figure 2 describes the appearance of the Sorghum Entrepreneur Marketplace which displays sorghum products that are ready to be marketed by sorghum entrepreneurs. On the left, buyers can choose the category of sorghum and the system can display products according to the category chosen by the buyer. In addition, buyers can also choose the required price limit so that the system can display products according to the price limit chosen by the buyer.
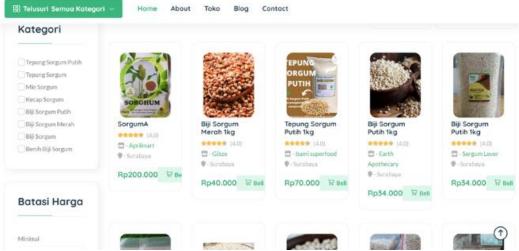

Fig. 2. Marketplace Entrepreneur Sorghum

In this study, the clustering method used was K-Medoid clustering. The purpose of K-Medoid clustering is to separate the data into a certain number of "k" clusters. The concept of this cluster is based on the level of data similarity based on the Euclidean distance. The result of this K-Medoid clustering is that each respondent has a data label according to the class formed. From the results of the questionnaire, obtained data on the level of desire for UES for each variable from 35 sorghum entrepreneurs as respondents. Quantitative data obtained is then entered and then proceed with the clustering stage. The data is divided into several clusters based on the type of sorghum, products sold and sales results. After that, the cluster performance calculation is carried out by calculating the k-fold value. In the K-Medoids algorithm, the most important aspects are the number of classes determined, the cluster center, and the number of iterations. For the clustering calculation process, it can be seen in Figure 3.

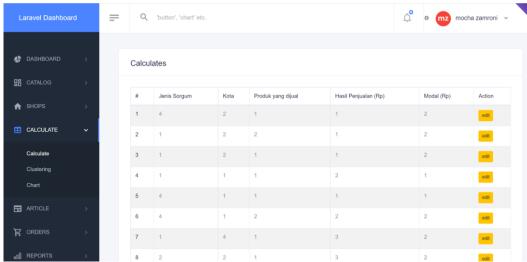*Sorghum Entrepreneur Marketplace Using K-Medoids Clustering (Nia Saurina)*

Fig. 3. Clustering Calculation Process in Sorghum Entrepreneur Marketplace

K-fold cross validation is a validation technique to assess how the results of statistical analysis generalize data sets. This technique is used as a predictive model which accommodates the estimated parameters of a model when it is run. One technique of cross validation is k-fold cross validation, which breaks data into k subsets of data sets of the same size. The use of k-fold cross validation to eliminate bias in the data [13].

Test 1	Test	Test	Test	Test	Test
Test 2	Test	Test	Test	Test	Test
Test 3	Test	Test	Test	Test	Test
Test 4	Test	Test	Test	Test	Test
Test 5	Test	Test	Test	Test	Test

Fig. 4. Calculation Example of K-fold cross validation [13]

Calculation steps for K-fold cross validation are:
1. Experiment 1, which is to make the first partition part of the test data and the other partitions the training data.
2. Experiment 2, which is to make the second part of the partition become data testing and the other partitions become data training.
3. Experiment 3, which is to make the third part of the partition become data testing and the other partitions become data training and so on

Testing the effect of the k-fold value is carried out using the value of k-fold = 2 to k-fold = 10. At each k-fold, input the values K=1 to K=10. The results of testing the effect of the k-fold value produce different accuracy values so that the average value is calculated to determine the best accuracy. The conclusion from the analysis of the results of testing the effect of the k-fold value is that the best percentage is found in the k-fold value = 6 with an accuracy value of 34.597%. In this test, the best k-fold value is not in one particular value because the graph results show that the percentage of accuracy has a value that fluctuates for each k-fold. The average result of testing the effect of the k-fold value can be seen in Figure 5. Based on the test of the effect of the K value and the effect of the k-fold value, there is no significant difference because the data is in the form of binary numbers. In addition, the symptoms in the training data do not show the dominant symptom equation even though the training data are in the same class. Conversely, the training data has many similarities in symptoms while the training data are in different classes

