Artificial Immune-Based Semi-supervised Learning for Classification

Yi Wang
School of Information Engineering, Hubei University of Economics, Wuhan 430205, China
Hubei Internet Finance Information Engineering Technology Research Center, Hubei University of Economics, Wuhan, 430205, China
yiwang@hbue.edu.cn

***Abstract*—Recently, artificial immune algorithms have attracted great attention of researchers and been widely used in function optimization, pattern recognition and classification. Traditional artificial immune algorithms for classification are applied to supervised learning problems, which require completely labeled data for training models. However, in many real-world scenarios, it is difficult to obtain all labeled samples. To solve this problem, an efficient semi-supervised artificial immune algorithm for classification tasks is proposed. It employs a clonal selection algorithm to generate memory cells used for classification, which is achieved via selection, cloning, and mutation procedures. Moreover, it utilizes the ensemble learning technique to extend the co-training paradigm and improves classification performance by adding the most confident unlabeled samples into the labeled set. In addition, the theory of learning from noisy examples is adopted to decide whether there are enough newly labeled samples that are used to reduce the negative effects caused by noises. Experimental results show that the proposed method achieves better or comparable performance than well-known semi-supervised and supervised methods on four datasets.**

***Index Terms*—semi-supervised learning, artificial immune system, clonal selection algorithm, machine learning, classification**

I. Introduction

In recent years, we have witnessed a proliferation of information generated from data communication, web applications, medical diagnosis and industrial manufacturing. How to discovery and extract useful information and knowledge has become the focus of current research. In the past ten years, inspired by biology, many researchers have explored the mechanism of bio-inspired algorithms and proposed various algorithms, such as Genetic Algorithm (GA), Artificial Immune System (AIS), Ant Colony Optimization (ACO) and Particle Swarm Optimization (PSO).

As one of the important solutions, AIS algorithms have been widely used in the data mining and other related fields, especially in classification. For example, artificial immune recognition system (AIRS) [1] based on immune network theory shows an effective performance on classification problems. [2] improved the clonal selection algorithm (CSA) with the local feature selection for classification problems. [3] used an ensemble of AIS-based classification models to detect mammography anomalies.

This work is supported by National Natural Science Foundation of China (Grant No. 62102146).

However, the traditional artificial immune-based algorithms are devised for supervised learning tasks, which require completely labeled data to train. Actually, class labels can only be provided for a subset of the data, which may limit the performance of the algorithms. To address this problem, a semi-supervised algorithm based on artificial immune system using clonal selection is proposed, named Semi-Clonal Selection algorithm (SCSA). It uses a substantial quantity of unlabeled data and a restricted amount of labeled data to improve the performance of the classifier.

In summary, there are several contributions that have been made: firstly, we extend traditional artificial immune algorithms to semi-supervised learning models, which use unlabeled samples to boost the performance of classifiers. Secondly, we adopt the ensemble learning technique to extend the co-training paradigm, and the classifiers are refined by using the most confident unlabeled samples. Finally, we introduce the theory of learning from noisy examples to decide whether there are enough newly labeled samples to ensure performance.

The rest of the paper is organized as follows: Section II reviews some related work. Section III discusses the SCSA algorithm. Experiments are described in Section IV. Finally, Section V draws some conclusions.

II. Related work

The artificial immune system and semi-supervised learning are briefly reviewed in this section.

A. Artificial immune system

The biological immune system defends the body against foreign pathogens using evolutionary learning mechanism, which has complexity, robustness and adaptability. Generally, it consists of two layers of defense mechanism. As the first-line barrier, invading organisms are recognized and eliminated by an unchanging mechanism in the innate immune system. The adaptive immune system generates immunological memory. When repeated infections by the same virus are met, it responds immediately to them. Inspired by the biological immune system, AIS, a novel artificial intelligence technique, consists of four immunological theories: negative selection [4], clonal selection [5], immune networks [6] and danger theory [7]. Many research results show that AIS is an effective solution in optimization, pattern recognition, classification and

2833-8898/23/$31.00 ©2023 IEEE
DOI 10.1109/CyberC58899.2023.00018

medical diagnosis [8], [9]. Here, we focus on reviewing the AIS algorithm for classification problems.

CLONALG [10] is a well-known clonal selection algorithm and provides a solution for machine-learning and pattern-recognition. It consists of initialization, selection, clonal proliferation, mutation, updating and replacement. In recent decades, variations have been developed to improve the classification performance. For example, [11] adopted a new affinity function which maximized the classification accuracy and minimized the misclassification accuracy in CLONALG for classification tasks. Sharma et al. [12] introduced KNN technique to improve the generation of memory cells. [13] combined clonal selection and negative selection to deal with multiclass anomaly detection.

