How to Cite:
Ramesh, G. R., Radha, D., Reddy, A. P., Kumar, C. R., & Kumar, P. R. (2022). An optimized time series data clustering method for NN network. *International Journal of Health Sciences*, *6*(S1), 4672–4683. https://doi.org/10.53730/ijhs.v6nS1.5887

An optimized time series data clustering method for NN network

Goli Raja Ramesh
Assistant Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Hyderabad – 500075, Telangana

D Radha
Associate professor, Department of Computer Science and Engineering, Raghu Engineering College, Visakhapatnam – 531162, Andhra Pradesh

Anantula Pranayanath Reddy
Assistant Professor, Department of Computer Science and Engineering, School of Technology, Gitam (Deemed to be University), Hyderabad – 502329, Telangana

Chanda Raj Kumar
Assistant Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Hyderabad – 500075, Telangana

Prathipati Ratna Kumar
Assistant Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Hyderabad – 500075, Telangana

***Abstract*---**Time series data is a common sort of data that can be found in a variety of fields. Clustering time series data has a wide range of uses and has drawn scholars from several fields. A unique approach for shape-based time series clustering is proposed in this research. Using the complex network principle, it may reduce the bulk of data, enhance efficiency, and not reduce the effects. To begin, a one-nearest neighbour network is constructed using time series items that are comparable. The similarity is measured in this step using the triangle distance. Each node in the neighbour network represents a single time series object, whereas each connection indicates a neighbour relationship between nodes. Second, high-degree nodes are selected and used to cluster. Dynamic time warping distance function and hierarchical clustering algorithm are used in the clustering process. Finally, several studies are carried out using both synthetic and actual data. The findings demonstrate that the suggested approach is both efficient and effective.

Keywords---Clustering, Time series data, Data mining, Distance measure, K-means.

Introduction

There was a rapid increase of IT in the digital world, as well as a massive amount of data collected from a variety of sources. The more difficult task for the business expert is to convert massive amounts of structured data into technical knowledge. Knowledge Discovery in Databases is used to complete this task (KDD). The Process model includes data mining[1]. Big data refers to the use of data analysis techniques to discover previously unknown, relevant patterns and correlations in large data collections. Grouping is one of the massive data mining operations. [2]. The act of collecting a number of observations in order to make the sustainability descriptions within that group much more comparable, as well as the data sets of other groups, is the study of clusters. Because groups are unknown, an unchecked strategy is used to cluster. The goal of grouping in a data collection [3] is to locate thick and vacant sections. Clusters are used in a variety of applications, including system modelling, model analysis, machine intelligence, image classification, image recognition, genomics, data recovery, and data discovery. As a result, it is a prominent topic of research in a variety of fields. Hierarchical techniques, partial approaches, cluster analysis techniques, location is strategic techniques, and modelling classification algorithms are all examples of data clustering. [5].

Hierarchical breakdown

The hierarchical breakdown of datasets is generated using the hierarchy technique. They could be going downhill or uphill. When each data item for one data set is placed in a specific unit, traditional top-down algorithms are split down into smaller groupings. A separate clustering is created between each data collection. Procedures that start from the bottom up begin. Once all groups have become one, it combines close data points consecutively.

Population methods

The population is divided into the number of nodes, as it was previously. They try to identify 'k' groups from a collection of 'N' datasets to meet the minimum requirements: each database should represent a group, and each person should have at least one.

Model-Based Clustering Methods

In tough clustering approaches, every item of data could belong to only one group. Approaches to design believe that a model is the best appropriate for each group, as well as that knowledge is better suited to a theory. They could be hierarchical or partial, depending on the design. With hybrid methodologies, a number of real performance predictions have grown more prevalent in recent years[1]. The euclidean distance measure has traditionally been utilized in research for various clustering algorithms. The efficiency of methods was investigated in this study

employing other significant measuring approaches, such as City Block and Chebyshev. In this paper, a clustering strategy based on K-means ++ and PSO algorithms (K++ PSO) is proposed for multiple measurement clustering. On four benchmark issues, such as the evaluation of teacher assistants, the heart, seedlings, breast cancer, and synthetic cancer, the K++ PSO approach produces strong group results.