Fig. 5. Effect testing graph k-fold value

Clustering results can be used by UES to provide assistance to sorghum entrepreneurs based on the type of sorghum, products sold and sales proceeds, as can be seen in Figure 6. There are 2 axes, namely the horizontal X axis showing the sorghum entrepreneur group, namely C1 - sorghum entrepreneurs who had the smallest business capital, C2 - sorghum entrepreneurs who had a large number of regular customers, C3 -sorghum entrepreneurs who had few employees, C4 - sorghum entrepreneurs who carry out promotions on social media, C5 - Entrepreneur sorghum that relieves the least. Meanwhile, the vertical Y axis shows the number of sorghum entrepreneurs who have been successfully grouped by clustering.
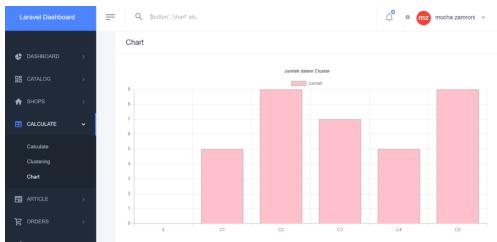

Fig. 6. Clustering Results in Sorghum Entrepreneur Marketplace

Function testing uses the Blackbox method approach. The test items are carried out on the main functions in the marketplace. The following table 2 displays the results of testing the sorghum entrepreneur marketplace function that has been developed.

Table 2. Marketplace Entrepreneur Sorghum Function Test Results

No	Test Item	Test Result
1	Saving sorghum entrepreneur data, buyer data and sorghum product data to the marketplace	Succeed
2	Adding transactions to the cart	Succeed
3	Displays sales reports per transaction, per date, per month and per year	Succeed
4	Sorghum entrepreneur grouping into 5 groups	Succeed
5	Displays the results of K-Medoids clustering	Succeed

IV. CONCLUSION

- The existence of a marketplace as a place of buying and selling, can facilitate and provide facilities for Sorghum Entrepreneurs in several cities in East Java Province in carrying out the process of buying and selling sorghum products to the wider community. This marketplace can also save costs because sorghum entrepreneurs get facilities to promote the sorghum products they sell. In addition, the presence of this marketplace can make it easier for the public to find out information on sorghum products and make the sorghum purchase process easy and real-time, thereby saving time in transactions.
- Testing the effect of the k-fold value is carried out using the value of k-fold = 2 to k-fold = 10. At each k-fold, input the values K=1 to K=10. The results of testing the effect of the k-fold value produce different accuracy values so that the average value is calculated to determine the best accuracy. The conclusion from the analysis of the results of testing the effect of the k-fold value is that the best percentage is found in the k-fold value = 6 with an accuracy value of 34.597%. In this test, the best k-fold value is not in one particular value because the graph results show that the percentage of accuracy has a value that fluctuates for each k-fold.
- This Marketplace can provide information on classification results using the K-Medoid method to classify Sorghum Entrepreneurs based on the income that UES has earned.
- There is mapping information based on clustering results from sales of sorghum entrepreneurs so that it helps UES to provide assistance to sorghum entrepreneurs.

ACKNOWLEDGMENT

REFERENCES

[1] Apriadi Deni, Saputra. "E-commerce Berbasis Marketplace Dalam Upaya Mempersingkat Distribusi Penjualan Hasil Pertanian". *Jurnal RESTI. Vol. 1 No. 2. Lubuklinggau: STMIK Bina Nusantara Jaya.* 2017. DOI: https://doi.org/10.29207/resti.v1i2.36

[2] Ayu Nandita Ashari, Ratna Mutu Manikam. "Analisa dan Perancangan Marketplace Bahan Bangunan Berbasis Website". *JUSIBI - (Jurnal Sistem Informasi Dan E-Bisnis), Volume 1, Issue 3.* 2019. DOI: https://doi.org/10.54650/jusibi.v1i3

