JOSAR, Vol.9 No.1 March 2024
ISSN: 2502-8521 (Print) / 2503-1155 (Online)
DOI: https://doi.org/10.35457/josar.v9i1.3105

APPLICATION OF K-MEANS CLUSTERING ALGORITHM FOR DETERMINING PIP SCHOLARSHIP RECIPIENTS AT SMPN 9 BLITAR

Andi Abrianto[1]

[1]University Students, Faculty of Informatics Engineering, Universitas Islam Balitar, Indonesia

Abstract
This research aims to determine the eligibility of recipients for the Program Indonesia Pintar (PIP) among students of SMPN 9 Blitar by applying the K-means clustering algorithm. PIP is an educational assistance program targeted at students from economically disadvantaged families in Indonesia. However, the selection process for recipients still needs improvement to be more efficient and accurate. In this study, data on students' socioeconomic status as input variables were collected with a sample size of 286. The K-means clustering method was used to divide students into groups based on similar characteristics, including family income level, access to facilities for underprivileged families, parents' or guardians' employment status, and household conditions. By grouping these characteristics, it becomes possible to accurately determine the eligibility of PIP recipients. The result of this research is that the implementation of the K-means clustering method can improve accuracy in determining eligibility for PIP recipients efficiently. Specifically, out of 286 students analyzed in this study, 201 were found eligible while 85 were deemed ineligible.
Keywords: *Eligibility, PIP, Algoritma ,K-means ,Clustering*

1. INTRODUCTION

The importance of education is the responsibility of the government in developing education for citizens to become human beings who have a healthy character and social life. The short specification can be said that education is the key to forming quality human resources.

Education is always changing and developing drastically so that the government must further improve the quality and quantity of education in the State of Indonesia. In this case the teaching (teachers), facilities, and education personnel are further improved, with the aim of advancing the country. PIP is a poor student assistance program that covers students from SD / MI, SMP / MTs, SMA / SMK / MA education levels. In this case, the provision of PIP assistance must be given to students who deserve this assistance, because many children who are still in education have the right to get and follow education.

[1*]Corresponding author, email: andiabrianto.05@gmail.com

Citation in APA style: Abrianto, A. (2024). Application Of K-Means Clustering Algorithm For Determining PIP Scholarship Recipients At Smpn 9 Blitar. *JOSAR (Journal of Students Academic Research)*, Vol. 9. No. (1):204-214.

Received:	**Revised:**	**Published:**
March, 1st 2024	March, 15th 2024	March, 31st 2024

DOI: https://doi.org/10.35457/josar.v9i1.3105

ISSN: 2502-8521 (Print) / 2503-1155 (Online)
DOI: https://doi.org/10.35457/josar.v9i1.3105

In this case, SMPN 9 Blitar is one of the schools that has problems related to determining PIP recipients. This school is located in Pakunden Village, Sukorejo Subdistrict, precisely on the border between Blitar City and Blitar Regency. Due to its location on the outskirts of the city or border, most of the jobs of parents or guardians of students include casual laborers, farmers and small traders. Related to this, from an economic point of view, the majority of parents of students in this school are classified as poor or vulnerable to poverty.

From this background, the problem faced is how the school policy determines or chooses which students really deserve PIP, so that the program run by the government is appropriate and right on target in the criteria for its provision.

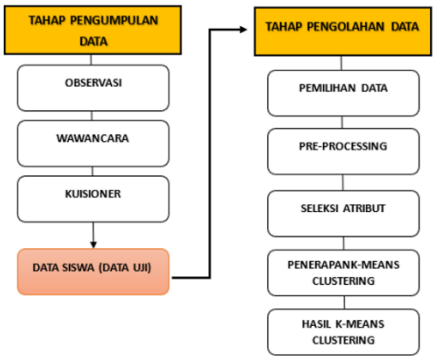
In solving this problem, the author has a solution that is in accordance with the problems faced, namely by applying the K-Means Algorithm using the Data Clustering Method. This method can help determine the group of students who are eligible to receive PIP and the group that is not eligible to receive PIP, by grouping the data variables taken from student data. Some of the data variables used are about parents' income, parents' occupation, number of siblings, living status, ownership of a certificate of incapacity and other data.

The reason for using the K-Means algorithm is because this algorithm can be used to divide a number of random data into partitions based on categories that exist in a set of data. This algorithm is one of the most popular and accurate algorithms to be applied and run for the data clustering process. Meanwhile, the clustering process is a series of processes to explore the added value of a data set in the form of unknown knowledge. Clustering is also a method that can group raw or random data into data sets that have a very close level of similarity. By applying K-Means Clustering, it can help schools group students who are eligible to be proposed for PIP and which students are not eligible to be proposed for PIP based on student data that has been processed. Thus K-Means Clustering is very useful to explore information from existing data, so that it can be used as a solution to solving a problem.

