Engineering Technology
Open Access Journal

Juniper PUBLISHERS key to the Researchers

Mini Review Article
Volume 6 Issue 1 - June 2024
DOI: 10.19080/ETOAJ.2024.06.555676

Eng Technol Open Acc

A Comparison of Machine Learning Algorithms (K-means) Using Data of Anterior Cruciate Ligament Reconstruction (ACLR) Patients

Bader E. Alabkary*[1], Galelah H. Alawami[2], Christopher J. Kobus[3], Suzan Kamel-ElSayed[4], Hoda S. Abdel-Aty Zohdy[1], Mohamed A. Zohdy[1]

[1]*Department of Engineering and Computer Science, Oakland University, Rochester, MI 48309*

[2]*Department of Teacher Education, University of Toledo, Toledo, OH 43606*

[3]*Department of Mechanical Engineering, Oakland University, Rochester, MI 48309*

[4]*Department of Foundational Medical Studies, Oakland University, Rochester, MI 48309*

Submission: May 30,2024; **Published:** June 24 ,2024

***Corresponding author:** Bader E. Alabkary, Department of Electrical and Computer Engineering, Oakland University, Rochester, MI 48309

Suzan Kamel-ElSayed, Department of Foundational Medical Studies, Oakland University, Rochester, MI 48309, Hoda S. Abdel-Aty Zohdy, Department of Electrical and Computer Engineering, Oakland University, Rochester, MI 48309 and, Mohamed A. Zohdy, Department of Electrical and Computer Engineering, Oakland University, Rochester, MI 48309.

Abstract

This research provides a comprehensive comparison of different machine learning (k-means clustering) algorithms as applied to datasets derived from Anterior Cruciate Ligament Reconstruction (ACLR) patients. The study aimed to find optimal clustering methodologies for ACLR patients' data, focusing on knee measurements and their association with demographic factors such as age and physical activities. The researchers gathered data from 20 ACLR patients and 20 normal participants in Saudi Arabia's eastern region. The researchers used machine learning algorithms (K-means clustering) for MATLAB programs. The study's findings highlight the differences in clustering results among k-means algorithms. A comparison of the impact of surgery and rehabilitation between normal and ACLR patient groups shows a definite variance in activity levels and knee functionality clusters. The normal group had more identical clustering, while the ACLR group revealed greater variability. This research advances our understanding of optimal clustering methodologies for ACLR patient data, offering valuable insights for precision medicine in orthopedics. The identified patterns and correlations contribute to the ongoing discourse in medical data analysis and evince implications for personalized rehabilitation strategies. Overall, the study provides a foundation for further exploration and refinement of clustering techniques in the context of anterior cruciate ligament reconstruction.

Keywords: AI, Machine learning, K-means, Elbow methods, Cruciate Ligament Reconstruction (ACLR)

Abbreviations: ACLR: Anterior Cruciate Ligament Reconstruction; ML: Machine Learning; MCID: Minimal Clinically Important Difference; IKDC: International Knee Documentation Committee; ENPLR: Elastic-Net Penalized Logistic Regression; LASSO: Least Absolute Shrinkage and Selection Operator; AUC: Area under the Curve; NKLR: Norwegian Knee Ligament Register; AM: Anteromedial; PL: Posterolateral; WT: Weight; HT: Height; SSE: Squared Errors; WCSS: Within-Cluster Sum of Squares

Introduction

Medical research has increasingly witnessed reliance on developing data analysis techniques to unravel intricate patterns within patient datasets [1]. Healthcare research has recently witnessed the use of machine learning (ML) techniques that can potentially improve predictive capabilities [1]. Machine learning is a type of artificial intelligence that has a major difference from standard statistical modeling, its methodology focuses on repeatable and accurate predictions rather than merely interpreting data [2]. Thus, ML applies techniques to navigate the complicated relationships between multiple variables to predict an outcome. It can process convoluted variable relationships more easily than traditional statistical techniques [1]. ML algorithms absorb information from available data and modify their internal parameters to strengthen associations with greater accuracy. This results in superior predictive ability [3]. Such accurate data processing can yield rich dividends in the field of orthopedics [4]. Specifically, the application of ML algorithms seems to be a promising approach for diagnosis and prediction in patients suffering from Anterior Cruciate Ligament (ACL) issues [5].

ACL injuries are common [6] and often occur in accidents connected to active sports such as football, soccer, basketball, handball, and skiing [7]. When ACL injuries occur, they may result in increased instability along the anterior-posterior translation at extension. Damage to the ACL in the knee may result in various injuries such as meniscal tears, chondral disruption, osteophytosis, and synovium hyperplasia [7]. Thus, it is ideal to opt for operative treatment instead of the more conservative options for active younger adults or patients who suffer from multiple ligament issues. Note that the ACL rarely self-repairs due to its low healing abilities [7]. Failure to address the surgical transection of the ACL can lead to posttraumatic osteoarthritis [6]. The number of orthopedic procedures performed in outpatient wards in the United States has steadily increased since the early 1980s, with ACLR accounting for more than 95% of such surgeries [3]. However, there is no guarantee that ACLR surgery is 100% foolproof. Sports players rely on ACLR surgery to quickly restore knee joint stability and resume their activities [8]. Other factors to consider include age, gender, pre-injury activity level, ability to overcome fear, psychological willingness, graft type, and type of sport, which may impact how soon a patient is able to return to sports [9].

Optimizing the analysis and interpretation of ACL injury data is critical to improving our understanding of the condition. This would ultimately lead to better outcomes for patients suffering from ACL issues. Utilizing an accurate predictive model would be of value to both orthopedic surgeons and patients undergoing ACLR surgeries. This would allow patient and surgical information to guide shared clinical decision-making when it comes to patient-specific management. As noted earlier, ML predictive algorithms are best suited to provide accurate ACL patient-specific predictive data because they can interpret complex datasets [1]. Clustering is one such machine learning method; it is a specific domain within the field of data mining that belongs to the category of unsupervised learning. Cluster analysis is a method that categorizes data by examining the underlying structures of items within a dataset and their connections [10]. Data classification, a fundamental method for organizing and comprehending data, involves categorizing or clustering data into different groups [11]. Adopting the K-means clustering ML algorithm approach offers a promising option to glean complex patterns emerging from ACLR patients' data. There is limited extant research on using the K-means algorithm to predict outcomes after ACLR based on knee data, especially in the Middle East.

Therefore, the main objective of this research paper is to examine the effectiveness and performance of K-means clustering. The researchers will conduct a comparative analysis of K-means clustering to identify discrete clusters and patterns of data obtained from ACLR patients. The dataset will focus on knee measurements and demographic factors such as age and activity levels. This study will enumerate the strengths and limitations of each algorithm for handling ACLR patients' data by assessing the efficacy and performance of different K-means variants. This study's significance will be beneficial for the medical and engineering fields, and such insights will assist clinicians, researchers, and healthcare practitioners in better understanding the critical nuances within ACLR patient profiles. The next section of this research will highlight the important studies done on machine learning algorithms with ACLR patients. Then, it will go into detail about the methodology used and the results of this study.

Application of Machine Learning Algorithms to ACLR Patients

Extant studies have explored applications of machine learning algorithms to orthopedic injuries - including ACL tears - with promising results, such as improved accuracy in diagnosis, personalized treatment plans, and enhanced predictive modeling of patient outcomes. That has helped lay the groundwork for this research to build upon and expand understanding of ACL patient care through machine learning. Kunze KN et al's study aimed to develop ML algorithms to predict whether patients undergoing ACLR would achieve the minimal clinically important difference (MCID) on the International Knee Documentation Committee (IKDC) score at a minimum 2-year follow up [2]. The researchers retrospectively analyzed data from an ACLR registry involving 442 patients. Thirty-six variables were considered, while the study population was divided between training and testing sets. Six ML algorithms were employed and evaluated for performance using various metrics. The results identified specific preoperative and intraoperative factors such as body mass index, medial collateral ligament examination grade, femoral tunnel fixation method, prior contralateral knee surgery, and preoperative knee extension, as significant predictors of achieving MCID. The elastic-net penalized logistic regression (ENPLR) algorithm exhibited the best predictive ability, emphasizing the utility of ML in assessing patient outcomes after ACLR.

