

Scoring System and K-Means Algorithm for Mutaba'ah Yaumiyah Activity

Ichsan Taufik¹, Nopianti Sa'adah², Neng Lisda Yuni Suparna³, Cecep Nurul Alam⁴, Popon Dauni⁵

{ichsan@uinsgd.ac.id¹, nopianti.saadah@gmail.com², lisdayuni02@gmail.com³,
cenura@uinsgd.ac.id⁴, popon.dauni@gmail.com⁵}

State Islamic University Sunan Gunung Djati of Bandung

Abstract. *Mutaba'ah yaumiyah* is a method of evaluating for daily religious activity. The method is usually implemented in Islamic boarding schools or integrated Islamic schools, one of them is the Daarut Tauhiid Islamic Unit Santri Siap Guna boarding school. Santri Siap Guna requires students to record the religious activities that has been done on a sheet called *mutaba'ah yaumiyah*. The sheet is then submitted to the accompanying trainer to have them examined one by one just before the training begins. This methods that have existed so far are still facing several obstacles including student who often forget to record the activities, forget to keep *mutaba'ah yaumiyah* sheets, or even losing it. Another obstacle is coming from the trainer'sside especially in the control process. The large number of student's and busy schedule of the training itself hinder the examination to be carried out optimally. Therefore, this study utilizes an Android smartphone by implementing the *Scoring System* method to assess student's activities and the *K-Means* algorithm for group it on a weekly basis. By using 50 students data, the *Scoring System* method is able to carry out the assessment process with an accuracy of 100% while the *K-Means* algorithm is able to group student's religious activity data on weekly basis with an accuracy rate of 48%.

Keywords: Mutaba'ah Yaumiyah, Santri Siap Guna, Scoring System, K-Means.

1 Introduction

Smartphone users in Indonesia are experiencing rapid growth. E-Marketer Research institute estimates in the year 2018 the number of active users of smartphones in Indonesia are more than 100 million people [1].

Smartphones can be used for Muslims to evaluate daily worship or be known as Mutaba'ah Yaumiyah. Yaumiyah Mutaba'ah is used in pesantren or integrated schools, one of them is a Santri Siap Guna Daarut Tauhiid. Santri Siap Guna Daarut Tauhiid (SSG DT) are units engaged in education and training (Diklat) [2].

The current system still has several obstacles, among them is the companion trainer facing difficulties in controlling the sheet of Mutaba'ah Yaumiyah students a lot. Building the application of Mutaba'ah Yaumiyah can help the escort trainer to monitor and evaluate students in charity and assist students in recording Yaumiyah practice. The built-in Android platform uses the Scoring System method and the K-Means algorithm. The Scoring System method is used to provide assessment of Yaumiyah students practices, while the K-Means

algorithm is used to classify the practice of Yaumiyah students into several scale categorisation.

The Scoring System method is a scoring method that has some advantages over other methods. The advantages of the Scoring System method are able to present data qualitatively and able to present information in the form of numbers [3]. The K-Means algorithm is the simplest and most common method of clustering. The advantage of the K-Means algorithm is that it has the ability to classify large amounts of data with relatively fast and efficient computing time [4].

2 Methodology

2.1 Mutaba'ah Yaumiyah

The *Mutaba'ah Yaumiyah* is an activity of daily charitable evaluation of both compulsory and Sunnah. The general term Yaumiyah Mutaba'ah is less muhasabah or devotional to pay attention to the quality of our faith. Doing a Mutaba'ah Yaumiyah is one way to check the quality of faith. Some of the daily practices include prayers in congregation in the mosque, Tilawah, Tahajud prayer, morning and evening Dhikr, Shaum Monday-Thursday, *Shaum Ayyamul bidh* (middle of the month), Tadabur, memorization of the Qur'an, and some other practices [5].

2.2 Scoring System Method

Scoring system is also known as a scale score, requiring a comparative norm in order to be qualitatively interpreted. Basically, the interpretation of the scale score is always normative, meaning that the score's meaning is eliminated in the relative position of the score in a group that has been restricted first. Define categorisation within a group with a level priority scale (ordinal) using the Scoring System method.

