An Analysis of Students' Learning Interest in Programming Language Based on Data Mining with Fuzzy C-Means Method

https://doi.org/10.3991/ijoe.v17i09.24595

Sitti Aisa(✉), Asmah Akhriana, Ahyuna Ahyuna, Andi Irmayana, Irmawati Irmawati, Nurul Aini
Dipa Makassar University, Makassar, Indonesian
Sitti.aisa@dipanegara.ac.id

Abstract—Students' learning interest and motivation in a programming language to date are determined by their scores and abilities to create applications. However, this is not sufficient to identify students' learning interest in programming language because some students got low scores. This study aims to identify students' learning interests in Dipanegara school of informatics management and computer (STMIK), Makassar, in Java programming language. The samples were 65 technical information students and 63 information system students. The data collection technique of this study was questionnaire and processed with data mining technique. Additionally, the Fuzzy C-Means clustering method was applied to Java Netbeans programming language to classify the level of students' interest in studying Java programming language. The result of the study was a web-based application that could determine students' learning interests. It was obtained through questionnaires and resulted as follows: 47 students had high learning interest, 45 students had moderate interest, and 36 students had fair interest out of 128 samples.

Keywords—data mining, clustering, fuzzy c-means, programming

1 Introduction

Programming language is a vital component of computer devices. The levels of programming language are machine code, low-level language, middle-level language, and high-level language. The higher the programming or computer language, the closer it is to the language applied by humans to communicate daily.

Dipa Makassar Of University is an Information Technology (IT) college where graduates could create programs or applications. The programming language is a compulsory and essential course since it is the communication media between programmer*s* and computer machines. Without programming language, we could not create a program or application that could help us with our works. Therefore, it is important to check out students' interest in learning a programming language.

Student learning motivation and interest in programming language courses can be identified through their scores and ability in making applications. However, this is not

really helpful for programming language interest since some students got the unsatisfying score for this course. Due to the problems above, the research problems are formulated as follow: (1) how to measure programming language learning interest, (2) how to determine the indicators affecting the programming language learning interest of students in Dipa Makassar Of University, and (3) how is the level of students' learning interest, especially java programming language by using the fuzzy c-means method for measurement.

This research selects some factors affecting students' learning interest in the programming language, i.e., internal and external factors. The internal factors are the physical factor that includes physical health while attending learning process, and psychological factors such as students' attention and hobbies. In comparison, the external factors, on the other hand, include campus facilities, lecturers' learning method or the behaviour of the lectures, family, peers, and involvement in the Students Activity Unit (SAU). The student questionnaires were then analyzed to find the most affecting factor on students' learning interest in the programming language. The data that had been processed with the clustering method, data mining, and Fuzzy C-Means algorithm was used to cluster students according to certain characteristics. They were grouped based on the quality rate, whether their learning interest was very good, good, or fair based on the analysis results of the previous factors in certain scores.

2 Literature Review

2.1 Learning

Witherington states that learning is an improvement within personality which states oneself as a new pattern of ability, behaviour, habit, intelligence, or a concept. Learning is a process or effort done by each individual to achieve a behaviour improvement, whether in the form of knowledge, skills, attitude, and positive value as an experience in achieving a certain sense from what is learned. Learning can be anywhere and anytime, depending on the individual's ability to penetrate the circumstances of the environment. Learning is a complex individual's action and behaviour. As an action, therefore, learning is experienced only by oneself and will determine the existent or inexistent of a learning process [1].

2.2 Interest

According to the Great Dictionary of the Indonesian Language, the word interest means heart's high tendency upon something. In short, interest means high tendency and passion or a great desire of something. States that "Interest is a preference and attraction to something or an activity, without being asked to." From the abovementioned opinions, it can be concluded that interest is one's fondness or being attracted to something or activity that motivates this person to take actions. Interest is also defined as a desire that comes from one's motivation to perform an activity without the force of others [2].

2.3 Programming Language

Programming language is a computer language used in writing programs." Programming language is a collection of certain meaningful and structured commands (syntax and semantic) that can be recognized by the computer and is useful in solving problems. For this reason, the programming language is divided into 4 (four) levels [3], which are:

1. Machine Language
2. Low-Level Language
3. Middle-Level Language
4. High-Level Language

2.4 Data mining

Data mining is a process of finding meaningful relations, patterns, as well as tendency and examine a great amount of data saved in storage by using pattern identification technique such as statistical and mathematical techniques.

Another definition of data mining is a term used to define the knowledge findings in a database. Data mining is a process that uses a statistical technique, mathematical technique, artificial intelligence, and machine learning to extract and identify useful information and knowledge related to the various right databases. Based on the definition above, it can be concluded that data mining is a process of information extraction and identification from a great database using the statistical technique, artificial intelligence, and machine learning [4].

2.5 Fuzzy C-Means

Clustering algorithm will be enforced in this study by using the Fuzzy C-means algorithm. This tool can be used to address its problem on clustering high dimensional datasets. Figure 1 shows the actual process of how Fuzzy C-Means Clustering works[5].

Fuzzy C-Means was first proposed by Dunn in 1973 and was renewed by Bezdek in 1981. This algorithm is one of the popular soft clustering techniques by implementing the data point approach, in which the centre point of the cluster will always be renewed with membership value from the existing data [6]. Besides, the fuzzy c-means algorithm is also an algorithm that works with a fuzz*y* model that enables all the data from the entire group member formed with different degree of membership between 0 and 1[7].

