International Journal of Management Science and Information Technology
IJMSIT

E-ISSN: 2774-5694 P-ISSN: 2776-7388

Volume 6 (1), January-June 2026, 366-376
DOI: https://doi.org/10.35870/ijmsit.v6i1.6902

The Use of the K-Means Algorithm as a Method for Grouping Major Interests of Class X Students of SMK Satrya Budi 1 Commerce

Handina Ananda Putri [1*], Herman Saputra [2], Yori Apridonal M [3]
[1*,2,3] Information Systems Study Program, Faculty of Computer Science, Universitas Royal, Asahan Regency, North Sumatra, Indonesia

Email: handinaanandaputri07@gmail.com [1*], hermansaputra4@gmail.com [2], yori.apridonal@gmail.com [3]

Article history:
Received April 6, 2026
Revised April 16, 2026
Accepted April 18, 2026

Abstract
In the modern era like today, education is one of the most important aspects in supporting the development of quality human resources. In addition, through the main discussion carried out in this study aims to be able to apply the K-Means algorithm as a method for grouping the interests of class X students of SMK Satrya Budi 1 Perdagangan systematically and based on data. In addition, to be able to determine the level of compatibility between student interests and available majors based on the results of the grouping using the K-Means algorithm. And in addition, for the Research Method section used in this study, namely a quantitative approach with the support of data analysis techniques using the K-Means Algorithm Method. The selection of this method is based on the need for research to be able to produce objective grouping of class 10 student majors based on numerical data and certain relevant attributes. So based on this, this study shows the results that the application of the k-means clustering algorithm is able to provide a more objective and systematic approach in the process of determining majors. Grouping is done by processing several assessment criteria such as academic grades, aptitude tests, interest tests, entrance exams and basic skills. Based on the final results of the grouping of 30 students, the system divides students into four main major groups, namely Motorcycle Engineering and Business (8 students), Automotive Light Vehicle Engineering (2 students), Heavy Equipment Engineering (11 students) and Industrial Chemistry (10 students).

Keywords:
K-Means Algorithm; Data Mining; Data Clustering; Major Interests; Vocational High School Students.

1. INTRODUCTION

In the modern era, education is a crucial aspect in supporting the development of quality human resources. Education is not only aimed at increasing academic knowledge, but also at equipping students with skills and competencies that meet the needs of the workforce (Darsono & Andrianti, 2022). As a medium for developing students' skills based on their interests and talents, vocational high school (SMK) education plays a strategic role in preparing students to enter the workforce and continue their education to a higher level (Dhamawan, W., 2023).

At the vocational high school level, one of the important decisions students must make is choosing a major or program of expertise that aligns with their individual interests, talents, and potential. This choice of major not only determines the academic path students will take, but also has a significant impact on their readiness to face the world of work and their ability to develop competencies relevant to current industry needs (Nasution & Saragih, 2021).

Satya Budi 1 Private Vocational School in Perdagangan is one of the most popular vocational high schools in the Perdagangan sub-district. With its strategic location and supportive educational environment, Satya Budi 1 Private Vocational School is one of the schools that are usually chosen to be able to pursue

vocational school for the surrounding community. As one of the most popular schools, Satya Budi 1 Private Vocational School certainly has a very large number of students. This can happen because this school provides 4 (four) majors/expertise programs, namely the automotive light vehicle engineering major (TKRO), the motorcycle engineering and business major (TBSM), the heavy equipment engineering major (TAB) and also the Industrial Chemistry major (KIN).

With the many choices of majors/expertise programs, the Satya Budi 1 Private Vocational School has a total number of students of around 756 students. It is recorded that for the 2025-2026 academic year, the vocational school has a total of 290 new students consisting of 235 male students and 55 female students. Even though this school has a very large number of students, when determining the major they want to choose, many students still make the wrong choice (Anwar et al., 2025).

Although this school has been established for a long time, there are still several problems, including how to place students in majors or expertise programs that suit their interests, talents, and potential. The process of determining the right major will not only affect students' learning motivation and academic achievement, but also impact their readiness to face the professional world after graduation (Permatasari & Tundjungsari, 2024). So far, the process of selecting majors at Satrya Budi 1 Private Vocational School is still done manually and is subjective. Where teachers or the school usually rely on direct observation and simple interviews to find out the interests of their students (Saputra et al., 2023).

In addition to the manual and subjective nature of the major selection process, there are also various other problems, such as identifying and accurately categorizing students' interests and potential. This problem stems from the limitations of the assessment methods used to determine the suitability of students' interests to the available majors. In the context of SMK Satrya Budi 1 Perdagangan, most of the interest identification process still relies on the perceptions of guidance and counseling (BK) teachers and the results of simple interviews without being supported by comprehensive data analysis. As a result, many students do not fully understand their potential, leading to a mismatch between their personal interests and the chosen major (Rofiq & Qoiriah, 2021).

