10

Performance Analysis of Nature-Inspired Algorithms in Breast Cancer Diagnosis

K. Hariprasath[1]*, S. Tamilselvi[2], N. M. Saravana Kumar [3], N. Kaviyavarshini[1] and S. Balamurugan[4]

[1]Department of Information Technology, Vivekanandha College of Engineering for Women, Namakkal, India
[2]Department of Biotechnology, Bannari Amman Institute of Technology, Erode, India
[3]Department of Artificial Intelligence and Data Science, M Kumarasamy College of Engineering, Karur, India
[4]Intelligent Research Consultancy Services (iRCS), Coimbatore, India

Abstract

Nature-inspired computing (NIC) is a fascinating computing paradigm that applies the methodology and approaches of nature that addresses various real-time complex problems ranging from how an organism finds its prey to genetic evolution. One of the unique features is that it has been provisioned with a decentralized control of computational activities naturally. In this chapter, to have a better insight of such NIC-based algorithms, the problem of identifying the breast cancer is used to exhibit their performances. Also, this chapter briefs about the application of stand-alone and hybridized approaches to identify the disease. Finally, it concludes with the experimental results and other statistical measures. The purpose of the chapter is to guide the new researcher in the area to get inspired from the conventional works and to bring out a new advanced approach that can perform further better. The three swarm algorithms are ant colony optimization, firefly, and particle swarm optimization algorithm. These swarm algorithms were used to optimize the support vector machine (SVM) which was trained to classify the malignant and benign images from the Wisconsin breast cancer dataset. With respect to the experimental results, it has been found that naïve algorithm with

**Corresponding author*: khariprasathit@gmail.com

S. Balamurugan, Anupriya Jain, Sachin Sharma, Dinesh Goyal, Sonia Duggal and Seema Sharma (eds.) Nature-Inspired Algorithms Applications, (267–294) © 2022 Scrivener Publishing LLC

10.1002/9781119681984.ch10 Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

PSO optimization demonstrates better discriminating property of the underlying conventional classifiers.

***Keywords*:** Nature-inspired computing, PSO optimization, Wisconsin breast cancer, malignant, benign

10.1 Introduction

Breast cancer is turning into a major and inevitable effect of women death across entire earth, in the interim, can be affirmed only at the initial disclosure in the pre-mature stages and an appropriate treatment of this deadly cancer will conform the prolong endurance in such sufferers. Around the world, breast cancer is considered as the second basic sort of disease by following lung carcinoma because 10.4% of disease rate also fifth basic reason for deaths due to cancer [22, 23]. In addition, it is, presently by a wide margin, the most well-known disease among ladies, the occurrence percentage is double than colorectal malignancy also and cervical malignant growth around multiple points than lung carcinoma. Anyway, the mortality rate of breast cancer overall is simply 25% more prominent than that of lung disease among women. Thus, it prompts many researchers who have been funded to model a system that automatically detects the breast cancer in its initial stages so that it can be treatable [19, 20].

The investigator has newly recommended numerous techniques by utilizing intelligence methodologies for infectious analysis; fewer sufferers fulfill the demand of good efficiency. More researches have been done about the analysis of breast cancer diagnosis in the past; this particular work confines in two aspects. First is to select the appropriate optimization technique that suits better and another is to ensure the best-in-class performance to detect the cancer autonomously from supervised to semi-supervised as a test-case analysis. The familiar swarm algorithms FA and PSO were used for analysis breast cancer. The statistical conclusion finds that swarm intelligent techniques are used for the purpose of breast cancer analysis along with high performance. In addition, PSO analyzes breast cancer for better efficiency compared to remaining swarm intelligent algorithms.

10.1.1 NIC Algorithms

Particle swarm optimization (PSO) is one of techniques that enhances a classifier to optimize itself and tends to pick an improved solution from

10.1002/9781119681984.ch10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

its alternative result set in accordance with known quality proportion and tries to optimize the candidate solution set of the particular iteration with the respective particle velocity and position by searching throughout the entire state space of particles. It is the most successful metaheuristic optimization algorithm used in variety of medical applications and one of the algorithms considered in the chapter to analyze the performance while hybridizing with other conventional classifiers [16–18].

Ant colony algorithm is widely deployed in data mining applications where ant colony clustering method is implemented in the synthesized correlated ant colony. One of the famous applications is that these types of artificial ant colony synthesizing approach has applied in solving traveling salesman problem used for local search in max-min ant systems. Also, to find minimum distance for data routing problem in heavily distributed networks and to exhibit accessing list for heavily decentralized database, systems deal with familiar ant swarm-based approach and scheduling. It has applications in manufacturing industry, market segmentation, feature selection, cloud computing, and resource allocation.

FA was developed for the purpose to tackle with complex issues of peculiar kind those were either does not fit into the true- or false-based conditions; perhaps, it deals with the values that are intermittent too. The basic idea of the algorithm relies on the natural characteristic trait of a firefly that flashes the guiding signals to other fireflies and the same phenomenon is followed by the other fireflies of the same group so as to follow the guiding signals to find prey or mates by minimal effort by the firefly group. The fireflies are closely having similar association in its behavior to that of swarms in terms of message communication, self-organization, and regionalized decision-attaining ability. It seems well suited for solving NP-hard problems, classification in complex datasets problems, multiobjective searches, parallel computational domains, combinatorial computational domains, and incessant search space domains. Because of ability of FA that would efficiently handle continuous data, it is widely used for solving manufacturing drugs, image compression, stock market price forecasting, load dispatch problems, and other biomedical researches studying about cell mutation at DNA sequence level.

The bat algorithm is one of the lately established NIC-based learning algorithms, which exploits the ability of exploring the new routes by using its ability to sense and follow the echo-based location prediction mechanism similar to that of bats and few other nocturnal animals to obtain solutions for a continuous indeterminist search space problems. Because of such ability, it is widely employed in obstacle tracking robots, autonomous cars, game planning, remote wireless multi-hop network management,

10.1002/9781119681984.ch10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

prioritization problems, feature selection problems, global search problems, and resource scheduling problems. Cuckoo search algorithm depends on generation of cuckoo fowls. Cuckoos are a group of winged creatures with one of a kind conceptive system progressively forceful contrasted with other flying creature's species. Some of cuckoo winged creature's species like Ani and Guira laid eggs on common homes and expel others eggs to build incubating likelihood of their own eggs. Different species use brood parasitism technique for laying their eggs in the homes of different flying creatures or host homes. The parasitic cuckoos are acceptable in donning homes where eggs have quite recently been laid and their planning of laying eggs is precisely. They lay one egg in the host home which will ordinarily incubate faster than different eggs.

Harmony Search Algorithm (HSA) is a fair mix of exploring in the continuous undesirable space search and exploiting the well-known trained cases to classify the cases, i.e., to pick the better candidate solution. Above all, it is very straightforward, robust, and simplicity to implement and ease of use. The HAS is enlivened by the procedure of melodic execution, comprises of three administering criterions: arbitrary inquiry, amicability memory thinking about the criteria of the particular harmony, and appropriate pitch modifying rule. These characteristic features make HSA a very special nature-inspired computing (NIC) algorithm that predominantly employed as a metaheuristic optimizing algorithm.

