# Weighted Clustering Ensemble with Base Clustering Frequency and Diversity

*Arko Banerjee*
*arko.banerjee@gmail.com*
*Biju Patnaik University of Technology, Odisha*
*Suvendu Chandan*
*Nayak suvendu2006@gmail.com*
*GITA Autonomous College Bhubaneswar, Odisha*
*Chhabi Rani Panigrahi*
*panigrahichhabi@gmail.com*
*Rama Devi Women's University, Bhubaneswar*

## ABSTRACT

*Clustering Ensemble, also referred as Consensus Clustering, is a tool for enhancing the reliability and stability of data clustering by aggregating the base clusterings obtained by different clustering algorithms in the input ensemble. This study introduces a novel ensemble selection strategy for establishing consensus clustering. Our strategy avoids looking at the entire population of base clusterings in the ensemble in order to establish a quality consensus by carefully selecting a few base clusterings. The experimental results reveal that the suggested method's consensus clustering surpasses some other well-known clustering ensemble methods in terms of clustering accuracy for diverse data sets.*

***Keywords*** *Ensemble Clustering; Consensus; Frequency; Diversity*

***Reference*** *to this paper should be made as follows**:** A. Ben-Hur, A. Elisseeff and A. Guyon, "A stability based method for discovering structure in clustered data", Pacific Symposium on Biocomputing, 2002, pp. 6-17.*

## I. INTRODUCTION

Consensus Clustering is a technique in which the base clusterings of an ensemble a data are optimally combined to generate a qualitatively enhanced clustering. The majority of the base clusterings that make up the ensemble can be formed overly, and not all of them accurately represent the data's underlying structure. As a result, when forming a consensus, all clusterings are not need to be considered. In such a setting, the question naturally arises as to whether certain of the ensemble's base clusterings should be used selectively. The purpose of this study is to address the issue. The aim is to develop a mechanism for picking a subset of base clusterings that has a consensus that is on a par with or better than that of the ensemble as a whole. Each clustering in this method can be thought of as an expert's judgment, representing the possibility of an element fitting into a cluster. In our work, we consider the weighting of expert judgments in terms of the diversity and frequency of the relevant base clustering. The opinions of experts who regularly concur are regarded to be more qualified than those who concur infrequently. On the other hand, if one expert's view differs from that of other experts, her judgement should not be dismissed because it may contain vital insights into the structure of the data that other experts have missed. As a result, frequently occurring base clusterings that are varied in comparison to the whole ensemble may include useful information for the ultimate consensus. This hypothesis has been explored empirically for several types of data in this paper. The remainder of the paper is arranged as follows: Section 2 contains an overview of the pertinent literature. Section 3 discusses the weighing technique for a clustering that takes frequency and diversity into account. Chapter 4 details the proposed ensemble selection technique. Experiments in Section 5 demonstrate that the approaches outperform other widely used consensus clustering algorithms. Finally, Chapters 6 and 7 summarise the contribution of the paper.



Published by: International Associations of Professionals and Technical Teachers (IAPATT)

