APPLICATION OF THE K-MEANS CLUSTERING METHOD FOR PERFORMANCE ASSESSMENT BASED ON EDUCATOR COMPETENCE

Paul Sihombing[1], Bobby Rahman Angkat[2], Eliza Christovel Yosua[3], Mutiara Sembiring[4], Marlince Nababan[5]
[1,2,3,4,5]Indonesian Prima University
Jl. Sampul No. 1, Sei Putih Tengah, Kec. Medan Petisah, Medan City, North Sumatra 20118
E-mail:marlince@unprimdn.ac.id

ABSTRACT- Performance appraisal is one thing to respect someone while working in an institution, one of which is a private higher education institution. To appreciate the performance of resources, there needs to be a value assigned to someone. Assessments carried out for one semester need to be reviewed again because during filling in, the student assessments do not fill in according to their understanding. Hence, a review needs to be carried out again. The evaluation was carried out using the K-Means method by applying the concept of the centroid value. There are 4 (four) variables used, namely academic competence, personal competence, and social and professional competence, with a value of K = 3. The maximum number of observations for Cluster 3 is 368, while the value of Distances Between Cluster Centroids shows two suitable clusters, namely Cluster 1 and Cluster 2, which is 1.7020. The author gives suggestions to remove outlier data before entering the data to be trained into the algorithm to improve visualization if the dataset is large.

Keyword: Performance Appraisal, Data Mining, K-Means.

1. INTRODUCTION

Every university should conduct a performance assessment, especially for lecturers or teaching staff, which aims to measure the ability of lecturers during the implementation of learning—the relationship between lecturers and students or vice versa. Implementation of the Independent Campus Program is conducted every semester, which requires an academic assistant lecturer (DPA)—to assess the performance of lecturers as traditional educational companions for quite a long time. Lecturer data was taken from as many as 28 data and found several groups of academic assistant lecturers with the same performance patterns that can help identify the strengths and weaknesses of each group. The K-Means Clustering method significantly contributes to optimizing performance as an academic assistant lecturer at the Merdeka Campus Program and taking appropriate steps to improve the quality of educational assistance. [1]. To get information and evaluation for lecturers, decision-makers with various algorithms applications are needed. The method used by the author is the K-Means Clustering algorithm with the stages carried out in analyzing and classifying lecturer performance using a random central centroid value.

To find out the level of lecturer performance, it is necessary to group permanent lecturers based on relevant assessment criteria and provide guidance for appropriate development actions with the suitable K-Means algorithm based on existing attributes. The resulting cluster results are evaluated and adjusted if necessary. By using the k-means clustering method, institutions can identify groups of lecturers with similar performance and determine appropriate development actions for each group.[2].

Data mining techniques can be used to classify lecturer performance. This study applies an algorithm for clustering lecturer performance assessment; k-means is a popular and efficient clustering algorithm for grouping data. The research methodology consists of the following steps: 1) lecturer performance assessment data uses student surveys, peer assessment, and internal evaluation; 2) the k-means algorithm will be applied to the data to form homogeneous groups based on the assessment attributes. So, by applying the k-means algorithm in clustering lecturer performance assessment[3].

Using the K-Means clustering method, the output is clustering the performance evaluation of teaching staff, which aims to see the ability of lecturers in terms of teaching by students/I, which can also be a reference for developing quality and increasing the capacity of lecturers in the teaching and learning process[4].

Clustering lecturer performance based on student assessments is an essential challenge in improving higher education quality. This study aims to apply the k-means clustering method in clustering lecturer performance based on student assessments of four critical competencies: material knowledge, delivery ability, responsiveness, and student interaction.[5].

They are implementing K-Means in the process of assessing questionnaires for lecturers to support mapping student satisfaction with lecturers. The research data was obtained through a lecturer questionnaire. The implementation of the K-Means algorithm is a questionnaire data from students. The research successfully built a clustering information system using the PHP programming language. The system created has succeeded in grouping 46 student satisfactions into categories. The information generated from the analysis of the questionnaire data can provide a better understanding of student needs and assist lecturers in improving the quality of their teaching[6].To analyze and group variables by determining the centroid value of the center point. The K-Means algorithm process will be finished if there is no change in the end centroid value[7]. Evaluation of lecturers whom Atma Jaya Yogyakarta students have assessed will then be clustered using the K-Means algorithm, which functions to group lecturers into four clusters[8].