Review of Literature

S. Akojwar and P. Kshirsagar. (2016)[24]PSO and Tabu searching K-Harmonic Data Optimization Algorithms were developed. For quantitative Principle Component Analysis (PCA), hybrid K-means grouping and Particle Swarm Optimization (PSO) techniques were presented (PCA-KPSO). PSO's global search functions and the fast K-mean technique completion are included in the model[4][8][12].

P. Kshirsagar and S. Akojwar, (2016)[32]K-Harmonic Means, Particle Swarm Optimization, and a GA-compatible numerical approach were developed. The hybrid solution not only helps with the global optimal problem, but it also helps with the slow performance constraint. [5] Chuang et al. (2012) Particle Swarm Optimization for Gaussian Chaotic Clusters has been suggested for improvement.. PravinKshirsagar et.al[4]To create a hybrid artificially intelligent application with optimizations to classify and forecast a diversified dataset with high accuracy, as well as to use numerous approaches for analysis and regression of benchmark functions, which have proven to be useful in all emerging areas. The methods used in numerous studies in computer security machine learning technology were effective in achieving more accurate conclusions using various quality factors.

Gavrilovet. al. (2005) The two time series being compared are usually taken at the same interval, although their lengths may or may not be identical. Raw data can be directly used by clustering algorithms. Prior to clustering, however, the time series data are usually preprocessed. Principle component analysis piecewise aggregate approximation, discrete Fourier transformation discrete wavelet transformation (sudhirakojwar et.al. 2014) The modified data are then fed into a clustering algorithm as inputs. Typically, raw data transformation can enhance efficiency by reducing data dimensions or improving clustering effects by smoothing the trend and emphasizing the characteristic traits. The function that measures the similarity between the data being compared is an important part of clustering. Many different measures have been used in the practice and research of clustering time series, including Euclidean distance, Pearson's correlation coefficient, short time series distance, dynamic time warping probability-based distance, KL distance, The two time series being compared are usually taken at the same interval, although their lengths may or may not be identical.

Time Series Clustering

Clustering is a form of unsupervised learning problem in which the goal is to detect similarities between distinct data points and group them together in such a way that data points in the same cluster are more similar to each other than data points in other clusters[4][8]. It's one of the most important aspects of exploratory

data mining, and it's employed in a variety of domains like bioinformatics, pattern recognition, image analysis, and machine learning, among others.Because each data point is an ordered sequence, clustering various time series into similar groups is a difficult process. Flattening time series into a table with a column for each time index (or aggregate of the series) and applying standard clustering algorithms like k-means is the most popular technique to time series clustering. (K-means clustering is a typical clustering approach that divides samples into k groups and minimizes the sum-of-squares in each cluster.). Frequency classification isa technique of grouping, and grouping is dependent on the type or objective of both the dataset and the proposed technique's decision, as well as the nature of the dataset. Whether the data are discerning or re-valuation, chosen, regular or non-uniform, simple or multimodal, but whether the historical data appear to be of similar or uniform duration, distinction may be formed in terms of logistic regression information. Data acquired from the sample should be changed to uniform data instead of using universal clustered techniques. This can be done using a wide range of alternatives. [10]. To compile different types of data from regression analysis, different methodologies were created. When their differences are considered, it is clear that they are all attempting to change current cluster methods in order to process current data sets or convert records to static data in order to make straighter use of current methods for clustered data format [34]. Actual data for the time are frequently utilized straight in the first methodology, thus the original data method, and a crucial alteration is the replacement of a distance metric for data structure with a data set measurement that is adequate [35]. To obtain selected features or modelling parameters, the last technique first translates raw data for the period into a lower-dimensional column or into particular design variables, and then employs a conventional clustering technique, nicknamed the features and prototype technique [10]. Figure 1 depicts the three techniques: raw, functionality, and design. It's worth noting that a moving splinter group of the prediction model learned the framework and also employed input variables for grouping [36] without requiring any other schedulingtechnique.