The study by D.S. Chen et al. introduced the novel approach of integrating unsupervised ML and supervised ML-derived radiomics in the context of ACL ruptures [12]. The research involves 68 patients, with demographic and radiomics features recorded and used as input for ML algorithms. The unsupervised algorithm classified patients into five groups revealing, among others, significant differences in ACL rupture incidences and left knee involvements. Utilizing a supervised ML algorithm, a radiomics model was constructed, and through feature selection methods like t-test and least absolute shrinkage and selection operator (LASSO), 43 radiomics features were identified. The combination of LASSO and random forest classifier demonstrated the highest sensitivity, specificity, accuracy, and area under the curve (AUC), highlighting its efficacy in predicting ACL rupture. The study validated the clinical application of unsupervised ML in ACL rupture and identified seven radiomics features as potential predictors for this condition, suggesting the promising utility of radiomics in predicting ACL injuries.

On their part, researchers R. K. Martin et al. aimed to enhance the prediction of ACLR revision by employing ML analysis on the Norwegian Knee Ligament Register (NKLR) [1]. Examining data of 24,935 patients with a mean follow-up of 8.1 years, the study developed four ML models, including Cox LASSO, survival random forest, generalized additive model, and gradient boosted regression. All models demonstrated moderate concordance and were well-calibrated, with the Cox LASSO model requiring only five

How to cite this article: B. Alabkary, G. Alawami, C. Kobus, S. Kamel-ElSayed, H. S. Abdel-Aty Zohdy, M. Zohdy Comparison of Machine Learning Algorithms (K-means) Using Data of Anterior Cruciate Ligament Reconstruction (ACLR) Patients. Eng Technol Open Acc 2024; 6(1): 555676.
DOI: 10.19080/ETOAJ.2024.06.555676

variables for accurate outcome prediction. A novel in-clinic calculator, the Revision Risk Calculator, was created to estimate individualized risk levels for ACL revision, ranging from near 0% for low-risk patients to 20% for high-risk patients within five years. The study concluded that ML analysis of a national knee ligament registry can moderately accurately predict the risk of ACLR revision, offering the potential for a practical in-clinic tool for personalized risk stratification based on minimal input variables.

Furthermore, Corban et al. conducted a systematic review to investigate the current and potential applications of artificial intelligence (AI) in managing ACL injuries within orthopedic surgery [4]. As technological innovation becomes integral to the field, the study explored the future role of AI, leveraging ML algorithms with adaptive learning and problem-solving capabilities to enhance accuracy. Nineteen publications were included in that systematic review, categorizing AI applications into prediction, diagnosis, intraoperative utilization, and postoperative care and rehabilitation. The applications spanned image interpretation, automated chart review, physical examination assistance through optical tracking, generation of predictive models, and optimization of postoperative care. While the review reflected a growing interest in AI among orthopedic surgeons and some studies showcasing comparable or superior outcomes, it acknowledged the existence of challenges that needed to be addressed before widespread adoption of this technology in orthopedic practice.

Uozumi, Y et.al conducted a study to introduce an automated segmentation approach for assessing the anteromedial (AM) and posterolateral (PL) regions of bone tunnels after double bundle ACLR [7]. The procedure, which typically involves creating two bone tunnels, is crucial for knee injuries. Utilizing K-means clustering, the study evaluated six patients (mean age 27 ± 7, four males/two females). The proposed method successfully segmented the bone tunnels for all patients, indicating its efficacy in delineating the regions after double bundle ACLR. In summary, the study suggested the introduced segmentation method was suitable for analyzing bone tunnels for double bundle ACLR. All the aforementioned studies have explored applications of ML algorithms in orthopedic injuries including ACL tears. The aim of the current study is to understand optimal clustering methodologies for ACLR patients' data, focusing on knee measurements and their association with demographic factors such as age, weight, height and physical activities.

Methodology

Machine learning

The concept of Artificial Intelligence (AI) was first propounded by John McCarthy in 1956, who argued it would reproduce human intelligence using computers [13]. Machine learning (ML) is a type of AI that uses computational algorithms which are adaptable and developed with experience. The two chief forms of ML are supervised and unsupervised algorithms [13]. Supervised learning involves training an algorithm to form associations using already labeled data, whereas unsupervised algorithms create complex new associations within unlabeled datasets [4]. K-means data clustering is a type of unsupervised machine learning that can divide unlabeled data into a set number of separate groups with equal variances. These groups are called clusters, and they are formed by similarities [14]. This algorithm was developed by MacQueen in 1967 [15]. The clusters are formed by classifying a given dataset. The main idea is to define K centers, one for each cluster. These centers should be placed carefully since discrete locations provide different results [16]. These functional objectives of K-means clustering apply to certain non-numeric data sets, too. The primary advantages of adopting the K-means clustering approach are brevity, efficiency, and promptness. This approach defines the total within-cluster variation as the sum of the squared distances between items and the `corresponding centroid [17].

$$W\left(C_k\right) = \sum_{x_i \in C_k} \left(x_i - \mu_k\right)^2$$

This formulation is a specification in which W (C_k) represents the within-cluster total; xi is a data point for a cluster, Ck indicates a cluster for each data point, and μ_k defines the mean value of the points that is assigned to the cluster C_k. Therefore, the sum of total within-clusters of the sum of squares measures compactness (TW) as follows:

$$TW = \sum_{K=1}^{K} W\left(C_k\right) = \sum_{K=1}^{K} \sum_{x_i \in C_k} \left(x_i - \mu_k\right)^2$$

In the current study, the researchers used the MATLAB program, a technical computer program that uses high-performance language. It integrates computation, visualization, and programming easily, and can solve problems using familiar mathematical notations [18]. The researchers applied k-means clustering, an unsupervised ML algorithm in MATLAB, to manage the data through the coding process.

Research Question

How does the k-means clustering algorithm expose distinct clusters within datasets containing information on age, activities, and knee size for both ACLR and normal groups? The purpose of this question was to compare the ACLR group with the normal knee function group. Furthermore, this study investigated the expected effects of age, activity level, and knee size on the formation of distinct clusters within the datasets of these two groups.

Research Hypotheses

Null Hypothesis (H0)

There is no significant difference in the clusters identified by the k-means clustering algorithm between the ACLR and normal groups based on variables such as age, activities, and knee size.

How to cite this article: B. Alabkary, G. Alawami, C. Kobus, S. Kamel-ElSayed, H. S. Abdel-Aty Zohdy, M. Zohdy Comparison of Machine Learning Algorithms (K-means) Using Data of Anterior Cruciate Ligament Reconstruction (ACLR) Patients. Eng Technol Open Acc 2024; 6(1): 555676.
DOI: 10.19080/ETOAJ.2024.06.555676

Alternative Hypothesis (H1)

There is a significant difference in the clusters identified by the k-means clustering algorithm between the ACLR and normal groups based on variables such as age, activities, and knee size.

Data Collection

The data was collected from a published thesis study titled "Knee Function Post Anterior Cruciate Ligament (ACLR)". The thesis was defended by Ali Mohammed Ismail Alyami [19]. The researchers used the following criteria to select this published thesis database for machine learning (k-means clustering): 1. Based on the age range of the population most frequently affected by ACLR injuries (20 to 45 years), the researchers chose data that indicated middle-age knee problems. 2. The researchers argued there is a lack of medical research on this population. 3. The researchers also argued that there is a lack of use of a machine learning k-mean clustering algorithm in ACLR patients' research, so further medical studies on the efficiency of k-mean clustering are important. In the current study, the researchers used Ali's data to apply k-means clustering to reveal distinct clusters within datasets containing information on the age, activities, and knee size of ACLR patients and normal adults.