The steps of determining categorization by level (ordinal) are as follows [6]:

1. Specifying statistical data in a descriptive form the minimum range (Xmin), maximum range (Xmax), level width, theoretical mean (σ) and standard deviation (μ).
2. Calculate the statistical data descriptively as follows:
 $X_{min} = \text{Number of questions} * \text{Minimum value}$
 $X_{max} = \text{Number of questions} * \text{Maximum value}$
Wide range of spreads = $X_{max} - X_{min}$
 $\sigma = \text{area of range spread} / 6$
 $\mu = \text{Number of questions} * \text{number of categories}$
3. Counting p by using the normal distribution table, first specify the Zmin and Zmax with the formula: $Z_{min} = (X_{min} - \mu) / \sigma$ $Z_{max} = (X_{max} - \mu) / \sigma$
4. Select P with the maximum value so that it can be found a priority scale range with 3 categories, namely:
 $X < (\mu - (P * \sigma))$ category low or less good
 $(\mu - (P * \sigma)) \leq X < (\mu + (P * \sigma))$ Medium or good category
 $(\mu + (P * \sigma)) \leq X$ category High or very good

2.3 K-Means Algorithm

K-Means is one of the algorithms in data mining that can be used to grouping or clustering a data. There are many approaches to creating clusters, including creating rules that dictate membership in the same group based on the level of equality among its members. Another approach is to create a set of functions that measure some properties of the grouping as a function of several parameters of a clustering. The K-Means method is a method included in the distance-based clustering algorithm that divides the data into a number of clusters and this algorithm only works on numeric attributes [7].

The data grouping by K-Means method is done with the algorithm [8]:

1. Specify the number of groups
2. Allocate data into groups randomly
3. Calculate the center of the group (centroid/average) of the existing data in each group. To calculate a new centroid using Equation 1.

$$C_i = \frac{1}{M} \sum_{j=1}^M x_j \quad (1)$$

4. Allocate each of the data to the nearest centroid/average. There are several ways you can measure the distance of the data to the center of the group, including Euclidean [9]. Distance measurement on distance space Euclidean can be searched using Equation 2.

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (2)$$

5. Go back to step 3, if there is still data that moved the group or if there is a change of the centroid value above the specified threshold value, or when the value change in the objective function is used still above the specified threshold value.

3 Result and discussion

3.1 Process Analysis

The analysis on this study will be divided through several phases. These stages include the assessment of daily practice of the students and the progress of practice using the Scoring System method and grouping into several scale categorisation such as good, adequate, and worse based on the value of the students ' practice Week using the K-Means algorithm. The following process analysis Scoring System method can be seen in **Figure 1**.

