

A NOVEL ASSOCIATION-BASED ENSEMBLE CLUSTER FOR HEART PREDICTION BASED ON ECG GRAPHS

G BALA KRISHNA¹, SIRISHA K L S², V PRADEEP KUMAR³, MUMMADI RAMACHANDRA⁴,
D SANDHYA RANI⁵, K GNANESHWAR⁶

¹Professor, Department of CSE, CVR College of Engineering, Ibrahimpatnam, Rangareddy, India

²Associate Professor, Department of CSE, CVR College of Engineering, Ibrahimpatnam, Rangareddy, India

³Associate Professor, Department of CSE, B V Raju Institute of Technology, Narsapur, Medak, India

⁴Associate Professor, Department of CSE, Sreenidhi Institute of Science and Technology, Hyderabad, India

⁵Professor, Department of CSE, CVR College of Engineering, Ibrahimpatnam, Rangareddy, India

⁶Assistant Professor, Department of CSE, Sreenidhi Institute of Science and Technology, Hyderabad, India

E-mail: govind.krishna83@gmail.com, klssirisha@gmail.com, pradeepkumar.v@bvrit.ac.in,
ramachandra.m@sreenidhi.edu.in, sandhyarani.hcu@gmail.com, tognaneshwar@gmail.com

ABSTRACT

Traditional image processing techniques usually process data frame-by-frame, which may lead to a loss of temporal information. Temporal aspects are crucial in understanding the ECG waveform's sequential events and dynamic changes. Social Network Analysis (SNA), or graph theory analysis, studies relationships and interactions between individuals or entities within a social network. This analysis can provide valuable insights into various social structures, dynamics, and behaviors. Analyzing electrocardiogram (ECG) data using social graphs is an innovative approach that can provide insights into the cardiac system's dynamics and relationships between different physiological parameters. Although ECG is traditionally used to study the heart's electrical activity, incorporating social graph analysis can offer new perspectives on the interactions and dependencies between various ECG components and their influence on overall cardiac health. Using clustering techniques in social graph analysis can offer several advantages, as it helps identify patterns, groups, and structures within the network.

Social networks often consist of diverse data types, such as textual content, user attributes, and interaction patterns. Ensemble clustering can effectively integrate information from different data sources and algorithms, enabling a more comprehensive analysis. The proposed methodology integrates DBSCAN with OPTICS and produces a consensus matrix with the help of an association matrix. DBSCAN and OPTICS are density-based clustering algorithms, but they have different characteristics. DBSCAN is more efficient and straightforward to implement. At the same time, OPTICS provides a hierarchical clustering structure and is more suitable for handling varying-density clusters. Integrating these two algorithms can enhance the overall clustering performance, especially when dealing with complex and varied density datasets

Keywords: *Density-based algorithms, Social Analysis, Ensemble Clustering, Association Matrix, Community Detection*

1. INTRODUCTION

An ECG could be helpful if a heartbeat is difficult to feel, too fast, or unpredictable to count correctly. An abnormally slow or rapid heartbeat can be found using an ECG, and a beating heart [10]. An ECG can be used to identify unusual heartbeats. Natural electrical impulses control the heart's numerous muscle contractions to ensure appropriate blood flow. An ECG records these electrical impulses to give data on the heart's pace, rhythm (normal or steady), and the length and intensity of the electrical signals as they travel through all of

the heart chambers. A range of cardiac disorders can be detected through ECG changes. ECG determines whether symptoms like excessive tiredness, shortness of breath, dizziness, or fainting may be heart-related. During an examination of the body to capture the heart's activity initially, this recording can be compared with subsequent ECGs to see if any changes have occurred. Clustering allows for identifying and categorising various ECG patterns and waveforms [19]. Different cardiac conditions can exhibit unique patterns in the ECG signal, and clustering can help distinguish these patterns for accurate diagnosis.

