World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:18, No:2, 2024

Identifying Autism Spectrum Disorder Using Optimization-Based Clustering

Sharifah Mousli, Sona Taheri, Jiayuan He

Abstract—Autism spectrum disorder (ASD) is a complex developmental condition involving persistent difficulties with social communication, restricted interests, and repetitive behavior. The challenges associated with ASD can interfere with an affected individual's ability to function in social, academic, and employment settings. Although there is no effective medication known to treat ASD, to our best knowledge, early intervention can significantly improve an affected individual's overall development. Hence, an accurate diagnosis of ASD at an early phase is essential. The use of machine learning approaches improves and speeds up the diagnosis of ASD. In this paper, we focus on the application of unsupervised clustering methods in ASD, as a large volume of ASD data generated through hospitals, therapy centers, and mobile applications has no pre-existing labels. We conduct a comparative analysis using seven clustering approaches, such as K-means, agglomerative hierarchical, model-based, fuzzy-C-means, affinity propagation, self organizing maps, linear vector quantisation – as well as the recently developed optimization-based clustering (COMSEP-Clust) approach. We evaluate the performances of the clustering methods extensively on real-world ASD datasets encompassing different age groups: toddlers, children, adolescents, and adults. Our experimental results suggest that the COMSEP-Clust approach outperforms the other seven methods in recognizing ASD with well-separated clusters.

Keywords—Autism spectrum disorder, clustering, optimization, unsupervised machine learning.

I. Introduction

ASD is a neurodevelopment disorder that affects brain functioning and causes difficulties in social communication and interaction. According to the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5), the diagnosis of autism is evident in two domains: a deficit in social communication and restricted repetitive behavior [1]. In clinical practice, different instruments, such as Autism Diagnostic Observation Schedule (ADOS), Autism Diagnostic Interview-Revised (ADI-R), and Childhood Autism Rating Scale (CARS), are used to diagnose ASD [2].

ASD is a heterogeneous disorder with diverse symptoms; any two autistic people may have drastically different symptoms due to the broad spectrum of ASD. As a result, among all child-onset psychiatric disorders, ASD is one of the most challenging disorders in diagnosis, as it can be present with different criteria, diversity in the intensity of autism, etiology, and response to therapy [3]. There exists extensive literature studying the detection of ASD at an early age [4], [5], as this will help a patient get early intervention services (e.g., speech, occupational therapy, and applied behavior analysis). The earlier a child receives intervention services, the better the outcome for the child's improvement in the future.

S. Mousli is a PhD candidate at the School of Sciences, RMIT University, Melbourne, Victoria, Australia (e-mail: s3635335@student.rmit.edu.au).

S. Taheri is a Lecturer at the School of Sciences, RMIT University, Melbourne, Victoria, Australia.

J. He is a Lecturer at the School of Computing Technologies, RMIT University, Melbourne, Victoria, Australia.

It has been acknowledged that machine learning techniques hold promise for automatic, faster, and potentially more accurate detection of ASD due to its data-driven nature and remarkable success in developing predictive models in various domains, including computer vision and natural language processing.

To date, various classification algorithms have been proposed and applied in identifying ASD. For instance, a comprehensive review of applying supervised learning models in ASD literature is provided in [6], [7]. In [6], the authors reviewed 35 ASD studies with a comparative analysis of several classification algorithms for ASD diagnosis, showing that Support Vector Machine (SVM) and Alternating Decision Tree (ADTree) produced the most accurate diagnoses of ASD. In [7], the recent works of ASD based on classification algorithms and feature selection techniques were described. This study showed that the issue with ASD diagnosis using classification methods can be addressed via efficient feature selection in data pre-processing.

Despite the prevalence of supervised learning models, using them for ASD diagnosis requires a good volume of labeled datasets, which is usually expensive to obtain in the medical domain. Acquiring high-quality datasets for training is usually non-trivial, in particular considering the extensive domain knowledge and time-consuming process of ASD diagnosis. Consequently, there is a trend of using unsupervised machine learning models for ASD detection, which do not require annotated data. These models seek to identify underlying patterns in data without relying on known values or labels. A series of studies present the application of unsupervised learning methods in ASD research. In particular, various clustering algorithms have been proposed to identify ASD patient groups with a focus on areas of behavioral symptoms and sensory profile data. However, these existing approaches only consider the compactness of clusters during training, while ignoring the inter-cluster distances. Hence, we apply the recently proposed optimization-based clustering approach to ASD detection, which has a two-term objective function guiding the model to not only minimize the intra-cluster variations but also maximize the inter-cluster distances.

The rest of the paper is structured as follows. In Section II, we provide a brief review of unsupervised learning in ASD, followed by our motivations to further study this research area. The datasets used in our experiments are discussed in Section III. Section IV presents brief definitions of the used clustering approach. Then a comparative analysis was done between these clustering approaches using performance measures to conclude the outperform approach. Therefore, in this section, we also report the numerical results. Finally, Section V concludes the paper.

II. Unsupervised Learning for ASD Diagnosis

In this section, we review some existing studies that utilized unsupervised learning for ASD diagnosis. In the literature, various types of unsupervised learning techniques have been investigated, including the classical K-means clustering algorithm, hierarchical clustering, model-based clustering, affinity propagation, self organizing maps, and linear vector quantisation. We review these studies based on the underlying clustering techniques that they utilized.

The K-means clustering is one of the most widely used method in ASD detection, see for example [8]–[12]. The analysis of challenging behavior in ASD children using machine learning was first studied by [8]. Using K-means this research was able to identify the dominant challenging behaviour for female and male clusters. In addition, they were able to identify some potential differences between male and female in their challenging behaviour profiles. The authors in [10] considered language abilities besides the different behavioral symptoms.

Hierarchical clustering is another approach in ASD studies; see, for example, a recent review study [13]. Although this approach offers flexibility in selecting the number of groups to work with, it performs well when the number of features was reduced before the approach was applied. For instance, the authors in [14] utilized first the principle component analysis (PCA) to reduce the dimension of the data. Then, using the methods of hierarchical clustering, the data were grouped into three latent groups.

