JURNAL TEKNOLOGI DAN OPEN SOURCE
Vol. 8, No. 2, December 2025, pp. 1257~1268
e-ISSN: 2622-1659, accredited Four Grade by Kemenristekdikti, Decree No: 152/E/KPT/2023
DOI: https://doi.org/10.36378/jtos.v8i2.5124

Analysis Of The K-Means Method For Classifying The Quality Of Palm Oil Fruit Bundles

Satrio Utomo Widodo[1], Samsudin[2]
[1,2]Information System Program, State Islamic University of North Sumatra

Article Info

Article history:

Received 12 07, 2025
Revised 12 16, 2025
Accepted 12 27, 2025

Keywords:

Fresh Fruit Bunches (FFB),
Palm Oil,
Data Mining,
K-Means Clustering,
Single Linkage,
Web-Based Information System

ABSTRACT

This study discusses the application of K-Means Clustering and Single Linkage algorithms in classifying the quality of Fresh Fruit Bunches (FFB) of oil palm based on four main variables: ripeness level, oil content, bunch weight, and water content. The dataset consists of ten harvest records representing variations in FFB characteristics. The analysis process began with Min–Max normalization, followed by clustering using K-Means (k=3) and Single Linkage (Hierarchical). The results showed that K-Means produced a lower Sum of Squared Error (SSE) (215.32), indicating more compact and homogeneous clusters, while Single Linkage achieved a higher Silhouette Score (0.914) compared to K-Means (0.892), reflecting clearer separation between clusters. This implies that K-Means is more suitable when compactness is prioritized, whereas Single Linkage performs better when the goal is natural separation. The implementation was carried out through a web-based system at PT. Teguh Karsa Wanalestari 2 Pom, Siak Regency, Riau, to address the problem of manual and subjective FFB quality assessment. The system enables real-time data input, automated calculations, evaluation with SSE and Silhouette Score, and presents results in tables, graphs, accuracy reports, and Gmail notifications. With this system, FFB quality can be consistently grouped into three grades (A, B, C), supporting decision-making in harvesting, distribution, and processing more objectively and efficiently to maintain Crude Palm Oil (CPO) quality.

Corresponding Author:

Satrio Widodo Utomo
Information System Program,
State Islamic University of North Sumatra
North Sumatra, Indonesia
Email: satrioutomowidodo193@gmail.com

1. Introduction

The increasingly complex development of technology has made artificial intelligence and data mining crucial in improving competitiveness, efficiency, and productivity. Educational institutions and companies in Indonesia are required to be adaptive to rapid digital transformation [1]. Companies now not only focus on operational efficiency, but must also be responsive to rapid technological changes . In the palm oil industry, data mining with cluster analysis is used to group the quality of Fresh Fruit Bunches (FFB) that are influenced by environmental factors such as monthly rainfall [2]. The goal is to produce high-quality FFB with maximum oil yield, low Free Fatty Acid (FFA) content, and efficient production costs [3].

PT. Teguh Karsa Wanalestari 2 Pom, located in Belutu Village, Kandis District, Siak Regency, Riau Province, is a research site engaged in the processing of oil palm fruit, often referred to as Fresh Fruit Bunches (FFB). To date, it still faces obstacles in the quality classification system for oil palm Fresh Fruit Bunches (FFB). Quality assessment of FFB at this company is still done manually based on harvest records without standardized quality measurement parameters. As a result, the assessment process is subjective, inconsistent, and increases the risk of errors in determining fruit quality, which impacts the potential for reduced oil yield, decreased CPO quality, failure in optimal FFB harvest and distribution planning, and decreased FFB and CPO selling prices in the market due to unmaintained quality. The problem of grading the quality of fresh fruit bunches (FFB) at PT. Teguh Karsa Wanalestari 2 Pom becomes increasingly complex during the dry season (July–September), when water shortages cause low moisture content, uneven ripeness, and many fruits drying before ripening. These quality variations make manual assessment difficult due to the lack of objective standards, resulting in inconsistent classification, risk of errors, and reduced production yields, as well as operational losses. Therefore, a more accurate, objective, and efficient technology-based system is needed to assess FFB quality. From the above problems, the verse that corresponds to this discussion is QS. Yusuf: 47.

Yusuf said, "*You will cultivate for seven years (consecutively) as usual. Then whatever you harvest, leave it (store it) on the branches, except for a little for you to eat*." As explained in QS. Yusuf verse 47, Prophet Yusuf emphasized the importance of careful management of harvests in order to face difficult times. This is relevant to the research conditions, where the quality of fresh fruit bunches (FFB) of oil palm tends to decline during the dry season, thus requiring an appropriate grouping system to support efficient and sustainable harvest management strategies. This study aims to develop a website-based system capable of processing FFB quality data in real-time, with integrated input from all oil palm land units of PT. Teguh Karsa Wanalestari 2 Pom. The data used includes ripeness level, oil content, bunch weight, and moisture content, which are then objectively grouped using the K-Means Clustering algorithm. This clustering aims to separate good and bad FFB quality to support decision-making in the harvesting, distribution, and processing processes so that oil yield is optimal and CPO quality is maintained. This system is equipped with graphical and tabular visualizations that help evaluate FFB quality systematically, accurately, and more consistently than manual methods, which are often subjective and time-consuming.

