Institut Riset dan Publikasi Indonesia (IRPI)
IJATIS: Indonesian Journal of Applied Technology and Innovation Science
Journal Homepage: https://journal.irpi.or.id/index.php/ijatis
Vol. 1 Iss. 1 February 2024, pp: 1-8
ISSN(P): 3032-7466 | ISSN(E): 3032-7474

Comparison of Density-Based Spatial Clustering of Applications with Noise (DBSCAN), K-Means and X-Means Algorithms on Shopping Trends Data

Vina Wulandari[1*], Yulia Syarif[2], Zhevin Alfian[3], Muhammad Adil Althof[4], Maylina Mufidah[5]

[1,2,3]Department of Information System, Faculty of Science and Technology, Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia
[1]Puzzle Research Data Technology, Faculty of Science and Technology, Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia
[4]Department of Political Economy, Faculty of Economics and Administrative Sciences, Sivas Cumhuriyet University, Turkey
[5]Department of Food Engineering, Faculty of Engineering, Sakarya University, Turkey

E-Mail: [1]12150321942@students.uin-suska.ac.id, [2]12150321439@students.uin-suska.ac.id, [3]12150311690@students.uin-suska.ac.id, [4]adilalthof0909@gmail.com, [5]maylinamufidahh22@gmail.com

Received Nov 12th 2023; Revised Dec 03th 2023; Accepted Jan 04th 2024
Corresponding Author: Vina Wulandari

Abstract

This study extensively compares the efficacy of three clustering algorithms of DBSCAN, K-Means, and X-Means in analyzing shopping trend data, utilizing the Davies-Bouldin Index (DBI) for group validity assessment. The dataset, sourced from Kaggle.com, encompasses various customer attributes. Results indicate that the DBSCAN algorithm demonstrates superior cluster validity, outperforming K-Means and X-Means. Specifically, with an Eps value of 0.3 and MinPts value of 3, DBSCAN achieves an optimal DBI value of 0.1973. K-Means follows with a DBI value of 2.2958, and X-Means attains its best value (2.5663) with k=3. This research underscores the pivotal role of clustering algorithms in understanding shopping trends and customer preferences, offering valuable insights into their comparative performance.

Keyword: Davies-Bouldin Index, DBSCAN, K-Means, Shopping Data, X-Means

1. INTRODUCTION

Shopping is an activity related to interaction, namely interaction with products that later have the potential to carry out purchasing activities, especially coming directly to retail stores. The popularity of a product can be measured through a number of factors, including frequency of purchase, customer reviews, and social engagement. The role of consumers has a significant impact on the advancement of economic levels, so manufacturers must prioritise their preferences as a primary focus. This involves decisions taken by individuals in choosing between available options, with the possibility of ranking them based on their level of happiness, satisfaction, excitement, pleasure and benefits gained [1]. Customer markets are getting more complex, creating an impetus to direct attention to homogeneous subpopulation groups within an overall heterogeneous market [2]. In this research, we will focus on data mining to do a comparison of algorithms for clustering data.

Finding important connections, patterns, and trends within large data repositories is the process of data mining. It is important for a variety of human endeavors because it may reveal, recognize, and extract previously undiscovered patterns or information. Due to its capabilities, data mining is becoming more and more important in a wide range of application sectors, such as bioinformatics, finance, retail, medical, and insurance. Clustering is one of the data mining techniques used in this study [3][4]. Within the domain of data science, clustering stands as a valuable tool, serving as a method utilized to discern cluster structures within a dataset. These structures are defined by a high degree of similarity among elements within the same cluster and significant dissimilarity between separate clusters. Initially utilized by biological and social scientists, hierarchical clustering was the first clustering technique adopted. However, cluster analysis has since evolved into a subfield within the broader domain of statistical multivariate analysis [5]. This research is important to the shopping world because data mining, especially clustering, can reveal customer buying patterns. By analyzing the cluster structure in shopping datasets, it can improve marketing strategies and provide new insights into consumer behavior. clustering in this study uses the K-Means algorithm, X-Means and DBSCAN.

