

On the Development of an Empirical Model for Carcinomic Treatment using Chaotic Hyper-Heuristic Algorithm



Prachi Vijayeeta, M. N. Das, B. S. P. Mishra

Abstract - A new dimension in the field of computational intelligence was introduced in the late nineties to comprehend a vivid combination of several multi-disciplinary areas. The coalescence biology along with data mining and statistical learning have given birth to Bioinformatics that provides various paradigms for studying the behaviour of unknown patterns at the micro level. In the present work, a recently developed human inspired optimization algorithm called search and rescue (SAR) optimization is employed with an improved version of parameters using Chaos theory. CSARO (Chaotic search and rescue optimization algorithm) unlike other existing algorithms has proven to be a better choice for optimising the gene selection mechanism as well as the control parameters of the learning model. This hyper heuristic algorithm obtained by the inclusion of chaos in SAR mainly aims at enhancement of its global search mobility and prevents from getting trapped in the local optimum. A comparative study with other existing techniques on seven benchmark datasets is performed. The performance of the algorithm is tested using evaluation metrics.

Keywords: Microarray, optimization, KPCA, KLDA, KSVM, k-fold.

I. INTRODUCTION

Hyper-heuristic models are the recent trailblazing approaches in the field of evolutionary computations. Researchers are sedulously trying to discover unknown facts and patterns about the cause as well as the relevant treatment for the occurrence of carcinogenic cells. Appropriate diagnosis at initial stage, followed by significant treatment (chemotherapeutic or immunotherapeutic) still remains a great challenge in the field of oncology. Keeping this in view, medical practitioners have expressed that the changes in the tumour morphology can be either a discrete or a continuous process, coupled with the time factor. Addressing the dichotomy between the cause of cancer at the genome level and its possible treatments [1] is the major focal point of our work.

Revised Manuscript Received on July 30, 2020.

* Correspondence Author

Prachi Vijayeeta*, Department of Computer Engineering, Kalinga Institute of Industrial Technology, [Deemed To Be University], Bhubaneswar, India. E-mail: pvijayeetafca@kiit.ac.in

M. N. Das, Department of Computer Engineering, Kalinga Institute of Industrial Technology, [Deemed To Be University], Bhubaneswar, India. E-mail: mndas_prof@kiit.ac.in

B. S. P. Mishra Department of Computer Engineering, Kalinga Institute of Industrial Technology, [Deemed To Be University], Bhubaneswar, India. E-mail: bsmishrafcs@kiit.ac.in

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Hyper heuristic methods have become the subject of many publications as it leads toward faster global optimal solution, reduces interference of irrelevant factors, and ensures better throughput. To improve the efficiency of the algorithm we have screened out the most contributing features from a high dimensional dataset.

The high-dimensional nature of biological microarray data is wider open to computational overhead [2]. We have employed kernel Principal component analysis (KPCA) and kernel Linear discriminant analysis (KLDA) for pre-selection of genes which has corroborated to the dimension reduction thus improving the computational accuracy. The present study employs optimal chaos theory to administer chemotherapeutic treatment for controlling the cell growth. Since chemotherapy behaves like two-sided sword, where a trade-off is involved in destroying the infected cells along with a negative impact on normal cells.

A chaotic search and rescue optimization (CSAR) algorithm is employed to maintain a right balance between the toxicity constraint and the fixed drug dosage.

The paper is organized as follows : section 1 gives a glimpse of the problem statement with its objectives, section 2 describes the related work pursued in this area, section 3 illustrates the work process model along with the workflow diagram. Section 4 provides a brief discussion about the significance of the applied technologies, feature reduction methodologies and hyper parameter tunings. Section 5 briefly explains the experimental set up with valid dataset descriptions, section 6 graphs are plotted and the behaviour of the variables are thoroughly depicted. Finally, section 7 concludes the paper with an authentic suggestion for future work.

II. RELATED WORK

The power of computational mechanism lies in its ability to reveal the significant features with the help of filter, wrapper and hybrid techniques on discrete as well as continuous data. Boln et al. [3] had employed a new distributed filtering approach on 8 microarray datasets. Dash et al. [4] designed a three-dimensional integrated system that could automatically recommend the most suitable algorithm for feature selection. Tang [5] used Least squares method with SVM (Support Vector Machine) and particle swarm optimization for identifying top-ranked homogeneous genes. Rough set theory was applied by M et al.[6] for efficient selection of genes from microarray gene expression patterns. Tibshirani et al.



[7] had implemented supervised principal component analysis for the first time to generate a subset of predictors from the given pattern. Further their contribution was extended for regression problems. Zhang [8] framed an ensembled classifier model along with kernel ridge regression with a good performance measures.

A new version of PSO, named as Geometric PSO was used by Jourdon et al.[9] to find small samples of informative genes. For solving complex optimization problem, a multiobjective evolutionary algorithm was designed by Petrovski [10],[11] which basically works on the optimization of set of conflicting constraints. Generalised fisher score for independent feature selection along with quadratically constrained linear optimization was initially proposed by Han et al. [12]. Evolutionary voting based extreme learning machine was applied by Cao et al.[13] to improve the efficacy of the of the learning algorithm. Integrating the benefits of artificial neural network and artificial bee colony optimization algorithm, a new hybrid algorithm was proposed by Garo et al. [15]. It was applied on three binary microarray dataset namely colon, leukaemia and lungs for choosing the most predictive and informative genes that contributes for cancer classification. Paul [16] had employed multi objective evolutionary algorithm for minimizing the sample misclassification frequency. He discovered multiple non-dominated solutions from multiple training sets with nearly cent percent reliable and accurate classification [17].

However, recent trends in research have mostly emphasized on building up hybridised algorithms by uniting the advantages of two or more methods that gives birth to a new approach [18]. Some them includes Intelligent dynamic genetic algorithm which integrates reinforcement learning, random restart hill climbing and genetic algorithm[27]. Zibakhsh [28] had proposed a novel memetic algorithm with multi-view fitness function for the first time.