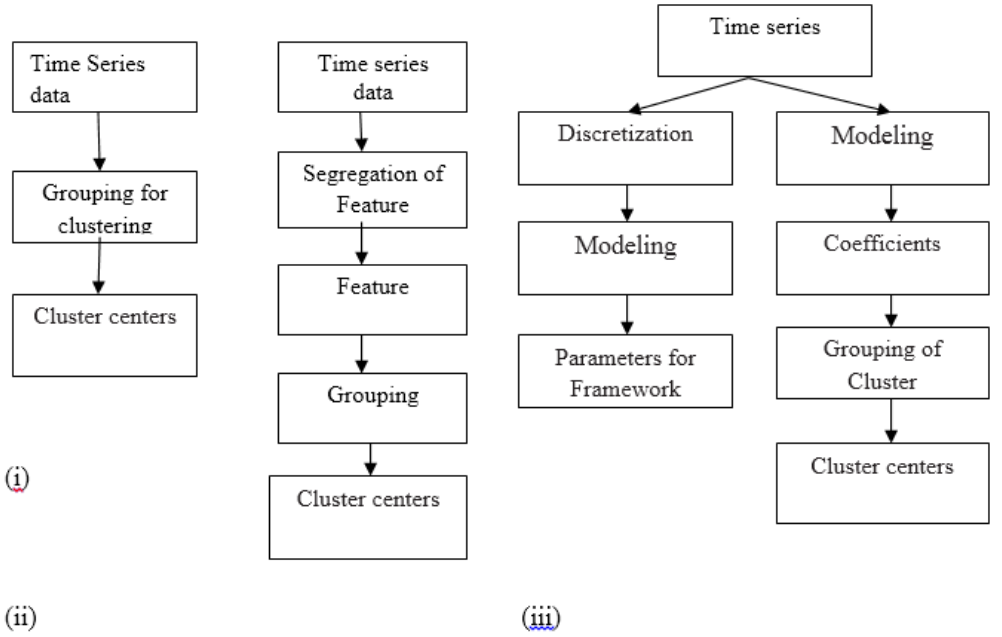

Figure 1. Clustering methods of 3 series: (i) raw-data-based, (ii) feature-based, (iii) model-based.

Distance Measures

An precise calculation in a time series cluster is debatable. When comparing the above methods as a means of determining similarity/differences, it becomes clear that dynamic programming (DP), despite its high cost of implementation, is by far the most reliable and timely strategy. Despite the fact that some constraints are frequently used to reduce waste for these proximity and cosine similarity, careful variable adjustment is required to be efficient and robust[20][24]. As a result, the application of these metrics should distinguish between power and agility. Another consideration is the extent to which distance measuring is effective in large time series analysis sets -. The majority of the studies examined are based on very small quantities of data [21], hence this topic is not from academia. The backdrop subtraction study investigates a number of issues relating to distance measurement. A key challenge is the inaccuracy of the distance measure when solely using the display technique[8] [12]. Several common strategies for time series research, for example, are based on wavelength, and it is feasible to specify similarities between episodes and build actual worth contrasts to use in clusters utilizing a different environment. Feature extraction and DTW[22] are the most widely used methods for establishing trust in time series clusters. According to research, the distances to Euclid are quite large. aggressive in the multiclass classification; nevertheless, the ability of DTW is not to be reduced in similarity measures. Using similarity metrics, clustering is used to measure the similarity or disagreement between any pair of items[23]. Background subtraction between data and centroids can be used to quantify range. The following are the primary characteristics of extracted features. d(a, b) 0 for all a and b d(a, b) = 0 only if a and b are equal d(a, b) = d(y, x) for every a and b d(a, c) d(a, b) + d(b, c) for every

a, b, and c d(a, b) + d(b, c) for every a, b, and c. On the basis of numerous feature vectors such as Euclidean, City Block, and Chebyshev[26], the productivity of K-means + + hybrid employing PSO clustering approaches was investigated in this work[28][32][36].

