K-means Clustering of Intracellular Calcium Signal Transients

Ivan Lorencin

Juraj Dobrila University of Pula, Faculty of Informatics, Rovinjska 14, 52100 Pula, Hrvatska, ivan.lorencin@unipu.hr

Domagoj Frank

Department of Computer Science and Informatics, University North Koprivnica, Trg dr. Žarka Dolinara 1, 48000 Koprivnica, domagoj.frank@unin.hr

Nikola Tanković

Juraj Dobrila University of Pula, Faculty of Informatics, Rovinjska 14, 52100 Pula, Hrvatska, nikola.tankovic@unipu.hr

Tomislav Horvat

Department of Electrical Engineering, University North Varaždin, Jurja Križanića 31b; 42000 Varaždin, tomislav.horvat@unin.hr

Abstract: Neurological diseases pose a significant public health challenge, leading to disability and mortality globally. Current diagnostics for neuroinflammatory diseases are complex and lack efficacy, necessitating invasive procedures. Calcium signaling dynamics in astrocytes and microglia play pivotal roles in central nervous system (CNS) function and dysfunction. This study curates time-series data on calcium transients in astrocytes and microglia, employing live-cell imaging techniques and preprocessing methodologies. Using k-means clustering, we analyze the data, revealing optimal clustering solutions between 2 to 6 clusters. This research offers valuable insights for understanding CNS disorders and highlights the potential of clustering techniques in neurological research.

Keywords: Artificial intelligence; Calcium signaling; K-means clustering; Unsupervised learning

1 Introduction

Neurological diseases are a major public health challenge worldwide, affecting millions of people and causing significant disability and death. According to the Global Burden of Disease (GBD) 2016 study, neurological disorders were the leading cause of disability-adjusted life years (DALYs) globally and the second cause of death after cardiovascular diseases [1].

Currently, the diagnosis of neuroinflammatory diseases is inadequate and generally entails complicated and costly procedures to rule out other conditions. These diagnostic methods are frequently invasive for the patient. Diagnosis typically occurs only after the disease's symptoms have become apparent. The diagnostic protocol for neuroinflammatory diseases is a long-term and demanding process that requires a multidisciplinary approach and the cooperation of various specialists. Currently, there is no single test that can confirm or rule out any of these diseases [2].

When astrocytes, a type of non-neuronal cells, fail to function normally, there is a disruption in the neuron-supportive activities they typically perform, leading to neuron degeneration. It is understood that signaling by intracellular calcium (Ca2+) plays multiple distinct roles in a wide range of physiological and pathophysiological processes within the nervous system [3-5]. In astrocytes, the concentration of intracellular calcium increases in response to a range of stimuli, a process enabled by the combined activity of various molecular pathways. Moreover, in the investigation of potential mechanisms underlying a specific neurodegenerative disease (with a neuroinflammatory component), researchers observed an increase in the transient levels of intracellular Ca2+ within cultured rat cortical astrocytes. This increase was a response to certain treatments, as highlighted in a previous study by our colleagues [6]. We employed a comparable set of raw data with the aim of conducting a more thorough analysis and identifying a method to cluster these data based on their temporal patterns.

The method for monitoring intracellular Ca2+ signaling relies solely on using indicators that emit fluorescence when they bind to the ion. These indicators are either chemical molecules, typically structured based on Ca2+ buffers, or proteins introduced into cells either through genetic expression or vectors [7]. The large somatic region of astrocytes is easily identifiable under a fluorescent microscope, designated as the so-called Region of Interest (ROI), simplifying the process of tracking Ca2+ fluctuations originating from the astrocytic soma [8]. In the observed field of view, every individual cell has the potential to react to a particular chemical treatment or set of successive treatments. The reaction of each cell is captured in the recording, creating a series of data points over time – a time-series, which represents each cell’s response.

Complexity of the recorded time-series call for the application of Artificial Intelligence (AI) methods. Astrocytes have been shown to respond very differently,

even to the same stimulus - only once immediately after stimulation, or exhibit Ca2+ oscillations with generally constant frequency and constant amplitude, or oscillations with decreasing frequency or decreasing amplitude [9]. A major challenge in studying astrocyte Ca2+ behavior is the spatial and temporal buildup of Ca2+ fluctuations. This accumulation manifests as intracellular and intercellular Ca2+ waves occurring over time frames that range from a few hundred milliseconds to several seconds. These waves are often observed, particularly in cell cultures, as noted in previous studies [10, 11].

The curation of time-series data involves a systematic process encompassing the collection, organization, and analysis of temporal changes in Ca2+ concentrations within these cell cultures. Following data acquisition, meticulous preprocessing steps, including background subtraction, noise reduction, and normalization, are employed to accurately represent calcium fluctuations over time. Subsequently, data analysis tools and algorithms extract critical information from the time-series, encompassing amplitude, frequency, duration, and spatial distribution of calcium transients.