Study Participants

The participants in this study were 40 adults who lived in the eastern region of Saudi Arabia. Participants were male, and their ages were between 18 to 45 years. The participants were divided into two groups - the ACLR patients' group and the normal group - each comprising 20 participants. The ACLR group had previously undergone ACLR surgery performed by the same surgeon using a consistent surgical technique, six to nine months after an isolated ACL tear. The normal group consisted of healthy adults who met the criteria for the study, with similar characteristics to the ACLR group. The Noyes Scale assessed the level of sports activity at not less than 70% [20]. The knee size was measured in millimeters. Their characteristics are shown in Table 1.

Table 1: Characteristics of Participants.

Normal cases	Age	WT lb	HT cm	Activities	R Knee mm	L Knee mm
normal 01	26	188.4	169	75%	90	80
normal 02	23	153	167	80%	70	80
normal 03	25	137	168	80%	70	80
normal 04	21	12.8	165	80%	90	75
normal 05	24	174.4	179	80%	90	80
normal 06	21	139.8	167	80%	80	70
normal 07	18	128.8	167	80%	80	70
normal 08	21	144.8	171	80%	70	80
normal 09	38	179.6	177	75%	70	90
normal 10	20	111.4	170	80%	80	70
normal 11	30	132.8	165	80%	90	70
normal 12	24	119	160	80%	80	80
normal 13	22	171.6	178	80%	80	80
normal 14	34	188.2	176	75%	80	95
normal 15	27	224.6	180	75%	80	85
normal 16	27	129.8	167	80%	70	75
normal 17	36	166.8	177	80%	70	80
normal 18	43	171.8	167	75%	80	80
normal 19	21	120	168	80%	80	60
normal 20	30	224.4	173	75%	95	80
ACLR Cases						
ACLR 1	23	170	180	80%	90	90
ACLR 2	21	129.8	171	80%	70	70
ACLR 3	20	144.2	166	80%	70	70
ACLR 4	25	194.6	173	75%	90	90
ACLR 5	25	205.4	174	75%	90	90

How to cite this article: B. Alabkary, G. Alawami, C. Kobus, S. Kamel-ElSayed, H. S. Abdel-Aty Zohdy, M. Zohdy Comparison of Machine Learning Algorithms (K-means) Using Data of Anterior Cruciate Ligament Reconstruction (ACLR) Patients. Eng Technol Open Acc 2024; 6(1): 555676.
DOI: 10.19080/ETOAJ.2024.06.555676

ACLR 6	20	137.2	160	80%	80	80
ACLR 7	25	126.8	159	80%	80	80
ACLR 8	25	163.4	173	80%	70	70
ACLR 9	37	157.2	166	80%	70	70
ACLR 10	26	145.2	164	80%	80	80
ACLR11	27	193.6	169	75%	90	90
ACLR 12	39	157.4	166	75%	80	80
ACLR 13	24	176.8	171	75%	80	80
ACLR 14	23	165.4	160	75%	80	80
ACLR 15	24	167.6	181	80%	80	80
ACLR 16	31	162.4	173	80%	70	70
ACLR 17	26	149.4	168	80%	70	70
ACLR 18	28	138.8	173	80%	80	80
ACLR 19	33	175.8	175	75%	80	80
ACLR 20	23	189.6	177	75%	95	95

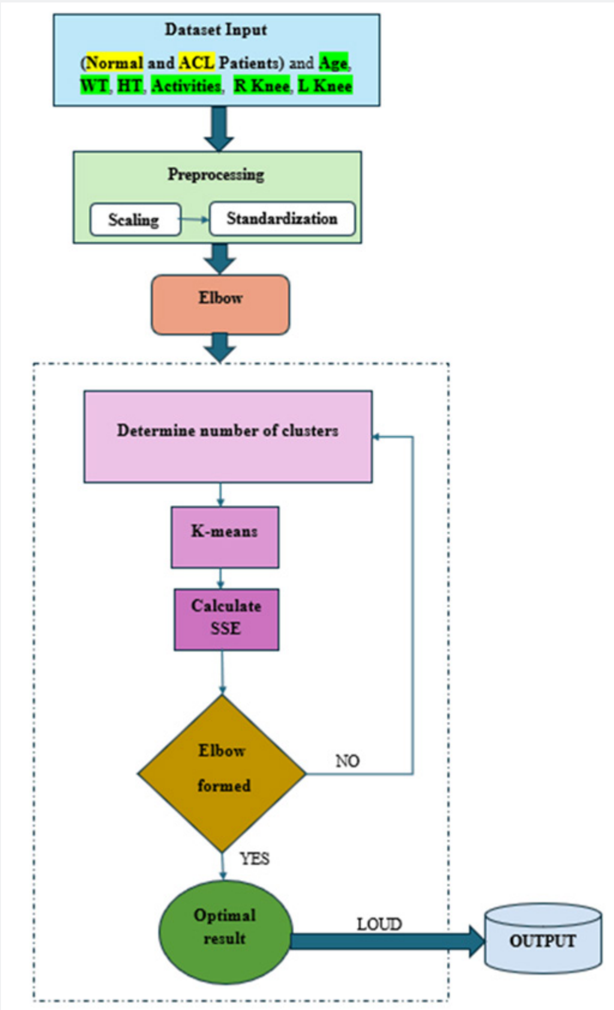

Figure 1: Process flowchart.

How to cite this article: B. Alabkary, G. Alawami, C. Kobus, S. Kamel-ElSayed, H. S. Abdel-Aty Zohdy, M. Zohdy Comparison of Machine Learning Algorithms (K-means) Using Data of Anterior Cruciate Ligament Reconstruction (ACLR) Patients. Eng Technol Open Acc 2024; 6(1): 555676.
DOI: 10.19080/ETOAJ.2024.06.555676

Processing

The process used was based on the MATLAB code the researchers created. The flowchart in Figure 1 shows the processing of a dataset using the Elbow and k-means clustering approaches to identify optimal clustering and provide results. The database for the ACLR group and normal group with their characteristics such as age, weight (WT), height (HT), activity level, and knee condition (right and left) was recorded to analyze and compare the difference between these two groups. The data set was processed in five steps. The first step was to input the datasets by extracting the formation from both normal and ACLR groups, including parameters such as age, weight, height, activities, and knee condition (right and left). The second step comprised data preprocessing through scaling and standardization, which ensured consistency and comparability across all parameters. The third step used the Elbow method to find the best number of clusters. This method involved adding up all the squared errors (SSE) for all the possible cluster numbers to find the point where adding more clusters stopped making a big difference in the SSE. The fourth step involved running the k-means clustering algorithm and using the dataset to determine the range of cluster numbers, as well as calculating the SSE to evaluate the compactness of each cluster. The fifth step was the optimal result, followed by output analysis, visualization, and presentation in an organized and meaningful way.

Results

To interpret the results of the k-means clustering algorithm, heatmaps were selected as one of the multiple ways to visualize the clustering results. The first result was the heatmap for the normal group, as shown in Figure 2. This showed the correlation matrix for normal group data with various measurements. Each cell determined the correlation coefficient between two variables on the axis, which ranges from -1 to 1. A coefficient close to -1 showed a negative correlation between variables, which means one variable increases as the other decreases. A coefficient close to 1 indicated a strong positive correlation, meaning one variable increases as the other increases. If the coefficient is zero, there is no correlation between the variables. Figure 2's heatmap revealed a strong correlation between the measurements of the right and left knees, displaying a dark color in range 1, a result consistent with a normal group's expectations. Also, the correlation coefficient was strongly positive between weight (WT) and height (HT), and between weight (WT) and both knees. Age and both knees, as well as age and weight (WT), had a correlation coefficient of around (0.5475) (0.4866), which is a moderately positive correlation. However, Activities & Weight (WT), Activities & Age, Activities & Height (HT), and Activities & Both Knees had a negative correlation as highlighted by the light blue color, implying the normal group had lower activity levels in this sample.

