International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-05, Issue-04, July 2019

Design of an ensemble classifier using Particle Swarm Optimization Algorithm

*Shampa Sengupta, #Sachin Balo, $Debabrata Chakroborty
*#$Department of Information Technology, MCKV Institute of Engineering, Liluah, Howrah , West Bengal, India.
*shampa2512@yahoo.co.in, #sachin.balo@gmail.com, $deba.c1999@gmail.com

Abstract **Classification is the determining of the class of any real world object which is a major research area in the field of data mining. Accurate classification analysis leads to better understanding of the underlying data structure of a particular object. In present situation, the structure of the data is difficult to understand directly, so many machine-learning approaches have been used for classification of dataset. Major motivation behind combining multiple classifiers is to achieve more classification accuracy compare to a single one. The algorithm has been designed for construction of an optimal ensemble classifier system (ECPSO) using the multiple important feature subsets generated by a rough set theory based feature selection algorithm and Particle swarm Optimization (PSO) algorithm. The method searches a particular combination of classifiers that produces the maximum classification accuracy for finding out the class of a real object. Proposed method compared with some popular existing classification method and experimental results on the datasets taken from uci repository prove the efficiency of the method.**

Keywords —***Data mining, Feature selection, Classification, Ensemble classifier, Rough Set Theory, Particle Swarm Optimization Algorithm***

I. INTRODUCTION

A single classifier may not always give good results as it depends on the training capability of the classifier on the data itself. The solution is to apply multiple classifiers whose combined decision often gives better result compared to a single one. The effectiveness of combined classification system is strongly dependent on proper selection of base classifiers. Various architecture of classifier combination [1, 2-4] is developed to achieve this goal. The techniques that combine multiple classifiers have received much attention, and this is now a standard approach to improving classification performance in machine learning [5]. The algorithm has been designed for construction of an optimal ensemble classifier system (ECPSO) using the multiple feature subsets generated by FSBR algorithm [6] and Particle swarm Optimization (PSO) [7-9]. In the method, suppose *N* number of feature subsets from a decision system is selected after applying FSBR algorithm [6], so *N* number of base classifiers is constructed each from one of *N* number of feature subsets. In the proposed ECPSO method the best combination of classifiers from *N* different base classifiers is determined using Particle swarm Optimization. The method searches a particular combination of classifiers that produces the maximum classification accuracy. The paper discusses a noble method of classification analysis based on the reduced dataset obtained by feature selection methods [6].The structure of the paper as follows: Literature review is given in section 2. Section 3 describes the proposed ensemble classification method and Section 4 shows the experimental result of the proposed method. Finally conclusion of the paper is presented in section 5.

II. LITERATURE RIVIEW

Researchers have already developed lots of ensemble methods including two widely used popular methods such as Bagging [10] and Boosting [11]. Many studies and research proposals discuss the way of developing a multiple classifier system (MCS) [12, 1, 2-4]. Ho (2000) [13] discussed coverage based optimization and decision based optimization techniques for combination of multiple classifier system. Gabrys et.al [14] described a classifier fusion method using Genetic Algorithm where a multidimensional selection of MCS is done. Zhou, Wu, and Tang (2002) [15] demonstrated that ensemble of selected classifiers give better result than all the classifiers used in a MCS. A genetic algorithm based ensemble classifier [16] is proposed for bankruptcy prediction where multi co-linearity problem of classifiers are resolved using the Variance influence factor (VIF).The paper [18] proposed an ensemble approach that attempts to obtain highly accurate classification system. In the paper [10], a bagging

 DOI : 10.35291/2454-9150.2019.0345

ISSN : 2454-9150 Vol-05, Issue-04, July 2019

(bootstrap aggregating) method is introduced. It is seen that the performance of the bagging depends on the stability of the base classifiers. In [11], a boosting method is introduced that produces a series of base classifiers. Here, a set of samples is chosen based on the outcome of prior classifiers in the series. Samples that are incorrectly classified by previous classifiers are giving further chances for classification. Ada-Boost [19] is currently used as a promising boosting technique [11]. Many methods for constructing ensembles of classifiers have been developed, several are universal and some are definite to particular algorithms [15][20][21]. For example, [20] uses 25 classifiers; the paper [21] uses 100 classifiers while it is extended up to 1000 in the paper [22]. To overcome such limitation, the paper [23] proposed a classification algorithm based on several decision tree classifiers using the concept of probability theory and graph theory (EOCDPG), where minimum number of rules is obtained to build an efficient ensemble classifier. The paper [24] integrated a multi-objective GA based feature selection scheme with an ensemble of classifiers (EOCASD) consisting of three basic classifiers: Artificial Neural Network (ANN) [5], Support Vector Machine (SVM) [5], and Decision Tree (DT) [5]. A genetic algorithm based ensemble classifier [17] is also proposed for development of an ensemble classification system for achieving higher classification accuracy with selected optimal base classifiers. But the proposed method provides better result in terms of classification accuracy in comparison with the GA based ensemble classification method [17].