A significant difference compared to previous studies that used the K-Means method. The study by Wahyuni Eka Sari entitled "Application of the K-Means Algorithm to Group the Maturity Levels of Oil Palm Fresh Fruit Bunches Based on Color" only used the color variable of fresh fruit bunches with 40 test data and produced an accuracy of 90% [4]. The study by Sibarani entitled "Comparison of K-Means and Hierarchical Clustering in Medication Stock Clustering" used medication stock data, where Single Linkage was superior (Silhouette Score 0.976) compared to K-Means (0.954) with an average time of 4 seconds and an accuracy of 90% [5]. Meanwhile, unlike previous studies, this study not only applied the K-Means algorithm for grouping the quality of palm oil fresh fruit bunches, but also expanded the data variables used, including fruit skin color, oil content, bunch weight, and moisture content. Additionally, this study applied two evaluation models, namely SSE (Sum of Squared Error) and Silhouette Score, to assess the accuracy level of the clustering results. The grouping results are displayed in the form of an interactive website that presents cluster graphs, analysis result tables, and accuracy comparison reports. The system is also equipped with automatic report notifications via Gmail to PT. Teguh Karsa Wanalestari 2 Pom.

2. Research Method

2.1. Research methods

This study employs a mixed-methods embedded design, which is a type of mixed-method approach that integrates qualitative and quantitative methodologies within a single research framework [6]. The embedded design is characterized by the deliberate combination of two different types of data, where one method plays a dominant role and the other is embedded to support, enrich, and clarify the primary findings. In this study, one approach is positioned as the main focus of analysis, while the complementary approach is used to strengthen interpretation, provide contextual depth, and enhance the validity of the results obtained [7]. By integrating these two approaches, the research is able to capture both numerical trends and contextual insights, allowing for a more holistic and comprehensive understanding of the phenomenon under investigation.

The use of a mixed-methods embedded design is particularly relevant for studies that aim to examine complex processes or systems, as it enables researchers to explore not only measurable outcomes but also the underlying conditions, perceptions, and operational practices that influence those outcomes. Through this approach, quantitative data provide objective evidence and patterns, while qualitative data help explain the reasons behind those patterns, thereby reducing potential bias and increasing the robustness of the conclusions. As a result, this design contributes to a more nuanced and reliable interpretation of research findings.

Third, secondary quantitative data in the form of TBS quality records for the period January–August 2025 were utilized as the primary quantitative data source. This dataset provided measurable indicators to support statistical analysis and trend evaluation. The integration of these three data collection techniques ensures that the findings are well-supported and systematically analyzed. The overall research stages and methodological flow are illustrated in the following figure.

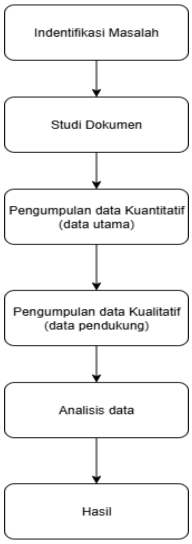

Figure 1. Research Stages

Research Stages

1. Problem Identification

PT. Teguh Karsa Wanalestari 2 Pom has difficulties in classifying the quality of palm oil fresh fruit bunches (FFB) because the assessment process is still manual and subjective, thereby affecting the quality of crude palm oil (CPO), oil yield, and selling price.

2. Document Study

Previous studies such as Implementation of the K-Means Algorithm for Clustering Oil Palm Productivity Data [8] were used as references to understand the application of clustering methods in assessing agricultural product quality.

3. Quantitative Data Collection (Main Data)

TBS harvest data from January to August 2025 was used, including variables such as maturity level, oil content, bunch weight, and moisture content. The data was analyzed using the K-Means Clustering algorithm to objectively group TBS quality.

4. Qualitative Data Collection (Supporting Data)

Field observations and interviews with plantation managers and sorting officers were conducted to understand the harvesting process, weather factors, and constraints in manual TBS quality assessment.

5. Data Analysis

The K-Means algorithm was used to group FFB into three grades: A (good), B (fair), and C (poor). The results were evaluated using SSE (Sum of Squared Error) and Silhouette Score to measure accuracy and consistency.

6. Research Results

The research produced a web-based information system capable of:

a. Grouping TBS quality in real-time using K-Means.
b. Displaying results in graphs and tables.
c. Providing harvesting and sorting recommendations to maintain CPO quality and operational efficiency.

The methods used to collect data in this study are described as follows:

1. Observation

Field observations were conducted at PT. Teguh Karsa Wanalestari 2 Pom, Belutu Village, Siak Regency, to understand the harvesting process and quality assessment of Fresh Fruit Bunches (FFB). The observations showed that quality assessment was still carried out manually by sorting officers without objective standards, relying only on experience and visual inspection. Harvest data recording also still used simple methods that were unable to provide automatic analysis or predict FFB quality. This situation causes inconsistent FFB grouping and has the potential to reduce oil yield and CPO quality. These observation findings emphasize the need for a web-based information system capable of analyzing FFB quality data in real-time, including considering weather and seasonal factors, to support more accurate and efficient harvesting and sorting decisions.

2. Interview

Interviews were conducted with plantation managers, sorting officers, and operational staff of PT. Teguh Karsa Wanalestari 2 Pom to gather information related to the harvesting process and quality assessment of Fresh Fruit Bunches (FFB). From the interviews, it was found that quality assessment is still subjective and highly dependent on the experience of the officers, without any standard measurement criteria. Respondents also said that variations in FFB quality often occur due to changes in weather, dry seasons, high rainfall, and other environmental factors that affect moisture content, ripeness, and oil yield. These conditions make manual management unable to maintain consistent FFB quality. These interview findings form an important basis for the design of a web-based system that integrates the K-Means algorithm to group TBS quality automatically and objectively. The system is expected to improve the accuracy of assessments, facilitate harvesting and sorting decisions, and maintain TBS quality and oil yield more consistently and efficiently.