## 2. RELATED WORK

The following is a description of the clustering ensemble problem: S = {s1, s2, …, sn} is the set of n data points.. A clustering P on S is described as a partition P = {C1, C2, …, Ck} where each set Ci $\subseteq$ S, also known as a cluster, is such that they are mutually exclusive i.e. Ci $\cap$ Cj = $\varnothing$ and exhaustive i.e. $\cup$Ci = S. The distinction between any random partition and a clustering of S is that the clustering is such a partition that attempts to group comparable (closer) data points together in a single set (cluster) and dissimilar data points together in a separate set (cluster). A base clustering can be obtained by running a clustering algorithm on S. A collection of base clusterings P = {P1, P2, …, PT} can be produced by repeatedly running multiple clustering algorithms on S several times (say T times) or by repeatedly running the same method with varying initial values. The objective is to discover a definite clustering P * = {C * 1, C * 2, …, C * k} that maximizes the consensus function mapping P to a final clustering P*. This P* is the consensus clustering, which is a median of P1, P2,..., and PT. In the literature, various works on consensus functions have been done. In this context, a great number of research papers on various areas of consensus clustering, such as novel algorithms, theoretical work [27], and applications, have recently been published [17] [18] [19] [20] [21] [22] [23] [24] [25] [26]. Fred et al. [4] generate an ensemble of clusterings using many iterations of the k-means method, which they then integrate into a co-association matrix. Following that, the coassociation matrix is used to construct the end clustering ensemble using a hierarchical single-link approach. According to Topchy et al. [12], the clustering ensemble problem can be formulated as a maximum likelihood estimation problem that can be solved using the EM algorithm. Caruana et al. [2] proposed that the ensemble be drawn from a trained model library. Gionis et al. [6] propose a theoretical framework for consensus clustering and a few consensus algorithms that are theoretically guaranteed to keep the final consensus's quality. Topchy et al. [13] describe two approaches that incorporate plurality voting and a partition space metric. Strehl and Ghosh [11] formulate the consensus clustering problem as a graph optimization problem and define the Normalized Mutual Information (NMI) of the ensemble of clusterings as a metric of consensus clustering quality. Nguyen and Caruana [16] propose the Iterative Voting Consensus (IVC) technique and its two variants for discovering clustering ensembles. By aiming to minimize/maximize a disagreement/agreement criterion function, the majority of the approaches listed above attempt to locate a clustering that is similar to the input ensemble of clusterings. Due to the disparate ways in which they generate the criterion function and the proposed heuristic for improving it, the results of various consensus clustering algorithms vary. All of these strategies take into account the entire ensemble of clusterings in order to arrive at a consensus clustering. However, many redundancies as well as diversity in constituent clusterings can be found in an ensemble. As a result, strategically selecting a subset of the whole ensemble of clusterings may improve the quality of the consensus outcome. Fern et al. refer to this as the Ensemble Selection Problem [3]. They provide three ensemble selection algorithms depending on the quality and variability of clusterings in the input ensemble. The first technique, Joint Criterion, develops a shared objective function based on diversity and quality. Cluster and Select (CAS) is the second approach, which determines a clustering of the clusterings in the ensemble based on a similarity measure. It discovers the consensus of all representative clusterings by selecting one representative clustering from each cluster of clusterings. The third technique, Convex Hull, uses the convex hull idea to determine the consensus of vertices in the ensemble expressed by average quality and diversity of clusterings. Cluster and Select outperforms the other two recommended approaches in the paper's experiments. III. CLUSTERING WEIGHT BASED ON DIVERSITY AND FREQUENCY The notions of variety and frequency of base clusterings in an ensemble are used in this part to present a novel approach of ensemble selection procedure. The ensemble of base clusterings P is supposed to be overly created by repeated iterations of the same or different algorithms. Because P is overly created, it's more possible that some of the clusterings will have numerous copies. Without losing generality, Assume that E is the set containing all (say, r) unique clusterings of P , where E= {P1, P2… Pr}. The terms listed below are defined. A. Definition 1 (Frequency) The number of copies of Pi in P is called the frequency of Pi (denoted as νi) in E. B. Validity Measures Numerous validity or similarity metrics have been proposed in the literature to compare two clusterings [1][5][7][8][10][11][12][14]. We use the Adjusted Rand Index as a measure of similarity between two partitions in our empirical section. It is important to note, however, that the strategy suggested in this paper works with any metric or non-metric similarity measure and is not dependent on any specific similarity measure. In most cases, similarity measurements return a maximum value of 1 in the event of two identical clusterings and a minimum value of 0 in the case of absolute disagreement between two clusterings. Definition 2 (Mean Similarity Index) The mean Similarity index (MSI) of a clustering Pi in E represents Pi 's average similarity to all clusterings in E and is indicated by MSI(Pi ,E) where, r i i j j=1 1 MSI(P ,E)= SI(P ,P ) (r -1)∑ (3) Here, SI(Pi ,E) is a similarity measure index such as the Adjusted Rand Index. C. Definition 3 (Diversity Index) In E, the diversity measure index of Pi is defined as follows. DI(P ,i i E) = 1- MSI(P ,E) (4) If Pi clusters in E as an outlier, DI (Pi , E) is usually large (near to 1), and if Pi tends to the median of E, DI (Pi , E) takes the smallest value (close to 0). D. Definition 4 (Weight of a clustering in E) In E, weight of clustering Pi is defined as, i i i j j ν w = DI(P ,E)×∑ν (5) Clustering weights wis are critical for prioritising certain clusterings within the ensemble, taking both diversity and frequency into account; in terms of the experts-analogy, this implies that a point of view that is endorsed by a sizable group of experts but deviates from the thoughts of certain other experts does have a great significance.