2. METHOD

2.1 The Flow of K-Means Clustering Process

The following is the flow or stages of the initial process, namely data collection to the calculation stage with K-Means Clustering.



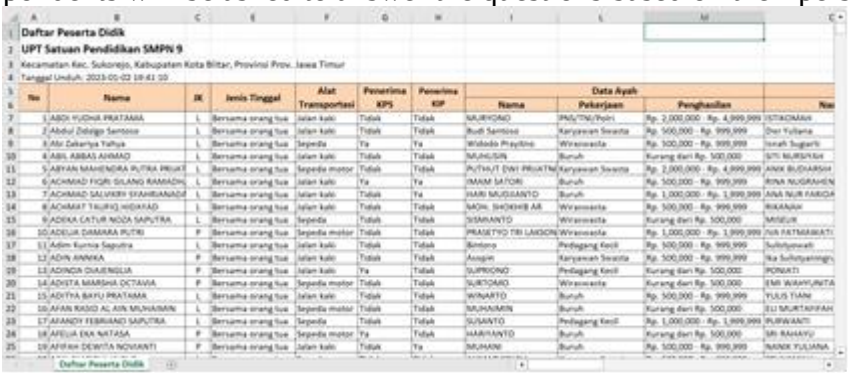
Picture. 2.1 Stages of data clustering process

2.2 Data Collection Stage

ISSN: 2502-8521 (Print) / 2503-1155 (Online)
DOI: https://doi.org/10.35457/josar.v9i1.3105

The research was located at SMPN 9 Blitar, which is located at Jalan Widas No. 29 Sukorejo District, Blitar City. As for the time of this research starting from February 1, 2023 to March 2023 by:

1. Observation
   Observation activities are carried out directly at SMPN 9 Blitar. From the results of observations obtained data that can be used in research.
2. Interview
   Interview activities were carried out with stakeholders related to the research case, in this case SMPN 9 Blitar.
3. Questionnaire
   The questionnaire was filled in by the student guardian by filling out the form that had been distributed by the school by answering a number of questions. Respondents will be asked to answer the questions based on their personal data.



