INTERACTIVE VISUALIZATION OF AI CLUSTERING ALGORITHMS WITH VR INTEGRATION

Anda-Teodora Cosma
Technical University of Cluj-Napoca
str. Memorandumului 28, 400114, Cluj-Napoca
andutza789@gmail.com

Dorian Gorgan
Technical University of Cluj-Napoca
str. Memorandumului 28, 400114, Cluj-Napoca
dorian.gorgan@cs.utcluj.ro

ABSTRACT
Understanding certain abstract concepts, as well as the operational mechanisms of specific algorithms, is a difficult task. Dynamic visualization of these algorithms in action is a real challenge, especially when trying to distinguish the differences and similarities between algorithms in the same category, such as clustering algorithms. This research aims to bridge this gap by developing an interactive tool that employs Virtual Reality (VR) technology and the Meta Quest 2 headset to facilitate a more enjoyable and engaging learning experience. It explores solutions that integrate a classification algorithm with key clustering algorithms in Artificial Intelligence (AI) to visualize and explore hidden patterns in multidimensional data in a 3D virtual space. Experiments highlight the issues and benefits brought by a certain level of usability of user interaction techniques.

Author Keywords
Virtual Reality, Classification Algorithm, Clustering Algorithm, Visual Analysis, Multidimensional data.

DOI: 10.37789/icusi.2025.12

INTRODUCTION
Fully understanding certain abstract concepts, as well as the operational mechanisms of specific algorithms, is not a simple task. Despite the wealth of resources available online, there is a lack of opportunities to visualize these algorithms in action, which rendered their understanding more challenging than necessary. This challenge is further intensified when attempting to distinguish the differences and similarities among algorithms of the same category, such as clustering algorithms.

The Virtual Reality (VR) based approach transforms complex concepts into an intuitive, game-like experience, fostering a deeper understanding as users manipulate and explore various clustering techniques. Unlike traditional methods that rely on static images, videos or basic 2D visualizations, which are limited in fully conveying the potential of clustering algorithms, VR offers an immersive 3D environment. Users can actively navigate the scene, manipulate data points, and observe changes in real-time as algorithms progress through iterations.

The purpose of this research is to explore solutions for interactive visualization tool that highlight various clustering and classification algorithms. By leveraging the capabilities of VR, the tool aims to enhance user experience and deepen the understanding of these machine learning techniques.

Main Objectives of the research
Specifically, this research aims to:

- Experiment the user interactions and visual analysis of *clustering algorithms* such as K-Means, Density-Based Spatial Clustering of Applications with Noise (DBSCAN), Affinity Propagation, and Mean Shift, as well as *classification algorithms* such as K-Nearest Neighbors.
- Leverage the capabilities of the Meta Quest 2 headset to provide *immersive 3D data visualization*.
- Explore *multi-dimensional data* (n-dimensional), both labelled and unlabelled, and map and visualize it into the 3D virtual space.
- Provide a *graphical user interface* within VR that enables users to manage algorithms, select specific clustering or classification techniques, adjust parameter values, choose datasets from a predefined list, or add new ones and select three features to serve as 3D coordinates for visualization.

Functional Challenges
These challenges refer to the issues that may affect the overall behavior of the application:

- Design of an *intuitive main menu* that allows the users to easily select datasets and apply algorithms.
- *Smooth interaction* with the user interface (UI) and the 3D objects using the VR controllers.
- Flexible yet controlled *configuration of parameters* to ensure user input remains within predefined intervals.
- Real-time updates on the *progress of the algorithms*.
- *Dynamic update* of 3D charts based on the combination of datasets, dimensions, algorithms, and parameters: As the user interacts with the system, the 3D charts are dynamically updated, allowing them to visually track the performance of various algorithms across different parameters. This enables the user to deduce which parameters yield the best performance for a given algorithm or dataset, helping the user to draw conclusions about how different combinations impact algorithm's result.

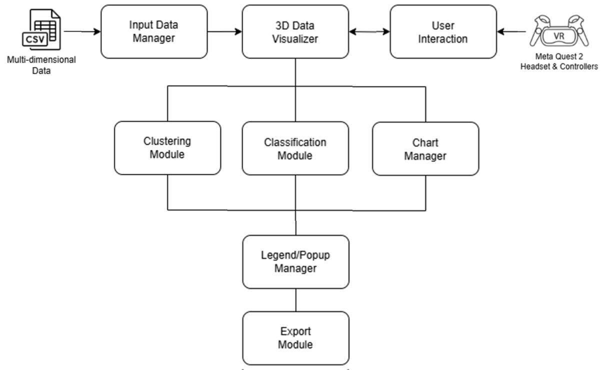

Figure 1. Conceptual architecture of the VR system.

