ORIGINAL SCIENTIFIC PAPER OPEN ACCESS

Support Vector clustering algorithm for cell formation in Cellular manufacturing systems

Prafulla C. Kulkarni[1*]

[1]Gokhale Education Society's R. H. Sapat College of Engineering, Management Studies and Research, Nashik, India

ARTICLE INFO

*** Correspondence:** prafulla.kulkarni@ges-coengg.org

DOI: 10.5937/engtoday2400012K

UDC: 621(497.11) **ISSN:** 2812-9474

Article history: Received 24 June 2024; Revised 12 August 2024; Accepted 22 August 2024

ABSTRACT

Support Vector Clustering (SVC) algorithm is presented in this study to resolve the cell creation problem in Cellular Manufacturing Systems, even if prior approaches have been used to solve this problem as well. A Gaussian kernel is used in the SVC technique to translate data points from data space to Hilbert space. The method then performs in Hilbert space for the smallest sphere that contains the data images. When the minimal sphere is transferred back to data space, it breaks up into multiple parts, each of which encloses a different point cluster. The scale, at which cluster formation in the data is examined, is determined by the Gaussian kernel's width. Managing overlapping clusters and outliers is made easier by the soft margin constant. The performance of the SVC algorithm is conducted using a collection of two test problems from the literature of Cellular manufacturing systems. The encouraging results are obtained.

KEYWORDS

Support Vector clustering, Cellular manufacturing systems, Cell formation

1. INTRODUCTION

The challenge of Cellular Manufacturing deals with the creation of component families based on shared processing needs and the arrangement of machines into cells according to their capacity to process particular part families [10], [11]. A machine cell is where all activities of a part family are finished in an ideal scenario. Data can be divided into homogenous groups based on features by using cluster analysis to find comparable and different features in the data. The existence of homogeneous clusters in unprocessed data is a fundamental premise. The methods of cell formation based on clustering forms part families and machine cells. Some of the earlier algorithms developed were ISNC, ZODIAC and GRAPHICS. These three algorithms obtained acceptable solutions to a number of problems. These algorithms were used to compare performances of algorithms developed by other methods. But the limitation of these algorithms is that it can be applied only to binary data, which restrict their applicability for manufacturing data considerations. The advancement in data analysis led to the development of other clustering algorithms. On a broader scale, cluster analysis identifies similar and dissimilar features in the data and segregates it into homogenous groups based on features. An underlying assumption is that homogeneous clusters exist in the raw data. Numerous studies have been conducted on clustering algorithms for the purpose of grouping data points based on different criteria [3]. Partitioning techniques (like k-means), hierarchical techniques (like agglomerative approach), density-based techniques (like Gaussian mixture models), and grid-based techniques (like self-organizing feature maps) are the categories into which the clustering algorithms can be divided. Despite the fact that these algorithms have been applied effectively in numerous contexts, there are a few things to keep in mind. Certain techniques are limited to a specific type of data (k-means, for example, can only be applied to interval-based data). Certain techniques result in poor clustering quality because they are vulnerable to outliers. The high-dimensional data set should be manageable for the clustering process [7]. In the end, the clustering technique ought to identify the clusters of any shape [4].

This work proposes Support Vector Clustering (SVC), a non-parametric method for group technology problems. Adoption is justified by its ability to effectively address the aforementioned problems and yield positive outcomes. This is how the rest of the paper is structured. In section 2 mathematical formulations of SVC are given. An overview of methodology is briefed in section 3 with the effect of parameters on clustering. Based approach is compared with other two problems from literature of Cellular manufacturing systems in section 4 with discussions and conclusions are in the Section 5.

2. THE SVC ALGORITHM

The paper of Ben-Hur et.al [1] presented Support Vector Compute (SVC) as a method for clustering data sets using the support vector machine (SVM) theory. Vapnik [2] developed the first support vector machines (SVMs) to minimize structural risk and address pattern classification and nonlinear regression issues. SVM has been extensively employed in many fields to handle classification, regression, and novelty detection based on the viewpoint of statistical learning theory. The SVC scheme starts with an optimization issue that must be solved in order to find the lowest possible space hyper sphere that contains images of the data points. Assume that the data space is a set of points in the data. The smallest enclosing using a nonlinear transformation from to some high dimensional feature-space