## 3. PROPOSED ALGORITHM

The proposed method begins with E, which is composed of the unique base clusterings of P. These unique clusterings are listed in non-increasing order of their weights wi . With the help of any known clustering ensemble algorithm, the first k sorted clusterings are picked for determining the consensus. The Ensemble Selection algorithm's pseudo code is shown below. Algorithm: ESFD INPUT: P (Collection of base clusterings), k (size of new ensemble) OUTPUT: Consensus Clustering P* METHOD: 1. Compute E, which is the collection of unique base clusterings in P and their associated frequencies, νi's according to definition-1. 2. Using definition-2, calculate the Mean Similarity Index (MSI) for each clustering in E. 3. For each clustering in E, compute the Diversity Index (DI) using definition-3. 4. Calculate the weight wi by definition-4 for each clustering Pi in E. 5. Arrange E's clusterings in descending order of wi 6. Assume E' is the collection of the top k sorted clusterings in E. 7. Using any clustering ensemble approach, calculate clustering ensemble on E'. If a clustering has a higher weight, it is more likely to provide reliable and innovative information in the consensus generation process. Clusterings with a lower weight, on the other hand, may be deemed outliers in the ensemble. The proposed approach is referred to as Ensemble Selection using Frequency and Diversity throughout the remainder of this paper (ESFD). The ESFD technique aims to create an ensemble with a high degree of variety and low redundancy. We use three examples to demonstrate the relevance of considering both frequency and diversity rather than these two parameters alone. Case1: ensemble selection based solely on diversity Case2: ensemble selection based solely on frequency Case3: ensemble selection based on both frequency and diversity (ESFD). Figure 1 illustrates the results of all three scenarios using yeast data [15]. We create an ensemble of 100 clusterings of yeast data and apply ESFD in the said order of weights. The validity of the consensus is measured using the Adjusted Rand index (AR) in each scenario, and the results are plotted by changing the ensemble size. The size of the ensemble and the value of AR are depicted on the x-axis and y-axis, respectively, in Figure 1. The subset of the ensemble considered by eliminating the last 'a' clusterings in the ensemble is denoted by x=a. Figure 1 shows that the weight that takes into account both frequency and diversity gets a greater maximum and average AR value than the other two situations. Other wellknown data, like as Ecoli, Segmentation, and Chart [15], have also shown similar traits, that are discussed in the experimental section

.

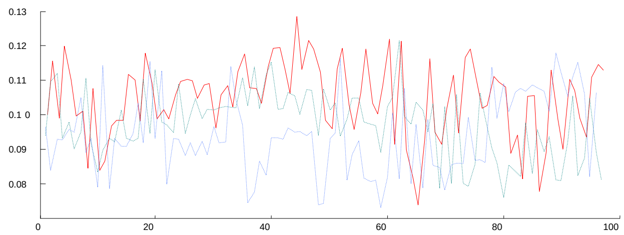



Fig 1: AR values plotted against different ensemble sizes. Green, blue, and red lines show ESFD weighted with diversity, frequency, and frequency-diversity, respectively.

### 3.1 EXPERIMENTAL RESULT

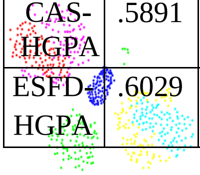
This part begins by describing the data sets used in the empirical evaluation and the basic experimental setups. The ESFD results are then compared to those obtained using the Cluster and Select (CAS) technique [3]. The experiments make use of real-world data sets obtained from the University of California, Irvine's machine learning data repository [15]. It's worth mentioning that all datasets have supervised categorization. The datasets are selected, since some are popular in the field of cluster analysis while others are less well-known but present a significant challenge to conventional clustering algorithms due to their high dimensionality[29]. Using various random initializations, the K-means algorithm is used to generate an ensemble of base clusterings of data. Individual set of data are divided into three different ensembles E1, E2 and E3, each of which contains 200 clusterings. Each set of data is subjected[30] to three different categories of testing. In the first category of test, CSPA and HGPA are used on the whole ensemble[31], whereas in the second and third category of tests, CAS and ESFD have been used for clustering selection[32] on the whole ensemble, and then CSPA and HGPA are used to obtain final consensus, respectively. The quality of experimental results is determined by their similarity to the ground truth class labels provided with the data. The adjusted rand index (AR) is applied to compare consensus and ground truth clustering . The AR is a variant of the Rand Index, a well-known validity metric that statistically corrects for the latter's lack of invariance when random clusterings are chosen [7]. AR has been chosen as a validity measure since it is straightforward to compute and has been used in a large number of previous studies. Let Pi and Pj be two clusterings. Pi = {C1, C2, …, Ck} and Pj = {C′1, C′2, …, C′k}. Then AR is defined as [11]