Technological Challenges

- *Input Handling*, in particular Controller Input: Since controllers are the primary means for users to engage with the virtual world, proper mapping of inputs to specific actions, such as selecting or grabbing, is essential for creating intuitive interactions. Ray interactors are used for the user to reach beyond their immediate physical space, allowing interaction with distant objects. Users should receive immediate visual, auditory, or haptic feedback, confirming their inputs have been recognized and processed by the system.
- *Moving around in VR*: This is a tricky matter because moving in a VR scene while standing still in reality can make some people feel uncomfortable or dizzy. With regular movement, you can use joysticks to walk forward, backward, left or right. Snap turn allows the user to rotate by fixed degrees with each button press while continuous turn lets users smoothly rotate their view by holding the joystick left or right.

Performance Related Challenges

- *Frame Drop*: It means that the scene freezes for a moment because the application fails to render frames at the expected rate. It may happen because of heavy computations if they are not properly handled in background threads. This can break the immersion and create discomfort for the user.
- *Flickering*: A visual phenomenon where surfaces appear to flash rapidly which is distracting and can cause eye strain.

Datasets Related Issues

- *High Dimensionality*: Choosing a set of three dimensions from a multi-dimensional dataset involves a trial-and-error process as it is not always obvious which combinations will yield meaningful results.
- *Large Coordinate Ranges*: Some datasets may contain coordinates with very large values or values that span a wide interval which makes it very difficult to visualize such points as they appear too scattered in the scene. Thus, a normalization or scaling step is required during preprocessing.
- *Large Datasets*: Performance can be affected when rendering datasets with a big number of data points because typically, each data point is mapped to a game object, such as a sphere, that must be instantiated.
- *Distinction between Labelled and Unlabelled Dataset*: When dealing with labelled datasets, it is important to distinguish the label column from the feature columns as it represents the class and should not be visualized alongside the features. This deduction is made during the preprocessing step as well. However, the class still needs to be visualized in some way, but it is not considered a coordinate column. For both types of datasets, only numerical values are valid for feature columns as these values are directly translated into coordinates in a 3D virtual world.

The paper is structured as follows: the next section highlights the state of the art in the field of virtual reality-based education and algorithm visualization. Then, the clustering and classification algorithms of interest are briefly presented. The next section details the system architecture and then the experimental evaluation of the solutions. The last section concludes on the achievements.

RELATED WORKS

VR as Extended World

Last decades, a new broader term, Extended Reality (XR), has been used to collectively describe the following immersive technologies: Augmented Reality (AR), Mixed Reality (MR) and Virtual Reality (VR), [1], [2]. XR environments significantly enhance students' engagement and motivation to learn by facilitating interactions with realistic objects and environments, thus making abstract concepts more tangible and interactive [1]. However, the full integration of XR into the learning process, alongside traditional methods like slides or lectures, remains a challenge. Also, head-mounted displays used in VR, for example, can cause discomfort or nausea after only a short period of time, with symptoms potentially becoming more severe during extended use making them less suitable for long lessons.

VR allows users to interact with virtual surroundings making use of a head-mounted display and controllers (e.g. Meta Quest 2 and 3). They come equipped with sensors having the role of tracking the hands, and the head movement and gaze, respectively. The VR technology has been of interest even before the year 2000, but technological constraints have prevented it from having widespread use until its recent advances [3].

Naranjo et al. [4] highlight the significance of VR in the field of psychology, as well as the field of surgery. The study reveals that VR can be particularly effective in the treatment of traumas or phobias without putting the patients in danger.

In the field of surgery and healthcare both students and professionals can perfect their skills through highly realistic simulations of complex procedures, which helps reduce the risk of errors in real-life situations.

Virtual Reality in Education

Radianti et al. [5] reviews the use of VR in higher education, highlighting its growing potential across multiple application domains such as Engineering, leading by 24%, Computer Science 10%, Astronomy, Medicine, Biology, Earth Science, each by 5%. As the authors, who are also lecturers, state, students prefer practice-oriented learning over memorization, which is exactly where VR's potential lies: offering immersive, hands-on experiences that traditional teaching methods cannot provide.