The basic concept of Fuzzy C-Means (FCM) is determining the cluster centre, which would mark the average of each cluster. The cluster centre is still inaccurate at the initial condition. Each data point has a degree of membership for each cluster. By repetitiously fixing the cluster center and degree of membership of each data point, it can be seen that the cluster center will move toward the right location. This repetition is based on objective function minimization that describes the distance from data

point given to cluster center which weighted by the degree of membership of the data point [8].

The Fuzzy C Means algorithm splits data to fuzzy groups with respect to some particular criterion. In FCM each data item belongs to two or more clusters [9].

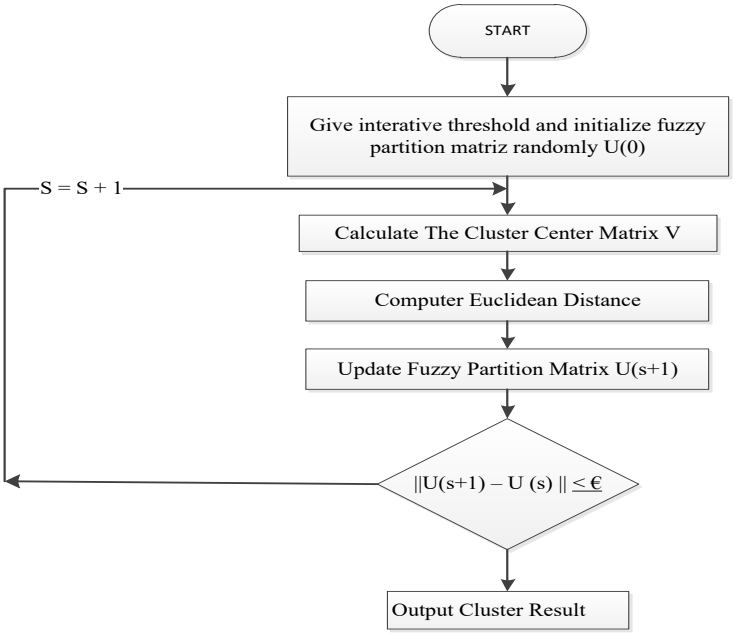

Fig. 1. The actual process of clustering using Fuzzy C-means algorithm

2.6 Related Work

Several studies using the Fuzzy C Means Method. Among other things, this article is in the form of Sucheta V. Kolekar et al. wrote an article entitled "Prediction of Learner's Profile Based on Learning Styles in Adaptive E-learning System" which in their paper tells about how students learn by capturing behavior in an e-learning portal, where the data is grouped with the fuzzy c-means algorithm and compares the results between Gravitation Search Algorithm and Back Propagation Neural Network Model [10]. Sumaya Hamad et al friends wrote a paper with the title "Clustering and Analysis of Dynamic Ad Hoc Network Nodes Movement Based on FCM Algorithm" published in 2017. It discusses the FCM algorithm used for simulation implementation and evaluation. Data set use NS2 simulator with optimized AODV protocol. That result from the algorithm application shows the maximally achieved technique Stability values for both cluster centre and node [11].

Abdullah W. Al Mutairi et al research entitled "IoT-Based Real-Time Monitoring System for Epidemic Diseases Patients: Design and Evaluation", in which this paper introduces a real-time monitoring system for monitoring pneumonia patients that allows doctors to monitor the health of their patients remotely. Remotely via a

smartphone or internet-connected device. To verify the functioning of this system, a real-time monitoring device was developed. The secure mechanism is designed to establish a wireless connection to the monitoring unit [12].

K.Geetha in his research entitled "Enhanced Web Log Data Mining using Probability Density-Based Fuzzy C Means Clustering" aims to Cluster into groups is one of the most important tasks in the action-bound areas of net record mining. It says without any doubt to grip the trouble of news given over-weight on the net while many users are connected on the meeting thing by which something is done. Clustering into groups is made use of for grouping news given into by comparison way in design for making the discovery of person for whom one does work interest. There are two bad points of the FCM algorithm, firstly the requirements of no. of clusters C and secondly giving to the first relation matrix. Because of, concerning these two bad points the FCM algorithm is hard to come to a decision about the right no. of mass, group and this algorithm is unsafe. The strong decision of desirable first stage mass, group is an important hard question, therefore a new expert way called PDFCM algorithm is made, was moving in [13].

3 Material And Methods

3.1 Research Method

This research applied observation and literature methods. The observation was done by collecting students' data and final score of java programming course from IT, information system, and informatics management study programs. The literature method was done by reading books that contain the Fuzzy C-Means method. Additionally, as well as reading the published papers discussing the Fuzzy C-Means method in various cases, such as making web-based applications to perform TOEFL test simulation and calculating the numbers and types of vehicles [13].

3.2 Application Architecture

Figure 2 is illustrates the constructed application architecture where two people are involved in the process. The first person is the user, which is a student as a sample, also the administrator who would witness the process of fuzzy c-means method implementation to assess the results of questionnaires filled by students that later will be stored in a database. A formulation is input inside to process all questionnaires.