Euclidean Distance

It is usually used to clusters apps using this distance measure. The L2 or Pythagoras measure is sometimes termed. The following rate is calculated:

$$d(a,c) = \|a - c\| = \sqrt{\sum_{i=1}^{n}(a_i - c_i)^2} \quad (1)$$

City Block Distance

It's also referred to as L1 or Distance measure. The length of the city block from strong functionality a and centre c

$$d(a,c) = \sum_{i=1}^{n}|a_i - c_i| \quad (2)$$

Chebyshev

The radius from Chebyshev is sometimes called the greatest range number. It is a specified measure in a subspace where the distance between any two equals the largest difference here between measurements of each component.

$$d(a_i, a_j) = \max|(a_{ik} - a_{jk})| \quad (3)$$

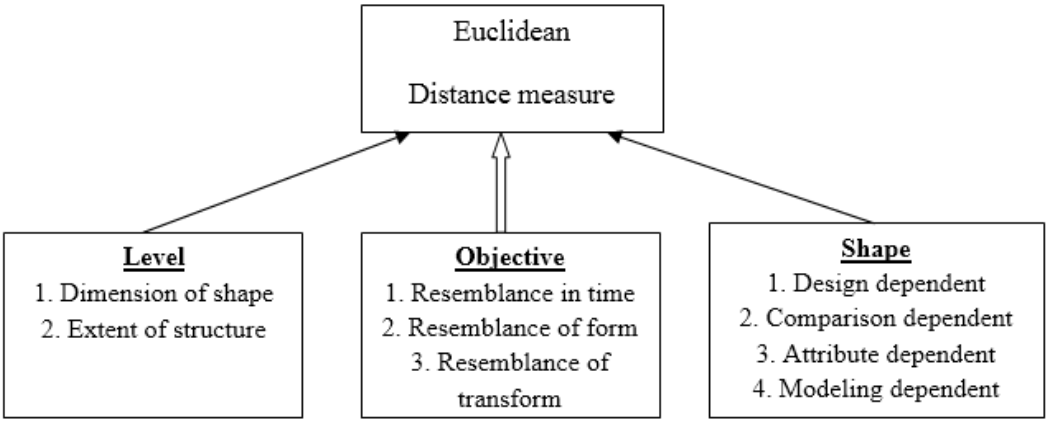

Figure 2. Distance measure approaches

Methodology
k-means clustering algorithm

The K-means method appears to be a non-hierarchical method for sorting items into clusters. The k-means method's basic premise is to calculate the size of groupings to be formed as quickly as possible [25]. In an initialization phase, an I J(i=1,...,v;j=1,...,u) is recognised, where v is the number of nodes to accomplish and u is the collection of attributes. In the application of information (t = 1,..., k; j= 1,..., u), the Centre of each clustered kj is chosen at random. Then, using the

centroid of each cluster, we measure the mobility of each piece of data. Calling ik calculates the distance between data-i and Centre k.
Sort the data into the appropriate clusters. A Centre value can be generated by utilizing Equations 4 to get the minimal value from data that constitute a clustered component.
n=Number of Cluster K c kj=((i=1)naij)/n (4)
n=Number of Cluster K Member
The k-means approach is used to describe cluster phases.
1. Assume that the data matrix A = an ij measures v for u, with i=1, 2... v and j=1, 2,..., m.
2. Compute the components (k), and set the Centre value at random.
3. Equation 1 can be used to calculate the distance between the information and the centroid.
4. Sort the data into the minority and majority clusters using Equation 2.
5. Calculate the new Centre using Equation 4.
6. Repeat steps 3–5 until all data has been transferred to a new group.