Research conducted by Rangarajan et al. [6], which involved interviews with 16 university-level instructors, offers insights into the use of VR in higher education. It emphasizes the need for adequate training, technical support, and seamless integration with existing educational tools for VR to be successfully adopted and optimally utilized in educational settings. Also, this study highlights the idea of tailoring VR applications to meet diverse educational needs while ensuring effectiveness in classroom.

In another study, Virvou and Katsionis [7] identify several usability issues encountered mainly by novice users. However, students are encouraged to use educational VR games. In fact, they are impressed by the sophisticated VR environments, motivating them to engage in these games both during school and in their free time.

Freina and Ott [8] reveal the main motivation behind the use of VR in education such as its ability to transport users to worlds and scenarios otherwise inaccessible in real life, by immersing them in virtual environment that simulate distant places, historical events, or abstract concepts. Moreover, the remarkable aspect of VR is that it provides these experiences within a completely safe environment, allowing learners to potentially engage with risky or complex situations without facing real-world dangers.

Data and Algorithms by VR

As modern datasets become increasingly complex and hyperdimensional, visualizing and understanding these complex patterns poses a significant challenge. To address this, Donalek et al. [9] explores the use of immersive VR tools for data visualization, particularly in the domain of Big Data. It is revealed that VR not only enhances understanding and the perception of complex data, but it also facilitates collaborative exploration in a shared virtual environment. The development of the data visualization tools was performed using the Unity 3D platform, which is increasingly becoming a dominant choice for VR applications since it supports VR headsets, such as the discontinued Oculus Rift or Meta Quest 2 or 3, as a display device.

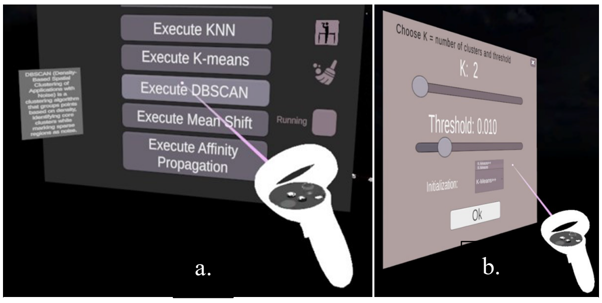

Figure 2. (a) Accessing DBSCAN menu; (b) K-Means configuration panel by K and Threshold.

Although VR technology seems promising for visualizing all types of artificial intelligence algorithms, it also comes with some limitations. Inkarbekov et al. [10] reviewed several works in this field and stated that although VR indeed has high potential, it has not yet been fully exploited. They also argue that a general solution for algorithm visualization may not be feasible and that more specialized systems may be needed.

The issue related with the high dimensionality of data was also discussed by Bobek et al. [11], which focused on visualizing clustering algorithms. The researchers proposed the solution of only displaying 3 dimensions of each data point at a time, so that they could be mapped on a scatter plot.

As Inkarbekov et al. [10] mentioned, there are still research gaps in literature, and there is still room to better harness the capabilities of VR. The system described in our paper aims to tackle some of the challenges. For instance, considering the call for more specialized systems, we focus only on the visualization of clustering algorithms. To prevent overwhelming the user, we limit the display to 3 parameters at a time, mapping them to spatial coordinates, while ensuring that the user can customize which parameters are shown.

CLASSIFICATION AND CLUSTERING ALGORITHMS

Classification and Clustering

While both Classification and Clustering aim to analyze and categorize data, they differ fundamentally in their approach and application domain. Classification relies on labelled data to predict outcomes from a predefined set of categories for new inputs and guidance from a supervisor, while clustering deals with grouping unlabelled data based on their similarity. Classification is typically used in image recognition and medical diagnosis, while clustering is used in data analysis, pattern recognition, being more suitable for exploratory tasks. There are totally different evaluation metrics since they address different types of problems. For classification tasks, the performance of the model is assessed on new, unseen data using common metrics such as accuracy, confusion matrix, or precision and recall. In contrast, for clustering tasks where the objective is to group data into

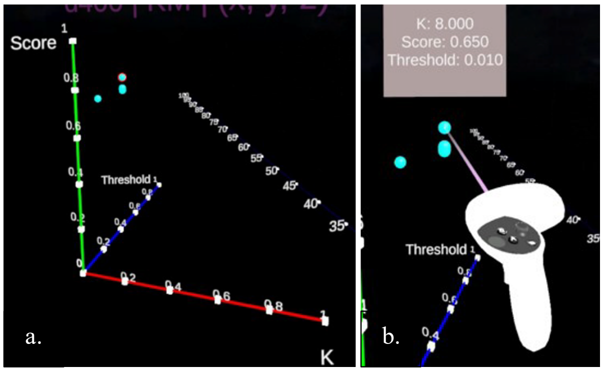

Figure 3. (a) Chart view; (b) Chart point inspection.

clusters, metrics include Silhouette Score, Davies-Bouldin Index and others.