III. WORKING PROCESS MODEL

Many models are designed by researchers to study the response of tumorous cells on application of chemotherapy. A very well-known model was proposed by Gompertz which linearly combines a logarithmic function along with tumour growth parameters and survivability time factors. The expression takes the form as:

$$\frac{dN}{dt} = N(t) \left[\lambda \ln \left(\frac{\Theta}{N(t)} \right) - \sum_{j=1}^d k_j \sum_{i=1}^n C_{ij} \{ H(t - t_i) - H(t - t_i + 1) \} \right] \quad (1)$$

Where $N(t)$ represents the number of tumour cells at time (t) . λ , Θ are the parameters of tumour growth. $H(t)$ is a Heavside steep function (function takes value 1 for +ve arguments and 0 for -ve arguments). C_{ij} is the concentration of anti-cancer drugs and k_j quantifies the drugs content. With the help of Gompertz model we try to establish an equilibrium between theoretical and clinical cognition. Optimal control theory is employed to basically optimize the control parameter C_{ij} as curative treatment for restricting the tumour growth. A wide range of cancer literature determines the control and state trajectories for a dynamic system over a period of time at variable stages. However, we can formulate an objective function using equation(1) to

maximize chemotherapeutic treatment as an attempt to extirpate the number of tumour cells at the end of the treatment period [32].

$$Max f_1(c) = \int_{t_1}^{t_n} \ln \left(\frac{\Theta}{N(t)} \right) . dt \quad (2)$$

Subject to the constraint:

$$g_1(c) = \{ C_{maxj} - C_{ij} \geq 0 : \forall i \in 1, n, \forall j \in 1, d \} \quad (3)$$

Where $g_1(c)$ function computes the maximum dosage of drugs to be provided instantaneously.

The primary objective of chemotherapeutic treatment is:

- To obliterate the tumour by keeping track of the growth ratio.
- To increase the patient survival time by controlling the growth and the spreading of infected cells.
- In an advanced stage when the above two conditions fails then we can palliate cancer by imposing chemotherapeutic treatment to increase the life span.

In our work we have employed Chaotic Search and Rescue Optimization (CSARO) a metaheuristic as a global search mechanism to generate the most contributing features and Emperor Penguin optimization algorithm as a local search mechanism to optimize the hyper parameters of the learning model.

IV. METHODS APPLIED

A systematic reduction of features is essential to be chosen to participate in the learning process to distinguish the normal cells from those of the infected ones. In the present work we have employed Kernel based PCA (Principal component analysis) and Kernel-LDA (Linear discriminant analysis) which seemingly has yielded better result than other dimensionality reduction techniques like RFE(Recursive feature elimination) [25], mRMR(minimum Redundancy Maximum Relevance) [26]. Some statistical methods like T-score, Z-score, and Fisher Score [14] have also been used for feature selections where the inter-class and intra-class distance are the vital parameters for deducing the class type [22].

4.1 Selection of Gene Subsets using kernel-PCA

Kernel-PCA extends conventional PCA to a high dimensional feature space using kernel trick. It has the potential to extract nearly ‘n’ non-linear principal components with minimal complexity [26] [27]. Among the popular kernels such as Gaussian Kernel, Polynomial kernel and Hyperbolic Tangent kernel, we have employed Radial basis function [28].

4.2 Selection of Feature Subsets using kernel-LDA

Kernel discriminant analysis also called as generalised discriminant analysis (GDA) is employed to map the data points into a new feature space [29]. The RBF Kernel function as:

$$k(x, y) = \exp(-\gamma \| x_i - x_j \|^2); \gamma > 0 \quad (4)$$

where $\sigma \in \mathbb{R}$ and x, y are the data points.

$$K_b K_b^T \alpha = \lambda K_w K_w^T \alpha \quad (5)$$

where K_b, K_w are the between class and within class squared matrix respectively.



4.3 Kernel-Support vector machine (KSVM)

A hyperplane is estimated for maximization of margin in Fig.1 [5]. The classification boundary for all values of x , such that $f(x) = 0$ is a hyperplane defined by $wx + b = 0$ where:-

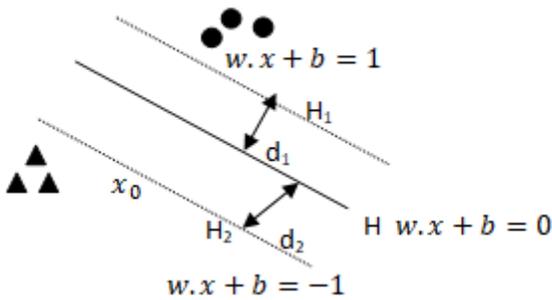


Fig 1. Geometrical interpretation of SVM

Hastie [14] had described the numerical interpretation of SVM in a very lucid manner. Mathematically, we can represent H in the form [20]:

$$W^T x + b = 0 \tag{6}$$

Where W represents the weight vector and b represents the bias that needs to be optimised.

$$\text{Min } z(w) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i \tag{7}$$

$$\text{Subject to the constraint:- } y_i((w \cdot x) + b) \geq 1 - \xi_i; \text{ for } i = 1, 2, \dots, n; \xi \geq 0 \tag{8}$$

Where C is the penalty parameter for no. of misclassified data and ξ_i is the distance measures of the points crossing the margin.

4.4 Chaotic Emperor Penguin Optimization

A recently developed meta heuristic algorithm named as Emperor Penguin Colony optimization algorithm is employed. Dhiman [33] had initially proposed this bio-inspired algorithm by simulating the huddling behaviour of penguins during the crowding process. A mathematical model has been designed by including chaos theory to the random parameters like the huddle boundary, the encircling temperature, their distance and the most active mover [24].

A. Phases of huddling behaviour of Penguins

The fundamental steps of the huddling behaviour of penguins was discussed in the previous work.

Part-I: The crowd boundary of emperor penguins is determined.

Part-II: The heat radiation is estimated from the temperature parameter

$$T = T'$$

Where the values are computed below:

$$T' = \left(T - \frac{\text{Iteration}_{\max}}{x - \text{Iteration}_{\max}} \right) \tag{9}$$

$$T = \begin{cases} 0, & \text{if } R > 0.5 \\ 1, & \text{if } R < 0.5 \end{cases} \tag{9.1}$$

Where R is the radius of the polygon which is modified using logistic chaotic map.

$$z_{t+1} = 4 \cdot z_t \cdot (1 - z_t) \tag{9.2}$$

Where z_t denotes the chaotic map value at t^{th} iteration.