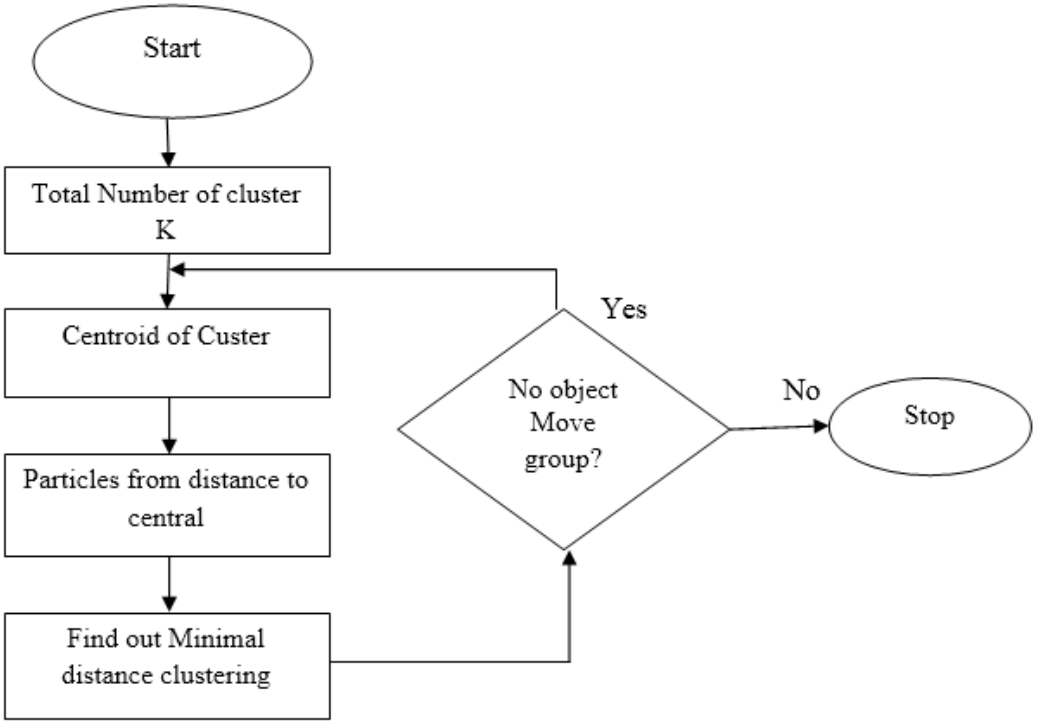

Figure 3. Flowchart of k-means clustering algorithm

Hybrid Algorithm

Hybrid algorithms incorporate even more constructivist approaches. Hybrid approaches are currently popular because they can handle a wide range of practical applications with varying levels of expense and risk. They use algorithms with different properties. The hierarchical clustering algorithms K-Means++ and

Support Vector Machine Algorithms (K++ PSO) were used throughout this investigation. The euclidean distance is the most common metric used in clustered approaches. We have, however, used alternate euclidean distance to examine alternative approaches, such as City Block and Chebyshev.

Result and Discussion

The suggested hybrid methodology on four reference sets consists of data from the Intern evaluation for teachers, hypothyroidism, seedlings, and breast cancer datasets, as well as clustering approaches. The data collection consists of 150 objects divided into three types of groups, each having four characteristics. Three typical seasons of academic performance evaluations and two consecutive semesters with 150 teaching staff are included in the figures. To create ordinary moments, the values were divided into three groups (low, medium, and high) of roughly equal size. In the endocrine data, there are 214 incidences. A T3-resin reception test, complete thyroid hormone replacement serum, total serum triodothyronine, baseline TSH, and picture intensity TSH-value change from the baseline level following 216 mg of testosterone are administered are all included in each case. Each specimen must be classified into one of three categories: Ordinary (151 occurrences), hyperactive (34 occurrences) are the two classes (31instances). There are 211 designs of three different types of wheat in the seed collection: Kame, Rose, and Canada. The area, circumference, compaction, and kernel lengths, as well as kernel thickness, factor of asymmetries, and kernel gap distance, are all geometrical properties of grain kernels in every design. The cancer data set contains 683 recordings with 9 features such as breadth, cell size uniformity, organogenesis sameness, margin adherence, unicellular size, naked nuclei, dull nucleosides, normal cytoplasm, and mitosis, among others. There are two types of instances: normal and anomalous (238 records) (445 records).