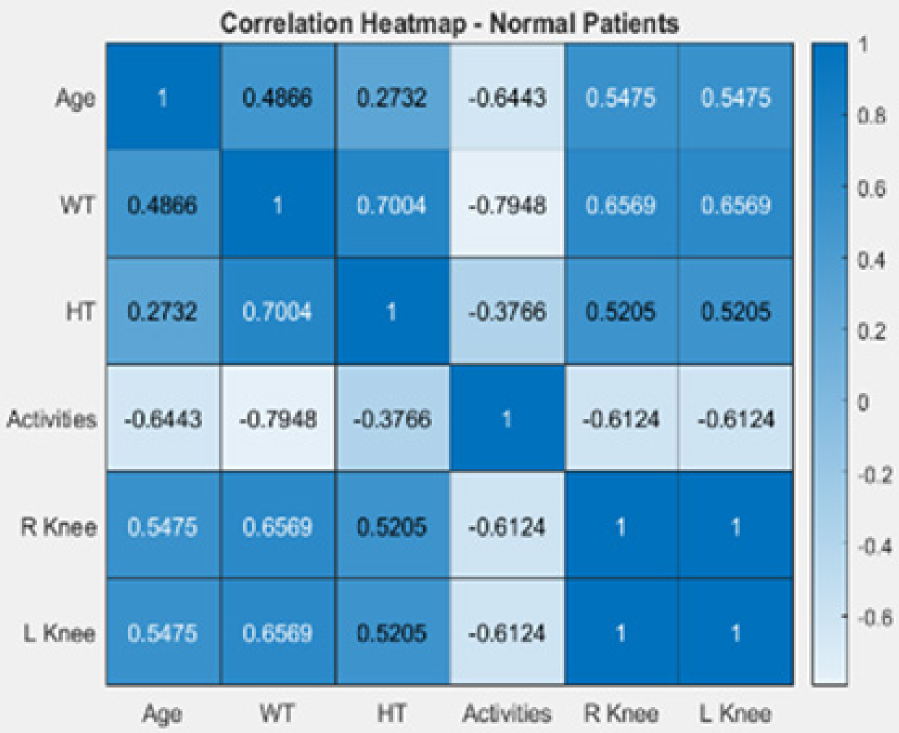

Figure 2: Heatmap for normal group.

How to cite this article: B. Alabkary, G. Alawami, C. Kobus, S. Kamel-ElSayed, H. S. Abdel-Aty Zohdy, M. Zohdy Comparison of Machine Learning Algorithms (K-means) Using Data of Anterior Cruciate Ligament Reconstruction (ACLR) Patients. Eng Technol Open Acc 2024; 6(1): 555676.
DOI: 10.19080/ETOAJ.2024.06.555676

The heatmap for the ACLR group, as shown in Figure 3, was the second result. The heatmap indicated the correlation matrix for ACLR group data, which contained multiple measurements for age, height (HT), weight (WT), activities, and both knees. Like the previous heatmap, each cell determines the correlation coefficient between two variables on the axis, which ranges from -1 to 1. A coefficient close to -1 showed a negative correlation between variables, which means one variable increase as the other decreases. A coefficient close to 1 indicated a strong positive correlation, meaning one variable increase while the other increases as well. If the coefficient is zero, there is no correlation between the variables. The heatmap indicated different patterns for the ACLR group compared to the normal group. There was a weak negative correlation between age and other variables; for example, the correlations between age and both knee measurements had very weak negative coefficients around -0.0466, which indicated there was no significant relationship between these variables in the ACLR group.

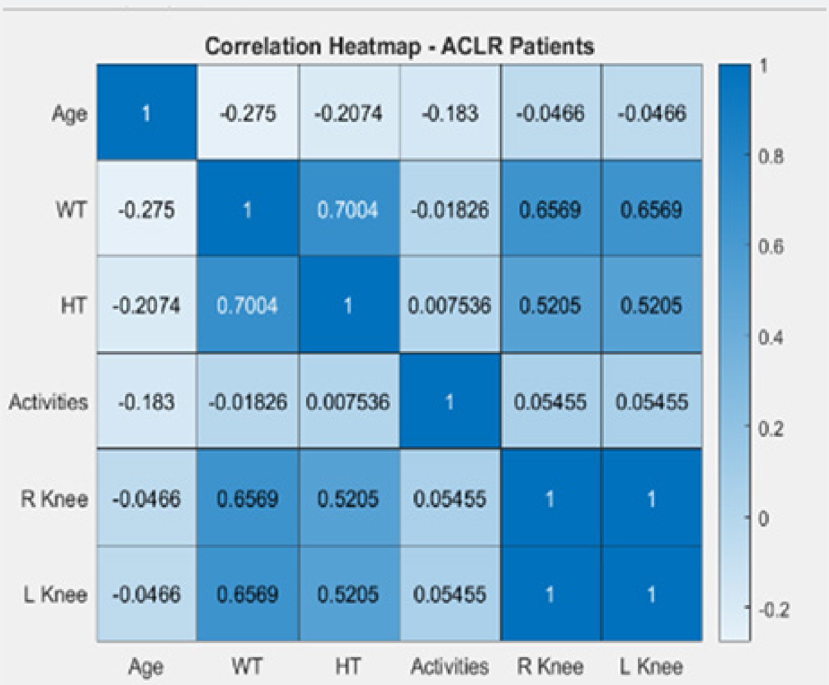

Figure 3: Heatmap for ACLR group.

Furthermore, weight (WT) and activities had a very weak negative correlation with the coefficients at -0.01826, indicating there was almost no relationship between weight and activities in the ACL group, which is a notable difference from the normal group. However, a strong positive correlation was revealed between the right and left knee measurements with a coefficient of 1, the same result as the normal group. Additionally, the correlation coefficient for weight (WT) and height (HT) was 0.7004, indicating a strong positive correlation, consistent with the normal group. The coefficient of activities and height (HT) was 0.007536, and the coefficient of activities and both knees was 0.05455, indicating a weak positive correlation.

The second way the researchers visualized k-means clustering was through scatter plots. To start the k-means procedure, it is important to specify the number of k, which represents the number of clusters in the data. This input is critically important to the quality of clusters, particularly in datasets with more than three variables [21]. In this research, the Elbow Method determined the optimal number of clusters (k) for k-means clustering. Figure 3 indicated the Elbow graph, and showed the within-cluster sum of squares (WCSS) for different numbers of clusters, from 1 to 10. The point at which the curve starts to stabilize (the elbow point) is where the convergence criterion is reached [21]. In this plot on Figure 4, the elbow appeared to be around curve flattened beyond K = 3, the WCSS continued to decrease the elbow metric until it reached zero (19). K = 3. As the curve flattened beyond K = 3, the WCSS continued to decrease the elbow metric until it reached zero [21].

K-means of Age with Activities for Normal Group and ACLR Group

In Figure 5, the scatter plot represented the results of k-means clustering for the normal group based on two variables: age and activities. The ages range was 15 to 45 years, with activity levels ranging from 75% to 80%. The red, green, and blue points indicate the identification of three distinct clusters. Each color represent-

How to cite this article: B. Alabkary, G. Alawami, C. Kobus, S. Kamel-ElSayed, H. S. Abdel-Aty Zohdy, M. Zohdy Comparison of Machine Learning Algorithms (K-means) Using Data of Anterior Cruciate Ligament Reconstruction (ACLR) Patients. Eng Technol Open Acc 2024; 6(1): 555676.
DOI: 10.19080/ETOAJ.2024.06.555676

ed a different cluster for the k-means algorithm. The red cluster contained adults who were younger (aged roughly between 20 and 30 years) and had higher activity levels (around 80%). The green cluster included older adults (approximately 30 to 40 years old); most of them had lower activity levels. The last blue cluster comprised adults who were older than 43 and had lower levels of activity, which might explain why this cluster had very few adults. The large 'X' markers represent the clusters' centroids, which are the average of all points within each cluster. These centroids are the points that minimize the total distance from all points within the cluster. In Figure 6, the scatter plot illustrates the results of a k-means clustering algorithm on data with two variables: age and activities. For the ACLR group, the age range was 20 to 40 years, and the activity levels for the previous group ranged between 75% to 80%. The red, green, and blue colors identified three clusters in the scatter plot. The green cluster included younger patients (20 to 25 years old) with similar activity levels (approximately 80%). The age range of patients in the red cluster was approximately 20 to 30 years; their activity levels ranged from 75% to 80%. The blue points were the third cluster, which represented older patients in terms of age (31 to 37 years) and activity levels (75% to 80%).