For classification, we will explore the K-Nearest Neighbors (KNN) algorithm. For clustering, we will consider several key algorithms such as K-Means, Density-Based Spatial Clustering of Applications with Noise (DBSCAN), Affinity Propagation and Mean Shift.

K-Means Algorithm

K-Means [12] is a widely used clustering algorithm within the field of Unsupervised Learning. Its objective is to partition an unlabelled dataset into K distinct, non-overlapping subgroups/subsets called clusters. The algorithm ensures that data points belonging to the same cluster exhibit strong *similarities* to one another, while being significantly different from points assigned to other clusters.

Density-Based Spatial Clustering of Applications with Noise

Density-Based Spatial Clustering of Applications with Noise (DBSCAN) is a fundamental clustering algorithm in the field of Unsupervised Learning [13]. Unlike traditional clustering methods such as K-Means, which rely on predefining the number of clusters and assume a certain shape of the clusters, DBSCAN identifies clusters based on the *density* of data points in each region. This makes it particularly useful for discovering clusters of arbitrary shapes.

Affinity Propagation

Affinity Propagation (AP) is an innovative clustering algorithm that fundamentally differs from more conventional clustering techniques like K-Means or DBSCAN discussed previously [14], [15]. Instead of requiring the number of clusters to be specified in advance, AP discovers clusters by means of *message passing* between data points. The main drawback of AP is that it can be computationally expensive especially for large datasets since message-passing involves many iterations.

Mean Shift

Mean Shift (MS) is a powerful, non-parametric clustering algorithm used in various machine learning and computer vision applications [16]. It does not require the specification of the number of clusters in advance. Instead, it identifies clusters by shifting data points towards the high-density regions called *modes* within a certain radius, thus being known as the mode-seeking algorithm.

At the core of the MS algorithm is the *kernel* function, which determines how much influence the neighbouring points have on the point of interest based on their distance. The kernel value is computed by a Gaussian Radial Function dependent on the squared Euclidean distance and on the standard deviation parameter, which controls the width of the kernel/sliding window, meaning the size of the neighbourhood around a point.

K-Nearest Neighbors

K-Nearest Neighbors (KNN) is one of the simplest and most widely used non-parametric, lazy learning algorithms in the category of Supervised Learning for *classification tasks* [17], [18]. The algorithm identifies the k nearest neighbors using a distance metric, such as the Euclidean distance and predicts, assigns a label to the query point by considering the label with the highest frequency among its neighbors.

SYSTEM ARCHITECTURE

The conceptual architecture of the VR-based application for visualizing clustering and classification algorithms on multi-dimensional data (Fig.1) outlines the main logical components that collaborate to support core functionalities such as dataset input, 3D visualization, immersive user interaction using the Meta Quest 2 headset, execution of machine learning algorithms based on user-defined parameters, with real-time visual updates and export of clustering results.

Architectural Modules

The main components of the system are the following:

- *Input Data Manager*: is responsible for receiving multi-dimensional datasets, typically in CSV format, which serve as the primary source of input. This module handles the loading, parsing and validation of the dataset, determining whether it is labelled or unlabelled. It applies normalization to ensure all values fall withing a consistent range, enabling proper visualization in the 3D scene.
- *3D Data Visualizer*: each data point is visualized as a distinct object, whose position is determined by the user-selected dimensions, and configured to support user interaction, enabling selection and manipulation within the VR environment. This module ensures real-time updates to the visualization when clustering or classification results are applied.
- *User Interaction Module*: facilitates the connection between the user and the system by capturing input from Meta Quest 2 headset and controllers. It interprets user actions such as button presses, trigger pulls, grab activations, and joystick movement to enable selection,

manipulation, interaction with both the User Interface (UI) elements and data points, as well as exploration of the virtual environment. Through mechanisms like ray interactors, it supports engagement with 3D objects, including drag-and-drop repositioning and detailed inspection of individual points within the scene. It also allows users to select datasets, choose dimensions, configure algorithm-specific parameters and initiate clustering or classification algorithms. Additionally, the module supports teleportation functionality, enabling users to quickly navigate to key areas for further analysis.