Therefore equation 9.1 becomes:

$$T = \begin{cases} 0, & \text{if } z_t > 0.5 \\ 1, & \text{if } z_t < 0.5 \end{cases} \tag{9.3}$$

Part-III: The distance and the velocity of spiral movement around the crowd according to the current best optimal solution is calculated as follows:

$$\overrightarrow{Dist_{ep}} = \text{Abs}(S(\vec{A}) \cdot \vec{P}(x) - \vec{C} \cdot \vec{P}_{ep}(x)) \tag{10}$$

Where $\overrightarrow{Dist_{ep}}$ represents the distance between the emperor penguin and the best optimal fitness penguin, \vec{A}, \vec{C} are the parameters to control the random search and to avoid collision among the penguins. $\vec{P}(x)$ represents the best fit penguin where as $\vec{P}_{ep}(x)$ denotes the position vector of penguin. $S(\vec{A})$ refers to the social forces that pulls down the penguins towards the optimal best value. If M is the movement parameter responsible to maintain gap among the penguins to avoid collision then, the values of the parameters \vec{A}, \vec{C} can be estimated as follows:

$$\vec{A} = (M * (T' + P(\text{Accuracy})) * z_t) - T' \tag{11}$$

$$P(\text{Accuracy}) = \text{Abs}(\vec{P} - \vec{P}_{ep}) \tag{12}$$

$$\vec{C} = z_t \tag{13}$$

z_t is a chaotic logistic map whose value lies in the range [0,1] and M is set to 2.

$$S(\vec{A}) = \left(\sqrt{f \cdot e^{-\frac{t}{l}} - e^{-t}} \right)^2 \tag{14}$$

Where e denotes an expression function, t is the iteration count. l, f are the control parameters for maintaining exploration and exploitation of penguins in such a manner that the lower and upper bounds of l, f are set to as follows: $l \in [1.5, 2]$ and $f \in [2, 3]$.

Part-IV: Re-compute and update to the best position of the penguin.

$$\vec{P}_{ep}(t + 1) = \vec{P}(t) - \vec{A} \cdot \overrightarrow{Dist_{ep}} \tag{15}$$

4.5 Chaotic Search and Rescue optimization (CSARO)

Searching is a mechanism to locate the existence of an object. The object may be in the form of food sources, any lost people or a hunter's prey etc. Rescue is a process of retrieving a person trapped in some problem and deliver them to a secured place. This nature inspired meta heuristic algorithm is usually a group activity that consists of several types such as mountain rescue, ground search and rescue, urban search and rescue, air-sea rescue and combat search. The vital parameter for searching is to explore a clue and then tracking the desired entity based upon the direction of the clue. These clues are stored in the form of $N \times D$ dimensional matrix C with N denoting the number of humans and D being the problem dimension [34].

$$C = \begin{bmatrix} X \\ M \end{bmatrix} \tag{16}$$

Where $X = \begin{bmatrix} x_{11} & \dots & x_{1D} \\ \vdots & \dots & \vdots \\ x_{N1} & \dots & x_{ND} \end{bmatrix}$ is the position matrix

and

$$M = \begin{bmatrix} M_{11} & \dots & M_{1D} \\ \vdots & \dots & \vdots \\ M_{N1} & \dots & M_{ND} \end{bmatrix}$$
 is the Memory Matrix

x_{N1} represents the position of 1st dimension for N^{th} human. Clues are categorised into two types:-



1. **Hold Clue:** It takes into account any one member of the search group and tries to explore its surroundings.
2. **Abandoned Clue:** It implies that even though the clue is applied for searching but it seems to be inappropriate and further hunts for some better path finding clues.

❖ **2-Phase operations of CSARO**

▪ **Phase-1: Social Phase:-**

It is basically a position specific searching where the probability of getting significant result is maximum.

Therefore, the search direction is computed as follows:

$$SD_i = (X_i - C_k), k \neq i \quad (17)$$

Where SD is the search direction

X_i is the position of i^{th} human

C_k is the position of k^{th} clue

k is a random integer ranging between 1 and $2N$. To prevent the movement from other clue we have chosen $k \neq i$.

The most important constraint in group search is that there is no repetition of search location and the pattern of movement among the group members are limited.

Mathematically, we can design the social phase equation as follows:-

$$X'_{ij} = \begin{cases} C_{kj} + r_1(X_{ij} - C_{kj}), & \text{if } f(C_k) > f(x_i) \\ X_{kj} + r_2(X_{ij} - C_{kj}), & \text{otherwise} \end{cases} \quad \forall j = 1 \text{ to } D$$

$$X_{ij}, \text{ otherwise} \quad (18)$$

Where X'_{ij} is the new position of the j^{th} dimension for the i^{th} human

C_{kj} j^{th} position of the k^{th} clue.

$f(C_k)$ $f(x_i)$ are the Objective function.

r_1 and r_2 are the random numbers between $[0,1]$.

▪ **Phase-2: Individual Phase:-**

Here the searching operation is independent of the position and the clues are usually proposed by others who are not the members of the group. The individuals carry out the search operation within the radius of the current position.

Therefore, the new position of the i^{th} person is numerically computed as follows:

$$X'_i = X_i + r_3(C_k - C_m), i \neq k \neq m \quad (19)$$

Where k and m are random integers between 1 and $2N$.

r_3 is a random number between $[0,1]$ with a uniform distribution.

❖ **Boundary Control**

Most of the meta heuristic algorithms expect the solutions of a problem to be confined within the given solution space. Therefore, the new position can be modified using the following equation:

$$X'_{ij} = \begin{cases} \frac{X_{ij} + X_j^{\max}}{2}, & \text{if } X'_{ij} > X_j^{\max} (\forall j = 1..D) \\ \frac{X_{ij} + X_j^{\min}}{2}, & \text{if } X'_{ij} < X_j^{\min} \end{cases} \quad (20)$$

Where X_j^{\max}, X_j^{\min} is the maximum and minimum threshold value of the j^{th} dimension.