This particular chapter can be foreseen as a common stage that used to exhibit the individual performances of such algorithms is by detecting the breast cancer malignant figures from Wisconsin breast cancer dataset. This work also quickly clarifies about the use of individual and hybridized ways to deal with the identification of the disease. At long last, it finishes up with the trial results and other factual measures. The purpose of the chapter is to guide the new researcher in the area to get inspired from the conventional works and to bring out a new advanced approach that can perform further better.

10.2 Related Works

Dhahri *et al.* [1] recommended a study based on algorithms of machine learning which has an objective to develop a model to compare the tumours of breast between benign and malignant and to learning algorithm is optimized. The perfect features and parameters of the classifier of machine learning are selected by the technique of genetic algorithm. This approach concentrates on the parameters like accuracy, sensitivity,

10.1002/9781119681984.ch10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

precision, ROC curves, and specificity. This study demonstrates that programming of genetic has the ability to automatically detect the preeminent approach with the help of algorithm of classifiers and methods of preprocessing. This approach has three important phases. In the first phase, three evolutionary algorithms have same result after the performance of effective configuration. In second phase, it focuses on feature selection method which will progress the performance of accuracy. In the third phase, the supervised classifier of the machine learning is utilized to design pattern that performs automatically.

In programming of genetic algorithm, it is resolved the problem of hyper parameter that performs the challenges of algorithm in machine learning. This proposed algorithm is used to select the suitable algorithm from different types of configurations and this research is done on the library of python. The estimation of ensemble method for the technique of machine learning using comprehensive method and the rate of time consumption is high. This model is best suitable for settings of parameter control of algorithms in machine learning at one side and automatic diagnosis of breast cancer on other side. Determination of model which is best suited for the data and designing the model by automatic and adjustment method of parameter are the two major challenges of this approach.

Kumar [2] proposed that breast cancer has high mortality rate as it is a significant reason between the women but diagnosis of breast cancer manually takes long time. Diagnosis of breast cancer automatically was designed to decrease the time through manually and spread rate of the cancer cells is decreased. This approach is trained as well as tested on dataset Wisconsin using four different types of machine learning algorithm, namely, support vector machine (SVM), naive Bayes, logistic regression, and k-nearest neighbor. This paper concentrates about the comparison of sensitivity, specificity, and accuracy. After comparison analysis is done by the result of accuracy and the finest algorithm is selected. The various types of hyper parameter used are selected manually. The paper presented results that SVM gives the best performance. The performance of machine learning algorithm is done on texture and cell radius with more accuracy than diagnosis done manually.

Idri *et al.* [3] recommended that breast cancer is considered as the primary sources in women death. Various result-based assisting methods were suggested to support cancer treating doctors to analyze meticulously the sufferer. These result-based assisting methods majorly used here is machine learning approach to segregate the analyzation of the mild or benignant cancer. This paper discussed the classification and investigation of the influence of attributes calibration on the accurate three well-defined machine learning approaches. They are SVM, multi-layer perception, and decision trees.

10.1002/9781119681984.ch10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The three attribute calibration methods used in this paper are PSO, grid search, and Weka software. The datasets are gathered from the machine learning repository. The global outcomes recommend that applying PSO and grid search helps to develop the specific classifiers, and thereby, the doctors can do the effective analyses.

Arafi *et al.* [4] introduced the machine learning approach about the breast cancer data evaluation and categorization, using PSO and SVM. This approach utilizes SVM as design for machine learning approach with the target of reducing standardization inaccuracy error. PSO is a technique for inevitable decision of top first element of two attributes in SVM. Its efficiency in resolving identification issues is analytically proved for realistic benchmark datasets. The final conclusion values are distinguished to the required four techniques using three various methods of efficiency.

This paper proposes the framework and a few exploratory decisions of machine learning techniques toward breast cancer. This technique is in accordance with SVM for categorizing the whole data with the motive of developing the efficiency of the out-coming classifier. Based on this paper, kernel is used also attributes that are systematically kept by applying PSO method.

This technique was executed in Linux PC utilizing LIBSVM library and C++ language. Subsequent programming device had tentatively tried on original benchmark of breast cancer dataset index openly accessible at popular VCI archive of ML. Data collection was partitioned into three distinct manners and into three subsets of various sizes, and substance was individually used for practicing, for approval, and for testing information. Test results had been contrasted with acquired of similar information models by four different strategies utilizing three various measures of efficiency. The observation values show that the proposed technique can additionally upgrade.

Purwaningsih *et al.* [5] suggested that there are a few investigations in the medical area that order information to analyze and examine choices. To investigate breast cancer, this examination analyzes two techniques: the neural network technique and the SVM technique depend on PSO which recommends deciding most efficiency result in the Coimbra dataset information. To execute the neural network and SVM strategy depending on PSO, Rapid Miner programming is utilized. At that point, the application result values are looked at utilizing Confusion Matrix and ROC curve. In light of the efficiency of the two models, it is realized that PSO-based neural network model has a highest efficiency estimation of 84.55% than the consequences of the PSO-based SVM with an efficiency estimation of 80.08%.

10.1002/9781119681984.ch10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The computation results and the precision of the AUC execution were got by the conclusion of the experiments; the two techniques are PSO-based neural network with AUC estimation of 0.885 and PSO-based SVM with an estimation of 0.819 comprised for the class of classification.

Data collecting systems utilized in this paper using auxiliary data, where the dataset utilized was included from the UCI machine learning repository and planned by leading PSO and SVM-based neural network method to decide most significant level of efficiency in breast cancer coordination.

The phases of researches for the categorization of breast cancer comprised of the following:

1. Preprocessing information is finished by erasing records that are missed elements and copies. Discretization systems are a piece of data transformed that is utilized to change the information types.
2. After the information is transformed, at that point, the information is handled and verified by Rapid Miner utilizing the neural network and SVM model which depends on PSO.
3. Moreover, this work creates an efficiency level, in particular Confusion Matrix, to analyze efficiency esteems in the SVM technique with the neural network technique depends on PSO.

The technique given in this paper is about the SVM technique and PSO-based neural network strategy. In the process of the underlying dataset, it is through verifying and studying the information, the dataset is changed into 0 and range 1, at that point, the dataset is split into 10-fold cross-validation. The dataset was verified utilizing the SVM and PSO-based neural network technique to approve the model efficiency. This technique has a high level of efficiency compared to the PSO-based SVM. In this way, the PSO-based neural network technique can be utilized for categorization in breast cancer investigation.

Mazen *et al.* [6] proposed model that is partitioned into two phases. From the start organize, firefly algorithm has applied; further, it produces utmost results as beginning populace of genetic algorithm in subsequent stage. At end of second stage, the utmost results created by genetic algorithm have viewed as utmost results at all and values produced by genetic algorithm for mean squared error are likewise studied low estimation of fitness function. This paper explored effectiveness of using genetic algorithm–based firefly algorithm which hybridizes the arrangement development technique of firefly algorithm and genetic algorithm is a

10.1002/9781119681984.ch10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

meta-heuristic enhancement method of practicing neural systems improving the loads among the layers and also biases of neuron organize in order to limit wellness function. The development pattern surveyed by various assessment criteria and contrasted and models improved utilizing meta-heuristic algorithms like biogeography-based optimization, PSO, and ant colony optimization (ACO). Final decision outcome demonstrates that optimized pattern utilizing the genetic algorithm–based firefly algorithm beats other metaheuristic algorithms in accomplishing increased efficiency and decreased mean squared error.

