Indonesian Journal of Electrical Engineering and Informatics (IJEEI)
Vol. 13, No. 1, March 2025, pp. 226~238
ISSN: 2089-3272, DOI: 10.52549/ijeei.v13i1.6008

Clustering the Addiction Levels of Drug Users Using Fuzzy C-Mean

Okfalisa[1], Pramulia Suliandri[2], Ola A. Alwesabi[3], Bolat Tassuov[4], Raimbayeva Saule[5], Kurdyukov Vitaliy[6], Zahidah Zulkifli[7], Lestari Handayani[8]

[1,2,8]Informatics Engineering Department, Faculty Science and Technology, Universitas Islam Negeri Sultan Syarif Kasim, 28293 Pekanbaru, Riau, Indonesia
[3]Faculty of Computer Studies, Arab Open University, Saudi Arabia
[4,5,6]Physics and Informatics Department, M.H. Dulati Taraz University, Taraz, Kazakhstan
[7]Information Systems Department, Kulliyyah of Information and Communication Technology, International Islamic University Malaysia, 53100 Kuala Lumpur, Malaysia

Article Info

Article history:

Received Oct 30, 2024
Revised Mar 2, 2025
Accepted Mar 21, 2025

Keywords:

Drug Addiction
Clustering
Data Mining
Fuzzy C-Means
Silhouette Coefficient Index

ABSTRACT

Recently, the number of drugs abused, such as Narcotics, Psychotropics, and Addictive Substances have been linear increased with the drug users. The increasing number of these cases triggers the difficulties for rehabilitation associations in diagnosticating the abuse level for medical and health prevention. Herein, data mining with a Fuzzy C-Mean clustering approach is employed to delve 506 drug users' addiction into three clusters by considering several indicators including age, urine test, duration of use, physical effects, and psychological effects. As a result, 215 data are recorded in clusters 1 as high optimum addiction, 105, and 186 data in clusters 2, and 3 as medium and regular addiction levels, respectively. The Silhouette Coefficient, Calinski-Harabasz Index, and Davies-Bouldin Index algorithms reveal high potential values to indicate the proper achievement of this clustering structural test. A clustering software has been successfully developed and tested to aid the calculation and analysis. Hence, the rehabilitation associations in Riau province as end-user of this case are aided in identifying the addiction level of drug users in order to ensure the proper therapeutic prevention and curative action.

Corresponding Author:

Okfalisa,
Informatics Engineering Department,
Faculty Science and Technology,
Universitas Islam Negeri Sultan Syarif Kasim,
28293 Pekanbaru, Riau, Indonesia.
Email: okfalisa@gmail.com

1. INTRODUCTION

The drug addiction develops with the continuous drug use which is particularly consequential in the first year of intake [1]. It is estimated that 255 million people used illegal drugs in 2015, which indicated to an annual prevalence of 5.3% illicit drug use [2]. It increased the number into 6 million in 2017 and about 37.8% of drug users reported augmented consumption due to various psychosocial pressures, while overall substance use behaviors fluctuated during the pandemic in 2020 in Indonesia. In 2008, the number of narcotics abusers had reached 1.5% of the Indonesia total population or ranged from 3.1 million to 3.6 million people. From total abusers, it found that 26% of them are a tried to use, 27% derived from the regularly used, 40% arose from non-injecting addicts, and 7% originate as injecting addicts. It is sophisticating, the drug abuse dominantly retrieved from non-student group up 40%. Meanwhile, the male gender user is 88% much greater than female. It also reported that the estimated loss in economic costs due to the narcotics crime in 2008 was around 32.4

trillion rupiahs which consisting of 26.5 trillion rupiah for individual losses and 5.9 trillion rupiah for social values [3].

Drugs (Narcotics, Psychotropics and Addictive Substances) retrieve natural or synthetic materials which trigger the physical changes function as well as addiction and the psychological negative consequences, including loss of consciousness and behavior thus lead into dependence; the brain's reward system which releases dopamine as drug used response that leading to a feeling of euphoria, long-term changes in brain function, affecting judgment, decision-making, memory, and learning abilities; physical health consequences viz. heart disease, high blood pressure, psychosis, reduced immune function, stomach issues, respiratory problems, liver damage, and kidney disease; withdrawal symptoms when the user need to stop and reduce the addiction; mental health issues such as emotional strains and cognitive impairment. Moreover, the social well-being affection lead into the health problems, financial difficulties, social isolation, legal issues, and strained relationships [4]. The treatment for this drug addiction may involve a combination of medical interventions, therapy, and support groups. Recovery is possible, nevertheless the requirement of such a long-term commitment and drug-free lifestyle maintenance are compulsory. In summary, the effects of drug addiction, including Narcotics, Psychotropics, and Addictive Substances, are profound and far-reaching. Early identification and comprehensive treatment are essential for managing these effects effectively. By addressing the addiction early can prevent long-term consequences and provide comprehensive care that includes therapy, medication, and support groups to aid in recovery.

Identifying the drug addiction in data mining from computer science perceives involve grouping individuals cluster based on their characteristics or behaviors related to drug addiction. This approach helps in identifying patterns and trends within the data, which can be useful for various purposes such as understanding the demographics and behavioral data of drug users, identifying high-risk groups, revealing distinct profiles of drug users, and developing targeted interventions. For instance, one study highlighted the effectiveness of k-means clustering in identifying subgroups of individuals based on drug distribution pattern using CRISP-DM to improve drug sampling planning decision making[5]. [6] deploys deep neural network-based clustering-oriented embedding algorithm to identify drug consumption patterns, including addiction, using data mining techniques. [7] investigated the patients latent clusters with addiction and misuse of opioid and undergone Covid-19 screening using FAMD preprocessing method and K-mean clustering. They revealed the distinct subgroups based on testing results and demographics. Herein, clustering algorithms equip as powerful tool aid in grouping data points according to rules into clusters based on similarities and differences, find regular sets of unlabelled data, encourage the arrangement of significant sub-parties or sub-classes, and furnish with broadly delegated, progressive and partition calculation [8] Moreover, the exploration of contextual factors influencing drug use become the potential showcasing of big data in addiction research. However, the limitations are identified on data collection as well as data privacy concerns and the need for high-quality datasets remain challenges in this field. Subsequently, the data mining is an invaluable tool for understanding drug addiction. It allows us to cluster drugs and develop personalized treatment strategies and public health initiatives.