$$AR(P,P_{i\,j}) = \frac{a^0 - a^3}{\frac{1}{2}(a_1 + a_2) - a_3} \qquad (1)$$

where, j i k k st s t 1 2 0 1 2 3 s,t r=1 s=1 M m m' a × a a = , a = , a = , a = 2 2 2 n 2 ⎛⎞⎛⎞⎛⎞| || | | |⎛⎞⎝⎠⎝⎠⎝⎠| |⎝⎠ ∑ ∑ ∑ (2) Here, Mst =| Cs ∩C′t |, ms=| Cs | and m′t=| C′t |, where Cs and C′t are clusters of the clusterings Pi and Pj , respectively. If Pi and Pj are identical clusterings, the AR returns a maximum value of 1, and a minimum value of 0 if Pi and Pj are random clusterings. In tables II, III, and IV, the term 'CSPA' refers to a clustering ensemble that uses CSPA across the entire ensemble; 'CAS-CSPA' refers to a clustering ensemble that uses CSPA on a subset of clusterings chosen by CAS from the the whole ensemble; similarly 'ESFD-CSPA' refers to a clustering ensemble that uses CSPA on a subset of clusterings chosen by ESFD from the whole ensemble. Likewise, terms for HGPA are defined. On to eliminate the user-defined parameter 'k', which specifies the number of sorted clusterings chosen in the ESFD technique, in 'ESFD-CSPA', the subset of clusterings with a weight (wi) greater than the ensemble mean weight is chosen. In the Chart, Segmentation, Ecoli, Yeast, and Iris datasets, ESFD surpasses CAS. ESFD performs similarly to CAS in the Glass and Wine datasets (not reported). CAS surpasses the ESFD when used with vehicle data. According to Table 1, ESFD beats CAS on Chart and Segmentation data utilising both CSPA and HGPA. Concerning ground-truth clustering, Figure 3 illustrates the consensus clusterings produced by ESFD and CAS using Chart data. Sammon Mapping is used to reduce the dimension of Chart data [9]. The accuracy of projection is determined using the stress function indicated at the beginning of Figure 3

TABLE 1. SEGMENTATION DATA AND CHART DATA OUTCOMES

| | **Chart** | | | **S egmentation** | | |
|---|---|---|---|---|---|---|
| | ***E1*** | ***E2*** | ***E3*** | ***E1*** | ***E2*** | ***E3*** |
| CSPA | .6392 | .6459 | .6520 | .3605 | .3551 | .4779 |
| CAS-CSPA | .7416 | .6730 | .6889 | .4907 | .3797 | .4805 |
| ESFD-CSPA | .7815 | .7210 | .7027 | .5169 | .4061 | .4869 |
| HGPA | .4572 | .4724 | .4819 | .1330 | .2479 | .2394 |
| CAS-HGPA | .5891 | .5511 | .5533 | .3458 | .3887 | .3503 |
| ESFD-HGPA | .6029 | .5881 | .5591 | .3567 | .3989 | .3563 |



The results of ESFD-CSPA and CAS-CSPA on Chart data are shown in Figure 4, where the X and Y axis represent the ensemble size and the value of the Adjusted Rand index, respectively. The subset of the ensemble considered by eliminating the last 'a' clusterings in the ensemble is denoted by x=a. Figure 4 shows that ESFD not only has a higher maximum AR value than CAS, but also has a higher average AR value over all X-values. With CSPA, ESFD also outperforms CAS on Segmentation data, as shown in Figure 5. ESFD likewise achieves greater maximum[29] and average AR values than CAS in this case.