The first step of the K-Means algorithm involves creating clusters by assigning a K value initially for clustering analysis. Different K values lead to different outputs. Similar characteristics group into one homogeneous cluster, while data with distinct characteristics form separate clusters [6]. X-means clustering addresses a key limitation of K-means, especially the need for prior knowledge of the number of clusters (K). Unlike K-means, X-means autonomously estimates the optimal K value in an unsupervised manner, relying solely on the dataset without predetermined input. Generally, K-Means is widely recognized and commonly used for data clustering [7].

Density-Based Spatial Clustering of Applications with Noise (DBSCAN), an unsupervised clustering technique, utilizes density as a key factor to detect clusters of diverse shapes [8]. A noteworthy example of a robust density-based algorithm is DBSCAN, which operates in two stages: firstly, identifying core points as clusters with high density, and secondly, grouping reachable points around these core points to form clusters. [9].

Previous research conducted in 2022, Comparative analysis of clustering algorithms for determining employee performance using Davies Bouldin Index shows that the K-Means algorithm is the most effective in determining employee performance data within the specified cluster criteria. This algorithm showed the lowest DBI value among the considered algorithms, with a value of 0.377, outperforming the second k-medoids algorithm (DBI = 0.930) and the x-means algorithm (DBI = 0.497) [10].

Then further research in 2020 tested the performance of the DBSCAN algorithm in density-based clustering, detailing the experimental results on three different data sets. The experimental analysis shows that the algorithm has a higher degree of homogeneity and diversity when applied to personalized clustering on data with non-uniform density, especially on data that has highly variable and gradually sparse density values[11].

From the explained details, a comparison of clustering algorithms, namely DBSCAN, k-means, and x-means, will be conducted on shopping data using the Davies-Bouldin Index (DBI) as a cluster validity measure. This aims to determine the best algorithm for grouping the shopping data.

2. MATERIAL AND METHOD

The research phase undergoes multiple steps, and the research flow is depicted in Figure 1. The research methodology is outlined in the same figure. Commencing with data collection, the study then proceeds to preprocess the data by removing irrelevant attributes and addressing noisy data records. Following this, the classification process utilizes the K-Means, DBSCAN, and X-Means algorithms, with the ultimate step involving an evaluation.

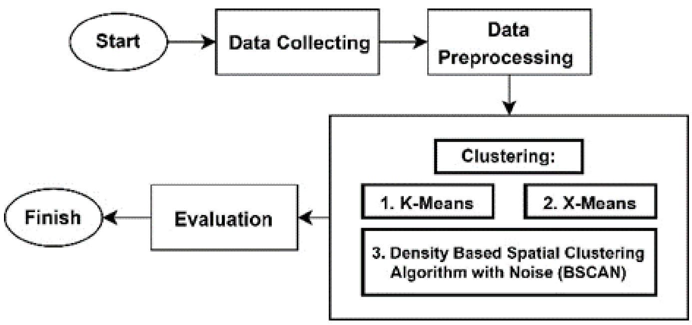

Figure 1. Research Methodology

2.1. Data Collecting

A key obstacle in the development of machine learning is data collecting, which is the subject of research being done by several communities [12]. The data collection process can be done directly or online by searching for datasets that are suitable for the research topic. In this research the dataset is sourced from Kaggle.com.

2.2. Data Pre-Processing

Data pre-processing is a crucial step in the information mining stage [13]. The process of preparing the data requires a significant amount of processing time [14]. Data preparation includes tasks like removing noise from the dataset and replacing any missing values. All of this is done in preparation for the next phase. Data cleansing and modification are among the activities included in the data preparation process [15].

2.3. Clustering

Cluster analysis is a commonly employed unsupervised pattern recognition method that aims to reveal inherent patterns within a dataset [16]. This approach seeks to assess anomalies in data objects by evaluating their relationships with clusters. Anomalies are identified among data objects that do not belong to any specific cluster or are included in distant clusters, categorizing them as anomaly points. In contrast to alternative approaches, clustering-based methods are versatile across various types of data and exhibit varying effectiveness. Detection tasks can be accomplished without requiring comprehensive knowledge of anomaly detection [17].