Model-based clustering is another commonly used clustering technique for ASD [13]. Sensory issue groups were identified using model-based clustering, and the results suggested three sensory types: over-sensitive, under-sensitive, and sensory seeking [15]. Besides the sensory issues, the latent pattern of the behavior issues was detected in autistic children resulting in 16 subgroups [16]. These groups then were combined into two behaviors with deficit profiles using hierarchical agglomerative clustering; each of the two groups had different levels of severity.

Affinity propagation clustering was applied to ASD; for instance, in [17]. The aim was to identify groups for vitamin B6 responsiveness treatment with two types (possible responders, fewer responders) by analyzing selected phenotype variables such as hypersensitivity to sound. The results from this study found five groups with one potential group for a possible respondent to vitamin B and the other four clusters with the low respondent groups to vitamin B6.

Self organizing maps were used prior to classification to create unbiased class labels without using scoring functions of the medical ASD screening [18]. These new labels, combined with clinician decisions, can help reduce biased decisions in ASD screening when using classification algorithms. Thus, it was achieved in two phases: first, by applying a SOM to identify new patterns and refine the dataset, and second, by using algorithms to create ASD classification models based on the refined dataset. Another study [19] utilized SOM to detect the latent pattern of response to a task for children who are typically developed (TD) and children who have ASD.

Linear vector quantisation (LVQ) was applied in ASD detection studies, for example, in [20] where an artificial intelligence called ”Medical AI” was introduced for screening and grading childhood autism. It used a possibilistic LVQ (Po-LVQ) approach and categorized autism into four possible grades (no autism, mild, moderate, and severe). The results of this method are compared with those of a standard LVQ and other existing models that have been previously used.

There also exist research works that perform comparative studies on different unsupervised learning techniques. In [21], six clustering methods, namely K-means, agglomerative hierarchical, model-based, partitioning around medoids, divisive hierarchical, and self-organizing tree algorithm, were applied in ASD detection. Further, the algorithms K-means, learning vector quantisation, fuzzy C-means, and self organizing maps were applied in the childhood autistic rating scale in [12]. In [11], the K-means and X-means clustering algorithms were utilized to the child ASDTest data.

In this paper, we focus on using unsupervised learning in ASD diagnosis due to the following main reasons:

1) The lack of available data or hard-to-access ASD data due to privacy issues, as these data can be identifiable and need ethical approval for research purposes. Thus, utilizing unidentifiable data which might be easier to find would give a better understanding of ASD. For example, some studies used the medical codes (ICD or PheWas) to train the unsupervised models, which create non-existed groups of the co-occurring condition with ASD [22]–[24];
2) The heterogeneity inheriting within ASD can be seen in a huge diversity from different etiological mechanisms, developmental trajectories, sex/gender, clinical comorbidities, cognitive and behavioral characteristics, the progression of language skills, and so forth. The list could be extended. As such, clustering is naturally considered to be an appropriate method for creating homogeneous sub-groups of ASD. For example, the study [8] aimed to develop personalized treatment plans or diagnoses for ASD patients, and they utilized clustering to divide the data into several meaningful sub-groups in order to receive the appropriate intervention;
3) Dealing with mixed data types such as brain image, genetics, clinical assessment, and behaviors would provide more homogeneous ASD groups in line with the DSM-5 [25]. In the study [26], new subgroups of different neurodevelopment disorders were identified, where they provide similar brain-behavior patterns across different disorders more than the groups within a single disorder.

The objectives of this study is to explore whether the information collected from ASDTest could reveal a potential clustering structure for ASD identification and to have a better insight into discovering interesting patterns through unsupervised learning. We apply the optimization-based clustering approach introduced recently [27] to achieve our goals.

Considering the reasons mentioned above, we explore different clustering methods for identifying ASD. Datasets for two screening tools called Autism Spectrum Quotient (AQ) and Quantitative Checklist for Autism in Toddlers (Q-CHAT) are used in our investigation. The clustering methods utilized for identifying ASD in [11] and [12] as well as the optimization-based clustering approach introduced in [27] to accomplish our objectives.

III. Data Description and Preparation

In this section, we first provide the description of ASD dataset. Then, we discuss the processing of the data to be used in our numerical experiments. The data are available via UCI Machine Learning Repository [28]. These data were proposed by Thabtah [29] and further discussed and utilized, for example in [18], [30], [31].

There are four age categories in these data: toddlers (up to 36 months), children aged (4-11 years), adolescents (12-16 years) and adults over 18 years. Using the behavioural screening for autism tool AQ [32], the last three age categories were developed. The first age category proposed utilising the Quantitative Checklist for Autism in Toddlers (Q-CHAT) [33].

The behavioral variables (Q1 to Q10) have been derived from ASD behavioral screening tools AQ and Q-CHAT. The respondent's answers are represented by coding them as '1' for "slightly agree" or "definitely agree", and '0' for "slightly disagree" or "definitely disagree" during the screening process using the screening systems. The latter features are collected after answering Q1 to Q10. More specifically, a respondent is given a score ranging from 0 to 10 based on the responses for these questions. Finally, the class is labeled as either YES or NO, depending on the score, where YES indicates "Autistic" and NO indicates "Non-autistic". The screening test classifies the response as YES if the score is above three for toddlers or above six for all other age categories. In addition to ten features from screening tools (Q1 to Q10), screening score, and Class (Output), there are six individuals characteristics and three questions about using the ASDTest app [31]. For information about the attributes/features, see Table I.

The ASD datasets contain 21 attributes, except for Toddler dataset, which has only 18. The number of samples in the produced datasets for Toddlers, Children, Adolescents, and Adults is 1054, 292, 104, and 704, respectively.

Data prepossessing was done to enhance the quality of clustering performance. First, the categorical features were transformed into numerical values because most clustering algorithms only work with numerical data. The imputation method by mean was used to fill in missing values in the variables ethnicity, born with jaundice, and who completed the test. Next, to evaluate the dataset more efficiently, Min-max normalization was performed to standardise data between 0 and 1. All features listed in Table I are included in the analysis, except for the screening method type, which presents the age group of the person taking the autism screening test, whether they are a toddler, child, adolescent, or adult. Thus, the numerical experiments were run separately for each category of data.

IV. Experiments and Results

In this section, we present and discuss the results of our experiments on datasets described in Section III, utilizing seven clustering algorithms applied in ASD, as well as COMSEP-Clust algorithm.