Another problem arises from the students themselves. Many tenth-grade students lack a clear understanding of their interests and potential. They tend to choose majors based on the influence of their peers, parents, or perceptions of job prospects, without considering their abilities and suitability to their personal capabilities. As a result, after some time in their chosen major, some students begin to lose motivation to study because they feel it is not a good fit for their chosen field (Asmana et al., 2022).

On the other hand, there has been a case experienced by one of the students when he was initially registered in the Automotive Light Vehicle Engineering (TKRO) department when he first entered grade 10. The decision to choose this major did not come entirely from his personal interests, but rather due to the encouragement of his parents and the influence of his friends, most of whom also chose the same major. Initially, the student tried to adapt to learning activities that focused heavily on the automotive field, such as vehicle engine maintenance, car electrical systems and workshop practice. However, as time went by, the student began to feel a mismatch between his personal interests and the chosen major. He felt less interested in workshop activities and engine analysis and also had difficulty understanding technical concepts related to light vehicle machining. His learning motivation also decreased, which was marked by a decrease in practical and theoretical grades. After an evaluation by the homeroom teacher and guidance counselor, it was discovered that the student actually had a stronger interest in the field of heavy equipment mechanization.

Based on the description of problems and case studies that have occurred, this is where the role of technology can provide significant solutions. With advances in information technology and the application of data-based methods, schools now have the opportunity to utilize technology as a tool in managing and analyzing student data more objectively (Ratih Yulia Hayuningtyas, 2022). Data regarding academic grades, talent assessment test results, interest assessment test results, entrance exam test results and basic skills assessment test results can be collected, processed and analyzed using various data analysis methods. The result is more accurate information that can be used as a basis for decision making in placing students in appropriate majors (Rusdianto et al., 2020).

One technology that can be applied is the clustering algorithm, specifically the K-Means algorithm. The K-Means algorithm is a popular method in the fields of data mining and machine learning, especially for grouping or clustering tasks. This algorithm works on the principle of dividing a set of data into several groups (clusters) based on the similarity of certain attributes. Each group will have a centroid, which is a center point that represents the average characteristics of the group members. Data with the highest similarity will be placed in the same group, while differences between groups will be kept significant (Rosadi et al., 2024).

By using the K-Means Algorithm method as a grouping method, the school can see the overall pattern of students' interests and talents, which can be used as a basis for recommendations for appropriate majors. This approach is much more systematic than manual methods which are often subjective and only rely on teacher intuition or the counselor's personal experience. On the other hand, by applying the K-Means algorithm, it is hoped that it can provide solutions to problems in determining students' interests and majors more accurately and objectively (Fahmi et al., 2025). From various problems that occur in the field, such as difficulties in determining appropriate majors and the lack of systematic analysis tools, the authors in this

study concluded to take the research title, namely " The Use of the K-Means Algorithm as a Method for Grouping Major Interests of Class X Students of SMK Satrya Budi 1 Perdagangan".

2. RESEARCH METHOD

The method used in this study is a quantitative approach supported by data analysis techniques using the K-Means Algorithm Method. The selection of this method is based on the research need to be able to produce objective grouping of 10th grade students' majors based on numerical data and certain relevant attributes. Thus, the research can produce decisions on major groupings that are measurable, logical and in accordance with the characteristics of each student. In this study also uses a quantitative approach because the entire analysis process involves processing numerical data such as academic grades, interest test scores, aptitude test scores, entrance test scores and basic skills assessments. This approach allows researchers to obtain more structured results and can be analyzed mathematically (Misbakhul Anam et al., 2025). And based on this, in this study the research framework used includes the following description and explanation.

2.1. Identification of Problems

This stage is the initial stage of the research, where the researcher must first determine what problem will be taken for research, for SMK Swasta Satrya Budi 1 Perdagangan itself, the researcher conducted observations on the process of selecting student majors which is still done manually through direct observation and interviews conducted by teachers or counselors with students. Although the school provides guidance and counseling services, the process does not utilize student data (Alawiyah et al., 2025).

2.2. Data Collection

The stages of data collection using interview, observation and questionnaire methods. The interviews conducted were by asking things that the researcher felt were the main problems currently being faced by the school, the observations conducted by the researcher were to directly observe how the process of determining majors had been running and how the concept of the guidance and counseling system had been implemented by the school guidance and counseling department and at this questionnaire stage the researcher gave a number of questions through the Google Form Platform which was given to the 10th grade students from each major (Amelia et al., 2025).

2.3. Research Data Set

The dataset in this study was compiled based on student grade data obtained from several important assessment aspects. Based on the table above, each student data has attributes in the form of academic grades, talents, interests, entrance exams, and skills. These attributes are used as the main variables in the grouping process because they reflect the abilities and tendencies of students' interests in certain majors. The amount of data used includes dozens of students with varying grades, thus providing a fairly representative picture. The data was then processed using the K-Means algorithm to form several groups (clusters) of students with similar characteristics. The results of this grouping are expected to be able to help schools in determining the appropriate major for each student more objectively, based on the pattern of grades they have, rather than just based on subjective assessments alone (Nurarofah et al., 2023).