- *Clustering Module*: handles the execution of unsupervised learning algorithms on unlabelled. Upon selection of one of the supported algorithms, such as K-Means, DBSCAN, Affinity Propagation or Mean Shift, and configuration of the respective parameters, the module initiates the clustering process. These algorithms operate iteratively until convergence is reached or a predefined maximum number of iterations has been met.
- *Classification Module*: execution of supervised learning algorithms, assigning labels to user-generated query points based on the existing labelled dataset. It predicts the class of new data points, updating their visual representation.
- *Chart Manager*: is responsible for generating and maintaining visual representations of the Silhouette Score within the 3D environment. It allows users to visually compare the impact of different configurations on the clustering quality, thus gaining deeper insights into algorithm performance and optimal parameter settings.
- *Legend/Popup Manager*: presents auxiliary information within the 3D scene, ensuring users are continuously informed about the current state of the system.
- *Export Module*: allows users to keep a permanent record of the outcomes for further analysis and, if needed, reload the exported files into the system as labelled datasets for use in classification tasks.

Development Tools and Technologies

The implementation of the experimental system uses the Unity engine as the primary development platform, Visual Studio 2022 as the integrated development environment (IDE), C# as the programming language, the Meta Quest 2 headset and controllers for VR interaction and the OpenXR plugin for ensuring cross-platform VR compatibility. Unity’s XR Interaction Toolkit and Input System provides core functionality for head and controller tracking, object interaction and user input handling in a VR environment.

EXPERIMENTAL EVALUATION AND VALIDATION

This section presents the experimental evaluation of the developed VR system for visualizing and interacting with clustering and classification algorithms. The validation process focuses on assessing functionality, performance, scalability, and usability regarding the proposed objectives.

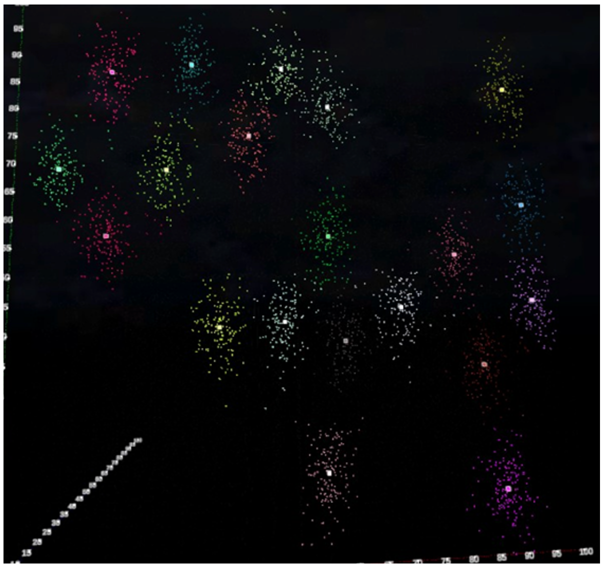

Figure 4. K-Means algorithm execution on *3000* dataset.

Experimental Datasets

The experimental evaluation employs numerous datasets that vary in terms of dimensionality, number of instances and labelling status, allowing for a comprehensive assessment of clustering and classification algorithms under different scenarios. All datasets consist of numeric features (integer or float) with no missing values.

Evaluation Metrics

The key metrics evaluate the system's correctness, scalability, and usability in the context of interactive visualization of clustering and classification algorithms within a virtual reality environment.

Silhouette Score Metrics

Evaluating clustering algorithms is a challenging task as they involve working with unlabelled datasets, where no ground truth is available for direct comparison. As a result, metrics such as accuracy or precision, which rely on known labels, cannot be computed.

One commonly used tool for assessing the quality of clustering results is the Silhouette Score [19]. It provides an indication of how well each data point fits within its own cluster compared to other clusters. This metric is calculated individually for each data point, then averaged to give an overall final score. For each data point i, two intermediary distances are computed:

1. *Mean Intra-Cluster Distance* a(i), representing the average of the distances between the point i and all the other points in the same cluster as point i.
2. *Mean Nearest-Cluster Distance* b(i), representing the minimum average of the distances between point i and all points in any other cluster.