Fallahzadeh *et al.* [7] proposed that the pathology as a typical diagnosis evaluation of cancer is an intrusive, tedious, and, in part, idiosyncratic technique. Thus, optical methods, particularly Raman spectroscopy, have engaged in consideration in the diagnosis of cancer analysts. Nonetheless, Raman spectra consists of various pinnacles associated with atomic bounds of the patterns, evaluating that the best attributes identified with dangerous changes can improve the efficiency of analysis. The current research endeavored to enhance the intensity in the cancer which is based on the Raman spectroscopy for the analysis to examine the utmost Raman attributes utilizing ant colony algorithm.

Based on current analysis, 49 spectra had been estimated among the ordinary, generous, and even the carcinogenic tissues of the breast tests, utilizing very low measurable Raman framework. Subsequent to preprocessing for the reduction in non-harmonious foundation luminance and force, about 12 significant Raman groups among the natural sampling were extricated from the attributes in every range.

At that point, the ACO algorithm was used to locate the optimal attributes for analysis. As the outcomes illustrated, by choosing five attributes, the categorization efficiency of the typical, favorable, and harmful class expanded by 14% and arrived at 87.7%. ACO attribute collection can improve the symptomatic efficiency of Raman-based analysis models.

10.3 Dataset: Wisconsin Breast Cancer Dataset (WBCD)

In experimental analysis, the algorithms were tested on the datasets of Wisconsin breast cancer from repository of UCI machine learning. The datasets of Wisconsin breast cancer are composed of 10 different attributes. The samples were records collected from 699 Fine Needle Aspirates (FNA) from tissue of human breast. They were simply categorized into two

10.1002/9781119681984.ch10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

classes, namely, Benign and Malignant. From acquired dataset of WBC, it is found that about 65.5% of the samples are benign and the rest of the 34.5% have been classified as malignant by the doctors while diagnosing. The attributes of the dataset are cell marginal adhesion, thickness, bare nuclei, normal metric and nucleoli, bland chromatin and shape of cell with homogeneity and size of cell with homogeneity, and single epithelial. It is also found to that the WBC dataset is in complete since 16 records were missing. In our experiments, the images of these 16 records were excluded and the remaining 683 samples were used.

10.4 Ten-Fold Cross-Validation

The procedure includes arbitrarily separating the dataset of Wisconsin breast cancer into 10 gatherings with around equivalent size. For testing, the main fold is kept and the prototype is prepared on 10 folds. The procedure is iterated for n intervals and in each interval diverse fold or an alternate gathering of data arguments is used for justification. Cross-validation is a resampling way implemented to assess machine learning on a restricted data model.

$$Cross-Validation(cv_{(n)}) = \frac{1}{n}\sum_{j=0}^{n} MSE_i \tag{10.1}$$

Cross-validations are essentially utilized in techniques of machine learning for estimation of the skill of a model in machine learning on hidden dataset. That is, to utilize a restricted example so as to gauge how the model is relied upon to perform all in all when used to make forecasts on data not utilized during the training of the sample.

10.4.1 Training Data

The procedure involves training the model in dataset; it can differ however ordinarily can utilize 60% of the accessible data to train.

10.4.2 Validation Data

The well-performed data is selected and it is made run in the validation data. Usually, 10% to 30% of data belong to subgroup. The model is said to be over fitting when it consists of more errors.

10.1002/9781119681984.ch10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

10.4.3 Test Data

Test data is also termed as holdout data. The dataset can hold data that will not be implemented in the data training. Test data advantage is that it has absolute sample estimation. This may range from 5% to 20%.

10.4.4 Pseudocode

Step 1: Train data is divided into 10 portions
Step 2: For i = 0 to 9
Step 3: Network is trained by 9 parts
Step 4: Accuracy is computed by 1 part
Step 5: Average of accuracy computed for 10 runs

10.4.5 Advantages of K-Fold or 10-Fold Cross-Validation

- Time of computation is diminished as it repeats the procedure just for 10 intervals when the estimation for n is 10.
- Minimum bias value.
- Information focuses acquire the chance that is tried precisely after and utilized in preparing 1-n intervals.
- There will be difference in subsequent gauge is diminished and n increments.

10.5 Naive Bayesian Classifier

Naive Bayesian is one of the techniques that belong to classification which was introduced by Thomas Bayes (1702–1761). Naive Bayesian classifier is said to be statistical classifier and it is supervised learning [24–28]. Naive Bayesian is one of inductive learning algorithm which is most effective and efficient learning algorithm for data mining and machine learning. Naive Bayesian classifier is a straightforward probabilistic type of classifier that is reliant on applying theorem of Bayes along with robust of objectivity prospects. The goal of the performance is to predict the test class instance with more accuracy. The Naive Bayesian classifier is based on two simple prospects. They are predictive attributes that are tentatively unconventional in the class and the result of non-categorical attribute is commonly separated inside the class. There are many different ways to treat the discrete and continuous attributes [21].

Naive Bayesian classifier is used explicitly to calculate the probabilities for hypothesis. Naive Bayesian classifier in input data is strong to sound.

10.1002/9781119681984.ch10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

In this case, the Naive Bayesian classifier can be realistic to huge dataset [8]. It will result in maximum precision and promptness. In Naive Bayesian classifier, the result of the parameter on the known class will not depend on the result of another parameter and the above statement is known as class-conditional independence. Naive Bayesian classifier has the ability to work better in any difficulty situation and most commonly in computer aided prediction. Naive Bayesian classifier study from train data, the conditional probability of a single variable I_Z, and the label of class are given by *C*. To calculate the probability by applying the classification through Naive Bayesian of particular instance I_1,....,I_m by the method:

$$P(C = c|I_1 = i_1,\ldots,I_m = i_m) \tag{10.2}$$

As the Naive Bayesian classifier is found on assumption, the variables are conditionally not dependent and subsequent of the class is given by the method

$$P(C = c|I_1 = i_1,\ldots,I_m = i_m) = P(C = c) * \prod_{I_z} P(I_z = i_z|C = c) \tag{10.3}$$

The effect of the Naive Bayesian classifier is discussed and it has the maximum subsequent probability

$$max_c \prod_{I_z} P(I_1 = i_1|C = c) \tag{10.4}$$

Naive Bayesian classifier has four conditions. They are as follows:

1. If class C is a class with his subsequent probability, then the method is

$$max_{c_x \in C} P(c_x|i)$$

2. To identify this class, the subsequent probabilities P(Cx|i) must be calculated. By the use of

$$P(c_x|i) = \frac{P(i|c_x)P(c_x)}{P(i)}$$

10.1002/9781119683964.ch10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119683964.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

This is more sufficient to calculate the probability of the class with given input.

3. The evaluation of probability $P(i|c_x)$ of a parameter to be conditionally not dependent known class

$$P(i|c_x) = \prod_{y=1}^{m} P(a_y|c_x)$$

4. The last step is to evaluate the probabilities for $P(C_x)$ and $P(a_y|c_x)$ by the train set. By giving all the parameters of dataset as continuous, the Gaussian distribution is used to calculate the probability

$$P(a_y|c_x) = \frac{1}{\sqrt{2\pi\sigma_{yx}}} e^{\left(-\frac{(a_y-\mu_{yx})^2}{2\sigma_{yx}^2}\right)}$$

10.5.1 Pseudocode of Naive Bayesian Classifier

Step 1: Read the train set.
Step 2: Estimate the mean and standard deviation of attribute in every class.
Step 3: Repeat
  Calculate the probability using density equation.
  Until all probability is calculated.
Step 4: Calculate the probability of every class.
Step 5: Get the greatest probability.