Therefore, in our comprehensive review article [12], we have identified a spectrum of methodologies currently in use for time series analysis within the Machine Learning paradigm and identified a clear gap between the application of state-of-the-art AI methods in the classification and clustering of Ca2+ signals and the application in the classification and clustering of other biomedical signals. In this article, we examine the possibility of applying unsupervised learning methods for k-means clustering of Ca2+ signals. The effectiveness of various time-series analysis methods, such as clustering, largely depends on choosing an appropriate distance measure. This selection is often deemed more crucial than the choice of the clustering algorithm itself in time-series analysis. Ideally, clustering based on shape should group patterns that are similar, regardless of differences in amplitude and phase. However, since “shape” is a vague concept, a variety of distance measures have been developed to accommodate different types of data distortions. Hence, employing appropriate clustering methods assists in comprehending which features of the data are most used by the ML algorithms to categorize time-series.

In this article we examine the possibility of applying unsupervised learning methods for k-means clustering of Ca2 signals.

2 Time-Series Curation

Primary cultures of astrocytes and microglia are pivotal tools in elucidating the pathophysiology of Amyotrophic Lateral Sclerosis (ALS), a progressive neurodegenerative disorder primarily impacting motor neurons. These primary cell cultures provide valuable insights into the intricate interplay between different cell

types within the central nervous system (CNS), shedding light on the underlying mechanisms associated with ALS.

Astrocytes, a type of glial cell, and microglia, the immune cells of the CNS, are essential for maintaining CNS homeostasis and responding to pathological stimuli. In ALS, alterations in these cell types contribute significantly to disease progression. Primary cultures derived from brain tissue offer an in vitro platform to isolate and propagate pure populations of astrocytes and microglia.

These primary cell cultures serve as critical models to study the roles of astrocytes and microglia in ALS pathogenesis. Researchers use these cultures to explore the cellular responses of astrocytes and microglia to ALS-related pathological factors, such as inflammation, oxidative stress, and excitotoxicity. Understanding how these cells release inflammatory mediators and neurotoxic factors provides crucial insights into their contributions to neuronal degeneration in ALS.

Furthermore, primary cultures enable investigations into the intricate interactions between astrocytes, microglia, and neurons. By deciphering the communication pathways and molecular signaling involved in these interactions, researchers can elucidate how astrocytes and microglia contribute to neuroinflammation, glial reactivity, and neuronal loss observed in ALS.

Additionally, these cell cultures are instrumental in drug screening and therapeutic development. Researchers use these models to assess potential therapeutic agents targeting astrocyte and microglial dysfunctions associated with ALS. Identifying compounds that modulate inflammatory responses or enhance neuroprotection holds promise for developing interventions that could potentially slow the progression of ALS or alleviate its symptoms.

Moreover, primary cultures provide a platform to explore the impact of genetic and environmental factors associated with ALS. Incorporating cells from ALS models or patient-derived samples allows researchers to investigate specific genetic mutations or environmental triggers, shedding light on their effects on astrocyte and microglial behavior and their contributions to ALS pathology.

Calcium signaling in primary cultures of astrocytes and microglia serves as a pivotal aspect in understanding the intricate dynamics of cellular responses within the central nervous system (CNS). These cultures, derived from astrocytes and microglia offer an invaluable platform for studying the dynamics of calcium signaling, which play crucial roles in intercellular communication, homeostasis maintenance, and responses to pathological stimuli [12].

3 Data Set Description

The curation of time-series data involves a systematic process encompassing the collection, organization, and analysis of temporal changes in calcium ion (Ca2) concentrations within these cell cultures. Initial acquisition of time-series data involves the use of live-cell imaging techniques like fluorescence microscopy, often using calcium-sensitive fluorescent dyes or genetically encoded calcium indicators. These techniques enable real-time visualization and recording of intracellular calcium dynamics in response to specific experimental conditions or stimuli. Following data acquisition, meticulous preprocessing steps, including background subtraction, noise reduction, and normalization, are employed to accurately represent calcium fluctuations over time. Subsequently, data analysis tools and algorithms extract critical information from the time-series, encompassing amplitude, frequency, duration, and spatial distribution of calcium transients.

In primary astrocyte cultures, calcium signaling often manifests as propagating waves or localized elevations in response to neurotransmitters or neuronal activity. Curating time-series data from astrocytic cultures enables the characterization of calcium wave propagation patterns and their influence on neuronal function. Similarly, calcium signaling in microglial cultures plays a crucial role in immune surveillance and responses to neuronal injury or inflammation. Curating time-series data from microglia allows the exploration of calcium dynamics associated with various cellular activities, such as microglial activation, migration, and cytokine release, in response to different stimuli. The curated time-series data serves as a fundamental resource for quantitative analysis, statistical modeling, and machine-learning approaches. It aids in deciphering spatiotemporal properties and functional implications of calcium signaling events in astrocytes and microglia. Furthermore, this curated data provides insights into cellular interactions within the CNS, contributing to a comprehensive understanding of CNS function and dysfunction in health and disease conditions, including neuroinflammatory disorders and neurodegenerative diseases like Alzheimer's

The dataset was collected from patients across diverse regions of interest, ensuring representation across different biological contexts. Each time series was meticulously recorded to capture the intensity of calcium transients over time. To facilitate comparative analyses and ensure data integrity, normalization procedures were applied to standardize the datasets, minimizing variations arising from experimental conditions and individual differences. The dataset comprises 670 time series, each representing the temporal evolution of calcium transients within specific biological samples. Examples of curated time-series are given in Figures 1, 2, 3, and 4.