A. Baseline Algorithms

We consider the following algorithms in our experiments:

- *K-means* divides the dataset into K clusters. The algorithm begins with a random set of K cluster centers (centroids), where each data point is assigned to the nearest centroid, and data points belong to the same centroid form a cluster. The distance between a data point and a centroid is computed using Euclidean distance. Then the centroid of each cluster is updated to the mean position of the data points within the cluster. Given the new set of centroids, the algorithm then re-assigns each data point to form new clusters. These steps are repeated until cluster membership does not change [34];
- *Agglomerative hierarchical clustering* is one of the main types of hierarchical clustering, which is particularly effective in identifying small clusters. It works from the bottom-up method, that is starting with each object or leaf as its separate group. Then, at each step, the two most similar groups are merged into a larger group. This process continues until all points belong to a single large group (representing as a root). The outcome is a tree structure, called a dendrogram, that shows how the groups are related to each other [35];
- *Model-based* clustering assumes that the data points fit with a single probability distribution or mixture of probability distributions, and each distribution indicates a distinct cluster within the data. In this method, if the data or an expert provides insight into the shape of the distribution, this information can be integrated into the MB clustering [36];
- *Fuzzy-C-means* is an extension of the K-means algorithm. It works similarly to K-means as it divides input data into C fuzzy groups and identifies a cluster center in each group that minimizes the dissimilarity metric. In addition, it uses fuzzy partitioning, which means that a data point can belong to multiple groups, and the degree of belongingness is expressed as membership grades between 0 and 1 [37];
- *Affinity propagation* is based on the idea of message passing between data points to perform clustering. It generates a representative sample, also known as an exemplar, for each cluster that it identifies. This method detects clusters of points by taking the similarity measure between pairs of data points as input while considering all

TABLE I
FEATURE DESCRIPTION OF ASD DATASET

Features	Description
Age	Age in years
Gender	Female or Male
Ethnicity	List of ethnicity
Born with jaundice	Whether the case have a jaundice from birth
Family member with PDD	History of family member has a Pervasive developmental disorders
Who is completing the test (User)	Parent, clinical, self , etc.
Country of residence	List of countries
Used the screening app before	If the user has used ASDTest screening app before
Screening method type	Type of screening chosen based on age category
Question 1 (A1)	A question from the screening tool
Question 2 (A2)	A question from the screening tool
Question 3 (A3)	A question from the screening tool
Question 4 (A4)	A question from the screening tool
Question 5 (A5)	A question from the screening tool
Question 6 (A6)	A question from the screening tool
Question 7 (A7)	A question from the screening tool
Question 8 (A8)	A question from the screening tool
Question 9 (A9)	A question from the screening tool
Question 10 (A10)	A question from the screening tool
Screening score	The final score determined using the screening method's scoring algorithm.
Class (Output)	The case was diagnosed with ASD = 1 or not = 0

data points as potential exemplars. It runs by exchanging messages between data points until an exemplar emerges and clusters are formed [38];

- *Self Organizing maps* is a form of Artificial Neural Network (ANN) where the data are trained via unsupervised learning. It is known that the neural network responds to multiple output units while training data. When the learning is based only on input data, then the network is forced to select a specific output unit for decision-making. Thus, an additional structure can be added; this process is called competition. The SOM is a competition-based neural network to cluster the data. Using the dissimilarity measure, we can determine how different the data are and the unit with the smallest difference is considered the winner [39];
- *Linear vector quantisation* is a standard statistical clustering technique that trains its network using a competitive learning algorithm similar to self-organizing maps. LVQ seeks to create representative prototypes or codebook vectors that capture the characteristics of the input data. This involves assigning input vectors to a set of predefined codebook vectors. The LVQ network consists of two layers: competitive and linear. The competitive layer learns to classify inputs by measuring the similarity between input vectors and creating subclasses. The linear layer then converts these subclasses into user-defined target classes [39];
- *COMSEP-Clust* is an unsupervised algorithm that constructs clusters incrementally. The COMSEP-Clust algorithm consists of two phases. In the first phase, the clustering model with the objective function, containing only the compactness term, is considered. In this phase, starting cluster centers are generated using an incremental approach, and data points located away from the cluster centers. In the second phase, starting from the best solution obtained in the first phase, the clustering model with the objective function, containing both compactness and separability terms, is applied. The aim of this phase is to improve the separability of clusters obtained in the first phase. We refer to [27] for more details about the COMSEP-Clust algorithm.

The first seven algorithms are implemented in R, using packages "Cluster" [40], "stats" [41], "MClust" [42], "PPClust"[43],"APCluster" [44], "kohonen" [45] and "LVQ" [46]. We use the default values for the parameters of these algorithms, as recommended in their respective references. The COMSEP-Clust algorithm is implemented using Fortran 77 and compiled with the GNU Fortran compiler. The source code of this algorithm is available at GitHub [47].

We set the number of clusters to be $k = 2$; that is, partitioning data into two clusters in order to differentiate between Autistic and non-Autistic. The computational experiments are carried out on a PC with Dual-core Intel Core(i5)CPU 1.8 GHz and RAM 8 GB.

B. Performance Measures

To evaluate the performance of clustering algorithms on the ASD datasets, we apply the measures accuracy and F-measure. Note that each data point in a dataset and its corresponding

TABLE II
RESULTS ON TODDLERS DATASET

Algorithm	Accuracy	F-measure	Silhouettes
K-means	0.7818	0.7392	0.6418
Agglomerative hierarchical	0.7287	0.6951	0.5947
Model based	0.8491	0.8039	0.6321
Fuzzy-C-means	0.7818	0.7392	0.6418
Affinity propagation	0.8696	0.8150	0.6325
Self-organizing maps	0.8624	0.8181	0.6315
Learning vector quantization	0.8994	0.8602	0.7228
COMSEP-Clust	0.9202	0.8791	1.0000

TABLE III
RESULTS ON CHILDREN DATASET

Algorithm	Accuracy	F-measure	Silhouettes
K-means	0.9041	0.9085	0.7262
Agglomerative hierarchical	0.7226	0.7652	0.6009
Model based	0.9760	0.9749	0.9286
Fuzzy-C-means	0.9315	0.9329	0.8014
Affinity propagation	0.8562	0.8627	0.5949
Self organizing maps	0.9110	0.9156	0.7094
Learning vector quantisation	0.8938	0.9003	0.6993
COMSEP-Clust	0.9808	0.9801	0.9999