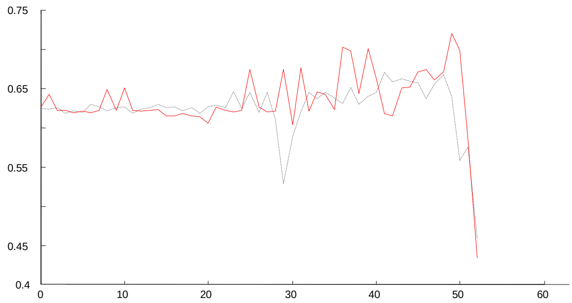



Fig 4. On the second ensemble ($E_2$) of chart data, the black line depicts CAS

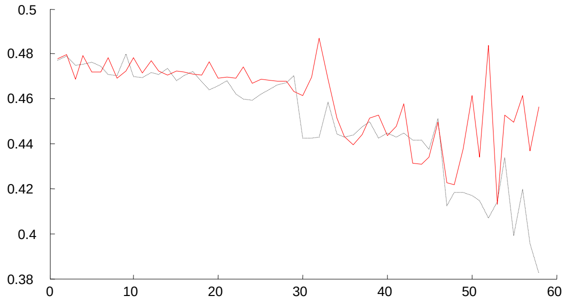



Fig 5. On the third ensemble (E3) of Segmentation data, the black line depicts CAS-CSPA and the red line shows ESFD-CSPA, respectively.

Table 2 shows that on both the Ecoli and Yeast datasets, ESFD-CSPA/HGPA produces superior results than CAS. Table 3 reveals that ESFD-CSPA/HGPA outperforms CAS on Iris data, whereas CAS outperforms ESFD on Vehicle data.

TABLE 2. YEAST DATA AND ECOLI DATA OUTCOMES

| | **Ecoli** | | | **Yeast** | | |
|---|---|---|---|---|---|---|
| | ***E1*** | ***E2*** | ***E3*** | ***E1*** | ***E2*** | ***E3*** |
| CSPA | .2104 | .2956 | .2827 | .0972 | .1008 | .0869 |
| CAS-CSPA | .2561 | .3284 | .3098 | .1018 | .1274 | .1111 |
| ESFD-CSPA | .2609 | .3366 | .3118 | .1128 | .1287 | .1127 |
| HGPA | .3454 | .3313 | .3116 | .1013 | .1153 | .099 |
| CAS-HGPA | .3623 | .3521 | .3327 | .1220 | .1226 | .1101 |
| ESFD-HGPA | .3673 | .3593 | .3338 | .1224 | .1244 | .1201 |

TABLE 3. IRIS DATA AND VEHICLE DATA OUTCOMES

| | **Iris** | | | **Vehicle** | | |
|---|---|---|---|---|---|---|
| | ***E1*** | ***E2*** | ***E3*** | ***E1*** | ***E2*** | ***E3*** |
| CSPA | .7139 | .7139 | .7139 | .1206 | .1319 | .1149 |
| CAS-CSPA | .7414 | .7414 | .7414 | .1451 | .1519 | .1211 |
| ESFD-CSPA | .8014 | .8014 | .8014 | .1344 | .1398 | .1159 |
| HGPA | .6151 | .6151 | .6151 | .0988 | .1010 | .0991 |
| CAS-HGPA | .7221 | .7221 | .7221 | .1133 | .1235 | .1011 |
| ESFD-HGPA | .7460 | .7460 | .7460 | .12015 | .1133 | .0996 |

## 4 CONCLUSION

We demonstrate in this research that carefully selecting a subset of the population of clusterings from an ensemble can result in a better consensus than evaluating the whole ensemble for consensus. According to our proposed method, weight of clustering in combination with frequency and diversity may encode considerable information for a meaningful consensus. When the actual number of clusters is considered, tests on a variety of benchmark datasets demonstrate that our method aids in improving clustering accuracy.

## 5 FUTURE WORK

We may test our proposed method with alternative diversity measures in the future [8], [1], [14], [5]. We would also like to learn more about the theory behind ESFD's good results on a variety of data with various distributions.

## 6 REFERENCES

[1]. A. Ben-Hur, A. Elisseeff and A. Guyon, "A stability based method for discovering structure in clustered data", Pacific Symposium on Biocomputing, 2002, pp. 6-17.

[2]. R. Caruana et al., "Ensemble selection from libraries of models", Proc. 21st International Conference on Machine Learning (ICML), 2004, pp. 137-144.

[3]. X.Z. Fern and W. Lin, "Cluster Ensemble Selection", Statistical Analysis and Data Mining, Vol.1, No.3, 2008, pp. 128-141.

