International Journal Of Health, Engineering And Technology (IJHET) *E-ISSN 2829 - 8683*
Volume 1, Number 2, July 2022, *Page. 208 - 216*
Email : editorijhess@gmail.com

Grouping Library Book Collection Based On Old Book Borrowers With Clustering Method (Case study: STMIK Kaputama)

Sumartika Br Hutasoit[1)], Budi Serasi Ginting[2)], Fuzi yustika manik[3)]
[1,2,3)]STMIK KAPUTAMA Binjai

*Coresponding Author
Email : tikahutasoiit@gmail.com

Abstract

Libraries are institutions that collect printed and recorded knowledge. Books can be borrowed at the library for a length of time according to library regulations. Based on observations that some types of books in the library of STMIK Kaputama have a very high ratio of the number of borrowers, while the availability of books is limited. The addition of new book collections cannot be done because the storage capacity of the collection is limited. Based on these conditions, the grouping of book collections based on the length of book borrowing is carried out to optimize the service time for borrowing books. This application was created to assist librarians in determining the optimal length of book borrowing. So in this case, we will design and build a system that will be used in grouping library book collections based on the length of borrowing and the variables determined using the clustering method. The purpose of this research is to design and build a system for grouping library book collections in order to produce information quickly about the availability of books in the library.

Keyword: Datamining, Library, K-Means Algorithm

INTRODUCTION

A library is an institution that collects printed and recorded knowledge. Books can be borrowed in the library with long borrowing according to library regulations. Based on the observation that most types of books in the library Stmik Kaputama has a very high ratio of the number of borrowers, while the availability of books is limited. The addition of a new collection of books can not be done because the capacity of the collection storage space is limited. based on these conditions, the grouping of book collections based on the length of book lending is carried out to optimize the service of book lending.

This application was created to assist librarians in determining the optimal length of book borrowing. So in this case will design and build a system that will be used in grouping library book collections based on the length of borrowing and variables determined by using the clustering method, in order to facilitate and shorten the time in providing fast and precise information to library visitors.

Based on the background of the above problems, the problems that can be formulated as follows:

1. How to analyze library book data based on book borrowing time using k-Means algorithm?
2. How to make an application that can provide information quickly about the data about borrowing library books?
3. How to apply k-Means algorithm in processing stack of data about library book borrowing?

Limitations made in this issue is as berkut:

1. The method used is clustering method with K-Means algorithm and software used MATLAB version 7.7.0 (2008).
2. Variables to be used are: Majors, types of books, long borrowed
3. The Data used is the data of 2016 - 2020

Based on the above explanation, the objectives of this study are as follows :

1. To determine the results of visitor data processing and generate library book collection information.

2. To determine the results of the application of the k-Means algorithm in the library book collection.
3. To design and build a system for grouping library book collections.

The benefits obtained from the preparation of this thesis is:

1. Can obtain information about the relationship of visitor data with the availability of book collections
2. Can generate information quickly about the availability of books in the library.
3. Can facilitate grouping (cluster) data by using clustering method and K-Means algorithm

RESEARCH METHODS

Definition Of Data Mining

Data mining is often known as the term used to find knowledge hidden in the database. Data mining is a process that uses statistical, mathematical, artificial intelligence and machine learning techniques to extract and identify useful information and related knowledge from various related databases. Meanwhile, according to Gorunescu (Prasetio, 2014, p.1) " Data mining is a job that utilizes data to be something important in various fields, ranging from academic, business, to medical"

Clustering

Clustering is a data analysis method that is often included as one of the data mining methods whose purpose is to Group data with the same characteristics. According to Eko Prasetyo (2012, h.178) " Clustering is the grouping of data (objects) based solely on the information found in the data that describes those objects and the relationships between them"

RESULTS AND DISCUSSION

Flowchart Design

The design process that will be built in the grouping of library book collections based on the length of borrowed books, student majors, types of books using the clustering method will have a flowchart description as follows :

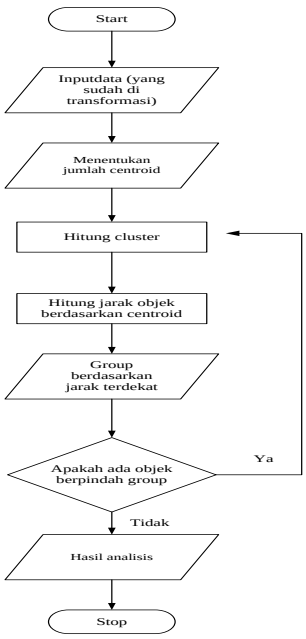

Figure 1. Flowchart System

Supporting Data

The Data used as supporting research adalahsebagai following :