Apply various clustering algorithms to the association matrix to obtain multiple clustering solutions. These algorithms could include traditional techniques like k-means, hierarchical clustering, or spectral clustering, as well as density-based methods like DBSCAN or OPTICS [16]. To make the data points within every group more similar to one another and distinct from the data points in the other groups, clustering divides the population or collection of data components into several groups [21]. In essence, it is a classification of things according to their similarities and differences.

a. Dense-based: According to these approaches, clusters are a dense part of the space with a few parallels to and differences from the less dense region.

b. Hierarchical based: The clusters produced by this method provide a structure mimicking a tree based

on the hierarchy [22]. To generate new clusters, the existing ones are utilised.

c. Partitioning methods: These techniques split the items into k clusters, so each split results in a new cluster. The goal criterion similarity function, such as when distance is a key parameter, is optimised using this technique.

d. Grid-based methods: This method divides the data space into several grid-like units. These grids enable quick and autonomous clustering operations.

e. Model-based clustering: Model-based clustering is an attempt to improve the fit between a mathematical model and the provided data [11]. It is predicated on the idea that data are produced by combining elements of a simple probability distribution. Each object can be assigned to a cluster based on weight. Figure 1 presents the different types of clustering

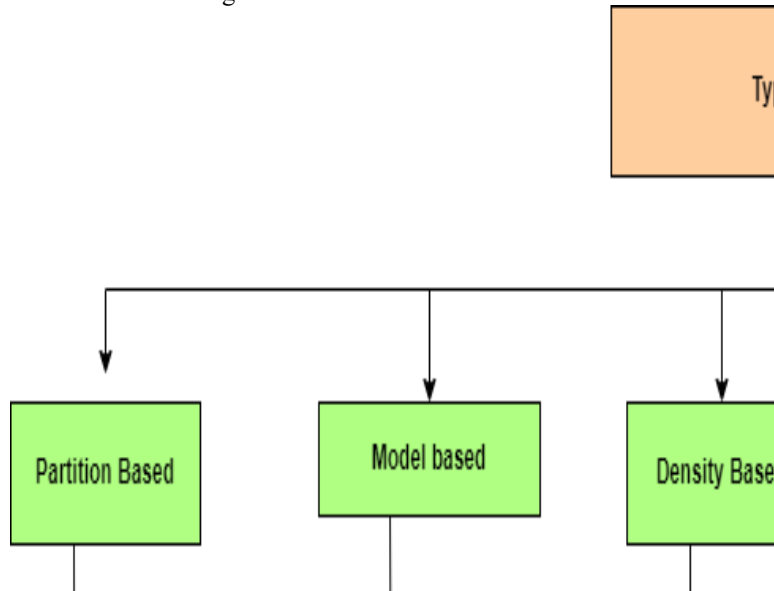


Figure 1: Clustering Categorisation Mechanisms

The various clustering methods frequently provide results that are quite distinct. This occurs due to the many criteria, such as cases, that are utilised to merge clusters. It is essential to assess which technique is appropriate for the subject carefully. It is pricey [12]. The cluster requires robust hardware and a strategy, which makes it more expensive than a non-clustered server administration solution. The fact that this design is not cost-effective is a serious flaw [19]. Since clustering takes additional computers and gear to set one up, it is challenging to monitor and manage [14].

As a consequence, expand the infrastructure. Interpreting the findings of an unclear or poorly defined cluster might be challenging. Clustering has many drawbacks. One is the temporal complexity of dealing with many attributes and data objects.

A clustering ensemble seeks to integrate various clustering models to provide more consistent and high-quality results than individual clustering methods [13]. In the ensemble learning approach, basic classifiers are created using the clustering technique, enhancing their diversity. Make the basic classifier pay greater attention to instances that are challenging to categorise while

doing this. Ensemble models incorporate a range of models to aid in prediction (decision-making) [20]. The possibility of excessive variation, low accuracy, noise, and bias increases if a single model cannot make the appropriate forecast for a given data set [18]. Models have a larger probability of increasing accuracy by merging numerous models.

2. LITERATURE SURVEY

Arezoo Dahesh et al. [1] introduced a SAECE-CD as a community detection method combining a stacked autoencoder for feature extraction and ensemble clustering to achieve final communities. In the pre-processing phase, the graph's adjacency matrix is reconstructed while preserving local and global neighbourhoods. The second phase uses a layered autoencoder to remove the reconstructed matrix's low-dimensional characteristics. Each node in the graph is represented as a vector and used as input to the autoencoder. The low-dimensional features for each node are obtained, and an embedding matrix H for the graph is created for clustering. The stacked autoencoder is trained using the mean square error (MSE) as the loss function. Hyperparameters of the autoencoder are tuned using Bayesian hyperparameter optimisation to minimise reconstruction error and enhance feature extraction quality. The final phase involves ensemble clustering, consisting of the generation mechanism and the consensus function. Different initial clustering algorithms or hyperparameter settings can be applied to the dataset to create base clustering results.

