Support Vector Clustering Algorithm for Cell Formation in Group Technology

Prafulla C. Kulkarni, PhD
G.E.S.'s R.H. Sapat College of Engineering,
Management Studies & Research, Nashik

ABSTRACT
In Group technology, parts with similar geometry, function, material and process are grouped into part families and the corresponding machines are grouped into machine cells. In cluster analysis, one seeks to find the natural groupings in the data. One searches for patterns in the data set by grouping it into clusters. The goal is to find an optimal grouping for which the data within clusters are similar, but the clusters are dissimilar to each other. Many techniques exist to group the data into clusters. Recently, Support Vector Clustering (SVC) has been employed for cluster analysis. In SVC algorithm, data points are mapped from data space to a high dimensional feature space using a Gaussian kernel. The smallest sphere enclosing in feature space is mapped back to data space where it forms cluster boundary. Thus clusters are formed. The scale parameter of Gaussian kernel and soft margin constant are the two parameters which determine clustering form. A data set of Group technology is considered for investigating performance of the algorithm. Part families are formed by SVC algorithm for the data set within a reasonable time as demonstrated.

Keywords
Support Vector Clustering, Gaussian kernel, Group Technology

1. INTRODUCTION
Group Technology (GT) is become a mature technique for classifying parts into groups and machines into machine cells. The advantages of grouping of parts and machines are numerous and include increased productivity, reduced costs, reduced material handling, reduced setup times and more. There are three main methods for part classification. The methods can be broadly classified as visual methods, coding and classification methods and production flow analysis. Similarities in the shape and geometry of parts are the main criteria for classification in visual methods. The part families obtained through visual methods highly depend on the experience and preference of classifier. However, visual methods are applicable for the classification of limited number of parts. Classification and coding systems are also the traditional tools used to implement Group Technology; Parts are grouped on the basis of design features such as geometric shape, dimensions, material, accuracy, etc. A numeric, alphabetic or coded symbol is considered for each feature. Arrangement of these symbols represents a unique part. But the disadvantage of this method is that it is time consuming. Also parts of similar size, shape and function may not use the same set of machine tools and other resources. There is no universal standard on the features to be considered in classification. Recent research has therefore emphasized the use of production flow analysis methods. Array based clustering operates by manipulation and reordering of rows and columns of machine-part incidence matrix in order to form a block-diagonal matrix. Bond Energy Algorithm, Rank Order Clustering, Direct Clustering Algorithm, MODROC, Cluster Identification Algorithm are some techniques proposed. Reference [1] compared the performance of these methods. Array based clustering techniques are efficient and simple but they are criticized for some limitations. They use a binary machine-part incidence matrix and do not consider other manufacturing data such as cost data, cell size, etc. these methods are not effective for ill-structured matrices where several exceptional elements exist. Hierarchical clustering techniques are one of the most frequently used methods for grouping. Part families are clustered and then method is reapplied once again to group the machines. These methods use similarity coefficient or distance clusters or distance measure to create a hierarchy of clusters. Given a similarity threshold level, different clusters can be identified. The similarity coefficient measure can incorporate different production-related data and therefore, too many similarity measures have been proposed in literature. Most of them used the generic Jacquard similarity coefficient. A Single Linkage Clustering approach is used to cluster machines. A similarity coefficient that incorporates machine-part incidence matrix, production volume of parts, sequence of operations and the processing times of each operation is used [2]. A comparison of the methods can be found in [3, 4, and 5]. Although, hierarchical clustering methods are easy to use and capable to incorporate different manufacturing data, these are criticized due to arbitrary selection of threshold level. Issues related to duplication of bottleneck machines cannot be easily handled by hierarchical clustering methods. Reference [6] proposed an ideal seed non-hierarchical clustering approach. The problem is initially formed as bipartite graph to obtain an upper limit on the maximum number of independent cells. A modified k-means method is adopted in which the last k data units are selected as the initial seed points. Reference [7] developed improved version called ZODIAC. The initial seed selection can be arbitrary, artificial, representative or natural. It is shown that natural seed forming provides better grouping than artificial and representative methods. The seed selection process is statistically analyzed to ensure that the selected seeds belong to different clusters. Reference [8] proposed improved method called GRAPHICS in which initial seeds are obtained through an assignment model maximizing the similarity between machines. Mathematical programming techniques are applied to get clustering. Mathematical programming provides a tool take into consideration different design and operational objectives and constraints in classifying parts and machines. Several mathematical programming techniques of linear as well as non linear nature such as assignment model, quadratic assignment model, mixed-integer programming, goal programming, multi-objective programming, etc. have been widely applied [9, 10, 11]. They are advantageous due to their potential to incorporate production-oriented data. But these methods are criticized for three limitations. First, these methods group parts and machines in a sequential manner rather than simultaneously due to the non-linearity of objective function.