2.4. K-Means

Data is grouped into one or more clusters using the non-hierarchical K-Means clustering technique depending on how comparable the attributes are between the individual data points. Different clusters are created from data with unique properties[18]. Samples are first allocated to the cluster's closest center. To begin, the sample's distance from the beginning center is determined. Next, an overall sample average is obtained for each cluster. Finally, the process is repeated until the centers converge and the difference between successive iterations hits a threshold that is small enough [19] The Euclidean distance between two points, p and q, in an n-dimensional space, equation 1 and 2.

$$D(p,q) = \sqrt{\sum_{i=1}^{n}(pi - qi)2} \tag{1}$$

Cluster Center (Centroid):

$$Ci = \frac{1}{|Si|}\sum_{x \in Si} \chi \tag{2}$$

The K-Means algorithm aims to minimize the variance within each cluster, measured as the sum of squared Euclidean distances between each data point and its cluster center. The objective function formula 3 optimized by K-Means.

$$J = \sum_{i=1}^{K} \sum_{x \in Si} D(x, Ci)2 \tag{3}$$

The primary objective of this algorithm is to minimize the value of the objective function J. The iterative process mentioned above converges to a solution where the cluster assignments and cluster centers do not change significantly.

2.5. DBSCAN

DBSCAN is a density-based data clustering method, where clustering is based on the number of data points (minimum points) contained within a radius Eps (ε) around each data point [20] The concept of density leads to three states for each data point: noise points (inaccessible by core points and not part of the boundary), core points (centers of clusters based on density, requiring a certain number of points within Eps), and boundary points (serving as boundaries around cluster centers). DBSCAN conducts data clustering based on input parameters epsilon and minpts. The resulting clusters from DBSCAN depend on the values set for these two parameters. In the clustering procedure of the DBSCAN method, Euclidean distance is employed to calculate the distance from a point to a randomly pre-determined centroid (C) [21]

Ester Martin is the developer of the data clustering technique known as the DBSCAN algorithm [22] and it focuses on clustering based on data density. DBSCAN can identify clusters with different shapes at particular density thresholds. The essential principles of DBSCAN include the following concepts [23]:

1. Start by setting up initial input parameters, such as the minimum number of points required in a cluster (MinPts) and the allowable distance between points within a cluster (Epsilon).
2. Next, choose a random starting point (p).
3. Compute the Eps distance (d_ij) between point p and all other points considered "density reachable" from p using the Euclidean distance formula 4.

$$d_{ij} = \sqrt{\sum_{a}^{p}(x_{ia} - x_{ja})^2} \tag{4}$$

Where the variable x_ia refers to the a-th attribute of object i, and d_ij is the result of the Euclidean distance.

4. If there are over MinPts points within the Eps distance, point p is designated as a "core point," resulting in the formation of a cluster.

The primary objective of this algorithm is to minimize the value of the objective function J. The iterative process mentioned above converges to a solution where the cluster assignments and cluster centers do not change significantly.

2.6. X-Means

X-means clustering addresses a primary limitation of K-means clustering, which necessitates prior knowledge about the number of clusters (K). Unlike K-means, X-means does not require a predetermined K value; instead, it estimates the optimal K value autonomously based solely on the dataset. The method employs Kmax and Kmin as the upper and lower boundaries for potential values of K. In the initial stage of X-means clustering, with X set to Xmin, the algorithm identifies the initial structure and centroids. Subsequently, each cluster within the estimated structure serves as a parent cluster, subject to further division into two groups in the subsequent step [24].

The X-Means algorithm equation is developed through two steps [25]:

1. Initialization of the cluster values (k).
2. Following this, the geometric distance formula is employed to compute the distance between every cluster center and each individual data point.

$$d_{ij} = \sqrt{\sum_{k}^{n} = 1\ (\mathcal{X}_{ik}\ -\ \mathcal{X}_{jk})} \tag{5}$$

3. The next step involves determining new cluster centers for each k value by calculating the average attribute values within the cluster.
4. The subsequent step is to optimize the values of each cluster obtained by calculating their characteristics.
5. In the final step, the process loops back to the second step until there is no further exchange with other clusters or until reaching the maximum iteration limit.