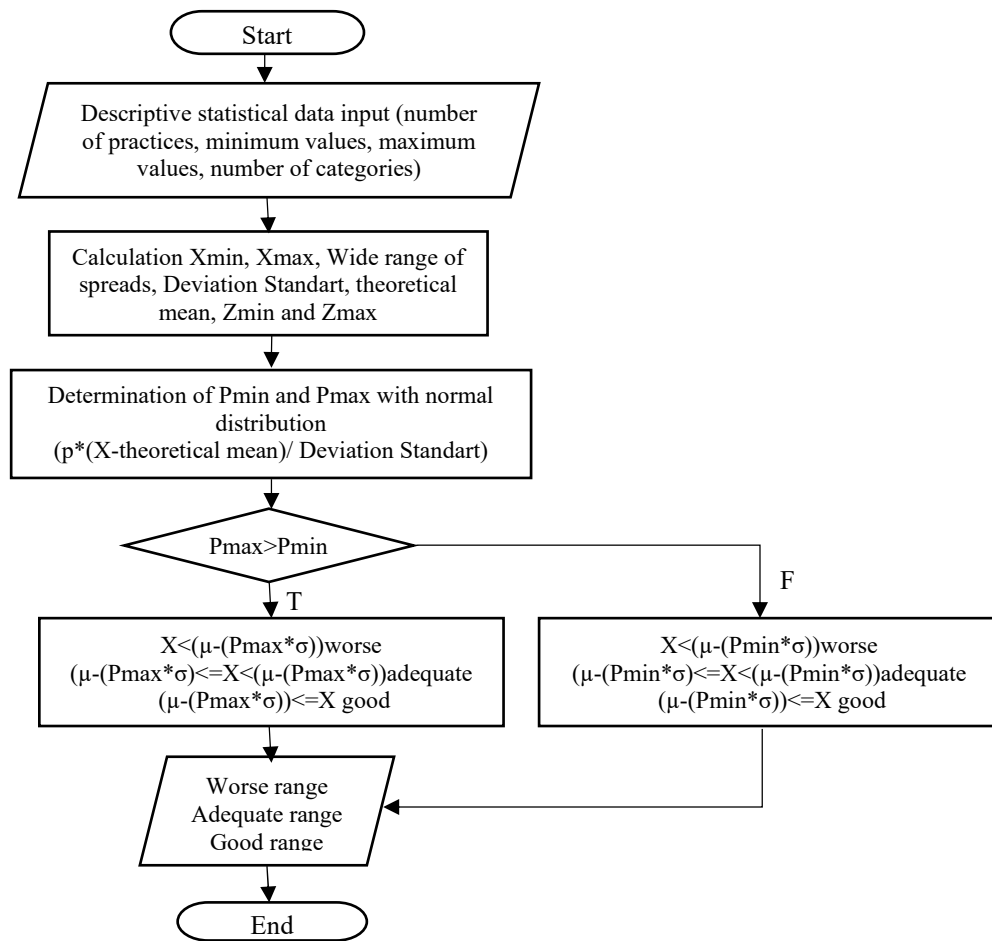


Fig. 1. Scoring System Method

The first stage of the Scoring System method is to define the category scale and valuation parameters can be seen in Table 1.

Table. 1 Scale Category

No	Category
1	Worse
2	Adequate
3	Good

Table. 2 Scoring System Method Parameter

Parameter	Values
Number of practice	31
Number of category	3
Highest values	5
Lowest values	1

After specifying the parameters in Table 2, the calculation is done using the equation above to determine the value of each category. To calculate Pmin and Pmax should use a normal distribution table and first specify Zmin and Zmax. Calculations can be seen in Table 3.

Table. 3 Normal Distribution

Z	0	0,01	0,02
-3,5	0,0002	0,0002	0,0002
-3,4	0,0003	0,0003	0,0003
-3,3	0,0005	0,0005	0,0005
-3,2	0,0007	0,0007	0,0006
-3,1	0,0010	0,0009	0,0009
-3,0	0,0013	0,0013	0,0013
0,0	0,5000	0,4960	0,4920
3,0	0,9987	0,9987	0,9988
3,1	0,9990	0,9991	0,9991
3,2	0,9993	0,9993	0,9994
3,3	0,9995	0,9995	0,9995
3,4	0,9997	0,9997	0,9997
3,5	0,9998	0,9998	0,9998

Table. 4 Calculation Scoring System Method

Variable	Calculation	Results
X _{min}	31 x 1	31
X _{maks}	31 x 5	155
LJS	155 - 31	124
μ	31 x 3	93
σ	124 / 6	20,67
Z _{min}	(31 - 93) / 20,67	-3
Z _{max}	(155 - 93) / 20,67	3
P _{min}	Z _{min} (dist table)	0,0013
P _{max}	Z _{max} (dist table)	0,9987
Category 1	$x < (93 - (0,9987 * 20,67))$	Worse ($x < 72,36$)
Category 2	$(93 - (0,9987 * 20,67)) \leq x < (93 + (0,9987 * 20,67))$	Adequate ($72,36 \leq x < 113,64$)
Category 3	$(93 + (0,9987 * 20,67)) \leq x$	Good ($113,64 \leq x$)

From Table 4, it can be noted that the worse category is obtained when the value of > 72.36 , sufficient if the value is $\geq 72.36 < 113.64$, and good if the value is ≥ 113.64 . The same equations and calculations are applied to 50 students ' practice data to obtain value and categorisation on a daily basis.

Perform the same process for one week from 26 August 2018 – 1 September 2018. Next, enter the K-Means algorithm calculation step. In Figure 2 is the K-Means flowchart algorithm.

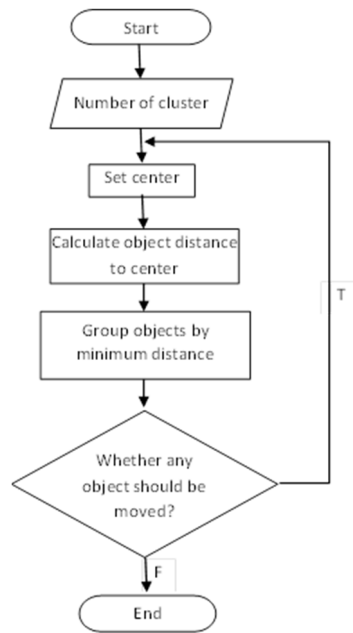


Fig. 2 K-Means Algorithm Flowchart

The Parameter used for the analysis of K-Means algorithm is the number of students ' practice value for one week from 26 August 2018 to 01 September 2018 as many as 50 data consisting of 25 female students ' data and 25 students ' data Men.