Table 1. Book Borrower Data

No	Department	Types Of Books	Long Borrowed
1	SI	Friendter	2 Day
2	SI	Facebook	1 Day
3	SI	Algoritma data mining	3 Day
4	SI	Algoritma data mining	7 Day
5	MI	Adobeflash profesional cs 5	1 Day
6	MI	Analisis perekonimian	7 Day
7	MI	Analisis perekonimian	7 Day
8	SI	VB	3 Day
9	KA	Gaul Via Internet	6 Day
10	KA	Gaul Via Internet	6 Day
11	KA	Gaul Via Internet	6 Day
12	KA	Gaul Via Internet	6 Day
13	KA	Gaul Via Internet	6 Day
14	KA	Gaul Via Internet	6 Day
15	SI	Google Earth	2 Day
16	MI	Komdat	5 Day
17	SI	Komdat	4 Day
18	SI	Sukses Membangun toko Online	3 Day
19	SI	Power Point 2010	2 Day
20	SI	PHP my sql	5 Day

Application Of The Method

To determine the group of a single object, the first thing to do is measure the Deuclidean distance between two points or objects or X and Y which is defined as follows:

Deuclidean (X,Y) =√((X1-Y2)^2)

Tabel 2 Data Book Lenders.

No	Department	Types Of Books	Long Borrowed
1	SI	Friendter	2 Day
2	SI	Facebook	1 Day
3	SI	Algoritma data mining	3 Day
4	SI	Algoritma data mining	7 Day
5	MI	Adobeflash profesional cs 5	1 Day
6	MI	Analisis perekonimian	7 Day
7	MI	Analisis perekonimian	7 Day
8	SI	VB	3 Day
9	KA	Gaul Via Internet	6 Day
10	KA	Gaul Via Internet	6 Day
11	KA	Gaul Via Internet	6 Day
12	KA	Gaul Via Internet	6 Day
13	KA	Gaul Via Internet	6 Day
14	KA	Gaul Via Internet	6 Day
15	SI	Google Earth	2 Day
16	MI	Komdat	5 Day
17	SI	Komdat	4 Day
18	SI	Sukses Membangun toko Online	3 Day
19	SI	Power Point 2010	2 Day
20	SI	PHP my sql	5 Day

Initialization of major criteria

Here below is a table of initialization criteria majors students in Stmik Kaputama which can be seen in the table below:

Table 3. Initialization of department criteria.

Department	Code (X)	Description
1	Komputerisasi Akutansi	KA
2	Manajemen Informatika	MI
3	Sistem Informasi	SI

Initialize book type criteria

Here below is a table of initialization criteria of the type of book that can be seen in the table below:

Table 4. Book type initialization.

Code	Type Book (Y)
1	Adobeflash profesional cs 5
2	Algoritma data mining
3	Analisis perekonimian
4	Facebook
5	Friendter
6	Gaul Via Internet
7	Google Earth
8	Komdat
9	PHP my sql
10	Power Point 2010
11	Sukses Membangun toko Online
12	VB

Initialization of the borrower's old criteria

Here below is a table of initialization criteria old borrower which can be seen in the table below:

Table 5. Old Initialization Of The Borrower.

Code	Long Borrowed (Z)
1	1-3 day
2	4-5 day
3	6-7 day

In order for the above data can be processed by using the K-means clustering algorithm, the data of nonnominal data such as majors, types of books and borrowing time must be initialized first in the form of numbers.

Table 6. Data Transformation

No	**Department (X)**	**Types Of Books (Y)**	**Long borrowed (Z)**
1	3	5	1
2	3	4	1
3	3	2	1
4	3	2	3
5	2	1	1
6	2	3	3
7	2	3	3
8	3	12	1
9	1	6	3
10	1	6	3
11	1	6	3
12	1	6	3
13	1	6	3
14	1	6	3
15	3	7	1
16	2	8	2
17	3	8	2
18	3	11	1

19	3	10	1
20	3	9	2

Perhitungan data berdasarkan Algoritma K-means clustering

Iterasi 1

Centroid 1 = (3, 5, 1) taken from at random from the data 1
Centroid 2 = (3, 4, 1) taken from at random from the data 2
Centroid 3 = (3, 2, 1) taken from at random from the data 3

Description:

The value of the centroid is taken randomly from the data that has been transformed.