Second, the number of cells is to be specified in advance. Third, the proposed mathematical models are computationally intractable for realistic applications. In the graph theoretic methods, the cell formation is represented by a graph such as network, bipartite graph, spanning tree, etc. Reference [12] reviewed the cell formation methods with respect to graph theoretic approaches and addressed that this approach is promising for a generalized grouping problem. The major weakness is that different manufacturing data cannot be incorporated in the graph. Various heuristics and Meta heuristics are applied. Researches employed general-purpose methodology like neural networks, genetic algorithms, simulated annealing and taboo search. Comprehensive review of the applications is available [13]. All these methods are developed for the grouping of parts and machines with a various objectives applicable to cellular manufacturing systems. A common weakness with all these methods is that, they implicitly assume that the part families are mutually exclusive and collectively exhaustive i.e. each part can only belong to one part family. In reality there may exist parts whose lineages are much less evident. Fuzzy clustering is one approach for more accurate presentation of uncertain or inexact information. Fuzzy clustering not only reveals the specific part family that a part belongs to, but also provides the degree or grade of membership of part associated with each part family. This information would give the users more flexibility in which part family a part should be actually produced so that the workload among machine cells can be balanced [14,15,19,22]. Fuzzy mathematics for clustering was attempted [21]. Clustering is the unsupervised classification of patterns into groups. But cluster analysis identifies similar and dissimilar features in data and segregate data into homogenous groups based on features. An underlying assumption is that homogeneous clusters exist in raw data. Clustering algorithms to group data points according to various criteria have been extensively studied [16]. Clustering may proceed according to some parametric model, as in the k-means algorithms or by grouping points according to some distance or similarity measure as in hierarchical clustering algorithms. Agglomerative algorithms, k-means algorithm, fuzzy algorithms, BIRCH and CLARANS are a few of the existing clustering algorithms. The k-means algorithm is popular because it is easy to implement. But the algorithm is sensitive to the choice of initial random seeds and the value of k. Many variants have been proposed for overcoming these problems. It is also found that the algorithm is not able to learn arbitrary clusters boundaries; for example, it fails to converge properly for data having concentric clusters. Kernel methods are the basis for Support Vector Clustering. In this paper, a non-parametric clustering algorithm based on support vector machine approach [17, 18] is presented. Data points are mapped to a high dimensional feature space using Gaussian kernel, where support vectors are used to define a sphere enclosing them. The boundary of the sphere forms in data space a set of closed contours containing the data. As the width parameters of the Gaussian kernel are decreased, these contours fit the data more tightly and splitting of contours occurs. The algorithms works by separating clusters according to valleys in the underlying probability distribution, and thus clusters can take on arbitrary geometrical shapes. As in other SV algorithm, outliers can be dealt with by introducing a soft margin constant leading to smoother cluster boundaries. The structure of the data is explored by varying the two parameters. Dependence of Support Vector Clustering Method on these parameters is investigated. The application of it is observed to a data set of Group technology. In section II methodology adopted for SVC algorithms is provided. SVC was applied recently for other problems [20]. Recently, SVC algorithm applied for cellular manufacturing systems [23] A problem from literature is solved by SVC algorithm in section III. Results are compared with other methods. Section IV discusses the usefulness of this SVC algorithm methodology and input parameters required with the conclusion at the end.