In order to facilitate the data processing, many algorithms and data mining techniques are utilized, including Fuzzy C-Mean, K-Mean, DBSCAN, Gaussian Mixture Models (GMM), Mean Shift, and Hierarchical Clustering. These above techniques have commonly used for handling many cases and offers several advantages and disadvantages. Fuzzy C-Mean provides as on one of the frequently used algorithms in clustering any types of large datasets, including text, numbers, or image processing within numerous field areas. [9] investigated that Fuzzy C-Means considers with the overlapping classes and shapes, improving data interpretation, unlike others clustering algorithm such as K-Means. Fuzzy C-Mean provides the weighted calculation techniques that reduces the ambiguity in membership grouping. Moreover, Fuzzy C-Mean allows data points to belong to multiple clusters with varying degrees of membership, suitable for vague data. This condition is in direct conflict with the K-mean, which is only assigns each point to a single cluster [10]. Thus, these above situation impact into the enhancement of clustering quality. Besides, Fuzzy C-Mean grows into outperformed traditional models that is more flexible and stabile for gene expression data but computationally intensive [11]. Fuzzy C-Mean contributes a soft clustering techniques that better handling the uncertainty data and develop it advantageous in scenarios where data points are not distinctly separable. Subsequently, the intensive computationally Fuzzy C-Mean can converge to local minima, which may affect its performance on complex datasets [12]. Moreover, K-Means can lead to suboptimal results based on its initial centroid selection, while Fuzzy C-Mean's membership degrees can provide more stability in chaotic datasets. Comparing to DBSCAN, Fuzzy C-Mean equally treats the entire data points, accommodating clusters of various shapes and densities without the need for density parameters that can complicate DBSCAN's use. Fuzzy C-Means effectively overcomes several limitations of density-based clustering methods like DBSCAN, especially in assigning the partial memberships to handle outliers more effectively, making it suitable for the diverse treatment and intervention needs found in drug addiction research [13].When GMM assumes that data points

are generated from a mixture of several Gaussian distributions and may struggle to find optimal solutions with non-Gaussian data distributions common in addiction profiles, Fuzzy C-Mean does not require underlying distribution assumptions, making it more flexible in dealing with the diverse patterns of addiction-related behaviors. Fuzzy C-Mean is simpler in terms of implementation, capturing the data without needing intricate assumptions regarding data distributions instead of GMMs can become computationally complex and require tuning several parameters (e.g., number of Gaussian distributions), potentially leading to challenges in practical implementation for real-world addiction data [14]. Mean Shift is a density-based clustering technique that clusters data points by identifying modes in the density of data points in a feature space. It shifts data points towards areas of higher density iteratively. Meanwhile, Fuzzy C-Mean uses a partitioning approach and enables each data point to belong to multiple clusters with varying degrees of membership thus it relies on the assumption that clusters are convex and isotropic, which can limit its application in datasets where clusters have irregular shapes whereby Mean Shift can identify clusters of arbitrary shapes, making it flexible in capturing complex cluster geometries. Unfortunately, Fuzzy C-Mean is computationally efficient and can handle larger datasets more effectively than Mean Shift, which may become computationally intensive, especially with a high number of iterations to find density peaks [15]. Regarding on the Hierarchical methods, it can be substantially sensitive to noise, potentially skewing results, particularly as they depend on metric choices for linkage. In scenarios where drug addiction symptoms exhibit high variability, Fuzzy C-Mean's robustness to noise could provide more reliable results [16]. By addressing the clustering gaps and weaknesses present in the above clustering techniques, Fuzzy C-Mean stands out as a suitable method for accurately analyzing the multifaceted nature of drug addiction, ultimately leading to better-targeted intervention strategies. Thus, the choice of Fuzzy C-Mean considers the specific characteristics of the data and the desired outcomes within the drug addiction case. Fuzzy C-Mean can determine the number of clusters earlier and adjust the addiction level according to the cluster. By detecting the high-level clusters, the relationship between different cluster patterns will be described for performing the further analysis. In this case, Fuzzy C-Mean excels in situations where data naturally falls on a continuum or when variables are imprecise and overlapping—both of which are characteristic of addiction amongst the clusters. Meanwhile, Hard clustering approaches like k-means can obscure these subtleties, making fuzzy clustering the more appropriate choice for classifying the addiction levels. Herein, the clustering considers the drug addiction parameters from the profile user including the prolonged use of the drug, the urine test, and the assessment of physical and psychological effect of addiction. Then, the clustering process groups the pattern addiction data into regular, medium and high optimum class. Fuzzy C-Means clustering is particularly suitable for drug addiction applications due to its ability to model overlapping clusters and address emotional and psychological factors in addiction. However, its requirement for an initial number of clusters presents challenges, especially in dynamic settings where addiction behaviors might change over time. This aspect can complicate the use of Fuzzy C-Mean in practical scenarios, where adaptability is crucial. Consequently, it becomes a limitation of this paper.

2. RESEARCH METHOD

2.1. Fuzzy C-Mean Clustering

Clustering grows into as a sophisticated tool for solving the complex issues that related to data analytical in effectively finding certain patterns of data interest within the knowledge discovery process. Generally, to optimize the clustering solution, the data is initially processed into specific techniques such as Knowledge Discovery Database (KDD). KDD defines as non trivial activities in identifying the valid, novel, potentially values, interesting and understandable patterns, and iterative nature data extraction from large datasets to advance in decision making and predictive modeling [17]. Numerous techniques applied in interpreting data mining, including KDD, CRISP-DM (Cross-Industry Standard Process for Data Mining), SEMMA (Sample, Explore, Modify, Model, Assess), and Agile Data Mining [18]; [19]; [20]; [21]. These above techniques work structurally depends on the nature of the project, the business goals, environment issues, the organizational requirements and the available data to ensure the successful of generating data mining outcomes. Herein, KDD is deployed due to the comprehensive and holistic framework of KDD in leveraging the entire aspects of the knowledge discovery process from data selection to knowledge representation [22]. KDD provides a clear structure in understanding the complexities of data mining projects, especially for large scale data projects. Moreover, the iterative nature approach of KDD allows the continues refinement of data and model based thus it enhancing the quality insights gained from data interpretation. Besides, the feedback loop of KDD techniques offers the learning and adjustment throughout the process thus can enhance the data effectiveness [23]. KDD follows the specific tasks in datamining including data selection task is conducted by filtering the data set and create the target analysis, pre-processing activity is carried out by removing noise or data outliers, the transformation function by converting data into a suitable format for analysis, including normalization and aggregation, the data mining activity through the application of specific data mining

algorithms to extract patterns or models from the prepared data, interpreting task by deriving actionable insights and validating the significance of mined results [24]. Fuzzy C-Mean (FCM) as one of clustering algorithm in data mining offers the multiple clusters handling by considering the membership of varying degrees of data. Despite K-Mean, the examination of each data point pertains exclusively to a single cluster thus referred to as hard clustering [25]. Fuzzy C-Mean provides more accurate representations of overlap cluster with complex dataset and not well defined of clusters boundaries. The flexibility of this algorithm in assigning membership values makes it the ideal solution for uncertain and ambiguous data modeling in various domain [26]. The fuzziness of this algorithm serves the control adjustment of clustering grouping that allows easier understanding of dataset characteristics and desired outcomes. FCM has been commonly applied in solving the healthcare issues with numerous data types involves medical diagnosis, heart disease classification, image analysis in medical imaging, disease pattern recognition, chemical analysis, and patient segmentation. Herein, FCM identify the patients' profiles in order to cluster the level addiction of them based on considerable variables include the onset age of drug consumption (age start to use the drugs), the prolonged use of the drug (the duration of drug use), the urine test (kinds of drugs detection in urine), and the assessment of physical (the physical examination report) and psychological effect of addiction (the psychiatrists report). Additionally, other attributes pertinent to the patient profile include code name, gender, age, the latest education, and the occupation. The data was collected from National Narcotics Boards in Riau Province with 506 dataset patients. The interviews from the specialist medical doctors in National Narcotics Boards in Riau Province determined the weighting values of each parameter as follows.