❖ **Updating new position and information**

After completion of each iteration if the new value obtained is better than the previous value then the algorithm replaces the memory value new better value otherwise it remains unchanged. The new value is computed using equation (36).

$$M_N = \begin{cases} X_i, & \text{if } f(x'_i) > f(x_i) \\ M_N, & \text{otherwise} \end{cases}$$

And $X_i = \begin{cases} X'_i, & \text{if } f(x'_i) > f(x_i) \\ X_i, & \text{otherwise} \end{cases} \quad (21)$

Where M_N is the position of the n^{th} stored clue in the memory matrix. N is a random number ranging between 1 and N .

Memory updating property of CSARO algorithm allows it to reach the global optimum at a faster rate.

❖ **Abandoned Clue computation**

For numerical computation of unsuccessful search number (USN) the equation is designed as follows:

$$USN_i = \begin{cases} USN_i + 1, & \text{if } f(x'_i) < f(x_i) \\ 0, & \text{otherwise} \end{cases} \quad (22)$$

If the searching yields better value then it is set to 0. Otherwise it will be increased by 1.

If USN_i exceeds maximum unsuccessful search then,

$$X_{ij} = X_j^{\min} + r_4(X_j^{\max} - X_j^{\min}), (\forall j = 1..D) \quad (23)$$

Where $r_4 \in [0,1]$.

❖ **Control parameters of SAR**

There are basically two control parameters of CSARO i.e SE (social effect) that ranges between $[0,1]$ and MU (Maximum unsuccessful number) whose values ranges between $[0, 2T_{\max}]$ where T_{\max} is the maximum number of iterations. SE is set to 0.05 so as to balance the global search ability and the convergence rate. The value of MU is computed using the expression

$$MU = 70 \times D \quad (24)$$

❖ **Optimal chaos theory**

If we will introduce chaos optimal theory to improve the parameter performance then a popular logistic function, used in chaotic map can be employed as follows:

$$z_{t+1} = 4 \cdot z_t(1 - z_t) \quad (25)$$

Where z_i denotes the chaotic map value obtained at ' t ' iteration.



The logistic function transforms a continuous value to a discrete binary value defined as:

$$\text{logsig}(x_j^m) = \frac{1}{1 + \exp(-x_j^m)} \quad (26)$$

$$x_j^m = \begin{cases} 1, & \text{if } \text{logsig}(x_j^m) \geq 0.5 \\ 0, & \text{otherwise} \end{cases} \quad (27)$$

Therefore, upon replacing the random variable by chaotic function z_t we can modify equation (33) as follows:

$$X'_{ij} = \begin{cases} \begin{cases} C_{kj} + z_t(X_{ij} - C_{kj}), & \text{if } f(C_k) > f(x_i) \\ X_{kj} + z_t(X_{ij} - C_{kj}), & \text{otherwise} \end{cases} \\ \forall j = 1 \text{ to } D \\ X_{ij}, & \text{otherwise} \end{cases} \quad (28)$$

Where z_t is the t^{th} chaotic value within the range [0,1].

4.6 Chaotic Search and Rescue optimization Algorithm

Step-1: Initialize the population(N) randomly.

Step-2: Generate 2N solutions uniformly distributed in the range $[X_j^{\text{max}}, X_j^{\text{min}}]$ ($\forall j = 1..D$).

Step-3: Sort the solution in the descending order and find the best position (X_{best}).

Step-4: Divide the solution into two parts such that: the first half of the sorted solution is dedicated for human position matrix (X) and the rest half for the memory matrix (M).

Step-5: Set the algorithm control parameters
 $SE(\text{social effect}) = 0.05$

$$MU(\text{maximum unsuccessful search dimension}) = [0, 2T_{\text{max}}]$$

Set USN_i (Unsuccessful search number dimension) = 0

Where $i=1..2N$.

Step-6: While stop criteria not satisfied
Do

Step-7: For $i = 1, 2 \dots N$.
Do

Step-8: Social Phase:

$$C = \begin{bmatrix} X \\ M \end{bmatrix} \quad \text{From equation (16)}$$

Step-9: $SD_i = (X_i - C_k)$, k is randomly selected in such a way that $i \neq k$.

Step-10: $j_{\text{rand}} = \text{rand int}[1, D]$

Step-11: $z_i = \text{rand}[0,1]$, the chaotic variable is assigned values between [0,1].

Step-12: For $j = 1, 2 \dots D$
Do

Compute equation (18).

Step-13: Update the new position using equation (20).

Step-14: End of For loop

Step-15: Compare the new value with the previous Value. Update the position as well as search information using equation (21).

Step-16: Evaluate the USN_i using equation (22).

Step-17: Individual Phase

$$C = \begin{bmatrix} X \\ M \end{bmatrix} \quad \text{from equation (16)}$$

Step-18: Compute equation (17) for new position.

Step-19: For $j = 1..D$ do

Evaluate equation (20)

Step-20: End For

Step-21: Calculate the variable values using equation (20), (21),(22) for individual phase once again.

Step-22: Check if $USN_i > MU$ do

For $j = 1..D$ do

$$X_{ij} = X_j^{\text{min}} + \text{rand}[0,1](X_j^{\text{max}} - X_j^{\text{min}}), (\forall j = 1..D)$$

Step-23: End For

Step-24: $USN_i = 0$

Step-25: End if

Step-26: End For

Step-27: Find the current best position and update X_{best} .

Step-28: End while

Step-29: Return X_{best} .

Step-30: End.

V. EXPERIMENTAL SET-UP

5.1 Dataset Description

The datasets belong to Kent Ridge Biomedical Dataset repository. Leukaemia, colon tumour, ovarian cancers are of binary class and the remaining dataset are of multi-class type [23]. The datasets are normalised using mi-max normalization to confine the values within the range of [0, 1].