[3] Yovi Apridiansyah, Ardi Wijaya. "Aplikasi E – Commerce Hasil Laut Bengkulu Dengan Penerapan Algoritma Seo Menggunakan Model Sematik Web". *JUSIBI - (Jurnal Sistem Informasi Dan E-Bisnis). Volume 1, Issue 5.* 2019. DOI: https://doi.org/10.54650/jusibi.v1i5

[4] Ghada Taher Al-Lami, Alhamzah Alnoor. "E-Commerce: Advantages and Limitations*". International Journal of Academic Research in Accounting Finance and Management Sciences Volume 11 Number 1.* 2021. DOI: https://doi.org/10.6007/IJARAFMS/v11-i1/8987

[5] Indra Mawardi, Nurdin Nurdin, Zulkarnaini Zulkarnaini. "Appropriate Technology Program of Postharvested Coffee: Production, Marketing, and Coffee Processing Machine Business Unit", *Indonesian Journal of Community Engagement*. Volume 5, Issues 2. 2019. DOI: https://doi.org/10.22146/jpkm.36470

[6] Rinny Yudowati, Dwi Handiatmoko. "Pengaruh e–marketing dan e-comerce terhadap pendapatan usaha umkm di tengah pandemi covid 19 (studi kasus kecamatan duren sawit) the effect of e - marketing and e commerce on MSME business revenue in the middle of the Covid-19 pandemic (case study of Duren Sawit District)". *JISAMAR – Journal of Information System, Applied Management, Accounting and Research. Volume 5 Issue 4. 2021.* DOI: https://doi.org/10.52362/jisamar.v5i4

[7] B. Wira, A. E. Budianto, and A. S. Wiguna, "Implementasi Metode K-Medoids Clustering Untuk Mengetahui Pola Pemilihan Program Studi Mahasiwa Baru Tahun 2018 Di Universitas Kanjuruhan Malang," *Jurnal Terapan Sains Teknol., vol. 1, no. 3.* 2019. https://doi.org/10.21067/jtst.v1i3.3046

[8] Evi Rahmah, Elin Haerani, Alwis Nazir, Siti Ramadhani. "Penerapan Algoritma K-Medoids Clustering Untuk Menentukan Srategi Promosi Pada Data Mahasiswa (Studi Kasus : Stikes Perintis Padang)". *JNKTI – Jurnal Nasional Komputasi dan Teknologi Informasi. Volume 5, Issue 3.* 2022. DOI: https://doi.org/10.32672/jnkti.v5i3

[9] Febiyanti Alfiah, Almadayani Almadayani, Danial Al Farizi, Edy Widodo. "Analisis Clustering K-Medoids Berdasarkan Indikator Kemiskinan di Jawa Timur Tahun 2020". *Jurnal Ilmiah Sains. Volume 22, Issue 1.* 2022. DOI: https://doi.org/10.35799/jis.v22i1.35911

[10] Han, J. & Kamber, M. "Data Mining, Concept and Techniques". Morgan Kauffman Publisher, Waltham. 2017.

[11] Dewi, D.A.I.C. & Pramita, D.A.K. Analisis Perbandingan Metode Elbow Dan Sillhouette Pada Algoritma Clustering K-Medoids Dalam Pengelompokkan Produksi Kerajinan Bali. *Jurnal Matrix, Volume 9 Issue 3.* 2019. https://doi.org/10.31940/matrix.v9i3.1662

[12] Daffa Alya Alodia, Astrika Praokta Fialine, Desy Endriani, Edy Widodo. "Implementasi Metode K-Medoids Clustering untuk Pengelompokan Provinsi di Indonesia Berdasarkan Indikator Pendidikan". *SEPREN – Journal of Mathematics Education and Applied. Volume 3, Issues 1.* 2021. DOI: https://doi.org/10.36655/sepren.v2i2.606

[13] F. Tempola, M. Muhammad, and A. Khairan, "Perbandingan Klasifikasi Antara KNN dan Naive Bayes pada Penentuan Status Gunung Berapi dengan K-Fold Cross Validation," *Jurnal Teknologi Informatika dan Ilmu Komputasi, vol. 5, no. 5.* 2018, doi: https://doi.org/10.25126/jtiik.201855983.