Table 1Cluster Count Value Data Table

No	Full name	Academic	Talent	Interest	Entrance examination	Skills
1	Alfino Siregar	82	78	80	84	79
2	Aditia Saputra	75	80	77	78	81
3	Adjie al fikri	88	85	83	86	84
4	Aga Tri Wibowo	70	74	72	75	73
5	Agung Kurniawan	85	82	84	83	86
6	Ahmad zain abiyyu nst	90	88	87	89	85
7	Akbar Maulana Makmur Sirait	78	76	79	80	77
8	Akhri Rizal Rizky	83	81	82	84	80
9	Ananda Aditya Pratama	87	86	88	85	87
10	Arief Ramadhan	79	77	75	78	76
11	Arya Lesmana	92	90	91	93	89
12	Azril Arizki	74	73	76	72	75
13	Baimsyah Ramadani	80	82	81	79	83
14	Bayu Gustiawan	76	75	74	77	78
15	Chairul Fiqri Coal	84	83	82	85	81
16	Danu Kurniawan	71	70	72	73	74
17	Dedek Pratama	77	78	79	76	80
18	Denni Admaja	82	80	81	83	82

19	Dhafin Septian Algadi	89	87	88	90	86
20	Dimas Arya Singgih	86	85	84	87	85
21	Dimas Panjaitan	73	72	74	71	75
22	Dio alfansya	78	79	77	80	78
23	Dirga Febriansyah Hutagaol	85	84	83	86	82
24	Duan parizi	81	82	80	83	81
25	Fadly Syahputra	88	87	89	90	86
26	Fahri Fatah Khiyar	84	85	86	83	87
27	Fery Febriansyah	79	78	80	77	81
28	Fitra Ramadhan	83	81	82	84	80
29	Gabriel Jericho Simatupang	91	89	90	92	88
30	Galu Setiawan	72	73	71	74	75

2.4. Research Variables

The research variables in this study consist of several indicators used to represent student characteristics. The main variables used are academic grades, talents, interests, entrance exams, and skills. Academic grades reflect students' abilities in formal learning aspects, while talents describe the natural potential possessed by students. Interest variables indicate students' tendency to be interested in certain fields, which is an important factor in determining majors. Entrance exam scores are used as an initial indicator of students' basic abilities when accepted into school. Meanwhile, skills indicate students' practical abilities in applying knowledge. All of these variables are numeric and are used as the basis for the clustering process using the K-Means algorithm to group students into several categories of appropriate major interests based on similarities in their characteristics (Amanda et al., 2023).

2.5. Analysis Stages

The analysis stage of this study begins with the student data collection process, which includes academic grades, talents, interests, entrance exams, and skills. Data preprocessing is then carried out, such as checking data completeness and normalizing values so that each variable has a balanced scale. After that, the number of clusters (k) that will be used as the basis for grouping is determined. The core process is carried out by applying the K-Means algorithm, namely by determining the initial centroid, calculating the distance of each data to the centroid, then grouping the data based on the closest distance. This step is repeated until the centroid position is stable. The final stage is the interpretation of the cluster results to determine the characteristics of each student group. The results of this analysis are used as a basis for providing recommendations for determining majors that suit the students' potential and interests (Lillah et al., 2023).

2.6. Utilization of the K-Means Clustering

The K-Means algorithm is one of the most widely used unsupervised learning techniques in data grouping or clustering. This algorithm works by grouping data into several clusters based on the level of similarity in characteristics between the data. Clustering is one of the main techniques in the field of data mining and machine learning that functions to group a set of data into several groups or clusters based on the level of similarity in characteristics between the data (Dhewayani et al., 2022). The following is a description of the method:

a. Determining the Number of Cluster
   In determining the centroid value for the initial centroid value, it is done randomly, whereas if determining the centroid value which is a stage of interaction, the following formula is used:
   $\overline{v_{ij}} = \frac{1}{N_i} \sum_{k=0}^{N_i} x_{kj}$, ... (1)
b. Calculate the distance between the centroid point and each object point
   $D_e = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2}$ .....
c. Grouping each object to determine cluster members is done by calculating the minimum distance of the object.
d. Continue to stage 2, repeat the process until the resulting centroid value remains constant and cluster members do not move to other clusters (Dhewayani et al., 2022).

3. RESULTS AND DISCUSSION

The results and discussion section of this study presents the results of data processing that has been carried out using the K-Means algorithm to group the interests of class X students' majors at SMK Satrya Budi 1 Perdagangan. This stage aims to show how the data that has been collected from students can be analyzed so as to produce more structured and easy-to-understand information. Through this data processing process, researchers attempt to obtain an overview of the tendencies of student interests in the majors

available at the school (Candra et al., 2021). The following is an explanation of the steps for completing the Linear Regression Method.