2. METHODOLOGY

2.1 Support Vector Clustering Model

This section deals with the research work related to Support Vector Clustering. The first step in the SVC algorithm involves solving an optimization problem to get the smallest hyper sphere in feasible space, which encloses images of the data points. This problem can be cast as a convex quadratic minimization problem.

$$\min_{R,a,\xi_i} \; \mathrm{R}^2 + C\sum \xi_i$$

$$\text{s.t. } \left\|\Phi(x_i) - a\right\|^2 \;\le\; \mathrm{R}^2 + \xi_i \;, \quad \xi_i \;\ge 0 \;\forall i, \qquad (1)$$

Where R is the radius, a is the center of the hyper sphere, ξ_i are the slack variables, Φ is the data space to feature space mapping and C is a hyper-parameter. Higher value of C is lesser the number of points whose images lie outside the hyper sphere and vice versa. The Wolfe dual of the optimization problem is

$$\underset{\alpha_i}{\text{Max}} \sum_i \alpha_i \Phi(x_i) \cdot \Phi(x_i) - \sum_{i,j} \alpha_i \alpha_j \Phi(x_i) \cdot \Phi(x_j)$$

$$\text{s.t. } \sum_i \alpha_i = 1, \quad 0 \le \alpha_i \le C \;\; \forall i, \qquad (2)$$

where α_i, the Lagrange multipliers, are the dual variables. The functional form of mapping $\Phi(x_i)$ does not need to be known since implicitly defined by the choice of kernel function $K(x_i, x_j) \equiv \Phi(x_i) \cdot \Phi(x_j)$ where "$\cdot$" represents dot product. Gaussian kernel used

$$K(x_i, x_j) = \exp(-q\left\|x_i - x_j\right\|^2). \qquad (3)$$

The distance between input pattern and spherical center is computed as

$$R^2(x) = \left\|\Phi(x) - a\right\|^2$$
$$= K(x,x) - 2\sum_i \alpha_i K(x_i, x) + \sum_{i,j} \alpha_i \alpha_j K(x_i, x_j). \qquad (4)$$

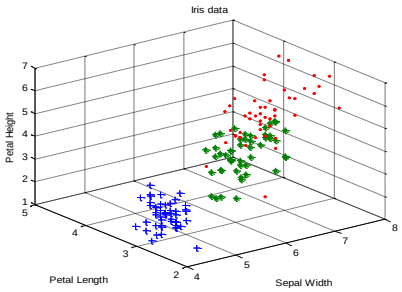
The Karush-Kuhn-Tucker (KKT) conditions are necessary and sufficient for the optimality of the dual, since it is a convex optimization problem [19]. The data points with $0 < \alpha_i < C$ are non-bound support vectors, $\alpha_i = C$ are bounded support vectors and $\alpha_i = 0$ are non-support vector points, which lie inside or on the hyper sphere. These three categories of the partitioned data are important. The non-bound support vectors lie on the hyper sphere and they are q in number. Bounded support vectors lie outside, or on the hyper sphere. For C > 1, there will be no Bounded support vectors. Both these are referred as support vectors denoted by p. In most of the cases p « m, where m are the total data points. Once the dual problem is solved, identification of clusters is done using a simple graphical connected-component method (GCM). Let *f* (x) denote the inference function which evaluates whether the image of a data point x lies inside, on, or outside the hyper sphere. Let LS [x_i, xj] denote the line segment joining xi and xj in the data space. If the image of the LS [x_i, xj] in the feature space does not exist in the hyper sphere then points x_i and xj are said to be connected. Otherwise, they are said to be disconnected. This leads to a natural definition of the adjacency matrix A: $A_{i,j}$ is 1 if the image of LS[x_i, xj] lies entirely within the hyper sphere and 0 otherwise. The matrix A is built taking all points except the Bounded support vectors since there images lie outside the hyper sphere. Clusters are now defined as the connected components of the graph induced by A and labeled with number.