3. Collection of Harvest History Data

Quantitative data collection was carried out to obtain measurable information about FFB quality based on numerical values that could be analyzed statistically. In this study, quantitative data was obtained from historical FFB harvest data from PT. Teguh Karsa Wanalestari 2 Pom for the period January–August 2025, covering the variables of maturity level, oil content, bunch weight, and moisture content. All data was then processed using the K-Means Clustering algorithm to produce an objective and systematic grouping of FFB quality. The results of this analysis provide accurate, structured, and accountable information, thereby supporting a more efficient decision-making process related to harvesting, sorting, and processing FFB.

2.2. Research Methodology

The Waterfall model is a system development method that runs in stages and sequentially, starting with requirements analysis, followed by design, then coding, and finally testing. This approach is known as one of the earliest software development methodologies and has a highly systematic work structure in the information technology industry [9].

This study uses a modified Waterfall model-based system development method tailored to the requirements. Waterfall was chosen because it has a structured workflow, starting from analysis, design, implementation, to testing. However, this study only focused on system design and testing, so the maintenance stage was not applied. This modification was made so that the development model was more in line with the research objectives, namely to build and test the system without a post-implementation management process [10].

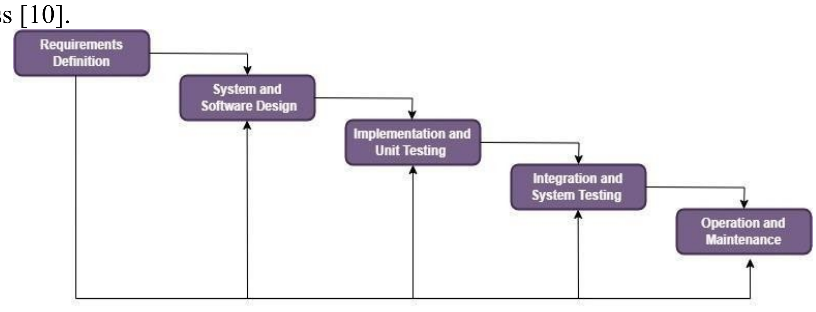

Figure 2. Waterfall Development Method

From the development method diagram above, there are five stages carried out by the author, including:

1. Requirement Definition

Requirements are the initial stage in the Waterfall method, which focuses on analyzing the requirements of the system to be developed so that all requirements can be clearly understood by users [11]. In this study, system requirements were obtained through direct observation at PT. Teguh Karsa Wanalestari 2 Pom, which included harvesting, sorting, and recording the quality of fresh fruit bunches (FFB). The results of the observation showed the need for a system that can manage FFB quality data in real-time, group FFB based on maturity, oil content, bunch weight, and moisture content, and present data visualization to support decision making.

2. System Software Design

At the System Software Design stage, the design focuses on the technical aspects of the system, including hardware management, software, and user interface design. In this study, the system design is described through Unified Modeling Language (UML) diagrams, such as use case diagrams, activity diagrams, and class diagrams, which show the flow of user interactions with the system, the TBS quality grouping process, and the relationships between database entities. This design aims to produce an information system that meets the needs of PT. Teguh Karsa Wanalestari 2 Pom, namely, one that is capable of classifying TBS accurately and consistently, supporting TBS quality evaluation, and assisting in the planning of CPO harvesting and distribution with optimal yields.

3. Implementation

The implementation stage includes coding the system using HTML, PHP, and CSS, as well as creating a database with MySQL [12]. The author used Visual Studio Code as a development tool and XAMPP as a local server.

After the system design was completed, the TBS grouping logic was integrated using the K-Means algorithm, which analyzed historical harvest data, such as maturity level, oil content, bunch weight, and moisture content, to form grade A, B, and C quality categories. This implementation resulted in a web-based system capable of providing real-time TBS grouping, supporting more efficient sorting decisions, harvest planning, and TBS distribution.

4. Integration and Tasting

At this stage, the author conducts testing on the developed program. Testing is carried out to ensure that every function in the system runs without errors and produces outputs that meet the previously defined requirements[13]. At this stage, the author applied the blackbox testing method to test the system. This test focused on examining the functionality of the system based on the specified specifications, without the need to view or analyze the program source code. The test was conducted by observing whether each feature in the system worked according to user expectations and previously defined requirements [14]. The main objective of this method is to ensure that the system is able to respond to inputs correctly and produce appropriate outputs, as well as to identify any errors in the application's functionality [14].

5. Opration and maintenance

This stage covers possible maintenance of the system after implementation. In this study, maintenance is not the main focus, and will only be carried out if there is a follow-up agreement with PT.Teguh Karsa Wanalestari 2 Pom regarding post-implementation system management. This phase is optional, depending on the needs and cooperation between the researchers and the company.

3. Result and Discussion

3.1. Collection of Palm Oil Fruit Bunch Data

The data for this study was obtained from records of fruit bunch harvests from January to August 2025, which contained information on the maturity level of fresh fruit bunches (on a scale of 1 to 5), oil content (%), bunch weight (kg), and water content (%). The following are the results of the fruit bunch data collection:

Table 1. Results Of Tbs Data Collection

Tanggal	Tingkat Kematangan	Kadar Minyak (%)	Berat TBS (kg)	Kadar Air (%)
Wednesday, 01 January 2025	Cukup matang (50% luar, 30% tengah, 15% dalam)	18,91	19,88	20,06
Monday, 06 January 2025	Sangat matang (65% luar, 20% tengah, 10% dalam)	20,88	11,32	18,71
Saturday, 11 January 2025	Kurang matang (45% luar, 28% tengah, 20% dalam)	20,55	10,11	18,38
Thursday, 16 January 2025	Matang (57% luar, 25% tengah, 16% dalam)	23,78	20,03	17,13