TABLE IV
RESULTS ON ADOLESCENTS DATASET

Algorithm	Accuracy	F-measure	Silhouettes
K-means	0.8942	0.9185	0.6954
Agglomerative hierarchical	0.8365	0.8794	0.5414
Model based	0.8750	0.9065	0.6450
Fuzzy-C-means	0.8750	0.8943	0.6420
Affinity propagation	0.8942	0.9185	0.6954
Self organizing maps	0.9231	0.8947	0.7754
Learning vector quantisation	0.8173	0.8690	0.6000
COMSEP-Clust	0.9611	0.9707	0.9990

TABLE V
RESULTS ON ADULTS DATASET

Algorithm	Accuracy	F-measure	Silhouettes
K-means	0.9190	0.8690	0.7756
Agglomerative hierarchical	0.9318	0.8681	0.8004
Model based	0.8849	0.8235	0.6890
Fuzzy-C-means	0.8267	0.7560	0.5478
Affinity propagation	0.8935	0.8344	0.7104
Self organizing maps	0.9290	0.9491	0.8011
Learning vector quantisation	0.9172	0.8602	0.8350
COMSEP-Clust	0.9559	0.9200	1.0000

predictive model can be classified into one of the following four categories:

- *True Positive (TP)*: an individual who has ASD and is correctly identified as having ASD;
- *True Negative (TN)*: an individual who does not have ASD and is correctly identified as not having ASD;
- *False Positive (FP)*: an individual who does not have ASD but is incorrectly identified as having ASD;
- *False Negative (FN)*: an individual who has ASD but is incorrectly identified as not having ASD.

We utilize these categories to calculate accuracy and F-measure. In addition, we utilize the widely used clustering index "Silhouette scores" to evaluate the quality of clustering solutions obtained by different algorithms [48].

C. Results

The performances of all compared algorithms on the four datasets are presented in Tables II-V, which show the results on Toddlers, Children, Adolescents, and Adults, respectively. In these tables, we summarize the performances of all algorithms in terms of accuracy, F-measure, and the Silhouette scores.

Table II illustrates that COMSEP-Clust obtains the highest performance in terms of all three metrics (Accuracy, F-measure, and Silhouettes), demonstrating the best effectiveness of this algorithm for this task. Specifically, COMSEP-Clust achieves a perfect Silhouettes score, 1.0, which is a solid improvement over other clustering methods with a Silhouettes score of below or around 0.7. COMSEP-Clust also obtains consistent improvement in terms of the other two metrics, Accuracy and F-measure, compared with all other algorithms.

We observe a similar phenomenon in the other three datasets: Children, Adolescents, and Adults. COMSEP-Clust consistently outperforms the other algorithms in all metrics, as shown in Tables III-V. All Silhouette scores have positive values; in particular, COMSEP-Clust achieves consistently the highest Silhouettes score, which is equal to or close to one, while other algorithms obtained values above 0.55. This is due to the fact that adding an extra term to the objective function of the clustering function enables this algorithm to produce well-separated clusters.

In terms of accuracy and F-measure, COMSEP-Clust shows competitive performance by achieving perfect accuracy (100%) on at least one of the ASD datasets. This means most of the data is correctly classified by COMSEP-Clust with less error. In addition, it is noted that SOM also achieves higher Silhouette scores following to COMSEP-Clust. Silhouette scores presented in Tables II-V have demonstrated the high-quality well-separated clusters formed by the COMSEP-Clust algorithm. Next, we present visualized cluster distributions learned by different algorithms to further illustrate this.

From Fig. 1, we see that the clusters obtained by COMSEP-Clust are well-separated than the other clusters. The clusters obtained by K-means, Model-based, Fuzzy-C-means, Affinity propagation, Self organizing maps, and Linear vector quantisation all contain a strong overlap between the two learned clusters, with some member data points of the two learned clusters are closely located. However, the COMSEP-Clust achieves perfect separation of data points, dividing them into two distinct clusters with a large inter-cluster distance. This confirms the superiority of the algorithm in finding the true clusters accurately, and thereby, well classify all data points. We can observe the same improvements on the clusters' separation for Children, Adolescents and Adults cluster in Figs. 2-4. Furthermore, the Adults and Toddlers data have larger sample size than other categories, which may suggest that our proposed algorithms work more efficiently with larger data sets.