10.5.2 Advantages of Naive Bayesian Classifier

- It is mainly based on statistical model.
- It produce efficient training algorithm.
- It is easy to understand.
- It is utilized across many domains.
- The order of instance in this has no effect.
- Predicting new data is easy.

10.1002/9781119681984.ch10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

10.6 K-Means Clustering

Given below are the phases that involved in the algorithm of k-means. Figure 10.1 represent the working of K-means clustering.

1. Initialization: K input vectors or data points are chosen for initializing the clusters.
2. Search for nearest neighbors: For every input vector, the cluster center that is the closest is found and that input is assigned to the cluster that is corresponding to it.
3. Updated mean: The centers of cluster in every center are updated by the value of mean or the centroid for the vectors of input that have been assigned to the cluster.
4. Stopping rule: steps 2 and 3 are repeated till such time there is no more change in the means and their values.

There are certain variants of the k-means that differ in the selection of the centroids of initial groups, their choosing of measure of similarity and also the strategies for calculation of cluster means. The usual procedure for the Euclidean data is using the mean as centroid and choosing initial centroids in a random way. This converges to one solution even though it is normally the local minimum [29, 30]. As it is only the vectors that are

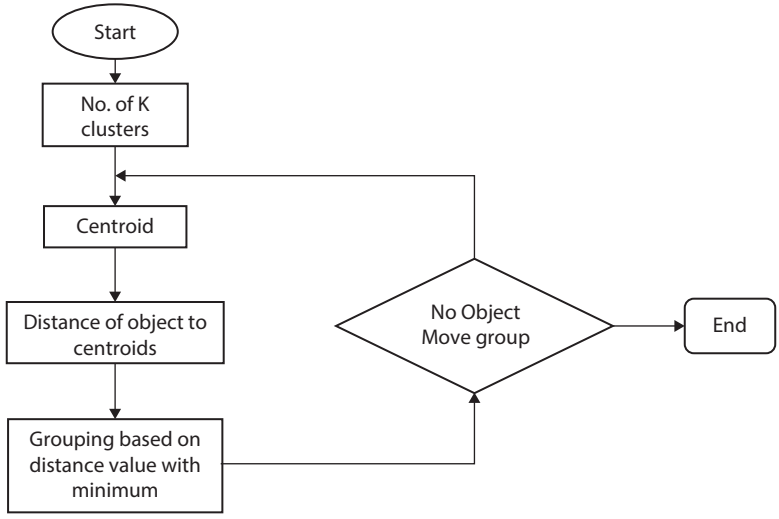


Figure 10.1 K-means centroid computation.

10.1002/9781119681984.ch10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

stored the requirement for space is O(m*n), in which the point number is referred as m and the attributes number is referred as n. The requirement for time is O(I*K*m*n), in which I is referred as the iteration number that is needed for convergence. The I is small and is bounded easily as most of the changes take place only in the initial iterations. So, the K-means is an effective and simple method for clusters that are less than m in number.

10.7 Support Vector Machine (SVM)

SVM is introduced by Vapnikthat is created on theory of Vapnik-Chervonenkis (VC) and principle of structural risk minimization (SRM). SVM discovers the adjustment among the reducing of train set and exhausts the possibilities of the boundary. This is done in for the achievement of the finest simplification capacity [14]. This also remembers hardy to over fitting. The sample for grouping is produced by the training set method along with the trained data. Next grouping is implemented on the model of training. The difficulties come across in position of the SVM prototype that is by choosing the function of kernel and its factor values. Settings of unfitting factor clue to reduce in grouping outcomes. This problem includes count of features and not all features is important for particular task. They may include both redundant and irrelevant, when some features are discarded the optimality of the performance is achieved and also result in decreasing the dimensionality space of input to preserve computation effort but might result in lower accuracy in classification. The smallest count of feature is used to make the classification procedure fast and exact. This procedure will be attained by the utilization of feature selection. This is feature selection schemes are inferred to discover the result of inappropriate features on the enactment of classifier systems [9].

The binary classification task is $\{i_x, j_x\}$, $x = 1 \ldots.n$, $j_x \in \{-1,1\}$, $i_x \in R^d$, in which i_x is point of data and j_x are labels of correspondence. They can be sliced along with the hyper plane denoted by $v^t i + k = 0$ where v is the coefficient of d-dimensional vector in typical vector to the hyper plane, and k is equalizer from origin. The linear type of SVM discovers an ideal splitting boundary by cracking the subsequent ideal activities:

$$\text{Decrease to: } g(v,\xi) = \frac{1}{2}||v||^2 + c\sum_{x=1}^{w} \xi_x$$

$$\text{Subject to: } j_x(v^t i_x + k) \geq 1 - \xi_x \xi_x \geq 0$$

10.1002/9781119681984.ch10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Lagrangian multipliers are introduced $\alpha_m (m = 1,2,\ldots \ldots w)$; the problem of primal can be decreased to a Lagrangian problem with dual property.

Increase to $\sum_x^w \alpha_i - \frac{1}{2}\sum_{x=1}^w \sum_{y=1}^m \alpha_x \alpha_y j_x j_y i_x^t i_y$

Subject to: $0 \le \alpha_x \le C, \sum_{x=1}^w \alpha_x j_{x=0}$

Quadratic optimization problem (QP) by means of linear constraints: The Karush-Kuhn-Tucker (KKT) states that

$$\alpha_x (j_x (v^t i_x + k) - 1) = 0 \tag{10.5}$$

If $\alpha_x > 0$, the points of the corresponding data are called support vector. This result receipts the method is $v = \sum_{x=0}^w \alpha_x y_x i_x$, in which w is number of SVs. k can be attained by $j_x(v^t i_x + k) - 1 = 0$, in which i_x is SVs. When v and k are identified, the functions of linear discriminate are assumed by

$$g(i) = sgn\left(\sum_{x=1}^w \alpha_x j_x i_x^t i + k\right)\Bigg) \tag{10.6}$$

The classes of vectors are not separated linearly, and to make the linear model of machine learning to perform better in nonlinear model, a new idea is implemented. The new idea in real space of input is mapped with the function where the dimensional space of feature is maximum for the trained dataset [10]. The decision with this mapped function can be represented as

$$g(i) = sgn\left(\sum_{x=1}^w \alpha_x j_x \varnothing(i_x)^t \varnothing(i) + k\right)\Bigg) \tag{10.7}$$

where $i_x^t i$ input space is represented in procedure of $\varnothing(i_x)^t \varnothing(i)$ in the space features. The practical procedure for planning $\varnothing(i_x)$ needs not be known because it is subliminally well-defined by the use of designated kernel, which is given by $K(i_x, i_y) = \varnothing(i_x)^t \varnothing(i)$. This function of decision can be conveyed as follows:

$$g(i) = sgn\left(\sum_{x=1}^w \alpha_x j_x K(i_x, i_y) + k\right)\Bigg) \tag{10.8}$$

10.1002/9781119683964.ch10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119683964.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

For positive functions for semi-definite which will fulfill the condition of Mercer's and the function of kernel, where function of kernel is included that can be utilized by SVM. The values of parameters can be selected in advance by

- C is a regularized parameter which controls the transaction among the decreasing the error of training as well as the model with complexity.
- The gamma parameter of the function of the kernel that describes the mapping functions of nonlinear space from the input to the dimensional space feature which is maximum.
- SVM utilizes the function of kernel that designs a hyper plane with decision for nonlinear function space of input.

