CLUSTERING OF SCREW PRESS MACHINE CONDITIONS USING THE K-MEDOIDS METHOD

TUGAS AKHIR

Disusun Sebagai Salah Satu Syarat
Untuk Memperoleh Gelar Sarjana Teknik
Pada Jurusan Teknik Informatika

Oleh

M. TAUFIK APRINALDO
NIM. 12150112222

UIN SUSKA RIAU

FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU
PEKANBARU
2025

LEMBAR PERSETUJUAN

CLUSTERING OF SCREW PRESS MACHINE CONDITIONS USING THE K-MEDOIDS METHOD

TUGAS AKHIR

Oleh

M. TAUFIK APRINALDO
NIM. 12150112222

Telah diperiksa dan disetujui sebagai Laporan Tugas Akhir
di Pekanbaru, pada tanggal 4 Desember 2025

Pembimbing I,	Pembimbing II,
Jasril, S.Si., M.Sc.	**Suwanto Sanjaya, S.T., M.Kom.**
NIP. 197102152000031002	**NIP. 198702072024211009**

LEMBAR PENGESAHAN

CLUSTERING OF SCREW PRESS MACHINE CONDITIONS USING THE K-MEDOIDS METHOD

Oleh

M. TAUFIK APRINALDO

NIM. 12150112222

Telah dipertahankan di depan sidang dewan penguji
sebagai salah satu syarat untuk memperoleh gelar Sarjana Teknik
pada Universitas Islam Negeri Sultan Syarif Kasim Riau

Pekanbaru, 4 Desember 2025

Mengesahkan,

Dekan,	Ketua Jurusan,
Dr. Yuslenita Muda, S.Si., M.Sc.	**Muhammad Affandes, S.T., M.T.**
NIP. 19770103 2000710 2 001	**NIP. 19861206 201503 1 004**

DEWAN PENGUJI

Ketua	: Muhammad Affandes, S.T., M.T.
Pembimbing I	: Jasril, S.Si., M.Sc.
Pembimbing II	: Suwanto Sanjaya, S.T., M.Kom.
Penguji I	: Dr. Lestari Handayani, S.T., M.Kom.
Penguji II	: Fitri Insani, S.T., M.Kom.

SURAT PERNYATAAN

Saya yang bertandatangan di bawah ini:

Nama	:	M. Taufik Aprinaldo
NIM	:	12150112222
Tempat/Tgl. Lahir	:	Pekanbaru, 10 April 2002
Fakultas	:	Sains dan Teknologi
Prodi	:	Teknik Informatika
Judul Skripsi	:	Clustering of Screw Press Machine Conditions using the K-Medoids Method

Menyatakan dengan sebenar-benarnya bahwa :

1. Penulisan jurnal dengan judul sebagaimana tersebut di atas adalah hasil pemikiran dan penelitian saya sendiri.
2. Semua kutipan pada karya tulis ini sudah disebutkan sumbernya.
3. Oleh karena itu jurnal saya ini, saya nyatakan bebas dari plagiat.
4. Apabila dikemudian hari terbukti terdapat plagiat dalam penulisan jurnal saya tersebut, maka saya besedia menerima sanksi sesua peraturan perundang-undangan.

Demikian surat pernyataan ini saya buat dengan penuh kesadaran dan tanpa paksaan dari pihak manapun juga.

Pekanbaru, 18 Desember 2025

Yang membuat pernyataan

METERAI TEMPEL 9C6ANX165843917

M. TAUFIK APRINALDO
NIM. 12150112222

Engineering and Technology Journal

Website: http://everant.org/index.php/etj

ISSN (print): 2456-3358
DOI: 10.47191/etj

Impact Factor : 8.482
ASI Score : 1.3 , IPI Value: 2.68

Letter of Acceptance

Dear Author: M. Taufik Aprinaldo, Jasril, Suwanto Sanjaya, Lestari Handayani, Fitri Insani,

Manuscript ID: ETJV10I11-22

Paper Title: "Clustering of Screw Press Machine Conditions Using the K-Medoids Method"

We are pleased to accept the same for publication in ETJ. Please send the scanned Copyright form (Can be downloaded from website) along with bank receipt of an online maintenance. Article will be online within 24 working hours after receiving all the necessary documents.

Payment details: 35 USD (For entire research paper for All Author).

You can pay by Credit Card or Debit card or net banking by using link

Payment Link: http://everant.org/index.php/etj/mod

Or

Stripe link :- https://buy.stripe.com/8wM7wsfi89A7a1adWh

In case of any query please do not hesitate to contact us. Early reply is appreciated. Sincerely,
E-mail ID: journaletj@gmail.com

With Regards,
Journal Manager
Engineering and Technology Journal
Website: http://everant.org/index.php/etj

LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL

Tugas Akhir yang tidak diterbitkan ini terdaftar dan tersedia di Perpustakaan Universitas Islam Negeri Sultan Syarif Kasim Riau adalah terbuka untuk umum dengan ketentuan bahwa hak cipta pada penulis. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau ringkasan hanya dapat dilakukan seizin penulis dan harus disertai dengan kebiasaan ilmiah untuk menyebutkan sumbernya.

