ALGORITHM K-MEANS ANALYSIS TO HELP MAJOR OF STUDENT AT ISLAMIC SENIOR HIGH SCHOLL COUNTRY BINONG SUBANG

Arim Wijaya

Universitas Komputer Indonesia
Jl. Dipati Ukur No.112 Bandung 40132
Email : arim.wijaya@yahoo.co.id

ABSTRAK

Islam senior high school country Binong, is one of the Islam senior high school country at Subang that conducted the majors students of class X (ten) who will continue the study to class XI (eleven). Constraints are often found in the process majors at Islam senior high school country Binong the difficulty determining where students who meet the criteria to occupy a specific majors. This is because the majors process are still done manually, thus requiring time-consuming and inefficient.

This research will be appointed a case that is analyzing the K-Means algorithm for assign majors students at Islam senior high school country Binong Subang. K-Means algorithm can segment (segments) that have the attribute data and have the amount of data that a lot, so it can be used in determining the student who majors in accordance with students' academic abilities.

Based on the analysis of K-Means algorithm for assign majors student, then the conclusion can be drawn is K-Means algorithm is less precise for decision support systems majors but the K-Means algorithm is more appropriate to categorize data based on the data value of students who can give an idea to student majors
Keyword : Analysis, K-Means, Majors, Value

1 INTRODUCTION

1.1 Probem Indentification

The difficulty of determining which students are eligible to occupy a specific majors.

1.2 Mean and Purpose

The purpose of this thesis is to whether the K-Means algorithm can help the student majors in MAN Subang Binong

The proposed objective is to test whether or not proper k-means algorithm in helping students' majors.

1.3 Benefits

The benefits in anlisis k-means algorithm in helping these students' majors are able to provide images for student majors

2. MODEL, ANALYSIS, DESIGN AND IMPLEMENTATION

2.1 Model

1. Stage of data collection
   a. Book study
   b. Field studies
      b.1 Interview
      b.2 Observasition
2. Stage of software development.
   a. *System engineering*
   b. *Requirement analysis*
   c. *System Design*
   d. *System Coding*
   e. *System Testing*
   f. *System Maintenance*

2.2 Problem Identification

Based on analysis and interviews on what is needed in assessing k-means algorithm is then to be evaluated the following: The decision making process on MAN Binong majors students are still done manually, ie by collecting the value of each semester from X to the homeroom teacher guidance counseling (the majors) and then they analyze and calculate the results and match it with a specific department criteria. It is difficult for teachers'guidance and counseling (the majors) MAN Binong in students' majors and require considerable time to analyze it.

2.3 Algorithm *K-Means* Analisys

K-Means clustering partitioning is included in every data should be entered in a particular cluster and allows for any data

that is included in a particular cluster at any stage of the process, at a later stage move to another cluster. K-Means to separate the data into separate parts of the region k, where k is a positive integer. K-Means algorithm is very famous for its ease and its ability to classify large data very quickly and outlier

Here are the steps of K-Means algorithm:

1. Determination of initial cluster centers

In determining the initial cluster centers n is the generation of random numbers that represent the input data sequence. Initial cluster centers obtained from the data itself rather than by defining a new point, namely the central menrandom beginning of the data.

2. The calculation of the distance to the cluster center

To measure the distance between cluster center used data with Euclidean distance. Distance calculation algorithm of data with the central cluster

1) Take the data value and the value of cluster centers
2) Calculate the Euclidean distance data with each cluster center

3. Grouping data

The distance calculation will be done comparisons and selected the closest distance between the data with the central cluster, this distance indicates that these data are in one group with the nearest cluster center.

Data clustering algorithm

1. Take the distance from the center of each cluster with the data
2. Find the smallest distance value
3. Group data with a central cluster that has the smallest distance..

4. Determination of a new cluster center

To get a new cluster center can be calculated from the average value of cluster members and cluster centers. The new cluster center is used to perform the next iteration, if the results obtained have not converged. Iteration process will stop if it meets the maximum iteration is entered by the user or the results achieved are convergent (the new cluster center equal to the old cluster center)**.**

New cluster center determination algorithms

1. Find the number of members of each cluster
2. Calculate the new center with the formula

   New cluster center = x1 + x2 + x3 + ... + xn + xp / Juml + 1

   Where:

   x1, x2, x3, ... xn = cluster members

   xp = old center interviews, and Psikotest and a crisp output of the department..