3.1. Data Processing

Data processing in this study was carried out systematically to produce data ready for analysis. The initial stage began with collecting student grades, covering academic aspects, talents, interests, entrance exams, and skills. After that, a data selection and cleaning process was carried out to ensure there were no blanks, duplicates, or inconsistencies. Next, the data was normalized so that all variables were on a comparable scale, preventing any single attribute from being more dominant in the distance calculation. The cleaned data was then processed using the K-Means algorithm by determining the desired number of clusters. This process involves calculating distances between data and repeatedly updating centroids until stable clusters are obtained. The final result of the data processing is a grouping of students with similar characteristics as a basis for determining major interests.

3.2. Utilization of the K-Means Clustering

To improve the accuracy of grouping student major interests, an approach capable of processing data objectively and systematically is required. One method that can be utilized is the K-Means Clustering algorithm, known as a technique for grouping data based on similarities in certain characteristics. By utilizing relevant data such as academic grades, aptitude assessment tests, interest assessment tests, entrance exam tests and basic skills assessment tests, the K-Means algorithm is able to divide students into several groups that have similar characteristics. And the following is a description of its use:

a. The first step is to determine the number of clusters (K). At this stage, researchers determine how many groups will be formed according to the analysis needs.
b. The second step is determining the initial centroid. The centroid is the center point of each cluster. In the initial stage, centroids are usually determined randomly from available data or based on specific values considered representative of each cluster.

Table 1Initial Centroid Value Data

C1	92	90	91	93	89
C2	92	92	92	93	89
C3	70	70	72	71	73
C4	72	73	71	73	75

c. The third step is to calculate the distance between each data point and the centroid. The system calculates the distance between each data point and all existing centroids. This distance calculation aims to determine the proximity of the data to each cluster. Generally, the distance used is the Euclidean distance.

Table 2Centroid C1 Value Data

C1	1	2	3	4	5
	10	12	11	9	10
	17	10	14	15	8
	4	5	8	7	5
	22	16	19	18	16
	7	8	7	10	3
	2	2	4	4	4
	14	14	12	13	12
	9	9	9	9	9
	5	4	3	8	2
	13	13	16	15	13
	0	0	0	0	0
	18	17	15	21	14
	12	8	10	14	6
	16	15	17	16	11
	8	7	9	8	8
	21	20	19	20	15
	15	12	12	17	9
	10	10	10	10	7
	3	3	3	3	3
	6	5	7	6	4
	19	18	17	22	14
	14	11	14	13	11
	7	6	8	7	7

11	8	11	10	8
4	3	2	3	3
8	5	5	10	2
13	12	11	16	8
9	9	9	9	9
1	1	1	1	1
20	17	20	19	14

d. The fourth step is iteration. The process of calculating distances, grouping data, and updating centroids will be repeated until the centroid positions no longer change or the changes are very small. This indicates that the clustering process has reached convergence.

Table 3Iteration Value Data

1	2	3	4	5	Amount	C1 (Root)
100,000	144,000	121,000	81,000	100,000	546,000	23,367
289,000	100,000	196,000	225,000	64,000	874,000	29,563
16,000	25,000	64,000	49,000	25,000	179,000	13,379
484,000	256,000	361,000	324,000	256,000	1681,000	41,000
49,000	64,000	49,000	100,000	9,000	271,000	16,462
4,000	4,000	16,000	16,000	16,000	56,000	7,483
196,000	196,000	144,000	169,000	144,000	849,000	29,138
81,000	81,000	81,000	81,000	81,000	405,000	20,125
25,000	16,000	9,000	64,000	4,000	118,000	10,863
169,000	169,000	256,000	225,000	169,000	988,000	31,432
0,000	0,000	0,000	0,000	0,000	0,000	0,000
324,000	289,000	225,000	441,000	196,000	1475,000	38,406
144,000	64,000	100,000	196,000	36,000	540,000	23,238
256,000	225,000	289,000	256,000	121,000	1147,000	33,867
64,000	49,000	81,000	64,000	64,000	322,000	17,944
441,000	400,000	361,000	400,000	225,000	1827,000	42,743
225,000	144,000	144,000	289,000	81,000	883,000	29,715
100,000	100,000	100,000	100,000	49,000	449,000	21,190
9,000	9,000	9,000	9,000	9,000	45,000	6,708
36,000	25,000	49,000	36,000	16,000	162,000	12,728
361,000	324,000	289,000	484,000	196,000	1654,000	40,669
196,000	121,000	196,000	169,000	121,000	803,000	28,337
49,000	36,000	64,000	49,000	49,000	247,000	15,716
121,000	64,000	121,000	100,000	64,000	470,000	21,679
16,000	9,000	4,000	9,000	9,000	47,000	6,856
64,000	25,000	25,000	100,000	4,000	218,000	14,765
169,000	144,000	121,000	256,000	64,000	754,000	27,459
81,000	81,000	81,000	81,000	81,000	405,000	20,125
1,000	1,000	1,000	1,000	1,000	5,000	2,236
400,000	289,000	400,000	361,000	196,000	1646,000	40,571

e. The fifth step is iteration. The process of calculating distances, grouping data, and updating centroids will be repeated until the centroid positions no longer change or the changes are very small. This indicates that the clustering process has reached convergence.