Table. 5 Initial Cluster Center

No	Category	Weekly Practice Values
1	Good	750
2	Adequate	650
3	Worse	450

The initial cluster Centre value in Table 5 is used to calculate the distance between the data with a centroid. The equation used to calculate the distance of this study is Euclidean Distance (Equation 2).

The same equation and calculation is applied in 50 data to get the distance of each data on each cluster.

After each of the data is counted by the distance of each cluster, the next step is to group the data accordingly. Cluster groups of data are retrieved from the shortest distance of the data against a cluster.

Once the data is grouped according to its cluster, the next step is to calculate the new centroid in each of the clusters using equation 1. The new centroid values on each cluster can be seen in Table 6.

Table. 6 New centroid values

No	Category	Weekly Practice Values
1	Good	719,00
2	Adequate	601,86
3	Worse	469,65

Once calculated next compared to the previous centroid value. If the value is the same then the iteration process is stopped. But if the value of different calculations is repeated again. The calculation process stops at the 4th iteration with the following result in Table 7.

Table. 7 Cluster Result

No	Students Name	Cluster	Distance between Cluster	No	Students Name	Cluster	Distance between Cluster
1	Abdul Aziz	2	52,90	26	Atin Janatin	1	18,17
2	Abdullah Nashih	2	50,10	27	Devie Prawitasari	2	26,90
3	Abdurrahman W	1	5,17	28	Dila Septiani	2	32,90
4	Affadil Yusral A	3	48,08	29	Erlina Martdhani	2	12,10
5	Ahmad Soleh	2	2,90	30	Fitri Fatmawati	2	3,10
6	Alrizqu Ahlu A	2	13,10	31	Fitri Yuniar	2	27,90
7	Anggi Suheri	3	2,08	32	Hanifa T.P	3	40,08
8	Dika Apriyana	3	57,08	33	Ibna Siti	3	3,08
9	Doni Rizaldi	2	32,10	34	Ika Riyanti	2	0,90
10	Endi Gusmanto	1	34,83	35	Inas Nida S	1	12,83
11	Fajar Maulana	3	3,08	36	Keissha Nuansa	1	2,17
12	Fajar Saepulloh	3	37,08	37	Kokom Nurul	3	4,92
13	Faqih Ainul B	3	36,92	38	Mega Deismasuci	2	4,90
14	Haekal Harana	2	8,10	39	Mira Nurseha	3	10,92
15	Irwan Munawar	3	32,92	40	Nenda	3	6,92
16	Muamar Firdaus	3	40,92	41	Nisa Nur Azizah	3	27,92
17	Rama Wahyu	3	30,08	42	Rafa Tsana	2	30,10
18	Rizal Agnimas	3	30,08	43	Resna Pujayantry	3	44,08
19	Tasim Maulana	3	9,08	44	Resnia Nur A	2	25,90
20	Tedi K	2	16,10	45	Santi Yuliawati A	1	22,17
21	Tedi Purwa A	3	62,92	46	Sinti Santi	2	13,10
22	Teguh Syam M	2	1,90	47	Siska Sri	3	125,92
23	Umarul Wahid	3	37,08	48	Siti Aminah	2	15,10
24	Veldy Qaedy R	2	15,90	49	Siti Aulia	3	9,92
25	Zaenal Mustafa	3	5,92	50	Zahro Azhar	3	25,08

4 Implementation

In Figure 3 is an implementation interface of home page Students.

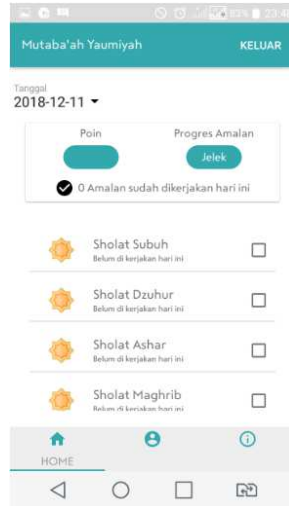


Fig. 3 Home Page Student

In Figure 4 is an implementation interface of home page Trainer.