[4]. A. Fred and A. Jain, "Data clustering using evidence accumulation", ICPR, Vol.4, 2002, pp. 276-280.

[5]. A.L.N. Fred and A. Jain, "Robust data clustering", IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Vol.2, 2003, pp. 128-136.

[6]. A. Gionis, H. Mannila and P. Tsaparas, "Clustering aggregation", 21st International Conference on Data Engineering (ICDE'05), 2005, pp. 341-352.

[7]. L. Hurbert and P. Arabie, "Comparing partitions", Journal of Classification, Vol.2, 1985, pp. 193-218.

[8]. W.M. Rand, "Objective criteria for the evaluation of clustering methods", Journal of the American Statistical Association, Vol. 66, No.336, 1971, pp. 846-850.

[9]. J.W. Jr. Sammon,"A Nonlinear Mapping for Data Structure Analysis", IEEE Transactions on Computers, Vol.C-18, No.5, 1969, pp. 401-408.

[10]. D. Steinley, "Properties of the Hubert-Arabic adjusted Rand index", Psychol Methods, Vol.9, 2004, pp. 386-396.

[11]. A. Strehl and J. Gosh, "Cluster ensembles – a knowledge reuse framework for combining partitionings", Proc AAAI, 2002, pp. 93-98.

[12]. Kumar, R., Singh, J.P., Srivastava, G. (2014). Altered Fingerprint Identification and Classification Using SP Detection and Fuzzy Classification. In: , et al. Proceedings of the Second International Conference on Soft Computing for Problem Solving (SocProS 2012), December 28-30, 2012. Advances in Intelligent Systems and Computing, vol 236. Springer, New Delhi. https://doi.org/10.1007/978-81-322-1602-5_139.

[13]. A. Topchy, M. Law, A. Jain and A. Fred, "Analysis of consensus partition in cluster ensemble", IEEE International Conference on Data Mining (ICDM), 2004, pp. 225-232.

[14]. D. L. Wallace, "A Method for Comparing Two Hierarchical Clusterings: Comment", Journal of the American Statistical Association, Vol. 78, No.383, 1983, pp. 569- 576.

[15]. David Aha et al.,"UC Irvine Machine Learning Repository,"1987; www.ics.uci.edu/~mlearn/MLRepository.

[16]. N. Nguyen and R. Caruana. Consensus Clusterings. Proc. 7th IEEE International Conference on Data Mining (ICDM), 2007, pp. 607-612.

[17]. H. Liu, R. Zhao, H. Fang, F. Cheng, Y. Fu,Y. Liu,"Entropy-based Consensus Clustering for Patient Stratification", Bioinformatics (Oxford, England), pp. 1-8, 2017, DOI 33. 10.1093/bioinformatics/btx167.

[18]. Y. Yan, T. Nguyen, B. Bryant, F. C. Harris Jr, ”Robust Fuzzy Cluster Ensemble on Cancer Gene Expression Data”, Proceedings of 11th International Conference on Bioinformatics and Computational Biology, vol 60, pages 120-128, 2019.

[19]. T.A. Geddes, K. Taiyun, N. Lihao, G.B. James, Y.H.Y. Jean, T. Dacheng, Y. Pengyi, “Autoencoder-based cluster ensembles for singlecell RNA-seq data analysis”, BMC Bioinformatics, vol. 20, no. 660, 2019.

[20]. S. Saha,R. Das, P. Pakray,”Aggregation of multi-objective fuzzy symmetrybased clustering techniques for improving gene and cancer classification”, Soft Computing, vol. 22, pp. 5935–5954, 2018.

[21]. J. Rasero, Javier, I. D´ıez, J. M. Cortes, D. Marinazz, S. Stramaglia, “Connectome sorting by consensus clustering increases separability in group neuroimaging studies.” Network Neuroscience, vol. 3, no. 2, pp.325-343, 2019.

[22]. S. Fouad, D. Randell, A. Galton, H. Mehanna, G. Landini,”Unsupervised Superpixel-Based Segmentation of Histopathological Images with Consensus Clustering”, Communications in Computer and Information Science(CCIS), vol. 723, pp. 767-779, 2017.