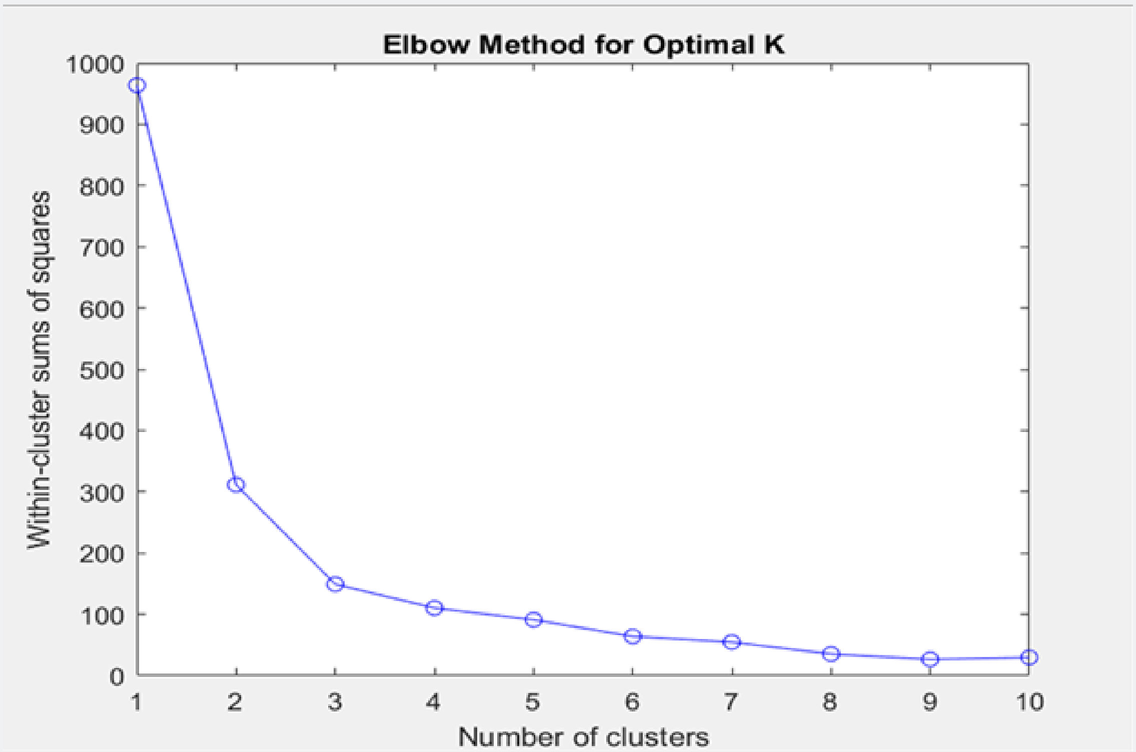

Figure 4: The Elbow Method.

K-means of Activities with Knees Size for Normal Group

Figure 7 is the scatter plot illustrating the results of a k-means clustering algorithm using data with two variables: activities and right knee measurements for the normal group. The red cluster had a narrow vertical spread, indicating an activity level of around 80%, and a small variation in right knee measurements from 60 mm to 75 mm. The green cluster had a broad spread across both distances, indicating a wide difference in both activity levels (75% to 80%) and right knee measurements (80 mm) among these patients. The blue cluster was at the top of the plot, and it showed an average spread along the activity level at 75% and right knee measurements between 85 mm and 95 mm. Figure 8 shows the results of k-means clustering on data with two features: normal group activities and normal group left knee measurements. The plot also shows how the algorithm grouped the data into three clusters, each represented by a different color (red, green, and blue) with large 'X' markers indicating the centroid or means of each cluster. The red cluster groups had high activity levels (around 80%) and a wide range of left knee measurements. The green cluster was also located at 75% activity level, and in the 60 mm to 70 mm range of knee measurements. The blue cluster had a lower level of activity (around 75%) but was spread slightly more in knee measurement than in activity levels.

How to cite this article: B. Alabkary, G. Alawami, C. Kobus, S. Kamel-ElSayed, H. S. Abdel-Aty Zohdy, M. Zohdy Comparison of Machine Learning Algorithms (K-means) Using Data of Anterior Cruciate Ligament Reconstruction (ACLR) Patients. Eng Technol Open Acc 2024; 6(1): 555676.
DOI: 10.19080/ETOAJ.2024.06.555676

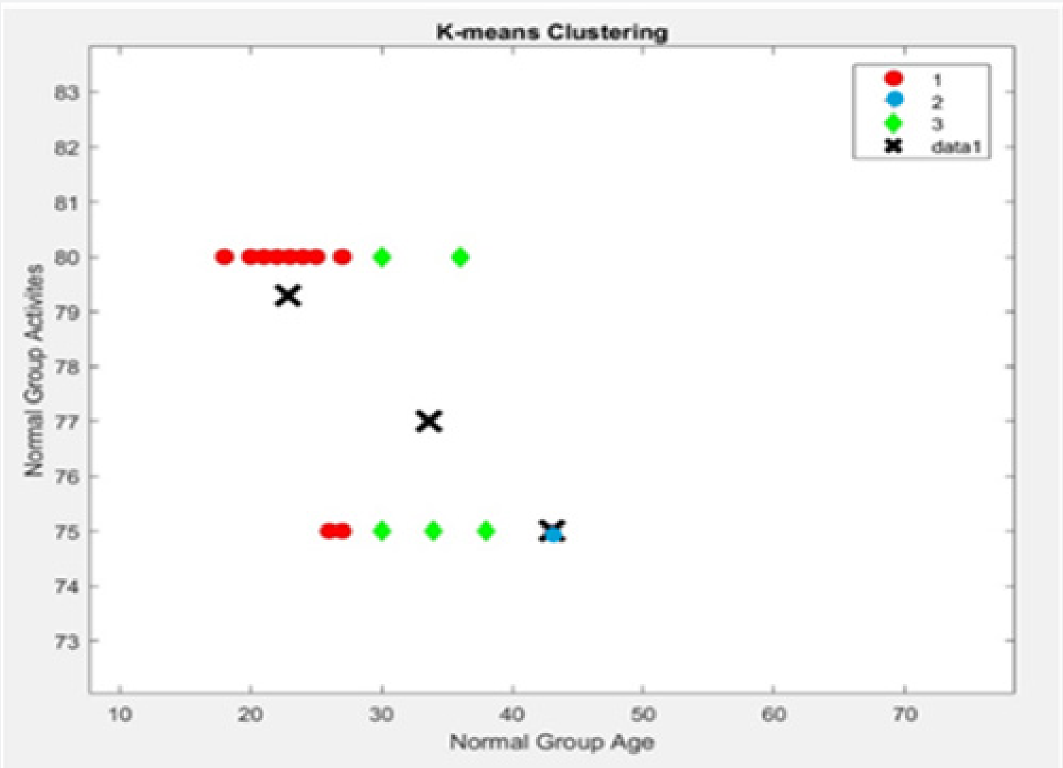

Figure 5: K-means of Age with Activities for Normal Group.