Penggandaan atau penerbitan sebagian atau seluruh Tugas Akhir ini harus memperoleh izin dari Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Perpustakaan yang meminjamkan Tugas Akhir ini untuk anggotanya diharapkan untuk mengisi nama, tanda peminjaman dan tanggal pinjam.

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa:

1. Tugas Akhir ini dengan judul "CLUSTERING OF SCREW PRESS MACHINE CONDITIONS USING THE K-MEDOIDS METHOD" adalah gagasan asli dari saya sendiri dan belum pernah dijadikan Tugas Akhir atau sejenisnya di Universitas Islam Negeri Sultan Syarif Kasim Riau maupun di perguruan tinggi lain.
2. Dalam Tugas Akhir ini TIDAK terdapat karya atau pendapat yang telah dipublikasikan orang lain, kecuali tertulis dengan jelas dan dicantumkan sebagai referensi di dalam Daftar Pustaka.
3. Dalam Tugas Akhir ini TIDAK terdapat penggunaan Kecerdasan Buatan Generatif (Generative AI) yang bertentangan dengan ketentuan dan peraturan yang berlaku.
4. Saya bersedia menerima sanksi sesuai ketentuan yang berlaku apabila di kemudian hari terbukti bahwa Tugas Akhir ini melanggar kode etik maupun peraturan yang berlaku, termasuk plagiat ataupun pelanggaran hak cipta.

Demikianlah pernyataan ini dibuat dengan sebenarnya.

Pekanbaru, 18 Desember 2025

Yang membuat pernyataan,

M. TAUFIK APRINALDO

NIM. 12150112222

LEMBAR PERSEMBAHAN

Tugas akhir ini penulis persembahkan sebagai bentuk semangat, usaha, serta ungkapan cinta dan kasih sayang kepada orang-orang terpenting dalam hidup penulis. Dengan ketulusan hati dan rasa terima kasih yang mendalam, tugas akhir ini penulis persembahkan untuk:

1. Kedua orang tua tercinta, Ayah Samual dan Ibu Asna, serta abang dan kakak, juga seluruh keluarga besar penulis yang telah memberikan dukungan moral, materil, serta doa dan restu, sehingga penulis dapat menempuh pendidikan hingga jenjang S1 di Jurusan Teknik Informatika, UIN Sultan Syarif Kasim Riau.
2. Dosen pembimbing, Bapak Jasril, S.Si., M.Sc. dan Bapak Suwanto Sanjaya S.T., M.Kom., yang telah memberikan rasa sabar dan ikhlas dalam memberikan arahan, dan motivasi hingga tugas akhir ini dapat terselesaikan dengan baik.
3. Seluruh dosen pengajar yang telah mendidik penulis dengan penuh kesabaran dan keikhlasan, sehingga ilmu yang diperoleh selama masa perkuliahan dapat menjadi bekal yang bermanfaat di masa depan.
4. Teman-teman seperjuangan di Program Studi Teknik Informatika, UIN Sultan Syarif Kasim Riau, atas kebersamaan dan dukungan selama menempuh perjalanan akademik.
5. Teman-Teman YK yang selalu membersamai saya dari awal perkuliahan serta selalu ada disaat saya butuh pertolongan selama pengerjaan Tugas Akhir ini.

Engineering and Technology Journal e-ISSN: 2456-3358
Volume 10 Issue 11 November-2025, Page No.-7916-7922
DOI: 10.47191/etj/v10i11.22, I.F. – 8.482

Clustering of Screw Press Machine Conditions using the K-Medoids Method

M. Taufik Aprinaldo[1], Jasril[2*], Suwanto Sanjaya[3], Lestari Handayani[4], Fitri Insani[5]
[1,2,3,4] Informatics Engineering, Faculty of Science and Technology, Sultan Syarif Kasim Riau State Islamic University.

ABSTRACT: The screw press plays an important role in the oil extraction process; thus, monitoring its condition is essential to maintain performance and prevent failures. This study aims to cluster screw press machine conditions using the K-Medoids method. The dataset consisted of 23,002 records from PT. XYZ was collected in April–May 2024 with two attributes: temperature and pressure. The data was processed through selection, pre-processing, and transformation stages using z-score normalization before clustering. Model evaluation employed the Silhouette Coefficient and the Davies-Bouldin Index (DBI). The results show that the best configuration was at K = 7, with a Silhouette value of 0.5494 and a DBI of 0.5521, indicating a reasonable structure and good separation. Thus, the K-Medoids method has been proven effective in clustering screw press machine conditions and useful in supporting machine maintenance decision-making.

KEYWORDS: Clustering, K-Medoids, Silhouette Coefficient, Davies-Bouldin Index, Screw Press.