$$Y = \frac{X - \text{min}}{\text{max} - \text{min}} \quad (29)$$

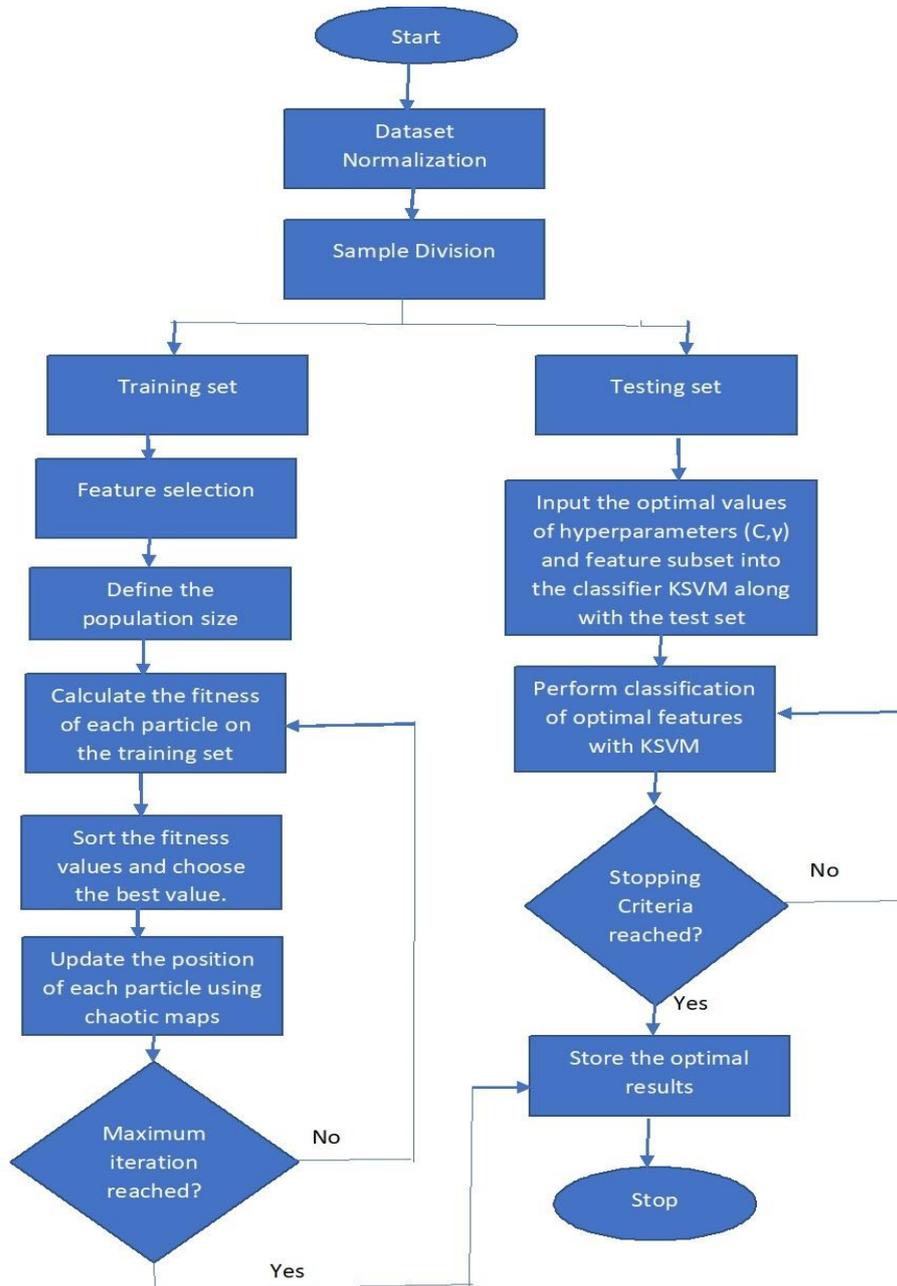


Fig:2 Working Model of Algorithm

Fig.2 Workflow Diagram of the Hyper heuristic learning model

Table-1 Dataset Description with class types

Sl_No.	Dataset Name	No. of Attributes	No. of Instances	Class Type
1	Colon	2000	62	2
2	Ovarian	15154	253	2
3	Leukaemia	7129	72	2
4	Lung cancer	12,600	203	5
5	ALL-AML	7129	72	3
6	SRBCT	2308	83	4
7	Lymphoma	4026	62	3

5.2 Performance Metrics

To judge the proficiency of applied model we need to evaluate various [25].

- Confusion Matrix:** It is a 2-dimensional representation for evaluating the accuracy, sensitivity, specificity, confusion matrix, F-measures of a classification model. The details are demonstrated in Table-3.

	True class		Measures	
	Positive	Negative		
Predicted class	Positive	True positive TP	False positive FP	Positive predictive value (PPV) $\frac{TP}{TP+FP}$
	Negative	False negative FN	True negative TN	
Measures	Sensitivity $\frac{TP}{TP+FN}$	Specificity $\frac{TN}{FP+TN}$	Accuracy $\frac{TP+TN}{TP+FP+FN+TN}$	

Fig.3 Confusion Matrix

- **Classification Accuracy:** It is the ratio of number of correct predictions to the total number of predictions made.

VI. SIMULATION AND RESULTS

A. Results of Feature Reduction

Kernel PCA and kernel LDA are employed to remove the redundant features as well noise.

Table-2 Feature Reduction Result

Sl_No	Dataset_Name	No. of Attributes	After KPCA	After KLD
1	Colon	2000	41	2
2	Ovarian	15154	72	2
3	Leukaemia	7129	55	3
4	Lung cancer	12,600	61	2
5	ALL-AML	7129	68	2
6	SRBCT	2308	83	3
7	Lymphoma	4026	62	2