2.7. Unsupervised Learning

Unsupervised machine learning (ML) techniques have a crucial role in analyzing raw datasets, assisting in the extraction of analytical insights from unlabeled data. Significant advancements in hierarchical learning, clustering algorithms, factor analysis, latent models, and outlier detection have substantially advanced the forefront of unsupervised ML methods. Notably, recent progress in unsupervised ML, exemplified by the introduction of "deep learning" techniques, has played a significant role in driving this evolution [26].

2.8. DBI

The Davies-Bouldin (DB) validity index calculates the average of the average values for each data point in a dataset. The computation involves summing up compactness values for each point and then dividing this sum by the distance between the respective cluster center points, serving as a measure of separation [27].

3. RESULTS AND DISCUSSION

3.1. Data Collecting

The dataset requirement for this research is fulfilled through Kaggle. The Customer Shopping Preference Dataset is the source of the study's data. Customer age, gender, purchase amount, preferred payment method, frequency of purchases, and feedback rating are some of its characteristics. Furthermore, information is provided on the type of items purchased, shopping frequency, preferred shopping seasons, and how they interact with special offers with a 3899-record library and 18 field indicated in Table 1.

Table 1. Dataset for research

Customer ID	Age	Gender	Item Purchased	Category	Purchase Amount (USD)	Location	…	Frequency of Purchases
1	55	Male	Blouse	Clothing	53	Kentucky	…	Fortnightly
2	19	Male	Sweater	Clothing	64	Maine	…	Fortnightly
3	50	Male	Jeans	Clothing	73	Massachusetts	…	Weekly
4	21	Male	Sandals	Footwear	90	Rhode Island	…	Weekly

Customer ID	Age	Gender	Item Purchased	Category	Purchase Amount (USD)	Location	…	Frequency of Purchases
5	45	Male	Blouse	Clothing	49	Oregon	…	Annually
…	…	…	…	…	…	…	…	…
3895	66	Female	Skirt	Clothing	78	Connecticut		Weekly
3896	40	Female	Hoodie	Clothing	28	Virginia		Bi-Weekly
3897	52	Female	Backpack	Accessories	49	Lowa		Quarterly
3898	46	Female	Belt	Accessories	33	New Jersey		Weekly
3899	44	Female	Shoes	Footwear	77	Minnesota	…	Quarterly

3.2. DBSCAN Algorithm

The clustering process is carried out through the implementation of the DBSCAN algorithm with experiments involving different values of epsilon (Eps) and MinPoints (MinPts). The range of Eps values explored in this study is between 0.3 and 0.5 with a MinPts value of 5, while for the range of 0.3 to 0.6, the MinPts value used is 3. Subsequently, cluster validity testing is conducted using the Davies-Bouldin Index (DBI). The DBI values based on the clustering results using DBSCAN can be observed in Figures 2 and 3.