I. INTRODUCTION

The screw press machine is a tool that plays an important role in the oil extraction process, where the selection of appropriate parameters is necessary to maintain the quality of valuable substances in the extracted oil [1]. In manufacturing industries such as palm oil processing, this machine has several important components, such as worm screws, extension shafts, bearings, press cages, and oil seals that require routine maintenance [2]. The performance of the production machine itself is highly dependent on the level of reliability and availability, where damage (downtime) to the screw press often causes the production process to be suboptimal and the company's targets to not be achieved [3]. Additionally, research on screw press hub damage found that friction and dynamic loads trigger material fatigue, initial cracks, and even breakage before four months of use [4].

Given the operational complexity of screw press machines and the importance of optimal parameters in their performance, a systematic approach is needed to monitor and analyze the condition of these machines. Clustering is an unsupervised data analysis technique used to group data based on specific patterns and similarities, which plays an important role in various fields, including machine condition monitoring [5]. This technique becomes particularly relevant when combined with Machine Learning (ML) technology to analyze machine operational data in greater depth. In its industrial application, ML can be used as a predictive maintenance method by utilizing data collected from IoT devices installed on machines to detect early faults and prevent major failures, as demonstrated in the case of knitting machines with an accuracy rate of 92% [6]. Specifically, clustering helps in analyzing unstructured and high-dimensional data in the form of sequences, expressions, text, and images [7].

Previous research has applied the Fuzzy C-Means algorithm to cluster screw press machine conditions based on temperature and pressure. The performance of the screw press machine significantly affects the quality and efficiency of palm oil production. The Back Pressure Vessel (BPV), which is responsible for distributing steam to various process stations, is an important part of this system [8]. Meanwhile, the *K-Means* algorithm has also been used on the same dataset with testing of up to fifteen cluster configurations. Based on *DBI,* the best quality was obtained with three clusters, while the *Elbow* method indicated four clusters as the optimal choice [9]. Both studies confirm that *clustering* techniques are effective in describing machine operational patterns, although they are still limited to the use of certain variables, thus requiring further study with other algorithms such as K-Medoids.

Therefore, the researcher will conduct research on clustering the conditions of screw press machines. The main focus of this study is to cluster screw press machine conditions using the k-medoids method. K-medoids is a partition-based clustering algorithm that uses actual points as medoids, which are the most central points in a cluster [10]. The selection of the k-medoids method is based on its advantages in handling machine condition data, where, in the context of machine condition analysis, K-medoids has been applied to model uncertainty in large-scale electrical power distribution systems, demonstrating high accuracy and scalability compared to other methods. Furthermore, with the application of K- medoids, data can be organized into more representative clusters, enabling more efficient decision-

making related to monitoring and improving system performance [11].

Previous studies have proven the effectiveness of the K-medoids algorithm in various application domains. This algorithm successfully grouped songs on Spotify based on their popularity and distribution in playlists with a Silhouette Score of 0.5014, which indicates good cluster separation and has practical implications for the music industry in designing more targeted marketing strategies [12]. Additionally, K-medoids has also been proven optimal in classifying earthquake vulnerability levels in Indonesia with k=2 and a maximum Silhouette Coefficient of 0.68016, successfully identifying 390 earthquake events with a very high vulnerability level in Eastern Indonesia and 179 events with a high vulnerability level in Western Indonesia, providing an important reference for the government in planning earthquake risk mitigation [13]. Furthermore, this algorithm is also effective in grouping violence-prone areas in the city of Padang using data from 2,434 cases spread across 11 subdistricts, producing 3 clusters that can categorize areas based on the level of violence cases from highest to lowest [14]. The success of K-medoids in these three studies demonstrates the consistency and adaptability of this algorithm in handling various types of data and clustering problems, making it a strong foundation for its application in analyzing the condition of screw press machines.

II. MATERIALS AND METHODS

A. Research Stages

This research follows a number of systematic steps designed to analyze and process data accurately. The initial step begins with selecting a dataset to determine the information to be analyzed. This is followed by a pre-processing stage to clean the data so that it is ready for transformation. After the data is cleaned, the process continues to the transformation phase, which includes data normalization. The next step is clustering using the k-medoids algorithm. The final stage is evaluation to determine the optimal number of clusters from the application of the algorithm used. Figure 1 shows the research flow.

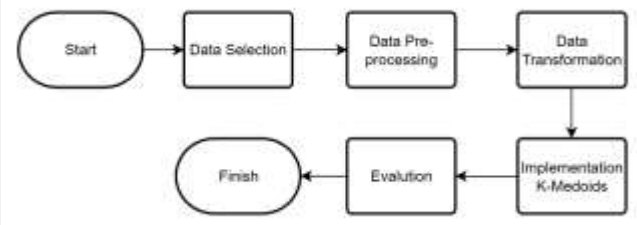

Figure 1: Research Flow

B. Data Selection

In this phase, data curation is carried out to ensure the suitability of the dataset with the research objectives, including the selection of relevant variables and the determination of data types that are in line with the methodology to be applied.