Wasim Khan et al. [2] use real-world attributed datasets BlogCatalog & Flickr. Autoencoders consist of two NN, an encoder & a decoder, utilised for compression and reconstruction of input data, respectively. AE aims to optimise parameters to recreate samples similar to the original. VAE is an extension of AE with a prior distribution added to the embedding learning process. It learns a posterior distribution of continuous latent variables, making it suitable for datasets with high dimensions, like social networks. GANs consist of two NN, a discriminator and a generator. The generator translates samples from an earlier distribution into the data space.

In contrast, the discriminator estimates the likelihood of a given data point being real or fake. Each model reconstructs a sample in its inference phase, and the anomaly scores for each model are computed using the mean square error. The final anomalous score is computed using an average weighted ensemble operation, with weights based on the correlation across model outputs. The

weighted total of each anomaly value from each model is used to get the overall anomaly score for a sample. The resulting score is then compared to a predetermined threshold to establish the sample's type.

Peng Zhou et al. [3] have introduced the CESH to address the problem of inconsistency and uncertainty present in individual clustering solutions by leveraging the hypergraph representation of the data and learning a structured hypergraph that captures the underlying clustering structure. The algorithm starts by constructing an initial hypergraph representation from the base clustering results. The quality scores of base clusters are evaluated based on their consistency with other base clusters. A cluster-wise is computed from the initial hypergraph representation. The main part of the algorithm involves iteratively optimising the hypergraph representation to obtain a structured hypergraph that reveals the underlying clustering structure. During each iteration, the algorithm optimises an auxiliary consisting of the eigenvectors with the least matching eigenvalues of the Laplacian. The optimisation steps are repeated until convergence is achieved.

Alexandru Petrescu et al. [4] have introduced EDSA, ML, and DL technics where the data obtained from Twitter data was listened from Twitter data and passed to the Pre-processing Module. Different models for text representation, including three modules. The event Detection Module Utilises three algorithms for detecting bursty chapters in clear text. Naïve Bayes includes a probabilistic classification algorithm for sentiment analysis with good results. Logistic Regression involves a Statistical model for binary classification within a probabilistic context. Ridge Classifier can learn an approach for several regressions in several classes. SVM can supervise learning algorithms for binary classification, which is particularly useful for sparse datasets. LSTM in an NN-related approach for write-up classification with the optimised outcome. EDSA-Ensemble Algorithm the algorithm combines Event Detection and Sentiment Analysis modules using an ensemble approach to determine individual and event sentiments for bursty topics on Twitter data.

Hang-Yuan Du et al. [5] have proposed a CEvGE framework that evaluated synthetic and real-world datasets, including six different fields. The proposed framework, called CEvGE, is designed for clustering ensembles and aims to create an agreement-based partition for the specified collection of data items. The coupling connections among data items are also taken into account from

an intra- and an inter-view standpoint, which results in creating an affinity matrix that organises the data's structural information. The clustering associated job is changed to a graph-based collected issue on the reorganised network using the CEvGE framework, which incorporates data attributes and structural information. The graph-represents text is mapped to unused embedded data using a GCN in the inference model. The learnt embeddings are utilised to rebuild the initial data, & the related model GMM is used to forecast the consensus group assignments.

Xuan Jiang et al. [6] have identified an over-smoothing problem in graph-based data analysis by exploring the diversity of the graph spectrum. Co-authorship graphs involved authors as nodes, and edges were determined by co-authorship. This includes goods as nodes and edges connected goods frequently bought together. Web network graphs consisted of webpages as nodes, connected by hyperlinks, with node features represented as bag-of-words from the webpages. The "sparse motif" method eliminates irrelevant connections and reduces the density of the motif-level graph Laplacian. A hierarchical motif aggregation mechanism that integrates early global features,

motif-level local features, and graph-level global features to enhance the representation. The theoretical analysis supports their claim that their method will not lead to over-smoothing if sufficient diversity across motifs exists. Evaluation is performed on seven real-world datasets, and experimental results demonstrate its superiority compared to existing approaches. Table 1 represents the metric analysis.

Table 1: Metric Analysis of the Existing Approach

Author	Year	Algorithm	Merits	Demerits
Arezoo Dahesh et al.	2023	Autoencoders, ensemble clustering	By reconstructing the adjacency matrix, the performance is efficient.	Based on votes, the output can be changed.
Wasim Khan et al.	2022	Autoencoders, VAE, GAN	The model was clustered on multiple occasions with optimised efficiency.	Identifying all the models may take a long time.
Peng Zhou et al.	2022	Clustering ensemble, hypergraph learning	It increases stability.	It will not work accurately.
Alexandru Petrescu et al	2023	Sentiment Analysis, EDSA, ML, & DL	It identifies accurately with text-based data.	The comparison was not proper
Hang-Yuan Du et al.	2022	GNN, GMM	The models are clustered accurately.	The real data was appropriate compared to reel data.
Xuan Jiang et al.	2022	GCN, SMEGCN	The diversity of the graph is identified with high preliminary.	The data should be labelled.

