E-ISSN : 2963-0568, P-ISSN : 2963-055X
Available online at:
https://ejournal.itbwigalumajang.ac.id/index.php/jid

Clustering of Lecturer Performance Using K-Means

Abdur Ro'uf[1], Marita Qori'atunnadyah[2], Hasyim Asy'ari[3], Maysas Yafi Urrohman[4]

Informatics Department, Institut Teknologi dan Bisnis Widya Gama Lumajang, Indonesia[1,2,3,4]

Corresponding Author: Abdur Ro'uf (abdurrouf.ar34@gmail.com)

ARTICLE INFO

Date of entry:
09 April 2024
Revision Date:
24 October 2024
Date Received:
29 October 2024

ABSTRACT

Lecturers serve as professional educators and scientists whose primary roles are knowledge transformation, development, and dissemination in fields such as science, technology, and the arts through education, research, and community service. They play a critical role in fostering an educated generation, and as such, must maintain high levels of integrity in their work. The academic position of a lecturer often reflects their involvement in research, community service, and scientific publications, indicating a broad scope of expertise. This study aims to cluster lecturers based on their academic positions, research activities, community service, and number of publications, using secondary data from the Community Service Research Institute, UPT Academic Positions and Lecturer Certification, and UPT Publications. The clustering was conducted using a non-hierarchical k-means method, which resulted in three clusters: Cluster 1 with 26 members showing minimal productivity in the tridharma tasks, Cluster 2 with 6 members demonstrating high engagement, and Cluster 3 with 20 members with moderate involvement. These findings suggest that universities need to monitor and support lecturers in Cluster 1 to improve their contributions to education, research, and community service. This clustering provides insights that can guide universities in promoting a balanced and active academic environment.

Keywords: Clustering, Lecturer Performance, K-Means.

Cite this as: Rouf, A., Qoritunnadyah, M., Asyari, H., & Urrohman, M. Y. (2024). Clustering of Lecturer Performance Using K-Means. *Journal of Informatics Development*, *3*(1), 27–33. https://doi.org/10.30741/jid.v3i1.1430

INTRODUCTION

Lecturers are professional educators and scientists with the main task of transforming knowledge, developing knowledge and disseminating science, technology and art through education, research, namely compiling scientific work and community service (Pemerintah Indonesia, 2021). Lecturers play an important role in the world of education with the main goal being to educate the nation's next generation, with professional teaching staff in their knowledge, the teaching and learning process in higher education will be able to achieve an intelligent generation of the nation.

In this case, lecturers are required to have academic qualifications, competencies, educational certificates, be physically and spiritually healthy and meet other qualifications which are supporting requirements for the creation of professional lecturers (Pemerintah Indonesia, 2005). In creating quality education, lecturers are also required to carry out the tri dharma of higher education, including education and teaching, research, community service and supporting elements. Apart from that, lecturers are very important in assessing study program accreditation and higher education accreditation. Universities or study programs that have lecturers with functional position qualifications will give grades according to the number of lecturer functional position qualifications in the university (Redaksi, 2023)

Apart from that, in higher education lecturers are required to have a functional position as low as an expert assistant (Ristekdikti, 2019), functional positions or what are called academic positions are regulated (Kementerian Pendidikan dan Kebudayaan, 2014). The academic position of lecturer is a position that indicates the duties, responsibilities, authority and rights of a lecturer in a higher education unit whose implementation is based on certain skills and is independent, therefore to become a professional teaching staff and have scientific work published in journals so that can increase lecturers' insight and provide up-to-date knowledge. In monitoring teaching staff in a tertiary institution, there needs to be data on lecturers who do not yet have a functional position or academic position as a lecturer, therefore it is necessary to group together lecturers who do not yet have a functional position and lecturers who already have a position functional by creating groupings or clusters, it can help lecturers resolve obstacles in the progress of applying for functional positions. In this grouping, the k-means algorithm is used for use in data mining. This algorithm is included in the non-hierarchical method which begins by first determining the desired number of clusters, then the clustering process is carried out. The term k-means was proposed by James McQueen (McQueen, 1967). The k-means algorithm is a partition-based cluster analysis method (Vora, 2013). The k-means working system is a method that can be used to partition objects into groups based on the closeness of the characteristics of each data, so that objects that have the same characteristics are grouped in one cluster and objects that have different characteristics will be grouped into different clusters others (Jhonson, 2007).