Table 1. Parameters weighting set (w1-w9)

The onset age of drug consumption (X1)		The urine test (X2)		Prolonged use of the drug (X3)	
Age (Years)	Weight(w)	Numbers of Drug Types	Weight(w)	Duration (Years)	Weight(w)
21 – 30	4	>3	4	>10	5
11-20	3	3	3	7 – 10	4
31-40	2	2	2	3 – 6	3
41-50	1	1	1	1 – 2	2
				< 1	1

The assessment of physical of addiction (X4)		The psychological effect of addiction (X5)	
Measures	Weight(w)	Measures	Weight(w)
Very high	8-9	Very high	8-9
High	6-7	High	6-7
Moderate	4-5	Moderate	4-5
Low	2-3	Low	2-3
Very low	0-1	Very low	0-1

The weighting of onset age of drug consumption indicates that the younger the onset age, the higher the risk of developing severe drug addiction, primarily due to neurodevelopmental vulnerability, social/behavioral factors, and cumulative exposure. Incorporating onset age as a key variable in clustering can enhance the accuracy of addiction-level classification and guide more targeted intervention strategies [27]. Urine test parameters is defined based on the indication that detecting more kinds of drugs in a urine test typically correlates with higher addiction risk and severity. This is supported by research indicating that polysubstance use leads to more complex medical, psychological, and social challenges. Consequently, individuals testing positive for multiple substances often require intensive, multifaceted treatment strategies to effectively address their addiction [28]. The weighting of parameter prolonged drug use generally reflects higher addiction risk due to cumulative physiological damage, escalating tolerance, and deepening behavioral patterns. Incorporating "duration of use" as a key variable in clustering or classification models helps identify individuals at greater risk and informs more targeted, long-term intervention strategies [29]. For parameter the assessment of physical of addiction weighting, it found that a higher physical assessment score—indicating greater physiological damage or more pronounced physical symptoms—often correlates with increased addiction severity. Recent research and clinical guidelines support the link between physical health deterioration and escalated substance dependence, underscoring the importance of integrated medical and psychological care for individuals exhibiting these signs [30]. The weighting values for parameter the psychological effect of addiction is defined by the stronger psychological effects found (e.g., severe anxiety, depression, cognitive deficits) are closely linked to greater addiction risk. Research consistently shows that co-occurring mental health issues, emotional dysregulation, and cognitive impairments amplify the severity of substance dependence and complicate treatment. Identifying and addressing these psychological factors is crucial for effective intervention and relapse prevention [31].

The tracing calculation of FCM algorithm follows the formula below [32].

1. Input the data, set into cluster *X* and performs the matrix *n* * *m* whereby *n* as the number of data samples, and *m* as the attributes of each data. Then, defines the number of clusters *(c)*, rank *(w)*, maximum iteration (*maxiter*), error rate (ξ), initial objective functions ($P_o = 0$), and initial iteration (*t=1*)
2. Calculating the attribute for each column using formula (1).
$$Q_i = \sum_{k=1}^{c} \mu_{ik} \quad (1)$$
3. Calculating the center of each cluster (*v*) with formula (2)
$$V_{kj} = \frac{\sum_{i=1}^{n}((\mu_{ik})^{w} * x_{ij})}{\sum_{i}^{n}(\mu_{ik})^{w}} \quad (2)$$
4. Calculating and mixing the matrix partition with formula (3)
$$\mu_{ik} = \frac{\left[\sum_{j=1}^{m}(x_{ik} - v_{kj})^2\right]^{\frac{-1}{w-1}}}{\sum_{j=1}^{m}\left[\sum_{j=1}^{m}(x_{ik} - v_{kj})^2\right]^{\frac{-1}{w-1}}} \quad (3)$$
5. Calculating the objective function at the iteration -*t* for *Pt* as defined in formula (4)
$$P_t = \sum_{i=1}^{n}\sum_{k=1}^{c}\left(\left[\sum_{j=1}^{m}(x_{ik} - v_{kj})^2\right](\mu_{ik})^{w}\right) \quad (4)$$
6. Checking the calculation criteria, if $(|P^t - P^{t-1}| \xi)$ or (*t* > *maxIter*) then stop if undefined then calculate *t* = *t+1*, repeat formula (2).

2.2. Evaluation Matrics

Silhouette coefficient (SC) is one of the evaluation matrices for measuring the cluster separation and cohesiveness using the average distance between data points in the similar cluster compared to other clusters [33]. Thus, this technique provides insights the optimal clustering configuration and overview dataset characteristics, calculation and significance. Silhouette coefficient aid in validating the clustering result effectively and plots visually represented data thus it easier to interpret the clustering quality [34]. The Silhouette width *s(xi)* for the point *xi* is defined as Equation (5). where *xi* is an element in cluster *πk*, *a(xi)* is the average distance of *xi* to all other elements in the cluster *πk* (within dissimilarity), and *b(xi) = min {dl(xi)}*, among all clusters *l ≠k.* where *dl(xi)* is the average distance from *xi* to all points in cluster *πl* for *l ≠ k* (between dissimilarity). From Equation (5) the value of the Silhouette width can vary between −1 and 1[29].

$$S(xi) = \left(\frac{b(xi) - a(xi)}{max\{b(xi), a(xi)\}}\right) \quad (5)$$

Calinski-Harabasz Index (CHI) provides the evaluation matric to assess the quality of clustering results by calculating the ratio between-cluster dispersion to within-cluster dispersion in order to identify the well-separated and compact of the clusters [34]. A compact and well-separated cluster configuration is expected to have high inter-cluster variance and relatively low intra-cluster variance [35]. CHI is sensitive to the number of clusters and capable for comparing the diversity of clustering solutions and determining the optimal numbers of cluster. CHI is valuable tool for examining the unsupervised learning context thus it widely adopted matric in data mining and machine learning application. The CHI values can be calculated by the following formula whereby *K* as the appropriate number of clusters, *B(K)* as inter-cluster divergence, also called inter-cluster covariance, and *W(K*) as intra-cluster divergence.

$$CH(K) = \frac{B(K)(N-K)}{W(K)(K-1)} \quad (6)$$

Davies-Bouldin Index (DBI) evaluates the efficacy of clustering by measuring the average similarity between each cluster and the most similar cluster [34]. Herein, intra-cluster similarity and inter cluster dissimilarity are measured to provide the information regarding on the well-separated and compact of the clusters. A lower DBI value indicate the better clustering performance thus the better-defined and more meaningful of clusters within dataset [36]. The DBI requires no labelled data and provides a balanced view of

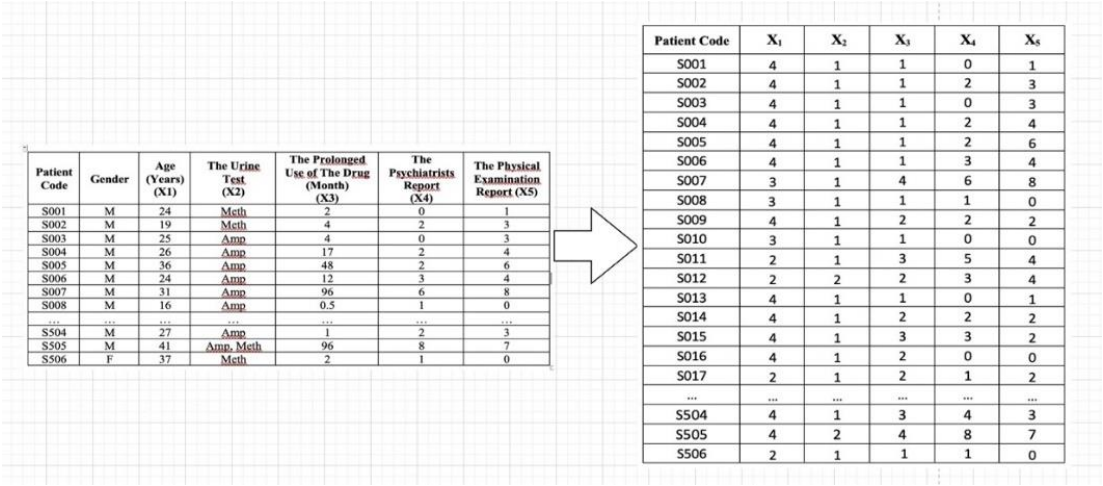
clustering quality thus rendering it appropriate for unsupervised learning contexts. The calculation of DBI can be depicted in Equation (7).

$$DBI = \frac{1}{K}\sum_{i}^{k} max_{i \neq j} R(i,j) \quad (7)$$

3. RESULTS AND DISCUSSION

3.1. KDD Analysis

Following the KDD stages that involves data selection, pre-processing, and data transformation, the result of dataset analysis can be depicted at Figure 1. The raw data which previously had 9 fixed attributes after selection became 5 main attributes (X1-X5). Parameter transformation based on weights was carried out to facilitate calculation and pre-processing of missing values, duplicates were not found, so the clean data of 506 data patients were ready for analysis. The percentage of dataset based on parameters after KDD analysis can be explained in Table 2.