Tanggal	Tingkat Kematangan	Kadar Minyak (%)	Berat TBS (kg)	Kadar Air (%)
Tuesday, 21 January 2025	Matang (57% luar, 25% tengah, 16% dalam)	23,52	16,89	22,7
Sunday, 26 January 2025	Matang (57% luar, 25% tengah, 16% dalam)	20,12	24,09	24,31
Friday, 31 January 2025	Kurang matang (45% luar, 28% tengah, 20% dalam)	20,96	16,76	21,23
Wednesday, 05 February 2025	Matang optimal (60% luar, 23% tengah, 15% dalam)	23,08	16,2	20,02
Monday, 10 February 2025	Cukup matang (50% luar, 30% tengah, 15% dalam)	18,91	19,88	20,06
Saturday, 15 February 2025	Cukup matang (50% luar, 30% tengah, 15% dalam)	23,09	11,32	26,83
Saturday, 13 September 2025	Kurang matang (45% luar, 28% tengah, 20% dalam)	22,79	16,04	27,09
Thursday, 18 September 2025	Sangat matang (65% luar, 20% tengah, 10% dalam)	21,15	23,61	24,37
Tuesday, 23 September 2025	Kurang matang (45% luar, 28% tengah, 20% dalam)	18,7	14,78	26,01
Sunday, 28 September 2025	Kurang matang (45% luar, 28% tengah, 20% dalam)	24,46	18,27	16,21
Friday, 03 October 2025	Matang optimal (60% luar, 23% tengah, 15% dalam)	23,08	16,2	20,02
Wednesday, 08 October 2025	Cukup matang (50% luar, 30% tengah, 15% dalam)	22,49	15,99	25,01
Monday, 13 October 2025	Cukup matang (50% luar, 30% tengah, 15% dalam)	18,91	19,88	20,06
Saturday, 18 October 2025	Cukup matang (50% luar, 30% tengah, 15% dalam)	20,51	21,16	24,74
Thursday, 23 October 2025	Matang optimal (60% luar, 23% tengah, 15% dalam)	19,93	11,75	26,25
Tuesday, 28 October 2025	Cukup matang (50% luar, 30% tengah, 15% dalam)	18,9	24,75	20,99

3.2. Data Transformation Maturity Level

The maturity level must be converted into numbers so that it can be calculated using K-Means. The following is a conversion table for maturity levels :

Table 2. Maturity Level

Tingkat Kematangan	Skor
Kurang matang	1
Cukup matang	2
Matang	3
Matang optimal	4
Sangat matang	5

3.3. Maturity Level Conversion Data

At this stage, the Fresh Fruit Bunch (FFB) maturity level category data, which was previously in qualitative form, was converted into quantitative values. This conversion process was carried out because the K-Means algorithm can only process numerical data in Euclidean distance calculations and centroid formation.

Table 3. Maturity Level Conversion Data

No	Kematangan	Skor	Kadar Minyak	Berat	Kadar Air

No	Kematangan	Skor	Kadar Minyak	Berat	Kadar Air
1	Cukup matang	2	18.91	19.88	20.06
2	Sangat matang	5	20.88	11.32	18.71
3	Kurang matang	1	20.55	10.11	18.38
4	Matang	3	23.78	20.03	17.13
5	Matang	3	23.52	16.89	22.70
6	Matang	3	20.12	24.09	24.31
7	Kurang matang	1	20.96	16.76	21.23
8	Matang optimal	4	23.08	16.20	20.02
9	Cukup matang	2	18.91	19.88	20.06
10	Cukup matang	2	23.09	11.32	26.83
11	Kurang matang	1	22.79	16.04	27.09
12	Sangat matang	5	21.15	23.61	24.37
13	Kurang matang	1	18.70	14.78	26.01
14	Kurang matang	1	24.46	18.27	16.21
15	Matang optimal	4	23.08	16.20	20.02
16	Cukup matang	2	22.49	15.99	25.01
17	Cukup matang	2	18.91	19.88	20.06
18	Cukup matang	2	20.51	21.16	24.74
19	Matang optimal	4	19.93	11.75	26.25
20	Cukup matang	2	18.90	24.75	20.99

3.4. Calculation of Initial Distance (Initial Centroid Distance) in K-Means.

This stage aims to calculate the distance of each data point from the initial centroid so that we can determine which cluster each data point belongs to (e.g., Cluster 1 or Cluster 2).
The formula is as follows:

$$D = \sqrt{(x_1 - c_1)^2 + (x_2 - c_2)^2 + (x_3 - c_3)^2 + (x_4 - c_4)^2} \quad (1)$$

Description:
x = data
c = centroid
Variables: Score, Oil Content, Weight, Water Content

Table 4. Initial Centroid Distance Calculation Results in K-Means.

No	Skor	Minyak	Berat	Air	D ke C1	D ke C2	D ke C3	Cluster
1	0.2500	0.0365	0.6673	0.3539	0.0000	1.0182	0.7828	C1
2	1.0000	0.3785	0.0827	0.2298	1.0182	0.0000	1.0055	C2
3	0.0000	0.3212	0.0000	0.1994	0.7828	1.0055	0.0000	C3
4	0.5000	0.8819	0.6776	0.0846	0.9219	0.9373	1.0182	C1
5	0.5000	0.8368	0.4631	0.5965	0.8965	0.8598	0.9423	C2
6	0.5000	0.2465	0.9549	0.7445	0.5847	1.1372	1.2102	C1
7	0.0000	0.3924	0.4542	0.4614	0.4961	1.0917	0.5292	C1
8	0.7500	0.7604	0.4160	0.3502	0.9150	0.5779	0.9753	C2
9	0.2500	0.0365	0.6673	0.3539	0.0000	1.0182	0.7828	C1
10	0.2500	0.7622	0.0827	0.9761	1.1206	1.1255	0.9311	C3
11	0.0000	0.7101	0.4051	1.0000	1.0013	1.3443	0.9778	C3
12	1.0000	0.4253	0.9221	0.7500	0.9673	0.9887	1.4712	C1
13	0.0000	0.0000	0.3190	0.9007	0.6959	1.2842	0.8347	C1
14	0.0000	1.0000	0.5574	0.0000	1.0622	1.2901	0.9007	C3
15	0.7500	0.7604	0.4160	0.3502	0.9150	0.5779	0.9753	C2
16	0.2500	0.6580	0.4016	0.8088	0.8148	1.0381	0.8418	C1
17	0.2500	0.0365	0.6673	0.3539	0.0000	1.0182	0.7828	C1
18	0.2500	0.3142	0.7548	0.7840	0.5195	1.1513	0.9869	C1
19	0.7500	0.2135	0.1120	0.9228	0.9557	0.7555	1.0535	C2
20	0.2500	0.0347	1.0000	0.4393	0.3435	1.2514	1.0964	C1