Several previous studies have shown the use of clustering methods to group regions or companies based on certain characteristics. Qori'atunnadyah (2024) used the C-Means Clustering method to group districts/cities in East Java based on the number of domestic and foreign tourists, resulting in three clusters for domestic tourists and five clusters for foreign tourists. This study provides insight into the distribution of tourists in the region. In addition, manufacturing companies were grouped based on factors that influence company value using the C-Means method, which formed two groups with medium and high company values, so that it can assist in strategic decision making (Qori'atunnadyah, Liyundira, & Indrianasari, 2024). Furthermore, Qori'atunnadyah (2023) grouped regions in East Java based on the Human Development Index (HDI) using the C-Means method and found four groups with different HDI levels, providing an overview of areas that require more attention. In another study, the Hierarchical Clustering method was applied to group regions in East Java based on road conditions. The results show three clusters, where the majority of areas have good roads, while one cluster shows damaged road conditions (Qori'atunnadyah & Rahmawati, 2022). Finally, the grouping of areas in Lumajang Regency was carried out based on the teacher-student ratio with the K-Means algorithm, resulting in three clusters where areas with high ratios need special attention to improve the quality of education (Qori'atunnadyah, 2022).

This study aims to cluster lecturer data based on key indicators such as functional positions, involvement in research, and contributions to community service using the C-Means Clustering algorithm. The data obtained from the Research and Community Service Institute at Institut Teknologi dan Bisnis Widya Gama Lumajang provides a comprehensive picture of the role of lecturers in carrying out the Tridharma of Higher Education, namely teaching, research, and community service. By applying the C-Means Clustering method, this study attempts to identify groups of lecturers with different levels of productivity and involvement in the three academic aspects. This approach allows institutions to understand the distribution of lecturer performance and identify areas that require further support or professional development. Lecturers who are included in the high-achieving cluster can be used as examples or mentors, while lecturers in the low-

E-ISSN : 2963-0568, P-ISSN : 2963-055X
Available online at:
https://ejournal.itbwigalumajang.ac.id/index.php/jid

achieving cluster can be targeted for strategic interventions to increase research output and involvement in community service. The results of this study are expected to be a guide in policy decision making and resource allocation, so that institutions are able to create a balanced and dynamic academic environment. Ultimately, this study aims to contribute to the continuous improvement of lecturer performance, support institutional goals, and improve the quality of education, research, and community service at Institut Teknologi dan Bisnis Widya Gama Lumajang.

METHOD

The data obtained as the basis for data processing from this research is secondary data obtained from the Community Service Research Institute, Institut Teknologi dan Bisnis Widya Gama Lumajang, which consists of the number of research studies and the number of community service services in 2022 - 2023, then data on lecturer academic positions was obtained. from the UPT Academic Positions and Lecturer Certification of Institut Teknologi dan Bisnis Widya Gama Lumajang, and data on the number of publications of scientific articles during 2022 – 2023 obtained from the UPT Publications of Institut Teknologi dan Bisnis Widya Gama Lumajang.

K-means will work by grouping objects or data into several desired groups, where one object is similar to other objects (Fitri et al., 2023), the k-means algorithm will continue to work with the same pattern by grouping data that is similar to other data according to groups determined into several parts. K-means will try to group lecturers with the criteria of having an academic position, conducting research in the last 2 years, doing community service in the last 2 years, having a number of publications in the last 2 years, with these criteria k-means will form clusters (Sartika & Jumadi, 2019) by producing members whose characteristics are similar to each other. In this research, the method used to group is non-hierarchical.

