Clustering for Recommendation of Further Studies for Lectures at Politeknik Negeri Media Kreatif, Indonesia

Ahmad Irfan Abdullah*, Arysespajayadi, Fadly Shabir, Tsaqila Yusuf Dahlan
Politeknik Negeri Media Kreatif, Jl. Srengseng Sawah Jagakarsa, Jakarta Selatan 12640, Indonesia

ORCID
Ahmad Irfan Abdullah: https://orcid.org/0009-0008-8826-3943

Abstract.
Professionally, lecturers must continue to develop themselves in their field of expertise and regularly update their knowledge and skills, in order to provide teaching that is up-to-date and relevant to the changing times. To improve the quality of lecturers, one of the strategies is to enhance their academic qualifications. In this regard, lecturers who meet the qualifications can pursue further studies at the doctoral level. The recommending authorities can consider various factors when evaluating the performance quality of a lecturer, including age, work experience, academic achievements and performance, specialization, lecturer-to-student ratio, availability and readiness of substitute lecturers, as well as the reputation and status of the university where the lecturer applies and is accepted.

In this research, the clustering of lecturer performance data is done using the K-Means method to address the aforementioned issues. This allows for faster determination of recommendations as the process is automated. The analysis and evaluation of the clustering results are conducted by determining the quality of the clustering using the Silhouette Coefficient method.

This research has successfully applied the K-Means Clustering method, providing 3 clusters of recommendations for further studies for Polimedia lecturers. The testing of this system using the Silhouette Coefficient obtained an average value of 0.78, indicating that the clustering results are in good condition.

Keywords: K-Means, evaluating, results

1. Introduction

Lecturers play a crucial role in implementing the national education system by teaching, transforming, developing, and disseminating knowledge, technology, and the arts through education, research, and community service, as stated in Article 1, paragraph 2 of Law No. 14 of 2005 on Teachers and Lecturers. As professionals, lecturers are
also responsible for elevating the dignity of their profession and advancing knowledge, technology, and the arts to improve the quality of national education [11].

Professionalism in education also encompasses responsibility towards students, educational institutions, and society. A professional lecturer must be capable of delivering quality instruction, providing guidance and support to students, and upholding ethics and integrity in their duties. Furthermore, a professional lecturer must continually self-improve in their field of expertise and periodically update their knowledge and skills, ensuring they can provide the most current and relevant instruction in line with the times’ developments [11].

For the purpose of improving the quality of lecturers, one of the methods is to enhance their academic qualifications, especially for lecturers who already meet the qualifications. Consequently, recommending authorities can consider factors that influence the assessment of a lecturer’s performance, such as age, work experience, academic achievements, specialization, the lecturer-to-student ratio, the availability and readiness of substitute lecturers, as well as the reputation and status of the university where the lecturer applies or is accepted. All these factors will be compiled in a database called the lecturer data. These factors will be treated as variables in a program, which will be grouped using the K-Means clustering method to obtain the best clusters for further consideration by recommending authorities in making recommendations for further studies. Many clustering methods have been developed for use in recommendation systems, such as K-Means, Hierarchical, ROCK, and various other methods [5].

Based on the above issues, to facilitate the recommendation process, an application has been created that can automatically group lecturer data, thereby reducing the time required to make recommendations for further studies for lecturers pursuing higher education.

2. Material and Methods

The experimental research in this case is focused on the application of the K-Means method to perform clustering of faculty data. A literature review was conducted to gather information from printed and electronic sources, such as books and modules, which can support the theories used as the basis for this research. The research was conducted using an experimental method, with the aim of testing the influence of specific treatments on other variables that can be controlled [12].