[23]. I.A. Pestunov, S.A. Rylov, Y.N. Sinyavskiy, V.B. Berikov,”Computationally efficient methods of clustering ensemble construction for satellite image segmentation”, 3rd International conference on Information Technology and Nanotechnology, pp. 194- 200, 2017, DOI: 10.18287/1613-0073-2017-1901-194-200.

[24]. A. Lourenco, C. Carreiras, S. R. Bulo, A. Fred, ”ECG analysis using consensus clustering,” 22nd European Signal Processing Conference (EUSIPCO), pp. 511-515, 2014.

[25]. M. Kheirkhahzadeh, M. Analoui, ”Community Detection in Social Networks Using Consensus Clustering”, Statistics, Optimization and Information Computing, vol. 7, pp. 864-884, 2019.

[26]. L.G.S. Jeub, O. Sporns, S. Fortunato,“Multiresolution Consensus Clustering in Networks.” Scientific Reports, vol. 8, no. 1, 2017.

[27]. S. Vega-Pons, J. Ruiz-Shulcloper, ”A Survey of Clustering Ensemble Algorithms”, IJPRAI, vol. 25, no. 3, pp.337-372

[28]. Anand Singh Rajawat; Romil Rawat; Kanishk Barhanpurkar, "Security Improvement Technique for Distributed Control System (DCS) and Supervisory Control-Data Acquisition (SCADA) Using Blockchain at Dark Web Platform," in Cyber Security and Digital Forensics: Challenges and Future Trends , Wiley, 2022, pp.317-333, doi: 10.1002/9781119795667.ch14.

[29]. S. Srivastava and R. Kumar, "Indirect method to measure software quality using CK-OO suite," 2013 International Conference on Intelligent Systems and Signal Processing (ISSP), 2013, pp. 47-51, doi: 10.1109/ISSP.2013.6526872.

[30]. A. S. Rajawat, K. Barhanpurkar, D. Mukhopadhyay and A. Ghosh, "Scheme Invisible Internet Protocol (I2P) using Blockchain techniques for tracking Darkweb User Activities," 2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON), 2021, pp. 1-5, doi: 10.1109/GUCON50781.2021.9573880.

[31]. Ram Kumar, Gunja Varshney , Tourism Crisis Evaluation Using Fuzzy Artificial Neural network, International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-1, Issue-NCAI2011, June 2011

[32]. A. S. Rajawat, K. Barhanpurkar, D. Mukhopadhyay and A. Ghosh, "Scheme Invisible Internet Protocol (I2P) using Blockchain techniques for tracking Darkweb User Activities,"

2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON), 2021, pp. 1-5, doi: 10.1109/GUCON50781.2021.9573880.

[33]. Ram Kumar, Jasvinder Pal Singh, Gaurav Srivastava, "A Survey Paper on Altered Fingerprint Identification & Classification" International Journal of Electronics Communication and Computer Engineering Volume 3, Issue 5, ISSN (Online): 2249–071X, ISSN (Print): 2278– 4209

[34]. Rajawat, A.S., Rawat, R., Mahor, V., Shaw, R.N., Ghosh, A. (2021). Suspicious Big Text Data Analysis for Prediction—On Darkweb User Activity Using Computational Intelligence Model. In: Mekhilef, S., Favorskaya, M., Pandey, R.K., Shaw, R.N. (eds) Innovations in Electrical and Electronic Engineering. Lecture Notes in Electrical Engineering, vol 756. Springer, Singapore. https://doi.org/10.1007/978-981-16-0749-3_58

[35]. Goyal, S.B., Bedi, P., Rajawat, A.S., Shaw, R.N., Ghosh, A. (2022). Multi-objective Fuzzy-Swarm Optimizer for Data Partitioning. In: Bianchini, M., Piuri, V., Das, S., Shaw, R.N. (eds) Advanced Computing and Intelligent Technologies. Lecture Notes in Networks and Systems, vol 218. Springer, Singapore. https://doi.org/10.1007/978-981-16-2164-2_25

[36]. Kumar, R., Singh, J.P., Srivastava, G. (2014). Altered Fingerprint Identification and Classification Using SP Detection and Fuzzy Classification. In: , et al. Proceedings of the Second International Conference on Soft Computing for Problem Solving (SocProS 2012), December 28-30, 2012. Advances in Intelligent Systems and Computing, vol 236. Springer, New Delhi. https://doi.org/10.1007/978-81-322-1602-5_139