3. PROPOSED METHODOLOGY

Cluster analysis is breaking up a collection of input data into subsets that may overlap and where the components in each subgroup are thought to be connected by some similarity metric. "graph analytics" refers to an analysis based on entities or "graph nodes," which can be things like goods, customers, processes, or even gadgets. In many real-life situations, graphs naturally occur, including social graphs or diffusion graphs in networking sites, citation graphs in academic fields, user interest graphs in e-commerce, knowledge graphs, and illness graphs or graphs for patient-to-patient interaction in infectious diseases like COVID-19. The hidden information is better understood when graphs are effectively analysed. Particularly, the notions of sub-graphs, network components, and communities have been used extensively in research on the structure of the graphs.

An ensemble approach aims to increase performance by combining the forecasts of many machine-learning approaches. It is a form of meta-learning where a third-party agent determines a final prediction after evaluating the individual forecasts of other agents. Clustering is more complicated than supervised techniques, where the ensemble may be created simply by piling some learners on top of others. If the point is in group 1 or 2, it makes no difference; the important thing is who else is in this group with the point. In clustering, the most recent label of the data set has no significance on its own. As a result, combining the data from two distinct clustering findings is more complex than just determining if all labels constitute the same, similar to classification. It is necessary to devise a method for calculating the degree of agreement among the various divisions. Consensus clustering, also known as ensemble clustering, has gained much attention recently to combine many basic clustering methods into a more effective ensemble clustering. Because of its strong performance, ensemble clustering is essential in several research fields, including community identification and bioinformatics. The ensemble's ultimate output is calculated by combining the results of many clustering techniques. The clustering algorithms' results are combined using a consensus function to enhance the ensemble's output. With prior information, choosing a consensus function can be easier since there are several varieties of consensus functions with different degrees of computational complexity. The

methodology adopted by the ensemble clustering is represented in Figure 2

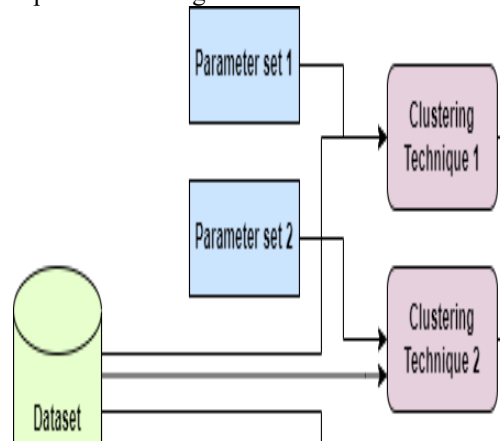


Figure 2: Generation of Consensus Partition using Ensemble Clustering

The DBSCAN approach is a valuable component in ensemble clustering for several reasons. DBSCAN is a density-based clustering algorithm that can find groups of arbitrary forms in the data. This can effectively identify clusters with complex geometries. They are designed to handle noise or outlier points effectively. It distinguishes noise points as those lying outside any cluster and does not assign them to any specific cluster. DBSCAN adapts to such variations by adjusting the size of neighbourhoods based on the local density of data points. DBSCAN's strengths complement those of other clustering algorithms. While traditional algorithms like k-means may struggle with non-convex clusters and outlier detection, DBSCAN excels in these areas. The time complexity is relatively low compared to other density-based algorithms, making it computationally efficient. Figure 3 presents the workflow of DBSCAN to perform clustering.

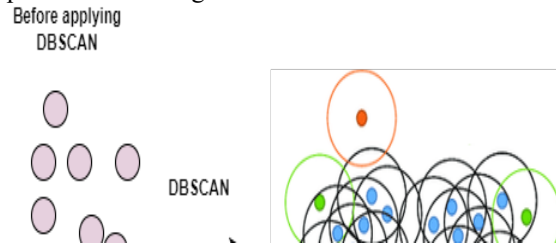


Figure 3: Workflow of DBSCAN Algorithm

OPTICS is a density-based clustering algorithm that identifies clusters in spatial data, such as geographical coordinates. However, it can also be applied to other types of data representations. OPTICS introduces the concept of "reachability distance" to measure the density of points in a dataset. For each data point, the reachability

distance is defined as the distance to the k th nearest neighbour, where k is a user-defined parameter. This distance is used to identify core points, which are points with a sufficient number of neighbours within a specified radius (epsilon). By analysing the reachability plot, OPTICS can identify clusters of different densities and shapes. Different density thresholds can be applied to extract clusters, allowing for greater flexibility in cluster identification compared to DBSCAN. The reachability plot visually displays the density-based hierarchy of clusters, making it useful for cluster exploration and parameter tuning. The procedure to generate clusters is presented in Figure 4.