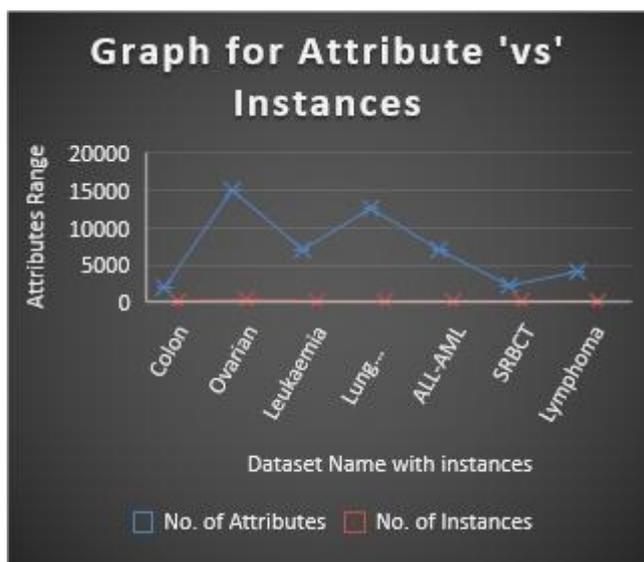


Fig. 4 Plotting of Attributes and samples

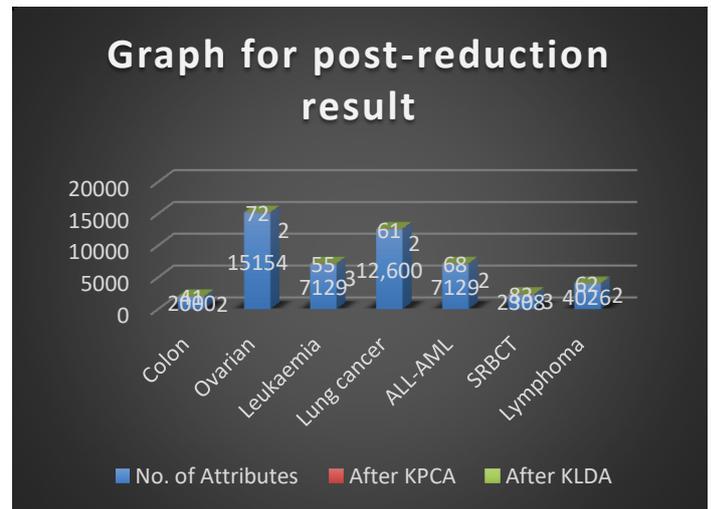


Fig.5 Representation of Attributes after feature reduction

B. Results of Classification

To analyse the effect of the suggested hyper heuristic model we have taken the help of the publicly available seven microarray datasets such as Colon, Ovarian, Leukaemia, Lung cancer, ALL-AML, SRBCT, Lymphoma. Table 3 represents the six performance The two parameters C, γ are represented in the search space and the diagrammatic view of the functionality of the learning model is also clearly mentioned. Table-4 represents the Parameter settings and the possible range of values for computing the performance metrics after 10 runs. Table-5 denotes the running time of the various algorithms implemented in this work. Fig. 4 and Fig.5 depicts the graph for attributes and number of features selected before and after reduction process. According to our simulation Lung cancer has an accuracy of 98.7% followed by Lymphoma and ALL-AML having 98.2% accurate classification. Fig.6 to Fig. 10 represents the graph for Leukaemia, Colon cancer, Ovarian cancer, ALL-AML and Lung Cancer.

VII. CONCLUSION AND FUTURE WORK

An attempt for integrating numerical computation with the help of a Hyper heuristic algorithm is designed and simulated using Scikitlearn. After certain repetitions of generations, it is observed that the rate of convergence and the stability of the results is uniformly maintained. The learning model optimises the control parameters (C, γ) of SVM using two optimization algorithm such as CSARO and EPO for local search and global search respectively.

The future work is directed towards employing of Deep learning, deep convolution network and some Quantum computing mechanism to improve the parameter efficiency. Instead of optimising the control parameters, many other parameters like population size, maximum iterations, movement velocity, energy loss, radiations of light source etc. can also be tried for optimization.

Table-4 Parameter setting

Sl-No.	Parameter Setting	Range of values
1	Exploration control parameter in EPO(l)	[1.5,2]
2	Exploitation control parameter in EPO(f)	[2,3]
3	Spiral movement parameter M	2
4	Population size (N) (both in EPO and CSARO)	100
5	Maximum Iterations (both in EPO and CSARO)	100
6	SE (social effect)	0.05
7	MU(max unsuccessful search dimension)	$[0, 2T_{max}]$
8	USN_i (unsuccessful search number)	$USN_i = 0$
9	KSVM control parameters(C, γ)	[0.5,1]

Table-5 Comparison of Running time (in seconds) of algorithms on seven datasets

		Run time (in seconds)		
		EPO+SVM	PSO+SVM	KSVM+CSARO+CEPO
1	Colon	164.45	179.32	153.213
2	Ovarian	78.32	82.42	68.323
3	Leukaemia	174.11	183.26	171.43
4	Lung cancer	255.47	259.32	233.41
5	ALL-AML	185.34	195.43	181.15
6	SRBCT	173.47	183.23	170.22
7	Lymphoma	74.53	84.32	70.34

Table-3 Performance Metrics Evaluation

Sl_No	Dataset Name	Accuracy	Recall	Precision	F-Score	Kappa measures	MCC
1	Colon	0.953	0.979	0.971	0.972	0.951	0.921
2	Ovarian	0.961	0.968	0.981	0.979	0.972	0.951
3	Leukaemia	0.931	0.934	0.878	0.902	0.939	0.914
4	Lung cancer	0.973	0.981	0.983	0.984	0.959	1
5	ALL-AML	0.971	0.973	0.969	0.971	0.979	0.982
6	SRBCT	0.972	0.974	0.975	0.977	0.966	0.955
7	Lymphoma	0.985	0.97	0.985	0.983	0.973	0.962