Patient Code	Gender	Age (Years) (X1)	The Urine Test (X2)	The Prolonged Use of The Drug (Month) (X3)	The Psychiatrists Report (X4)	The Physical Examination Report (X5)
S001	M	24	Meth	2	0	1
S002	M	19	Meth	4	2	3
S003	M	25	Amp	4	0	3
S004	M	26	Amp	17	2	4
S005	M	36	Amp	48	2	6
S006	M	24	Amp	12	3	4
S007	M	31	Amp	96	6	8
S008	M	16	Amp	0.5	1	0
...	...	...	...	...	...	...
S504	M	27	Amp	1	2	3
S505	M	41	Amp, Meth	96	8	7
S506	F	37	Meth	2	1	0

Patient Code	X_1	X_2	X_3	X_4	X_5
S001	4	1	1	0	1
S002	4	1	1	2	3
S003	4	1	1	0	3
S004	4	1	1	2	4
S005	4	1	1	2	6
S006	4	1	1	3	4
S007	3	1	4	6	8
S008	3	1	1	1	0
S009	4	1	2	2	2
S010	3	1	1	0	0
S011	2	1	3	5	4
S012	2	2	2	3	4
S013	4	1	1	0	1
S014	4	1	2	2	2
S015	4	1	3	3	2
S016	4	1	2	0	0
S017	2	1	2	1	2
...	...	...	...	...	...
S504	4	1	3	4	3
S505	4	2	4	8	7
S506	2	1	1	1	0

Figure 1. KDD dataset analysis for patients profiles

Table 2. The distribution of KDD analysis dataset based on parameters

X1	%	X2	%	X3	%	X4	%	X5	%
21-30	59,68%	>3	0,59%	>10	6,32%	Very high	7,31%	Very high	8,30%
11-20	9,68%	3	6,13%	7 – 10	8,50%	High	14,23%	High	15,42%
31-40	28,26%	2	52,17%	3 – 6	32,81%	Moderate	24,90%	Moderate	26,88%
41-50	2,37%	1	41,11%	1 – 2	35,38%	Low	33,99%	Low	29,25%
				< 1	7,00%	Very low	19,57%	Very low	20,16%

3.2. FCM Analysis

Initial dataset for FCM analysis defines as *c=3*, *w=2*, *maxiter=50*, error rate *(ξ)=10^{-5}*, initial objective functions ($P_o = 0$), and initial iteration *(t=1)*. Thus, the analysis reveals the center calculation of each cluster as in Table 3. Table 3 explains the result calculation of center cluster for cluster 1 as following Equation (2). It then adheres by cluster 2 and 3. As the recapitulation of the center cluster calculation for iteration#1, Table 4 is elucidated. Following the Equation (4), the result calculation of the objective function at the iteration -1 is explained in Table 5. The process will persist till the 50th iteration is attained. Finally, the FCM clusters the dataset successfully as distributed in Figure 2 and Figure 3. Figure 2 records 215 data patients in cluster 1 (green colour), 105 data in cluster 2 (yellow colour), and 186 data patient in cluster 3 (purple colour). The distribution of FCM clusters is visualized through the application of Principle Component Analysis (PCA). PCA offers the significant contribution in visualizing the FCM data distribution including the dimensionality reduction, enhanced interpretability, noise reduction, and improve identification of cluster separation and outliers [37]. Figure 3 describes the percentage (%) distribution of FCM clustering based on the parameters (X1, X2, …, X5) within cluster C1, C2, and C3 by considering the weighting of each parameters (w1,w2,..,w9).

Table 3. The center cluster calculation for cluster 1

Center cluster calculation											
The Cluster Degree -1	Parameters					$(\mu_{i1})^2$	$(\mu_{i1})^2 * X_1$	$(\mu_{i1})^2 * X_2$	$(\mu_{i1})^2 * X_3$	$(\mu_{i1})^2 * X_4$	$(\mu_{i1})^2 * X_5$
	X_1	X_2	X_3	X_4	X_5						
0,6066	4	1	1	0	1	0,367940	1,471762	0,367940	0,367940	0,000000	0,367940
0,2646	4	1	1	2	3	0,070003	0,280013	0,070003	0,070003	0,140006	0,210009
0,4563	4	1	1	0	3	0,208210	0,832839	0,208210	0,208210	0,000000	0,624629
0,1382	4	1	1	2	4	0,019094	0,076378	0,019094	0,019094	0,038189	0,076378
0,4502	4	1	1	2	6	0,202711	0,810845	0,202711	0,202711	0,405422	1,216267
0,4561	4	1	1	3	4	0,208063	0,832254	0,208063	0,208063	0,624190	0,832254
0,5448	3	1	4	6	8	0,296792	0,890376	0,296792	1,187167	1,780751	2,374335
0,3956	3	1	1	1	0	0,156489	0,469466	0,156489	0,156489	0,156489	0,000000
0,0778	4	1	2	2	2	0,006052	0,024207	0,006052	0,012103	0,012103	0,012103
0,7461	3	1	1	0	0	0,556613	1,669840	0,556613	0,556613	0,000000	0,000000
…..	…	…	…	…	…	…..	…..	…..	…..	…..	…..
0,3784	4	1	1	0	1	0,143174	0,572695	0,143174	0,143174	0,000000	0,143174
0,1118	4	1	2	2	2	0,012506	0,050023	0,012506	0,025011	0,025011	0,025011
0,5251	4	1	3	3	2	0,275711	1,102846	0,275711	0,827134	0,827134	0,551423
		∑				**83,635597**	**275,360137**	**139,488116**	**207,045696**	**296,765569**	**315,934593**
						The center cluster	**3,292380**	**1,667808**	**2,475569**	**3,548317**	**3,777513**

Table 4. The recapitulation of center cluster iteration#1

The Centre cluster					
	X_1	X_2	X_3	X_4	X_5
Cluster 1	3,422136	1,028467	1,708091	1,910725	2,710647
Cluster 2	3,613499	1,137565	1,572049	2,060831	2,970372
Cluster 3	3,647378	1,019377	1,738682	2,022508	2,647707