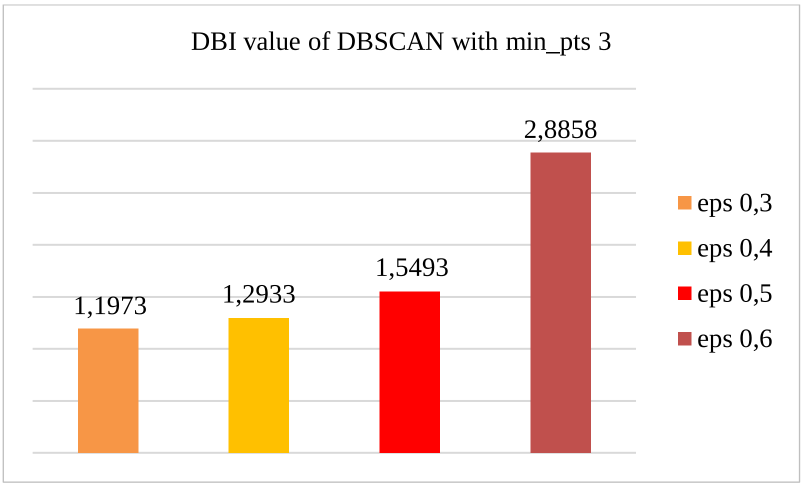

Figure 2. DBI Value with MinPts 3

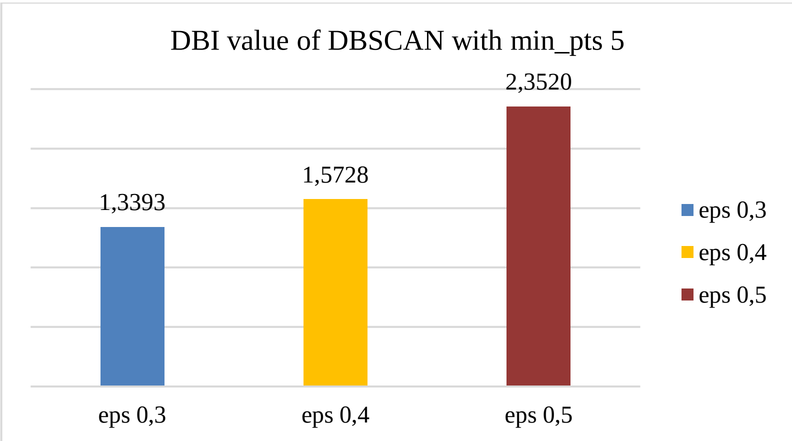

Figure 3. DBI Value with MinPts 5

Figures 2 and 3 show the test results of the DBSCAN algorithm using the Davies-Bouldin Index (DBI) to evaluate the validity of the best cluster results. The experiment results show that the DBI reaches the best value of 1.1973 when using Min_Pts 3 and eps value 0.3 and and eps value 0.3 and can be seen in Figure 2.

3.3. K-Means Algorithm

Based on the K-Means algorithm calculations, cluster validity is assessed using the Davies-Bouldin Index (DBI) method with cluster experiments ranging from k=2 to k=11. The DBI values from the calculations using the K-Means algorithm are presented in the following Figure 4.

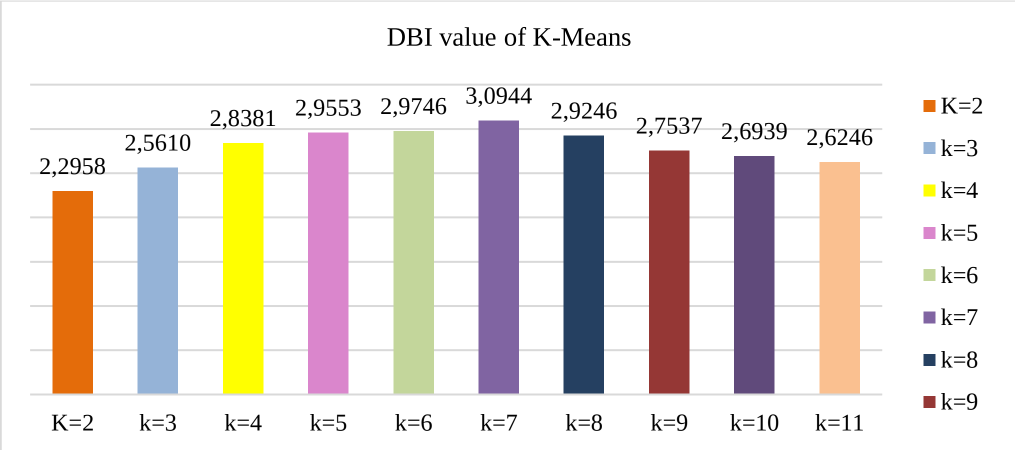

Figure 4. DBI Value of K-Means

In Figure 4, it can be seen testing the validity of cluster results in the k-means algorithm using DBI, the experiment was carried out from cluster 2 to cluster 10, from 9 times the experiment obtained the best cluster validity results, namely in cluster 2 with a validity value of 2.2958.

3.4. X-Means

Data clustering experiments were conducted using the X-means algorithm. The clustering process was carried out from cluster 3 to cluster 11 with a total of 9 experiments. The experiment was validated using DBI to see the best clusters that could be generated. Of the nine experiments, the best cluster was cluster 3 with a DBI value of 2.5663. The value of X-Means can be shown in Figure 5.