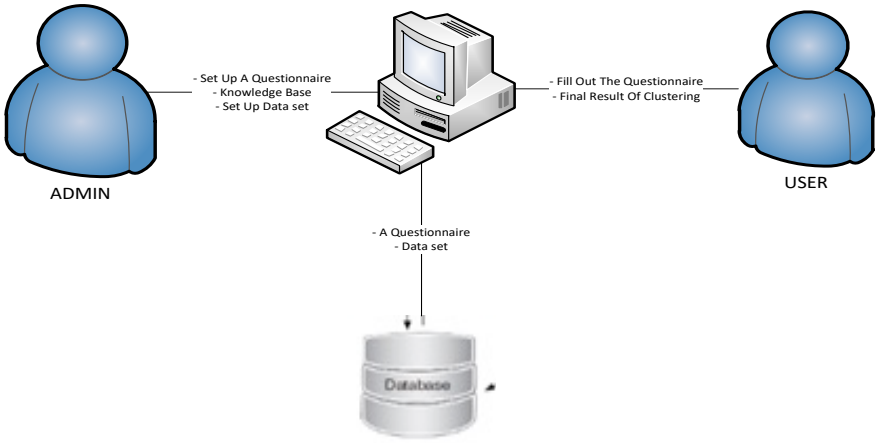

Fig. 2. Application Architecture

3.3 Systems Design

Use Case Diagram. Use Case Diagram describes the functionality expected from a system, it also describes the system requirement from users' point of view [14].

Figure 3 shows the system of use case diagram that consists of two actors, i.e., administrator and student. The use case above describes a system in which the administrator holds the most crucial role. An administrator could log in, enter the main menu, input questions or manage questionnaires and dataset. Students then log in to the main menu and fill in the questionnaire. After the students fill in the questionnaire, the fuzzy c-means method is applied to process the input questionnaire data, and then the clustering result is displayed after being processed.

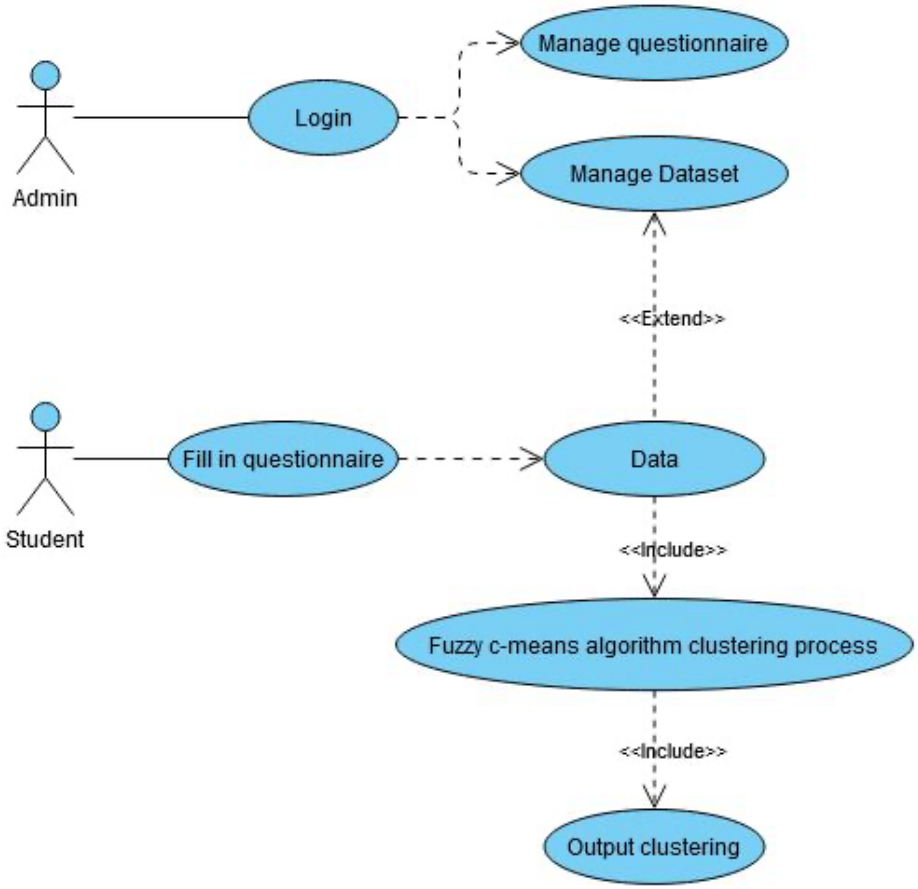

Fig. 3. Use Case Diagram

Sequence Diagram. Figure 4 shows how the sequence diagram describes the process in the system. The process starts with students filling out questionnaires, the administrator then processes and stores it in the dataset. The next step is calculation, it is done by generating random numbers as the initial degree of membership, calculating cluster center and objective function, then update the random numbers or degree of membership until it meets the requirements. The last process is application shows the clustering result.

Activity Diagram. Figure 5 shows class diagram, which is an object description representing the system, used as the basic in creating system in programming and database that consists of several class, i.e., Dataset, Random Numbers, Cluster Center, Objective Function, Random Number Update, and Clustering Result.