$$\|\Phi(x_j)-a\|^2 \le R^2 \forall j=1,...,n,$$

where $\|.\|$ is the Euclidean norm and a is the center of the sphere. In order to deal with problem of outliers, the soft constraints are incorporated by adding slack variables ξ_j :

$$\|\Phi(x_j)-a\|^2 \le R^2+\xi_j \tag{1}$$

With $\xi_j \ge 0$. The problem above can be solved by introducing Lagrangian as follows:

$$L=R^2-\sum_j (R^2+\xi_j-\|\Phi(x_j)-a\|^2 \beta_j-\sum \xi_j\mu_j+C\sum \xi_j, \tag{2}$$

where $\beta_j \ge 0$ and $\mu_j \ge 0$ denote the Lagrange multipliers. C is the user defined constant, and $C\sum \xi_j$ is the penalty term. To solve the equation, the derivative of L with respect to R, a and ξ_j respectively are as follows:

$$\sum_j B_j = 1, \tag{3}$$

$$a=\sum_j \beta_j \Phi(x_j), \tag{4}$$

$$\beta_j = C-\mu_{j,} \tag{5}$$

By adopting Karush-Kuhn-Tucker (KKT) complimentarily conditions (Fletcher, 2003),

$$\xi_j\mu_j = 0, \tag{6}$$

$$(R^2+\xi_j-\|\Phi(x_j)-a\|^2)\beta_j = 0, \tag{7}$$

It follows from Eq. (7) that the image of a point x_i with $\xi_i > 0$ and $\beta_i > 0$ lies outside the feature space sphere. Eq. (6) states that such a point has $\mu_i = 0,$ hence it is concluded from Eq. (5) that $\beta_j = C$. This will be called as a bounded support vector or BSV. A point x_i with $\xi_i > 0$ is mapped to the inside or to the surface of the feature space sphere. If it is $0 < \beta_j < C$ then Eq. (7) implies that its image $\Phi(x_i)$ lies on the surface of the feature space sphere. Such a point will be referred to as a support vector or SV. SV lies on cluster boundaries, BSV lie outside the boundaries and all other points lie inside them. When C ≥ 1, no BSV exist because of the constraints (3). Using these relations, the variables R, a and μ_j are eliminated. By turning the Lagrangian into the Wolf dual form that is a function of the variables β_j. The image of a point with and lies outside the feature space sphere, as per Eq. (7). Since Eq. (6) indicates that such a point possesses, Eq. (5)'s conclusion is that... This will be referred to as a BSV, or bounded support vector. A point with is mapped to either the feature space sphere's interior or outside. If so, Eq. (7) suggests that the surface of the feature space sphere contains its image. We'll refer to this point as a support vector, or SV. All other points lie inside cluster borders, SV lies on them, and BSV lies outside of them. There are no BSV when C ≥ 1 due to the limitations (3). Making use of these relations, the variables are eliminated.

$$W = \sum_j \Phi(x_i)^2 \beta_j - \sum_{i,j} \beta_i \beta_j \Phi(x_i) \cdot \Phi(x_j). \tag{8}$$

Since the variables μ_j do not appear in the Lagrangian, and subjected to the constraints:

$$0 \le \beta_j \le C, j = 1,....,N. \tag{9}$$

The dot product $\Phi(x_i) \cdot \Phi(x_j)$ should be satisfied by Gaussian Kernel $K(x_i, x_j)$.

$$K(x_i, x_j) = \Phi(x_i) \bullet \Phi(x_j) = e^{-q\|x_i - x_j\|^2} \tag{10}$$

where q denotes the width parameter. The Lagrangian W is now written as:

$$W = \sum_j K(x_j, x_j)\beta_j - \sum_{i,j} \beta_i \beta_j K(x_i, x_j) \tag{11}$$

At each point x define its distance, when mapped to feature space, from center of sphere. The radius of support vector is R where x_i is a support vector.

$$R^2(x) = \|\Phi(x) - a\|^2 \tag{12}$$

$$R^2(x) = K(x, x) - 2\sum_j \beta_j K(x_j, x) + \sum_{i,j} \beta_i \beta_j K(x_i, x_j) \tag{13}$$

The contour that encloses the cluster in data space is $\{x \mid R(x) = R\}$. The data point x_i is a bounded Support Vector if $R(x_i) > R$. Three types of the inner-product kernels can be described as shown in Table 1. The cluster assignment can be determined as follows:

Let a segment of points y, the clustering rule can be represented as the adjacency matrix:

A_{ij} = 1, for all y on the line segment connecting x_i and x_j and 0 otherwise (14)

All data points are checked to assign a specific cluster. The outliers are unclassified since their feature space lie outside the enclosing sphere [8].