Recap of cluster distribution (Iteration 1)

1. Cluster C1: No. 1, 4, 6, 7, 9, 12, 13, 16, 17, 18, 20
2. Cluster C2: No. 2, 5, 8, 15, 19
3. Cluster C3: No. 3, 10, 11, 14

Table 5. Recap Of Cluster Distribution (Iteration 1)

No	Skor	Minyak	Berat	Air	D ke C1	D ke C2	D ke C3	Cluster
1	0.25	0.0365	0.6673	0.3539	0.3182	0.8433	0.8211	C1
2	1.00	0.3785	0.0827	0.2298	0.9664	0.4703	1.0544	C2
3	0.00	0.3212	0.0000	0.1994	0.8296	0.8988	0.5772	C3
4	0.50	0.8819	0.6776	0.0846	0.7827	0.6754	0.7806	C2
5	0.50	0.8368	0.4631	0.5965	0.6280	0.4026	0.5040	C2
6	0.50	0.2465	0.9549	0.7445	0.3843	0.8227	0.9576	C1
7	0.00	0.3924	0.4542	0.4614	0.4160	0.7917	0.3762	C3
8	0.75	0.7604	0.4160	0.3502	0.7267	0.2500	0.7334	C2
9	0.25	0.0365	0.6673	0.3539	0.3182	0.8433	0.8211	C1
10	0.25	0.7622	0.0827	0.9761	0.8825	0.7499	0.5079	C3
11	0.00	0.7101	0.4051	1.0000	0.7531	0.9212	0.4825	C3
12	1.00	0.4253	0.9221	0.7500	0.7651	0.7395	1.1969	C2
13	0.00	0.0000	0.3190	0.9007	0.6583	1.0391	0.7889	C1
14	0.00	1.0000	0.5574	0.0000	0.9686	1.0188	0.6916	C3
15	0.75	0.7604	0.4160	0.3502	0.7267	0.2500	0.7334	C2
16	0.25	0.6580	0.4016	0.8088	0.5424	0.6059	0.3559	C3
17	0.25	0.0365	0.6673	0.3539	0.3182	0.8433	0.8211	C1
18	0.25	0.3142	0.7548	0.7840	0.2585	0.7882	0.6957	C1
19	0.75	0.2135	0.1120	0.9228	0.8087	0.6031	0.9347	C2
20	0.25	0.0347	1.0000	0.4393	0.4220	1.0265	1.0160	C1

Rekap Clustre (Iterasi 2)

a. Cluster C1: 1, 6, 9, 13, 17, 18, 20 (7 anggota)
b. Cluster C2: 2, 4, 5, 8, 12, 15, 19 (7 anggota)
c. Cluster C3: 3, 7, 10, 11, 14, 16 (6 anggota)

3.5. Determination of Final Grades A–B–C Based on K-Means Clustering Results

After the clustering process was carried out until stable iteration, three final clusters were formed that represented the quality of the fresh fruit bunches. Cluster 2 had the highest average scores for ripeness, oil content, and bunch weight, and was therefore designated as Grade A. Cluster 1 was in the medium range and was classified as Grade B. Meanwhile, Cluster 3 had the lowest average scores for the variables analyzed and was designated as Grade C. This grading system provides a basis for objective and consistent assessment of TBS quality.

Table 6. Final A–B–C Grade Determination Based on K-Means Clustering Results

No	Skor (Normalisasi)	Kadar Minyak	Berat	Kadar Air	Cluster	Grade
1	0.25	0.0365	0.6673	0.3539	C1	B
2	1.00	0.3785	0.0827	0.2298	C2	A
3	0.00	0.3212	0.0000	0.1994	C3	C
4	0.50	0.8819	0.6776	0.0846	C2	A
5	0.50	0.8368	0.4631	0.5965	C2	A
6	0.50	0.2465	0.9549	0.7445	C1	B
7	0.00	0.3924	0.4542	0.4614	C3	C
8	0.75	0.7604	0.4160	0.3502	C2	A
9	0.25	0.0365	0.6673	0.3539	C1	B
10	0.25	0.7622	0.0827	0.9761	C3	C
11	0.00	0.7101	0.4051	1.0000	C3	C
12	1.00	0.4253	0.9221	0.7500	C2	A
13	0.00	0.0000	0.3190	0.9007	C1	B

14	0.00	1.0000	0.5574	0.0000	C3	C
15	0.75	0.7604	0.4160	0.3502	C2	A
16	0.25	0.6580	0.4016	0.8088	C3	C
17	0.25	0.0365	0.6673	0.3539	C1	B
18	0.25	0.3142	0.7548	0.7840	C1	B
19	0.75	0.2135	0.1120	0.9228	C2	A
20	0.25	0.0347	1.0000	0.4393	C1	B