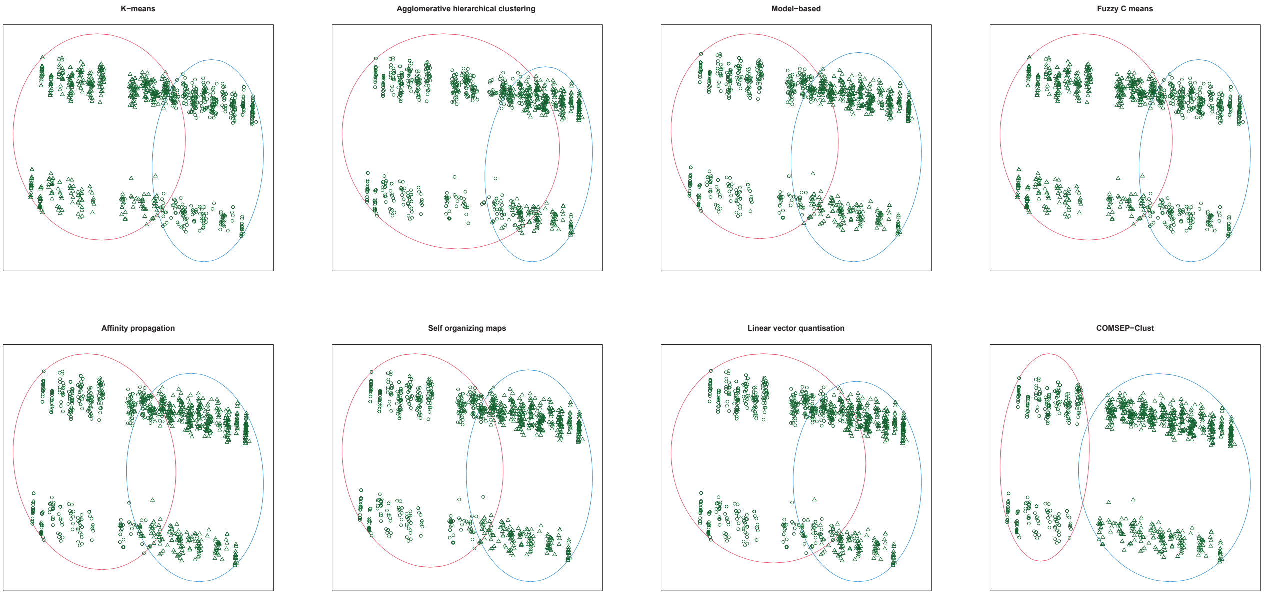

Fig. 1 Two clusters in Toddlers dataset obtained using algorithms K-means, Agglomerative hierarchical, Model-based, Fuzzy-C-means, Affinity propagation, Self organizing maps, Linear vector quantisation, and COMSEP-Clust

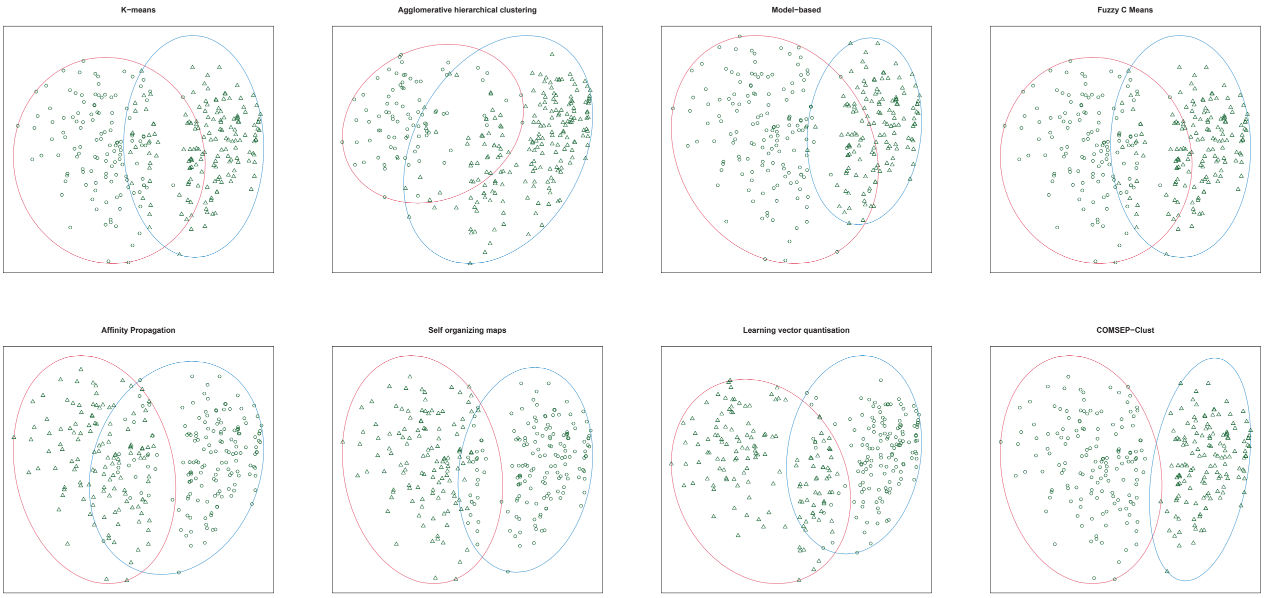

Fig. 2 Two clusters in Children dataset obtained using algorithms K-means, Agglomerative hierarchical, Model-based,Fuzzy-C-means, Affinity propagation, Self organizing maps, Linear vector quantisation, and COMSEP-Clust

V. Conclusions

In this paper, we explored and analyzed the application of unsupervised learning approaches in identifying ASD. We studied the application of the recently developed optimization-based clustering approach to ASD detection and compared it with seven traditional unsupervised learning techniques. We evaluated the performances of these clustering algorithms using four ASD data sets, with results showing that the optimization-based clustering approach strongly and consistently outperforms the other clustering methods concerning all metrics. In particular, the optimization-based clustering approach demonstrates robust capability in terms of producing well-separated cluster distributions, supported by its high Silhouette score and the visualization of leaned cluster distributions. The ability to learn high-quality clusters of this method has led to stronger results in ASD detection. Based on the results from this study, we plan to extend this work and investigate the usage of unsupervised learning techniques for detecting the severity of ASD in the future. We believe that this