2.2 Effect of parameter in Support Vector Clustering algorithm

SVC algorithm was applied on the Iris data, which is a standard benchmark in the pattern recognition literature, and can be obtained from Blake and Merz (1998) at UCI repository of machine learning databases [24]. The data set contains 150 instances, each composed of four features of an Iris flower. There are three types of flowers, Iris Setosa, Iris Versicolour and Iris Virginica. The four features, Sepal length, Sepal width, Petal length and Petal width are measured in centimeters. The data is depicted in Figure1, Figure 2 and Figure 3 before clustering. The shape of the enclosing contours in input space is governed by two parameters, Gaussian kernel q and soft margin constant C. As q is increased, the enclosing contours form tighter fits to the data. For C = 1, all the data points are surrounded by the boundary of support vectors. For fixed q, as C is decreased the number of support vectors decreases since ignoring outliers gives a smoother shape. In iris data, the numbers of clusters are three and there is overlap between two of them. One of the clusters is linearly separable from the other two at q=0.5 with no bounded Support Vectors. The remaining two clusters have significant overlap, and are separated at q= 4.2, 1/ (NC) = 0.55, with 4 misclassifications. As the numbers of principal components are increased from 3 to 4, the number of misclassifications increases from 4 to 14. The increase in number of support vectors and bounded support vectors required to obtain contour splitting. As the dimensionality of the data increases a larger number of support vectors are required to describe the contours, given in Table 2. The clustering obtained is shown in Figure 1 for first three principal components. http://archive.ics.uci.edu/ml/

Table 2: Performance of SVC on Iris data for principal components

Iris data



Principal Components	q	1/(NC)	Support Vectors	Bounded Support Vectors	Misclassified
1-2	4.2	0.55	20	72	4
1-3	7.0	0.70	23	94	4
1-4	9.0	0.75	34	96	14

Figure 1. Three principle components of Iris data set

3. EXAMPLE FROM G. TECHNOLOGY

To implement it in cellular manufacturing on a large scale data, an example is selected having the data for group technology (Xu & Wang, 1989). Fuzzy mathematics is applied to capture the uncertainty in the measurement of similarities between parts. Two different fuzzy pattern recognition approaches, namely fuzzy classification and fuzzy pattern recognition are used. It is shown that fuzzy classification method is very sensitive to the initial part families. Results obtained from fuzzy equivalence relation method depend upon the threshold value. In the example, 25 valve spools are used. These design specifications are provided by a company in the precision control industry. The first step is to select part features to be classified. Generally, no restrictions are put on feature selection. In this example, after careful review and discussion with engineers from the company, the following 15 (But the author has reported 14) features are restricted. They are all related to machining. (1) Overall length (L); (2) maximum diameter (D_{max}); (3) (L/D_{max}) ratio; (4) number of grooves; (5) minimum diameter (D_{min}); (6) tightest dimensional tolerance; (7) best surface finish; (8) perpendicularity; (9) cylindricity; (10) parallelism; (11) round out; (12) position; (13) straightness; and (14) symmetry (15) flatness. The following are data & result of the fuzzy clustering. Similarity value λ and number of desired part families C are parameters. The part features are as below.