C. Data Pre-processing

This pre-processing stage is implemented to filter and purify the dataset from elements that can hinder the analysis process. The stages carried out include removing redundant data, addressing missing values, correcting writing errors, and harmonizing inconsistent data formats.

D. Data Transformation

In data transformation, there are various commonly used transformation techniques such as min-max scaling, z-score, decimal scaling, robust scaling, one-hot encoding, and others. In this study, the author will use the z-score technique in the data transformation section [15]. The z-score Formula [16] can be seen in equation (1):

$$X' = \frac{Xi - \text{mean(x)}}{\text{std(x)}} \qquad (1)$$

Where $X'$is the normalization result value, Xi is the normalized data, mean(x) is the mean value of an attribute, and std(x) is the standard deviation of an attribute.

E. K-Medoids Implementation

K-medoids is a data clustering method that can provide more balanced results in terms of group distribution compared to other methods, such as K-Means [17]. K-medoids is a partition-based clustering algorithm that uses actual points as medoids, which are the most central points in a cluster [10]. The steps in the K-Medoids method are [18] :

a. Initialize medoids by randomly selecting k objects from the dataset as initial medoids.
b. Next, each object is assigned to *the cluster* with the nearest medoid based on the Euclidean distance using equation (2).

$$d(A,B) = \sqrt{[(x_1 - x_2)^2 + (y_1 - y_2)^2]} \qquad (2)$$

Explanation :

$d(A,B)$ = Euclidean distance between point A and point B
A = Object to be calculated
B = Cluster center
x_1 = The first-dimensional coordinate value of point A
x_2 = The first-dimensional coordinate value of point B
y_1 = The value of the second dimension coordinate of point A
y_2 = The value of the second dimension coordinate of point B

c. Then, a new medoid is selected where, for each cluster, the total distance from all objects to other objects in the same cluster is calculated.
d. Calculate the total deviation (S) by calculating the total new distance – total old distance. If S< 0, then swap the object with the cluster data to form a new set of k objects as the medoid. If S> 0, then the total distance increases and no replacement is made (use

the old medoid as a reference for the next iteration), using equation (3).

$$S = \sum D_{new} - \sum D_{old} \quad (3)$$

Where S is the total deviation, $\sum D_{new}$ is the total distance to the new medoid, and $\sum D_{old}$ is the total distance from the old medoid.

e. Repeat steps C and D until there is no change in the medoid (S=0).

B. Model Evaluation

The evaluation methods to be used are the Silhouette Coefficient and Davies-Bouldin Index (DBI).

a. Silhouette Coefficient

The Silhouette Coefficient describes the quality of a data point's placement within its cluster by comparing its proximity to its own cluster and other clusters. This coefficient ranges from -1 to 1. Evaluation using the Silhouette method [19] is calculated using equation (4).

$$S(x_i) = \frac{b(x_i) - a(x_i)}{Max\{b(x_i), a(x_i)\}} \quad (4)$$

Where $S(x_i)$ is the silhouette coefficient value for the i-th data point, $a(x_i)$ is the average distance between the i-th object and other objects in the same cluster (intra-cluster), and $b(x_i)$ is the average distance between the i-th object and objects in the nearest different cluster (inter-cluster).

b. Davies-Bouldin Index

In order to maximize the number of clusters formed, the Davies-Bouldin Index (DBI) is employed to calculate how many clusters should be formed [20]. Unlike the Silhouette Coefficient, a lower DBI value indicates better clustering quality. DBI has a value range from 0 upwards, with values close to 0 indicating optimal cluster separation and high internal cohesion. The DBI evaluation guidelines are as follows: values less than 1.0 indicate good clustering, values between 1.0 and 2.0 indicate acceptable clustering, while values above 2.0 indicate suboptimal clustering with high overlap between clusters. There are four steps in the DBI evaluation [21] :

The evaluation begins with calculating *SSW* and *SSB* using the initial centroid as a reference, followed by calculating the *DBI* value. The Formula applied to obtain *the Sum of Square Within cluster* (SSW) is :

$$SSW = \frac{1}{m}\sum_{j=i}^{m_i} d(x_j, C_i) \quad (5)$$

Where m_i is the number of data points in cluster i, C_i is the centroid of cluster i, x_j is the data in the cluster, and $d(x_j, C_i)$ is the distance between the data and the centroid.

The Sum of Squares Between-Cluster (SSB) has a function to analyze the level of separation between data groups. The SSB metric serves to evaluate how separated the clusters are by calculating the distance between the center point of each cluster (centroid) and the center point of the entire dataset. The higher the SSB value obtained, the more optimal the separation between clusters. The mathematical Formula applied in the calculation of SSB is :

$$SSB_{ij} = d(C_i, C_j) \quad (6)$$

Where $d(C_i, C_j)$ is the distance between clusters.