The collected data will undergo data analysis using predefined methods. The analysis method to be used may vary depending on the research objectives, the type of data
<table>
<thead>
<tr>
<th>No.</th>
<th>Name</th>
<th>Age</th>
<th>Masa Kerja</th>
<th>Pengalaman</th>
<th>Kinerja</th>
<th>Spesifikasi</th>
<th>Rasio</th>
<th>Dosen Pengganti</th>
<th>Reputasi ( \text{dan Status PT Tujuan} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ahmad Irfan A</td>
<td>39</td>
<td>4</td>
<td>Cukup</td>
<td>Baik</td>
<td>Sesuai</td>
<td>Cukup</td>
<td>Sedia</td>
<td>Sangat Baik</td>
</tr>
<tr>
<td>2</td>
<td>Fitriani Halik</td>
<td>27</td>
<td>2</td>
<td>Cukup</td>
<td>Baik</td>
<td>Kurang</td>
<td>Sama</td>
<td>Sama</td>
<td>Sangat Baik</td>
</tr>
<tr>
<td>3</td>
<td>Junaedi</td>
<td>36</td>
<td>5</td>
<td>Cukup</td>
<td>Berpengalaman</td>
<td>Sangat Baik</td>
<td>Kurang</td>
<td>Sama</td>
<td>Sangat Baik</td>
</tr>
<tr>
<td>4</td>
<td>Anyespa Ajayadi</td>
<td>32</td>
<td>1</td>
<td>Cukup</td>
<td>Berpengalaman</td>
<td>Baik</td>
<td>Sesuai</td>
<td>Sedia</td>
<td>Sangat Baik</td>
</tr>
<tr>
<td>5</td>
<td>H. Suardi</td>
<td>61</td>
<td>33</td>
<td>Sangat</td>
<td>Berpengalaman</td>
<td>Sangat Baik</td>
<td>Kurang</td>
<td>Sama</td>
<td>Sangat Baik</td>
</tr>
<tr>
<td>6</td>
<td>Moh Djazman Addin</td>
<td>51</td>
<td>17</td>
<td>Berpengalaman</td>
<td>Baik</td>
<td>Sesuai</td>
<td>Cukup</td>
<td>Sedia</td>
<td>Sangat Baik</td>
</tr>
<tr>
<td>7</td>
<td>Elok Faiqoh</td>
<td>29</td>
<td>4</td>
<td>Cukup</td>
<td>Berpengalaman</td>
<td>Mencukup</td>
<td>Kurang</td>
<td>Sama</td>
<td>Sangat Baik</td>
</tr>
<tr>
<td>8</td>
<td>Besse Ima Tawaddud</td>
<td>34</td>
<td>4</td>
<td>Cukup</td>
<td>Berpengalaman</td>
<td>Mencukup</td>
<td>Sesuai</td>
<td>Sedia</td>
<td>Sangat Baik</td>
</tr>
<tr>
<td>9</td>
<td>Fadly Shabir</td>
<td>35</td>
<td>1</td>
<td>Cukup</td>
<td>Berpengalaman</td>
<td>Baik</td>
<td>Sesuai</td>
<td>Sedia</td>
<td>Sangat Baik</td>
</tr>
<tr>
<td>10</td>
<td>Fitriani Halik</td>
<td>27</td>
<td>2</td>
<td>Cukup</td>
<td>Berpengalaman</td>
<td>Baik</td>
<td>Kurang</td>
<td>Sama</td>
<td>Sangat Baik</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>137</td>
<td>Sofi Andriyanti, S.Pd., M.Sn.</td>
<td>33</td>
<td>4</td>
<td>Cukup</td>
<td>Berpengalaman</td>
<td>Baik</td>
<td>Sesuai</td>
<td>Sedia</td>
<td>Sangat Baik</td>
</tr>
</tbody>
</table>

After data analysis is performed, the research results need to be interpreted and analyzed to obtain accurate and relevant conclusions related to the research problem. At this point, the data is ready to be processed into the K-Means clustering algorithm. This stage will be carried out by the research team leader and members with the achievement indicator being that the clustering of candidate faculty members for further study recommendations is of good quality.

Clustering data with the K-Means method is often performed using a 5-step algorithm as follows [8].

Determination of the number of clusters.

Data will be randomly allocated into clusters.

Calculation of centroids for the data within each cluster.

Allocation of data to the nearest centroid/average.

Return to step three, which is to calculate centroids, if there is still data cluster movement in the previous clustering results, or if the centroid values exceed the
predetermined maximum value, or if the objective function being used still exceeds
the threshold value. The measurement of data distance to the cluster center is done
using Euclidean distance, thus obtaining the shortest distance between two points.
The application flow design is depicted as shown in Figure 1. The Euclidean distance
equation between two points $\mathbf{x}_j$ and $\mathbf{c}_j$ in p dimensions is as follows [2]:

$$d (x_j, c_j) = \sqrt{\sum_{j=1}^{n} (x_j - c_j)^2}$$

Where :
- $d$ : distance
- $n$ : number of objects
- $j$ : (ranging from 1 to n)
- $x_j$ : feature of object j with respect to x
- $c_j$ : centroid of feature j

Figure 1: K-Means Flow Diagram.

3. Result and Discussion

The lecture data displayed in Table 1 were processed using the K-Means algorithm,
resulting lecturer clusters. Three clusters were obtained and labeled as follows: less
recommended, recommended, and highly recommended. Out of the 139 lecturers,
90 were clustered as less recommended, 28 as recommended, and 21 as highly
recommended for further study. These cluster results can be used by policymakers
to provide further study recommendations to lecturer at Politeknik Negeri
Media Kreatif. These cluster results shown in Figure 2.
The obtained cluster results were then tested using the Silhouette Coefficient method, which yielded a value of 0.78. This test result indicates that the clustering results are in good condition since it is above the standard threshold of 0.6 for cluster testing using the K-Means method. Visualisation of lecturer clusters result in scatter graphic and shilouette score shown in Figure 3.
4. Conclusion

This research has successfully applied the K-Means clustering method, providing 3 clusters of study recommendations for faculty members at Polimedia. The system's trial results, using the Silhouette Coefficient, obtained an average value of 0.78, indicating that the clustering results are in good condition.

Furthermore, for the development of this research and to achieve even better results, future research will apply various other methods and make changes to variables.

Acknowledgements

I would like to express my sincere gratitude to Politeknik Negeri Media Kreatif for providing the platform and resources essential for the successful completion of this [thesis/project/research]. The institution's commitment to fostering creativity and innovation has greatly enriched my academic journey. I am also thankful for the vibrant and collaborative learning environment that Politeknik Negeri Media Kreatif cultivates, which has inspired me to explore new horizons and push the boundaries of my creative endeavors.

In conclusion, my heartfelt thanks to Politeknik Negeri Media Kreatif for the transformative experience and for contributing to my growth as a creative individual.

References