B. Semi-supervised learning

Semi-supervised learning is a special classification method that employs a restricted amount of labeled data and a larger amount of unlabeled data for training. The often-used learning methods include: generative models, graph-based methods, self-training, co-training and transductive support vector machines [14]. In the semi-supervised generative models, Fujino et al. [15] proposed a semi-supervised method for text classification that combined naive Bayes multinomial models with AUC optimization. In [16], a discriminative principle of max-margin learning was employed to enhance performance of deep generative models. The Graph-based semi-supervised models follow the manifold assumption, where nodes are used to represent samples, and edges are used to reflect the similarity between samples. [17] provides systematic reviews of them, including generalized taxonomy, valuable resources and future research directions. [18] proposed a scalable graph-based semi-supervised learning framework, which defined a graph-based sparse prior to create the sparse Bayesian mode and introduced incremental learning technology to deal with large-scale datasets. Lee et al. [19] combined graph construction with cutting label propagation to reduce the influence of noises on the performance.

Self-training and Co-training are other popular semi-supervised algorithms, and have been applied in a variety of fields, such as pattern recognition [20], sentiment analysis [21] and fault diagnosis [22]. In addition, transductive support vector machines [23] have also become the research hotspot, which assign labels for unlabeled samples based on clustering assumptions and use iteratively local search to obtain the optimal solution. Gu et al. [24] extended Semi-Supervised Support Vector Machine (S3VM) [25] by employing a new incremental strategy into large-scale applications. Chen et al. [26] combined transductive support vector machines with an infinitesimal annealing algorithm to decrease the risk of falling into a local optimum. Moreover, they introduced incremental learning to reduce the training time.

III. The proposed Semi-Clonal Selection algorithm

Inspired by [27], [28], we propose a new semi-supervised learning algorithm based on the clonal selection algorithm. It exploits unlabeled samples to boost the performance of classifiers and employs the theory of learning from noisy examples to decide whether there are enough newly labeled samples to ensure performance. In this section, we describe definition, the analysis of the stop condition, and a system framework of the SCSA algorithm.

A. Definition

Definition 1: Antigens

Antigens: Antigens include a set of training samples, denoted by

$$ag = \{ag^{(i)} | ag^{(i)} = (x_i, y_i), x_i \in R^d, y_i \in [\phi, 1, .., C]\} \quad (1)$$

where x_i represents a d-dimensional feature vector, y_i represents its class label, and C is the total number of classes. When $ag^{(i)}$ is an unlabeled sample, y_i is set to ϕ.

Definition 2: Antibodies

Antibodies: Antibodies represent the common features of the training samples for each class.

$$ab = \{ab^{(i)} | ab^{(i)} = (x_i, y_i), x_i \in R^d, y_i \in [1, .., C]\} \quad (2)$$

Definition 3: Ensemble of CSA

H^* denotes an ensemble of CSA-based classifiers with fixed size M. The symbol h_i $(i = 1, 2, ..., M)$ represents the ith component classifier of H^*, and its concomitant ensemble is denoted by H_i, which contains all the component classifiers of H^* except h_i.

B. Analysis of the stop condition

Some studies [29] found that the mitigation of the adverse impact caused by noise can be achieved through the augmentation of the labeled set by incorporating enough newly labeled samples under specific circumstances. Moreover, [30] gives a relation between the worst-case error rate of hypothesis (ϵ), the noise rate (η) and the training sample size (m).

$$m = \frac{c}{\epsilon^2 (1 - 2\eta)^2} \quad (3)$$

The variable c is a constant. If the equation is satisfied, then the learned hypothesis h_i will minimize the disagreement on training examples affected by noise and converge to the true hypothesis h^* with probability one.

By reforming formula (3), the following equation can be obtained.

$$\frac{c}{\epsilon^2} = m(1 - 2\eta)^2 \quad (4)$$

The noise rate in the tth learning iteration is calculated using the following formula.

$$\eta_{i,t} = \frac{\eta_0 W_0 + \hat{e}_{i,t} W_{i,t}}{W_0 + W_{i,t}} \quad (5)$$

where η_0 is the noise rate of initial hypothesis on the original labeled training set L, W_0 is the total predictive confidence of the set L, $\hat{e}_{i,t}$ is the out-of-bag error rate of its concomitant ensemble H_i on the newly labeled set $L_{i,t}$, and $W_{i,t}$ is the total predictive confidence of the set $L_{i,t}$ estimated by H_i.

Then, formula (4) can be written as formula (6) by substituting η and m with $\eta_{i,t}$ and $m_{i,t}$.

$$\frac{c}{\epsilon^2} = (W_0 + W_{i,t})(1 - 2\frac{\eta_0 W_0 + \hat{e}_{i,t} W_{i,t}}{W_0 + W_{i,t}})^2 \qquad (6)$$

where the size of augmented training set $m_{i,t}$ equals to $W_0 + W_{i,t}$.