III. Ensemble Classifier Design using Multiple Feature Subsets (ECPSO)

Major motivation behind combining multiple classifiers is to achieve more classification accuracy compare to a single one. The algorithm has been designed for construction of an optimal ensemble classifier system (ECPSO) using the multiple feature subsets generated by FSBR algorithm [6] and Particle swarm Optimization (PSO) algorithm [7]. In the method, suppose N number of feature subsets from a decision system is selected after applying FSBR algorithm [6], so N number of base classifiers is constructed each from one of N number of feature subsets. It is noted that some base classifiers may perform well individually on the training dataset but others may show poor performances. In the proposed ECPSO method, the best combination of classifiers from N different base classifiers is determined using Particle swarm Optimization algorithm. The method searches a particular combination of classifiers that produces the maximum classification accuracy. In the first phase of ECPSO, in the first phase, FSBR algorithm [6] is used to select the important feature subsets from the dataset. Thus, a dataset is considered as a combination of multiple sub-datasets, each corresponding to a feature subset called reduct. Now, from each reduct, rule based classifier is constructed using the concept of association rule mining [25, 26]. In this way, base classifier models, one for each reduct are generated. In the second phase, base classifiers are fused and an optimal ensemble classifier system (ECPSO) is developed using Particle swarm Optimization (PSO) algorithm and performance of the classifier is measured to express its effectiveness. Here, combination of the best performing classifiers performs better compare to a single one, as objects which are not classified by one classifier may be classified by another classifier. For a particular dataset, the ECPSO combines the classifiers with the objectives to maximize the classification accuracy of the ensemble classification system. The overall flow diagram of the ECPSO method is briefly given in Figure 3.1.

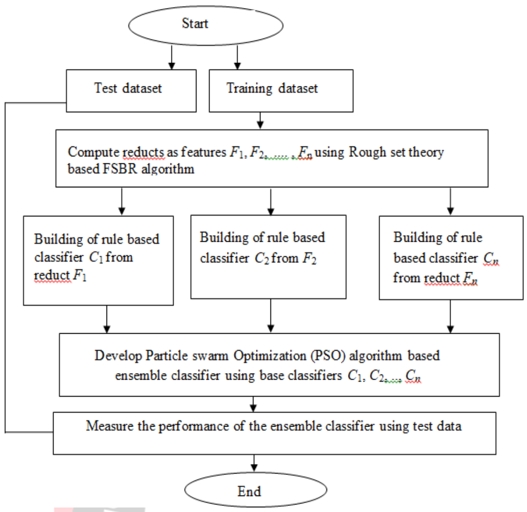

Figure 1: ECPSO workflow diagram

A Construction of Base Classifiers (BC)

This section discusses the construction of rule base classification model BC which is the base classifier of the proposed ECPSO system. Here, the base classifiers are designed based on the set of important feature subsets called reducts selected using FSBR algorithm [6]. The rules of the base classifiers are defined using two interesting measures namely support and confidence, the terms used for association rule mining. The whole process of development of the rule based classifier model BC is divided into two phases such as (i) Feature subsets selection or reducts generation and (ii) Classification rule set Generation.

i. Reducts generation

Feature subset selection or reducts generation using FSBR algorithm [6].

ii. Classification rule set Generation

The method develops rule-based classifiers based on the generated reducts. Initially, many possible rule items are generated based on the core and noncore attributes values for each reduct. The core attribute values are independently

 DOI : 10.35291/2454-9150.2019.0345

ISSN : 2454-9150 Vol-05, Issue-04, July 2019

considered first and set as rule items and if a rule item is not an actual rule in the rule set (which is decided by two interesting measures namely, support and confidence of association rule mining concept), then combined it with a noncore attribute (in reduct) values and a new rule item is formed, which is checked to determine if it is a true rule in the rule set or not. The process is continued until all the attribute values in the reduct are exhausted. Thus the method gives us a rule based classifier with a set of rules for a reduct. The same process is performed for all generated reducts and a set of rule-based classifiers is designed for the decision system. To determine if a rule item is a true rule in the rule set, a weighted value is calculated for each rule item r using the support and confidence measures, concept of association rule mining [25, 26]. The formula of calculating rule weight is given in Equation (1).