Parts A (3, 5, 1)

K=3, centroid 1 = (3, 5, 1), centroid 2 = (3, 4, 1), centroid 3 = (3, 2, 1)

Distance from C1 (X) = $\sqrt{(3-3)^2+(5-5)^2+(1-1)^2}$ = 0

Distance from C2 (Y) = $\sqrt{(3-3)^2+(5-4)^2+(1-1)^2}$ = 1

Distance from C3 (Z) = $\sqrt{(3-3)^2+(5-2)^2+(1-1)^2}$ = 3

Table 7 Results Of Iteration 1 Calculation

No	(X)	(Y)	(Z)	Distance from C1	Distance from C2	Distance from C3	Group
	X	Y	Z				
1	3	5	1	0	1	3	1
2	3	4	1	1	0	2	2
3	3	2	1	3	2	0	3
4	3	2	3	3.61	2.83	2	3
5	2	1	1	4.12	3.16	1.41	3
6	2	3	3	3	2.45	2.45	2
7	2	3	3	3	2.45	2.45	2
8	3	12	1	7	8	10	1
9	1	6	3	3	3.46	4.90	1
10	1	6	3	3	3.46	4.90	1
11	1	6	3	3	3.46	4.90	1
12	1	6	3	3	3.46	4.90	1
13	1	6	3	3	3.46	4.90	1
14	1	6	3	3	3.46	4.90	1
15	3	7	1	2	3	5	1
16	2	8	2	3.32	4.24	6.16	1
17	3	8	2	3.16	4.12	6.08	1
18	3	11	1	6	7	9	1
19	3	10	1	5	6	8	1
20	3	9	2	4.12	5.10	7.07	1

Description:

1. If the centroid 1 is smaller then the cluster results in Group 1.
2. If the centroid 2 is smaller then the cluster results in Group 2.
3. If the centroid 3 is smaller then the cluster results in Group 3.

Old Group: {0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0}
New Group : {1,2,3,3,3,2,2,1,1,1,1,1,1,1,1,1,1,1,1,1}

Group changes occur, then proceed to the next iteration.

For group 1 there are 14 data;
C1 1 = (3+3+1+1+1+1+1+1+3+2+3+3+3+3)/14 = 2.07
C1 2 = (5+12+6+6+6+6+6+6+7+8+8+11+10+9)/14 = 7.57
C1 3 = (1+1+3+3+3+3+3+3+1+2+2+1+1+2)/14 = 2.07
For group 2 there are 3 data;
C1 1 = (3+2+2)/3 = 2.33
C1 2 = (4+3+3)/3 = 3.33
C1 3 = (4+3+3)/3 = 2.33
For group 3 there are 9 data;
C1 1 = (3+3+2)/3 = 2.67
C1 2 = (2+2+1)/3 = 1.67
C1 3 = (1+3+1)/3 = 1.67

Iterasi 2

Centroid 1 = (2.07, 7.57, 2.07)
Centroid 2 = (2.33, 3.33, 2.33)
Centroid 3 = (2.67, 1.67, 1.67)

Part A (3, 5, 1)

K=3, C1 = (2.07, 7.57, 2.07), C2 = (2.33, 3.33, 2.33), C3 = (2.67, 1.67, 1.67)