It can be seen from the formula (6) that the squared worst-case error rate ϵ^2 is inversely proportional to the right hand side of formula. To decrease the error rate during each iteration, the following condition should be satisfied.

$$W_{i,t} > W_{i,t-1} \quad and \quad \hat{e}_{i,t} W_{i,t} < \hat{e}_{i,t-1} W_{i,t-1} \qquad (7)$$

More details are presented in [27].

C. The framework of semi-clonal selection algorithm

In SCSA, the training set consists of the labeled set L and the unlabeled set U. The set L is used to train the initial CSA-based classifiers, and these classifiers are improved by using the newly labeled set $L_{i,t}$ whose class labels are decided by corresponding concomitant ensemble. Here, in order to achieve excellent prediction performance of the ensemble, we use Bagging technique and inherent characteristic of CSA to keep the diversity of component classifiers. Specifically, bootstrap sampling are employed to generate a differently labeled set L_i for each CSA classifier. On the other hand, CSA classifiers have the stochastic characteristic in the antibody generation.

The general framework and the pseudocode are shown in Fig.1 and Algorithm 1, respectively. First, we use the bootstrap sample method to generate the training set L_i and the estimation set E_i (see line 11). Then, each CSA classifier in the ensemble H^* is initiated from the bootstrap sample dataset L_i (see line 12). In the improvement process, we estimate error rate $e_{i,t}$ of the concomitant ensemble H_i on the estimation set E_i. If error rate is reduced in each iteration and condition (7) is met, then the CSA classifier h_i is refined with most confident unlabeled samples selected by H_i. Here, we adopt incremental technology to update the CSA classifier. Each newly sample is sequentially presented to the antibodies in the CSA classifier, instead of rebuilding h_i; When none of CSA classifiers have been updated, we obtain the final training model H^* (see lines 15-36).

D. Several key implementation details in CSA

The primary steps of improved CSA are briefly discussed below.

Initialization: An antibody pool M with fixed size N consists of N_c memory antibody subpool $M_{\{mc\}}$, which is generated by randomly added antigens belonging to the class c. N_c is the number of classes of antigens. The size of $M_{\{mc\}}$

Algorithm 1 SCSA Algorithm

Input: training set T,
confidence threshold θ,
ensemble size N_c
Output: predicted label $h_{fin}(x)$ of x

1: set $L = \phi, U = \phi$
2: **for** each example (x_t, y_t) in the training set T **do**
3: **if** $y_t = \phi$ **then**
4: $U = U \cup (x_t, y_t)$
5: **else**
6: $L = L \cup (x_t, y_t)$
7: **end if**
8: **end for**
9: **for** $i \in \{1, 2, ..., N_c\}$ **do**
10: set $e_{i,0} = 0.5, W_{i,0} = 0$
11: $(L_i, E_i) = BootstrapSample(L, N_c)$
12: $h_i = BuildCSA(L_i)$
13: **end for**
14: $t = 1$
15: **repeat**
16: **for** each CSA classifier h_i **do**
17: $e_{i,t} = MeasureErr(H_i, E_i)$
18: set $L_{i,t} = L$
19: **if** $(e_{i,t} < e_{i,t-1})$ **then**
20: $U_{i,t} = SubSampled(U, \frac{e_{i,t-1} W_{i,t-1}}{e_{i,t}})$
21: **for** each $x_u \in U_{i,t}$ **do**
22: **if** $(GetConfidence(H_i, x_u) > \theta)$ **then**
23: $L_{i,t} = L_{i,t} \cup (x_u, H_i(x_u))$
24: $W_{i,t} = W_{i,t} + GetConfidence(H_i, x_u)$
25: **end if**
26: **end for**
27: **end if**
28: **end for**
29: **for** each CSA classifier h_i **do**
30: **if** $(e_{i,t} W_{i,t} < e_{i,t-1} W_{i,t-1})$ **then**
31: $h_i = UpdateCSA(L_{i,t})$
32: **end if**
33: **end for**
34: $t = t + 1$
35: **until** none of the classifiers in the ensemble changes
36: **return** $h_{fin}(x) = \arg\max_{y \in Y} \sum_{t=1}^{M} I(h_t(x) = y)$

is proportional to the number of the antigens with class c in the training set, which is computed as following.

$$N_{m_c} = \frac{|Ag_c|}{\sum_{c=1}^{N_c} |Ag|} * N \qquad (8)$$

Selection and Exposure: Select an antigen at random, and present it to the antibody pool M_c that has the same class with the antigen. Then calculate an affinity value of each antibody in M_c against the antigen. Here, the affinity function relies on Euclidean distance and remains within the range of [0, 1].