5.514 1.1600 .5600 4.7534 2.0 .00080 8.000 .00010
.00000 .000 .00005 .0100 .0000 .000 .000
5.960 .9996 .3900 5.9624 3.0 .00010 32.000 .00050
.00010 .000 .00000 .0000 .0000 .000 .000
5.960 .9996 .3900 5.9624 3.0 .00010 32.000 .00050
.00010 .000 .00000 .0000 .0000 .000 .000
5.960 .9996 .3900 5.9624 3.0 .00010 32.000 .00050
.00010 .000 .00000 .0000 .0000 .000 .000
4.687 .7180 .3762 6.5279 6.0 .00050 3.000 .00000
.00005 .000 .00020 .0000 .0000 .000 .000
4.687 .7180 .3762 .6279 6.0 .00005 3.000 .00000
.00005 .000 .00020 .0000 .0000 .000 .000
4.090 .5040 .3290 8.1150 2.0 .00020 32.000 .00080
.00003 .020 .00500 .0030 .0000 .000 .000
3.928 .7509 .2900 5.2310 4.0 .00010 8.000 .00020
.00000 .000 .00000 .0000 .0000 .000 .000
6.312 1.0600 .8750 5.9547 .0 .00005 63.000 .00100
.00000 .000 .00003 .0300 .0000 .000 .005
2.510 .3165 .1590 7.9305 4.0 .00000 .000 .00060
.00000 .000 .00200 .0000 .0000 .010 .000
5.960 .9997 .3900 5.9618 .0 .00010 16.000 .00050
.00010 .000 .00000 .0000 .0000 .000 .000
11.481 .8780 .5000 13.0763 5.0 .00100 125.000 .00000
.00000 .000 .00010 .0000 .0000 .000 .000
4.687 .7180 .3762 6.5279 6.0 .00050 5.000 .00000
.00000 .000 .00010 .0000 .0000 .000 .000
11.281 .8750 .5000 12.8926 7.0 .00010 8.000 .00200
.00000 .000 .00000 .0000 .0003 .002 .000
3.700 .3800 .2000 9.7368 2.0 .00010 5.000 .00000
.00000 .000 .00020 .0000 .0000 .000 .000
3.700 .5900 .2750 6.2712 2.0 .00005 5.000 .00050
.00005 .000 .00020 .0000 .0000 .015 .000
2.174 .1800 .1090 12.0778 .0 .00010 5.000 .00000
.00000 .000 .00050 .0020 .0000 .010 .000
3.700 .6252 .3000 5.9181 2.0 .00100 5.000 .00050
.00005 .000 .00050 .0000 .0000 .000 .000
5.512 .7500 .3500 7.3493 2.0 .00015 16.000 .00010
.00005 .000 .00000 .0000 .0000 .000 .000
3.600 .3849 .1880 9.3531 2.0 .00010 5.000 .00080
.00000 .000 .00000 .0030 .0000 .000 .000
4.076 .5040 .4000 8.0873 2.0 .00010 5.000 .00080
.00003 .020 .00020 .0000 .0000 .000 .000
5.512 1.1873 .5500 4.6425 2.0 .00010 8.000 .00010
.00000 .000 .00005 .0100 .0000 .000 .000
2.174 .3125 .1590 6.9568 4.0 .00005 5.000 .00000
.00000 .000 .00500 .0000 .0000 .010 .000
6.388 .9377 .4400 6.8124 2.0 .00020 5.000 .00000
.00000 .000 .00050 .0020 .0000 .000 .000
4.687 .7180 .3550 6.5279 6.0 .00020 16.000 .00000
.00005 .000 .00010 .0000 .0000 .000 .000

4. RESULTS AND ANALYSIS

The clustering with similarity value λ=0.7 is obtained for six part families as shown below.

P.F. No. 1.	P. F. No . 2.	P.F. No. 3	P. F. No. 4	P. F. No. 5	P.F. No. 6
1,2,3,4,5,6,8,11,13, 15,18,19,22,24,25	7, 21	9	10, 16, 17,23	12, 14	20

The clustering with seven part families is obtained as below.