Advantages of Support Vector Machine

- SVM is utilized when there is no idea about the data.
- SVM unstructured and semi structured data like text, Images and trees.
- The real trick of SVM is kernel strength that is with kernel we can solve complex problem.
- The over-fitting risk is minimum in SVM.
- SVM utilizes programming in technique of convex quadratic that delivers merely minima in global way to circumvent the confined in minima in local way.

10.8 Swarm Intelligence Algorithms

In 1989, Gerardo Beni and Jing Wang acquainted Swarm Intelligence (SI) in connection with cell automated frameworks. Piece of a huge number of animals that have chosen their own will for focus on a typical objective is called swarm. How do swarms (winged animals, fish, and so on) figure out how to move so well altogether. How do ants locate the best wellsprings of nourishment in their surrounding? To respond to these inquiries, new incredible streamlining technique, i.e., SI, is planned.

SI is simply the order that manages advanced collection, self-managed, collaborate, flexible, and incredible behavior of class which observes the straightforward principles. The idea of swarm insight depends on artificial intelligence. Individual can be considered as idiotic yet numerous specialists' display self-association conduct and in this manner can act

10.1002/9781119681984.ch10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

like community-oriented intelligence. SI–based calculation is famous and broadly utilized. A biologic staging arrangement of nature gives the motivation [11]. Case of SI incorporates ant colony, bird congregation, animal or bird grouping, bacterial development, and fish tutoring. Swarm telerobotics is the methodology of SI, alludes to increasingly regular setup of algorithms. "Swarm forecasting" is utilized for determining issues.

10.8.1 Particle Swarm Optimization

PSO is presented by Kennedy and Eberhart in 1995 aspired from the behavior of social creatures in gatherings, for example, flying creature and fish schooling or subterranean ant colonies. This algorithm imitates the communication between individuals to share data. PSO is realistic to different types of fields for development as well as in coordination with other existent calculations.

Swarm optimization (PSO) is a strategy of computational which reduces an issue by regularly and attempting to expand an individual answer based on a specified value of proportion. This understands an issue by the way of having a populace of individual response that is named particles here and particles movement around in the space of searching as per the normal statistical principle from the particle's location and promptness. The development of each particle is attacked by its near most popular location but at the similar period which is guided toward is the most popular situations by the seeking environment that are restored with correct location is sort by particles of different types [12]. Figure 10.2 explains the working of PSO.

PSO will make not maximum or no presumptions about the advanced issue and that can stare over massive spaces of individual solutions. Nonetheless, metaheuristic algorithm like PSO will not ensure an ideal solution at all times. Additionally, PSO will not utilize the changing of the issue that is improved and implies that PSO will not impose that the issue of optimization can be differentiated as it is required by strategies of classical development.

Its applications include combination with a back engendering calculation, to prepare a neural system framework structure, multi-target optimization, classification, image clustering and image clustering, image processing, automated applications, dynamic, pattern recognition, image segmentation, robotic applications, time frequency analysis, decision-making, simulation, and identification [15].

Pseudocode for Particle Swarm Optimization

Step 1: Initialize particle

10.1002/9781119681984.ch10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

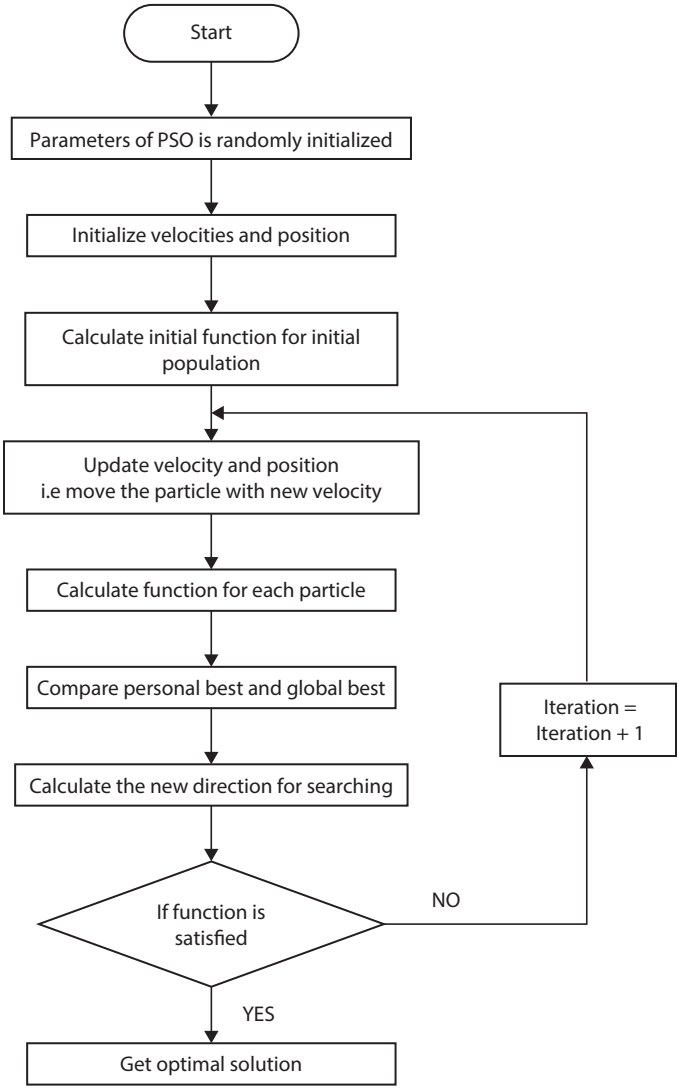


Figure 10.2 Flow chart of particle swarm optimization process.

Step 2: Do till iterations reaches maximum value or error criteria reaches minimum value
Data value for fitness is calculated
Case 1: If value of fitness is better than personalBest then
Set personalBest = currentvalue of fitness
Case 2: If personalBest is superior to globalBest

10.1002/9781119681984.ch10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Set globalBest = personalBest
Step 3: Velocity of particle is calculated and updated

Advantages

- PSO technique is a simple method for execution.
- PSO technique can perform parallel computation.
- PSO has high performance in finding global optima.
- PSO coverage of particle is fast.
- PSO time of computation is short.

10.8.2 Firefly Algorithm

Firefly algorithm is the swarm-based metaheuristic approach which is introduced by Xin. The behavior such as flashing lights of the fireflies is inspired and utilized in the algorithm. The algorithm utilizes the concept that fireflies are always both sex and implies any firefly can be engrossed by some fireflies, and the ability of the desirability of the firefly is directly relational to the ability of its brightness which depends upon the goal work. A firefly will be pulled in to the firefly with more brightness [31, 32].

The working function firefly algorithm has the following steps. Figure 10.3 explains the working of firefly algorithm.

1. Objective function is initialized by absorbing the light intensity.
2. Initial population of the firefly is generated.
3. For every firefly, the light intensity is determined.
4. Attractiveness of the firefly is calculated.
5. The firefly which has brightness level of minimum is moved toward the firefly which as brightness level of maximum.
6. Light intensity of the firefly is updated.
7. Fireflies are ranked based on the intensities and best solution is found.