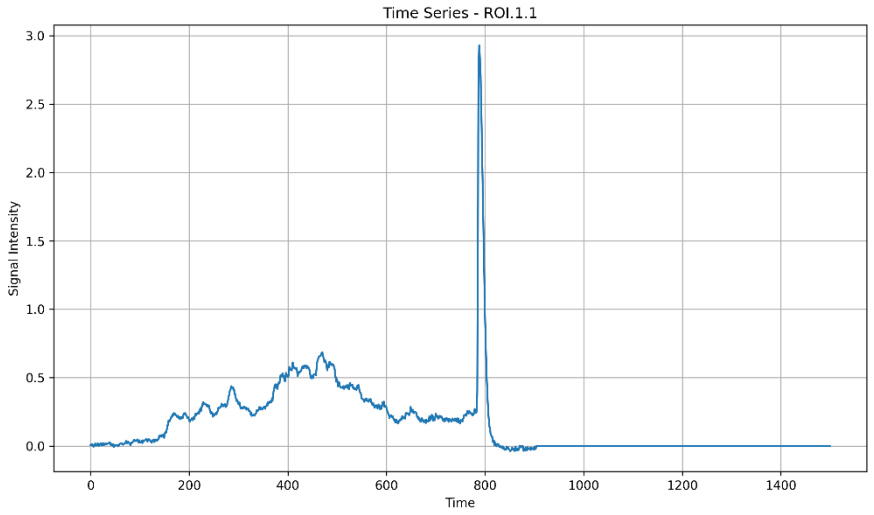

Figure 1
Plot of Time series marked as ROI 1.1

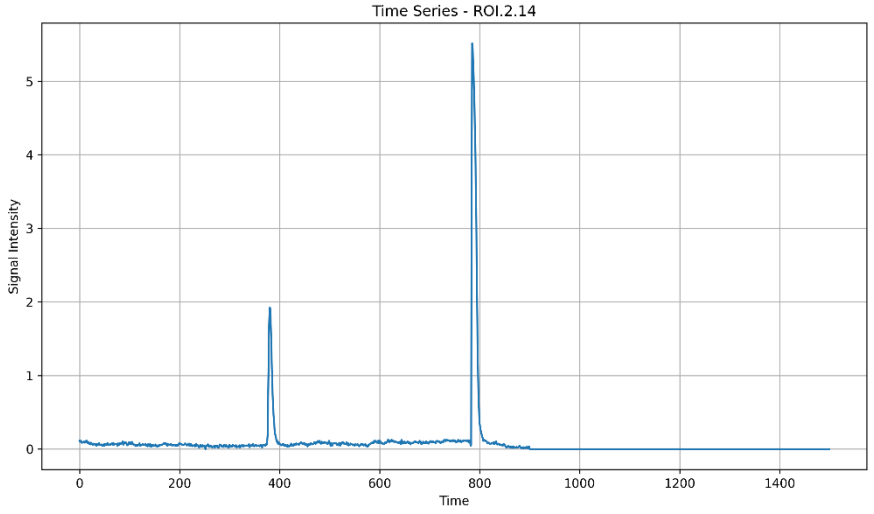

Figure 2
Plot of Time series marked as ROI 2.14

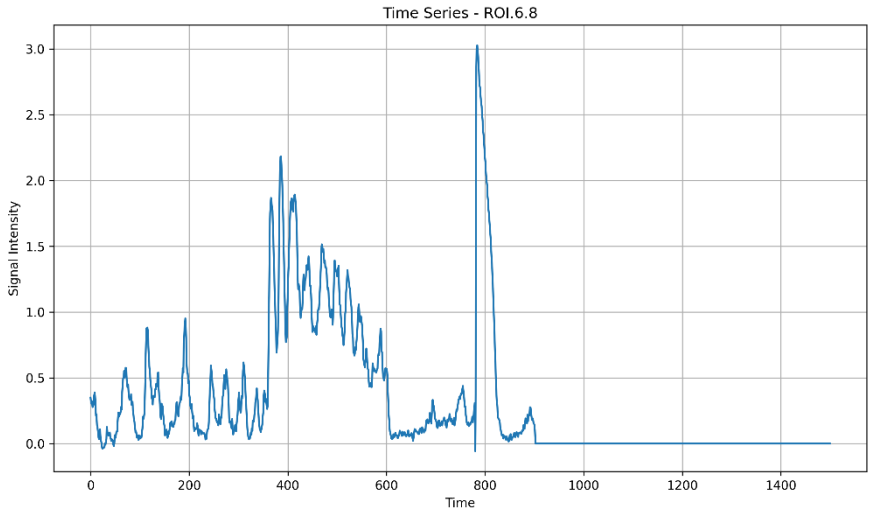

Figure 3
Plot of Time series marked as ROI 6.8

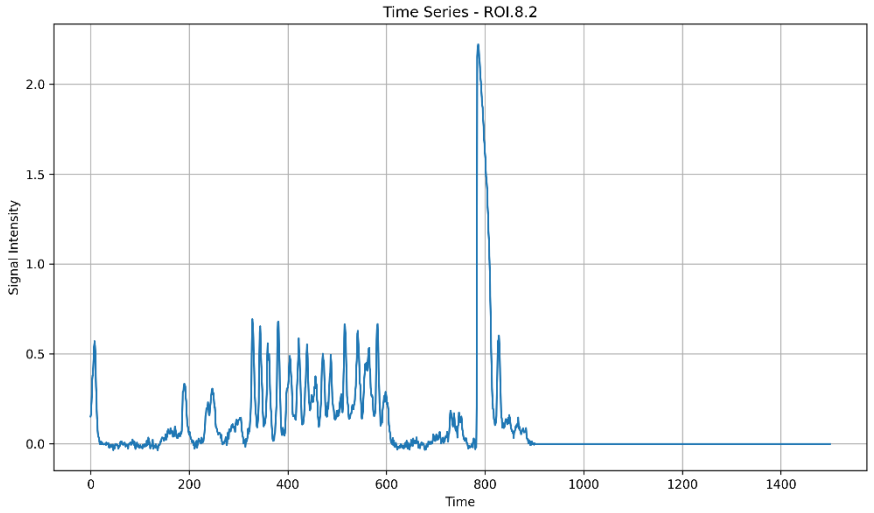

Figure 4
Plot of Time series marked as ROI 8.2

When examining various time series, it becomes evident that they exhibit diverse patterns and behaviors. Each time series may demonstrate unique trends, fluctuations, or periodicities, reflecting the underlying dynamics and characteristics of the phenomena being observed.

4 K-means Clustering

Clustering stands as a foundational technique within machine learning and data mining, serving the purpose of grouping data points based on similarities. Among the clustering algorithms, K-means stands out due to its popularity and efficiency. The primary objective of the K-means algorithm involves partitioning a dataset into k clusters, assigning each data point to the nearest cluster centroid [13]. Initially, k centroids representing cluster centers are placed randomly in the feature