P.F. No. 1.	P.F. No. 2.	P. F. No. 3	P. F. No. 4	P. F. No. 5	P.F. No.6	P.F. No.7
10,15,17, 20,21	2,3, 4,7	14	11,1 9,25	12	1,5,6,8,13,16,1 8,22,23,24	9

SVC as a "divisive" clustering algorithm is used, starting from a small value of q and increasing it. The initial value of q is chosen to result in a single cluster. Choosing C=1, as at this value no outliers are formed. As q is increased, bifurcations of clusters are found. Start out with p= 1/ (NxC), or C=1, which do not allow for any outliers. As q is increased, single or few points' breaks off, or cluster boundaries become very rough. Therefore, p should be increased in order to investigate when BSVs are allowed. The clustering with using SVC algorithms is shown for six part families as below.

P.F. No. 1.	P.F. No. 2.	P.F . No . 3	P. F. No. 4	P. F. No . 5	P.F. No. 6
1,5,6,8,10,13,15,1 6, 17,18,20,21,22,23, 24	2,3,4, 7	9	11,19,2 5	12	14

The clustering with 7 part families using SVC algorithms is shown below

P.F. No. 1.	P.F. No. 2.	P.F . No. 3	P. F. No. 4	P. F. No. 5	P.F . No. 6	P.F . No. 7
1,5,8,10,13,15 , 16,17, 18,20,21,22,2 3,24	2,3,4 ,7	6	9	11,19, 25	12	14

Table 3: Performance of SVC on the data for C= 0.50

Sr. No.	Value of q	Number of clusters	Time (Sec.)
1	0.001	2	0.672
2	0.002	3	0.906
3	0.005	4	1.672
4	0.03	5	1.922
5	0.05	6	2.344
6	0.08	7	2.594
7	0.10	8	2.640
8	0.50	18	3.391

Table 3 gives the values of parameter q, number of clusters formed and the time in seconds required for forming clusters by SVC algorithm. More time is required for as the number of clusters goes on increasing. The clusters are formed within a reasonable time period. The dataset typically contains the data which forms a numbers of clusters for smaller values of both the parameters.

5. CONCLUSIONS

Support Vector Clustering (SVC) method is applied for cellular manufacturing for grouping of parts into part families. It is methodology for group technology. There is no need to convert the data into part machine incidence matrix. The values of parameter q, number of clusters are input parameters. In fair amount of time, it clusters in decent time. For even large size datasets, it clusters within reasonable time.

6. ACKNOWLEDGMENTS

Author thanks Dr. Kripa Shanker, Emeritus Professor, Indian Institute of Technology, Kanpur for guidance & kind support.

7. REFERENCES

[1] C. Chu, and M. Tsai, 1990. "A comparison of three array-based clustering techniques for manufacturing cell formation," International Journal of Production Research, 28(8), 1417-1433. https://doi.org/10.1080/00207549008942802

[2] T. Gupta, and H. Seifoddini, 1990. "Production data based similarity coefficient for machine-component grouping decisions in the design of cellular manufacturing system," International Journal of Production Research, 28(7), 1247-1270.

[3] T. Gupta, 1991. "Clustering algorithms for design of cellular manufacturing system- an analysis of their performance," Computers & Industrial Engineering, 20(4), 461-468. https://doi.org/10.1016/0360-8352(91)90018-2

[4] B. R. Sarkar, 1996."The resemblance coefficients in group technology: A survey and comparative study on relational metrics," Computers & Industrial Engineering, 30(11), 103–116. https://doi.org/10.1016/0360-8352(95)00024-0

[5] C. T. Mosier, J. Yelle, and G. Walker, 1997. "Survey of similarity coefficient based methods as applied to group technology configuration problem," Omega: the International Journal of Management Science, 25(1), 65-79. https://doi.org/10.1016/S0305-0483(96)00045-X

[6] M. P. Chandrasekharan and R. Rajagopalan 1986, "An ideal seed non-hierarchical clustering algorithm," International Journal of Production Research, 24(2), 451-464. https://doi.org/10.1080/00207548608919741

[7] M. P. Chandrasekharan and R. Rajagopalan, 1987 "ZODIAC-an algorithm for concurrent formation of part families and machine cells," International Journal of Production Research, 25(6), 835-850. https://doi.org/10.1080/00207548708919880

[8] G. Srinivasan, and T. T. Narendran, 1991, "GRAPHICS: a non-hierarchical clustering algorithm for group technology problem," International Journal of Production Search, 29(3), 463–478. https://doi.org/10.1080/00207549108930083

[9] A. Kusiak, and W. S. Chow, 1987, "Efficient solving of the group technology problem," Journal of Manufacturing Systems, 6(2), 117-124.