Weight of rule r (W_r) $= w * Confidence_r + (1-w) * Support_r$ (1)
Where the value of w is set experimentally.
An association rule r is of the form ($C_1 = p1 \wedge C_2 = p2$) => ($D = d$), where $C_1, C_2, \ldots$, are the conditional attributes, $p1$, $p2$, …, are the values of $C1, C2$, …, respectively, and $D = d$ is a class with label d. So a rule is a mapping from $C \rightarrow D$ i.e. $r: C \rightarrow D$.
Then the support ($Support_r$) and confidence ($Confidence_r$) of an association rule r can be calculated using the Equation 2 and Equation 3.

If C and D are two item sets corresponding to a database T and $r: C \rightarrow D$, an association rule then

$$Support\ (C \rightarrow D) = \frac{\text{tuples containing both } C \text{ and } D}{\text{Total number of tuples in } T} \quad (2)$$

$$Confidence(C \rightarrow D) = \frac{\text{tuples containing both } C \text{ and } D}{\text{Total number of tuples containing } C} \quad (3)$$

If the weighted value of a particular rule item is more than the experimental threshold value then it is selected as a classification rule and stored in the rule set. Otherwise, the rule item is combined with the next noncore attributes values and the same process is repeated to decide if the new rule items are actually the rules of the final rule set or not. In this way, possible rules are obtained in the rule set. Now, to get the more compact rule set, rule pruning is also done to remove irrelevant rule components from the rules without affecting the rule quality. The process is quite simple as iteratively one component at a time is removed and rule quality is recomputed. If the quality of a rule is not decreased after removing any component then the component is permanently removed from the rule. Thus, all unnecessary rule components of the rules are deleted and finally, a more compact rule set is generated. In this way for each reduct of the reduct set a base classifier is constructed. These base classifiers are used to design ensemble classifier ECPSO.

B Ensemble of classifiers

Here, classical particle swarm optimization algorithm (PSO) is used [27] to construct an optimal ensemble classification System (ECPSO). In PSO, a swarm of candidate solutions contains particles and the size of the particle is set as the number of base classifiers (N) to be combined. Each particle in the initial swarm represents the combinations of some base classifiers randomly selected.

As fitness function determines quality of solution (i.e., particle) in the swarm, so a strong global best fitness function is imperative for obtaining good result. For the classifier, the classification accuracy is likely to be the best performance measure, so in the method Combined Classification Accuracy is used to define the fitness function.

A particle is evaluated by its fitness value computed as that is the accuracy of the associated classifier defined in Equation (4) on the training dataset on which the model is learned.

$$\text{Classification accuracy} = \frac{TP+TN}{P+N} \quad (4)$$

Where TP is number of the positive object classified as positive, FP is the number of negative object classified as positive, P is the total number of positive objects and N is the total number of negative object.

IV. EXPERIMENTAL RESULTS

Performance evaluation of the ECPSO algorithm and comparative study with some state- of-the-art classification methods are discussed using real world experimental datasets [27]. Major motivation behind ensemble classifier is to achieve more classification accuracy in comparison with a single one.

A PSO Parameter Setup and preprocessing

The parameters used in ECPSO are given below. These parameters are selected after several test evaluation of the proposed algorithm until reaches to the best configuration in terms of the quality of solutions.

PSO parameter setting:

Input parameters:

- a = 0.14 (weight for individual particle)
- b = 0.16 (weight for individual best particle)
- c = 0.18 (weight for global best particle)
- Population Size: 100

Applied methods:

- Selection type: best
- Replace if better gbest: true

DOI : 10.35291/2454-9150.2019.0345

- Replace if better pbest: true
- Velocity updation: true

Termination criteria:

Search stops in one run when the average fitness of a swarm does not change for 2 consecutive generations.

No of independent run : 30

The ECPSO method uses base classifiers obtained from each reduct using BC method for designing ensemble classifier. In the experiments, '*10-fold cross validation*' is used to evaluate classification performance where in each iteration 80% samples are used for training and 20% other samples are used for test purpose.

B Comparative Study

The method ECPSO is compared with individual base classifier BC from which ECPSO is generated and with some popular ensemble classification methods, Bagging [10], Boosting [11], the classifiers proposed by Das et al. [23], and Zhang [24], where the last two are named here for reference as EOCDPG[23] and EOCASD[24] respectively and the GA based ECS[17] method.