The advantages of the firefly algorithm are that it has an ability to death with nonlinear more effectively, optimization of multimodal problem can be solved naturally, there is no need of velocity as it is needed in PSO, solution to the global optimization problem can be found as soon as possible, flexible to integrate with other technique of optimization, and initial solution is not required [13]. The disadvantage of firefly algorithm is it consumes more time to reach the optimal solution. Firefly algorithm are used in the field of semantic web composition, classification and clustering problems, neural network, fault detection, digital image compression,

10.1002/9781119681984.ch10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

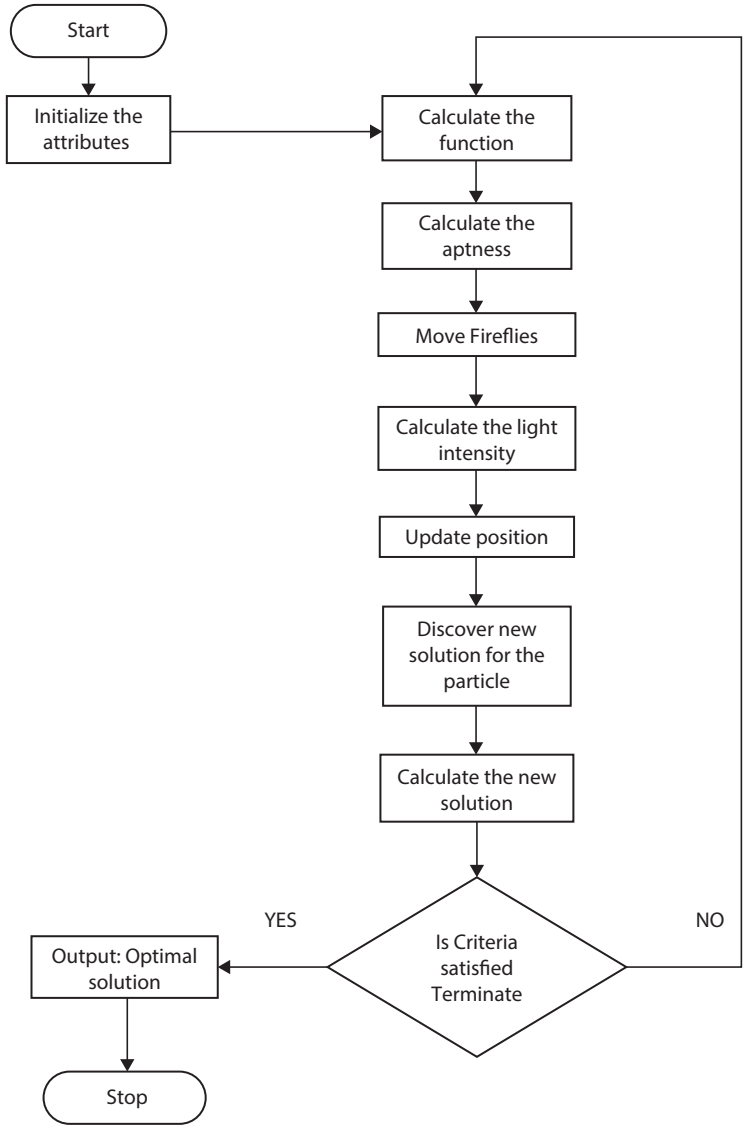


Figure 10.3 Flow chart of firefly algorithm process.

feature selection, digital image processing, scheduling problems, and traveling sales man problem.

Pseudocode for Firefly

Step 1: Randomly generate m fireflies as initial population

10.1002/9781119681984.ch10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

```
Step 2: Calculating appropriateness of each fireflies
    While condition terminates do
    for x = 1: N do (fireflies)
            for y = 1: N do (nflreflies)
                    if(aptness(jx) > aptness(jy))
                            swap(jy.jx)
            differ attractiveness with distance d between fireflies
    arrange the fireflies and global best in ascending order
```

10.8.3 Ant Colony Optimization

ACO is a populace-oriented method of metaheuristic which is utilized for discovering inexact results for troublesome enhancement issues. This method is probabilistic in resolving the problems of issues computational that is diminished with the help of discerning new ways through plans. In ACO, a lot of software transmitter referred as artificial ants will probe for respectable answers for optimal for a given issue of appreciation. For the use ACO, the issue of optimization can be transformed into the issue for identifying the best way on a pattern with weight. The artificial ants gradually built by proceeding onward the pattern.

Artificial ants represent multi-agent techniques roused by the behavior of ordinary ants. The pheromone-based correspondence of natural ants is regularly the overwhelming prototype used. Combinations of artificial ants and neighborhood search algorithms have become a technique for decision for various development jobs including a type of graph, e.g., vehicle steering and web directing. The expanding movement right now prompted conferences devoted exclusively to artificial ants and to various business applications by particular organizations, for example, AntOptima [7].

This algorithm is hidden for an individual from the ant algorithms, but in SI techniques, it comprises some approach of metaheuristic developments. It is make known by Marco Dorigo in 1992; the primary algorithm was in the family way to look for an ideal result in an illustration, supported by the ant's behavior of observing for path between the portion as well as the feed root. The major assumption is that it has improved to explain a maximum class of extensive for issues if numeric, and as a result, little issues have been developed and illustration on various types of the ant's behavior. ACO plays out a model-based searching and offers a few reproductions technique with over assessment of circulation algorithms.

Its application includes the problem with generalized assignment and the set covering, classification problems, Ant Net for organize directing, and multiple knapsack problem.

10.1002/9781119681984.ch10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

10.9 Evaluation Metrics

Accuracy: The accuracy is referred as the exact total of positive and the negatives and separated through the number of classification attributes by total [True Positive (TP) + True Negative (TN) + False Positive (FP) + False Negative (FN)] in the subsequent formula.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{10.9}$$

Sensitivity: Sensitivity is referred as the classification of positive fraction and calculated as below:

$$Sensitivity = \frac{TP}{TP + FN} \tag{10.10}$$

Specificity: Specificity is referred as the classification of negative fraction and calculated as below:

$$Specificity = \frac{TN}{TN + FP} \tag{10.11}$$

F-measure: F-measure is referred by the mean value of harmonic by recall as well as precision. It takes the final call in the adjustment among the recall as well as precision.

$$F - Measure = \frac{2 \times (\mathrm{Pr}\,ecision \times \mathrm{Re}\,call)}{\mathrm{Pr}\,ecision \times \mathrm{Re}\,call} \tag{10.12}$$

In which,

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

10.1002/9781119681984.ch10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$$Misclassification\ rate = 1 - \left[\frac{TP + TN}{TP + TN + FP + FN} \right]$$

10.10 Results and Discussion

Figure 10.4 show the accuracy of various clustering methods performed in Wisconsin dataset. It is observed that, the accuracy of PSO-SVM possesses higher value of 90.41 than PSO– K-means having accuracy of 88.41, while the accuracy of PSO-Bayesian is 86.84. Figures 10.5 and 10.6 show the

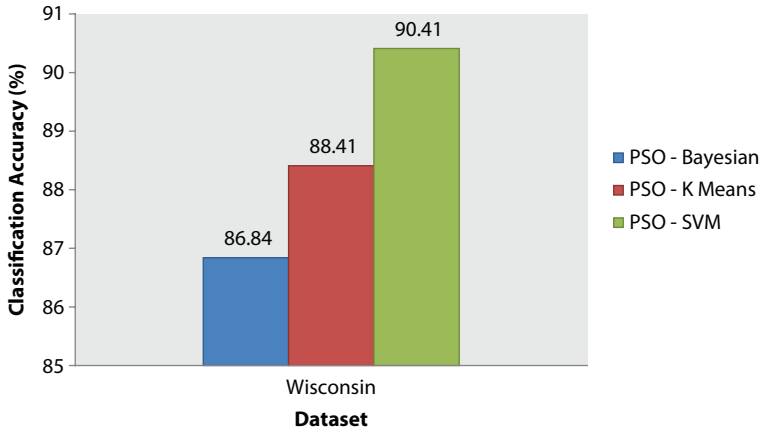


Figure 10.4 Classification accuracy for various clustering methods.