The next step is to calculate the ratio value of each cluster using the following Formula:

$$R_{i,j} = \frac{SSW_i + SSW_j}{SSB_{i,j}} \quad (7)$$

Where $R_{i,j}$ is the ratio of SSW and SSB between cluster i and cluster j, SSW_i is the Sum of Squares Within-cluster for cluster i, SSW_j is the Sum of Squares Within-cluster for cluster j, and $SSB_{i,j}$ is the Sum of Squares Between-cluster between cluster i and j.

The final phase in the evaluation process involves calculating the DBI by taking the average value of the maximum ratio found between each data group and the other data groups. The DBI index serves as a parameter to measure the level of internal compactness in each group and the quality of separation formed between data groups. The mathematical Formula for calculating the Davies-Bouldin Index can be formulated as :

$$DBI = \frac{1}{k}\sum_{i}^{k} max(R_{i,j}) \quad (8)$$

Where $R_{i,j}$ is the SSW and SSB ratio between cluster i and cluster j, and k is the number of clusters formed.

III.RESULT AND DISCUSSION

A. Data Selection

This study will process the original dataset from screw press devices with BPV device identification during the period from April to May 2024, obtained from PT.XYZ, with a total of 23,002 data rows. The initial data consists of no, rowstamp, device code, date, temperature, pressure, pH, current, created at, created by, and moved at history. Then, data selection is carried out, and two important variables are taken, namely temperature and pressure.

In the next stage, the main characteristics of the existing data were collected and processed to form a dataset of 23,002 data points used for system analysis and modeling. This study applied a feature selection method by taking two main parameters, namely Temperature and Pressure, as input variables for further data processing. These two attributes were selected because they have a significant correlation with the operational performance and condition status of the screw press device. The selected data results are shown in Table 1.

Table 1. Dataset after data selection

No	Temperature	Pressure
0	123	3,19
1	123	3
...	...	...
23000	100	2,12
23001	100	2,03

B. Data Pre-processing

In the next stage, important attributes from the dataset were sorted and processed for checking.

Table 2. Data verification results

No	Attribute	Count	Null Count	Data Type
1	Temperature	23002	Not-Null	Float64
2	Pressure	23002	Not-Null	Float64

After checking the data, the results show that there are no null values in the Temperature and Pressure attributes, which contain 23002 data points and have the same data type, Float64, as shown in Table 2.

C. Data Transformation

At this stage, data transformation is performed to standardize the scale between variables with different value ranges, namely *temperature* and *pressure*, so that no variable is more dominant in the distance calculation process in the K-Medoids algorithm. The results of the dataset transformation in Table 1 using equation (1) can be seen in Table 4 below.

Table 3. All data after transformation

No	Temperature	Pressure
0	0,693464	0,826668
1	0,693464	0,588746
...	...	...
23000	0,201872	-0,513208
23001	0,201872	-0,625908

D. K-Medoids Implementation

In this section, based on Table 3, distance calculations were performed using equation (2) and Total Deviation using equation (3) until there were no changes in the medoids. The calculation results can be seen in Figure 2 and Table 4 .

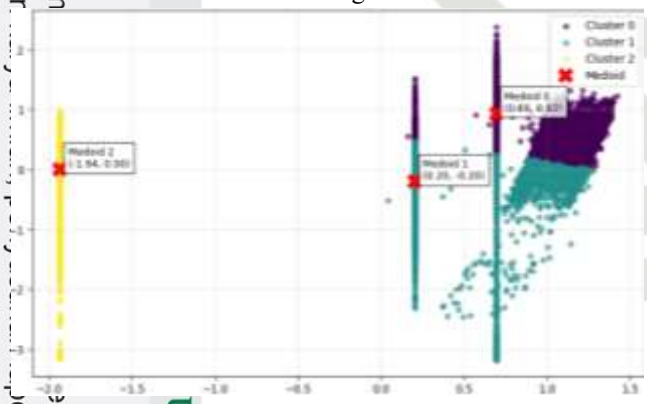

Figure 2: Visualization of clustering with medoid centers

The results of applying the K-Medoids method to the entire dataset can be seen in Figure 2, which shows the positions of the medoids and the distribution of clusters after the process reaches a stable condition.

Table 4. Distribution of the number of members in each cluster

No	Temperature	Pressure	Cluster
0	0,693464	0,826668	0
1	0,693464	0,588746	0
...	...	...	...
22987	0,201872	0,676402	0
22988	0,693464	0,263169	0
5	0,693464	0,225602	1
6	0,693464	0,213080	1
...	...	...	...
23000	0,201872	-0,513208	1
23001	0,201872	-0,625908	1
743	-1,935483	0,000203	2
744	-1,935483	0,050291	2
...	...	...	...
22996	-1,935483	0,313258	2
22997	-1,935483	0,576224	2

In addition, the *clustering* results showing the number of members in each *cluster* can be seen in Table 5, namely *cluster* 0 contains 7,356 data, *cluster* 1 contains 11,173 data, and *cluster* 2 contains 4,473 data.