Daftar Peserta Didik

UPT Satuan Pendidikan SMPN 9

Kecamatan Kec. Sukorejo, Kabupaten Kota Blitar, Provinsi Prov. Jawa Timur

Tanggal Unduh: 2023-01-02 10:41:10

No	Nama	JK	Jenis Tinggal	Alat Transportasi	Penerima KPS	Penerima KIP	Data Ayah			
							Nama	Pekerjaan	Penghasilan	
1	ABDI YUDHA PRATAMA	L	Bersama orang tua	Jalan kaki	Tidak	Tidak	MURYONO	PNS/TNI/Polri	Rp. 2,000,000 - Rp. 4,999,999	ISTIKOMAH
2	Abdul Zidqiqi Sembodo	L	Bersama orang tua	Jalan kaki	Tidak	Tidak	Budi Sembodo	Karyawan Swasta	Rp. 500,000 - Rp. 999,999	Dwi Yuliana
3	Ahi Zakariya Yafiya	L	Bersama orang tua	Sepeda	Ya	Ya	Widodo Prayitno	Wiraswasta	Rp. 500,000 - Rp. 999,999	Innah Sugiarti
4	ABIL ABBAS AHMAD	L	Bersama orang tua	Jalan kaki	Tidak	Tidak	MUNGSIN	Buruh	Kurang dari Rp. 500,000	SITI NURSIYAH
5	ABYAN MAHENDRA PUTRA PRIAT	L	Bersama orang tua	Sepeda motor	Tidak	Tidak	PUTHUT DWI PRIATNO	Karyawan Swasta	Rp. 2,000,000 - Rp. 4,999,999	ANIK BUDIARSIH
6	ACHMAD FIQRI SILANG RAMADH	L	Bersama orang tua	Jalan kaki	Ya	Ya	IMAM SATORI	Buruh	Rp. 500,000 - Rp. 999,999	RINA NUGRAHENI
7	ACHMAD SALVIKRY SYAHBANADI	L	Bersama orang tua	Jalan kaki	Tidak	Ya	HARI MUGIANTO	Buruh	Rp. 1,000,000 - Rp. 1,999,999	ANA NUR FARIDA
8	ACHMAT TAUFIQ HIDAYAD	L	Bersama orang tua	Jalan kaki	Tidak	Tidak	MOH. SHOKHIB AR	Wiraswasta	Rp. 500,000 - Rp. 999,999	RAMANA
9	ADEKA CATUR NOZA SAPUTRA	L	Bersama orang tua	Sepeda	Tidak	Tidak	SISMANTO	Wiraswasta	Kurang dari Rp. 500,000	MISEUR
10	ADELIA DAMARA PUTRI	P	Bersama orang tua	Sepeda motor	Tidak	Tidak	PRASETYO TRI LAKSONO	Wiraswasta	Rp. 1,000,000 - Rp. 1,999,999	IVA FATMAWATI
11	Adim Kurnia Saputra	L	Bersama orang tua	Jalan kaki	Tidak	Tidak	Bintoro	Pedagang Kecil	Rp. 500,000 - Rp. 999,999	Sulistyowati
12	ADIN ANNIKA	P	Bersama orang tua	Jalan kaki	Tidak	Tidak	Asepin	Karyawan Swasta	Rp. 500,000 - Rp. 999,999	Ika Sulistyaningsi
13	ADINDA DWIENGELIA	P	Bersama orang tua	Jalan kaki	Ya	Tidak	SUPRIONO	Pedagang Kecil	Kurang dari Rp. 500,000	PONIATI
14	ADISTA MARSHA OCTAVIA	P	Bersama orang tua	Sepeda motor	Tidak	Tidak	SUBTOARO	Wiraswasta	Kurang dari Rp. 500,000	ENI WAHYUNITA
15	ADITYA BAYU PRATAMA	L	Bersama orang tua	Jalan kaki	Tidak	Tidak	WINARTO	Buruh	Rp. 500,000 - Rp. 999,999	YULIS TIANI
16	AFAN RIXIO AL ASN MUHAIMIN	L	Bersama orang tua	Sepeda motor	Tidak	Tidak	MUHAIMIN	Buruh	Kurang dari Rp. 500,000	ZU MURTAFIFAH
17	AFANDY FEBRIANO SAPUTRA	L	Bersama orang tua	Sepeda	Tidak	Tidak	SUSANTO	Pedagang Kecil	Rp. 1,000,000 - Rp. 1,999,999	PURWANTI
18	AFELIA EKA NATASA	P	Bersama orang tua	Sepeda motor	Ya	Tidak	HARIYANTO	Buruh	Kurang dari Rp. 500,000	SRI RAHAYU
19	AFIFAH DEWITA NOVIANTI	P	Bersama orang tua	Jalan kaki	Tidak	Ya	MUHAINI	Buruh	Rp. 500,000 - Rp. 999,999	WANIK YULIANA

Daftar Peserta Didik

Picture 2.2 Student Preliminary Data

2.3 Data Processing Stage

1. Data Selection
   Data selection is the process of determining data for research, after which determining attributes related to the research is carried out. The data used from the Dapodik Application of SMPN 9 Blitar which is individual student data containing student names, number of siblings, type of residence, means of transportation, KPS recipient status, KIP recipient status, certificate of incapacity, father and mother's occupation, and father and mother's income and others.
2. Data Pre-Processing
   At this stage the data will be adjusted before starting the K-Means Clustering calculation process, the Pre-Processing processes carried out in this study include:
   - Data Cleaning Process
     In the data cleaning process is a process to eliminate empty data, data that is not needed or irrelevant (noise) that has no relationship to research.
   - Data Transformation Process
     In the data transformation process is the process of changing the form of data into the appropriate form as needed in the mining process. In this study, all data used is numeric so polynomial data must be transformed into numeric form.
3. Attribute Selection
   This selection is to determine which attributes are used, therefore an attribute selection stage is needed in order to know which attributes are relevant and have

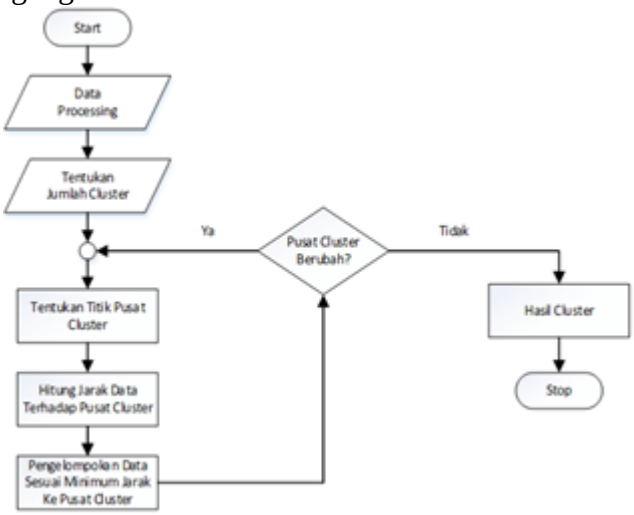
ISSN: 2502-8521 (Print) / 2503-1155 (Online)
DOI: https://doi.org/10.35457/josar.v9i1.3105

a relationship between fellow attributes. The following is attribute data that has been selected to be used as research data.