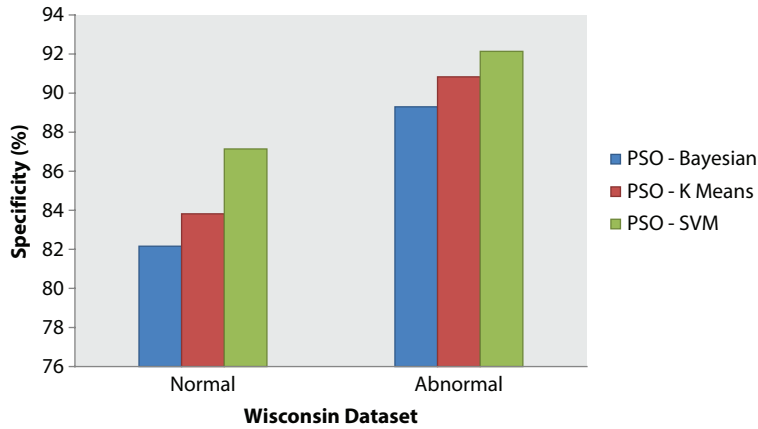


Figure 10.5 Specificity for various clustering methods in Wisconsin dataset.

10.1002/9781119681984.ch10. Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

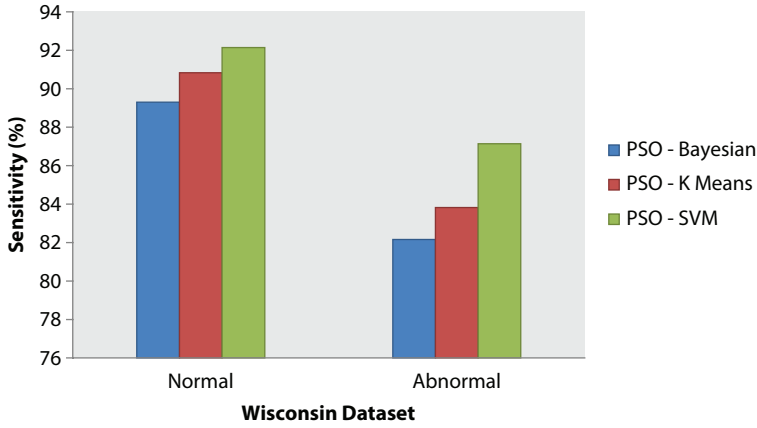


Figure 10.6 Sensitivity for various clustering methods in Wisconsin dataset.

sensitivity of various clustering methods performed in Wisconsin dataset. It is observed that for normal, the sensitivity of PSO-SVM possesses higher value of 0.9214 than PSO–K-means having sensitivity of 0.9083, while the sensitivity of PSO-Bayesian is as lower as 0.893. Also, the PSO-SVM for abnormal has improved sensitivity in Wisconsin dataset having 0.8714 than PSO–K-means having sensitivity of 0.8382, while the sensitivity of PSO-Bayesian is as lower as 0.8216.

Table 10.1 and Figure 10.7 show the F-measure of various methods of clustering performed in dataset of Wisconsin. It is observed that for normal, the F-measure of PSO-SVM possesses a higher value of 0.8625 than 0.833 of PSO–K-means and 0.8115 of PSO-Bayesian. Also, the PSO-SVM

Table 10.1 Performance of various PSO optimized classifiers.

Classifiers	PSO-Bayesian	PSO–K-Means	PSO-SVM
Accuracy (%)	86.84	88.41	90.41
Specificity for Normal (%)	82.16	83.82	87.14
Specificity for Abnormal (%)	89.3	90.83	92.14
Sensitivity for Normal (%)	89.3	90.83	92.14
Sensitivity for Abnormal (%)	82.16	83.82	87.14
F-Measure for Normal (%)	81.15	83.3	86.25
F-Measure for Abnormal (%)	89.89	91.13	92.65

10.1002/9781119681984.ch10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

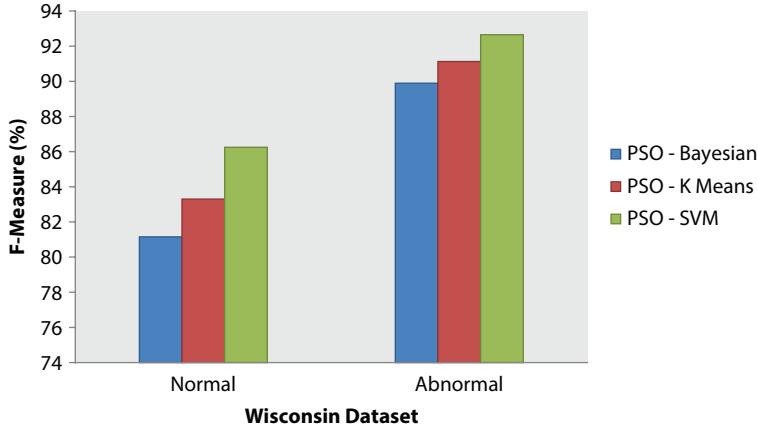


Figure 10.7 F-measure for various clustering methods in Wisconsin dataset.

for abnormal has improved F-measure having 0.9265 than PSO–K-means having F-measure of 0.9113, while the F-measure of PSO-Bayesian is as lower as 0.8989.

10.11 Conclusion

The evaluation metrics plays a key role to assess the quality of a strategy among which the PSO tops in sensitivity, specificity, and precision. For any medical diagnosing model that possess higher sensitivity and precision implies that it enjoys a higher diagnosing power to identify the disease accurately more often. We believed that a very important characteristic of a diagnosing system lies in rejecting the normal images and to identify and pick the cysts. This ideology motivates to carry out the experimental result analysis in all the possible ways considering both to pick the normal as well as abnormal images. In this particular case of breast cancer, if the disease has been diagnosed in the beginning stage, then the percentage of survival of patient will be increased by providing with an appropriate treatment. The obtained result is very promising and claims that PSO-based classifier outperforms the other conventional algorithms.

As a future scope, with respect to the obtained results, it is seen that the accuracy can be further improved to propose an ideal diagnosing model of the disease by employing other NIC algorithms as well. The attempts will be made to implement the most successful proposed model in a physical hospital over a period of time to get trained and the model shall be tested with real-time cases to assess the performance. Also, the metaheuristic approach can be considered to improve further.