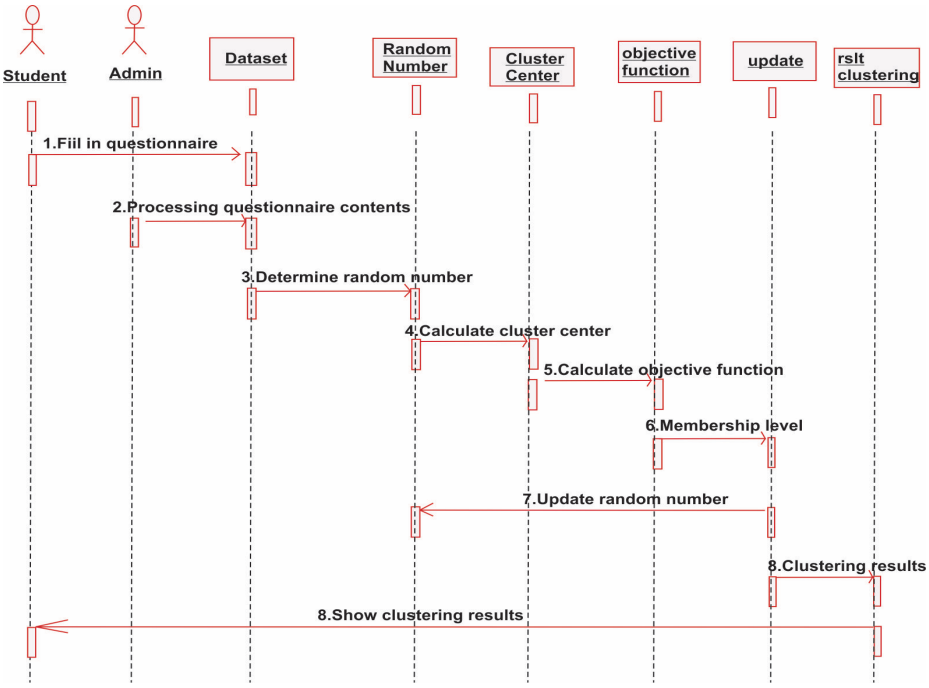

Fig. 4. Sequence Diagram

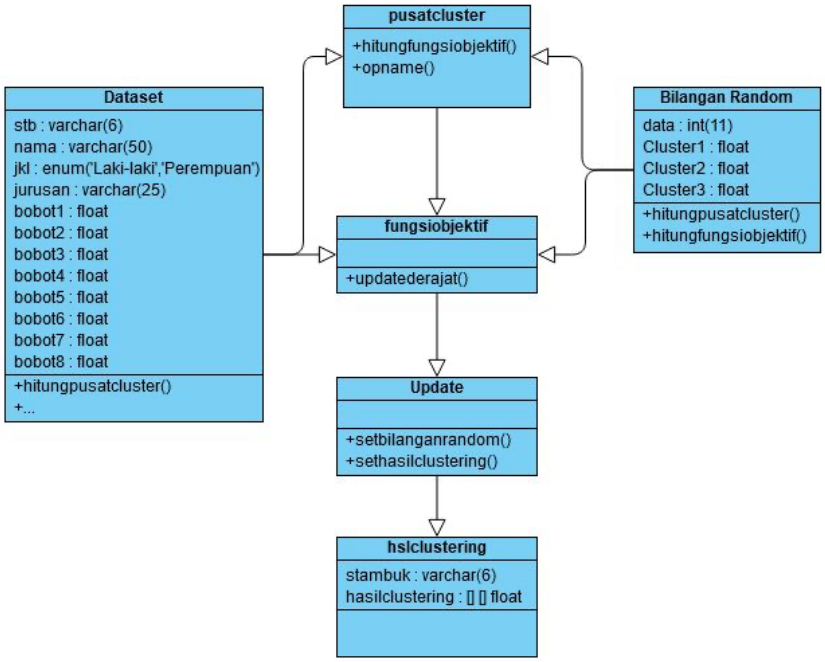

Fig. 5. Class Diagram Application

4 Results And Finding

4.1 The Implementation of Fuzzy C-Means Method

If a fuzzy clustering technique is to be employed in an application-oriented task to significantly cluster groups of objects, then the first aspect to encounter is assessing clustering tendency [15].

This is a desktop application that uses java NetBeans language. The applied method, which later would be inserted into java Netbeans programming is the Fuzzy C-Means method. The steps of data processing by applying Fuzzy C-Means method.

The questionnaire consists of 30 questions; it is 5 (five) optional questions that come with point 1-5.

- Question 1 – 4 = First factor
- Question 5 – 8 = Second factor
- Question 9 – 12 = Third factor
- Question 13 – 16 = Fourth factor
- Question 17 – 19 = Fifth factor
- Question 20 – 22 = Sixth factor
- Question 23 – 26 = Seventh factor
- Question 27 – 30 = Eighth factor

The score average of the questions would be calculated and set as the score of each factor. The score of each factor would be the attribute which then would be processed with Fuzzy C-Means method.

The steps of Fuzzy C-Means in processing the data

1. Determine matrix A_{5x8}

Number of students (sample) = 128
Number of attributes (factor)= 8
Information table 1:
Xi1, Xi2, Xi3, Xi4, Xi5, Xi6, Xi7, Xi8: Attribute data 1 to 8
Table 1 contains the questionnaire results filled by the students as sample from three study programs which are IT, information system, and informatics management.

Table 1. Attribute Score

id Number	Attribute (Factor)							
	Xi1	***Xi2***	***Xi3***	***Xi4***	***Xi5***	***Xi6***	***Xi7***	***Xi8***
132024	3	3	2	3	4	2	3	3
131036	3.25	3	4	3.25	3.667	3.667	3	4.75
132280	3	2	2	2	3	3	2	3
132314	3.5	2.75	3	3.75	3.33	3.33	3	3.5

2. Determining

Number of cluster (c) = 3
Weight (w) = 2
Maximum Iterations (MaxIter) = 40
Error (e) = 0.01
Objective function (P0) = 0
Initial Iteration (iter) = 1

3. Generating random numbers with matrix μ_{128x3}

Number of data (sample) = 128
Number of clusters = 3
Information table 2:
C1, C2, C3: cluster 1 to 3
Ui1, Ui2, Ui3: degree of membership 1, 2, and 3
Table 2 displays the results of value randomization from the matrix in which the number of samples is 128 students with three clusters.