Distance from C1 (X) = $\sqrt{(3-2.07)^2+(5-7.57)^2+(1-2.07)^2}$ = 2.94

Distance from C2 (Y) = $\sqrt{(3-2.33)^2+(5-3.33)^2+(1-2.33)^2}$ = 2.24

Distance from C3 (Z) = $\sqrt{(3-2.67)^2+(5-1.67)^2+(1-1.67)^2}$ = 3.42

Table 8. Calculation Results Iteration 2

No				Distance from C1	Distance from C2	Distance from C3	Group
	X	Y	Z				
1	3	5	1	2.94	2.24	3.42	1
2	3	4	1	3.84	1.63	2.45	2
3	3	2	1	5.75	2	0.82	3
4	3	2	3	5.72	1.63	1.41	3
5	2	1	1	6.66	2.71	1.15	3
6	2	3	3	4.67	0.82	2	2
7	2	3	3	4.67	0.82	2	2
8	3	12	1	4.65	8.79	10.36	1
9	1	6	3	2.12	3.06	4.83	1
10	1	6	3	2.12	3.06	4.83	1
11	1	6	3	2.12	3.06	4.83	1
12	1	6	3	2.12	3.06	4.83	1
13	1	6	3	2.12	3.06	4.83	1
14	1	6	3	2.12	3.06	4.83	1
15	3	7	1	1.53	3.96	5.39	1
16	2	8	2	0.44	4.69	6.38	1
17	3	8	2	1.03	4.73	6.35	1
18	3	11	1	3.71	7.81	9.36	1
19	3	10	1	2.81	6.83	8.37	1
20	3	9	2	1.71	5.72	7.35	1

Old Group: {1,2,3,3,3,2,2,1,1,1,1,1,1,1,1,1,1,1,1}
New Group: {1,2,3,3,3,2,2,1,1,1,1,1,1,1,1,1,1,1,1}

Because in the 1st and 2nd iterations the cluster position does not change or there is an equation, the iteration calculation is stopped.

CONCLUSION

Of the 20 data there are 3 groups, namely Group 1 there are 14 data and Group 2 there are 3 data and Group 3 there are 3 data. As for the explanation, there are 3 groups as follows:

Cluster 1 there are 14 Data with the value :

1. Based on the above calculation, it can be seen that in cluster 1 group of book collections based on the most borrowed length with majors (X) is MI (Information Management), and for the type of Book Group (Y) borrowed is Google Earth, and the length of borrowed books (Z) is 1-3 days.
2. Cluster 2 There are 3 Data with the value :
   Based on the above calculations can be seen that the cluster 2 is a collection of books based on the length of borrowing the least with the Department (X) is MI (Information Management), and for the type of Book Group (Y) borrowed is Economic Analysis, and the length of borrowed books (Z) is 1-3 days.
3. Cluster 3 there are 3 Data with the value :
4. Based on the above calculation, it can be seen that the cluster of 3 groups of book collections based on the length of borrowing with the Department (X) is MI (Information Management), and for the type of Book Group (Y) borrowed is Adobeflash professional cs 5, and the length of borrowing books (Z) is 1-3 days.

REFERENCES

Adi Nugroho. (2005), Analisis dan perancangan sistem informasi dengan metodologi berorientasi objek.Bandung: Informatika

B, Indra Yatini. (2010). Flowchart, Algoritma dan Pemrogaman Menggunakan Bahasa C++ Builder. Yogyakarta : Penerbit Graha Ilmu.

Enke, D. & Mehdiyev, N.. (2014). A Hybrid Neuro Fuzzy Model to Forecast Inflation. Procedia Computer Science. 36. (254-260).

Hend. (2006). Pengertian Unified Modeling Language (UML). Dipetik Oktober 05, 2015, dari http://adwintaactivity.blogspot.co.id/2012/04/definisiunified-modeling-language-uml.htML.

Maulia, S., Serasi Ginting, B., & Sihombing, A. (2021). IMPLEMENTASI DATA MINING PENGELOMPOKAN JENIS PENYAKIT PASIEN MENGGUNAKAN METODE CLUSTERING (STUDI KASUS : PUSKESMAS SAMBIREJO). *JIK)*, *5*(1).

Nugraha, D. D. C., Naimah, Z., Fahmi, M., & Setiani, N. (2014). Klasterisasi Judul Buku dengan Menggunakan Metode K-Means. *Seminar Nasional Aplikasi Teknologi Informasi (SNATI) Yogyakarta*, *21*(1).

Prasetyo, Eko. 2014. Data Mining. Yogyakarta: Andi Offset.

Prasetyo, Eko. (2013). Data Mining Konsep Dan Aplikasi Menggunakan Matlab. Yogjakarta: Penerbit Andi.

Santosa, B. 2007. Data Mining Terapan dengan MATLAB. Graha Imu. Yogyakarta

Silalahi, N. (2020). Penentuan Strategi Promosi Universitas Budi Darma Menggunakan Algoritma

K-Means Clustering. *Terapan Informatika Nusantara*, *1*(1), 40–46.

Yahyono. Yoyo. 2007. Makalah Layanan Perpustakaan dan Informasi (Layana Bahan Pustaka). Jakarta : Perpustakaan Nasional RI.