10.1002/9781119681984.ch10. Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

References

1. Dhahri, H., Maghayreh, E., Mahmood, A., Elkilani, W., Nagi, M., Automated breast cancer diagnosis based on machine learning algorithms. *J. Healthcare Eng.*, 2019, 4253641, 2019, 10.1155/2019/4253641.
2. Kumar, N., *Application of Machine Learning for Breast Cancer Diagnosis*, Malaviya National Institute of Technology, Jaipur, 2019.
3. Idri, A., Hosni, M., Abnane, I., *Optimizing Parameters of Machine Learning techniques for Breast Cancer Diagnosis*, Software Project Management Research Team, ENSIAS, University Mohammed V Rabat, Morocco.
4. Arafi, A., Fajr, R., Bouroumi, A., Breast cancer data analysis using support vector machines and particle swarm optimization. *2014 Second World Conference on Complex Systems (WCCS)*, Agadir, pp. 1–6, 2014.
5. Purwaningsih, E., Application of the support vector machine and neural network model based on particle swarm optimization for breast cancer prediction. *SinkrOn.*, 4, 66, 2019, 10.33395/sinkron.v4i1.10195.
6. Mazen, F., Abul Seoud, R., Gody, A., Genetic algorithm and firefly algorithm in a hybrid approach for breast cancer diagnosis. *Int. J. Comput. Trends Technol. (IJCTT)*, 32, 62–68, 2016.
7. Fallahzadeh, O., Dehghani-Bidgoli, Z., Assarian, M., Raman spectral feature selection using ant colony optimization for breast cancer diagnosis. *Lasers Med. Sci.*, 33, 1799–1806, 2018, https://doi.org/10.1007/s10103-018-2544-3.
8. Kharya, S., Agrawal, S., Soni, S., Naive Bayes classifiers: A probabilistic detection model for breast cancer. *Int. J. Comput. Appl.*, 92, 10, 0975–8887, April 2014.
9. Furey, T.S., Cristianini, N., Duffy, N., Bednarski, D.W., Schummer, M., Haussler., D., Support vector machine classification and validation of cancer tissue samples using microarray expression data. *Bioinformatics*, Oxford Acad., 16, 10, 906–914, 2000.
10. Chao, C.-F. and Horng, M.-H., The construction of support vector machine classifier using the firefly algorithm. *Comput. Intell. Neurosci.*, Hindawi, 2015, 1–7, 2015.
11. Chen, H.-L., Yang, B., Wang, G., Wang, S.-J., Liu, J., Liu., D.-Y., Support Vector Machine Based Diagnostic System for Breast Cancer Using Swarm Intelligence. *J. Med. Sys.*, 36, 4, 2505–2519, 2011.
12. Lin, S.-W., Ying, K.-C., Chen, S.-C., Lee, Z.-J., Particle swarm optimization for parameter determination and feature selection of support vector machines. *Expert Syst. Appl.*, Elsevier, 35, 4, 1817–1824, 2008.
13. Olatomiwa, L., Mekhilef, S., Shamshirband, S., Mohammadi, K., Petkovic, D., Sudheer, Ch., A support vector machine–firefly algorithm-based model for global solar radiation prediction. *Solar Energy*, Elsevier, 115, 632–644, 2015.

10.1002/9781119681984.ch10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

14. Tuba, E., Mrkela, L., Tuba, M., Support vector machine parameter tuning using firefly algorithm, in: *26th Conference Radioelektronika*, IEEE, 2016.
15. Selvaraj, C., Siva Kumar, R., Karnan, M., A survey on application of bio-inspired algorithms. *IJCSIT*, 5, 1, 366–370 2014.
16. Sousa, T., Neves, A., Silva, A., Swarm optimisation as a new tool for data mining. *Proceedings of the International Parallel and Distributed Processing Symposium (IPDPS'03)*, IEEE Computer Society, 2003.
17. Wang, Z., Sun, X., Zhang, D., Classification rule mining based on particle swarm optimization. *ICIC*, Springer-Verlag Berlin Heidelberg, 2006.
18. Wang, Z., Sun, X., Zhang, D., A PSO-based classification rule mining algorithm. *ICIC*, Springer-Verlag Berlin Heidelberg, 2007.
19. Aalaei, S., Shahraki, H., Rowhanimanesh, A., Eslami, S., Feature selection using genetic algorithm for breast cancer diagnosis: Experiment on three different datasets. *Iran. J. Basic Med. Sci.*, 2016.
20. Zamani, H. and Nadimi-Shahraki, M.-H., Swarm intelligence approach for breast cancer diagnosis. *Int. J. Comput. Appl.*, 151, 1, 40–44, October 2016.
21. Dumitru, D., Prediction of recurrent events in breast cancer using the Naive Bayesian classification. Annals of University of Craiova, 36, 2, 92–96, 2009.
22. Aruna, S., Rajagopalan, S.P., Nandakishore, L.V., Knowledge based analysis of various statistical tools in detecting breast cancer at CS&IT – CCSEA, 2, 37–45, 2011.
23. Soria, D., Garibaldi, J.M., Biganzoli, E., Ellis, I.O., A comparison of three different methods for classification of breast cancer data. *Seventh International Conference on Machine Learning and Applications*, Computer Society, IEEE, 2008.
24. Aci, M., Inan, C., Avci, M., A hybrid classification method of k nearest neighbor, Bayesian methods and genetic algorithm. *Expert Syst. Appl., Elsevier*, 37, 7, 5061–5067, 2010.
25. Shah, C. and Jivani, A.G., Comparison of data mining classification algorithms for breast cancer prediction, in: *4th ICCCNT 2013*, IEEE, 31661, 2013.
26. Al-Aidaroos, K.M., Bakar, A.A., Othman, Z., Medical data classification with Naïve Bayes approach. *Inf. Technol. J.*, Asian Network for Scientific Information, 11, 9, 1166–1174, 2012.
27. Singh, S. and Gupta, P., Comparative study ID3, cart and C4.5 decision tree algorithm: A survey. *Int. J. Adv. Inf. Sci. Technol. (IJAIST)*, 27, 27, 97–103, July 2014.
28. Xiaoliang, Z., Jian, W., Hongcan, Y., Shangzhuo, W., Research and application of the improved algorithm C4.5 on decision tree. *2009 International Conference on Test and Measurement*, IEEE, 2009.
29. Ture, M., Tokatli, F., Kurt, I., Using Kaplan–Meier analysis together with decision tree methods (C&RT, CHAID, QUEST, C4.5 and ID3) in

10.1002/9781119681984.ch10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

determining recurrence-free survival of breast cancer patients. *Expert Syst. Appl. Sci. Direct*, Elsevier, 5, 1, 366–370, 2009.

30. Surya Prasanthi, L. and Kiran Kumar, R., ID3 and its applications in generation of decision trees across various domains- Survey. *Int. J. Comput. Sci. Inf. Technol.*, 6, 6, 5353–5357, 2015.
31. Fister, I., Fister Jr, I., Yang, X.-S., Brest, J., A comprehensive review of firefly algorithms. *Swarm Evol. Comput.*, 13, 34–46, 2006.
32. Atashpaz-Gargari, E. and Lucas, C., Imperialist competitive algorithm: An algorithm for optimization inspired by imperialistic competition, in: *2007 IEEE Congress on Evolutionary Computation*, pp. 4661–4667, 2007.

10.1002/9781119681984.ch10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119681984.ch10 by Tsinghua University Library, Wiley Online Library on [03/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License