space. These centroids can be chosen randomly from the dataset or using techniques like k-means++ for improved convergence and result quality. The subsequent step involves assigning each data point to the nearest centroid, typically calculated through Euclidean distance. Based on these assignments, data points are clustered accordingly. Iteratively, centroids are updated by computing the mean of the data points within each cluster, recalculating their positions [14]. This assignment and update process continues until convergence, where either there is no significant change in centroids' positions or the maximum iteration limit is reached. An illustration of the K-means clustering procedure is given in Figure 6.

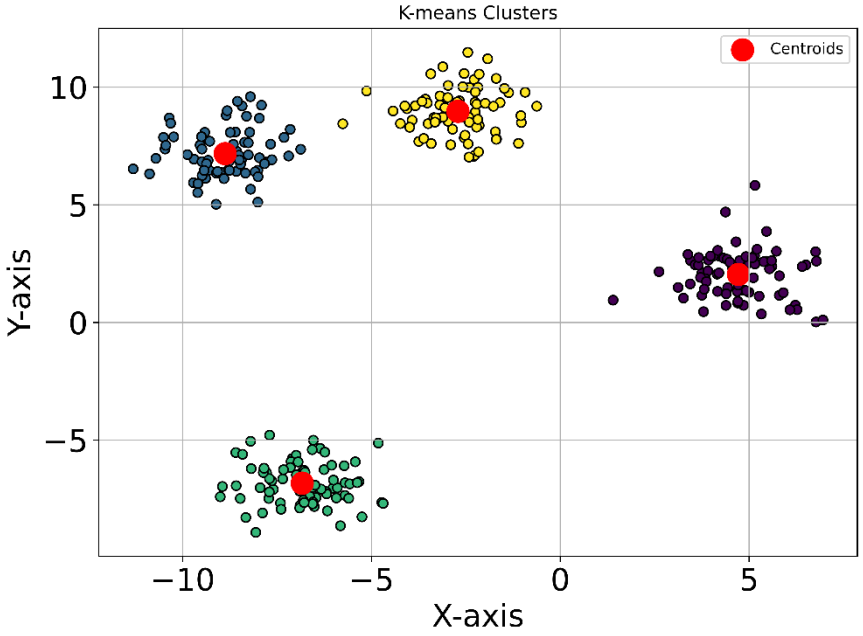

Figure 5
Illustration of k-means clustering procedure

5 K-Means Clustering for Time-Series Data

Clustering time-series data involves categorizing sequential observations recorded over successive time intervals into meaningful groups or clusters. K-means clustering, primarily designed for static data has been adapted to address the unique characteristics of time-series data [15].