Non-hierarchical methods are approach methods used in various contexts such as grouping that do not rely on layered structures or hierarchies. In general, non-hierarchical methods emphasize dividing or grouping data or objects according to the characteristics of these objects (Jhonson, 2007). The advantage of the k-means method is that it is very efficient for use in large amounts of data, the weakness of the k-means method is that the number or number of clusters can be determined in advance by the researcher.

The k-means algorithm is a method used in data clustering (unsupervised learning) which works to divide a number of data into groups or what are called clusters which are mutually exclusive based on the similarity or closeness of the values which are based on these features. In general, this algorithm tries to minimize the distance between points and the center of the object or what is called the centroid. The way k-means works starts by first randomly selecting k objects within D or centroids which initially represent the average or cluster center. For the remaining objects, a new cluster will be formed for each object according to the similarity of the object's character based on the distance between the object and the cluster average. Then k-means iteratively increases the variance into clusters. each cluster, will be calculated using the objects assigned to the cluster in the previous iteration (Gorunescu, 2011). This iteration is carried out continuously until it is stable, and the clusters formed at this time are the same as the previous process.

In this study, cluster analysis is used, which is a method for processing data and aims to group various objects based on the similarity of the characteristics of the objects. The k-means algorithm works with the following stages:

1. Input:
   k : Number of clusters
   D : Collection of (n) objects

E-ISSN : 2963-0568, P-ISSN : 2963-055X
Available online at:
https://ejournal.itbwigalumajang.ac.id/index.php/jid

2. Determine the number of clusters *k*
3. Determine the initial position of the centroid for each cluster randomly
4. Assign each data point to a cluster based on the distance closest to the centroid
5. Recalculate the newly formed centroid
6. Repeat this iteration until you get a stable value

Calculating the distance between objects in the centroid continues by inserting objects into clusters based on their closest distance to the centroid. In this case, generally calculate the distance with the following equation:

$$dist = \sqrt{\sum_{j=1}^{k} \sum_{i=c_j} \left(x_i - z_j\right)^2} \quad (1)$$

Where, c_j is the jth cluster and z_j is the centroid of cluster c_j and x_i is the object value (Eliyanto & Surono, 2022).

RESULTS AND DISCUSSION

In cluster analysis using the no-hierarchy method with k-means as the method. Researchers can determine the number of clusters as many as 3 clusters, therefore cluster analysis uses the k-means method with k divided by 3 clusters. The results of cluster analysis using the k-means method are as follows:

Table 1. Initial Cluster Center

	Cluster		
	1	**2**	**3**
Academic position	0	2	1
Research	0	2	2
Community service	0	3	1
Number of publications	0	20	10

Source: Processed data (2024)

The Table 1 illustrates the initial centers of three clusters formed based on several indicators: academic position, research, community service, and number of publications. In the first cluster, lecturers generally have low academic positions or even no formal academic positions, do not conduct research or community service, and have no publications. This shows that lecturers in this cluster have low productivity in fulfilling the tridharma tasks of higher education. The second cluster consists of lecturers with a relatively high level of academic position, active in research and community service, and have a fairly large number of publications (20 publications). Lecturers in this cluster show high productivity and significant involvement in the tridharma tasks.

While in the third cluster, lecturers have moderate academic positions, are quite active in research with several community service activities, and have a moderate number of publications (10 publications). This cluster illustrates productive lecturers with a more balanced focus between research and publication, although their contribution to community service is still limited. The results of this clustering can help universities to identify and understand the productivity profile of lecturers in carrying out the tridharma of higher education, as well as provide special attention to lecturers in the first cluster to be more active in academics, research, and community service. The results can be seen in the table below.