No	Data Awal	Atribut
1	Penghasilan Orang Tua	X1
2	Jumlah Saudara	X2
3	Memiliki Kartu Pra Sejahtera	X3
4	Jenis Surat Keterangan	X4

Picture 2.3 The attributes used

4. Application of K-Means Clustering
   The K-Means Clustering method is to group some data into groups that explain the data in one group has the same characteristics and has different characteristics from the data in other groups. Here is the process flow or flowchart of the K-Means Clustering algorithm:



Picture. 2.4 Flow Chart K-Means Clustering

Here are the steps of K-Means:

1. Determine K (the value is free) as the number of clusters to be formed.
   In this research, the value of K = 2, namely C1 students who are eligible to receive PIP and C2 students who are not eligible for PIP.
2. Generate a random value for the initial cluster center (centroid) as many as K.
3. Calculate the distance of each input data to each centroid using the Euclidean Distance formula until the closest distance of each data to the centroid is found.
   Here is the Euclidian Distance equation:

   $$\boldsymbol{d(xi, \mu j)} = \sqrt{\sum(\boldsymbol{xi} - \boldsymbol{\mu j})^2} \quad \text{..... (2.1)}$$

   xi : initial centroid point
   μj : many number of attributes
4. Classify each data based on its proximity to the centroid (smallest distance).
5. Update the new centroid values obtained from the average of the cluster concerned using the formula:

   $$\boldsymbol{\mu j\ (t+1)} = \frac{1}{Nsj}\sum_{j \in sj} \boldsymbol{xj} \quad \text{..... (2.2)}$$

   $\mu j\ (t+1)$: *centroid* baru pada iterasi ke-(t+1),

JOSAR, Vol.9 No.1 March 2024
ISSN: 2502-8521 (Print) / 2503-1155 (Online)
DOI: https://doi.org/10.35457/josar.v9i1.3105

Nsj : banyak data pada *cluster Sj*.

6. Repeat from steps 3 to 5, until the results of the members of each cluster have not changed (converged).

3. RESULTS AND DISCUSSION

3.1 Research Data

The research data here uses 7th grade student data for the 2022/2023 academic year which has gone through the cleaning process. The following is the data obtained:

Daftar Peserta Didik Tahun Pelajaran 2022-2023
UPT Satuan Pendidikan SMPN 9
Kecamatan Kec. Sukorejo, Kabupaten Kota Blitar, Provinsi Prov. Jawa Timur
Tanggal Unduh: 2023-01-02 09:41:10

No	Kelas	Nama Siswa	Penghasilan Orang Tua	Jumlah Saudara	Kartu Pra Sejahtera	Surat Ket Tidak Mampu
1	7 A	AKA DEEPSKY ALIYU HENANDRA	Rp.500,000-Rp.999,999	1 Saudara	KIS	Dinas Sosial
2	7 A	Amanda Dea Fitria	Rp.1,000,000-Rp.1,999,999	4 Saudara	KIS	Dinas Sosial
3	7 A	ARDIAN WAHYU PRATAMA	Rp.500,000-Rp.999,999	3 Saudara	KIS	Dinas Sosial
4	7 A	AURIKA ARUM MASHIROH	Rp.500,000-Rp.999,999	1 Saudara	PKH	Dinas Sosial
5	7 A	AYU CANTIKA	Rp.500,000-Rp.999,999	3 Saudara	KIS	Dinas Sosial
6	7 A	AYUN AURA SHELOW	Rp.500,000-Rp.999,999	2 Saudara	PKH	Organisasi/ Partai
7	7 A	AZAFI VIBRANEO MARZUQI ABDULLAH	Rp.500,000-Rp.999,999	1 Saudara	PKH	Dinas Sosial
8	7 A	DIMAS ADITYA DWI NUGROHO	Rp.500,000-Rp.999,999	1 Saudara	KIS	Dinas Sosial
9	7 A	Dzeeya Rizwana Saputro	Rp.500,000-Rp.999,999	1 Saudara	KKS	Dinas Sosial
10	7 A	ELISABETH KARTIKA	≥ Rp. 5,000,000	1 Saudara	Tidak Memiliki	Tidak Memiliki
11	7 A	Faiz Ridho Fergiano	Rp.1,000,000-Rp.1,999,999	1 Saudara	KIS	Organisasi/ Partai
12	7 A	FIRMAN AFIF ARDIANSYAH	Rp.500,000-Rp.999,999	1 Saudara	PKH	Dinas Sosial
13	7 A	Fitri Az Zahra Asfianto	Rp.1,000,000-Rp.1,999,999	1 Saudara	KKS	Kelurahan
14	7 A	MOHAMMAD UBAIDILLAH ZUHRI MUSTOFA	Rp.1,000,000-Rp.1,999,999	1 Saudara	KIS	Dinas Sosial
15	7 A	Muhammad Nabil Wafi	Rp.1,000,000-Rp.1,999,999	1 Saudara	PKH	Dinas Sosial