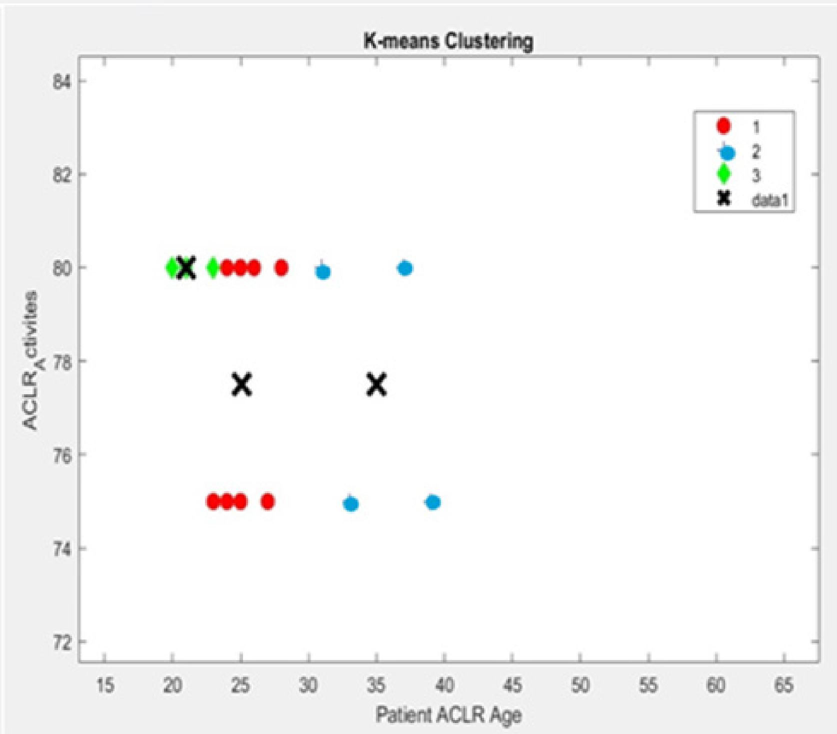

Figure 6: K-means of Age with Activities for ACLR Group.

How to cite this article: B. Alabkary, G. Alawami, C. Kobus, S. Kamel-ElSayed, H. S. Abdel-Aty Zohdy, M. Zohdy Comparison of Machine Learning Algorithms (K-means) Using Data of Anterior Cruciate Ligament Reconstruction (ACLR) Patients. Eng Technol Open Acc 2024; 6(1): 555676.
DOI: 10.19080/ETOAJ.2024.06.555676

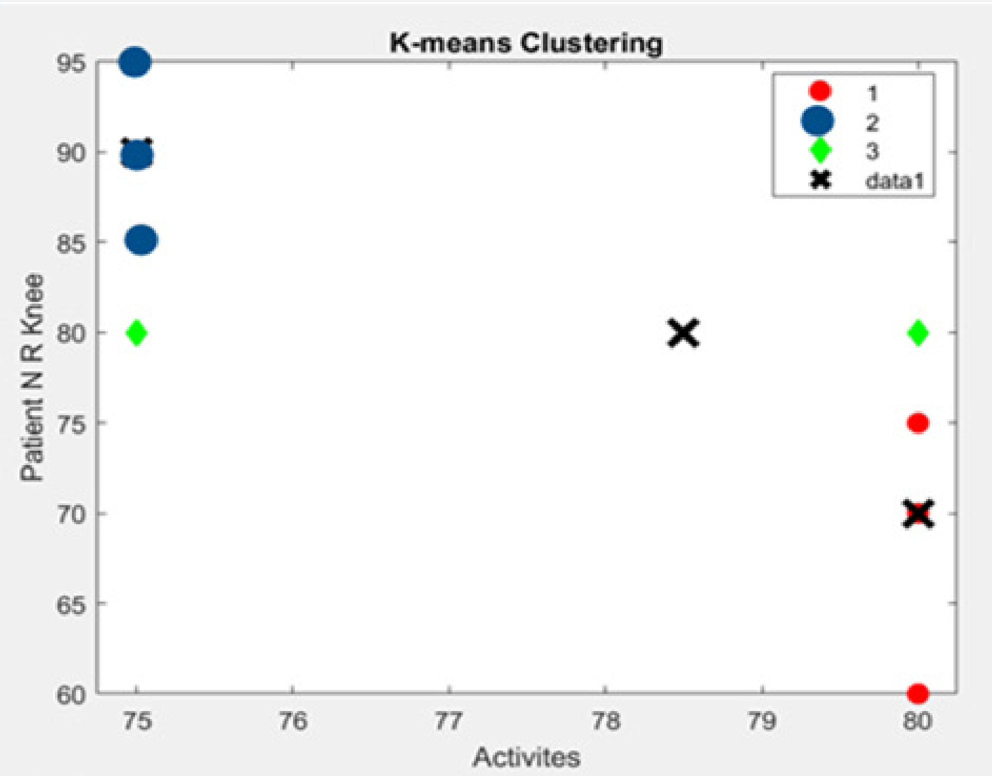

Figure 7: K-means of Activities with R knee N group.

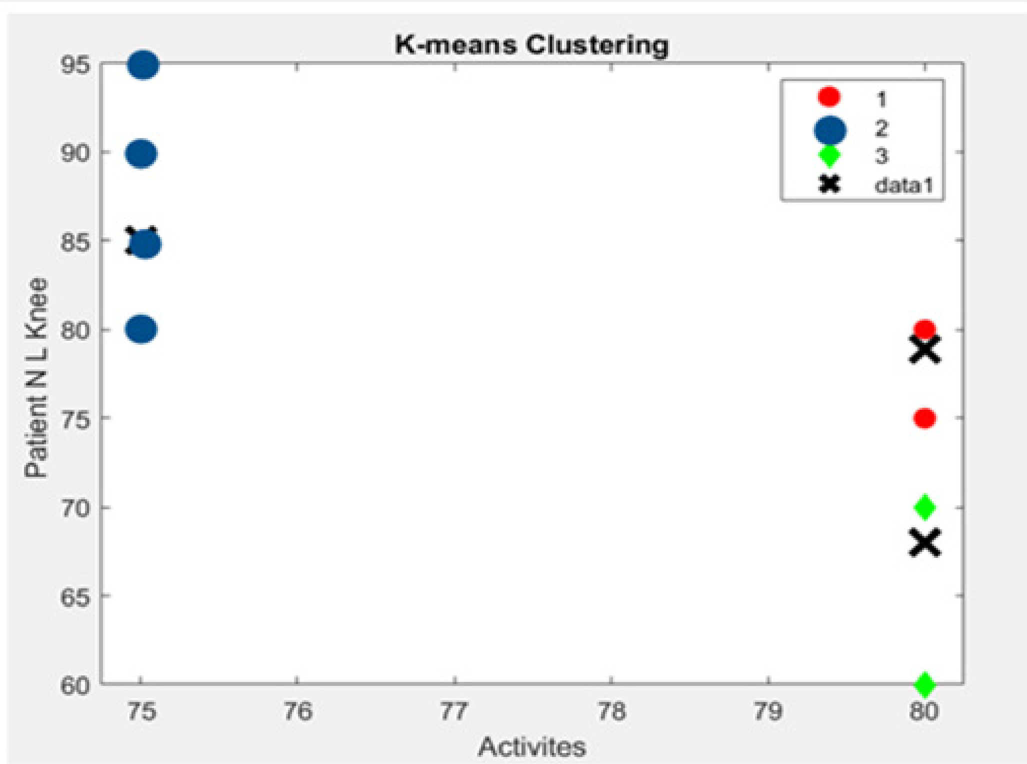

Figure 8: K-means of Activities with L knee N group.