This study aims to evaluate the performance using the Euclidean distance matrix K-Means algorithm. The K-Means algorithm is a popular method used in clustering analysis, and the Euclidean distance matrix is used to measure the distance between data points. Within the scope of education, student assessments are not only based on academic and non-academic values but also reviews of the student's character; with student incident records, all student activities will be recorded. Determining problem students is very important in education because it can help identify students who need special attention. The data is processed by clustering students with problems or not with the K-Means method of the Euclidean distance matrix. Centroid data value is not problematic at 2.165 and inappropriate at 7.372, and the Davies Bouldin index score gets a score of 0.397[9] [10] [11].

To measure the ability of lecturers during the implementation of learning. The relationship between lecturers and students or vice versa. The assessment has been carried out every semester and according to the course taught by the lecturer. This assessment was carried out on each student portal with 20 questionnaires. Four competency variables are tested on teaching staff, namely 1) Pedagogic, 2) Social, 3) Personality, and 4) Professional. Some students only filled out a questionnaire to fulfill the requirements for filling out the KRS (Semester Plan Card). Therefore, it needed to be more suitable for assessing the lecturer's final results. Thus, the authors are interested in analyzing Lecturer Competency Assessment using K-Means titled "Application of the K-Means Method for Clustering Performance Assessment Based on Educator Competency."

1.2. Formulation of the problem
Based on the problems above, several issues can be formulated: How to do data analysis on educator performance assessments? How do we analyze performance appraisal using the K-Means algorithm?

1.3. Scope of problem
The limitations of the problem in this study are: Only discussing the performance of teaching staff for 1 (one) semester, Only consulting the implementation of teaching staff with 4 competencies, namely social, academic, social, and professional. To analyze using the K-Means Clustering algorithm

1.4. Research purposes
This research aims to increase understanding of semester lecturer teaching competencies, provide feedback for lecturer development, and classify the competence of teaching staff.

2. RESEARCH CONTENT

2.1 Literature Study
    This research starts from a literature study through interviews and direct observation, as shown in Figure 2.1. Research methods.
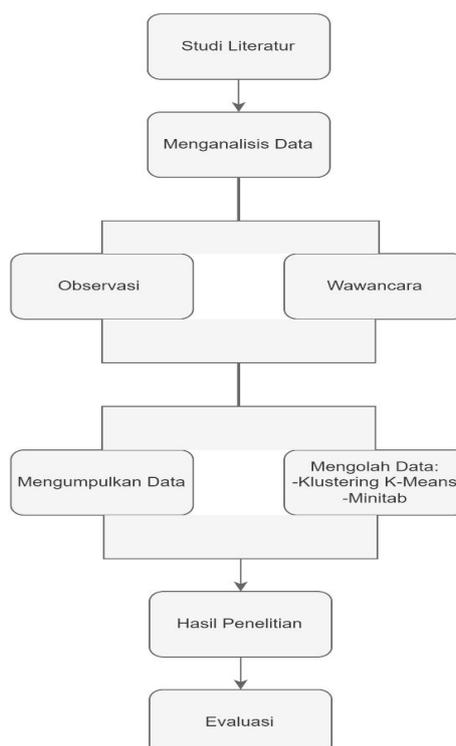
Figure 2.1. Research methods
    This study collects the data needed for the data mining process. The data used in this study comes from the Center for Information Systems, namely the lecturer evaluation dataset. The author also retrieves information based on facts from previous research, namely national journals. Lecturers' evaluation is carried out according to what is taught in class. After this, data will be used in mining using K-Means Clustering.
    The algorithm used is K-Means Clustering, which is the shortest distance value system[10]. The stages in clustering the K-Means method are:

1. Determine the value of K, where K is the number of clusters to be formed.
2. Determine the initial center point of each group.
3. Calculate the distance for each input data for each centroid using the Euclidean Distance formula:

$$D(x,y) = \sqrt{((X_{(1-)} Y_1) + (X_2 - Y_2))}2$$
(1)

Information;

  D = Distance
  x = Data
  y = Centroids

4. Clustering data based on their proximity to the centroid.
5. Recalculate the cluster center with the current cluster members.
6. Compute each object again using the new cluster center.