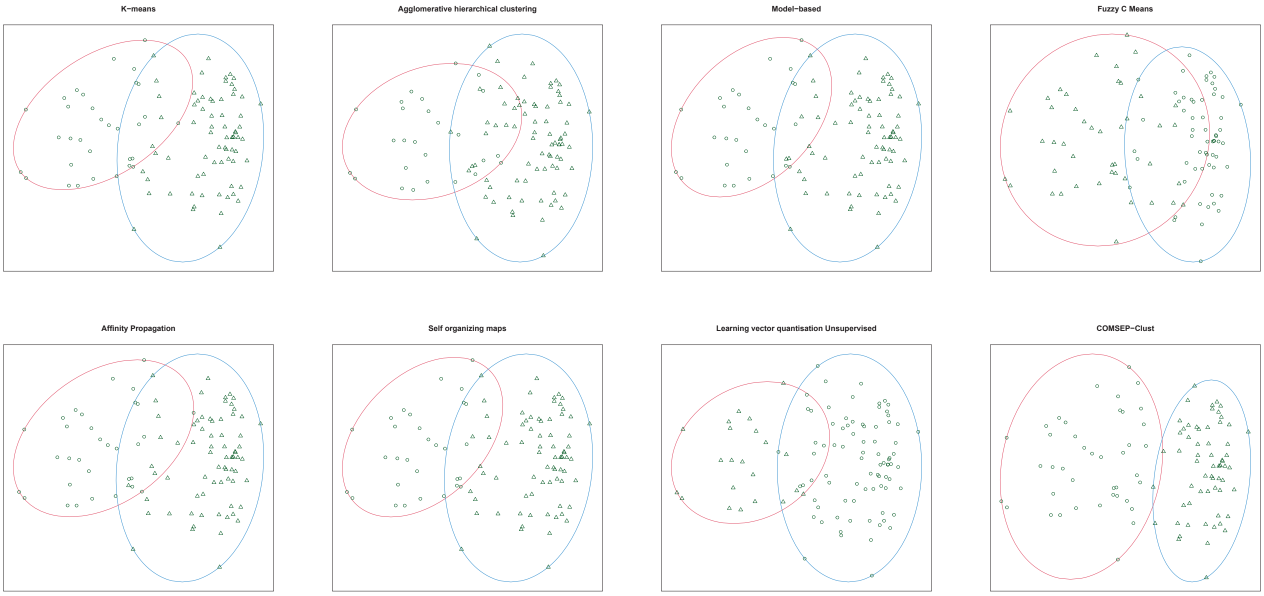

Fig. 3 Two clusters in Adolescents dataset obtained using algorithms K-means, Agglomerative hierarchical, Model-based, Fuzzy-C-means, Affinity propagation, Self organizing maps, Linear vector quantisation, and COMSEP-Clust

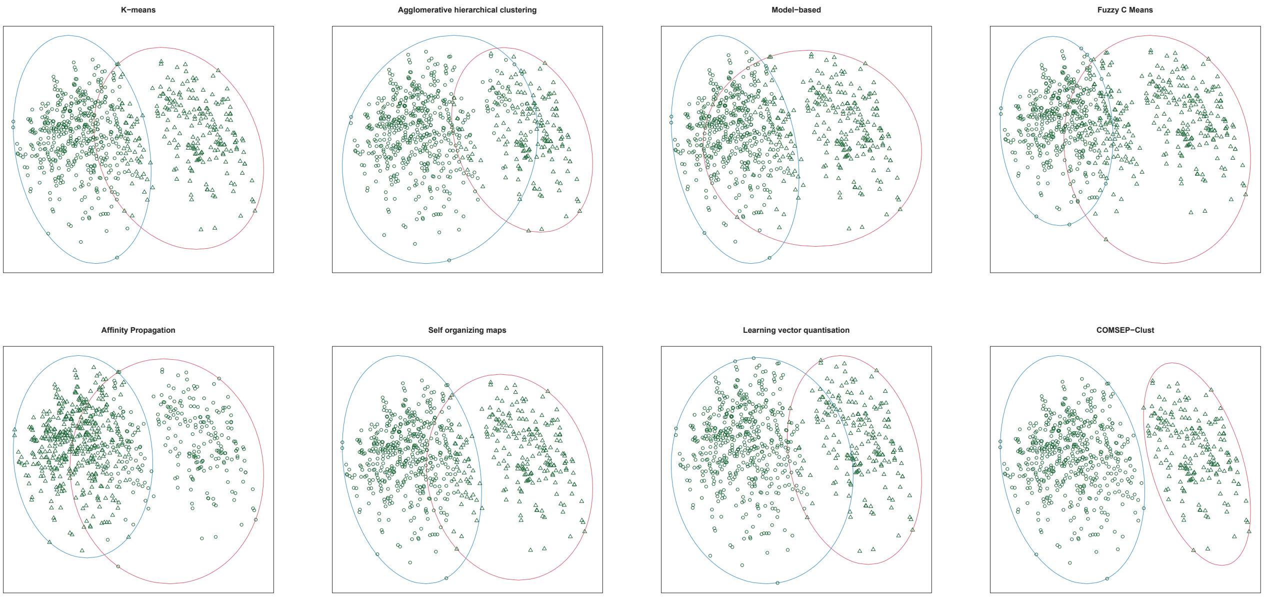

Fig. 4 Two clusters in Adults dataset obtained using algorithms K-means, Agglomerative hierarchical, Model-based, Fuzzy-C-means, Affinity propagation, Self organizing maps, Linear vector quantisation, and COMSEP-Clust

research direction has the potential to provide clinicians and patients with valuable information about the level of severity, and thus, managing ASD treatment effectively.

References

[1] A. P. Association *et al.*, "American psychiatric association: Diagnostic and statistical manual of mental disorders, arlington," 2013. [Online]. Available: https://doi.org/10.1176/appi.books.9780890425596

[2] H. S. Park, S. Y. Yi, S. A. Yoon, and S.-B. Hong, "Comparison of the autism diagnostic observation schedule and childhood autism rating scale in the diagnosis of autism spectrum disorder: a preliminary study," *Journal of the Korean Academy of Child and Adolescent Psychiatry*, vol. 29, no. 4, p. 172, 2018. [Online]. Available: https://doi.org/10.5765/jkacap.180015

[3] C. Lord, S. Risi, L. Lambrecht, E. H. Cook, B. L. Leventhal, P. C. DiLavore, A. Pickles, and M. Rutter, "The autism diagnostic observation schedule—generic: A standard measure of social and communication deficits associated with the spectrum of autism," *Journal of autism and developmental disorders*, vol. 30, pp. 205–223, 2000. [Online]. Available: https://doi.org/10.1023/A:1005592401947

[4] C. Lord and C. Corsello, "Diagnostic instruments in autistic spectrum disorders." *Handbook of Autism and Pervasive Developmental Disorders, Fourth Edition*, 2005. [Online]. Available: https://doi.org/10.1007/978-1-4419-1698-3_908

[5] F. R. Volkmar, C. Lord, A. Bailey, R. T. Schultz, and A. Klin, "Autism and pervasive developmental disorders," *Journal of child psychology and psychiatry*, vol. 45, no. 1, pp. 135–170, 2004. [Online]. Available: https://doi.org/10.1046/j.0021-9630.2003.00317.x

[6] K. K. Hyde, M. N. Novack, N. LaHaye, C. Parlett-Pelleriti, R. Anden, D. R. Dixon, and E. Linstead, "Applications of supervised machine learning in autism spectrum disorder research: a review," *Review Journal of Autism and Developmental Disorders*, vol. 6, pp. 128–146, 2019. [Online]. Available: https://doi.org/10.1007/s40489-019-00158-x