Table 1: Three common types of the kernels

Sr. No.	Type of SVM	$K(x_i, x_j)$	Parameter
1.	Polynomial	$(1 + x_i' x_j)^p$	Where p denotes the power and is specified by user
2.	Radial-basis function	$\exp(-\frac{1}{2\sigma^2}\|x_i - x_j\|^2)$	Where σ denotes the width and is specified by user
3.	Multilayer perceptron	$\tanh(\beta x_i' x_j + b)$	Where β, b denote the coefficient and are specified by user

3. EFFECT OF PARAMETERS

The Iris data, a well-known benchmark in the pattern recognition field, was used to test the SVC algorithm. It is available from Blake and Merz [5] at the UCI library of machine learning resources. Each of the 150 occurrences in the data set is made up of four characteristics of an iris flower. Iris Virginica, Versicolor, and Setosa are the three varieties of iris blossoms. Centimeters are used to measure the four features: sepal length, sepal breadth, petal length, and petal width. Figure 1 shows the data clustering. Two parameters, the soft margin constant C and the Gaussian kernel q, control the shape of the enclosing contours in input space. With an increase in q, the enclosing outlines 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. Three clusters are present in the iris data, with two of them overlapping. At q=0.5, one of the clusters may be separated from the other two linearly and does not have any bounded support vectors. With four misclassifications and a large overlap, the remaining two clusters are divided at q=4.2, 1 / (NC) = 0.55. The number of misclassifications rises from 4 to 14 when the number of major components increases from 3 to 4. Contour splitting necessitates a rise in the number of support vectors and bounded support vectors. As the data becomes more dimensional, more support vectors—as seen in Table 2—are needed to characterize the contours.

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

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 displays the resulting clustering for first three principal components. Support vector clustering is applied for many datasets and problems [9], [12].

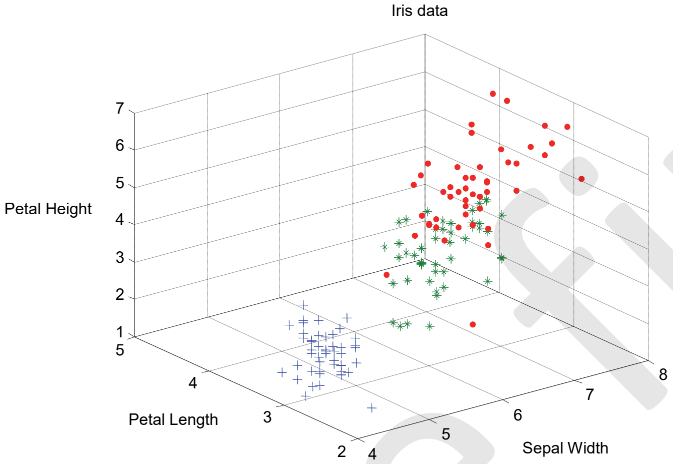

Figure 1: First three principle components are considered of Iris data set.

4. EXAMPLE FROM CELLULAR MANUFACTURING SYSTEMS

An example with the data of manufacturing technology is chosen in order to apply it to cellular production on a broad scale. Choosing which component features to classify is the initial step. Generally, feature selection is unrestricted. In this example [8], after careful review and discussion with engineers from the company, the following 15 features are restricted. They are all related to machining process and inspection. (1) Overall length (L); (2) maximum diameter (Dmax); (3) (L/Dmax) ratio; (4) number of grooves; (5) minimum diameter (Dmin); (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 part feature data are:

Table 2: Input of the dataset 1

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
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

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
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

The following are results of the clustering with number of desired part families. Using SVC as a "divisive" clustering algorithm, q is increased from a tiny starting point. One cluster is produced by selecting q's starting value. Selecting C=1 because there are no outliers created at this value. Cluster bifurcations are observed as q is increased. To begin, set p= 1/(NxC) or C=1, which precludes the possibility of outliers. Single or few points fall off or the boundaries between clusters become extremely ragged as q increases. Thus, p ought to be raised in order to study the conditions under which BSVs are permitted. The values of parameter q, the number of clusters created, and the amount of time in seconds needed for the SVC algorithm to produce clusters are shown in Table 3. Additional time is needed as the number of clusters increases.