Table 5. The objective function at the iteration #1

The objective function							
	$(\mu_{i1})^2$-i			L_4	L_5	L_6	L_7
No	$(\mu_{i1})^2$	$(\mu_{i2})^2$	$(\mu_{i3})^2$				
1	0,350735	0,301464	0,347801	2,727659	0,108277	0,591875	3,427811
2	0,070003	0,145909	0,124919	0,064951	0,072973	0,099308	0,237232
3	0,20821	0,145909	0,026153	0,95167	0,692113	0,127759	1,771541
4	0,019094	0,344374	0,075616	0,047861	0,537011	0,189007	0,773879
5	0,202711	0,11354	0,045287	2,364412	1,098829	0,539316	4,002556
6	0,208063	0,133008	0,032097	0,766731	0,324236	0,110882	1,201849
7	0,296792	0,03282	0,075103	14,87855	1,545985	3,75521	20,17975
8	0,156489	0,161025	0,041263	1,386088	1,718303	0,372237	3,476629
9	0,006052	0,12765	0,319142	0,005646	0,165533	0,195645	0,366824
10	0,556613	0,020629	0,012167	6,50062	0,284535	0,14681	6,931966
11	0,234402	0,016656	0,149609	3,492202	0,239178	2,244025	5,975405
…….	…….	…….	…….	…….	…….	…….	…….
…….	…….	…….	…….	…….	…….	…….	…….
505	0,030415	0,541696	0,008010	1,007870	16,219932	0,278363	17,506165
506	0,176736	0,302610	0,000870	4,428485	8,260474	0,020684	12,709643
						P_1 ∑	**3088,942448**

Figure 2. The distribution of FCM clustering using PCA

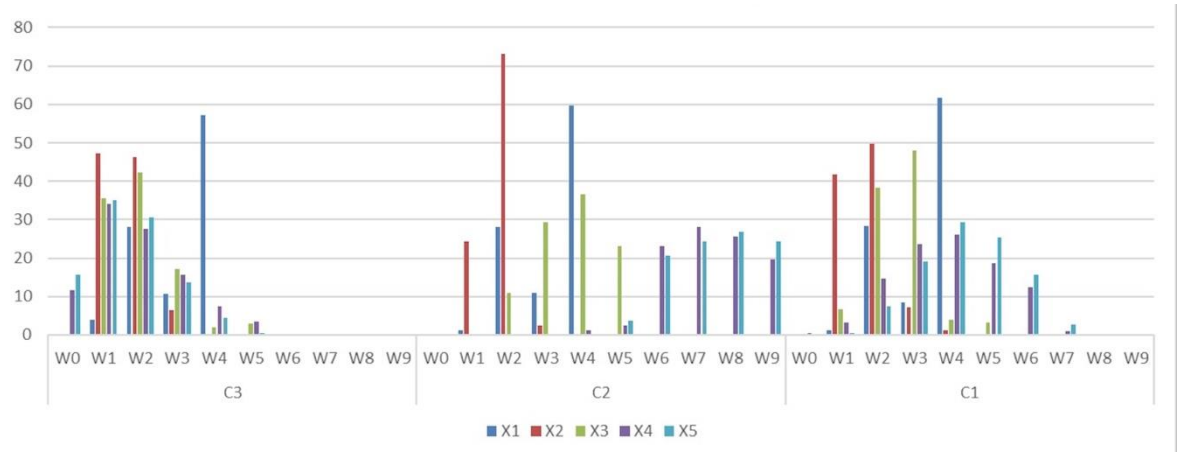


Figure 3. The percentage (%) dataset distribution of FCM clustering based on the parameters

The comparison evaluation matrices analysis is explained in Table 6. This analysis shows that CHI, DBI, and SI provide the optimum cluster structured with number cluster of 3. Herein, SI shows a calculated value of 0.199267157 and is categorized as low (0.2-0.3 as a minimum acceptable range). This is triggered by the membership overlaps amongest the clusters that have similar feature values thus the data points in different clusters are not well-separated, the average distance between points and their cluster centers won't differ much from the distance to other cluster centers. These membership overlaps frequently occur in fuzzy clustering methods, such as fuzzy c-means, as points partially belong to multiple clusters. Moreover, in many real-world datasets such as the drug addiction data, especially those involving human behavior or complex phenomena, the distinct boundaries are hard to define thus impact naturally lower silhouette scores. Unfortunately, when it compared to the calculated values of other metrics, such as CHI (195,5899065 that categorized as high) and DBI value 1.196654107 (indicates better cluster separation whereby 1.0-1.5 can be acceptable), thus the choice of cluster 3 as the most optimum cluster can be acceptable. Moreover, the comparison of SI calculation results in the other clusters (4-5) are even lower, the recommendation of cluster 3 is expected to be the best choice [38].

Table 6. The comparison of cluster evaluation

Numbers of clusters	CHI	DBI	SI
2	257,6285422	1,255475287	0,30569592
3	195,5899065	1,196654107	0,199267157
4	156,6324182	1,865339450	0,190636962
5	125,1797499	2,176182867	0,136076017

The assessment metrics for FCM indicates that the patients' profiles dataset structured and patterns (See Figure 3) whereby cluster 1 as high optimum class addiction is demonstrated by conditional parameters including the onset age of drug consumption (x1) between 21-30 years old (w4) achieved 61.78%; the urine test (x2) contains two type of drug detection (w2) within 49.78%; the prolonged use of the drug (x3) around 3-6 year at 48%; the assessment of physical effect of addiction (x4) is categorized in moderate (w4) at 26.22%; and the psychological effect of addiction (x5) is reported as moderate (w4) indication at 29.33%. Meanwhile, cluster 2 as medium drugs dependency is specified by the patients pattern into the onset age of drug consumption (x1) between 21-30 years old at 59.8%; the urine test (x2) contains two type of drug detection at 73.2%; the prolonged use of the drug (x3) between 7-10 years at 36.6%; the assessment of physical effect of addiction (x4) is clustered in High at 28.1%; and the psychological effect of addiction (x5) is reported as very high indication at 26.8%. Cluster 3 as regular class addiction that shows by the parameters indication perceives within the onset age of drug consumption (x1) between 21-30 years old at 57%; the urine test (x2) contains one type of drug detection (w1) at 47%; the prolonged use of the drug (x3) between one up to two years (w2) at 42%; the assessment of physical effect of addiction (x4) is clustered in very low(w1) at 34%; and the psychological effect of addiction (x5) is reported as very low(w1) indication at 35%. This useful information aid the medical and rehabilitation doctors in taking and decide the curatives and preventives action. [39] explains that handling drug addiction effectively requires a tailored approach based on the level of addiction whether it's in low, medium, or high addiction. The three stages of the addiction cycle emerge as a consequence of the disruption of brain networks involved with reward and motivation (reward network), executive function (executive control network), mood and stress reactivity (salience and emotion networks), and self-awareness (interoceptive and default mode networks). The length of the cycle and the prominence of each stage varies as a function of the severity of the pharmacological characteristics of the drug consumed [40]. Understanding these situation effectively provides the appropriate strategies handling at each level ranging from education and outpatient counselling for low addiction to intensive inpatient treatment for high addiction. The patients as individuals can achieve better recovery outcomes and improve their overall quality of life. A comparison of the results and the clustering process of drug addiction conducted in this study (Fuzzy C-Mean clustering) with other techniques, such as the Diagnostic and Statistical Manual of Mental Disorders (DSM 5) and the International Classification of Diseases (ICD), was undertaken. It was found that both DSM-5 and ICD provide a global framework for the classification of diseases and health conditions based on diagnostic criteria that allow clustering based on severity into three categories: mild, moderate, and severe. The severity criteria delineated by DSM-5 and ICD encompass impaired control over substance use, social impairment, risky substance use, and pharmacological criteria (tolerance and withdrawal) [41 and 42]. However, the substance use disorder classification and their categorical approach in DSM-5 and ICD may not fully leverage the complexity of individual experiences of addiction. Fuzzy C-Means clustering offers a more nuanced and flexible methodology, accommodating overlapping symptoms and dynamic assessment, ultimately enhancing understanding and treatment outcomes in drug addiction contexts. This combination of structured classification and adaptive clustering could lead to more effective treatment pathways and better outcomes for individuals struggling with addiction. Regarding on the parameters, both DSM-5 and ICD consider the age of onset of substance use in understanding the course and severity of addiction. Early onset is associated with a higher risk of developing severe substance use disorders. Herein, Fuzzy C-Mean can utilize the precise age of onset as a continuous variable, allowing for nuanced membership gradations in clustering. For example, individuals who started using drugs at a very young age may cluster differently from those who began later, leading to better-tailored interventions. Although the DSM-5 and ICD do not mandate urine tests for diagnosis, substance presence can be indicative of severity and support the clinician's overall assessment when considering the number of criteria met. Fuzzy C-Mean can analyze urine test results more dynamically, incorporating them as a continuous variable impacting cluster membership. For instance, variations in test results (e.g., frequency, intensity) could facilitate the differentiation of low, middle, and high addiction levels, allowing for multiple memberships based on substance use frequency and patterns. The DSM-5 and ICD consider prolonged use as part of the diagnostic criteria for substance use disorder. Specifically, it looks for patterns of continued use despite negative consequences, along with withdrawal symptoms or tolerance development. Meanwhile, Fuzzy C-Mean allows prolonged use to be treated as a gradient rather than a binary assessment. For example, an individual may have different levels of membership in "middle" and "high" addiction categories based on how long and consistently they have used the substance. This flexibility can highlight the severity of addiction more accurately than categorical approaches. Both DSM-5 and ICD identify physical effects (e.g., tolerance and withdrawal symptoms) as criteria contributing to substance use disorder diagnoses. These effects are considered significant indicators of addiction severity. The Fuzzy C-Mean framework can evaluate physical health impacts continuously across different dimensions of severity, allowing for profiling individuals based on specifics regarding physical health impacts and their degrees of association with clusters representing addiction severity levels. Lastly, the psychological effects, such as anxiety, depression, and cognitive