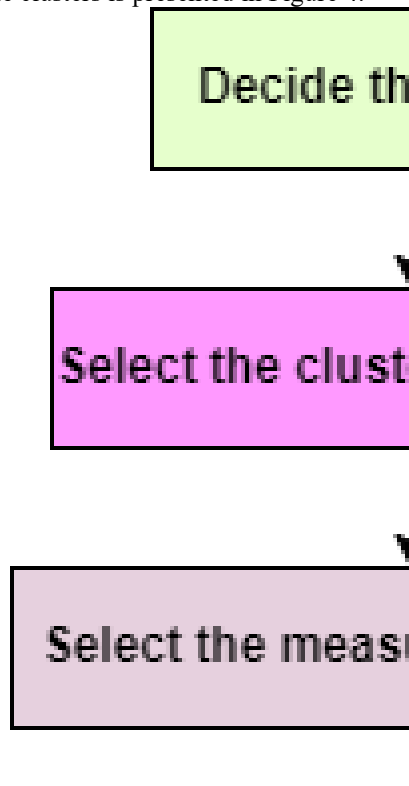


Figure 4: Generation & Validation of Clusters

Generating the association matrix is crucial in various data mining and machine learning tasks, especially within the association rule-mining industry. The association matrix is typically constructed from a transactional dataset, where each column denotes an item, and each row denotes a transaction. The matrix's components are binary values that denote whether an item is present or absent in a particular transaction. Before constructing the association matrix, the transactional dataset often undergoes pre-processing. In pre-processing, each unique item in the dataset is assigned a unique identifier or index. The set of unique items forms the columns of the

association matrix. While encoding each transaction in the dataset is encoded in a way that corresponds to the presence or absence of items. The association matrix is constructed by arranging the encoded transactions into rows. A row and each item by a column represent each transaction. Several association rule extraction methods, like Apriori or FP-Growth, use the resulting association matrix as input to discover common item groups and extract useful association rules. The textural cells are the recurring patterns in the texture picture. Since textural cells vary in size, mining topographical association rules can be applied at several levels and with multiple resolutions. It is necessary to create an effective tool for extracting a set of relationships from photos that capture the needed information. It is required to model the visual in terms of objects and transactions and mine the common patterns to utilise association rules to collect this information.

Each pair's likelihood of showing up in the same cluster is recorded in the association matrix. This frequency can teach about the neighbourhood. Thus, it can measure how similar two things are. The association matrix may also be considered the original data's kernel transformation. Particularly, an existing co-association matrix cannot be used to generate the original base partitions. This implies that some partition information might need to be noticed during the translation from the segments to the matrix. The link between an observation and an initial cluster is the main emphasis of the binary matrix. The proposed flow diagram is presented in Figure 5

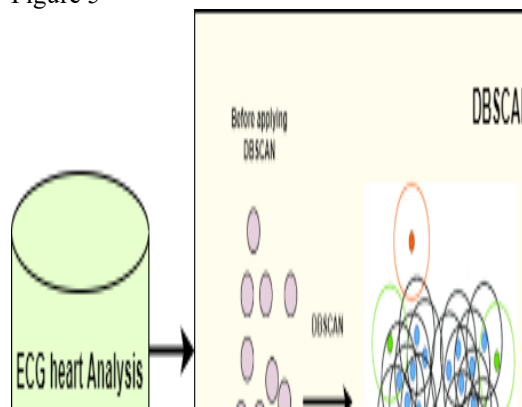


Figure 5: Generation of Association Matrix from DBSCAN

4. RESULTS & DISCUSSION

Heart attack datasets are often used in machine learning and data analysis to predict the likelihood of a heart attack based on various medical features and risk factors. These datasets can help researchers

and medical professionals better understand the factors associated with heart attacks and develop predictive models for early detection. Table 2 provides the attributes available in the dataset and this dataset in public repositories like UCI, Kaggle