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
1,5,6,8,10,13,15,16,17,18,20,21,22,23,24	2,3,4,7	9	11,19,25	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
1,5,8,10,13,15,16,17,18,20,21,22,23,24	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

The part families are obtained with the parameters given in the Table 4 for the dataset 1. Similarity of parts is observed in the data set by close observation. Part 2, part 3 and part 4 are having the same manufacturing requirements, therefore all form one part family. Similarly, part 5 and part 13 have same requirements. Both the parts are clustered in one part family. Part 15, part 20 and part 21 have nearly similar requirements, which form a separate group. Similarly, part 11, part 19 and part 25 have similar requirements, which initially forms one group but when the number of clusters increases they split to form separate groups. Part 1 and part 22 are also having quite similar requirements. Therefore, both are in one part family. Part 12 is different from the others in the length parameters; therefore it forms a separate group. Similarly, part 9 forms a separate group as it is not having grooves and surface finish is high. Accordingly, distinct parts families are formed.

Table 4: Part families formed by SVC algorithm of dataset 1

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

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

In another example, dataset 2 is selected (Masnata and Settineri) [6] for 25 parts. The dataset 2 is provided in Table 5. Each part has 16 features of dimensional, geometrical and technological nature. The features are denoted as **A**: Maximum external diameter; **B**: Minimum external diameter; **C**: Maximum internal diameter; **D**: Minimum internal diameter; **E**: Grooves; **F**: Auxiliary holes; **G**: Spherical surfaces; **H:** Planer surfaces; **I**: Screw threads; **J:** Toothing; **K:** Length; **L**: Tolerances; **M**: Surface finish; **N**: Complexity; **O**: Material; and **P**: Length-to-maximum external diameter ratio.

Table 5: Input of the dataset 2

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
30.9	30.9	0	0	4	0	0	1	0	1	144	0.05	1.3	3	31	4.66
123	123	0	0	4	8	0	2	2	0	344	0.045	1.2	3	30	2.8
381	381	0	0	3	0	2	0	2	0	594	0.024	0.8	6	49	1.56
141	141	0	0	3	9	1	0	3	0	509	0.059	1.4	4	39	3.6
315	315	0	0	1	24	0	3	1	0	249	0.08	2	5	32	0.79
206	206	0	0	4	12	2	0	0	0	625	0.02	0.6	4	41	3.03
451	451	304	304	4	20	0	3	0	0	503	0.102	3	6	49	1.12
231	231	0	0	0	0	0	1	1	2	663	0.106	3	3	33	2.87
478	382	382	382	2	0	2	2	0	0	581	0.165	3	8	44	1.21
411	411	169.6	116.6	0	0	1	3	0	0	257	0.028	0.7	6	2	0.63
228	228	0	0	4	0	2	1	3	0	146	0.044	1	4	21	0.64
518	358	0	0	1	0	0	3	0	0	810	0.02	0.6	6	43	1.57
102.4	102.4	70.7	70.7	0	0	0	2	1	0	236	0.051	1.2	2	20	2.3
504	504	0	0	2	0	2	0	0	0	631	0.02	0.8	8	39	1.25
449	449	0	0	0	25	0	4	1	0	316	0.097	2.8	9	37	0.7
45.7	45.7	0	0	2	0	0	1	2	0	133	0.04	1	4	33	2.91
476	476	350	180	0	30	0	3	2	0	779	0.09	2.5	7	30	1.64
482	482	0	0	0	0	0	2	1	1	354	0.141	4	8	25	0.73
387	387	175	175	0	22	0	0	0	0	481	0.1	3.2	9	40	1.24
502	502	205	120	3	30	0	2	1	0	295	0.075	1.7	7	21	0.59
285	285	0	0	3	12	0	0	2	3	729	0.095	1.7	6	27	2.56
271	271	0	0	1	0	0	3	2	1	701	0.072	1.5	5	20	2.59
445	445	0	0	0	0	2	0	3	0	612	0.018	0.8	9	31	1.38
231	149	0	0	4	0	1	1	0	1	393	0.063	1.2	3	45	1.71
189.7	102.9	0	0	3	10	1	0	0	0	519	0.05	1.9	3	48	2.74

Table 6 shows ten instances of the experiments conducted by varying parameters q and p .