Picture 3.1 Research Data

3.2 Transformation of each data and attribute

The first step is to give a symbol to each attribute used to make it easier to read the data, in this study all attributes are likened to the X symbol,

No	Data Awal	Atribut
1	Penghasilan Orang Tua	X1
2	Jumlah Saudara	X2
3	Kartu Pra Sejahtera	X3
4	Surat Ket Tidak Mampu	X4

Picture 3.2 Attribute Data

Next, perform data transformation, which aims to change the form of the initial data into a suitable form as needed in the mining process.

Transformation of parent income data

No	Data Awal	Transformasi Data
1	Tidak Berpenghasilan	1
2	≤ Rp. 499.999	2
3	Rp. 500,000 - Rp. 999,999	3
4	Rp. 1,000,000 - Rp. 1,999,999	4
5	Rp. 2,000,000 - Rp. 4,999,999	5
6	≥ Rp. 5,000,000	6

Ket : Data on parents' income is transformed with a number scale of 1 to 6. Number 1 is for the smallest income while number 6 is for the largest income.

Data transformation of number of siblings

JOSAR, Vol.9 No.1 March 2024
ISSN: 2502-8521 (Print) / 2503-1155 (Online)
DOI: https://doi.org/10.35457/josar.v9i1.3105

No	Data Awal	Transformasi Data
1	≥ 5 Saudara	1
2	4 Saudara	2
3	3 Saudara	3
4	2 Saudara	4
5	1 Saudara	5
6	Anak Tunggal	6

Ket: The number of siblings is converted to a number scale of 1 to 6. The number 1 is for the largest number of siblings while the number 6 is for those with no siblings or only children.

Ownership Transformation (Pre-Prosperity Card)

No	Data Awal	Transformasi Data
1	Memiliki KPKH	1
2	Memiliki KKS	2
3	Memiliki KIS	3
4	Dalam Proses Pembuatan	4
5	Tidak Memiliki	5

Ket : Data on the ownership of pre-prosperity cards was changed to a scale of 1 to 5. 1 is for the highest priority card, PKH, while 5 is for those who do not have one at all.

Transformation of Certificate of Disability

No	Data Awal	Transformasi Data
1	Surat dari Dinas Sosial	1
2	Surat dari Partai	2
3	Surat dari Kelurahan	3
4	Surat dari Yayasan	4
5	Surat dari RT/RW	5
6	Tidak Memiliki	6

Ket : Data on the ownership of a certificate of incapacity is converted on a scale of 1 to 6. The number 1 is for the highest priority card from the Social Affairs Office while the number 6 is for those who do not have one at all.

Here are the overall results of the test data that has gone through the data transformation process.

JOSAR, Vol.9 No.1 March 2024
ISSN: 2502-8521 (Print) / 2503-1155 (Online)
DOI: https://doi.org/10.35457/josar.v9i1.3105