5.1 Characteristics of Time-Series Data

Observations in time-series data are not independent; they exhibit dependencies across time points. This characteristic requires clustering methods capable of considering these temporal relationships while forming clusters. Furthermore, time-series sequences often have varying lengths, which traditional k-means struggles to handle. Adapting clustering techniques or employing distance measures accommodating variable lengths becomes essential. Time-series data may contain irregular patterns or noise, influencing the clustering process. Ensuring robustness in clustering algorithms to handle such irregularities is crucial.

5.2 Approaches in Clustering Time-Series Data

To apply k-means clustering to time-series data, appropriate representations and adaptations are necessary [16].

Feature-based representations involve extracting pertinent features from time-series data, such as statistical measures like mean, variance, skewness, and kurtosis, spectral features including Fourier coefficients or wavelet energies, or domain-specific attributes like trend slopes or periodicity indicators. These features are used to transform the raw time-series data into feature vectors suitable for clustering analysis, enabling the representation of complex temporal patterns in a format conducive to clustering algorithms.

Symbolic Aggregate Approximation (SAX) is a technique used to represent time-series data as sequences of symbols based on predefined breakpoints. By employing SAX, time-series patterns can be effectively represented while reducing the dimensionality of the data. This reduction in dimensionality aids in mitigating the curse of dimensionality and computational complexity often encountered in clustering large-scale time-series datasets, thus facilitating more efficient and scalable clustering analysis.

Shapelets, a concept derived from time-series data mining, are discriminative subsequences identified within time-series data that capture significant temporal characteristics. These shapelets serve as representatives for clustering purposes, enabling the identification of meaningful patterns and clusters within complex time-series datasets. By leveraging shapelets, clustering algorithms can effectively capture the intrinsic structure and variability present in time-series data, leading to more accurate and interpretable clustering results.

6 Used Metrics

Assessing the quality and performance of clustering algorithms is essential to determine their effectiveness in partitioning data into meaningful clusters. Various evaluation metrics are employed to measure the goodness of clustering results.

6.1 Silhouette Score

A widely used metric for clustering is the silhouette score, which quantifies the cohesion and separation of clusters [17]. The Silhouette Score is a metric used to calculate the goodness of a clustering technique on a dataset. It quantifies how well each data point fits into its assigned cluster, based on both the distance from the data point to other points in the same cluster (cohesion) and the distance from the data point to points in other clusters (separation). The Silhouette Score ranges from -1 to 1, where a high value indicates that the data point is well matched to its cluster and poorly matched to neighboring clusters. It is calculated for each data point *i* as [18]:

$$silhouette = \frac{b(i)-a(i)}{(a(i),b(i))} \qquad (1)$$

where:

- *a(i)* represents the average distance of point *i* to other data points within the same cluster, and
- *b(i)* is the smallest average distance of point *i* to points in any other cluster, minimizing the inter-cluster distance for point *i*.

The silhouette score is a measure used to evaluate the quality of clustering results by quantifying how similar an object is to its cluster compared to other clusters. It ranges from -1 to 1, where a score close to +1 indicates good clustering, around 0 suggests boundary cases, and close to -1 indicates possible misclassification. A high positive score suggests strong and compact cluster structures, while low negative scores indicate potential misclassifications or poor clustering performance. Interpreting silhouette scores should consider specific dataset characteristics, distance metrics, and clustering algorithms employed.

6.2 Davies-Bouldin Index

The Davies-Bouldin Index (DB) is a measure of cluster separation and cohesion. It evaluates the quality of clustering by considering the average similarity between each cluster and its most similar cluster, normalized by the average dissimilarity within clusters [19]. Lower values of the DB indicate better clustering, with each cluster being well-separated from others and internally cohesive. It can be defined as:

$$DB = \frac{1}{n}\sum_{i=1}^{n} \left(\frac{\sigma_i+\sigma_j}{d(c_i,c_j)}\right), \qquad (2)$$

where:

- n is the total number of clusters,
- σ_i is the average distance between each point in cluster *i* and the centroid c_i of that cluster, and
- $d(c_i, c_j)$ is the distance between centroids c_i and c_j, typically calculated using Euclidean distance.

The Davies-Bouldin Index is a metric used to evaluate clustering results by measuring the average similarity between each cluster and its most similar cluster, relative to the compactness of each cluster. A lower Davies-Bouldin Index indicates better clustering performance, suggesting well-separated clusters with high intra-cluster similarity. Conversely, a higher index suggests poorer clustering, indicating less distinct clusters with lower intra-cluster similarity or potential overlap. When comparing multiple clustering solutions, the solution with the lower Davies-Bouldin Index is generally preferred as it signifies more cohesive and well-separated clusters.

6.3 Calinski-Harabasz Index

The Calinski-Harabasz Index (CH), also known as the Variance Ratio Criterion, is a measure of cluster separation and compactness. It calculates the ratio of the sum of between-cluster dispersion to the sum of within-cluster dispersion for all clusters. Higher values of CH indicate better-defined and more compact clusters [20]. This index is particularly useful for determining the optimal number of clusters in a dataset. CH index can be defined as:

$$CH = \frac{B}{W} \times \frac{n-k}{k-1} \qquad (3)$$

where:

- B is the between-cluster dispersion,
- W is the within-cluster dispersion,
- k is the number of clusters, and
- n is the total number of data points.