Table 6: Performance of SVC on the dataset 2

Sr. No.	Value of q	Value of p	Number of clusters	Time (Sec.)
1	0.030	0.001	8	2.8590
2	0.070	0.001	14	3.3910

3	0.045	0.0001	15	3.5990
4	0.051	0.0001	16	3.4220
5	0.001	0.0002	17	3.4070
6	0.050	0.0001	18	3.8280
7	0.001	0.0005	19	3.4690
8	0.030	0.0020	22	3.4750
9	0.500	0.0025	24	3.5160
10	0.500	0.005	25	3.5160

The numbers of clusters are formed within a reasonable time period as given in Table 7.4. The dataset 2 typically contains the data which forms eight numbers of clusters for smaller values of both the parameters. The part families are obtained with the parameters given in the Table 7 for dataset 2.

Part 1 and part 16 are similar and distinct from other parts. For a number of instances, the dataset is distinctly giving eight clusters. However, by increasing the value of q and decreasing the value of p, more number of part families are formed.

Table 7: Part families formed by SVC algorithm of dataset 2

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

5. CONCLUSIONS

Support Vector Clustering (SVC) method is applied for cellular manufacturing for grouping of parts into part families. SVC method has been widely applied in other areas. It is a robust algorithm. For application in cellular manufacturing, it is proposed for forming part families in cell formation. The values of parameter q, the number of clusters created, and the amount of time in seconds needed for the SVC algorithm to produce clusters. As the number of clusters increases, more time is needed. The clusters form in a decent time. Typically, the dataset includes data that clusters into several groups based on lower values of both parameters. The data does not require conversion into a part machine incidence matrix. In a fair amount of time, it clusters even for massive amounts of data. Even for large size data, SVC algorithm clusters within a reasonable time.

ACKNOWLEDGEMENTS

Author acknowledge and thanks the timely guidance of Prof. Dr. Kripa Shanker, Emeritus Professor, IME Department at I. I. T. Kanpur during which the experiments were conducted.

REFERENCES

[1] A. Ben-Hur, D. Horn, H. T. Siegelmann, and V. Vapnik, “Support vector clustering”, The Journal of Machine Learning Research, Vol. 2, pp. 125–137, (2001)

[2] V. N. Vapnik, “The Nature of Statistical Learning Theory”, Springer, New York (USA), (1995)

[3] A. K. Jain and R. C. Dubes, "Algorithms for clustering data", Prentice Hall, New Jersey, (USA), (1988)

[4] J. Han and M. Kamber, "Data mining: concepts and techniques", Morgan Kaufmann Publishers, Waltham (USA), (2000)

[5] C. Blake, E. Keogh, and C. Merz, "UCI repository of machine learning databases", http://archive.ics.uci.edu/ml/, (1998)

[6] A. Masnata and L. Settineri, "An application of fuzzy clustering to cellular manufacturing", International Journal of Production Research, Vol. 35(4), pp. 1077–1094, https://doi.org/10.1080/002075497195551, (1997)

[7] J. S. Nath and S. K. Shevade, "An efficient clustering scheme using support vector methods", Pattern Recognition, Vol. 39(8), pp. 1473–1480, https://doi.org/10.1016/j.patcog.2006.03.012, (2006)

[8] H. Xu and H. P. (Ben) Wang, "Part family formation for GT applications based on fuzzy mathematics", International journal of production research, Vol. 27(9), pp. 1637–1651, https://doi.org/10.1080/00207548908942644 (1989)

[9] P.C. Kulkarni, "Developments and analysis of heuristics for manufacturing cell formation", Ph. D. Thesis, I. I. T. Kanpur (India), (2009)

[10] P. C. Kulkarni and K. Shanker, "Genetic algorithm approach for solving intercellular layout problems in cellular manufacturing systems", Proceedings of the IEEE International Conference on Industrial Engineering and Engineering Management, Bangkok (Thailand), pp. 1587–1591, https://doi.org/10.1109/IEEM.2013.6962677, (2013)

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

[12] H. Li and Y. Ping, "Recent advances in Support Vector Clustering: Theory and Applications", International Journal of Pattern Recognition and Artificial Intelligence, Vol. 29(1), p. 1550002, https://doi.org/10.1142/S0218001415500020, (2015)