Table 2. Matriks μ_{128x3}

id Number	Random Number						
	µi1	***µi2***	***µi3***	***Total***	***µi1^2***	***µi2^2***	***µi3^2***
132024	0.27971419	0.40285384	0.31743198	1	0.07824003	0.16229121	0.10076306
131036	0.2851273	0.39371607	0.3211312	1	0.08131208	0.15501235	0.10312524
…….	…….	…….	…….	…….	…….	…….	…….
132280	0.17168329	0.19524075	0.63307596	1	0.02947515	0.03811895	0.40078517
132314	0.21315545	0.45424152	0.33260302	1	0.04543525	0.20633536	0.11062477

4. Calculate Cluster Center

With formula:
Description:
Vkj: cluster center
Ui2: degree of member 2
Ui1: degree of member 1
Ui3: degree of member 3
Xi1, Xi2, Xi3, Xi4, Xi5, Xi6, Xi7, Xi8: attribute data 1 to attribute 8
Table 3 provides the results of matrix *U* power for each column and would be applied on multiplication with matrix *X*. The score of Σ is resulted from the addition of each column that will be used to calculate the center of cluster 1, 2, and 3.
Table 4 below is the multiplication results of column $\mu i1w$ in Table 3 and each column on matrix *X*, and the score of Σ is the results of adding up each column. The value of Σ will be applied in calculating the cluster center 1.

Table 3. Cluster Center Squared Random

Id Student	µi1	µi2	µi3	µi1^2	µi2^2	µi3^2
132024	0.27971	0.402854	0.317432	0.07824	0.162291	0.100763
131036	0.28515	0.393716	0.321131	0.08131	0.155012	0.103125
.....	.....	.....	.....	.....	.....	.....
132280	0.17168	0.195214	0.633076	0.02948	0.038119	0.400785
132314	0.21316	0.454242	0.332603	0.04544	0.206335	0.110625
∑				15.966	19.20046	18.01107

Table 4. Weight* Squared_Random 1

Xi1*µi1^2	Xi2*µi1^2	Xi3*µi1^2	Xi4*µi1^2	Xi5*µi1^2	Xi6*µi1^2	Xi7*µi1^2	Xi8*µi1^2
0.23472008	0.23472008	0.15648005	0.23472008	0.312960102	0.15648005	0.23472008	0.23472008
0.26426426	0.24393624	0.32524832	0.26426426	0.298171398	0.29817140	0.24393624	0.38623238
.....	.....	.....	.....	.....	.....	.....	.....
0.08842555	0.05895031	0.05895031	0.05895031	0.088425461	0.08842546	0.05895031	0.08842546
0.15902337	0.17038218	0.13630574	0.17038218	0.151435681	0.15143568	0.13630574	0.15902337
48.2349221	54.5926775	54.710118	56.1855771	55.81241031	51.0353363	53.0433083	52.2920562

Table 5 below is the multiplication results of column $\mu i2w$ in Table 3 and each column on matrix X, and the score of Σ is the results of adding up each column. The value of Σ will be applied in calculating the cluster center 2 result.

Table 5. Weight* Squared_Random 2

Xi1*µi2^2	Xi2*µi2^2	Xi3*µi2^2	Xi4*µi2^2	Xi5*µi2^2	Xi6*µi2^2	Xi7*µi2^2	Xi8*µi2^2
0.48687364	0.48687364	0.32458243	0.48687364	0.64976486	0.32458243	0.48687364	0.48687364
0.50379013	0.46503704	0.62000494	0.50379013	0.568430275	0.56843027	0.46503704	0.73630865
.....	.....	.....	.....	.....	.....	.....	.....
0.11435685	0.0762379	0.07623790	0.07623790	0.114356846	0.11435685	0.07623790	0.11435685
0.72217376	0.7737576	0.61900608	0.77375760	0.687715752	0.68771575	0.61900608	0.7221738
60.65534950	65.6471129	63.81891720	65.47885510	67.94196823	61.4750588	61.0750051	62.5549680

Table 6 below is the multiplication results of column $\mu i3w$ in Table 1 and each column on matrix X. The score of Σ is the results of adding up each column. The value of Σ will be applied in calculating the cluster center 3.

Table 6. Weight* Squared_Random 3

Cluster Centre								
Vkj	***V1***	***V2***	***V3***	***V4***	***V5***	***V6***	***V7***	***V8***
K1	3.0211	3.419306	3.426661	3.519074	3.495701	3.196499	3.322264	3.2752108
K2	3.159057	3.419038	3.323822	3.410275	3.538559	3.201749	3.180914	3.2579931
K3	3.12455	3.385155	3.308263	3.404402	3.476422	3.272663	3.255518	3.1379397

Table 7 provides the calculation results of cluster center 1, 2, and 3. The result for cluster 1 in the first row is obtained from the row in the sum of each Σ in Table 4 divided into Σ value in column $\mu i1w$. While the result of cluster 2 in the second row is generated from the row in the sum of each Σ in Table 5 divided into Σ value in column $\mu i2$, same goes to cluster 3.