A higher Calinski-Harabasz Index indicates better clustering performance, suggesting more distinct and well-separated clusters. Conversely, a lower index may indicate less effective clustering, potentially with clusters that are less distinct or more overlapping. When comparing multiple clustering solutions, the one with the higher Calinski-Harabasz Index is generally preferred as it signifies stronger and more compact clusters. Overall, the Calinski-Harabasz Index helps in assessing the effectiveness of clustering algorithms and parameter settings, aiding in the identification of solutions that yield more meaningful and interpretable clusters.

7 Results and Discussion

When the silhouette score is around 0.5 for cluster numbers ranging from 2 to 6, it suggests that the clusters are well-separated and distinct, with objects exhibiting relatively high cohesion within their clusters and significant separation from neighboring clusters. As the number of clusters increases beyond 6, the silhouette score begins to decline, dropping to around 0.35, indicating a decrease in cluster quality. This decline suggests that increasing the number of clusters may lead to clusters that are less well-defined or more overlapping. Furthermore, as the number of clusters exceeds 10, the silhouette score decreases further to approximately 0.2, as presented in Figure 7.

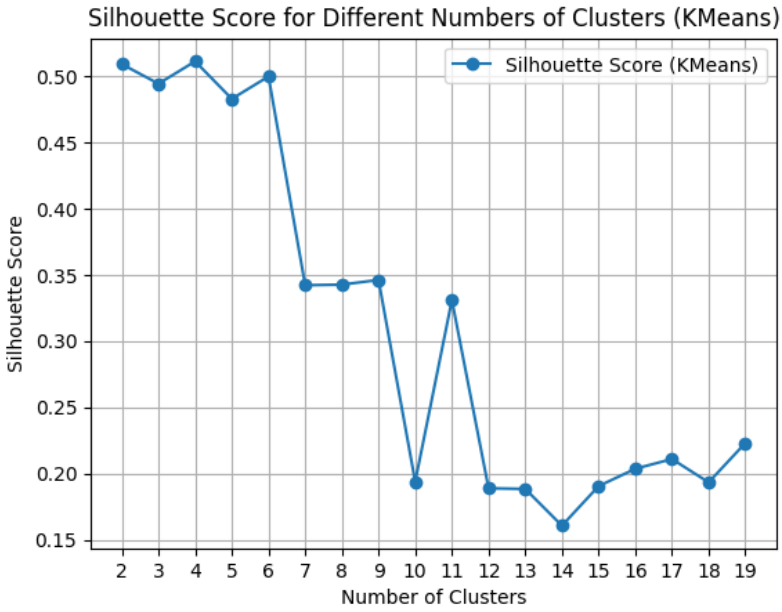

Figure 6

Change of sillouethe score with an increasing number of clusters

This decline indicates diminishing cluster quality, with clusters becoming even less distinct and exhibiting reduced cohesion and separation. Overall, these results suggest that the optimal number of clusters lies within the range of 2 to 6, where clusters are well-separated and exhibit high cohesion, while increasing the number of clusters beyond this range may result in less effective clustering with reduced cluster quality.

When the Davies-Bouldin Index ranges between 1.5 and 1.2 for cluster numbers up to 5, it suggests that the clustering solution may not be optimal, with clusters exhibiting moderate to high overlap or insufficient separation. However, as the number of clusters increases to 5 and 6, the Davies-Bouldin Index drops to 1 and 0.9, respectively, indicating improved clustering quality. This decrease suggests

that the clustering solution with six clusters is more favorable, with clusters showing greater distinctiveness and separation. On the other hand, for a larger number of clusters beyond 6, the Davies-Bouldin Index starts to rise again, reaching values between 1.2 and 1.55, as presented in Figure 8.

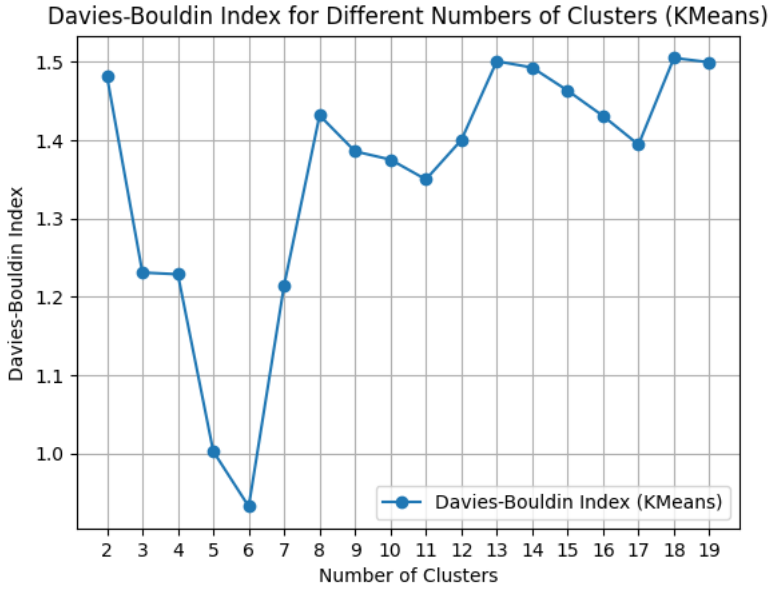

Figure 7
Change of Davies-Bouldin Index with an increasing number of clusters

This increase implies a deterioration in clustering quality as the number of clusters increases further, possibly due to increased overlap or fragmentation of clusters. Davies-Bouldin Index provides insights into the effectiveness of the clustering solution, with lower values indicating better cluster separation and distinctiveness, particularly evident when the index decreases to 0.9 for six clusters before rising again with additional clusters.