(i) Comparison based on classification accuracy with the single classifier

Here, classification accuracy of each individual base classifier and Ensemble Classifiers are calculated using test data for each experimental benchmark dataset. For different dataset maximum accuracy value achieved by each individual base classifier is compared with the accuracy value achieved by proposed ensemble classifier, which is listed in Table 1. It is seen from Table 1 that the ensemble classification system provides more accuracy than individual base classifiers in all the cases.

Table 1: Comparison of ensemble classifier with individual base classifier

Data set	No of classifier	Classifier Used	Maximum accuracy by individual base classifiers (%)	ECPSO
Wine	4	BC	96.00	96.79
Zoo	10	BC	95.00	96.04
Heart	8	BC	83.15	85.54
Dermatology	10	BC	96.89	98.09
Mushroom	10	BC	96.32	97.98

(*ii) Comparison with popular ensemble classifiers*

The classification accuracies of the ECPSO method and other popular compared ensemble classifier methods are shown in Table 2.

Table 2: Comparison of ECS with other ensemble classifiers

Data set	Bagging	Boosting	EOCDPG	EOCASD	ECS	ECPSO
Wine	**97.09**	95.05	94.74	94.74	96.52	96.79
Zoo	95.48	95.03	94.56	95.01	95.55	**96.04**
Heart	84.52	82.62	83.06	84.09	84.89	**85.54**
Dermatology	98.51	96.06	94.39	**98.59**	97.57	98.09
Mushroom	**98.78**	96.34	96.10	97.40	97.69	97.98

The Table 2 shows that the ECPSO method gives the best results for two datasets whereas EOCASD[24] method gives the best results for dermatology dataset and Bagging method gives the best result for wine and mushroom dataset. This proves that the proposed method is comparable with other existing classification methods and efficient also.

V. CONCLUSION

In recent era of big data, lots of data are being generated in every moment and at the same time data are not very structured. This has inspired the researchers for developing many classification algorithms to analyze the static data and finding out the class of an object. Ensemble classifier construction with optimal base classifiers is the main objective of the paper. As the single classifier system is not always the universal learner for different data mining job, in that case ensemble classifier system improves the performance over single classifier. So the proposed ensemble classifier (ECPSO) is also designed based on the classifiers generated from the important feature subset obtained using single objective particle swarm optimization algorithm. The objective of developing ECPSO is to maximize the classification accuracy, as it is the main target of a classifier. Test results confirm that the proposed classification method is an efficient method. In future, statistical analysis of the method can be done. While designing fitness function, other parameters such as the number of classifiers parameter can also be used to define more powerful fitness function for the development of multiple classifiers system.

REFERENCES

[1] S. Saha, A. Ekbal, U. K. Sikdar, "Named entity recognition and classification in biomedical text using classifier ensemble", Int. J. of Data Mining and Bioinformatics, Vol. 11 (4), pp.365- 391, 2015.

[2] O. Abbaszadeh, A. Amiri, A. R. Khanteymoori, "An ensemble method for data stream classification in the presence of concept drift.", Frontiers of Information Technology & Electronic Engineering, Vol. 16 (2), pp.1059-1068.

[3] S. K. Pati, A. K. Das, "Ensemble Classifier Design Selecting Important Genes", International Journal of Data mining and Bioinformatics, Inder Science, 2017.

 DOI : 10.35291/2454-9150.2019.0345

[4] R. Bryll, R. G. Osuna, F. Quek, "Attribute bagging: improving accuracy of classifier ensembles by using random feature subsets", Pattern Recognition, Vol. 36 (6), pp.1291-1302, 2003.

[5] T. Mitchell, "Machine Learning", New York: McGraw-Hill, 1997.

[6] A. K. Das, S. Chakrabarty, S. Sengupta, "Formation of a Compact Reduct Set Based on Discernibility Relation and Attribute Dependency of Rough Set Theory", Proceedings of the Sixth International Conference on Information Processing, Wireless Network and Computational Intelligence Springer, pp. 253-261, 2012.