Table 4Iteration Value Data

C1	C2	C3	C4	Average value
23,367	73,163	51,980	55,408	81.6
29,563	68,214	44,560	52,681	75
13,379	80,811	63,865	63,335	87
41,000	63,032	32,458	52,385	80.4
16,462	78,180	60,581	60,368	77
7,483	85,902	71,962	70,033	89.6
29,138	69,220	45,137	52,644	84
20,125	75,345	55,668	57,307	72
10,863	82,774	67,981	65,862	86
31,432	68,334	41,843	52,371	79
0,000	93,032	83,039	80,946	83
38,406	64,758	34,807	51,908	76

23,238	72,751	51,796	55,154	88
33,867	66,751	39,524	52,001	72
17,944	76,989	58,231	59,077	85
42,743	63,212	30,506	52,762	81
29,715	68,607	44,269	52,400	89
21,190	74,510	54,452	56,663	78
6,708	86,604	73,481	71,297	74
12,728	81,226	65,167	64,216	84
40,669	64,027	32,423	52,358	80
28,337	69,572	45,634	53,091	87
15,716	78,758	61,107	61,058	75
21,679	73,997	53,611	56,353	82
6,856	86,417	73,563	71,120	88
14,765	79,314	62,976	61,854	79
27,459	70,176	46,878	53,085	91
20,125	75,345	55,668	57,307	73
2,236	90,841	79,810	77,567	85
40,571	63,803	32,567	52,367	80

3.3. Grouping Results

The results section on the grouping of student majors in this study begins with a general overview of the analysis process conducted using the K-Means algorithm. This stage is important because it serves as a bridge between the data processing process and the interpretation of the results obtained. Through an average-based approach, previously heterogeneous student data is processed into more structured groups, thus providing a clearer and more objective picture of the pattern of major interests. This process not only considers the numerical aspect but also attempts to represent the tendencies of students' academic abilities as reflected in their average scores. The following grouping results are outlined in the table 6.

Table 5. Data Grouping of Majors Based on Average Grades

No	Student Name	Class	Suitable Major	Average value
1	Alfino Siregar	A	Teknik Alat Berat	81,6
2	Aditia Saputra	A	Teknik dan Bisnis Sepeda Motor	75
3	Adjie Al Fikri	A	Kimia Industri	87
4	Aga Tri Wibowo	A	Teknik Alat Berat	80,4
5	Agung Kurniawan	A	Teknik dan Bisnis Sepeda Motor	77
6	Ahmad Zain Abiyyun NST	A	Kimia Industri	89,6
7	Akbar Maulana Makmur Sirait	A	Teknik Alat Berat	84
8	Akhri Rizal Rizky	A	Teknik dan Bisnis Sepeda Motor	72
9	Ananda Aditya Pratama	A	Kimia Industri	86
10	Andika Pratama	A	Teknik Alat Berat	79
11	Angga Saputra	A	Teknik Alat Berat	83
12	Ardiansyah	A	Teknik dan Bisnis Sepeda Motor	76
13	Arif Rahman	A	Kimia Industri	88
14	Arifin	A	Teknik dan Bisnis Sepeda Motor	72
15	Bagas Pratama	A	Kimia Industri	85
16	Bima Sakti	A	Teknik Alat Berat	81
17	Daffa Ramadhan	A	Kimia Industri	89
18	Dimas Saputra	A	Teknik Alat Berat	78
19	Egi Pratama	A	Teknik dan Bisnis Sepeda Motor	74
20	Fajar Nugroho	A	Teknik Alat Berat	84
21	Fikri Maulana	A	Teknik Alat Berat	80
22	Gilang Ramadhan	A	Kimia Industri	87
23	Hafiz Alfarizi	A	Teknik dan Bisnis Sepeda Motor	75
24	Ihsan Maulana	A	Teknik Alat Berat	82
25	Ilham Prasetya	A	Kimia Industri	88
26	Irvan Hidayat	A	Kimia Industri	79
27	Kurniawan	A	Kimia Industri	91
28	M. Rizki Pratama	A	Teknik dan Bisnis Sepeda Motor	73
29	M. Yusuf	A	Teknik Kendaraan Ringan Otomotif	85
30	Nanda Saputra	A	Teknik Kendaraan Ringan Otomotif	80

3.4. Evaluation of Grouping Through the Elbow Method

In the clustering evaluation stage of this study, the Elbow method was used as an approach to determine the most optimal number of clusters (k) in the application of the K-Means algorithm. This evaluation is an important part to ensure that the results of the grouping of interests of class X students of SMK Satrya Budi 1 Perdagangan truly represent naturally formed data patterns. Based on the processed student data, which includes various major choices such as Heavy Equipment Engineering, Motorcycle Engineering and Business, Automotive Light Vehicle Engineering, and Industrial Chemistry, the Elbow method helps identify the point were increasing the number of clusters no longer provides a significant decrease in error values.