Table 7. Cluster Center Result

Xi1*µi3^2	Xi2*µi3^2	Xi3*µi3^2	Xi4*µi3^2	Xi5*µi3^2	Xi6*µi3^2	Xi7*µi3^2	Xi8*µi3^2
0.30228918	0.30228918	0.20152612	0.30228918	0.40305224	0.20152612	0.30228918	0.30228918
0.33515704	0.30937573	0.41250098	0.33515704	0.378160272	0.37816027	0.30937573	0.48984491
.....	.....	.....	.....	.....	.....	.....	.....
1.20235551	0.80157034	0.80157034	0.80157034	1.202355514	1.20235551	0.80157034	1.20235551
0.38718670	0.41484290	0.33187432	0.41484290	0.368712370	0.36871237	0.33187432	0.3871867
56.27649630	60.97027280	59.58537440	61.31694190	62.3169419	58.9441712	58.6353725	56.5176632

5. Objective Function Calculation

The cluster center results in the table above would be used to find the objective function value in the next step. The fifth step is to calculate the objective function value ($P1$) with the following formula.

Description:

Pt = iteration objective function

m = attribute of each data

The new cluster center value is required to calculate the iteration of each cluster. Table 8 shows the new cluster center value with the sum of all the calculation of value for each cluster.

Table 8. Squared Random Numbers

Id Student	µ1^2	µ2^2	µ2^3
132024	0.279714	0.402854	0.317432
131036	0.285153	0.393716	0.321131
.....	.....	.....	.....
132280	0.171683	0.195241	0.633076
132314	0.21355	0.454242	0.332603
	15.96601	19.20046	18.01107

The results in Table 9 column $C1$ are obtained from each row on matrix X minus cluster 1 in Table 7 then raised to power 2. The sum of each row in column $C1$ can be seen in column $\Sigma \boldsymbol{C1}$. In column $\Sigma \boldsymbol{C1} * \mu i1w$, the results are achieved from the multiplication of column $\Sigma \boldsymbol{C1}$ and column $\mu i1w$ in Table 3. The results in column $\Sigma \boldsymbol{C1}$ would be used to determine the new matrix U. The results of column $\Sigma \boldsymbol{C1} * \mu i1w$ would be applied to calculate the objective function value.

The results shown in Table 10 column $C2$ are obtained from the subtraction of cluster 2 in Table 7 from each row in matrix X, then raised to power 2. Next, the sum up of each row in column $C2$ can be seen in column $\Sigma \boldsymbol{C2}$. Then, the results in column $\Sigma \boldsymbol{C2} * \mu i2w$ are achieved from the multiplication of column $\Sigma \boldsymbol{C2}$ by column $\mu i2w$ in Table 3. The results in column $\Sigma \boldsymbol{C2}$ would be used to determine the new matrix U.

The results of column $\Sigma \boldsymbol{C2} * \mu i 2w$ would be applied to calculate the objective function value.

Table 9. Calculation Process of Objective Function of Cluster 1

(X11-V11)^2	(X11-V12)^2	(X11-V13)^2	(X11-V14)^2	(X11-V15)^2	(X11-V16)^2	(X11-V17)^2	(X11-V18)^2	C1
0.00044522	0.17581731	2.03536274	0.26943711	0.25431716	1.43160895	0.1038541	0.075741	4.346583585
0.05239514	0.17581731	0.32871715	0.07240076	0.02934324	0.22137155	0.1038541	2.17500312	3.158902366
.....	.....	.....	.....	.....	.....	.....	.....	.....
0.00044522	2.01442879	2.03536274	2.3075855	0.2457198	0.03861171	1.74838215	0.075741	8.466276912
0.22934507	0.109358869	0.18203995	0.05332687	0.02647172	0.1863263	0.1038541	0.05053017	0.94125305

Table 10. Calculation Process of Objective Function of Cluster Center 2

(X11-V21)^2	(X11-V22)^2	(X11-V23)^2	(X11-V24)^2	(X11-V25)^2	(X11-V26)^2	(X11-V27)^2	(X11-V28)^2	C2
0.02529913	0.17559321	1.7525052	0.16832576	0.212927533	1.44420106	0.03272971	0.06656043	3.878142033
0.00827063	0.17559321	0.45721643	0.02568815	0.016497017	0.21645834	0.03272971	2.22608462	3.158538107
.....	.....	.....	.....	.....	.....	.....	.....	.....
0.02529913	2.01367008	1.7525052	1.9888762	0.290046104	0.04070273	1.39455678	0.06656043	7.572216654
0.11624213	0.10953556	0.10486081	0.11541293	0.04225462	0.01722648	0.03272971	0.05856734	0.59682958

The results written in Table 11 column $C3$ are obtained from the subtraction of cluster 3 in Table 7 from each row in matrix X, then raised to power 2. Next, the sum up of each row in column $C3$ can be seen in column $\Sigma \boldsymbol{C}3$. Then, the results in column $\Sigma \boldsymbol{C}3 * \mu i 3w$ are obtained from the multiplication of column $\Sigma \boldsymbol{C}3$ by column $\mu i 3w$ in Table 3. The results in column $\Sigma \boldsymbol{C}3$ would be used to determine the new matrix$U$. The results of column $\Sigma \boldsymbol{C}3 * \mu i 3w$ would be applied to calculate the objective function value.