When the Calinski-Harabasz Index ranges between 400 and 500 for 2 to 4 clusters, it indicates that the clustering solution exhibits high cohesion within clusters and significant separation between clusters. As the number of clusters increases to 5, the Calinski-Harabasz Index rises to 575, suggesting even stronger clustering performance with improved cluster separation. Similarly, for 6 clusters, the index remains relatively high at 548, indicating continued effectiveness in clustering. However, with further increases in the number of clusters, the index gradually declines with each additional cluster, eventually reaching 300 for 19 clusters, as presented in Figure 9.

This decline implies diminishing clustering quality as the number of clusters increases beyond a certain point, likely due to increased fragmentation or overlap between clusters. Overall, the Calinski-Harabasz Index provides insights into the

optimal number of clusters, with higher values indicating better cluster separation and cohesion, particularly evident in the range of 2 to 6 clusters before declining with additional clusters.

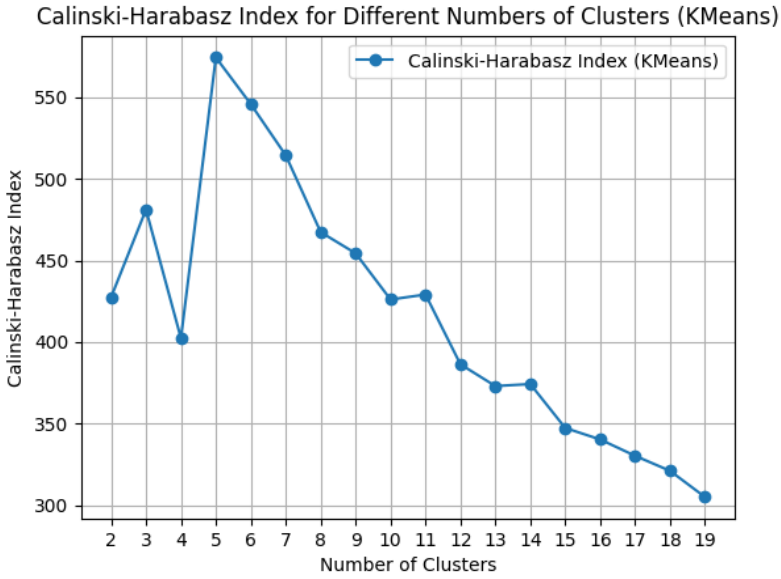

Figure 8
Change of Calinski-Harabasz Index with an increasing number of clusters

Conclusions

In conclusion, the analysis conducted in this research sheds light on the effectiveness of the k-means clustering method for the clustering of Intracellular Calcium Signal Transients, a crucial aspect in the study of neurological diseases. Through the evaluation of clustering performance using silhouette score, Davies-Bouldin Index, and Calinski-Harabasz Index, key insights have been gained regarding the optimal number of clusters and the quality of the clustering solution. The findings suggest that for the dataset under examination, the optimal number of clusters lies within the range of 2 to 6, where clusters exhibit high cohesion and separation. Beyond this range, the clustering quality diminishes, with clusters becoming less distinct and showing increased overlap. These results provide valuable guidance for researchers and practitioners in the field of neuroscience, enabling more effective analysis and interpretation of Intracellular Calcium Signal Transients data. Given the substantial burden imposed by neurological diseases on global health, the utilization of robust clustering techniques such as k-means holds promise for advancing our understanding of disease mechanisms and facilitating the development of targeted interventions to improve patient outcomes. Further research and refinement of clustering methodologies are warranted to enhance the accuracy and reliability of clustering analyses in the context of neurological research. The insights from our study extend beyond neuroscience to various fields.

In neuroimmunology, understanding calcium signaling dynamics in glial cells informs research on neuroinflammatory processes. Pharmacologically, identifying distinct clusters of calcium transients aids in targeted therapeutic interventions for CNS disorders. Our findings also contribute to computational biology by offering a model for analyzing similar cellular systems. Moreover, in systems neuroscience, elucidating calcium dynamics enhances understanding of brain function. Overall, our research has broad implications across disciplines, fostering interdisciplinary collaborations and innovative approaches to address biomedical challenges beyond the scope of neuroscience. Furthermore, the insights gained from our study have practical implications in clinical settings. By elucidating the clustering patterns of calcium signaling dynamics, our research may facilitate the development of biomarkers for diagnosing and monitoring neurological disorders. Additionally, the identification of novel therapeutic targets based on these clustering patterns holds promise for the development of more effective treatments, ultimately improving patient outcomes in real-world healthcare settings.