1. Composition Rules

   In determining fuzzynya general inference can be written as follows:

   μf (xi) = max (μf (xi), μk (xi))

   Description:

   μf (xi): the fuzzy membership value to the rule to - i

   μk (xi): membership value of fuzzy rule consequent to - i

2. The assertion (defuzzyfication)

Defuzzyfication process can be written as follows:

$$Z = \frac{\sum_{K=1}^{M} \mu_f X_i}{\sum_{K=1}^{M} \mu_f}$$

Description:

Z = value - average weighted

μf = fuzzy membership value as a result of the composition rules

Xi = value domain

2.3.1 Proses *Clustering* Algorithm *K-Means*

At this phase will be the main process of segmenting the data value accessed from a database which is a method of K-Means clustering algorithm. Here's a flowchart diagram of K-Means algorithm with the assumption that the input parameter is the number of data sets of n data and the number of initial centroids K = 2 corresponds to the number of existing majors in MAN Binong of Science and Social Science

Figure 1 is a picture of K-Means process flowchart for the application based on the value determination majors semester.

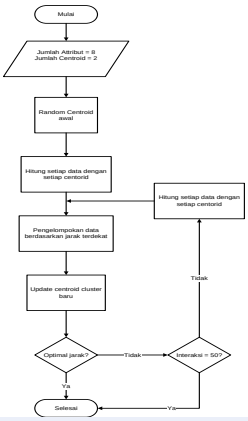


Figure 1 Flow Chart Process K-Means Algorithm

2.4 Desain

1. *Entity Relationship Diagram* (ERD)

ERD can be illustrated in Figure 2.

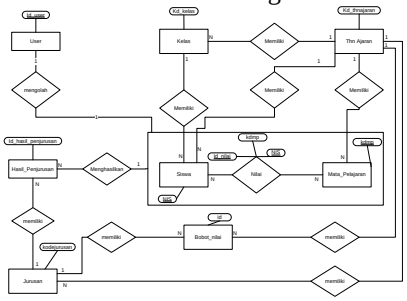


Figure 2. *Entiry Relationship Diagram*

2. Table Relasi

Relation schema analysis of k-means algorithm to help the majors can be seen in figure 3.

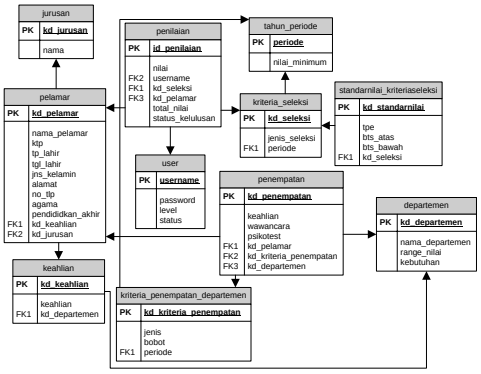


Figure 3. Relation Schema

3. Context Diagram

Context diagram can be seen in figure 4

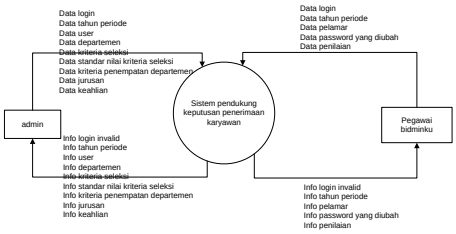


Figure 4. Context Diagram

4. *Data Flow Diagram* (DFD)

DFD can be seen in figure 5

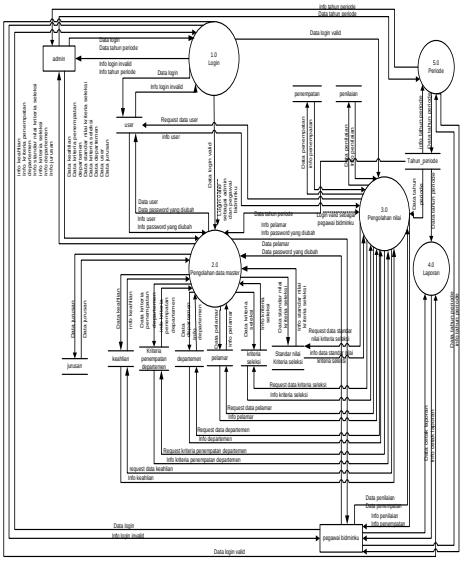


Figure 5. DFD Level 1

5. Perancangan Antarmuka

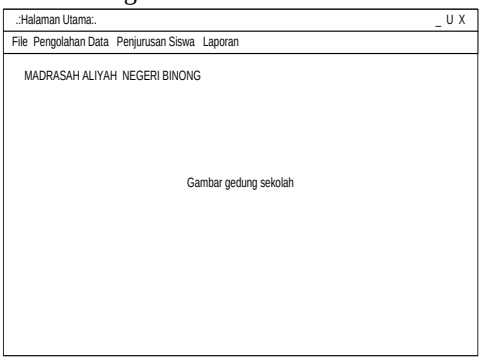


Figure 6. Antarmuka Menu Utama

1.4 Implemention

Implementation phase is done after doing the analysis stage of design in the system

1. Implementation hardware

The hardware needed to implement software from can be seen in table 1.