E. Model Evaluation

In the evaluation stage, the clustering results were tested using two methods, namely the Silhouette Coefficient and the Davies-Bouldin Index (DBI). The Silhouette Coefficient serves to assess the quality of separation between clusters, while the Davies-Bouldin Index (DBI) measures the degree of similarity between clusters.

a. Silhouette Coefficient

The first step in the evaluation process is the application of the *Silhouette Coefficient* method, as shown in equation (4). This method provides a quantitative measure of how well each data point is placed in the appropriate *cluster*, taking into account the comparison of its proximity to its own *cluster* and its distance from other *clusters*. The results of the evaluation using *the silhouette* can be seen in Table 5, the graph is attached in Figure 3, and the cluster distribution with 7 clusters is shown in Figure 4.

Table 5. Silhouette Coefficient evaluation results

Cluster	Silhouette Score
2	0,2627
3	0,4572
4	0,4339
5	0,3917
6	0,5225
7	0,5494
8	0,5349
9	0,5395
10	0,5299

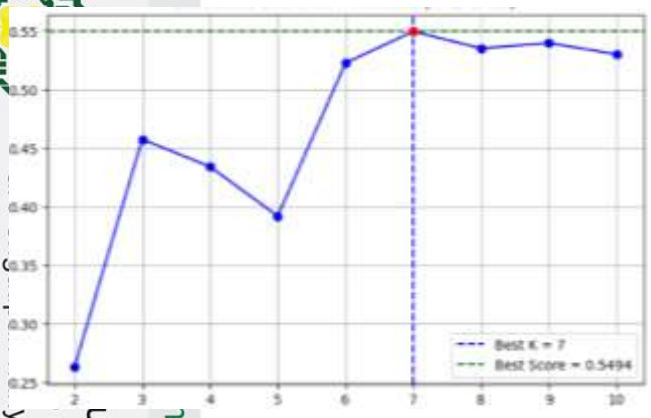

Figure 3: Silhouette Coefficient evaluation graph

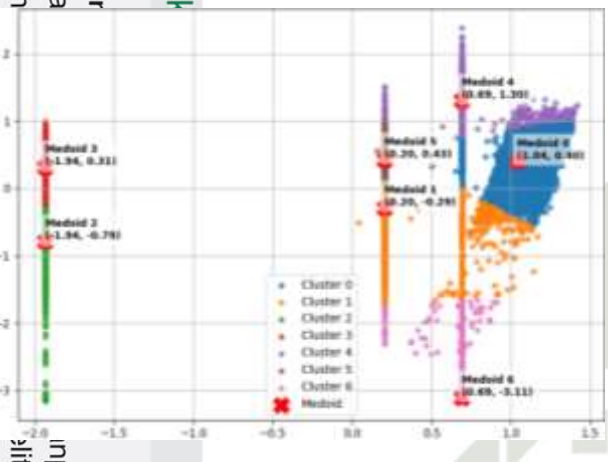

Figure 4: Cluster distribution with k=7

The Silhouette Coefficient approach was used to assess the quality of clustering with different numbers of clusters (K = 2 to K = 10). The test results showed that the highest Silhouette Coefficient value was obtained at K = 7 with a score of 0.5494, while the lowest value was found at K= 2 with a score of 0.2627. Based on the evaluation criteria, a Silhouette Coefficient value range between 0.5 and 0.7 indicates that the clustering structure formed is in the fair/good category, even though it is not completely separated.

b. Davies-Bouldin Index

The next evaluation stage was conducted using the *Davies-Bouldin Index (DBI)* method. This index provides a quantitative measure of cluster quality by calculating the average level of similarity between clusters. The *DBI* value is obtained from a comparison between the distance between cluster centers and the distribution of data within the cluster. The smaller the *DBI* value, the better the clustering quality, as it indicates that the data within a cluster is sufficiently dense and clearly separated from other clusters. The results of testing using *DBI* can be seen in Table 6 and the graph for the *DBI* method is attached in Figure 5.

Table 6. DBI evaluation results

Cluster	DBI Value
2	1,0537
3	0,7719
4	0,722
5	0,8586
6	0,6157
7	0,5521
8	0,5529
9	0,5635
10	0,5707

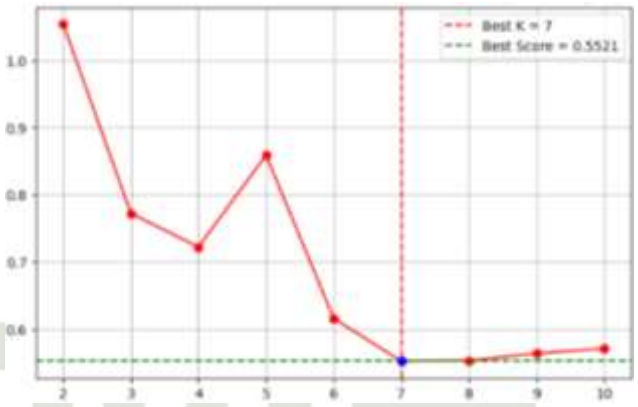

Figure 5: Davies-Bouldin Index evaluation graph

Based on the DBI calculation results for the number of clusters (K) between 2 and 10, the index value variations are shown in Table 7 and Figure 4 above. The highest DBI value occurs at K = 2 with 1.0537, while the lowest DBI value is obtained at K = 7 with 0.5521. This indicates that the configuration with seven clusters produces the most optimal separation, because each cluster is sufficiently compact within itself and has a relatively large distance from other clusters. Thus, the best number of clusters according to the Davies-Bouldin Index evaluation is K = 7.