Table1
Lists specified characteristics of the data set

Data set	Total Samples	No. of classes	No. of attributes	Size of classes
Intern assessment for teachers	300	3	5	59,50,51
Thyroid	322	3	5	15934,31
Seeds	286	3	7	80,70,70
Tumor of the breast	784	2	9	348,445

Table 2
Evaluation of the objective function score of the 7 clustering methods in the set of data

Distance	K-means	K++	PSO	K_PSO	K++_PSO
Euclidean	1603.563	1605.563	1605.122	1599.193	1594.049
City block	2466.833	2766.627	2678.158	2309.712	2284.583
Chebyshev	1354.935	1330.370	1248.709	1206.684	1201.851

Table 3
Correlation of the best fitness in hypothyroidism set of data for the various cluster methods

Distance	K-means	K++	PSO	K_PSO	K++_PSO
Euclidean	2001.638	2001.638	2250.460	1962.504	1930.335
City block	2985.350	2985.350	3463.442	2929.858	2925.507
Chebyshev	1678.178	1678.178	1752.755	1632.318	1622.337

Table 4
Evaluation of the efficiency of the 7 techniques inside the set of data seedlings

Distance	K-means	K++	PSO	K_PSO	K++_PSO
Euclidean	324.218	310.218	348.983	320.162	310.160
City block	540.622	543.591	662.035	542.684	541.590
Chebyshev	251.506	251.502	296.536	248.017	246.988

Table 5
Evaluation of the efficiency of the 7 techniques inside thebreast cancer data set

Distance	K-means	K++	PSO	K_PSO	K++_PSO
Euclidean	3088.429	3088.429	3841.142	2867.179	2866.432
City block	7126.376	7126.376	7843.117	6612.535	6354.469
Chebyshev	1833.128	1833.128	2079.060	1786.596	1780.629

The programmes with the best results are as follows: The particle frequency (p) is equal to ten. The emotional (c2) and cerebral (c1) components are both 2.0. The momentum has a weight of 0.91 to 1.41. • ALS decreases linearly from 0.91 to 0.41 during the search procedure. The dawn is calculated using the preceding Eq.

$$I\ max =\_max-(max- min)/I\ max \quad (5)$$

Where _max and _min are the starting and end values of the weighted coefficients, respectively; _max= 0.91 and _min= 0.41; Imax appears to be the highest current iteration; Imax appears to be the present current iteration. The maximum number of iterations is 100. For each approach, the trials are conducted in ten different runs. In iteration, the error is 0.00001. (p) The goal of this study is to see how different distance measures are used by data collectors to affect hybrid algorithms. Every strategy is evaluated over the course of 100 rounds and ten separate trials. In this study, basis functions are used to assess the success of clustering clustered dataclusters.

Conclusion

The K-means algorithms are easily restricted to a local optima and are vulnerable to starting values and injuries. The findings of K-means++ are superior to those of other approaches. PSO is a population of probabilistic supervised classifiers. This hybrid technique improves the efficiency of grouping. The distance measure is frequently employed in clustering algorithms. In this thesis, K++ PSO is built

using a variety of distance measures, including City Block and Chebyshev. The best fitness is used to evaluate the comparability of individual approaches. The suggested technique is compared to four established scheduling son techniques, including teacher assistant evaluation, hypothyroidism, seedlings, and breast tumors through the application of various distance measures in a fictitious dataset. Experiments have shown that the K++ PSO method's genetic algorithm is superior than the other techniques K-means, K-means++, PSO, K-PSO, and K-med PSO. In comparison to other feature vectors, the proposed approach produces outstanding results for the Chebyshev length.

References

[1] Aghdasi, T., J. Vahidi and H. Motameni, 2014. Kharmonicmeans data clustering using combinationof particle swarm optimization and tabu search. Int.J. Mech. Electr. Comput. Technol., 4(11): 485-501.

[2] Sethi, C. and G. Mishra, 2013. A linear PCA basedhybrid K-means PSO algorithm for clustering largedataset. Int. J. Sci. Eng. Res., 4(6): 1559-1566.

[3] Danesh, M., M. Naghibzadeh, M.R.A. Totonchi,M. Danesh, B. Minaei and H. Shirgahi, 2011. Dataclustering based on an efficient hybrid of Kharmonicmeans, PSO and GA. In: Nguyen,N.T. (Ed.), Transactions on CCI IV. LNCS 6660,Springer-Verlag, Berlin, Heidlberg, pp: 125-140.