How to cite this article: B. Alabkary, G. Alawami, C. Kobus, S. Kamel-ElSayed, H. S. Abdel-Aty Zohdy, M. Zohdy Comparison of Machine Learning Algorithms (K-means) Using Data of Anterior Cruciate Ligament Reconstruction (ACLR) Patients. Eng Technol Open Acc 2024; 6(1): 555676.
DOI: 10.19080/ETOAJ.2024.06.555676

K-means of Activities with Knees Size for ACLR Group

Figure 9 shows the k-means clustering results for two features: ACLR group activities and right knee measurements. There were three clusters colored red, green, and blue. The blue cluster had a few points, but higher knee measurements and activity levels. The green cluster featured lower activity levels and lower knee measurements. The red cluster exhibited a dispersed pattern, boasting a higher number of points; its activity ranged from 75% to 80%, and its knee measurements fell between 75 mm and 80 mm. The large 'X' markers represent the central point of each cluster. Figure 10 indicates the results of a k-means clustering for two features: ACLR activities and ACLR left knee measurements. The large 'X' marks represent the centroids of each cluster. Also, there were three clusters: red, green, and blue. The red cluster included more points with a range of knee measurements from 70 mm to 80 mm but consistently high activity levels, with most points at 80%. The green cluster was at the lower end of the activity scale, with a lower knee measurement. The blue points stood alone, far from the other data points, with higher knee measurement ranging between 90 mm and 95 mm.

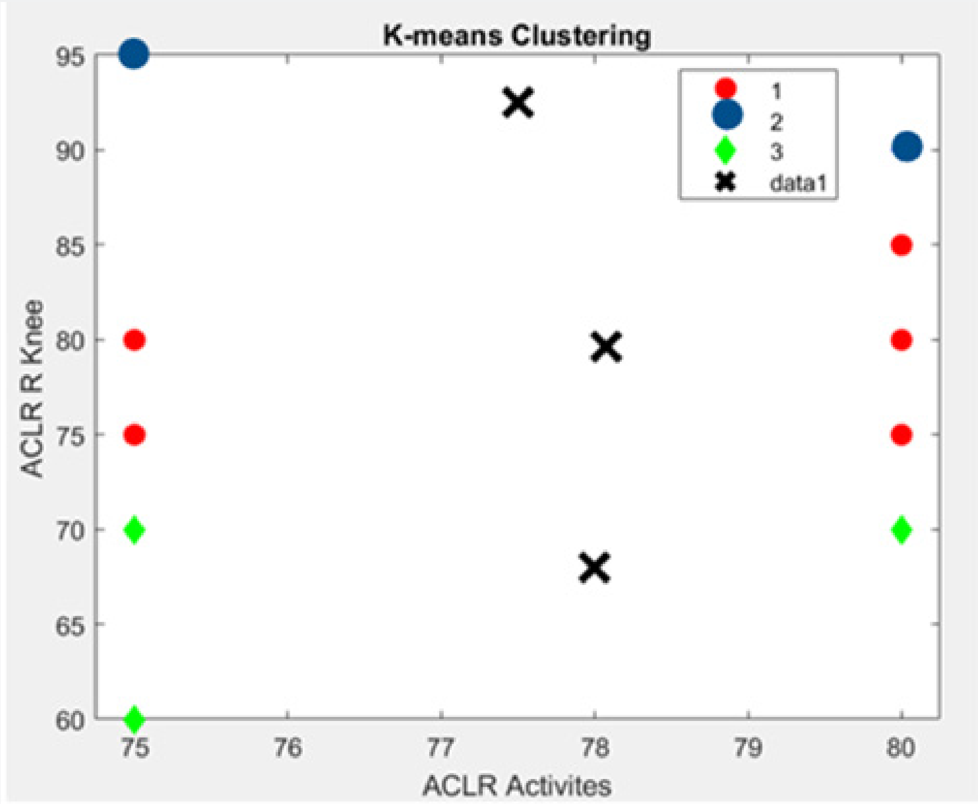

Figure 9: K-means of Activities with R knee ACLR group.

Discussion

This study conducted k-means clustering analysis on both normal and ACLR groups, focusing on comparing variables such as age, activities, and knee measurements. Boyer and Andriacchi argued age plays a significant role in differences in knee function during activities [22]. Other extant research too has stated movement techniques differ among age groups [22-24]. The data from the normal group in this study indicated that younger adults were generally more active and had larger knee measurements. This may reflect higher levels of physical activity or a healthier lifestyle, contributing to stronger knee health among the younger adults in that cohort. Therefore, this result aligns with the findings of Boyer and Andriacchi, who observed that younger adults were more active and walked faster than older adults [22,23]. On the other hand, the ACLR group in this study exhibited a different clustering pattern, influenced by the recovery and rehabilitation process post-surgery. The results showed a weak correlation between age and activity levels in the ACLR group, suggesting that surgery and the recovery path had a greater impact than age alone.

Csapo et al.'s study yielded a different result: the researchers argued that younger individuals exhibited higher post-operative physical activity levels, while older adults struggled more to recover to the average performance level of healthy individuals [25]. However, there has been no study comparing knee size between

How to cite this article: B. Alabkary, G. Alawami, C. Kobus, S. Kamel-ElSayed, H. S. Abdel-Aty Zohdy, M. Zohdy Comparison of Machine Learning Algorithms (K-means) Using Data of Anterior Cruciate Ligament Reconstruction (ACLR) Patients. Eng Technol Open Acc 2024; 6(1): 555676.
DOI: 10.19080/ETOAJ.2024.06.555676

ACLR patients and normal individuals. When comparing our findings for the normal and ACLR groups, the impact of surgery and rehabilitation on activity levels and knee functionality clusters became evident. The ACLR group had smaller knee measurements, while the normal group had larger ones. The normal group displayed a positive correlation for most variables, whereas the ACLR group exhibited more negative correlations between variables. Additionally, the normal group exhibited more consistent clustering, while the ACLR group showed greater variability, likely due to the individualized nature of their recovery process, as some individuals may require more time to recover. These findings are consistent with the results of Dunn WR et al., who asserted that 45% of ACLR patients returned to the same or a higher level of activity two years after surgery [26]. However, Csapo et al. reported that 50% of highly active ACLR patients experienced weakness in strength and functional fitness six months post-surgery [25,27].

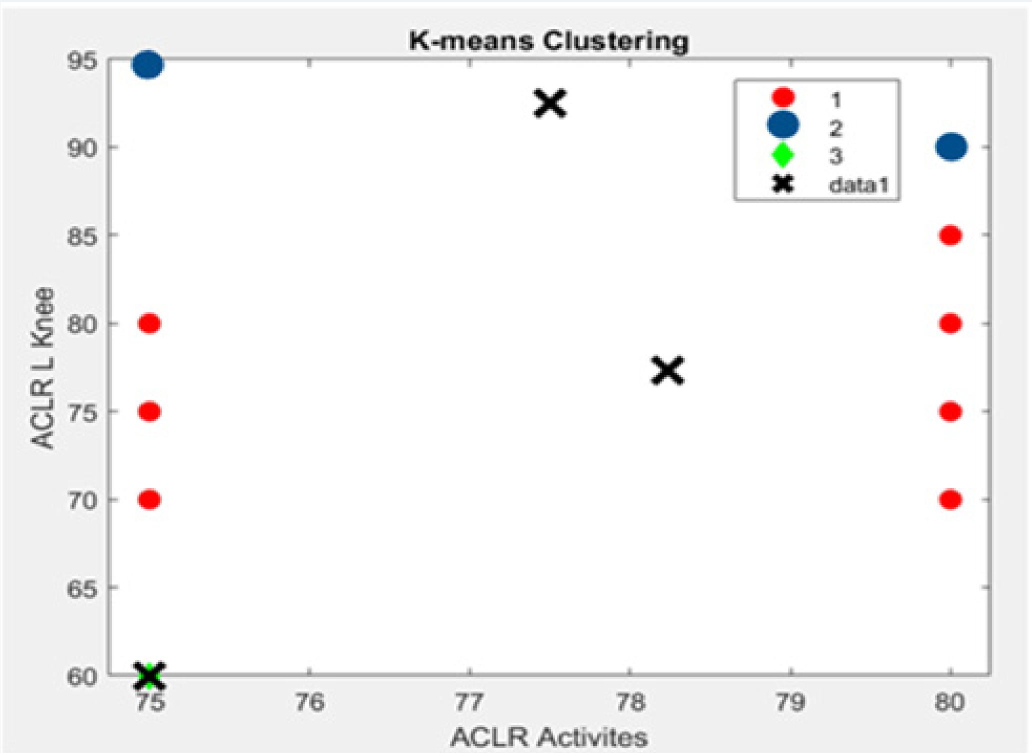

Figure 10: K-means of Activities with L knee ACLR group.