disturbances, are critical components of the DSM-5 and ICD classifications. These criteria for substance use disorders include behavioral patterns reflecting psychological distress. Both classifications techniques emphasize the interplay between mental health and substance use, recognizing that psychological effects can exacerbate addiction severity. Meanwhile, Fuzzy C-Mean in this case captures the continuum of psychological impacts by allowing individuals to belong to multiple clusters representing varying psychological severities. For instance, someone may show high psychological distress (anxiety/depression) alongside moderate addiction levels, thus informing a more multifaceted treatment approach. In nutshells, while the DSM-5 and ICD provide structured criteria for assessing drug addiction severity, Fuzzy C-Means clustering introduces flexibility and adaptability in evaluating complex addiction profiles. Integrating both traditional diagnostic criteria and advanced clustering techniques could significantly improve the treatment and understanding of substance use disorders. This result shows the efficacy of Fuzzy C-Means as soft clustering in handling the overlapping categories in each cluster and reflect the continuous nature of addiction. Thus, it can provide more nuanced subgroup definitions, proper guiding clinical interventions and psychological treatment based on the characterictics of data for each clustering level.

4. CONCLUSION

This study has been successfully cluster the patients profiles of the drug addiction by considering the several parameters including the onset age of drug consumption, the prolonged use of the drug, the urine test, and the assessment of physical and psychological effect of addiction. 506 dataset patients from National Narcotics Boards in Riau has been optimum structured into 3 clusters viz., 215 data are recorded in clusters 1 as high optimum class addiction, 105, and 186 data in clusters 2, and 3 as medium and regular addiction levels, respectively. Three evaluation approaches as well as CHI, DBI, and SI test the potentially structured of dataset clustered at cluster 3. This analysis provides the pattern of patients' profiles for the drug addiction. Thus, the National Narcotics Boards in Riau, who are the ultimate beneficiaries of this system, determine the severity of drug addiction in order to implement effective treatment programs, both preventative and curative.

ACKNOWLEDGMENTS

This research was supported by Universitas Islam Negeri Sultan Syarif Kasim Riau in Pekanbaru, Indonesia, in partnership with M.H. Dulati Taraz University in Taraz, Kazakhstan and Arab Open University, Saudi Arabia. We extend our profound gratitude to the medical professionals and therapists at the National Narcotics Boards in Riau province for their interviews and recommendations during data collection, which offered invaluable insights and knowledge that significantly enhanced our research

REFERENCES

[1] K. D. Ersche *et al.*, "Brain networks underlying vulnerability and resilience to drug addiction," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 117, no. 26, pp. 15253–15261, 2020, doi: 10.1073/pnas.2002509117.

[2] Z. Zhang, A. K. Mitra, J. A. Schroeder, and L. Zhang, "The Prevalence of and Trend in Drug Use among Adolescents in Mississippi and the United States: Youth Risk Behavior Surveillance System (YRBSS) 2001–2021," *Int. J. Environ. Res. Public Health*, vol. 21, no. 7, 2024, doi: 10.3390/ijerph21070919.

[3] M. B. Saberi Zafarghandi *et al.*, "Indicators of Drug-Related Community Impacts of Open Drug Scenes: A Scoping Review," *Eur. Addict. Res.*, vol. 28, no. 2, pp. 87–102, 2022, doi: 10.1159/000519886.

[4] S. Rehman, E. Mustafa, A. A. Faiz, M. Kanwal, F. Yasmin, and A. Fatima, "Impacts of Drug Addiction on Psychological and Emotional Health and Role of Medicinal Plants in Treatment of Drug Addiction," *J. Innov. Sci.*, vol. 8, no. 2, 2022, doi: 10.17582/journal.jis/2022/8.2.311.325.

[5] L. V. Panlilio *et al.*, "Beyond abstinence and relapse: cluster analysis of drug-use patterns during treatment as an outcome measure for clinical trials," *Psychopharmacology (Berl).*, vol. 237, no. 11, pp. 3369–3381, 2020, doi: 10.1007/s00213-020-05618-5.

[6] Y. Deng, X. Xu, Y. Qiu, J. Xia, W. Zhang, and S. Liu, "A multimodal deep learning framework for predicting drug-drug interaction events," *Bioinformatics*, vol. 36, no. 15, pp. 4316–4322, 2020, doi: 10.1093/bioinformatics/btaa501.

[7] W. W. Xu *et al.*, "Unmasking the Twitter Discourses on Masks During the COVID-19 Pandemic: User Cluster-Based BERT Topic Modeling Approach," *JMIR Infodemiology*, vol. 2, no. 2, pp. 1–29, 2022, doi: 10.2196/41198.

[8] S. Park *et al.*, "Improving unsupervised image clustering with robust learning," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, no. 2, pp. 12273–12282, 2021, doi: 10.1109/CVPR46437.2021.01210.

[9] O. N. Kenger, Z. D. Kenger, E. Ozceylan, and B. Mrugalska, "Clustering of Cities Based on Their Smart Performances: A Comparative Approach of Fuzzy C-Means, K-Means, and K-Medoids," *IEEE Access*, vol. 11, no. October, pp. 134446–134459, 2023, doi: 10.1109/ACCESS.2023.3333753.