Rekap Grade Final

a. Grade A (Cluster C2) 2, 4, 5, 8, 12, 15, 19 (8 anggota)
b. Grade B (Cluster C1) 1, 6, 9, 13, 17, 18, 20 (7 anggota)
c. Grade C (Cluster C3) 3, 7, 10, 11, 14, 16 (6 anggota)

3.6. Clustering Results Evaluation

Based on the clustering process using the K-Means algorithm on 20 samples of Fresh Fruit Bunch (FFB) quality data, three main clusters were obtained, representing Grades A, B, and C. The centroid results show that Cluster C2 has the highest maturity score (0.75), high oil content (0.6081), moderate weight (0.4414), and stable moisture content (0.4692), thus it is designated as Grade A. Meanwhile, Cluster C1 had a moderate ripeness score (0.25), low oil content (0.1007), high weight (0.7187), and fairly high moisture content (0.5615), so it was represented as Grade B. Cluster C3 has a low maturity score (0.0833), the lowest weight (0.3168), and high oil content (0.6407), but unstable water content (0.5743), so it is classified as Grade C.

The model quality evaluation shows a Total SSE value of 4.114, which indicates that the clusters are quite compact, and a Silhouette Score of 0.2684, indicating that the separation between clusters is quite good for a small dataset. Overall, these results prove that the K-Means method successfully grouped TBS into three quality grades based on a combination of four main variables: maturity score, oil content, weight, and moisture content.

```
SSE per cluster:
C1: 0.864889
C2: 1.725654
C3: 1.523732
Total SSE: 4.114275

Silhouette Score (overall): 0.2684
```

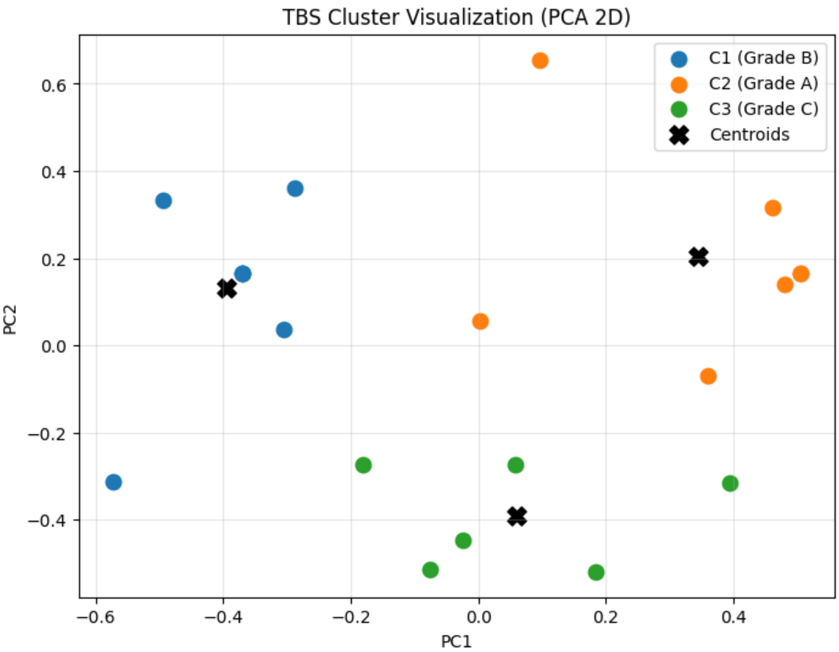

Figure 3. TBS Cluster Visualization

3.7. System Implementation

System implementation is the stage of realizing the previously made design into a web form that can be used by users. The main process in the system is the grouping of FFB quality using the K-Means and Single Linkage methods. FFB data, which includes oil content, water content, bunch weight, and harvest age, is input by users through a data entry form.

1. Calculation Display

The Calculation Display is a page that displays the weight of the assessment criteria (maturity, oil content , bunch weight, and moisture content) as well as the previously inputted TBS quality data. On this page ,users can view the original data table and then press the Perform Calculation button to run the normalization process and generate a final score according to the specified weight.

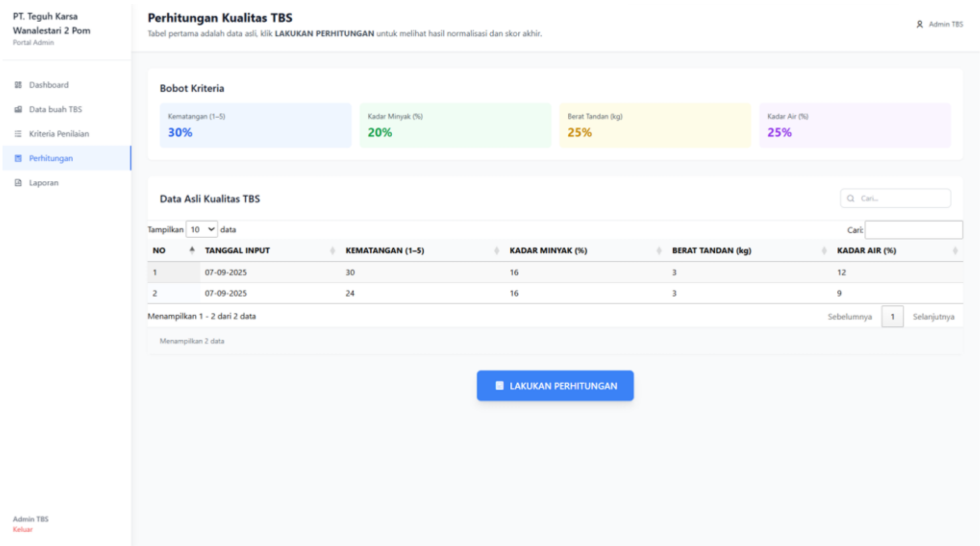

Figure 4. Calculation Display

2. Calculation Results Display

The Calculation Results Display is a page that shows the output after the TBS quality calculation process is complete. The top section still displays the weight of each criterion, followed by the original TBS quality data table. After the calculation is complete, the system displays the TBS Quality Calculation Results table containing the ranking, input date, value of each criterion, final score, and grade for each data point.