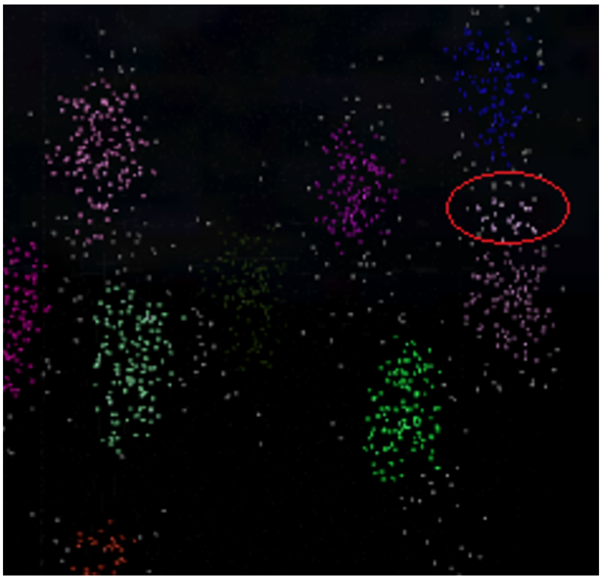

Figure 5. DBSCAN clustering on 3000 dataset. Over clustering highlighted by the red contour.

Using these two values, the Silhouette Score s(i) for point i is computed as:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}$$

The result is a number that ranges from -1 and 1 and can be interpreted as follows:

Close to 1: data point fits well within its cluster and is distant from others.

Around 0: data point is ambiguously clustered as it lies near the boundary of two clusters.

Close to -1: data point fits poorly in its assigned cluster which suggests it might belong to a different cluster.

The Silhouette Score is a very helpful tool that provides insightful feedback and can help determine the optimal number of clusters. In algorithms like K-Means, this is done through trial and error by testing different values of k and choosing the one with the highest score. Similarly, for algorithms such as DBSCAN or Mean Shift, it can guide the selection of parameters like epsilon (neighborhood radius) or bandwidth by revealing the quality of the resulting clusters.

It is worth mentioning that its reliability can be questionable when it comes to cluster of diverse, complex shapes, for example moons or spirals, since it might fail to accurately reflect the quality of clusters. The Silhouette Score tends to favor well-separated, spherical clusters.

Functionality Evaluation

The virtual menu allows the selection of the algorithm and to configure the specific parameters (Fig.2).

The multi-dimensional data is mapped into the 3D virtual space, allowing the user to specify the three coordinate axes and their meaning. The visualization is chosen appropriately to highlight important patterns and relationships between data (Fig. 3).

The system ensures the correct execution of both clustering and classification algorithms through visual and interactive validation.

The functionality of the clustering algorithms is assessed by visually inspecting whether the expected cluster structures are discovered, particularly for synthetic datasets with apparent and well-separated groups. This evaluation acknowledges the inherent algorithmic limitations, including K-Means assumption of spherical clusters leading to suboptimal performance on more complex, non-convex shapes, which is a constraint rather than an implementation deficiency.

The functional validation of the K-Nearest Neighbors classification is conducted by allowing users to spawn query points within the 3D scene. The system employs k-dimensional tree (KD-tree) optimization to efficiently retrieve the k closest neighbors, determines the frequency of each class label among them and assigns the majority label to the query point.

Performance and Scalability Evaluation

Performance is measured based on the total execution time (i.e. milliseconds) of each clustering or classification algorithm.

Scalability is assessed by testing the system with datasets of varying size, ranging from smaller datasets of approximately 400 instances to larger datasets of up to 3000 instances. The evaluation criteria include frame rate stability and responsiveness of the user interface during algorithm execution.

Usability Evaluation

The development team analyzed and validated technical and technological solutions that allow for appropriate usability of user interaction techniques.

- *Hands-on testing* across *multiple scenarios* to evaluate the ease/intuitiveness of user interaction and the effectiveness of the visual feedback within the immersive environment.
- *Main menu* is organized into sections that facilitate a clear interaction flow, following a logical top-down approach from dataset selection and addition through dimension selection and reset option, leading to the visualization button and then to the five algorithm-specific buttons.
- *Auxiliary controls*, including dataset addition, dimension reset, color reset, execution toggle, go to viewpoint and go to chart functions, are strategically placed on the side to

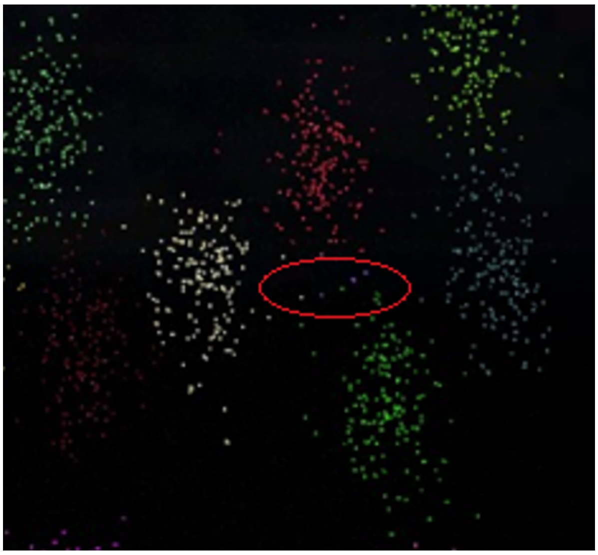

Figure 6. Mean Shift - Accurate clustering with minor outliers on 3000 dataset.

maintain separation from the primary workflow while remaining easily accessible.