[10] B. Lund and J. Ma, "A review of cluster analysis techniques and their uses in library and information science research: k-means and k-medoids clustering," *Perform. Meas. Metrics*, vol. 22, no. 3, pp. 161–173, 2021, doi: 10.1108/PMM-05-2021-0026.

[11] O. Al-Janabee and B. Al-Sarray, "Evaluation Algorithms Based on Fuzzy C-means for the Data Clustering of Cancer Gene Expression," *Iraqi J. Comput. Sci. Math.*, vol. 3, no. 2, pp. 27–41, 2022, doi: 10.52866/ijcsm.2022.02.01.004.
[12] V. Rayala and S. R. Kalli, "Big data clustering using Improvised Fuzzy C-Means clustering," *Rev. d'Intelligence Artif.*, vol. 34, no. 6, pp. 701–708, 2021, doi: 10.18280/RIA.340604.
[13] Wani, A. A. (2024). Comprehensive analysis of clustering algorithms: exploring limitations and innovative solutions. *PeerJ Computer Science*, *10*, e2286.
[14] Gatea, A. N., & AL-Asadi, H. A. A. (2024). Soft Clustering Techniques: An In-Depth Analysis of GMM and FCM Algorithms and Comparative Performance. *Basrah Researches Sciences*, *50*(2), 223-238.
[15] Hajihosseinlou, M., Maghsoudi, A., & Ghezelbash, R. (2024). Intelligent mapping of geochemical anomalies: Adaptation of DBSCAN and mean-shift clustering approaches. *Journal of Geochemical Exploration*, *258*, 107393.
[16] Hashemi, S. E., Gholian-Jouybari, F., & Hajiaghaei-Keshteli, M. (2023). A fuzzy C-means algorithm for optimizing data clustering. *Expert Systems with Applications*, *227*, 120377.
[17] X. Shu and Y. Ye, "Knowledge Discovery : Methods from data mining and machine learning ☆," *Soc. Sci. Res.*, vol. 110, no. April 2022, p. 102817, 2023, doi: 10.1016/j.ssresearch.2022.102817.
[18] I. T. Ayorinde, "Applying Machine Learning Technique for Knowledge Discovery in Network Database," no. July, 2024, doi: 10.35444/IJANA.2024.16109.
[19] V. Plotnikova, M. Dumas, and F. P. Milani, "Applying the CRISP-DM data mining process in the financial services industry: Elicitation of adaptation requirements," *Data Knowl. Eng.*, vol. 139, no. March, p. 102013, 2022, doi: 10.1016/j.datak.2022.102013.
[20] L. S. Okfalisa, Elvia Budianita, Rezi Yuliani *et al.*, "Forecasting company financial distress: C4.5 and adaboost adoption," vol. 49, no. 3, pp. 300–307, 2022, doi: 10.14456/easr.2022.31.
[21] K. Gjini *et al.*, "The use of AGILE 3+ and AGILE 4 for the prediction of advanced fibrosis and cirrhosis in patients with Non-Alcoholic Fatty Liver Disease," *Dig. Liver Dis.*, vol. 55, pp. S31–S32, 2023, doi: 10.1016/j.dld.2023.01.060.
[22] M. Jansevskis and K. Osis, "Securing the Future: The Role of Knowledge Discovery Frameworks," *Artif. Intell. Secur.*, pp. 85–101, 2024, doi: 10.1007/978-3-031-57452-8_5.
[23] A. H. Al-Faouri, "Adopting Data Mining as a Knowledge Discovery Tool: The Influential Factors from the Perspectives of Information Systems Managers," *Inf. Sci. Lett.*, vol. 12, no. 5, pp. 1851–1861, 2023, doi: 10.18576/isl/120529.
[24] R. Zebari, A. Abdulazeez, D. Zeebaree, D. Zebari, and J. Saeed, "A Comprehensive Review of Dimensionality Reduction Techniques for Feature Selection and Feature Extraction," *J. Appl. Sci. Technol. Trends*, vol. 1, no. 1, pp. 56–70, 2020, doi: 10.38094/jastt1224.
[25] E. Pekel Özmen and T. Küçükdeniz, "Two-stage clustering and routing problem by using FCM and K-means with genetic algorithm," *Sigma J. Eng. Nat. Sci.*, vol. 42, no. 4, pp. 1030–1038, 2024, doi: 10.14744/sigma.2023.00096.
[26] S. E. Hashemi, F. Gholian-Jouybari, and M. Hajiaghaei-Keshteli, "A fuzzy C-means algorithm for optimizing data clustering," *Expert Syst. Appl.*, vol. 227, no. April, p. 120377, 2023, doi: 10.1016/j.eswa.2023.120377.
[27] Aşut, G., Alıcı, Y. H., Ceran, S., Danışman, M., & Şahiner, Ş. Y. (2024). Affective neuroscience personality traits in opioid use disorder patients: the relationship with earlier onset of substance use, the severity of addiction, and motivational factors to quit opiate use. Brain and Behavior, 14(9). https://doi.org/10.1002/brb3.70050
[28] Compton, W. M., Valentino, R. J., & DuPont, R. L. (2020). Polysubstance use in the u.s. opioid crisis. Molecular Psychiatry, 26(1), 41-50. https://doi.org/10.1038/s41380-020-00949-3
[29] Stringfellow, E. J., Lim, T. Y., Humphreys, K., DiGennaro, C., Stafford, C., Beaulieu, É., … & Jalali, M. S. (2022). Reducing opioid use disorder and overdose deaths in the united states: a dynamic modeling analysis. Science Advances, 8(25). https://doi.org/10.1126/sciadv.abm8147.
[30] Udoh, E., Omorere, D. E., Olarewaju, S. O., Osasu, O. S., & Amoo, B. (2021). Psychological distress and burden of care among family caregivers of patients with mental illness in a neuropsychiatric outpatient clinic in nigeria. Plos One, 16(5), e0250309. https://doi.org/10.1371/journal.pone.0250309
[31] Ma, Z., Liu, Y., Wan, C., Jiang, J., Li, X., & Zhang, Y. (2022). Health-related quality of life and influencing factors in drug addicts based on the scale qlicd-da: a cross-sectional study. Health and Quality of Life Outcomes, 20(1). https://doi.org/10.1186/s12955-022-02012-x
[32] Z. Liang, C. Wang, S. Han, K. Ullah Jan Khan, and Y. Liu, "Classification and susceptibility assessment of debris flow based on a semi-quantitative method combination of the fuzzy C-means algorithm, factor analysis and efficacy coefficient," *Nat. Hazards Earth Syst. Sci.*, vol. 20, no. 5, pp. 1287–1304, 2020, doi: 10.5194/nhess-20-1287-2020.
[33] M. Shutaywi and N. N. Kachouie, "Silhouette analysis for performance evaluation in machine learning with applications to clustering," *Entropy*, vol. 23, no. 6, pp. 1–17, 2021, doi: 10.3390/e23060759.
[34] I. F. Ashari, E. Dwi Nugroho, R. Baraku, I. Novri Yanda, and R. Liwardana, "Analysis of Elbow, Silhouette, Davies-Bouldin, Calinski-Harabasz, and Rand-Index Evaluation on K-Means Algorithm for Classifying Flood-Affected Areas in Jakarta," *J. Appl. Informatics Comput.*, vol. 7, no. 1, pp. 89–97, 2023, doi: 10.30871/jaic.v7i1.4947.
[35] S. P. Lima and M. D. Cruz, "A genetic algorithm using Calinski-Harabasz index for automatic clustering problem," *Rev. Bras. Comput. Apl.*, vol. 12, no. 3, pp. 97–106, 2020, doi: 10.5335/rbca.v12i3.11117.
[36] W. Chang *et al.*, "Analysis of university students' behavior based on a fusion K-means clustering algorithm," *Appl. Sci.*, vol. 10, no. 18, 2020, doi: 10.3390/APP10186566.
[37] S. Gowda, C. S. Nandan, M. A. Jayaram, A. Gupta, and R. S. Jaya, "Unsupervised Clustering of Asphalt Pavement Conditions Using Fuzzy C-Means Algorithm with Principal Component Analysis Aided Dimensionality Reduction,"