Conclusion

The findings of this paper indicate that utilizing ML algorithms such as the k-means method facilitates the identification of unique clusters and the relationship between several variables. In addition, it can handle multiple databases. However, the study is not without any limitations. In particular, the participant size was small, based on the availability of 40 people from the eastern province of Saudi Arabia. Thus, it is not possible to generalize the results with such a small sample size from the same geographical area. Future research should focus on a large population of ACLR patients and compare this across discrete physical locations.

Yet, this study may be beneficial for clinical practices as it has identified patterns that might not be apparent through standard medical evaluation. For ACLR patients, the results of this study could help in faster identification of patients who may be at risk of slower recovery or complications, which allows for early intervention and support.

References

1. Martin RK, Solvejg W, Ayoosh P, Andreas P, Håvard V, et al. (2022) Predicting Anterior Cruciate Ligament Reconstruction Revision: A Machine Learning Analysis Utilizing the Norwegian Knee Ligament Register. J Bone and Joint Surg 104(2): 145-153.

2. Kunze KN, Polce EM, Ranawat AS, et al. (2021) Application of Machine Learning Algorithms to Predict Clinically Meaningful Improvement After Arthroscopic Anterior Cruciate Ligament Reconstruction. Orthopedic J Sports Med 9(10): 23259671211046575.

3. Lu Y, Forlenza E, Cohn MR, Ophelie Lavoie-G, Ryan RW, et al. (2021) Machine learning can reliably identify patients at risk of overnight hospital admission following anterior cruciate ligament reconstruction. Knee Surg Sports Traumatol Arthrosc 29(9): 2958-2966.

How to cite this article: B. Alabkary, G. Alawami, C. Kobus, S. Kamel-ElSayed, H. S. Abdel-Aty Zohdy, M. Zohdy Comparison of Machine Learning Algorithms (K-means) Using Data of Anterior Cruciate Ligament Reconstruction (ACLR) Patients. Eng Technol Open Acc 2024; 6(1): 555676.
DOI: 10.19080/ETOAJ.2024.06.555676

4. Corban J, Justin-Pierre L, Carl L, Jason K, Gil R, et al. (2021) Artificial Intelligence in the Management of Anterior Cruciate Ligament Injuries. Orthopaedic J Sports Med 9(7): 23259671211014206.
5. Taborri J, Luca M, Adriano S, Mario V, Maria CV, et al. (2021) A Machine-Learning Approach to Measure the Anterior Cruciate Ligament Injury Risk in Female Basketball Players. Sensors 21(9): 3141.
6. Kiapour AM, Sieker JT, Proffen BL, Lam TT, Fleming BC, et al. (2019) Synovial Fluid Proteome Changes In ACL Injury-Induced Posttraumatic Osteoarthritis: Proteomics Analysis Of Porcine Knee Synovial Fluid. PLoS One 14(3): e0212662.
7. Uozumi Y, Nagamune K, Araki D, Hoshino Y, Kuroda R, et al. (2013) An Automatic Segmentation of Bone Tunnels after Anterior Cruciate Ligament Reconstruction in MDCT Image Using K-means Clustering. Procedia Comput Sci 22: 590-598.
8. Andriolo L, Di Matteo B, Kon E, Filardo G, Venieri G, et al. (2015) PRP Augmentation for ACL Reconstruction. Biomed Res Int 2015: 371746.
9. Bauer M, Feeley BT, Wawrzyniak JR, Pinkowsky G, Gallo RA (2014) Factors affecting return to play after anterior cruciate ligament reconstruction: a review of the current literature. Phys Sportsmed 42(4): 71-79.
10. Mantas Lukauskas, Tomas Ruzgas (2022) A New Clustering Method Based on the Inversion Formula. Mathematics 10(15): 1-16.
11. Syakur MA, Khotimah BK, Rochman EMS, Satoto BD (2018) Integration K-Means Clustering Method and Elbow Method for Identification of the Best Customer Profile Cluster. IOP Conf Ser Mater Sci Eng 336(1): 2017.
12. Chen DS, Tong-Fu W, Jia-Wang Z , Bo Z, Zeng-Liang W, et al. (2021) A Novel Application of Unsupervised Machine Learning and Supervised Machine Learning-Derived Radiomics in Anterior Cruciate Ligament Rupture. Risk Manage Healthc Policy 14: 2657-2664.
13. Han XG, Tian W (2019) Artificial Intelligence In Orthopedic Surgery: Current State And Future Perspective. Chin Med J (Engl) 132(21): 2521-2523.
14. Logunova I (2023) K-Means Clustering in Machine Learning. Serokell.
15. Kumar J, Vashistha R (2017) Estimation of Inter-Centroid Distance Quality in Data Clustering Problem Using Hybridized K-Means Algorithm, in 2017 2nd International Conference on Electrical, Computer and Communication Technologies (ICECCT), Coimbatore.
16. Mahesh B (2020) Machine learning algorithms-a review. International Journal of Science and Research (IJSR) 9(1): 381-386.
17. Omar T, Alzahrani AK, Zohdy M (2020) Clustering approach for analyzing the student's efficiency and performance based on data. J Data Analysis Info Processing 8(3): 171-182.
18. What is Matlab? (2023) What is Matlab.
19. Alyami AMI (2024) Knee Function Post Anterior Cruciate Ligament Reconstruction (ACLR), Imam Abdulrahman Bin Faisal University.
20. Noyes FR, Barber-Westin SD, Roberts CS (1989) Use of modified International Knee Documentation Committee forms to assess ligament injury of the knee. J Orthopaedic and Sports Physical Therapy 29(10): 639-644.
21. Duarte TSM (2018) Learning Analytics of K-means Clustering: A Pilot Study.
22. Boyer KA, Andriacchi TP (2016) The Nature of Age-Related Differences in Knee Function during Walking: Implication for the Development of Knee Osteoarthritis. PLoS One 11(12): e0167352.
23. Boyer TP, Andriacchi, Beaupre GS (2012) The role of physical activity in changes in walking mechanics with age. Gait Posture 36(1): 149-153.
24. Favre J, Erhart-Hledik JC, Andriacchi TP (2014) Age-related differences in sagittal-plane knee function at heel-strike of walking are increased in osteoarthritic patients. Osteoarthritis Cartilage 22(3): 464-471.
25. Csapo R, Pointner H, Hoser C, Gföller P, Raschner C, et al. (2020) Physical Fitness after Anterior Cruciate Ligament Reconstruction: Influence of Graft, Age, and Sex. Sports (Basel) 8(3): 30.
26. Dunn WR, Spindler KP (2010) MOON Consortium. Predictors of activity level 2 years after anterior cruciate ligament reconstruction (ACLR): a Multicenter Orthopaedic Outcomes Network (MOON) ACLR cohort study. Am J Sports Med 38(10): 2040-2050.
27. Yoo H, Marappa-Ganeshan (2023) Anatomy, Bony Pelvis and Lower Limb, Knee Anterior Cruciate Ligament, in StatPearls, Treasure Island (FL): StatPearls.

DOI: 10.19080/ETOAJ.2024.06.555676

How to cite this article: B. Alabkary, G. Alawami, C. Kobus, S. Kamel-ElSayed, H. S. Abdel-Aty Zohdy, M. Zohdy Comparison of Machine Learning Algorithms (K-means) Using Data of Anterior Cruciate Ligament Reconstruction (ACLR) Patients. Eng Technol Open Acc 2024; 6(1): 555676.
DOI: 10.19080/ETOAJ.2024.06.555676