No	Nama Siswa	Penghasilan Ortu (X1)	Jumlah Saudara (X2)	Kartu Pra-Sejahtera (X3)	SKTM (X4)
1	Aka Deepsky Aliyu H	3	5	3	1
2	Amanda Dea Fitria	4	5	1	1
3	Ardian Wahyu Pratama	3	3	3	1
4	Aurika Arum Mashiroh	3	5	1	1
5	Ayu Cantika	3	3	3	1
6	Ayun Aura Shelow	3	4	1	2
7	Azafi Vibraneo Marzuqi	3	5	1	1
8	Dimas Aditya Dwi N	3	5	3	1
9	Dzeeya Rizwana Saputro	3	5	2	1
10	Elisabeth Kartika	6	5	5	6
11	Faiz Ridho Fergiano	4	5	3	2
12	Firman Afif Ardiansyah	3	5	1	1
13	Fitri Az Zahra Asfianto	4	5	2	3
14	Mohammad Ubaidillah	4	5	3	1
15	Muhammad Nabil Wafi	4	5	1	1
16	Putra Wahyu Purnawan	3	5	1	1
17	Radisty Okta Putri R	3	5	2	1
18	Ragil Prayoga	3	5	1	1
19	Rangga Adi Saputra	5	4	3	6
20	Revalina Kanaya Azzura	3	5	2	2
21	Revan Aditya Pratama	3	5	3	1
....	....	....	....	....	....
284	Yavik Vivian Efendi	3	5	1	1
285	Yofi Athaya Adiwitya	5	5	5	6
286	Zalfa Myrrabel Jonata	4	5	3	2

Picture 3.3 Data after the transformation process

1. Determining the Initial Centroid Center

In this study, 2 clusters were determined, namely:
Cluster Eligible to Receive PIP (C1)
Cluster Not Eligible to Receive PIP (C2)
Next, determine the cluster point or initial centroid which is taken randomly from student data. From all student data, here we take the initial centroid data, namely :

CENTROID	x1	x2	x3	x4
C1	1	5	3	1
C2	5	4	3	6

2. Calculating the distance from the centroid

The next step calculates the distance of all data or values to each initial centroid, using the eucludian distance formula. This calculation process is called iteration. The method is as follows:

Calculating the 1st student data with centroid 1

D1 = $\sqrt{(1-\mathbf{3})^2+(5-\mathbf{5})^2+(3-\mathbf{3})^2}+(1-\mathbf{1})^2$

D1 = $\sqrt{(-2)^2+(0)^2+(0)^2+(0)^2}$

D1 = $\sqrt{4+0+0+0=4}$

D1 = **2,00**

Calculating the 1st student data with centroid 2

D2 = $\sqrt{(5-\mathbf{3})^2+(4-\mathbf{5})^2+(3-\mathbf{3})^2}+(6-\mathbf{1})^2$

D = $\sqrt{(2)^2+(-1)^2+(0)^2+(5)^2}$

D2 = $\sqrt{4+1+0+25=30}$

D2 = **5,48**

JOSAR, Vol.9 No.1 March 2024
ISSN: 2502-8521 (Print) / 2503-1155 (Online)
DOI: https://doi.org/10.35457/josar.v9i1.3105

3. Calculating the 1st Iteration

The next step continues to calculate the distance for all students in the same way as the 1st student with the eucludian distance formula then we will get the results of the 1st Iteration.

No	Nama Siswa	X1	X2	X3	X4	D1	D2	Cluster
1	Aka Deepsky Aliyu H	3	5	3	1	2,00	5,48	1
2	Amanda Dea Fitria	4	2	3	1	4,24	5,48	1
3	Ardian Wahyu Pratama	3	3	3	1	2,83	5,48	1
4	Aurika Arum Mashiroh	3	5	1	1	2,83	5,83	1
5	Ayu Cantika	3	3	3	1	2,83	5,48	1
6	Ayun Aura Shelow	3	4	1	2	3,16	4,90	1
7	Azafi Vibraneo Marzuqi	3	5	1	1	2,83	5,83	1
8	Dimas Aditya Dwi N	3	5	3	1	2,00	5,48	1
9	Dzeeya Rizwana Saputro	3	5	2	1	2,24	5,57	1
10	Elisabeth Kartika	6	5	5	6	7,35	2,45	2
11	Faiz Ridho Fergiano	4	5	3	2	3,16	4,24	1
12	Firman Afif Ardiansyah	3	5	1	1	2,83	5,83	1
13	Fitri Az Zahra Asfianto	4	5	2	3	3,74	3,46	2
14	Mohammad Ubaidillah	4	5	3	1	3,00	5,20	1
15	Muhammad Nabil Wafi	4	5	1	1	3,61	5,57	1
16	Putra Wahyu Purnawan	3	5	2	1	2,24	5,57	1
17	Radisty Okta Putri R	3	5	1	1	2,83	5,83	1
18	Ragil Prayoga	3	5	2	1	2,24	5,57	1
19	Rangga Adi Saputra	3	5	1	1	2,83	5,83	1
20	Revalina Kanaya Azzura	5	4	3	6	6,48	0,00	2
21	Revan Aditya Pratama	2	5	2	2	1,73	5,20	1
....	....	....	....	....	....	....	....	....
284	Yavik Vivian Efendi	2	5	1	3	3,00	4,80	1
285	Yofi Athaya Adiwitya	5	5	5	6	6,71	2,24	2
286	Zalfa Myrrabel Jonata	4	5	3	2	3,16	4,24	1