Lect. Notes Networks Syst., vol. 831, pp. 35–45, 2024, doi: 10.1007/978-981-99-8135-9_4.
[38] Guo, X., Li, J., & Xu, H. (2021). *An improved silhouette coefficient method for determining the optimal number of clusters.* Expert Systems with Applications, 177, 114921. doi**:** 10.1016/j.eswa.2021.114921
[39] N. D. Volkow and C. Blanco, "Substance use disorders : a comprehensive update of classification , epidemiology , neurobiology , clinical aspects , treatment and prevention," no. June, pp. 203–229, 2023, doi: 10.1002/wps.21073.
[40] S. Marinelli, G. Basile, R. Manfredini, and S. Zaami, "Sex- and Gender-Specific Drug Abuse Dynamics: The Need for Tailored Therapeutic Approaches," *J. Pers. Med.*, 2023, doi: 10.3390/jpm13060965.
[41] Petry, J. (2021). A Critical Look at the Concept of Addiction in DSM-5 and ICD-11. *American Journal of Psychiatry and Neuroscience*, *9*(1), 7-14.
[42] Reichert, R. A., da Silva, E. A., De Micheli, D., Noto, A. R., Skinstad, A. H., & Galduróz, J. C. F. (2025). Substance Use Disorders: History, Theoretical Models and Diagnostic Criteria (ICD-11 e DSM-5-TR). In *Neuropsychology and Substance Use Disorders: Assessment and Treatment* (pp. 3-49). Cham: Springer Nature Switzerland.

BIOGRAPHY OF AUTHORS

Dr. Okfalisa is Professor of Informatics Enggineering Department, Faculty Science and Technology, Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia. She completed her PhD in Information System and Computer Science at Universiti Teknologi Malaysia in 2012. Her research interests are performance measurement, strategy execution, management information system, knowledge management, knowledge management system, decision support system, data science, machine learning, and data mining. She was once involved in several research projects and publications, national (cooperation with UIN Jakarta; UIN Bandung; UIN Malang; UGM; Amikom Yogya and international (cooperation with Universiti Teknologi Malaysia- Johor Bahru Malaysia; Price of Songkla University-Hat Yai; Universiti Teknologi MARA-Perlis, Malaysia; International Islamic University Malaysia (IIUM)-Kuala Lumpur, Malaysia, University De Haute-Alsace-France; Ostbayerische Technische Hochschule Amberg-Weiden (OTH-AW)-German; University of Antwerp-Belgium; and Sheffield Hallam University-UK; Anglia Ruskin University Cambridges-UK; Birmingham City University, Birmingham-UK; M.H. Dulati Taraz University, Taraz, Kazakhstan; Arab Open University, Saudi Arabia) from UIN Suska Riau, Indonesia Islamic Higher Education Minister and Bank Indonesia Institute, Indonesia). She reviewed several local and international journals. She was awarded two silvers and one bronze medal during the paper presentation and exhibition at the 2nd International Indonesia, Malaysia, and Thailand Symposium on Innovation and Creativity 2018. She also involves in many certifications as an assessor for Computer field expertises. She is a life member of the Association Higher Education Informatics and Computers (Aptikom) and the Institute of Electrical and Electronics Engineers (IEEE).

Pramulia Suliandri is a full-stack software developer with over four years of experience in the field of information technology. He is a graduate of Informatics Engineering from UIN Suska Riau since 2020, and has also obtained a certificate of competence as a programmer from the National Professional Certification Agency. He has expertise in the development of data science-based systems and data mining, which are utilized to provide decision-makers with valuable assistance during their decision-making processes. He currently serves as the head of the IT department of an educational company.

Dr. Ola A. Alwesabi is an Assistant Professor at the Faculty of Computer Studies, Arab Open University, Saudi Arabia. She received her PhD in Computer Science from Universiti Sains Malaysia (USM) in 2018, specializing in Distributed Multimedia and Multimedia Communication. She has published numerous research papers in high-reputation international journals and conferences. Her research interests include Distributed Multimedia, Computer Networks, Information Security, Cryptography, Cloud Computing, and Artificial Intelligence.

Dr. Bolat Tassuov is an Associate Professor of Physics and Informatics Department, M.H. Dulati Taraz University, Taraz, Kazakhstan. He graduated from Zhambyl University for mathematics and computer science degree. Currenlty he is the dean of faculty natural science and his research interest are on Information and Communication, Computer Architecture, Information Security based on the pedagogy. He published many scientific works national and international, especially in Kazakhstan.

Raimbayeva Saule is senior lecture at Physics and Informatics Department, M.H. Dulati Taraz University, Taraz, Kazakhstan. Currently, she is an acting master for head of department physics and informatics at M.H. Dulaty Taraz University. Her research interest are educational based information technology, programming, and computer graphics. She graduated from Zhambyl Pedagogical Institute with a degree in Computer Science and Mathematics Teacher and continued her master's degree at the Kyrgyz State University as master of Pedagogy.

Kurdyukov Vitaliy is senior lecture at Physics and Informatics Department, M.H. Dulati Taraz University, Taraz, Kazakhstan. His reseach interests are on educational based information and communication technology, and computer architecture.

Dr. Zahidah Zulkifli is an Associate Professor at the Kulliyyah (Faculty) of Information Communication and Technology (KICT) at the International Islamic University Malaysia (IIUM). With a distinguished career in academia and research, her expertise lies at the intersection of information sciences, IT management, and governance. Dr. Zahidah earned her PhD in Information Sciences from Universiti Malaya (UM) in 2013. Prior to that, she completed both her Master's in IT Management and Bachelor's in Computer Science at Universiti Teknologi Malaysia (UTM) between 2003 and 2007, laying a strong foundation in technology and management. Her research focuses on areas such as information systems governance, ethical practices in information system development, and the role of ICT in organizational processes. As a faculty member at IIUM, Dr. Zahidah plays an active role in teaching, research, and mentoring students. She is also a key member of the Information System, Internet, and Governance Research Unit, contributing to initiatives aimed at advancing knowledge in ICT and its responsible use. In addition to her academic work, Dr. Zahidah is passionate about fostering ethical leadership and collaboration within system development methodologies. She actively engages in interdisciplinary projects and collaborations, contributing to the academic community through conferences, publications, and research projects. Her dedication to teaching, research excellence, and academic service reflects her commitment to nurturing future generations of ICT professionals

Dr. Lestari Handayani received the Ph.D. degree from INSA Centre Val de Loire, France, since 2022. She is currently an Assistant Professor at the Department of Informatics Engineering, Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia. She is also the head of the Artificial Intelligent of Things (AIoT) research team. Her research areas include the study of Prognostics and Health Management, pattern recognition, signal processing techniques, image processing, data mining, and Artificial Intelligent.