F. Best Cluster Result

The optimal number of clusters was found at k=7 based on the evaluation findings utilizing the Silhouette and Davies Bouldin Index (DBI) methodologies. The highest silhouette value and lowest DBI both consistently indicate optimal cluster separation quality. Therefore, further analysis uses 7 clusters as the best configuration.

Table 7. Range of values for each cluster

Cluster	Min Temperature	Max Temperature	Min Pressure	Max Pressure
0	121	155,50	2,10	3,36
1	92,5	145,25	1,15	2,58
2	0	0	0,02	2,32
3	0	0	2,36	3,32
4	100	157	3,16	4,43
5	98	114	2,59	3,33
6	100	133,25	0	1,21

Table 7 shows the temperature and pressure ranges for each cluster. From the table, it can be seen that Cluster 0 has a high

temperature range of 121–155.5°C with moderate pressure of 2.10–3.26. Clusters 1 and 4 have relatively high temperatures reaching 155.5–157°C with considerable pressure, where Cluster 4 shows maximum operating conditions with pressure up to 4.43. Meanwhile, Clusters 2 and 3 have a temperature value of 0 and very low pressure, even reaching 0, with pressure variations of 0.02-2.32 and 2.36-3.32, which indicates that the pressure in Cluster 3 is higher. Other clusters show variations in range that describe different operating conditions. For example, Cluster 5 shows stable operation in a medium level with temperatures of 98–114°C and pressures of 2.59–3.33, while Cluster 6 has normal temperatures of 100–133.25°C but low pressure.

CONCLUSIONS

This study successfully grouped screw press machine conditions using the K-Medoids method with a dataset of 23,002 records from PT. XYZ for the period April–May 2024. The research process began with data selection, pre-processing, transformation using Z-Score Normalization, implementation of K-Medoids clustering, and evaluation using the Silhouette Coefficient and Davies-Bouldin Index (DBI). The evaluation results showed that the best number of clusters was obtained at K=7, with the highest Silhouette Coefficient value of 0.5494, which is included in the fair good clustering structure category, and the lowest DBI value of 0.5521, which indicates optimal cluster separation quality. In other words, the seven-cluster configuration produced the most representative and stable data separation compared to other cluster number variations. Further analysis of the temperature and pressure ranges in each cluster shows that each cluster describes different machine operating conditions, ranging from idle conditions with almost zero pressure, normal operation in the medium temperature range, to maximum operating conditions with high temperatures and pressures. These findings indicate that the K-Medoids method can be used effectively to monitor the condition of screw press machines.