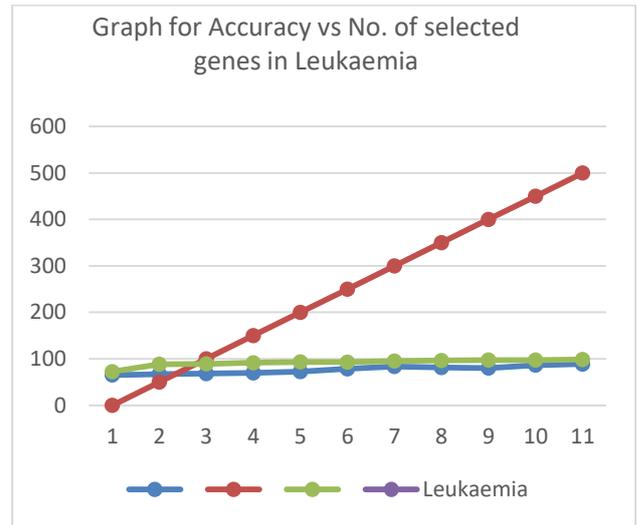


Fig:6 Representation of Accuracy ‘vs’ No. Of selected genes in Leukaemia

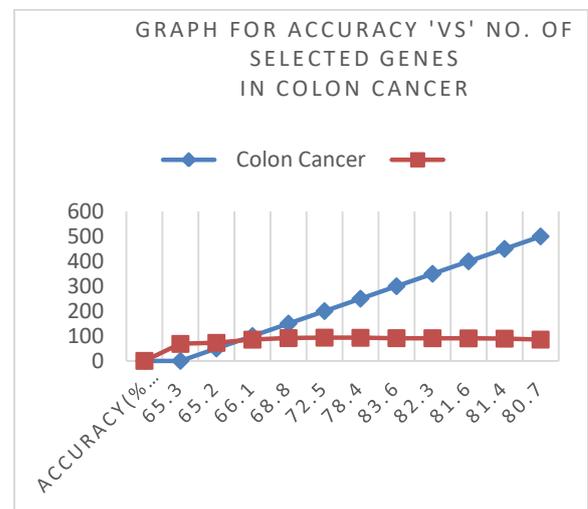


Fig:7 Representation of Accuracy ‘vs’ No.of selected genes in Colon cancer

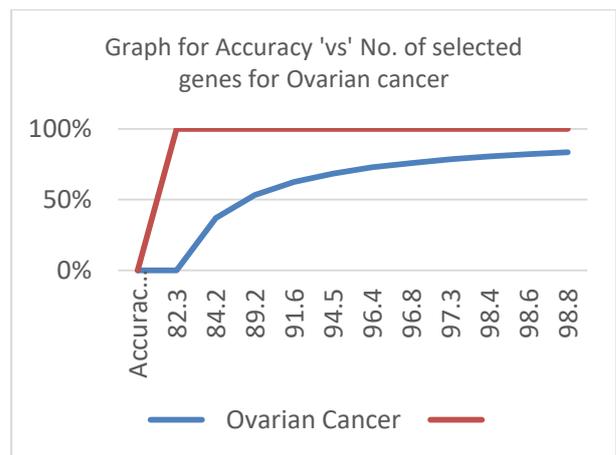


Fig:8 Representation of Accuracy ‘vs’ No.of selected genes in Ovarian cancer

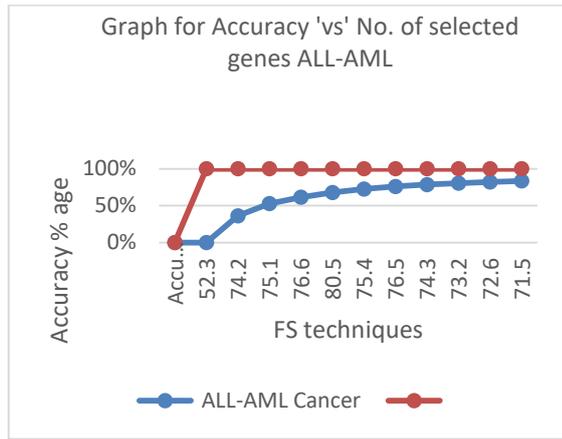


Fig:9 Representation of Accuracy 'vs' No. of selected genes in ALL-AML

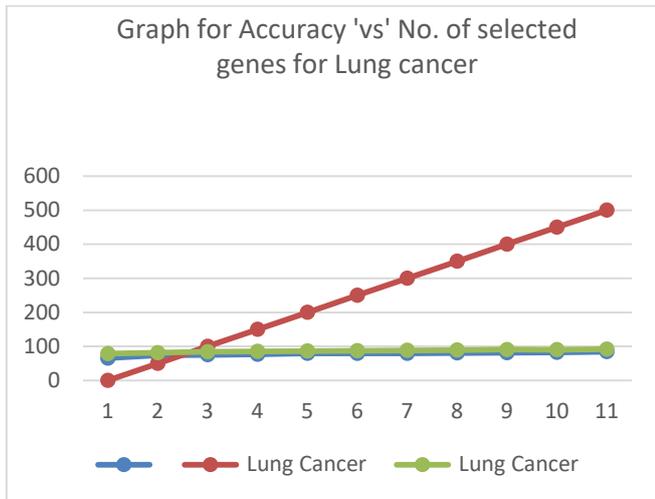


Fig:10 Representation of Accuracy 'vs' No. of selected genes in Lung cancer

REFERENCES

1. Maulik, Ujjwal. "Analysis of gene microarray data in a soft computing framework." *Applied Soft Computing* 11, no. 6 (2011): 4152-4160.
2. Tom M. Mitchell, McGraw-Hill, 1997, ISBN 0071154671, 9780071154673.
3. V. Boln-Canedo, N. Snchez-Maroo, A. Alonso-Betanzos, Distributed feature selection: an application to microarray data classification, *Appl. Soft Comput.*30 (2015) 136–150.." Artificial intelligence in medicine 31, no. 2 (2004): 91-103.
4. Dash, Manoranjan, and Huan Liu. "Feature selection for classification." *Intelligent data analysis* 1, no. 3 (1997): 131-156.
5. E.K. Tang, P.N. Suganthan, X. Yao, Feature selection for microarray data using least squares SVM and particle swarm optimization, in: 2005 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology, IEEE, 2005, pp. 1–8.
6. Momin, Bashirahamad F., Sushmita Mitra, and Rana Datta Gupta. "Reduct generation and classification of gene expression data." In *2006 International Conference on Hybrid Information Technology*, vol. 1, pp. 699-708. IEEE, 2006.
7. Bair, Eric, and Robert Tibshirani. "Semi-supervised methods to predict patient survival from gene expression data." *PLoS Biol* 2, no. 4 (2004): e108.
8. L. Zhang, P.N. Suganthan, Benchmarking ensemble classifiers with novel co-trained kernel ridge regression and random vector functional link ensembles[research frontier], *IEEE Comput. Intell. Mag.* 12 (4) (2017) 61–72.
9. E.-G. Talbi, L. Jourdan, J. Garcia-Nieto, E.Alba, Comparison of population based metaheuristics for feature selection: Application to microarray data classification, in: 2008 IEEE/ACS International Conference on Computer Systems and Applications, IEEE, 2008, pp. 45–52, <http://dx.doi.org/10.1109/AICCSA.2008.4493515>.