- *Color-coding* scheme where the main workflow maintains a different color scheme than the auxiliary controls, creating clear visual distinction between essential operations and supplementary functions (Fig.4).
- *Tooltips*, each algorithm button incorporates tooltip functionality that provides concise descriptions and upon selection, dedicated parameter configuration panels appear, allowing real-time adjustment of key parameters through intuitive sliders and buttons.
- *Real-time visual updates* during algorithm execution, showing clustering progress through dynamic visualization of centroid updates, cluster boundary changes or point shifts, complemented by iteration progress popups that deliver status updates and guidance for interactive features such as spawning and deleting query points (Fig.5, Fig.6).
- *Contextual legends* map colors to cluster labels for labelled datasets or colors to frequency distributions for KNN visualization are also present, along with coordinate popups that reveal precise spatial information when hovering over data points.
- *Teleportation*. Users can move freely throughout the scene using joystick controls, while for ease of navigation, two teleportation options are provided, including teleportation to the latest updated chart and to the best viewpoint for optimal cluster inspection.
- *Interactive data manipulation* and enables successive algorithm execution, either by running the same algorithm with different parameters or testing different algorithms on the same dataset and dimensions.

Clustering quality

- *Silhouette Score* is used as the primary quantitative measure of the clustering quality, with higher scores indicating well-defined clusters with greater separation and cohesion.
- *Number of iterations* required for convergence is tracked, indicating the algorithm stability and convergence speed. Whether the algorithm has converged and how many iterations it takes is displayed through the real-time status update popups.
- *Interactive charts* are integrated within the VR environment as means of visualizing and assessing the evolution of Silhouette Score against the two algorithm-specific parameters.
- *Chart point inspection*. Each chart can be further inspected by hovering over it to reveal detailed information about the specific parameter values and the corresponding Silhouette Score, to gain deeper insights into their relationship as well as deduce the best combination of parameters.
- *Navigation* to the newly updated chart by offering teleportation functionality and highlighting both the chart title and the most recently added point for immediate identification.

CONCLUSIONS

The VR environment allows users to interactively explore multi-dimensional datasets mapped into the 3D space. The system integrates multiple clustering algorithms (K-Means, DBSCAN, Mean Shift, Affinity Propagation) with real-time visual feedback. The K-Nearest Neighbors classification algorithm enables users to interactively spawn query points in the VR scene, to classify them based on the labelled dataset and remove them optionally from the scene.

Providing visual aids, such as Silhouette Score charts, highlights how different combinations of parameters influence clustering quality.

Teleportation shortcuts allow users to quickly move to key points in the VR scene, such as viewing the optimal dataset or the most recently processed subset of data, increasing navigation efficiency.

Seamless immersive integration with the Meta Quest 2 headset provides user interaction and navigation in the VR space. Up to this stage, the development team has experimentally assessed the possibility of achieving a high level of usability through the adopted technical and technological solutions. This stage will be extended by the actual experimental validation of usability by a group of target users.

REFERENCES

1. D. Pimentel, G. Fauville, K. Frazier, E. McGivney, S. Rosas, and E. Woolsey", An Introduction to Learning in the Metaverse," Meridian Treehouse, 2022. [Online].

Available: https://scholar.harvard.edu/files/mcgivney/files/introductionlearningmetaverse-april2022-meridiantreehouse.pdf