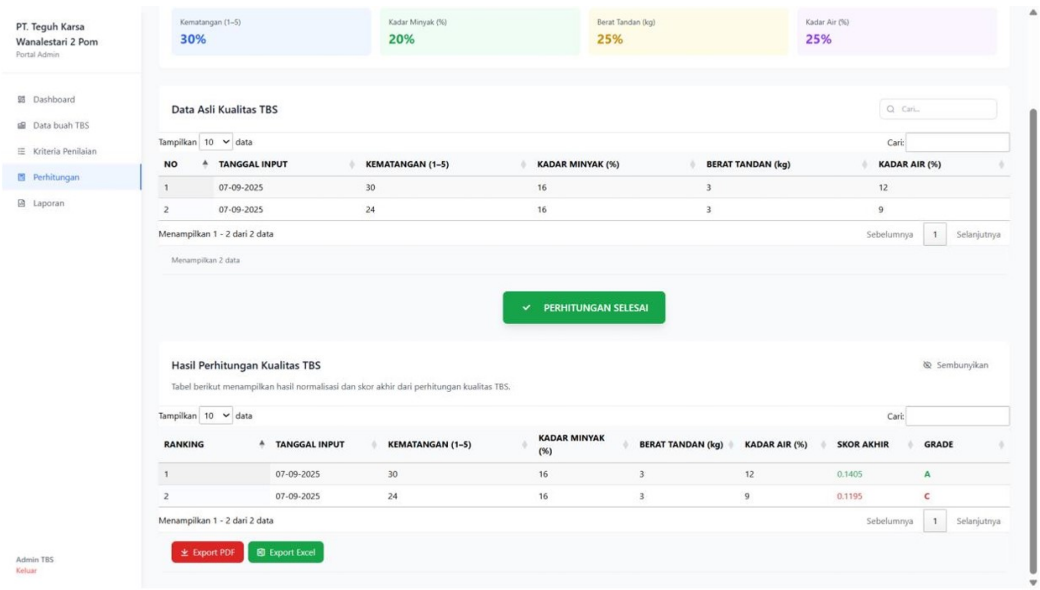

Figure 5. Calculation Results Display

4. Conclusion

Based on the clustering results using the K-Means algorithm, three stable final clusters were obtained in the second iteration. Cluster C2 showed the highest average values, namely a maturity score of 0.71, oil content of 0.63, bunch weight of 0.58, and moisture content of 0.53, and was therefore designated as Grade A, with data members numbers 2, 4, 5, 8, 12, 15, and 19. Meanwhile, cluster C1 has a medium average value with a maturity score of 0.25, oil content of 0.19, bunch weight of 0.72, and water content of 0.50, so it is categorized as Grade B, with data members 1, 6, 9, 13, 17, 18, and 20. Cluster C3 shows the lowest average values, namely a maturity score of 0.08, oil content of 0.64, bunch weight of 0.30, and moisture content of 0.73, thus being designated as Grade C with members 3, 7, 10, 11, 14, and 16. These results show that K-Means can objectively cluster FFB based on the proximity of characteristic values, so that the FFB quality determination process can be carried out more consistently, measurably, and free from subjectivity.

References

[1]. S. Suharyo, S. Subyantoro, and R. Pristiwati, "Kecerdasan buatan dalam konteks kurikulum merdeka pada jenjang pendidikan dasar dan menengah: Membangun keterampilan menuju Indonesia emas 2045," *Humanika*, vol. 30, no. 2, pp. 208–217, 2024.

[2]. A. Al Masykur, "Penerapan Metode K-Means Clustering untuk Pemetaan Pengelompokan Lahan Produksi Tandan Buah Segar," *Penerapan Metod. K-Means Clust. untuk Pemetaan Pengelompokan Lahan Produksi Tandan Buah Segar*, vol. 10, no. 1, pp. 92–100, 2023.

[3]. R. A. Sirait and G. Supriyanto, "Pengaruh Kematangan Buah Terhadap FFA dan Besarnya Kandungan Minyak di Dalamnya di Pabrik Kelapa Sawit," vol. 1, no. Gapksi 2022, pp. 676–684, 2023.

[4]. A. K-means, W. E. Sari, A. Franz, and P. Sugiartawan, "Deteksi Tingkat Kematangan Tandan Buah Segar Kelapa Sawit Dengan," vol. 5, no. 2, pp. 154–164, 2022.

[5]. M. Sibarani, G. Susrama, and S. Sugiarto, "PENGGUNAAN K-MEANS DAN HIERARCHICAL CLUSTERING SINGLE LINKAGE DALAM PENGELOMPOKKAN STOK OBAT," *J. Lebesgue J. Ilm. Pendidik. Mat. Mat. dan Stat.*, vol. 5, pp. 1286–1294, Aug. 2024, doi: 10.46306/lb.v5i2.715.

[6]. R. Justan, M. Margiono, A. Aziz, and S. Sumiati, "Penelitian Kombinasi (Mixed Methods)," *ULIL ALBAB J. Ilm. Multidisiplin*, vol. 3, pp. 253–263, Jan. 2024, doi: 10.56799/jim.v3i2.2772.

[7]. M. Dzulhijj and M. Albina, "Model Penelitian Campuran (Mixed Method)," *J. Arjuna Publ. Ilmu Pendidikan, Bhs. dan Mat.*, vol. 3, pp. 80–91, Jun. 2025, doi: 10.61132/arjuna.v3i4.2187.