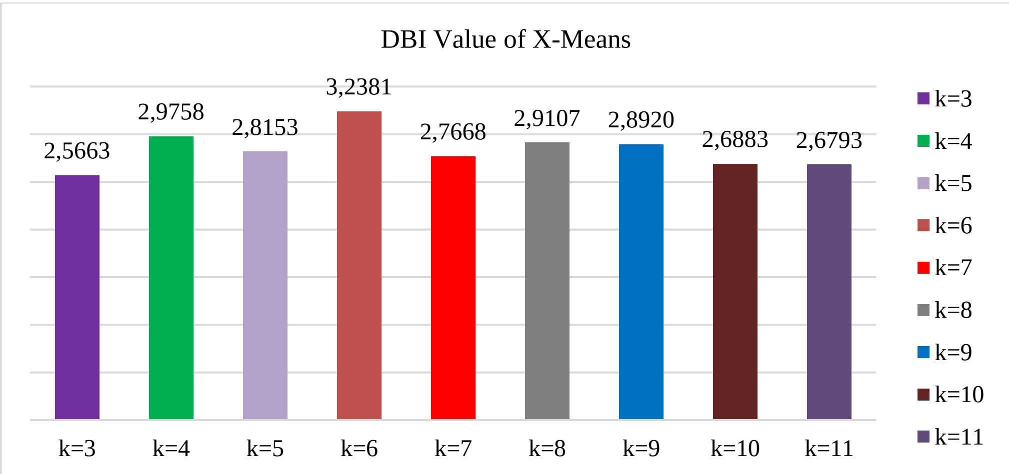

Figure 5. DBI Value of X-Means

3.5. Comparison Results of the Algorithms

Comparison of the DBSCAN, K-Means, and X-Means algorithms is conducted based on the quality of the clusters formed, as determined by the calculation of the Davies-Bouldin Index (DBI) values. The following are the results of the graph comparison of DBI values to determine the best cluster quality. The value of comparison algorithm can be shown in Figure 6.

The best cluster validity results between the DBSCAN, K-Means and X-Means algorithms, using the Davies-Bouldin Index (DBI) method, reveal that the DBSCAN algorithm attains its optimum value in the experiment with Eps value of 0.3 and a MinPts value of 3, yielding a value of 0.1973. Meanwhile, the K-Means algorithm obtained the best value of 2.2958. On the other hand, the X-Means algorithm achieves its best value in the experiment with k=3 yielding a value of 2,5663. Therefore, in this study, it is evident that the DBSCAN algorithm exhibits more optimal cluster validity results compared to theK-Means and X-Means algorithm.

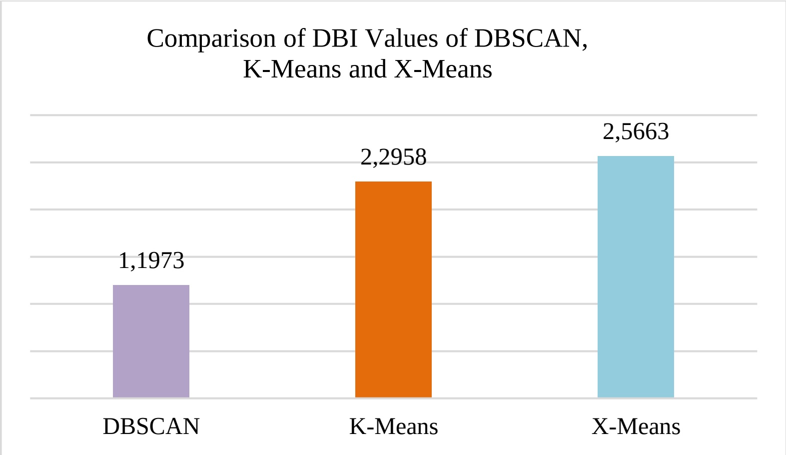

Figure 6. Graph of Comparison Results of the Algorithms

4. CONCLUSION

This study compares DBSCAN, K-Means, and X-Means algorithms based on cluster validity using the Davies-Bouldin Index (DBI) method. The DBSCAN algorithm showed the most optimal cluster validity results, followed by the K-Means algorithm and then the X-Means algorithm. The optimal DBI value for DBSCAN algorithm was obtained with Eps value of 0.3 and a MinPts value of 3, while K-Means algorithm achieved the best value with k=2. The X-Means algorithm achieved the best value with k=3. This study provides insight into the performance of these clustering algorithms.