[10] G. Srinivasan, T. T. Narendran, and B. Mahadevan, 1990 "An assignment model for the part families problem in group technology," International Journal of Production Research, 28(1), 145-152. https://doi.org/10.1080/00207549008942689

[11] N. Singh, 1993, "Design of cellular manufacturing systems: an invited review", European Journal of Operational Research, 69(3), 284-291. https://doi.org/10.1016/0377-2217(93)90016-G

[12] K. Shanker, and A. K. Agrawal, 1997, "Models and solution methodologies for the generalized grouping problem in cellular manufacturing," International Journal of Production Research, 35(2), 513-538. https://doi.org/10.1080/002075497195885

[13] V. Venugopal, 1999, "Soft-computing based approaches to the group technology problem: a state-of-the-art review," International Journal of Production Research, 37(14),3335-3357. https://doi.org/10.1080/002075499190310

[14] D. Ben-Arieh, E. Triantaphyllou, 1992, "Quantifying data for group technology with weighted fuzzy features," International Journal of Production Research, 30(6), 1285-1299 http://dx.doi.org/10.1080/00207549208942957

[15] C. Chu, and J. Hayya, 1991, "A fuzzy clustering approach to manufacturing cell formation", International Journal of Production Research, 29(7), 1475-1487. https://doi.org/10.1080/00207549108948024

[16] A. K. Jain, M. Murthy, and P. J. Flynn, 1999, "Data clustering: A review", ACM Computing Surveys, 31, 264-323. http://archive.ics.uci.edu/ml/

[17] A. Ben-Hur, D. Horn, H. T. Siegelmann and V. Vapnik 2001, " Support Vector Clustering", *Journal of Machine Learning Research*, 2(1), 125-137. http://dx.doi.org/10.1162/15324430260185565

[18] J. Sakethanath, and S. K. Shevade, 2006 "An efficient clustering scheme using support vector method", Pattern Recognition, *39*, 1473-1480. http://dx.doi.org/10.1016/j.patcog.2006.03.012

[19] P. C. Kulkarni, 2024, "Multi-objective machine cell formation by NSGA-II", Engineering Today Journal, Vol.3(2), 45-53. https://doi.org/10.5937/engtoday2400007K

[20] H. Li and Y. Ping, 2015 "Recent advances in Support Vector Clustering: Theory and applications", International Journal of Pattern Recognition and Artificial Intelligence", Vol.29, No. 01, 1550002. https://doi.org/10.1142/S0218001415500020

[21] Xu, H. and Wang, H., 1989, "Part family formation for GT applications based on fuzzy Mathematics", International Journal of Production Research, 27 (9), 1637-1651. http://dx.doi.org/10.1007/s001700170036

[22] P. C. Kulkarni 2021, "Solving generalized grouping problems in Cellular Manufacturing Systems by genetic algorithms", Journal of Science and Technology, Vol. 6, Issue. 5, 82-88. https://doi.org/10.46243/jst.2021.v6.i05.pp82-88

[23] P.C. Kulkarni 2024, "Support vector clustering algorithm for cell formation in Cellular manufacturing systems", Engineering Today Journal, Volume 3(3), DOI: 10.5937/engtoday2400012K

[24] Blake, C., and Merz, C. UCI repository of machine learning databases.(1998) https://doi.org/10.1111/1740-9713.01589