Picture 3.4 Calculation results of iteration 1

In the 1st iteration calculation, the group members of cluster 1 and cluster 2 are obtained, namely:

C1 members: 201 (PIP Eligible Students)

C2 consists of : 85 (Students Not Eligible for PIP)

4. Calculating the new Centroid

After getting the results of the 1st iteration, the next step is to calculate the new centroid. To calculate the new centroid is by the way. Sum all data on each attribute X1 through X4 which are members of each cluster (C1 and C2). Next is to divide each attribute data that has been summed by the total number of cluster members.

Atribut Siswa	x1	x2	x3	x4
Jumlah Data Atribut x	616	967	487	287
Jumlah Anggota C1	201	201	201	201
Hasil Centroid Baru C1	3,06	4,81	2,42	1,43

÷

Atribut Siswa	x1	x2	x3	x4
Jumlah Data Atribut x	429	406	318	481
Jumlah Anggota C2	85	85	85	85
Hasil Centroid Baru C2	5,05	4,78	3,74	5,66

÷

So until this stage the new centroid has been obtained from the calculation of Cluster 1 and Cluster 2, namely :

CENTROID	x1	x2	x3	x4
C1	3,06	4,81	2,42	1,43
C2	5,05	4,78	3,74	5,66

5. Calculating the 2nd Iteration

JOSAR, Vol.9 No.1 March 2024
ISSN: 2502-8521 (Print) / 2503-1155 (Online)
DOI: https://doi.org/10.35457/josar.v9i1.3105

In this step, it is the same as the previous step, namely calculating the distance of the data value to the centroid using the Eucludian Distance formula but this time using the latest centroid data, then the result:
In the 2nd Iteration calculation, the member groups of cluster 1 and cluster 2 are obtained, namely:
C1 members: 199 (PIP Eligible Students)
C2 consists of : 87 (Students Not Eligible for PIP)

From these results the data is still not the same or convergent, so we do the calculation until the data does not change.

ITERASI	Cluster	Titik Centroid				Jumlah Anggota Cluster
		X1	X2	X3	X4	
Iterasi ke 3	C1	3,06	4,81	2,42	1,43	201
	C2	5,05	4,78	3,74	5,66	85
Iterasi ke 4	C1	3,06	4,81	2,42	1,43	201
	C2	5,05	4,78	3,74	5,66	85

Picture 3.5 Comparison Results of 3rd and 4th Iterations

After doing up to the 4th iteration, do the same comparison as before, namely comparing the calculation results of the 3rd iteration and the 4th iteration.

No	Nama Siswa	X1	X2	X3	X4	Layak/Tidak
1	Aka Deepsky Aliyu H	3	5	3	1	Layak
2	Amanda Dea Fitria	4	2	3	1	Layak
3	Ardian Wahyu Pratama	3	3	3	1	Layak
4	Aurika Arum Mashiroh	3	5	1	1	Layak
5	Ayu Cantika	3	3	3	1	Layak
6	Ayun Aura Shelow	3	4	1	2	Layak
7	Azafi Vibraneo Marzuqi	3	5	1	1	Layak
8	Dimas Aditya Dwi N	3	5	3	1	Layak
9	Dzeeya Rizwana Saputro	3	5	2	1	Layak
10	Elisabeth Kartika	6	5	5	6	Tidak Layak
11	Faiz Ridho Fergiano	4	5	3	2	Layak
12	Firman Afif Ardiansyah	3	5	1	1	Layak
13	Fitri Az Zahra Asfianto	4	5	2	3	Layak
14	Mohammad Ubaidillah	4	5	3	1	Layak
15	Muhammad Nabil Wafi	4	5	1	1	Layak
....	....	....	....	....	....	....
284	Yavik Vivian Efendi	2	5	1	3	Layak
285	Yofi Athaya Adiwitya	5	5	5	6	Tidak Layak
286	Zalfa Myrrabel Jonata	4	5	3	2	Layak

Picture 3.6 Results of List of Students Eligible and Not Receiving PIP

From the table above, both the number of members of each cluster C1 and C2 and the centroid point of each cluster are exactly the same. It can be concluded that the results of the 4th iteration have converged. So with this the iteration process has stopped and the desired results have been obtained:
C1 consists of : 201 (PIP Eligible Students)
C2 consists of : 85 (PIP Ineligible Students)

ISSN: 2502-8521 (Print) / 2503-1155 (Online)
DOI: https://doi.org/10.35457/josar.v9i1.3105

4. CONCLUSION

The results of this study indicate that the use of the K-Means Clustering Algorithm is very effective in grouping students into two groups, namely the group of students who are eligible to receive (PIP) which amounts to 201 students and the group that is not eligible to receive 85 students. So that the application of K-Means clustering provides a clear picture and makes it easier for schools to determine aid recipients based on the desired criteria.