REFERENCES

[1] R. Kirk, “Lifestyle centers, the next boom and bust after shopping malls? Governance, public-private partnerships, and Guy Debord’s spectacle in Dallas-Fort Worth,” *Cities*, vol. 133, Feb. 2023, doi: 10.1016/j.cities.2022.104155.

[2] M. Akkaya, “Understanding the impacts of lifestyle segmentation & perceived value on brand purchase intention: An empirical study in different product categories,” *European Research on Management and Business Economics*, vol. 27, no. 3, Sep. 2021, doi: 10.1016/j.iedeen.2021.100155.

[3] M. K. Gupta and P. Chandra, “A comprehensive survey of data mining,” *International Journal of Information Technology*, vol. 12, no. 4, pp. 1243–1257, Dec. 2020, doi: 10.1007/s41870-020-00427-7.

[4] P. D. Inas Azizah, “Penerapan Probabilistic Neural Network pada Klasifikasi Berat Bayi Baru Lahir,” *Jurnal Riset Statistika*, vol. 1, no. 2, pp. 152–159, Feb. 2022, doi: 10.29313/jrs.v1i2.524.

[5] K. P. Sinaga and M.-S. Yang, “Unsupervised K-Means Clustering Algorithm,” *IEEE Access*, vol. 8, pp. 80716–80727, 2020, doi: 10.1109/ACCESS.2020.2988796.

[6] M. Ahmed, R. Seraj, and S. M. S. Islam, “The k-means algorithm: A comprehensive survey and performance evaluation,” *Electronics (Switzerland)*, vol. 9, no. 8. MDPI AG, pp. 1–12, Aug. 01, 2020. doi: 10.3390/electronics9081295.

[7] M. Mughnyanti, S. Efendi, and M. Zarlis, “Analysis of determining centroid clustering x-means algorithm with davies-bouldin index evaluation,” *IOP Conf Ser Mater Sci Eng*, vol. 725, no. 1, p. 012128, Jan. 2020, doi: 10.1088/1757-899X/725/1/012128.

[8] K. S. Pranata, A. A. S. Gunawan, and F. L. Gaol, “Development clustering system IDX company with k-means algorithm and DBSCAN based on fundamental indicator and ESG,” in *Procedia Computer Science*, Elsevier B.V., 2022, pp. 319–327. doi: 10.1016/j.procs.2022.12.142.

[9] S. Chowdhury, N. Helian, and R. Cordeiro de Amorim, “Feature weighting in DBSCAN using reverse nearest neighbours,” *Pattern Recognit*, vol. 137, May 2023, doi: 10.1016/j.patcog.2023.109314.

[10] J. Penerapan, T. Informasi, D. Komunikasi, G. B. Kaligis, and S. Yulianto, “IT-Explore Analisa Perbandingan Algoritma K-Means, K-Medoids, dan X-Means Untuk Pengelompokkan Kinerja Pegawai (Studi Kasus: Sekretariat DPRD Provinsi Sulawesi Utara),” 2022.

[11] D. Deng, “DBSCAN Clustering Algorithm Based on Density,” in *2020 7th International Forum on Electrical Engineering and Automation (IFEEA)*, IEEE, Sep. 2020, pp. 949–953. doi: 10.1109/IFEEA51475.2020.00199.

[12] Y. Roh, G. Heo, and S. E. Whang, “A Survey on Data Collection for Machine Learning: A Big Data - AI Integration Perspective,” *IEEE Trans Knowl Data Eng*, vol. 33, no. 4, pp. 1328–1347, Apr. 2021, doi: 10.1109/TKDE.2019.2946162.

[13] H. Huang, B. Wei, J. Dai, and W. Ke, “Data Preprocessing Method For The Analysis Of Incomplete Data On Students In Poverty,” in *2020 16th International Conference on Computational Intelligence and Security (CIS)*, IEEE, Nov. 2020, pp. 248–252. doi: 10.1109/CIS52066.2020.00060.