Table 2. Iteration History

Iteration	Change in Cluster Centers		
	1	2	3
1	2.164	2.920	1.071
2	.707	.000	.682
3	.651	.000	.752
4	.168	.000	.219
5	.000	.000	.000

Source: Processed data (2024)

The Table 2 shows the change in cluster centers at each iteration during the clustering process. In the first iteration, there was a significant change in cluster centers, especially in cluster 2 which experienced a change of 2.920, while clusters 1 and 3 experienced changes of 2.164 and 1.071, respectively. In the second iteration, the change in cluster centers decreased significantly, with cluster 2 experiencing no change at all (0.000), indicating stabilization. The third and fourth iterations showed a further decrease in changes in cluster centers, until finally in the fifth iteration all clusters had reached stability, indicated by a change of 0.000 in all clusters. This process indicates that the clustering algorithm has reached convergence in the fifth iteration, where the cluster centers no longer experience changes, indicating that the clusters are stable. This process produces the results of the closest distance between objects as follows.

Table 3. Distance Between Final Cluster Centers

Cluster	1	2	3
1		14.015	7.655
2	14.015		6.370
3	7.655	6.370	

Source: Processed data (2024)

The Table 3 shows the distances between cluster centers after the clustering process is complete. Cluster 1 has the largest distance from cluster 2, which is 14.015, indicating a significant difference between the two clusters. The distance between cluster 1 and cluster 3 is 7.655, while the distance between cluster 2 and cluster 3 is 6.370. These distances indicate that the clusters have different characteristics, where cluster 2 and cluster 3 are more similar to each other than to cluster 1. This can help in interpretation, where cluster 1 may have the most different characteristics compared to the other two clusters. From the results of k-means, all variables used have a significant effect.

Table 4. ANOVA

	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
Academic Position	4.002	2	0.322	47	12.442	0.000
Research	6.671	2	0.354	47	18.821	0.000
Community Service	7.519	2	0.380	47	19.808	0.000
Number of Publications	594.642	2	4.313	47	137.869	0.000

Source: Processed data (2024)

The Table 4 shows the results of the analysis of variance for the variables Academic Position, Research, Community Service, and Number of Publications among the formed clusters. The results show that all variables have significant F values ($p < 0.05$), as indicated by the Sig. value of 0.000 for each variable. This indicates that there are significant differences among the clusters in each of these variables. The high F value for the variable Number of Publications (137.868) indicates that this variable has the most significant differences among the clusters, followed by Community Service (19.808), Research (18.821), and Academic Position (12.442). Thus, these variables significantly differentiate the formed clusters, with Number of Publications being the strongest factor in cluster separation.

Table 5. Observations for Each Cluster

Cluster	1	24.000
	2	6.000
	3	20.000
Valid		50.000
Missing		.000

Source: Processed data (2024)

The Table 5 shows the distribution of the number of observations in each cluster resulting from the cluster analysis. Cluster 1 has the largest number of members, namely 24 observations, followed by cluster 3 with 20 observations, while cluster 2 has the smallest number of members, namely 6 observations. The total valid observations are 50, with no missing data. This distribution shows that the clustering produces three groups of different sizes, where the majority of observations are collected in clusters 1 and 3. Cluster 2, with a smaller number of members, may have more specific or different characteristics than the other clusters.

CONCLUSION

Based on the research results, the clustering of lecturers at the Institut Teknologi dan Bisnis Widya Gama Lumajang for the period 2022 to 2023, based on academic position, research, community service, and number of publications, resulted in three distinct clusters. Cluster 1, with the largest membership of 24 lecturers, predominantly includes those with lower academic positions, minimal engagement in research and community service, and a limited number of publications, indicating lower productivity in fulfilling the tridharma of higher education. Cluster 2, consisting of 6 lecturers, includes those with relatively higher academic positions, actively involved in research and community service, and possessing the highest average number of publications, demonstrating high productivity and engagement. Meanwhile, Cluster 3, with 20 lecturers, represents those with moderate academic positions, balanced involvement in research and community service, and a moderate number of publications.