References

[1] W. M. Carroll, “The global burden of neurological disorders,” The Lancet Neurology, Vol. 18, No. 5, pp. 418-419, 2019

[2] C. C. Gross, A. Schulte-Mecklenbeck, L. Madireddy, M. Pawlitzki, C. Strippel, S. Rauber, J. Kr ¨ amer, L. Rolfes, T. Ruck, C. Beuker et al., “Classification of neurological diseases using multi-dimensional csf analysis,” Brain, Vol. 144, No. 9, pp. 2625-2634, 2021

[3] Milošević M, Stenovec M, Kreft M, Petrušić V, Stević Z, Trkov S, et al. Immunoglobulins G from patients with sporadic amyotrophic lateral sclerosis affects cytosolic Ca2+ homeostasis in cultured rat astrocytes. Cell Calcium. 2013;54(1):17-25, DOI: 10.1016/j.ceca.2013.03.005

[4] Alexej Verkhratsky, Richard K Orkand, and Helmut Kettenmann. Glial calcium: homeostasis and signaling function. Physiological reviews, 78(1):99-141, 1998

[5] Narges Bazargani and David Altwell. Astrocyte calcium signaling: the third wave. Nature neuroscience, 19(2): 182-189, 2016

[6] Murali Prakriya. Calcium and cell function. The Journal of physiology, 598(9): 1647, 2020

[7] Annamaria Lia, Vanessa Jorge Henriques, Micaela Zonta, Angela Chiavegato, Giorgio Carmignoto, Marta Gomez-Gonzalo, and Gabriele Losi. Calcium signals in astrocyte microdomains, a decade of great advances. Frontiers in cellular neuroscience, 15:673433, 2021

[8] Ana Covelo, Anais Badoual, and Audrey Denizot. Advocating for interdisciplinary collaborations to unravel the astrocyte” calcium code”. Advocating for interdisciplinary collaborations to unravel the astrocyte “Calcium Code.", 2022

[9] Yinxue Wang, Guilai Shi, David J Miller, Yizhi Wang, Congchao Wang, Gerard Broussard, Yue Wang, Lin Tian, and Guoqiang Yu. Automated functional analysis of astrocytes from chronic time-lapse calcium imaging data. Frontiers- in neuroinformatics, 11:48, 2017

[10] Audrey Denizot, Misa Arizono, U Valentin Nagerl, Hedi Soula, and Hugues Berry. Simulation of calcium signaling in fine astrocytic processes: Effect of spatial properties on spontaneous activity. PLoS computational biology, 15(8):el(X)6795, 2019, DOI: 10.1371/journal.pcbi.1006795

[11] Wolfgang J Nett, Scott H Oloff. and Ken D McCarthy. Hippocampal astrocytes in situ exhibit calcium oscillations that occur independent of neuronal activity. Journal of neurophysiology, 87(l):528-537, 2002, DOI: 10.1152/jn.00268.2001

[12] A. Korenić, I. Lorencin, N. Tanković, P. Andjus, and D. Frank, "Review of the machine learning based methods applicable for analysis of intracellular calcium transients imaging," in IEEE 21st International Symposium on Intelligent Systems and Informatics (SISY 2023), 2023, pp. 35-41

[13] K. P. Sinaga and M.-S. Yang, "Unsupervised k-means clustering algorithm," IEEE access, Vol. 8, pp. 80 716-80 727, 2020

[14] M. Ahmed, R. Seraj, and S. M. S. Islam, "The k-means algorithm: A comprehensive survey and performance evaluation," Electronics, Vol. 9, b No. 8, p. 1295, 2020

[15] A. Javed, B. S. Lee, and D. M. Rizzo, "A benchmark study on time series clustering," Machine Learning with Applications, Vol. 1, p. 100001, 2020

[16] G. Li, B. Choi, J. Xu, S. S. Bhowmick, K.-P. Chun, and G. L.-H. Wong, "Efficient shapelet discovery for time series classification," IEEE transactions on knowledge and data engineering, Vol. 34, No. 3, pp. 1149-1163, 2020

[17] K. R. Shahapure and C. Nicholas, "Cluster quality analysis using silhouette score," in 2020 IEEE 7th international conference on data science and advanced analytics (DSAA). IEEE, 2020, pp. 747-748

[18] M. Shutaywi and N. N. Kachouie, "Silhouette analysis for performance evaluation in machine learning with applications to clustering," Entropy, Vol. 23, No. 6, p. 759, 2021

[19] Mughnyanti, M., Efendi, S., & Zarlis, M. (2020) Analysis of determining centroid clustering x-means algorithm with davies-bouldin index evaluation. In IOP Conference Series: Materials Science and Engineering (Vol. 725, No. 1, p. 012128) IOP Publishing

[20] da Costa, J. P., & Garcia, A. (2021) New confinement index and new perspective for comparing countries-COVID-19. Computer Methods and Programs in Biomedicine, 210, 106346