Table 1 The hardware used

Perangkat	Server	Client
Processor	Kecepatan 800 MHz	Kecepatan 500 MHz
RAM	1 GHz	512 MHz
Harddisk	80 Gb	80 Gb
Monitor	Resolusi 1024 x 768	Resolusi 1024 x 768
VGA	128 Mb	128 Mb
Lan Card	10/100 Mbps	10/100 Mbps

2. Implementation software

The software needed to implement software from can be seen in table 2.

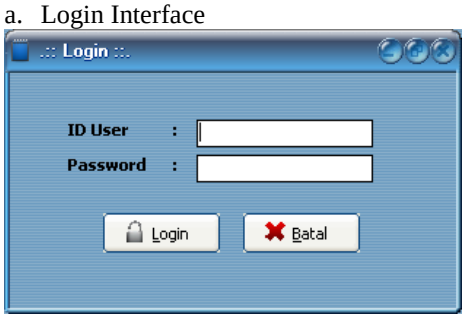
Table 2 The software used

Perangkat Lunak	Keterangan
Operation Sistem	Windows XP Profesional SP III
Programming Languages	Delphi 7
Database Server	MySQL 5.0.51b

3. Implementation interface

Implementation interface is as follows.

a. Login Interface



b. Views main page



3. RESULTS AND DISCUSSION

Based on the results of interviews with employees of the majors can be concluded that a system built to help the majors

4 CONCLUSIONS AND RECOMMENDATIONS

2.1 Conclusions

After doing the analysis, design, implementation and testing, then we concluded that the student majors in MAN Binong Subang using K-Means algorithm will take up a lot of time, so that K-Means algorithm less than optimal in the majors to help students in MAN Binong Subang. K-Means algorithm is more suitable for determining the majors of students who have close proximity to the center of the cluster, but for students who have long distances to the clusters easier to use manual processes majors.

2.2 Recommendation

Based on the above conclusion, it is expected in the future is in the process of student majors in MAN Subang Binong selectable algorithms / methods are more suited to the concept of the student majors in MAN Binong Subang, so that the process of majors in MAN Subang Binong.

5 BIBLIOGRAPHY

[1] Agus, M, (2002), *Belajar Sendiri Borland Delphi 7*, Jakarta : Exex Media Komputido.

[2] Agusta, Yudi. Pebruari 2007. "*K-Means Penerapan, Permasalahan dan Metode Terkait*". Jurnal Sistem dan Informatika Vol.3 : 47-60.

[3] Barakhbah, Ali Ridho. 2006. *Workshop Data Mining - Cluster Analysis.*: PENS-ITS Surabaya.

[4] Forta, Ben, (2000), *Belajar Sendiri SQL*, Yogyakarta :Andi

[5] Fathansyah (2007), *Basis Data*, Informatika, Bandung.

[6] http://www-2.cs.cmu.edu/~awm/tutorials/kmeans.html pada tanggal 21 Maret 2010 Pukul 21.15WIB

[7] Jogiyanto, HM. (2005), *Analisis dan Desain Sistem Informasi: Pendekatan*

Terstruktur Teori dan Praktek Aplikasi Bisnis, Andi, Yogyakarta.

[8] Rismawan, Tedy. Juni 2008 . "*Aplikasi K-Means untuk Pengelompokan Mahasiswa Berdasarkan Nilai Body Mass Index (BMI) & Ukuran Kerangka*". Jurnal Seminar Nasional Aplikasi Teknologi Informasi ISSN: 1907-5022

[9] Prayitno (2009), *Penentuan Bidang Konsentrasi Studi Tugas Akhir Berdasarkan Nilai Matakuliah Dengan Klasterisasi Kmeans,* Proyek Akhir, PENS-ITS Surabaya

[10] Pressman, Roger S. (2002), *Rekayasa Perangkat Lunak: Pendekatan Praktisi Buku 1*, Andi Offset, Yogyakarta.