[14] S.-A. N. Alexandropoulos, S. B. Kotsiantis, and M. N. Vrahatis, “Data preprocessing in predictive data mining,” *Knowl Eng Rev*, vol. 34, p. e1, Jan. 2019, doi: 10.1017/S026988891800036X.

[15] M. N. Sidqi, D. P. Rini, and S. Samsuryadi, “Optimization of Deep Neural Networks with Particle Swarm Optimization Algorithm for Liver Disease Classification,” *Computer Engineering and Applications*, vol. 12, no. 1, 2023.

[16] A. Gere, “Recommendations for validating hierarchical clustering in consumer sensory projects,” *Curr Res Food Sci*, vol. 6, Jan. 2023, doi: 10.1016/j.crfs.2023.100522.

[17] J. Fang, Z. Xie, H. Cheng, B. Fan, H. Xu, and P. Li, “Anomaly detection of diabetes data based on hierarchical clustering and CNN,” in *Procedia Computer Science*, Elsevier B.V., 2021, pp. 71–78. doi: 10.1016/j.procs.2022.01.010.

[18] R. Cordeiro de Amorim and V. Makarenkov, “On k-means iterations and Gaussian clusters,” *Neurocomputing*, vol. 553, Oct. 2023, doi: 10.1016/j.neucom.2023.126547.

[19] Q. Huang, S. Chen, and Y. Li, “Selection of seismic noise recording by K-means,” *Case Studies in Construction Materials*, vol. 19, Dec. 2023, doi: 10.1016/j.cscm.2023.e02363.

[20] M. M. Putri, C. Dewi, E. Permata Siam, G. Asri Wijayanti, N. Aulia, and R. Nooraeni, “Comparison of DBSCAN and K-Means Clustering for Grouping the Village Status in Central Java 2020 Komparasi DBSCAN dan K-Means Clustering pada Pengelompokan Status Desa di Jawa Tengah Tahun 2020,” *Jurnal Matematika, Statistika dan Komputasi*, vol. 17, no. 3, pp. 394–404, 2021, doi: 10.20956/j.v17i3.11704.

[21] M. Civera, L. Sibille, L. Zanotti Fragonara, and R. Ceravolo, “A DBSCAN-based automated operational modal analysis algorithm for bridge monitoring,” *Measurement (Lond)*, vol. 208, Feb. 2023, doi: 10.1016/j.measurement.2023.112451.

[22] W. Jing, C. Zhao, and C. Jiang, “An improvement method of DBSCAN algorithm on cloud computing,” *Procedia Comput Sci*, vol. 147, pp. 596–604, 2019, doi: 10.1016/j.procs.2019.01.208.

[23] N. Nurhaliza and Mustakim, “Clustering of Data Covid-19 Cases in the World Using DBSCAN Algorithms,” *IJIRSE: Indonesian Journal of Informatic Research and Software Engineering*, vol. 1, no. 1, pp. 01–08, 2021.

[24] M. Mughnyanti, S. Efendi, and M. Zarlis, “Analysis of determining centroid clustering x-means algorithm with davies-bouldin index evaluation,” in *IOP Conference Series: Materials Science and Engineering*, Institute of Physics Publishing, Jan. 2020. doi: 10.1088/1757-899X/725/1/012128.

[25] R. Adhitama, A. Burhanuddin, and A. Febriani, “Penerapan X Means Clustering Pada UMKM Kab Banyumas Yang Mendukung Mega Shifting Consumer Behavior Akibat Covid-19,” *Journal of Informatics, Information System, Software Engineering and Applications*, vol. 4, no. 1, pp. 71–80, 2021.

[26] M. Usama *et al.*, “Unsupervised Machine Learning for Networking: Techniques, Applications and Research Challenges,” *IEEE Access*, vol. 7, pp. 65579–65615, 2019, doi: 10.1109/ACCESS.2019.2916648.

[27] I. Rizky Mahartika and A. Wibowo, “Data Mining Klasterisasi dengan Algoritme K-Means untuk Pengelompokkan Provinsi Berdasarkan Konsumsi Bahan Bakar Minyak Nasional,” in *Prosiding Seminar Nasional Sisfotek (Sistem Informasi dan Teknologi)*, 2019, pp. 89–91.