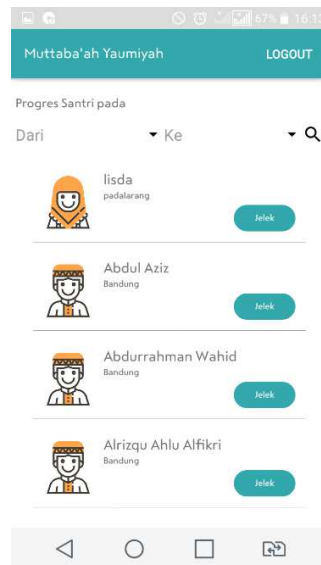


Fig. 4 Trainer Page

In Figure 5 is an implementation interface of type practices page

#	Amalan	Nilai	Aksi
1	Barda Isya	4	[Edit] [Delete]
2	Barda Maghrib	4	[Edit] [Delete]
3	Barda Zuhur	4	[Edit] [Delete]
4	Dzikir Pagi	5	[Edit] [Delete]
5	Dzikir Petang	5	[Edit] [Delete]
6	Tikar	5	[Edit] [Delete]
7	Kajian Al-Hikmah	4	[Edit] [Delete]
8	Kajian Marfatullah	5	[Edit] [Delete]
9	Kajian MQ Pagi	4	[Edit] [Delete]
10	Khidmat untuk orang sekitar	5	[Edit] [Delete]

Fig. 5 Type Practices Page

5 Conclusion

The conclusions that can be drawn from this study are:

1. Implementation of the Scoring System method in the application of Mutaba'ah Yaumiyah applied when the students have entered the practice data Yaumiyah done, then the data will be processed and displays the progress of daily practice students. While the K-Means algorithm is applied to the practice of Yaumiyah students for a week, then the practice is grouped based on the categorization scale and the result of the grouping can be seen on the trainer's page.
2. Results of performance testing method Scoring System and algorithm K-Means in the activity of Mutaba'ah Yaumiyah to test as much as 50 data. Obtained an accuracy value of 100% for the scoring system and obtained an accuracy value of 48% for the K-Means algorithm. From the test results the accuracy is said that the algorithm K-Means is not suitable to be implemented in the activity of Mutaba'ah Yaumiyah.

Acknowledgements

Authors wishing to acknowledge Research and Publication Centre of UIN Sunan Gunung Djati Bandung that supports and funds this research publication.

References

- [1] "Indonesia Raksasa Teknologi Digital Asia," 2015. [Online]. Available: https://www.kominfo.go.id/content/detail/6095/indonesia-raksasa-teknologi-digital-asia/0/sorotan_media. [Accessed: 16-Jan-2018].
- [2] Tim IT SSG DT, "Diklat SSG," 2016. [Online]. Available: <https://ssg-dt.org/>. [Accessed: 01-Mar-2018].

- [3] D. R. Anamisa, "Penerapan Metode Scoring System Untuk Penilaian Latihan Pemahaman Materi Ibadah Sholat Fardhu Dan Sunnah," pp. 9–10, 2015.
- [4] T. Alfina, B. Santosa, and R. Barakbah, "Analisa Perbandingan Metode Hierarchical Clustering , K-means dan Gabungan Keduanya dalam Cluster Data (Studi kasus : Problem Kerja Praktek Jurusan Teknik Industri ITS)," *J. Tek. ITS*, vol. 1, no. Data Mining, pp. 1–5, 2012.
- [5] F. Arifin, "Pentingnya Mutaba'ah Yaumiyah," 2015. [Online]. Available: <http://www.firman-its.com/2015/06/19/pentingnya-mutabaah-yaumiyah/>. [Accessed: 16-Jan-2018].
- [6] P. Hermayanti, A. Budimansyah, and U. T. Lenggana, "Implementasi Metode Scoring System Sebagai Paramater dalam Memahami Kajian Ilmu Tasawuf Berbasis Android," *J. Online Inform.*, vol. 2, no. 2, p. 92, 2018.
- [7] I. H. Witten, *Data Mining Practical Machine Learning Tools and Technique*, 2nd ed. San Francisco: Elsevier, 2005.
- [8] E. Prasetyo, *Data Mining : Konsep dan Aplikasi Menggunakan MATLAB*. Yogyakarta: ANDI, 2012.
- [9] R. S. Pressman, *Rekayasa Perangkat Lunak*. Yogyakarta: ANDI, 1992.