Table 2: Attributes Information for Heart Attack Prediction

Name	Description	Type
Age	Patient Age	Numerical
Sex	Patient Gender	Categorical
Chest Pain: Type	Impact of pain on chest	Categorical
Resting BP	Blood Pressure	Numerical
Cholesterol	Fat Levels	Numerical
FBS	Sugar Levels during fasting	Boolean
Resting ECG	Waveforms type	Categorical
Heart Rate	During exercise, heartbeat levels	Numerical
EIA	Angina Impact	Boolean
ST	Depression rate due to exercise	Boolean
Slope	Segmentation slope	Categorical

Figure 6 is a popular dimensionality reduction technique in visualisation and exploratory data analysis. It is particularly effective for visualising high-dimensional data in a lower-dimensional space, aiming to preserve local structures and capture complex patterns that may not be easily apparent in the original data.

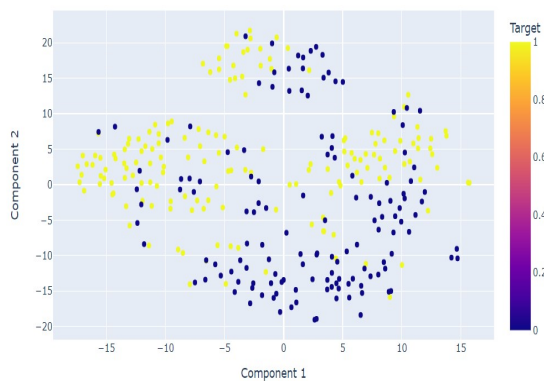


Figure 6: Component t-SNE presentation

The similarity or dissimilarity between data points is calculated based on the attribute values. Various distance or similarity measures are used depending on the type of attributes (e.g., Euclidean distance for numeric attributes, Jaccard similarity for categorical attributes). After clustering, the

attributes of the data points within each cluster can be analysed to interpret the characteristics of the clusters and understand what features define each group.

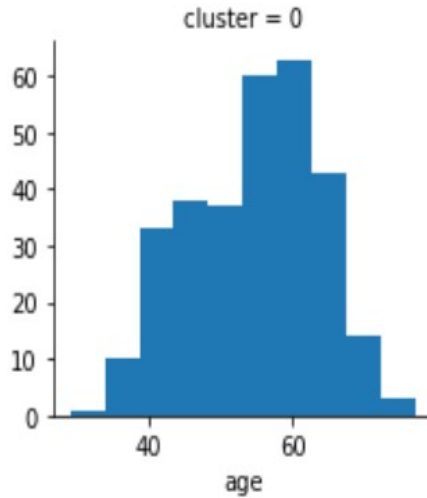


Figure 7: Cluster Analysis of Attributes

i. Similarly Matrix: A binary similarity matrix records the piecewise similarity among any two items for each input partition; a similarity of one denotes the presence of a cluster for two objects, while zero similarity indicates their absence. Only half of the matrix mentioned above is required because it is symmetrical. One of the metrics mentioned before calculates how similar all item pairings are. This results in the similarity or distance matrix, depending on how you define (dis)similarity. Clustering like pictures is a crucial task in image processing and artificial intelligence. To quantify comparisons between pictures, a measure is needed. The selected similarity metric will affect how well the clustering method performs. The proposed methodology computes the correlation coefficient:

$$R = \frac{\sum_{i=1}^n (y_i - \bar{Y})(x_i - \bar{X})}{\sum_{i=1}^n (y_i - \bar{Y})^2 (x_i - \bar{X})^2} \quad - (1)$$

X and Y are respective means of data points (X, Y).

Where X_i and Y_i are respective values in the ith data point.

ii. Dice Coefficient: Other names for the Dice similarity coefficient, a statistical instrument used

to compare two data sets, are the Srensen-Dice index and simply the Dice coefficient. This index may be the most popular tool for evaluating photo segmentation algorithms developed by AI. However, it is a much more general concept that can be used to assess data sets for various reasons, including NLP. The Dice coefficient may assess the pixel-by-pixel agreement of a predicted segmentation with the corresponding ground truth. To put it another way, the Dice score is determined by dividing the dimension of the intersection by the total sizes of the two groups. As a result, normalised by the quantity of the sets, the Dice score represents the percentage of overlap across the two sets. The formula is shown in equation (2)

$$\text{Dice Coefficient} = \frac{2 * |X \cap Y|}{|X| + |Y|} \quad - (2)$$

where Y is the actual ground truth, and X is the expected collection of pixels. When X and Y are both empty, the Dice coefficient is said to be 1. The leaderboard score is the average of the Dice coefficients per picture in the test set.