10. Petrovski, A., McCall, J. A.: Multi-objective optimization of cancer chemotherapy using evolutionary algorithms. *Evolutionary Multi-Criterion Optimization*, Lecture Notes in Computer Science, Volume 1993, Zitzler E., et al. (Eds.) Springer, Berlin, Heidelberg (2001), pp.531-45.
11. Z. Zhu, Y.-S. Ong, and J. M. Zurada, "Identification of full and partial class relevant genes," *IEEE/ACM Trans. Comput. Biol. Bioinformatics.* 7, no. 2, pp. 263–277, Apr. 2010.
12. Q. Gu, Z. Li, J. Han, Generalized Fisher score for feature selection, in: *Proceedings of the Twenty-Seventh Conference on Uncertainty in Artificial Intelligence*, AUAI Press, 2011, pp. 266–273.
13. J. Cao, Z. Lin, G.-B. Huang, and N. Liu, "Voting based extreme learning machine," *Information Sciences*, vol. 185, no. 1, pp. 66–77, 2012.
14. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Trevor Hastie, Robert Tibshirani, Jerome Friedman Springer Science & Business Media, 2013, ISBN 0387216065, 9780387216065.
15. Garro, Beatriz A., Katya Rodríguez, and Roberto A. Vázquez. "Classification of DNA microarrays using artificial neural networks and ABC algorithm." *Applied Soft Computing* 38 (2016): 548-560.
16. Paul, Topon Kumar, and Hitoshi Iba. "Selection of the most useful subset of genes for gene expression-based classification." In *Proceedings of the 2004 Congress on Evolutionary Computation (IEEE Cat. No. 04TH8753)*, vol. 2, pp. 2076-2083. IEEE, 2004.
17. A. Wang, N. An, G. Chen, J. Yang, L. Li, G. Alterovitz, Incremental wrapper based gene selection with Markov blanket, in: *Bioinformatics and Biomedicine (BIBM)*, 2014 IEEE International Conference on, IEEE, 2014.
18. Xia, Yong, Zexuan Ji, Andrey Krylov, Hang Chang, and Weidong Cai. "Machine learning in multimodal medical imaging." *BioMed research international* 2017 (2017).
19. El Naqa I, Ruan D, Valdes G, Dekker A, Mc-Nutt T, Ge Y, et al. Machine learning and modeling: Data, validation, communication challenges. *Med Phys.* 2018 Oct; 45(10): 834–40.
20. H. Motieghader, A. Najafi, B. Sadeghi, A. Masoudi-Nejad, A hybrid gene selection algorithm for microarray cancer classification using genetic algorithm and learning automata, *Inf. Med. Unlocked* 9 (2017) 246–254.
21. M.M. Mafarja, S. Mirjalili, Hybrid Whale Optimization Algorithm with simulated annealing for feature selection, *Neurocomputing* 260 (2017).302–312, <http://dx.doi.org/10.1016/j.neucom.2017.04.053>.
22. Z.Y. Algamal, M.H. Lee, A two-stage sparse logistic regression for optimal gene selection in high-dimensional microarray data classification, *Adv. Data Anal. Classif.* (2018) 1–19, <http://dx.doi.org/10.1007/s11634-018-0334-1>.
23. Sharifi, Maryam, Arta A. Jamshidi, and Nazanin Namazi Sarvestani. "An adaptive robust control strategy in a cancer tumor-immune system under uncertainties." *IEEE/ACM transactions on computational biology and bioinformatics* 16, no. 3 (2018): 865-873.
24. Q. Zhang, Z. Li, C. Zhou, X. Wei, Bayesian network structure learning based on the chaotic particle swarm optimization algorithm, *Genet. Mol. Res.* 12 (4) (2013) 4468–4479.
25. Vellido, Alfredo, Elia Biganzoli, and Paulo JG Lisboa. "Machine learning in cancer research: implications for personalised medicine." In *ESANN*, pp. 55-64. 2008.
26. Zhang, Xuedong, Changzhi Wu, Jing Li, Xiangyu Wang, Zhijing Yang, Jae-Myung Lee, and Kwang-Hyo Jung. "Binary artificial algae algorithm for multidimensional knapsack problems." *Applied Soft Computing* 43 (2016): 583-595.
27. M. Dashtban, M. Balafar, Gene selection for microarray cancer classification using a new evolutionary method employing artificial intelligence concepts, *Genomics* 109 (2) (2017) 91–107.
28. A. Zibakhsh, M.S. Abadeh Gene selection for cancer tumor detection using a novel memetic algorithm with a multi-view fitness function *Eng. Appl. Artif. Intell.*, 26 (4) (2013), pp. 1274-1281.
29. A. Rosales-Pérez, S. García, H. Terashima-Marin, C.A.C. Coello, F. Herrera, MC2ESVM: Multiclass classification based on cooperative evolution of support vector machines, *IEEE Comput. Intell. Mag.* 13 (2) (2018) 18–29, <http://dx.doi.org/10.1109/MCI.2018.2806997>.
30. Zhu, Zexuan, Yew-Soon Ong, and Jacek M. Zurada.



- "Identification of full and partial class relevant genes." *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 7, no. 2 (2008): 263-277.
31. M. Lozano, C. Garca-Martnez, Hybrid metaheuristics with evolutionary algorithms specializing in intensification and diversification: overview and progress report, *Comput. Oper. Res.* 37 (3) (2010) 481–497.
 32. Martin, R., Teo, K.: *Optimal Control of Drug Administration in Cancer Chemotherapy*. World Scientific, Singapore New Jersey, London Hong Kong (1994).
 33. G.I. Sayed, G. Khoriba, M.H. Haggag, A novel chaotic salp swarm algorithm for global optimization and feature selection, *Appl. Intell.* (2018) 1–20.
 34. Shabani, Amir, Behrouz Asgarian, Saeed Asil Gharebaghi, Miguel A. Salido, and Adriana Giret. "A New Optimization Algorithm Based on Search and Rescue Operations." *Mathematical Problems in Engineering* 2019 (2019).
 35. J. Lv, Q. Peng, X. Chen, Z. Sun, Amulti-objective heuristic algorithm for gene expression microarray data classification, *Expert Syst. Appl.* 59 (2016) 13–19.
 36. Yang, Ruixin, John Tan, and Menas Kafatos. "A Pattern Selection Algorithm in Kernel PCA Applications." In *International Conference on Software and Data Technologies*, pp.374-387. Springer, Berlin, Heidelberg, 2006.