[7] M. M. Rahman, O. L. Usman, R. C. Muniyandi, S. Sahran, S. Mohamed, and R. A. Razak, "A review of machine learning methods of feature selection and classification for autism spectrum disorder," *Brain sciences*, vol. 10, no. 12, p. 949, 2020. [Online]. Available: https://doi.org/10.3390/brainsci10120949

[8] E. Stevens, A. Atchison, L. Stevens, E. Hong, D. Granpeesheh, D. Dixon, and E. Linstead, "A cluster analysis of challenging behaviors in autism spectrum disorder," in *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*. IEEE, 2017, pp. 661–666. [Online]. Available: https://doi.org/10.1109/ICMLA.2017.00-85

[9] M. Elbattah, R. Carette, G. Dequen, J.-L. Guérin, and F. Cilia, "Learning clusters in autism spectrum disorder: image-based clustering of eye-tracking scanpaths with deep autoencoder," in *2019 41st Annual international conference of the IEEE engineering in medicine and biology society (EMBC)*. IEEE, 2019, pp. 1417–1420. [Online]. Available: https://doi.org/10.1109/EMBC.2019.8856904

[10] S. Silleresi, P. Prévost, R. Zebib, F. Bonnet-Brilhault, D. Conte, and L. Tuller, "Identifying language and cognitive profiles in children with asd via a cluster analysis exploration: Implications for the new icd-11," *Autism Research*, vol. 13, no. 7, pp. 1155–1167, 2020. [Online]. Available: https://doi.org/10.1002/aur.2268

[11] A. A. Abdulrazzaq, S. S. Hamid, A. T. Al-Douri, A. Mohamad, and A. M. Ibrahim, "Early detection of autism spectrum disorders (asd) with the help of data mining tools," *BioMed Research International*, vol. 2022, 2022. [Online]. Available: https://doi.org/10.1155/2022/1201129

[12] A. Pratap and C. Kanimozhiselvi, "Predictive assessment of autism using unsupervised machine learning models," *International Journal of Advanced Intelligence Paradigms*, vol. 6, no. 2, pp. 113–121, 2014. [Online]. Available: https://doi.org/10.1504/IJAIP.2014.062174

[13] C. M. Parlett-Pelleriti, E. Stevens, D. Dixon, and E. J. Linstead, "Applications of unsupervised machine learning in autism spectrum disorder research: a review," *Review Journal of Autism and Developmental Disorders*, pp. 1–16, 2022. [Online]. Available: https://doi.org/10.1007/s40489-021-00299-y

[14] S. Zheng, K. A. Hume, H. Able, S. L. Bishop, and B. A. Boyd, "Exploring developmental and behavioral heterogeneity among preschoolers with asd: A cluster analysis on principal components," *Autism Research*, vol. 13, no. 5, pp. 796–809, 2020. [Online]. Available: https://doi.org/10.1002/aur.2263

[15] M. Uljarević, A. Lane, A. Kelly, and S. Leekam, "Sensory subtypes and anxiety in older children and adolescents with autism spectrum disorder," *Autism Research*, vol. 9, no. 10, pp. 1073–1078, 2016. [Online]. Available: https://doi.org/10.1002/aur.1602

[16] E. Stevens, D. R. Dixon, M. N. Novack, D. Granpeesheh, T. Smith, and E. Linstead, "Identification and analysis of behavioral phenotypes in autism spectrum disorder via unsupervised machine learning," *International journal of medical informatics*, vol. 129, pp. 29–36, 2019. [Online]. Available: https://doi.org/10.1016/j.ijmedinf.2019.05.006

[17] T. Obara, M. Ishikuro, G. Tamiya, M. Ueki, C. Yamanaka, S. Mizuno, M. Kikuya, H. Metoki, H. Matsubara, M. Nagai *et al.*, "Potential identification of vitamin b6 responsiveness in autism spectrum disorder utilizing phenotype variables and machine learning methods," *Scientific reports*, vol. 8, no. 1, pp. 1–7, 2018. [Online]. Available: https://doi.org/10.1038/s41598-018-33110-w

[18] F. Thabtah, R. Spencer, N. Abdelhamid, F. Kamalov, C. Wentzel, Y. Ye, and T. Dayara, "Autism screening: an unsupervised machine learning approach," *Health Information Science and Systems*, vol. 10, no. 1, p. 26, 2022. [Online]. Available: https://doi.org/10.1007/s13755-022-00191-x

[19] Á. E. Tovar, A. Rodríguez-Granados, and N. Arias-Trejo, "Atypical shape bias and categorization in autism: Evidence from children and computational simulations," *Developmental Science*, vol. 23, no. 2, p. e12885, 2020. [Online]. Available: https://doi.org/10.1111/desc.12885

[20] C. Kanimozhiselvi and A. Pratap, "Possibilistic lvq neural network-an application to childhood autism grading," *Neural Network World*, vol. 26, no. 3, p. 253, 2016. [Online]. Available: https://doi.org/10.14311/NNW.2016.26.014

[21] O. Veatch, J. Veenstra-VanderWeele, M. Potter, M. Pericak-Vance, and J. Haines, "Genetically meaningful phenotypic subgroups in autism spectrum disorders," *Genes, Brain and Behavior*, vol. 13, no. 3, pp. 276–285, 2014. [Online]. Available: https://doi.org/10.1111/gbb.12117

[22] T. Lingren, P. Chen, J. Bochenek, F. Doshi-Velez, P. Manning-Courtney, J. Bickel, L. Wildenger Welchons, J. Reinhold, N. Bing, Y. Ni *et al.*, "Electronic health record based algorithm to identify patients with autism spectrum disorder," *PloS one*, vol. 11, no. 7, p. e0159621, 2016. [Online]. Available: https://doi.org/10.1371/journal.pone.0159621

[23] T. Vargason, R. E. Frye, D. L. McGuinness, and J. Hahn, "Clustering of co-occurring conditions in autism spectrum disorder during early childhood: A retrospective analysis of medical claims data," *Autism Research*, vol. 12, no. 8, pp. 1272–1285, 2019. [Online]. Available: https://doi.org/10.1002/aur.2128