Figure 8 Begin by visually inspecting the graph and the clustering results. Many ensemble clustering methods provide visualisations, such as colour-coded nodes or edges representing cluster assignments. Look for patterns, clusters, and any variations in the cluster assignments.

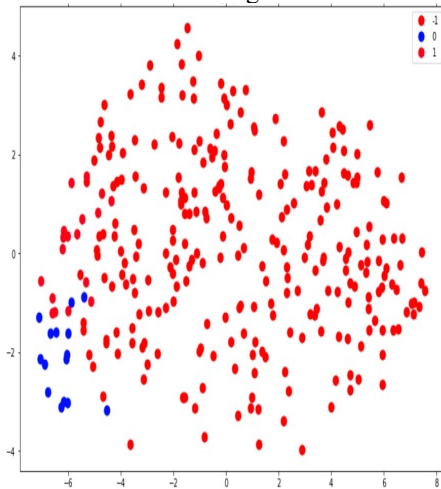


Figure 8: Clustering of Attributes using the Ensemble Approach

Table 3 presents the metrics necessary to compare the performance of the clustering approaches. The Silhouette Score is a metric commonly used to evaluate the quality of clusters produced by various clustering algorithms. ARI measures the agreement between the true labels of the data and the labels

assigned by the clustering algorithm. NMI considers both the mutual information and the entropy of the labels, providing a measure of how much information is shared between the true labels and the cluster labels while accounting for the size of the clusters.

Table 3: Metric Values on Clustering Algorithms

	Silhouette Score	ARI	NMI
DBSCAN	0.05	0.56	0.33
OPTICS	0.49	-0.16	0.501
Hierarchical	-0.75	-0.28	0.04
Ensemble	0.98	1	0.97

5. CONCLUSION

Clustering algorithms like k-means are sensitive to the initial placement of cluster centres, which can lead to different results depending on the initial random seed. Clustering results from some traditional algorithms might not be easily interpretable, particularly in high-dimensional spaces, making it challenging to understand the cluster assignments. Ensemble clustering with DBSCAN and OPTICS for ECG (Electrocardiogram) analysis can be advantageous in addressing some of the challenges associated with clustering cardiac data. DBSCAN and OPTICS are density-based clustering algorithms, and their integration in an ensemble can provide complementary benefits for ECG analysis. Ensemble clustering can provide insights into the stability of the clustering results. Suppose clusters are consistently detected across different DBSCAN and OPTICS runs. In that case, they will likely be more reliable representations of the underlying cardiac patterns

REFERENCES:

- [1] Dahesh, Arezoo and Naderi Dehkordi, Mohammad, Saece-Cd: A Hybrid Community Detection Algorithm Based on Stacked Autoencoder and Ensemble Clustering. Available at SSRN: <https://ssrn.com/abstract=4331048> or <http://dx.doi.org/10.2139/ssrn.4331048>