3.5. System Implementation

The system implementation in this study was carried out by building an application capable of automatically processing student data. The system is designed to accept input in the form of academic grades, talents, interests, entrance exams, and skills, which are then stored in a database. Next, the system runs a calculation process using the K-Means algorithm to group students into several clusters based on similar characteristics. The grouping results are displayed in an easy-to-understand format, allowing schools to see recommendations for appropriate majors for each student. Furthermore, the system also supports the testing process by entering new student data to obtain classification results quickly. With this implementation, the process of determining majors becomes more effective, objective, and structured compared to manual methods.

3.5.1. Login Page View

The user login page display is the website page display that will be accessed by the user to enter the main page of the website, where the user will carry out the account verification process to log in to the main page display of the website.

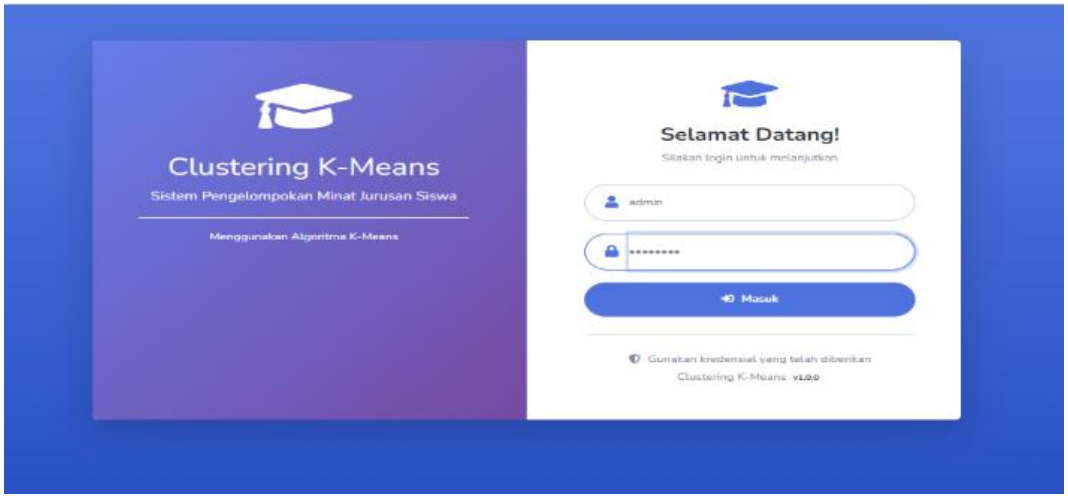

Figure 1. Login Page Display

3.5.2. K-Means Clustering Calculation Process Data Detail Page View

C. The K-Means Clustering Calculation Process Data Detail Page is a website page that users can access to view the results of the K-Means Clustering method calculation process that has been run on the system based on the assessment that has been carried out.