5. ACKNOWLEDGEMENT

The author realizes that the preparation of this thesis will not be completed without the help of various parties. Therefore, on this occasion we would like to thank:

1. Dr. Drs. H. Soebiantoro, M.Si as the Rector of Balitar Blitar Islamic University.
2. Mr. Abdi Pandu Kusuma, S.Kom, M.T as the Dean of the Faculty of Information Technology at Balitar Blitar Islamic University.
3. Mr. Saiful Nur Budiman, S.Kom., M.Kom as Head of the Informatics Engineering Study Program at Balitar Blitar Islamic University.
4. Mrs. Sri Lestanti, S.Kom., M.T. as the main Supervisor for the guidance, advice, and motivation provided.
5. Mr. Rizki Dwi Romadhona, S.S.T., M.Tr.T. as the second Supervisor for direction and correction so that this Thesis Report can be completed.
6. Mr. Widada Adi, S.Pd., M.Pd. as the principal who has allowed the SMPN 9 Blitar Education Unit as a research site.
7. All Lecturers of Informatics Engineering Department of Balitar Islamic University who have provided their knowledge so that this study can be completed.
8. All Employees of Balitar Blitar Islamic University who have provided support for the completion of this Thesis Report.
9. Parents who never stop giving prayers and love and support given so far.
10. The big family of Informatics Engineering class 2019 employees, especially brothers Ahmad Farid Arwani, Inun Pujianto and Yusuf Pramono for all their support, enthusiasm, and cooperation.
11. All parties who have helped and provided input for the preparation of this thesis.

6. REFERENCES

Devi Tri Yuliana, M. Ivan Ariful Fathoni, Naning Kurniawati. (2022). Penentuan Penerima Kartu Indonesia Pintar KIP Kuliah dengan Menggunakan Metode K-Means Clustering.

Rini Gustini, RZ. Abdul Aziz. (2019). Pengembangan Model Pengambilan Keputusan Penerima Kartu Indonesia Pintar (KIP) Dengan Metode K-Means Dan Average Linkage Clustering (Studi Kasus : SMA Negeri 1 Kotagajah).

Yosep Filki. (2022). Algoritma K-Means Clustering dalam Memprediksi Penerima Bantuan Langsung Tunai (BLT) Dana Desa.

ISSN: 2502-8521 (Print) / 2503-1155 (Online)
DOI: https://doi.org/10.35457/josar.v9i1.3105

Darlinda, Joy Nashar Utamajaya. (2022). Sistem Pendukung Keputusan Penerima Beasiswa Program Indonesia Pintar Menggunakan Metode Algoritma K-Means Clustering.

Muktiadi, R. and A. Y. Badharudin (2020). "Metode K-Means untuk Mengelompokkan Beasiswa Sekolah." Sainteks 16(1).

Rahma, I., et al. (2019). "Penerapa Metode Hierarchical Clustering Dan K-Means Clustering Untuk Mengelompokkan Potensi Lokasi Penjualan Linkaja." Jurnal Informatika Polinema 6(1): 15-22.

Kumar, Sumit, Vijender Kumar-Solanki, Saket Kumar Choudhary, Ali Selamat, and Rubén Gonzalez-Crespo. (2020). "Comparative Study on Ant Colony Optimization (Aco) and K-Means Clustering Approaches for Jobs Scheduling and Energy Optimization Model in Internet of Things (Iot)." International Journal of Interactive Multimedia and Artificial Intelligence 6, no. 1.

Aditya, Agil, Ivan Jovian, and Betha Nurina Sari. (2020). "Implementasi K-Means Clustering Ujian Nasional Sekolah Menengah Pertama Di Indonesia Tahun 2018/2019." Jurnal Media Informatika Budidarma 4, no. 1.

Fauzi, A., & Musa, M. (2018). Implementasi Algoritma K-Means Clustering pada Sistem Pendukung Keputusan Penentuan Penerima Beasiswa Bidikmisi Menggunakan Bahasa Pemrograman PHP dan MySQL. Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK), 5(3), 227-234.