Table 11. Calculation Process of Objective Function of Cluster Center 3

(X11-V31)^2	(X11-V32)^2	(X11-V33)^2	(X11-V34)^2	(X11-V35)^2	(X11-V36)^2	(X11-V37)^2	(X11-V38)^2	C3
0.01551265	0.14834425	1.71155299	0.16356413	0.274133822	1.61967073	0.06528941	0.01902737	4.017012524
0.01573775	0.14834425	0.4784994	0.0238401	0.036319938	0.15550179	0.06528941	2.59873831	3.522271139
.....	.....	.....	.....	.....	.....	.....	.....	.....
0.01551265	1.91865394	1.71155299	1.97234609	0.226978013	0.07434503	1.57632524	0.01902737	7.514741325
0.14096285	0.13311199	0.09502629	0.1194377	0.020569897	0.00364057	0.06528941	0.13108764	0.709126341

Table 12 provides the results of the objective function calculation for the first iteration. The results are obtained from the sum of column $\Sigma C1 * \mu i 1w$ in Table 8 and 9 to column $\Sigma C2 * \mu i 2w$ in Table 10 and 11. Then, the sum of values that would be the value of $P1$ is 242.235927.Σ

Table 12. Objective Function Result

C1*U1^2	C2*U1^2	C3*U1^2	Total
0.34007686	0.62938838	0.40477252	1.37423776
0.25685692	0.48961240	0.36323507	1.1097044
0.24954481	0.28864494	3.01179689	3.54998664
0.03514685	0.12314711	0.07844694	0.2367409
		P1-P0	242.235927
			242.235927

6. Degree of Membership Update

   Description:
   Uk = degree of membership
   W = Fuzziness

Table 13. New Degree of Membership

New Degree Membership						
C1^(-1)	*C2^(-1)*	*C3^(-1)*	*Total*	*μ1*	*μ2*	*μ3*
0.34007686	0.62938838	0.40477252	1.37423776	0.279714185	0.40285384	0.31743198
0.25685692	0.48961240	0.36323507	1.1097044	0.285152731	0.39371607	0.32113120
0.24954481	0.28864494	3.01179689	3.5499866	0.171683295	0.19524075	0.63307596
0.03514685	0.12314711	0.07844694	0.2367409	0.213155454	0.45424152	0.33260302

7. Next, we scan the state condition. Because $|p1 - p0| = 242.235927 - 0 = 242.235927\ means\ |p1 - p0| > \varepsilon$) and iteration=1<MaxIter(=40), therefore, it is continued to the second iteration (t=2). Back to the FCM algorithm of step 3 and so on.
8. The following results are obtained after 40 iterations:

Table 14. Objective Function

Cluster Centre								
Vkj	*V1*	*V2*	*V3*	*V4*	*V5*	*V6*	*V7*	*V8*
K1	3.33619	3.618633	3.745307	3.859148	3.77991	3.483809	3.36679	3.443086
K2	3.100169	3.530663	3.621262	3.659858	3.489727	3.407267	3.343536	3.279045
K3	2.817578	2.974294	2.49348	2.545357	3.128007	2.535126	2.848944	2.83891

$$|p40 - p39| = 41.44524(< \varepsilon) \quad (4)$$

$$\mu ik = \begin{pmatrix} 0.15319 & 0.181456 & 0.665353 \\ 0.446982 & 0.405036 & 0.147982 \\ \dots\dots\dots & \dots\dots\dots & \dots\dots\dots \\ 0.414582 & 0.171481 & 0.686937 \\ 0.405761 & 0.493491 & 0.100748 \end{pmatrix} \quad (5)$$

Through the matrix μik (new degree of membership), information regarding the tendency of data to be categorized as one of the clusters could be obtained.

Table 15. Clustering Result

NUMBER	ID STUDENT	K1	K2	K3
1	132024			X
2	131036	X		
.....	.....			
127	132280		X	
128	132314			X

Through the cluster result table above, it could be concluded that:

1. The student with registration number 131036 had a high learning interest. The calculation above showed that 47 out of 128 students had high learning interests.
2. A student with registration number 132280 had a fair learning interest. The calculation above showed that 45 out of 128 students had fair learning interests.
3. Students with registration numbers 132024 and 132314 had low learning interest. The calculation above showed that 36 out of 128 students had low learning interest.

5 Conclusion

Students' interest and motivation in a learning programming language can be determined with the system that is built by implementing the Fuzzy C-Means method. Based on the data, which were processed through several steps in Fuzzy C-Means, it could be concluded that out of 128 students, 47 of them had high learning interest, 45 students had fair learning interest, and 36 students had low learning interest.

6 Reference

[1] M. A. Firmansyah, 'Analisis hambatan belajar mahasiswa pada mata kuliah statistika', *JPPM J. Penelit. Dan Pembelajaran Mat.*, vol. 10, no. 2, 2017. https://doi.org/10.30870/jppm.v10i2.2036

[2] A. A. Slameto and E. Pramono, 'Inovasi Peta Digital Lokasi Bengkel Tambal Ban Untuk Smartphone', *Respati*, vol. 12, no. 3, 2017.

[3] Munir, Rinaldi. 2011;"Algoritma dan Pemrograman dalam bahasa Pascal dan C", Bandung: Informatika Bandung

[4] E. T. L. Kusrini, 'Algoritma data mining', *Yogyak. Andi Offset*, 2009.