[8]. Y. A. Sandria, M. R. A. Nurhayoto, L. Ramadhani, R. S. Harefa, and A. Syahputra, "Penerapan algoritma selection sort untuk melakukan pengurutan data dalam bahasa pemrograman PHP," *Hello World J. Ilmu Komput.*, vol. 1, no. 4, pp. 190–194, 2023.

[9]. W. Kurniawan and R. Kurniawan, "PENERAPAN ALGORITMA K-MEANS CLUSTERING DALAM MENENTUKAN PELUANG MASUK SISWA KE UNIVERSITAS NEGERI," *J. Inform. Teknol. dan Sains*, vol. 7, no. 1, pp. 386–393, 2025.

[10]. N. E. Christine, B. Priyatna, A. L. Hananto, and S. S. Hilabi, "Penerapan Metode Agile Scrum Pada Sistem E-Posyandu Berbasis Web," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 2, pp. 2013–2019, 2024.

[11]. F. Mahmudah and S. Suendri, "Perancangan Sistem HRM Pada PT. Pelindo Terminal Petikemas Menggunakan Metode ERP Berbasis Web," *Jutisi J. Ilm. Tek. Inform. dan Sist. Inf.*, vol. 13, no. 3, 2025.

[12]. R. V. Alif, M. R. Alhafiz, and M. Fakhriza, "Perancangan Sistem Informasi Dokumen P44 (Hasil Putusan) Berbasis Web Studi Kasus Kejaksaan Negeri Deli Serdang," *J. Informatics Busisnes*, vol. 1, no. 4, pp. 233–238, 2024.

[13]. D. K. Maulana and S. Samsudin, "Perancangan Sistem Informasi Presensi Non Pegawai Berbasis Mobile Menggunakan QR Code di PT Perkebunan Nusantara IV Medan," *J. Ilm. Sains dan Teknol.*, vol. 8, no. 2, pp. 209–219, 2024.

[14]. E. Novalia and A. Voutama, "Black Box Testing dengan Teknik Equivalence Partitions Pada Aplikasi Android M-Magazine Mading Sekolah," *Syntax J. Inform.*, vol. 11, pp. 23–35, Jun. 2022, doi: 10.35706/syji.v11i01.6413.

[15]. E. Oktafanda, N. W. Al-Hafiz, A. Latif, and F. Santosa, "Analysis and Design of Monolithic System Architecture Migration to Microservices at PT. MALINDO Conceptual Approach", *JTOS*, vol. 8, no. 1, pp. 54 - 63, May 2025.

[16]. R. Hardiartama, A. A. Arifiyanti, and Ana Wati3S. F., "Application of Ensemble Machine Learning Methods for Aspect-Based Sentiment Analysis on User Reviews of the Wondr by BNI App", *JTOS*, vol. 8, no. 1, pp. 97 - 111, Jun. 2025.

[17]. Mirza Khazim Nugraha, Nur Cahyo Wibowo, and Iqbal Ramadhani Mukhlis, "Design and Development of the Boarding House Management Information System (SIMKO) Using Laravel with Agile Methodology", *JTOS*, vol. 8, no. 1, pp. 112 - 122, Jun. 2025.

[18]. M. A. Sifaul Anam, A. Faroqi, A. Faroqi, and T. Lathif Mardi Suryanto, "Analysis of MyPertamina Application Acceptance Using a Modified Technology Acceptance Model (TAM)", *JTOS*, vol. 8, no. 1, pp. 133 - 141, Jun. 2025.

[19]. D. Handoko Putra and N. Rahditiantoro, "Implementation of Forward Chaining in a Web-Based Expert System for Passport Services", *JTOS*, vol. 8, no. 1, pp. 217 - 277, Jun. 2025.

[20]. R. A. P. Alisia, Anita Wulansari, and Rafika Rahmawati, "Analysis of Factors Influencing the Acceptance of the Indodax Application Using the UTAUT 2 Model", *JTOS*, vol. 8, no. 1, pp. 317 - 327, Jun. 2025.

[21]. H. Nopriandi, N. W. Al-Hafiz, and S. Chairani, "Analysis and Modeling of the Internal Quality Audit Information System Islamic University of Kuantan Singingi", *JTOS*, vol. 8, no. 1, pp. 398 - 408, Jun. 2025.

[22]. M. Yusfahmi, Febri Haswan, Nofri Wandi Al-Hafiz, Elgamar Syam, Helpi Nopriandi, Jasri, Aprizal, Harianja, Erlinda, Sri Chairani, Gunardi Hamzah, & Morine Delya Octa. (2025). SOSIALISASI DAN PENERAPAN APLIKASI BERBASIS TEKNOLOGI INFORMASI UNTUK MENDUKUNG TRANSFORMASI DIGITAL BUMDes TEBING TINGGI. BHAKTI NAGORI (Jurnal Pengabdian Kepada Masyarakat), 5(2), 712 - 719. https://doi.org/10.36378/bhakti_nagori.v5i2.4910

[23]. R. Nazli, A. Amrizal, H. Hendra, and S. Syukriadi, "Modeling User Interface Design E-Business Applications for Marketing Umkm Products in Payakumbuh City Using Pieces Framework", *JTOS*, vol. 7, no. 2, pp. 55 - 66, Nov. 2024.

[24]. M. Syamsul Azis and H. Basri, "Building A Powerfull File Specification Application For Database System Design", *JTOS*, vol. 7, no. 2, pp. 103 - 109, Dec. 2024.

[25]. Nofri Wandi Al-Hafiz, Helpi Nopriandi, and Harianja, "Design of Rainfall Intensity Measuring Instrument Using IoT-Based Microcontroller", *JTOS*, vol. 7, no. 2, pp. 202 - 211, Dec. 2024.