2.2 Data analysis

Data processing is by processing performance data in the form of numerical values . Then, the data is processed using the K-means clustering method to classify the competencies of educators. Before entering into the data processing stage, it is necessary to analyze the data that has been taken to be able to check the clarity and completeness of the data in the research. The purpose of data analysis is to determine the quality of the data to be tested and to draw hypotheses related to the data to be tested. The variables used are 4 (four) variables in Table 2.

NO.	Variable
1	Pedagogic Competence
2	Professional Competence
3	Social Competence
4	Personality Competence

Table 2. Educator Competency Variables
Source: Head of HR Department

The steps in analyzing the data are:
 A. *Data Preprocessing*
  This stage is the stage for cleaning the results of data selection. The data used is the results of student/I assessments for 1 (semester) in 2021. Before the data mining process can be carried out, it is necessary to carry out the data cleaning stage.
 B. Minitab Software
  Minitab is a computer program designed to perform statistical processing. Minitab combines the ease of use of Microsoft Excel with the ability to perform complex statistical analysis. Minitab is an open-source software.

BIBLIOGRAPHY

[1]   S. Fathuroh, "The K-Means Clustering Method in Optimizing the Performance of Academic Assistant Lecturers in the Merdeka Campus Program,"*J. Inf. and Technol.*, vol. 5, pp. 5–9, 2022, doi: 10.37034/jsisfotek.v5i2.172.

[2]   L. Hakim and YP Santoso, "Permanent Lecturer Performance Assessment Using K-Mean Clustering at Xyz University,"*J. Inform. and Computing*, vol. 13, no. August, pp. 87–94, 2019.

[3]   D. Sartika and J. Jumadi, "National Seminar on Computer Technology & Science (SAINTEKS) Clustering Lecturer Performance Assessment Using the K-Means Algorithm (Case Study: Dehasen University Bengkulu)," pp. 703–709, 2019, [Online]. Available: https://seminar-id.com/semnas-sainteks2019.html

[4]   D. Sukrianto, "Application of Data Mining for Lecturer Performance Using the K-Means Clustering Method (Case Study at Amik Mahaputra Riau),"*J. PI-Cache*, vol. 5, No. 1, no. Dm, pp. 54–63, 2016.

[5]   KK Silalahi, "Simki-Techsain Vol. 01 No. 01 of 2017 ISSN : XXXX-XXXX,"*Simki-Techsin*, vol. 01, no. 01, pp. 1–7, 2017.

[6]   B. Parlambang and Fauziah, "Implementation of the K-Means Algorithm in the Questionnaire Assessment Process for Lecturers to Support Student Satisfaction with Lecturers,"*J. Ilm. Technol. and Engineering*, vol. 25, no. 2, pp. 161–173, 2020, doi: 10.35760/tr.2020.v25i2.2719.

[7]   NLPP Dewi, IN Purnama, and NW Utami, "Application of Data Mining for Lecturer Performance Assessment Clustering Using the K-Means Algorithm (Case Study: STMIK Primakara),"*J. Ilm. Technol. inf. Asia*, vol. 16, no. 2, p. 105, 2022, doi: 10.32815/jitika.v16i2.761.

[8]   S. Hendrawan, F. Kartika, and S. Dewi, "Clustering Evaluation of Atma Jaya University Yogyakarta Lecturers Using the K-Means Method," pp. 1–8, 1965.

[9]   NA Widiastuti*et al.*, "PERFORMANCE EVALUATION OF THE K-MEANS ALGORITHM USING THE MATRIX," vol. 13, no. 1, pp. 1–8, 2022.

[10]   Normah, B. Rifai, S. Vambudi, and R. Maulana, "Sentiment Analysis of Vtuber Development Using the SMOTE-Based Support Vector Machine Method,"*J.Tek. computer. AMIK BSI*, vol. 8, no. 2, pp. 174–180, 2022, doi: 10.31294/jtk.v4i2.

[11]   N. Afiasari, N. Suarna, and N. Rahaningsih, "Implementation of Sales Transaction Data Mining Using the Clustering Algorithm with the K-Means Method E-commerce K-Means analyzes the application of Data Mining in classifying quantities," vol. 9, pp. 100–110, 2023.