REFERENCES

1. B. M. Корендій and В. Б. Гавран, "Analysis of the oil extraction process and prospects of automation of screw press operation," *Sci. Bull. UNFU*, vol. 34, no. 1, pp. 85–90, 2024, doi: 10.36930/40340112.
2. F. Pohan, I. Saputra, and R. Tua, "Scheduling Preventive Maintenance to Determine Maintenance Actions on Screw Press Machine," *J. Ris. Ilmu Tek.*, vol. 1, no. 1, pp. 1–12, 2023, doi: 10.59976/jurit.v1i1.4.
3. S. Yasir and A. Saputra, "Analisa Reliability Dan Availability Mesin Screw Press Kelapa Sawit (Studi Kasus di PT. Ujong Neubok Dalam)," *Unistek*, vol. 9, no. 2, pp. 83–94, 2022, [Online]. Available: http://ejournal.unis.ac.id/index.php/UNISTEK
4. I. Anwar, A. Afdal, D. Denur, and D. Dermawan, "Analisis Penyebab Kerusakan Hub Screw Press dan Optimasi Teknik Pemeliharaan Terhadap Mesin Pabrik Pengolahan Kelapa Sawit," *J. Surya Tek.*, vol. 10, no. 1, pp. 733–737, 2023, doi: 10.37859/jst.v10i1.5003.
5. C. Tortora and F. Palumbo, "Clustering mixed-type data using a probabilistic distance algorithm[Formula presented]," *Appl. Soft Comput.*, vol. 130, p. 109704, 2022, doi: 10.1016/j.asoc.2022.109704.
6. S. Elkateb, A. Métwalli, A. Shendy, and A. E. B. Abu-Elanien, "Machine learning and IoT – Based predictive maintenance approach for industrial applications," *Alexandria Eng. J.*, vol. 88, no. January, pp. 298–309, 2024, doi: 10.1016/j.aej.2023.12.065.
7. M. R. Karim *et al.*, "Deep learning-based clustering approaches for bioinformatics," *Brief. Bioinform.*, vol. 22, no. 1, pp. 393–415, Jan. 2021, doi: 10.1093/bib/bbz170.
8. J. Jasril, M. F. Al Fiqri, S. Sanjaya, L. Handayani, and F. Insani, "Pengelompokan Data Kondisi Mesin Screw Press Menggunakan Algoritma Fuzzy C-Means," *Inf. Syst. J.*, vol. 8, no. 01, pp. 60–70, 2025, doi: 10.24076/infosjournal.2025v8i01.2133.
9. F. K. Rahman, J. S. Sanjaya, L. Handayani, and F. Insani, "Penerapan Algoritma K-Means Clustering pada Kinerja Mesin Screw press," *Bull. Inf. Technol.*, vol. 6, no. 2, pp. 59–70, 2025, doi: 10.47065/bit.v5i2.1783.
10. N. Sureja, B. Chawda, and A. Vasant, "An improved K-medoids clustering approach based on the crow search algorithm," *J. Comput. Math. Data Sci.*, vol. 3, no. July 2021, p. 100034, 2022, doi: 10.1016/j.jcmds.2022.100034.
11. A. Sobrinho Campolina Martins, L. Ramos de Araujo, and D. Rosana Ribeiro Penido, "K-Medoids clustering applications for high-dimensionality multiphase probabilistic power flow," *Int. J. Electr. Power Energy Syst.*, vol. 157, no. February, 2024, doi: 10.1016/j.ijepes.2024.109861.
12. Alfia Nurlaili Tahiyat, Bima Maulana, Ade Eka Saputra, Lusiana Efrizoni, and Rahmaddeni Rahmaddeni, "Klasterisasi Lagu Populer dan Eksplorasi Subgenre Spotify 2024 dengan K-Medoids," *J. Inform. Dan Tekonologi Komput.*, vol. 5, no. 1, pp. 34–48, 2025, doi: 10.55606/jitek.v5i1.5699.
13. J. Inayah, A. Fanani, and W. D. Utami, "Klasterisasi Data Kejadian Gempa Bumi di Indonesia Menggunakan Metode K-Medoids," *J. Sist. dan Teknol. Inf.*, vol. 12, no. 2, p. 271, 2024, doi: 10.26418/justin.v12i2.73594.
14. M. Minarni, M. Mahendra, A. Anisya, D. W. T. Putra, G. Y. Swara, and I. Warman, "Klasterisasi

Wilayah Rawan Kekerasan Anak Menggunakan Algoritma k-Medoids di Kota Padang," *J. Minfo Polgan*, vol. 13, no. 2, pp. 2309–2319, 2025, doi: 10.33395/jmp.v13i2.14447.

15. D. Sitanggang, M. Kom, and A. Apriori, *Buku Monograf Algoritma Apriori*. 2023.

16. S. E. Saqila, I. P. Ferina, and A. Iskandar, "Analisis Perbandingan Kinerja Clustering Data Mining Untuk Normalisasi Dataset," *J. Sist. Komput. dan Inform.*, vol. 5, no. 2, p. 356, 2023, doi: 10.30865/json.v5i2.6919.

17. E. Herman, K. E. Zsido, and V. Fenyves, "Cluster Analysis with K-Mean versus K-Medoid in Financial Performance Evaluation," *Appl. Sci.*, vol. 12, no. 16, 2022, doi: 10.3390/app12167985.

18. U. Linarti, A. Rahmawati, A. Hendri Soleliza Jones, L. Zahrotun, and U. Ahmad Dahlan, "Penerapan Metode K-Medoids Guna Pengelompokan Data Usaha Mikro, Kecil dan Menengah (UMKM) Bidang Kuliner Di Kota Yogyakarta," *J. Ilmu Komput. dan Sist. Inf.*, vol. 7, no. 1, pp. 37–45, 2024.

19. K. Dbscan and Y. Hasan, "Pengukuran Silhouette Score dan Davies-Bouldin Index pada Hasil Cluster K-Means dan DBSCAN," vol. 06, no. 01, pp. 60–74, 2024.

20. H. Henderi *et al.*, "Optimization of Davies-Bouldin Index with k-medoids algorithm," *AIP Conf. Proc.*, vol. 3065, no. 1, 2024, doi: 10.1063/5.0225220.

21. M. D. Kartikasari, "Self-Organizing Map Menggunakan Davies-Bouldin Index dalam Pengelompokan Wilayah Indonesia Berdasarkan Konsumsi Pangan," *Jambura J. Math.*, vol. 3, no. 2, pp. 187–196, 2021, doi: 10.34312/jjom.v3i2.10942.