[24] A. Verma, L. Bang, J. E. Miller, Y. Zhang, M. T. M. Lee, Y. Zhang, M. Byrska-Bishop, D. J. Carey, M. D. Ritchie, S. A. Pendergrass *et al.*, "Human-disease phenotype map derived from phewas across 38,682 individuals," *The American Journal of Human Genetics*, vol. 104, no. 1, pp. 55–64, 2019. [Online]. Available: https://doi.org/10.1016/j.ajhg.2018.11.006

[25] T. Obafemi-Ajayi, D. Lam, T. N. Takahashi, S. Kanne, and D. Wunsch, "Sorting the phenotypic heterogeneity of autism spectrum disorders: A hierarchical clustering model," in *2015 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*. IEEE, 2015, pp. 1–7. [Online]. Available: https://doi.org/10.1109/CIBCB.2015.7300337

[26] A. Kushki, E. Anagnostou, C. Hammill, P. Duez, J. Brian, A. Iaboni, R. Schachar, J. Crosbie, P. Arnold, and J. P. Lerch, "Examining overlap and homogeneity in asd, adhd, and ocd: a data-driven, diagnosis-agnostic approach," *Translational Psychiatry*, vol. 9, no. 1, p. 318, 2019. [Online]. Available: https://doi.org/10.1038/s41398-019-0631-2

[27] A. M. Bagirov, N. Hoseini-Monjezi, and S. Taheri, "A novel optimization approach towards improving separability of clusters," *Computers & Operations Research*, vol. 152, pp. 106–135, 2023. [Online]. Available: https://doi.org/10.1007/978-3-030-37826-4

[28] D. Dua and C. Graff, "UCI Machine Learning Repository," 2019. [Online]. Available: http://archive.ics.uci.edu/ml

[29] F. Thabtah, "Autism spectrum disorder screening: machine learning adaptation and dsm-5 fulfillment," in *Proceedings of the 1st International Conference on Medical and health Informatics 2017*, 2017, pp. 1–6. [Online]. Available: https://doi.org/10.1145/3107514.3107515

[30] J. Alwidian, A. Elhassan, and R. Ghnemat, "Predicting autism spectrum disorder using machine learning technique," *International Journal of Recent Technology and Engineering*, vol. 8, no. 5, pp. 4139–4143, 2020. [Online]. Available: https://doi.org/10.35940/ijrte.E6016.018520

[31] F. Thabtah, "An accessible and efficient autism screening method for behavioural data and predictive analyses," *Health informatics journal*, vol. 25, no. 4, pp. 1739–1755, 2019. [Online]. Available: https://doi.org/10.1177/1460458218796636

[32] S. Baron-Cohen, S. Wheelwright, R. Skinner, J. Martin, and E. Clubley, "The autism-spectrum quotient (aq): Evidence from asperger syndrome/high-functioning autism, malesand females, scientists and mathematicians," *Journal of autism and developmental disorders*, vol. 31, pp. 5–17, 2001. [Online]. Available: https://doi.org/10.1023/a:1005653411471

[33] C. Allison, S. Baron-Cohen, S. Wheelwright, T. Charman, J. Richler, G. Pasco, and C. Brayne, "The q-chat (quantitative checklist for autism in toddlers): a normally distributed quantitative measure of autistic traits at 18–24 months of age: preliminary report," *Journal of autism and developmental disorders*, vol. 38, pp. 1414–1425, 2008.

[34] J. MacQuuen, "Some methods for classification and analysis of multivariate observation," in *Proceedings of the 5th Berkley Symposium on Mathematical Statistics and Probability*, 1967, pp. 281–297.

[35] F. Murtagh and P. Legendre, "Ward's hierarchical agglomerative clustering method: which algorithms implement ward's criterion?" *Journal of classification*, vol. 31, pp. 274–295, 2014. [Online]. Available: https://doi.org/10.1007/s00357-014-9161-z

[36] C. Fraley, A. Raftery, L. Scrucca, T. Murphy, and M. Fop, "mclust: Normal mixture modeling for model-based clustering, classification, and density estimation," *R package version*, vol. 4, no. 7, 2014.

[37] J. C. Bezdek, "Cluster validity with fuzzy sets," *Journal of Cybernetics*, 1973. [Online]. Available: https://doi.org/10.1080/01969727308546047

[38] B. J. Frey and D. Dueck, "Clustering by passing messages between data points," *science*, vol. 315, no. 5814, pp. 972–976, 2007. [Online]. Available: https://doi.org/10.1126/science.1136800

[39] T. Kohonen and T. Kohonen, "Learning vector quantization,"

Self-organizing maps, pp. 175–189, 1995. [Online]. Available: https://doi.org/10.1007/978-3-642-97610-0_6

[40] M. Maechler, "cluster: "finding groups in data": Cluster analysis extended rousseeuw et al." 2022, last accessed 21 April 2023. [Online]. Available: https://cran.r-project.org/web/packages/cluster/index.html

[41] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2023. [Online]. Available: http://www.R-project.org/

[42] C. Fraley, "mclust: Gaussian mixture modelling for model-based clustering, classification, and density estimation," 2022, last accessed 21 April 2023. [Online]. Available: https://cran.r-project.org/web/packages/mclust/index.html

[43] Z. Cebeci, "Partitioning cluster analysis using fuzzy c-means," 2017, last accessed 21 April 2023. [Online]. Available: https://cran.r-project.org/web/packages/ppclust/vignettes/fcm.html

[44] U. Bodenhofer, "Package 'apcluster'," 2022, last accessed 20 April 2023. [Online]. Available: https://cran.r-project.org/web/packages/apcluster/apcluster.pdf

[45] R. Wehrens and J. Kruisselbrink, "Package 'kohonen'," 2022, last accessed 07 June 2023. [Online]. Available: https://cran.r-project.org/web/packages/kohonen/kohonen.pdf

[46] V. Nikolaidis, "Package 'nnlib2rcpp'," 2023, last accessed 14 June 2023. [Online]. Available: https://cran.r-project.org/web/packages/nnlib2Rcpp/index.html

[47] S. Taheri, "Comsep-clust," 2021, last accessed 21 April 2023. [Online]. Available: https://github.com/SnTa2019/Clustering-via-Nonsmooth-Optimization

[48] P. J. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," *Journal of computational and applied mathematics*, vol. 20, pp. 53–65, 1987. [Online]. Available: https://doi.org/10.1016/0377-0427(87)90125-7