[7] B. Xue, M. J. Zhang, W. N. Browne, "Particle swarm optimization for feature selection in classification: a multi-objective approach", IEEE Transactions on Cybernetic, Vol. 6, pp. 1656-1671, 2013

[8] S. Sengupta, A. K. Das, "Optimal rule set generation using pso algorithm" , Fourth International Conference on Artificial Intelligence, Soft Computing and Applications Volume Editors: Dhinaharan Nagamalai, Sundarapandian Vaidyanathan Vol. 4, pp. 301-306, DOI : 10.5121/csit.2014.4530,ISBN-978-1-921987-22-9, 2014

[9] S. Sengupta , A. K. Das, "Particle Swarm Optimization based incremental classifier design for rice disease prediction", Computers and Electronics in Agriculture, Vol. 140, pp. 443–451, 2017

[10] E. Bauer, R. Kohavi, "An empirical comparison of voting classification algorithms: bagging, boosting, and variants", Machine Learning, Vol. 36 (1), pp 105–139, 1999.

[11] X. Sun, H. Zhou, "Experiments with Two new boosting algorithms", Intelligent Information Management, Vol. 2, pp. 386-390, 2010.

[12] L. I. Kuncheva, "Combining Pattern Classifiers, Methods and Algorithms", New York, NY: Wiley Interscience, 2005.

[13] T. K. Ho, "Complexity of classification problems and comparative advantages of combined classifiers", Int.Workshop on Multiple Classifier Systems, lecture Notes on Computer Science, Vol. 1857, pp. 97-106, Springer verlag, 2000B.

[14] Gabrys, D. Ruta, "Genetic algorithms in classifier fusion", Applied Soft Computing, Vol. 6 (4), pp. 337- 47, 2006.

[15] Z-H. Zhou, J-X. Wu, W. Tang, "Ensembling neural networks: Many could be better than all", Artificial Intelligence, Vol. 137 (1–2), pp. 239–263, 2002.

[16] Myoung-Jong. Kim, Dae-Ki. Kang, "Classifiers selection in ensembles using genetic algorithms for bankruptcy prediction", Expert Systems with Application Vol. 39, pp. 9308-9314, 2012.

[17] S. Sengupta , A. K. Das, "An Approach to Development of an Ensemble Classification System", Second IEEE International Conference on Research in Computational intelligence and Communication Networks (ICRCICN), pp. 218-223, 978-1-5090-1047-9/16/$31.00 ©2016 IEEE, 2016

[18] S. Kim, F. Scalzo, D. Telesca, X. Hu, "Ensemble of sparse classifiers for high dimensional biological data", Int. J. of Data Mining and Bioinformatics, Vol. 12 (2), pp. 167-183, 2015.

[19] E. Bauer, R. Kohavi, "An empirical comparison of voting classification algorithms: bagging, boosting, and variants", Machine Learning, Vol. 36 (1), pp 105–139, 1999.

[20] X. Sun, H. Zhou, "Experiments with Two new boosting algorithms", Intelligent Information Management, Vol. 2, pp. 386-390, 2010.

[21] J. R. Quinlan, "Bagging, Boosting and C4.5", Proceedings of the thirteenth national conference on Artificial (AAAI'96), Vol. 1, pp. 725-730, 1996.

[22] E. Bauer, R. Kohavi, "An empirical comparison of voting classification algorithms: bagging, boosting, and variants", Machine Learning, Vol. 36 (1), pp 105-139, 1999.

[23] X. Sun, H. Zhou, "Experiments with Two new boosting algorithms", Intelligent Information Management, Vol. 2, pp. 386-390, 2010.

[24] R. E. Schapire, Y. Freund, "Boosting of margin: A new explanation for the Effective of voting methods", Vol. 26 (5), pp. 1651-1686, 1998.

[25] A. K. Das, J. Sil, "An efficient classifier design integrating rough set and set oriented database operations", Applied Soft Computing, Vol. 11, pp. 2279-2285, 2011.

[26] Z. Zhang, P. Yang, "An ensemble of classifiers with genetic algorithm based feature selection", The IEEE Intelligent Informatics Bulletin, Vol. 9(1), pp. 18-24, 2008.

[27] R. Agrawal, T. lmielinsk, A. Swami, "Mining Association Rules between Sets of Items in Large Databases", Proceedings of the ACM SIGMOD International Conference on the Management of Data, pp. 207 – 216, 1993.

[28] R. Agrawal, R. Srikant, "Fast Algorithms for Mining Association rules", Proc. 20th VLDB conference, Santiago, Chile, 1994.

[29] P. Murphy, W. Aha, "UCI repository of machine learning databases (1996)", http://www.ics.uci.edu/mlearn/MLRepository.html.

 DOI : 10.35291/2454-9150.2019.0345