- [2] Khan, W., & Haroon, M. (2022). An unsupervised deep learning ensemble model for anomaly detection in static attributed social networks. *International Journal of Cognitive Computing in Engineering*, 3, 153-160. <https://doi.org/10.1016/j.ijcce.2022.08.002>
- [3]. Zhou, P., Wang, X., Du, L., & Li, X. (2022). Clustering ensemble via structured hypergraph learning. *Information Fusion*, 78, 171-179. <https://doi.org/10.1016/j.inffus.2021.09.003>
- [4]. Petrescu, A., Truică, C., Apostol, E., & Paschke, A. (2023). EDSA-Ensemble: An Event Detection Sentiment Analysis Ensemble Architecture. *ArXiv*. /abs/2301.12805
- [5]. Du, Hang-Yuan, and Wen-Jian Wang. 2022. "A Clustering Ensemble Framework with Integration of Data Characteristics and Structure Information: A Graph Neural Networks Approach" *Mathematics* 10, no. 11: 1834. <https://doi.org/10.3390/math10111834>
- [6]. Jiang, Xuan, Zhiyong Yang, Peisong Wen, Li Su, and Qingming Huang. "A sparse-motif ensemble graph convolutional network against over-smoothing." In *Proc. 31st Int. Joint Conf. Artif. Intell.*, pp. 2094-2100. 2022
- [7] P. Makam and G. Janardhan, "Survival Analysis of Heart Failure Patients Using Advanced Machine Learning Techniques," *2023 International Conference on Advanced & Global Engineering Challenges (AGEC)*, Surampalem, Kakinada, India, 2023, pp. 91-95, doi: 10.1109/AGEC57922.2023.00029..
- [8] Bhavani, S., Subhash Chandra, N. (2023). Histogram Based Initial Centroids Selection for K-Means Clustering. In: Goswami, S., Barara, I.S., Goje, A., Mohan, C., Bruckstein, A.M. (eds) *Data Management, Analytics and Innovation. ICDMAI 2022. Lecture Notes on Data Engineering and Communications Technologies*, vol 137. Springer, Singapore. https://doi.org/10.1007/978-981-19-2600-6_38
- [9] Kohane, I. S., Lander, E. S. (2020). Machine learning as an alternative to risk scores for identifying individuals at risk of heart attack. *JAMA Cardiology*, 5(2),127-128. doi:10.1001/jamacardio.2019.4787
- [10] Kandukuri Ratna Prakarsha, Gaurav Sharma, Time series signal forecasting using artificial neural networks: An application on ECG signal, *Biomedical Signal Processing and Control*, Volume 76, 2022, 103705, ISSN 1746-8094, <https://doi.org/10.1016/j.bspc.2022.103705>.
- [11] Pazzani M. J., LeBlanc M., Marler M. (2020). A framework for collaborative, hierarchical human interpretation of machine learning outputs: Application to a study of heart attack risk in Type 2 diabetes. *Journal of Biomedical Informatics*, 102, 103357. doi:10.1016/j.jbi.2019.103357
- [12] Godi, R. K., & Janakiraman, S. (2022). Border Collie optimisation algorithm-based node clustering technique in vehicular ad hoc networks. In *Transactions on Emerging Telecommunications Technologies* (Vol. 33, Issue 5). Wiley. <https://doi.org/10.1002/ett.4457>
- [13] Motwani, M., Dey, D., Berman, D. S., Germano, G., Achenbach, S., Al-Mallah, M. H., ... & Slomka, P. J. (2020). Machine learning for predicting all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis. *European Heart Journal*, 41(22), 2125-2134. doi:10.1093/eurheartj/ehz766
- [14] Soppari, K., Chandra, N.S. Automated digital image watermarking based on a multi-objective hybrid meta-heuristic-based clustering approach. *Int J Intell Robot Appl* 7, 164-189 (2023). <https://doi.org/10.1007/s41315-022-00241-3>
- [15] Chatterjee, S., & Dey, N. (2020). Deep learning approach for prediction of heart attack. *International Journal of Computational Intelligence and Informatics*, 9(1), 1-13.
- [16] Mahmud, M., Islam, M. T., & Khan, M. S. (2020). Early Prediction of Heart Attack Using Deep Learning Techniques. 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 1-5.
- [17] Banerjee, R., Bhowmick, P., & Ghosh, A. (2020). Heart attack prediction using deep learning. *International Journal of Computer Applications*, 174(31), 15-20.
- [18] Sharma, S., Srinivasan, D., & Pani, S. K. (2020). Prediction of heart attack using deep learning neural networks. 2020 6th International Conference on Signal Processing and Integrated Networks (SPIN), 535-539.
- [19] Radha, Marepalli, Maragoni Mahendar, and Pinnapureddy Manasa. "A comparative analysis for deep learning-based approaches for image forgery detection." *International Journal of Systematic Innovation* 8.1 (2024): 1-10..

- [20] D. Sandhya Rani, T. Sobha Rani and S. Durga Bhavani “Feature Subset Selection using Consensus Clustering” in the proceedings of “The Eighth International Conference on Advances in Pattern Recognition(ICAPR-2015)” published by IEEE Computer Society.
- [21] D. Sandhya Rani, T. Sobha Rani and S. Durga Bhavani “Consensus clustering for dimensionality reduction” in the proceedings of IC3-2014 published by IEEE Computer Society, pp: 148-153
- [22] Balakrishna, G., V. Radha, and K. Rao. "ESMP: EXPLORATORY SCALE FOR MALWARE PERCEPTION THROUGH API CALL SEQUENCE LEARNING." journal of theoretical & applied information technology 95.10 (2017).
- [23] Krishna, G. Bala, V. Radha, and K. Venugopal Rao. "ELC-PPW: ensemble learning and classification (LC) by positional patterns weights (PPW) of API calls as dynamic N-grams for malware perception." International Journal of Simulation-Systems, Science & Technology 18.1 (2017).