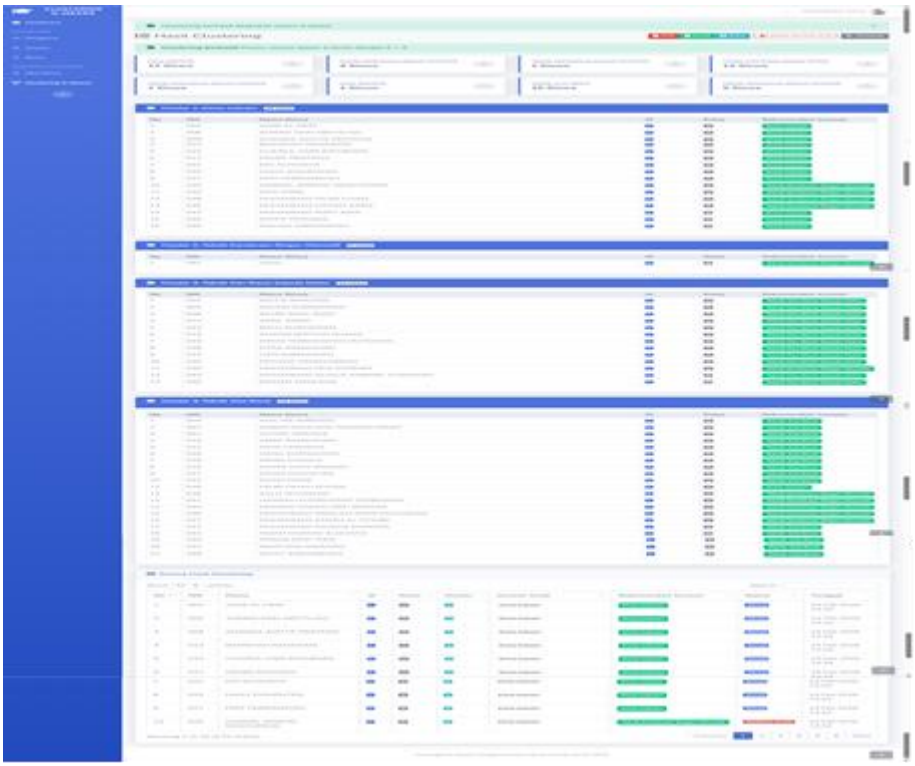

Figure 2. Detail Page View of Calculation Process Data Details K-Means Clustering

4. CONCLUSION

Based on the results of research on the use of the K-Means algorithm as a method for grouping the interests of class X students of SMK Satrya Budi 1 Perdagangan, it can be concluded that this method is able to provide a more objective and systematic approach in the process of determining majors. Grouping is done by processing several assessment criteria such as academic grades, aptitude tests, interest tests, entrance exams and basic skills. Through an iterative process in determining the centroid and calculating the proximity distance between data, the system successfully forms student clusters based on the similarity of the characteristics of the values they have, thus producing more measurable major recommendations compared to manual methods. Based on the final results of the grouping of 30 students, the system divides students into four main major groups, namely Motorcycle Engineering and Business (11 students), Automotive Light Vehicle Engineering (2 students), Heavy Equipment Engineering (11 students), and Industrial Chemistry (10 students). The largest number of students are in the Industrial Chemistry major, which indicates that most students have a tendency to be able to analyze and process industrially. Meanwhile, the Motorcycle Engineering and Business and Automotive Light Vehicle Engineering majors are filled by students with more dominant value characteristics in automotive technical abilities and the Heavy Equipment Engineering major is filled by students with a tendency towards strong technical abilities and field work readiness.

REFERENCES

Alawiyah, A. (2025). Implementasi Clustering Algoritma K-Means Pada Penjualan Beras Di CV Tangguh Bumi Perkasa. Jurnal Komputer dan Sistem Informasi, 2(2), 64-71.

Mariska Dwi Amanda, Santoso, G., Ari Metalin Ika Puspita, & Frischa Amalia Imanda. (2023). Kontribusi Masyarakat dalam Perspektif Ketahanan Nasional Indonesia di Era Digital. Jurnal Pendidikan Transformatif (JPT), 2(6), 45–63. https://doi.org/10.9000/jpt.v2i5.1049

Amelia, M., Faqih, A., & Rinaldi, A. R. (2025). Penerapan Metode K-Means Clustering Dalam Pemetaan Kemiskinan Kabupaten/Kota Di Indonesia Untuk Perencanaan Kebijakan Yang Tepat. Jurnal Informatika Dan Teknik Elektro Terapan, 13(2). https://doi.org/10.23960/jitet.v13i2.6231

Anwar, M., Firmansyah, E., & Sutara, B. (2025). Analisis Clustering Menggunakan Metode K-Means dalam Penentuan Prioritas Operasional Penagihan Pajak Bumi dan Bangunan di Desa Gunasari. Data Sciences Indonesia (DSI), 5(2), 101–110. https://doi.org/10.47709/dsi.v5i2.7464

Asmana, A., Arie Wijaya, Y., & Martanto, M. (2022). Clustering Data Calon Siswa Baru Menggunakan Metode K-Means Di Sekolah Menengah Kejuruan Wahidin Kota Cirebon. JATI (Jurnal Mahasiswa Teknik Informatika), 6(2), 552–559. https://doi.org/10.36040/jati.v6i2.5236

Candra, H. K., Bahit, M., & Sabella, B. (2021). Penerapan Metode Klustering Fuzzy C-Means Untuk Penentuan Peminatan Pemilihan Jurusan Pada Sekolah Menengah Tingkat Atas. POSITIF: Jurnal Sistem Dan Teknologi Informasi, 7(2), 108–119. https://doi.org/10.31961/positif.v7i2.1106

Darsono, V., & Andrianti, A. (2022). Penerapan Data Mining Algoritma K-Means Untuk Rekomendasi Pemilihan Bidang Studi Perguruan Tinggi Pada Siswa SMKN 1 Kota Jambi. Jurnal Informatika Dan Rekayasa Komputer (JAKAKOM), 2(2), 161–171. https://doi.org/10.33998/jakakom.2022.2.2.80