[4] P. R. Kshirsagar, H. Manoharan, F. Al-Turjman and K. Kumar, "DESIGN AND TESTING OF AUTOMATED SMOKE MONITORING SENSORS IN VEHICLES," in *IEEE Sensors Journal*, doi: 10.1109/JSEN.2020.3044604.

[5] Chuang, L.Y., Y.D. Lin and C.H. Yang, 2012. Animproved particle swarm optimization for dataclustering. Proceeding of InternationalMultiConference of Engineers and ComputerScientists (IMECS, 2012). Hong Kong, Vol. 1,March 14-16.

[6] Li, Y.R., Z.Y. Yong and Z.C. Na, 2013. The K-meansclustering algorithm based on chaos particleswarm. J. Theor. Appl. Inform. Technol., 48(2):762-767.

[7] Rai, P., & Singh, S. (2010). A Survey of Clustering Techniques. *International Journal of Computer Applications IJCA*, *7*(12), 1–5.

[8] S. Sundaramurthy, S. C and P. Kshirsagar, "Prediction and Classification of Rheumatoid Arthritis using Ensemble Machine Learning Approaches," *2020 International Conference on Decision Aid Sciences and Application (DASA)*, 2020, pp. 17-21, doi: 10.1109/DASA51403.2020.9317253.

[9] Zakaria, J., Rotschafer, S., Mueen, A., Razak, K., & Keogh, E. (2012). Mining Massive Archives of Mice Sounds with Symbolized Representations. *SIGKDD* (pp. 1–10).

[10] M. Kumar, N.R. Patel, J. Woo, Clustering seasonality patternsin the presence of errors, Proceedings of KDD '02, Edmonton,Alberta, Canada.

[11] H.Kremer,P.Kranen,T.Jansen,T.Seidl,A.Bifet,G.Holmes, B. Pfahringer,Aneffectiveevaluationmeasureforclusteringon evolvingdatastreams,in:Proceedingsofthe17thACMSIGKDD internationalconferenceonKnowledgeDiscoveryandDataMining, 2011,pp.868–876.

[12] H. Manoharan, et al., "Examining the effect of aquaculture using sensor-based technology with machine learning algorithm", Aquaculture Research, 51 (11) (2020), pp. 4748-4758.

[13] C.-P.P.Lai,P.-C.C.Chung,V.S.Tseng,Anoveltwo-levelclustering method fortimeseriesdata analysis,ExpertSyst.Appl.37(9)(2010) 6319–6326.

[14] X. Zhang,J.Liu,Y.Du,T.Lv,Anovelclusteringmethodontimeseries data, ExpertSyst.Appl.38(9)(2011)11891–11900.

[15] J.Zakaria,A.Mueen,E.Keogh,Clusteringtimeseries using unsupervised-shapelets, in:Proceedingsof2012IEEE12thInternational Conference onDataMining,2012,pp.785–794.

[16] PravinKshirsagar et.al (2016), "Brain Tumor classification and Detection using Neural Network", DOI: 10.13140/RG.2.2.26169.72805.

[17] R. Darkins,E.J.Cooke,Z.Ghahramani,P.D.W.Kirk,D.L.Wild, R.S.Savage,AcceleratingBayesianhierarchicalclusteringoftimeseries datawitharandomizedalgorithm,PLoSOne8(4)(2013)e59795.

[18] O.Seref,Y.-J.Fan,W.A.Chaovalitwongse,Mathematical programming formulationsandalgorithmsfordiscretek-medianclustering of time-seriesdata,INFORMSJ.Comput.26(1)(2014)160–172.

[19] S. Ghassempour,F.Girosi,A.Maeder,Clusteringmultivariatetime series usinghiddenMarkovmodels,Int.J.Environ.Res.PublicHealth 11(3)(2014)2741–2763.

[20] Sudhir G. Akojwar, Pravin R. Kshirsagar, "Performance Evolution of Optimization Techniques for Mathematical Benchmark Functions". *International Journal of Computers,* **1**, 231-236,2016.