[5] J. A. Tolentino and B. D. Gerardo, 'Enhanced manhattan-based clustering using fuzzy c-means algorithm for high dimensional datasets', *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 9, pp. 766–771, 2019. https://doi.org/10.18517/ijaseit.9.3.6005
[6] E. Hasmin and N. Aini, 'Data Mining For Inventory Forecasting Using Double Exponential Smoothing Method', 2020, pp. 1–5. https://doi.org/10.1109/icoris50180.2020.9320765
[7] N. Aini and M. S. Mustafa, 'Data Mining Approach to Predict Air Pollution in Makassar', 2020, pp. 1–5.
[8] R. F. Kemala, I. F. Astuti, and S. Maharani, 'Penerapan metode Fuzzy C-Means pada aplikasi simulasi TOEFL (test of English as a foreign language) berbasis web (studi kasus: fakultas MIPA Universitas Mulawarman)', *J. Ilm. Ilmu Komput.*, vol. 14, no. 1, p. 17, 2019. https://doi.org/10.30872/jim.v14i1.1954
[9] W. M. S. Mohammed and M. M. Saraee, 'Sematic Web Mining Using Fuzzy C-means Algorithm', *J. Adv. Math. Comput. Sci.*, pp. 1–16, 2016. https://doi.org/10.9734/bjmcs/2016/25471
[10] S. V. Kolekar, R. M. Pai, and M. P. MM, 'Prediction of Learner's Profile Based on Learning Styles in Adaptive E-learning System.' *Int. J. Emerg. Technol. Learn.*, vol. 12, no. 6, 2017. https://doi.org/10.3991/ijet.v12i06.6579
[11] S. Hamad, K. Alheeti, Y. Ali, and S. Shaker, 'Clustering and Analysis of Dynamic Ad Hoc Network Nodes Movement Based on FCM Algorithm', 2020. https://doi.org/10.3991/ijoe.v16i12.16067
[12] A. Al-Mutairi, K. Al-Aubidy, and F. Al-Halaiqa, 'IoT-Based Real-Time Monitoring System for Epidemic Diseases Patients: Design and Evaluation', 2021. https://doi.org/10.3991/ijoe.v17i01.18849
[13] Enhanced Web Log Data Mining using Probability Density Based Fuzzy C Means Clustering', *Int. J. Recent Technol. Eng.*, vol. 8, no. 6, pp. 2933–2939, Mar. 2020. https://doi.org/10.35940/ijrte.f8366.038620
[14] A. Nugroho, *Rekayasa perangkat lunak berorientasi objek dengan metode USDP*. Penerbit Andi, 2010.
[15] K. V. Rajkumar, A. Yesubabu, and K. Subrahmanyam, 'Fuzzy clustering and Fuzzy C-Means partition cluster analysis and validation studies on a subset of CiteScore dataset.', *Int. J. Electr. Comput. Eng. 2088-8708*, vol. 9, no. 4, 2019. https://doi.org/10.11591/ijece.v9i4.pp2760-2770

7 Authors

Sitti Aisa is a lecturer in the Department Of Information System, Dipa Makassar University. Her Current Research has been carried out since 2011. The focus of the research carried out includes information systems, data mining and mobile-based applications. She has had about 18 publications in both national journal or national conference proceedings.

Asmah Akhriana was born in Bantaeng on September 4th, 1986, is a lecturer at Dipa Makassar University in the Department of Informatics. The subject taught include computer networks, electronics and Information Technology. Her current research has been carried out since 2014. Her research interests are Microcontrollers, computer networks, the Internet of Things and mobile-based applications. There are as many as 17 research publications that have been carried out which have been pub-

lished in the form of national and international journals as well as national and international proceedings. Contact person via Email: asmah.a@dipanegara.ac.id

Ahyuna Ahyuna is a lecturer in the Department of informatics, Dipa Makassar University. Her current research Information Technology and Computer Vision. In 2017 I participated in competitive grant research that I carry out for two years. By participating in the research, I gained much experience. Since becoming a lecturer at the University of Dipa Makassar (2004), the research we raise is usually about websites, expert systems, design, Promethee Method, Augmented Reality and others. (email: ahyuna@ dipanegara.ac.id).

Andi Irmayana is a lecturer in the department of informatics, Dipa Makassar University. The subject taught include Algorithms and programming, website programming and database. Her interesting research is decision support systems and website-based information technology design (email: irmayana.andi@dipanegara.ac.id).

Irmawati Irmawati is a lecturer at Dipa Makassar University, Department of Information Systems; she teaching research methods courses since 2008 and she is interested in computer science and systems management. She actively participates in community service activities every semester in the city of Makassar. The articles that have been published include decision support system methods, android-based system design and software development (email: irmawaty@dipanegara.ac.id).

Nurul Aini is a lecturer in the Department of Software Engineering, Dipa Makassar University, was born in Barru city on December 6, 1986. He has completed a bachelor's program in Information Systems at the University of Dipa Makassar in 2009 and a master's program in Electrical Engineering at Hasanuddin University in 2013. In 2010 she started teaching at Dipa Makassar University in the Department of Software Engineering, and Since then, she has studied and taught programming language and data mining courses. Become the administrator and editorial team at KONIK 2021-2021 in Indonesia. Her currentresearch interests include rough sets and big dataanalysis. (email: nurulaini.m11@dipanegara.ac.id)

Article submitted 2021-06-05. Resubmitted 2021-07-03. Final acceptance 2021-07-05. Final version published as submitted by the authors.