This clustering analysis provides a valuable overview of lecturer productivity profiles, identifying specific groups where support and interventions may be needed, particularly for lecturers in Cluster 1 to enhance their contributions. For future research, it is recommended to explore the use of alternative clustering methods such as Fuzzy C-Means, which allows for partial membership in clusters and may offer deeper insights into the varying levels of lecturer productivity. Additionally, hierarchical clustering methods (e.g., single linkage, complete linkage) could be considered to compare the accuracy and effectiveness of different approaches in categorizing lecturers based on these key academic criteria.

REFERENCES

Eliyanto, J., & Surono, S. (2022). The suitable distance function for fuzzy C-means clustering. AIP Conference Proceedings, 060006. https://doi.org/10.1063/5.0106185

Fitri, Suryono, R. R., & Wantoro, A. (2023). Klasterisasi data penjualan berdasarkan wilayah menggunakan metode K-means pada PT XYZ. Jurnal Komputasi, 11(2), 157–168. https://doi.org/10.23960/komputasi.v11i2.12582

Gorunescu, F. (2011). Data mining: Concepts, models, and techniques. Springer.

Johnson, R. A., & Wichern, D. W. (2007). Applied multivariate statistical analysis (6th ed.). Pearson.

Kementerian Pendidikan dan Kebudayaan. (2014). Peraturan Menteri Pendidikan dan Kebudayaan No. 92 Tahun 2014 tentang Jabatan Fungsional Dosen.

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability (pp. 281-297). University of California Press.

Pemerintah Indonesia. (2005). Undang-Undang No. 14 Tahun 2005 tentang Guru dan Dosen.

Pemerintah Indonesia. (2021). Peraturan Menteri Pendidikan dan Kebudayaan Nomor 14 Tahun 2021 tentang Pedoman Penetapan Daerah Khusus dalam Pelaksanaan Kebijakan Pendidikan Nasional.

Qori'atunnadyah, M. (2024). Mapping domestic and foreign tourists in East Java using C-means clustering. Jurnal Statistika dan Aplikasinya, 8(1), 54-62.

Qori’atunnadyah, M. (2023). Metode C-means untuk pengelompokkan kabupaten/kota Provinsi Jawa Timur berdasarkan indikator Indeks Pembangunan Manusia (IPM). Journal of Informatics Development, 1(2), 51-58.

Qori’atunnadyah, M. (2022). Pengelompokkan wilayah berdasarkan rasio guru-murid pada jenjang pendidikan menggunakan algoritma K-means. Journal of Informatics Development, 1(1), 33-38.

Qori’atunnadyah, M., Liyundira, F. S., & Indrianasari, N. T. (2024). Grouping manufacturing companies based on factors affecting firm value using C-means clustering. In Conference on SDGs Transformation through the Creative Economy: Encouraging Innovation and Sustainability (TCEEIS 2023) (pp. 28-31). Atlantis Press.

Qori’atunnadyah, M., & Rahmawati, F. D. (2022). Pengelompokkan kabupaten dan kota berdasarkan kondisi infrastruktur jalan menggunakan hierarchical clustering. Journal of Informatics Development, 1(1), 1-5.

Redaksi. (2023). Jabatan akademik dosen berpengaruh pada nilai akreditasi program studi. Dunia Dosen. Retrieved from https://duniadosen.com/jabatan-akademik-dosen-berimplikasi-langsungpada-nilai-akreditasi-program-studi/

Ristekdikti. (2019). Pedoman Operasional Penilaian Angka Kredit (PO PAK) Dosen.

Sartika, D., & Jumadi, J. (2019). Clustering penilaian kinerja dosen menggunakan algoritma K-means (Studi kasus: Universitas Dehasen Bengkulu). Jurnal Teknologi Informasi dan Komunikasi, 4(2), 25–32.

Vora, P. (2013). A survey on K-means clustering and particle swarm optimization. International Journal of Science and Research, 1(3), 96-102.