[21] S. Aghabozorgi,T.YingWah,T.Herawan,H.A.Jalab,M.A.Shaygan, A.Jalali,Ahybrid algorithm forclusteringoftimeseriesdatabased on affinitysearchtechnique,Sci.WorldJ.2014(2014)562194.

[22] S.Aghabozorgi,T.Y.Wah,A.Amini,M.R.Saybani,Anewapproachto present prototypesinclusteringoftimeseries,in:Proceedingsofthe 7th InternationalConferenceofDataMining,vol.28,no.4,2011, pp. 214–220.

[23] F.PetitjeanA.Ketterlin,P.Gançarski,Aglobalaveragingmethodfor dynamic timewarping,withapplicationstoclustering,Pattern Recognit44(3)(2011)678–693.

[24] S. Akojwar and P. Kshirsagar, "A Novel Probabilistic-PSO Based Learning Algorithm for Optimization of Neural Networks for Benchmark Problems", Wseas Transactions on Electronics, Vol. 7, pp. 79-84, 2016.

[25] X.-G. Xia, X. Ye, J. Zhang, Optimal metering plan of measurement andverification for energy efficiency lighting projects, in: Energy EffciencyConvention (SAEEC), Southern African, 2012, pp. 1–8.

[26] J. P. Noonan, "Time series expression analyses using RNA-seq:a statistical approach,"*BioMed Research International*, vol. 2013,Article ID 203681, 16 pages, 2013.

[27] T.Rakthanmanon, E. J. Keogh, S. Lonardi, and S. Evans, "MDLbasedtime series clustering," *Knowledge and Information Systems*,vol. 33, pp. 371–399, 2012.

[28] PravinKshirsagar, NagarajBalakrishnan&Arpit Deepak Yadav "Modelling of optimised neural network for classification and prediction of benchmark datasets", Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 8:4, 426-435, DOI: 10.1080/21681163.2019.1711457,2020.

[29] K. Xu, Y. Jiang, M. Tang, C. Yuan, and C. Tang, "PRESEE: anMDL/MML algorithm to time-series stream segmenting," *TheScientific World Journal*, vol. 2013, Article ID 386180, 11 pages,2013.

[30] S. Aghabozorgi, T. Y. Wah, T. Herawan, H. A. Jalab, M. A.Shaygan, andA. Jalali, “Ahybrid algorithm for clustering of timeseries data based on affinity search technique,” *The ScientificWorld Journal*, vol. 2014, Article ID 562194, 12 pages, 2014.

[31] N. Madicar, H. Sivaraks, S. Rodpongpun, and C. A. Ratanamahatana,“Parameter-free subsequences time series clusteringwith various-width clusters,” in *Proceedings of the 5th InternationalConference on Knowledge and Smart Technology (KST ’13)*,pp. 150–155, 2013.

[32] P. Kshirsagar and S. Akojwar, “Optimization of BPNN parameters using PSO for EEG signals,” ICCASP/ICMMD-2016. Advances in Intelligent Systems Research. Vol. 137, Pp. 385-394,2016.

[33] S. Aghabozorgi and Y. W. Teh, “Stock market co-movementassessment using a three-phase clustering method,” *ExpertSystems with Applications*, vol. 41, no. 4, part 1, pp. 1301–1314,2014.

[34] V. Niennattrakul, D. Srisai, and C. A. Ratanamahatana, “Shapebasedtemplatematching for time series data,” *Knowledge-BasedSystems*, vol. 26, pp. 1–8, 2012.

[35] Liu Ni &Jinhang,"The analysis and research of clustering algorithmbased on PCA", IEEE International conference on electronicmeasurement and instrument,20-22,Oct 20 17.

[36] Kshirsagar, P.R., Akojwar, S.G., Dhanoriya, R, “Classification of ECG-signals using artificial neural networks”, In: Proceedings of International Conference on Intelligent Technologies and Engineering Systems, Lecture Notes in Electrical Engineering, vol. 345. Springer, Cham (2014).