Dhewayani, F. N., Amelia, D., Alifah, D. N., Sari, B. N., & Jajuli, M. (2022). Implementasi K-Means Clustering untuk Pengelompokkan Daerah Rawan Bencana Kebakaran Menggunakan Model CRISP-DM. Jurnal Teknologi Dan Informasi, 12(1), 64–77. https://doi.org/10.34010/jati.v12i1.6674

Fahmi, M. F., Helilintar, R., & Farida, I. N. (2025, July). Implementasi Metode K-Nearest Neighbor untuk Klasifikasi Tingkat Kepuasan Pelanggan pada Toko Buku Anugerah Ilmu. In Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi) (Vol. 9, No. 2, pp. 1471–1480). https://doi.org/10.29407/x4mg2w61

Lillah, M. R. R. L., Maylawati, D. S.., Zulfikar, W. B.., Uriawan, W.., & Wahana, A.. (2023). Implementasi Algoritma K-Nearest Neighbor (KNN) untuk Analisis Sentimen Pengguna Aplikasi Tokopedia. Intellect: Indonesian Journal of Learning and Technological Innovation, 2(2), 171–184. Retrieved from https://www.journal.makwafoundation.org/index.php/intellect/article/view/296

Misbakhul Anam, Annisa Maulana Majid, & Ermanto. (2025). Analisis Pola Pengangguran Menggunakan Metode Clustering Algoritma K-Means Di Wilayah Kabupaten Cirebon. Jurnal Teknologi Dan Manajemen Industri Terapan, 4(4), 1691–1701. https://doi.org/10.55826/jtmit.v4i4.743

Nasution, R. A., Saragih, I. S., & Sormin, R. K. (2021). Pelatihan Dasar Kepemimpinan (Leadership) Pada Anggota OSIS SMK Satrya Budi 2 Perdagangan Kab. Simalungun The Basic Training of Leadership on Member OSIS At SMK Sarta Budi 2 Perdagangan Simalungun District. Jurnal Ilmiah Pengabdian Kepada Masyarakat Jurnal TUNAS, 3(1), 147–151. https://dx.doi.org/10.30645/jtunas.v3i1.45

Nurarofah, E., Herdiana, R., & Dienwati Nuris, N. (2023). Penerapan Asosiasi Menggunakan Algoritma Fp-Growth Pada Pola Transaksi Penjualan di Toko Roti. JATI (Jurnal Mahasiswa Teknik Informatika), 7(1), 353–359. https://doi.org/10.36040/jati.v7i1.6299

Permatasari, W. I., & Tundjungsari, V. (2024). Implementasi Algoritma K-Means untuk Mengetahui Minat Siswa SMA Terhadap Mata Pelajaran Teknologi Informasi dan Komunikasi. MALCOM: Indonesian Journal of Machine Learning and Computer Science, 5(1), 168–174. https://doi.org/10.57152/malcom.v5i1.1639

Ratih Yulia Hayuningtyas, R. S. (2022). Implementasi Data Mining Dengan Algoritma Multiple Linear Regression Untuk Memprediksi Penyakit Diabetes. Jurnal Teknik Komputer AMIK BSI, 8(1), 40–44. https://doi.org/10.31294/jtk.v4i2

Rofiq, M. A., & Qoiriah, A. (2021). Pengelompokan Kategori Buku Berdasarkan Judul Menggunakan Algoritma Agglomerative Hierarchical Clustering dan K-Medoids. Journal of Informatics and Computer Science (JINACS), 2(03), 220–227. https://doi.org/10.26740/jinacs.v2n03.p220-227

abriel Rosadi, A., Helilintar, R., & Farida, I. N. (2025, July). Penerapan Algoritma K-Means untuk Pengelompokan Minat dan Prestasi Siswa SMK PGRI 1 Kediri. In Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi) (Vol. 9, No. 2, pp. 1183–1190). https://doi.org/10.29407/aspqv204

Rusdianto, D., & Zaelani, L. (2020). Implementasi Data Mining Menggunakan Algoritma Apriori Untuk Mengetahui Pola Peminjaman Buku di Perpustakaan Universitas Bale Bandung. J-SIKA| Jurnal Sistem Informasi Karya Anak Bangsa, 2(02), 1–10. https://ejournal.unibba.ac.id/index.php/j-sika/article/view/376

Saputra, I. P. Y., Siswanto, S., & Fredricka, J. (2023). Penerapan Metode K-Means Clustering Dalam Klasterisasi Peminatan Siswa Terhadap Mata Pelajaran Simulasi Digital (Simdig). Jurnal Media Infotama, 19(1), 20–26. https://doi.org/10.37676/jmi.v19i1.3308

Dhamawan, W. (2023, March 14). Algoritma K-Means Clustering Pada Pengelompokan Minat Bakat Siswa SMK PGRI 2 Karawang. JATISI, 10(1), 38-50. https://doi.org/https://doi.org/10.35957/jatisi.v10i1.2469