2. Çöltekin, Arzu et al. "Extended Reality in Spatial Sciences: A Review of Research Challenges and Future Directions". SPRS Int. J. Geo-Inf., 2020. [Online]. Available: https://www.mdpi.com/2220-9964/9/7/439
3. P. L. Weiss and A. S. Jessel, "Virtual reality applications to work", WORK, 1998. [Online]. Available: https://content.iospress.com/download/work/wor11-3-05?id=work%2Fwor11-3-05
4. J. E. Naranjo, D. G. Sanchez, A. Robalino-Lopez, P. Robalino-Lopez, A. Alarcon-Ortiz, and M. V. Garcia, "A scoping review on virtual reality-based industrial training", Appl. Sci., 2020. [Online]. Available: https://doi.org/10.3390/app10228224
5. J. Radianti, T. A. Majchrzak, J. Fromm, and I. Wohlgenannt, "A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda", Computers & Education, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360131519303276
6. V. Rangarajan, A. Shahbaz Badr, and R. De Amicis, "Evaluating Virtual Reality in Education: An Analysis of VR through the Instructor’s Lens", Multimodal Technologies and Interaction, 2024. [Online]. Available: https://doi.org/ 10.3390/mti8080072
7. M. Virvou and G. Katsionis, "On the usability and likeability of virtual reality games for education: The case of VR-ENGAGE", Computers & Education, 2008. [Online].Available: https://doi.org/10.1016/j.compedu.2006.04.004
8. L. Freina and M. Ott, "A literature review on immersive virtual reality in education: State of the art and perspectives", The International Scientific Conference eLearning and Software for Education, 2015. [Online]. Available: https://www.researchgate.net/publication/280566372_A_Literature_Review_on_Immersive_Virtual_Reality_in_Education_State_Of_The_Art_and_Perspectives
9. C. Donalek et al., "Immersive and Collaborative Data Visualization Using Virtual Reality Platforms", 2014 IEEE International Conference on Big Data (Big Data), 2014. [Online]. Available: https://www.researchgate.net/publication/267515366_Immersive_and_Collaborative_Data_Visualization_Using_Virtual_Reality_Platforms
10. M. Inkarbekov, R. Monahan, and B. A. Pearlmutter, “Visualization of AI Systems in Virtual Reality: A Comprehensive Review”, International Journal of Advanced Computer Science and Applications (IJACSA), 2023. [Online]. Available: http://dx.doi.org/10.14569/IJACSA.2023.0140805
11. S. Bobek, S. K. Tadeja, Ł. Struski, P. Stachura, T. Kipouros, J. Tabor, G. J. Nalepa, and P. O. Kristensson, "Virtual Reality-Based Parallel Coordinates Plots Enhanced with Explainable AI and Data-Science Analytics for Decision-Making Processes", Applied Sciences, 2022. [Online]. Available: https://doi.org/10.3390/app12010331
12. H. He, Y. He, F. Wang, and W. Zhu, "Improved K-Means Algorithm for Clustering Non-Spherical Data", Expert Systems, 2022. [Online]. Available: https://www.researchgate.net/publication/362175498_Improved_K-means_algorithm_for_clustering_non-spherical_data
13. S. H. Nagesh, "An introduction to the DBSCAN algorithm and its implementation in Python". KDnuggets on April 4, 2022 in Machine Learning. [Online]. Available: https://www.kdnuggets.com/2020/04/dbscan-clustering-algorithm-machine-learning.html
14. T. Qiu and Y. Li, "A Generalized Affinity Propagation Clustering Algorithm for Nonspherical Cluster Discovery", 2015. [Online]. Available: https://www.researchgate.net/publication/271140657_A_Generalized_Affinity_Propagation_Clustering_Algorithm_for_Nonspherical_Cluster_Discovery
15. D. Dueck, "Affinity Propagation: Clustering Data by Passing Messages", PhD thesis, 2009. [Online]. Available: https://utoronto.scholaris.ca/server/api/core/bitstreams/909ffca0-d1c7-4eb4-9f44-84f03b5fe927/content
16. E. Cecchini, "Comparison between a sequential and a multithreading version of the mean shift clustering algorithm", [Online]. Available: https://github.com/sinecode/MeanShift/blob/master/Paper/MeanShift.pdf
17. V. Tiwari, "Developments in KD Tree and KNN Searches," International Journal of Computer Applications, 2023. [Online]. Available: https://www.researchgate.net/publication/371770725_Developments_in_KD_Tree_and_KNN_Searches
18. W. Hou, D. Li, C. Xu, H. Zhang, and T. Li, "An Advanced k Nearest Neighbor Classification Algorithm Based on KD-tree," in Proc. IEEE Int. Conf. on Intelligent Computing and Signal Processing (IICSPI), 2018. [Online]. Available: https://www.researchgate.net/publication/332434248_An_Advanced_k_Nearest_Neighbor_Classification_Algorithm_Based_on_KD-tree
19. J. Tushar, "Evaluating Clustering Algorithm - Silhouette Score", [Online]. Available: https://tushar-joshi-89.medium.com/, 2021