$$affinity = \frac{1}{(1 + dist(ag, ab))} \qquad (9)$$

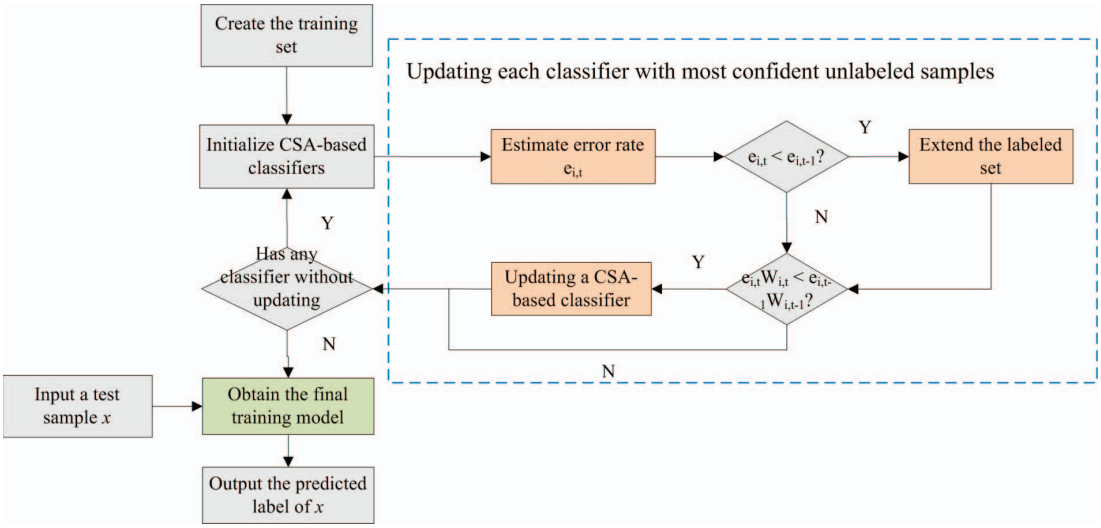

Fig. 1. The general framework of SCSA

Clone and affinity maturation: The n antibodies with the highest affinities in $M_{\{mc\}}$ are selected and cloned proportionally to their affinities. Then, the clones are subjected to an affinity maturation process which adopts Gaussian mutations operator.

Clone Exposure: After exposing the mutated antibodies to the antigen, their affinity values are computed.

Update antibody pool: Select the matured clone with the highest affinity as a candidate. If the affinity value is higher than that of the highest stimulated antibody in $M_{\{mc\}}$, then the candidate will replace it. In addition, the r antibodies with the lowest affinities will be replaced with new antibodies generated at random.

If there is no affinity improvement in t successive iterations or the average affinity value of $M_{\{mc\}}$ is no less than a predefined threshold σ, then the stopping criterion of the above process is reached.

IV. Experimental results

In this section, we aim to provide evidence that the performance of the SCSA classifier can be improved by using the unlabeled examples, and then compare the proposed algorithm with several state-of-the-art semi-supervised and supervised classification methods on benchmark datasets.

A. Experimental setup

The comparison methods include MPSVM [31], KNN-WDM [32], CLONALG [10], CSCA [11], AIRS1 [1], AIRS2 [33], KNN and SVM. The first two methods are both semi-supervised algorithms. MPSVM is based on the transductive support vector machines, a nonparallel proximal classifier was built by using geometric information, and particle swarm optimization was employed to optimize parameters. KNN-WDM is a weighted distance-based KNN classifier, which used the information of labeled training and unlabeled samples to form a metric space. CLONALG, CSCA, AIRS1, AIRS2 are immune-based supervised algorithms. The Gaussian kernel is used in the SVM classifier. The kernel width of SVM and the parameter k in KNN are tuned by 10-fold cross validation strategy.

The SCSA is developed in the python software, the parameter values of the ensemble size, antibody population size, selection pool size, replacement size in the antibody pool, and maximum number of successive iterations without affinity improvement T is set to 10, 150, 10, 10% and 10, respectively. Other comparison methods use suggested parameters in the original literatures.

In addition, four well-known datasets are employed, of which the detail is shown is Table I. RandomRBF is a synthetic dataset generated by the MOA tool [34]. All the attribute values of samples are normalized to the interval [0, 1]. Because MPSVM is a binary classifier, the samples for target classes 2 with target classes 3 are merged in the Wine dataset. The experiments are conducted using 10-fold cross validation, where a dataset is randomly divided into ten folds. Nine folds are used as training data, and the rest one is used as test data. For the semi-supervised methods, the training data T includes a labeled set L and an unlabeled set U, where U is equal to the unlabeled rate μ multiplied by T. For the supervised methods, all training samples are given class labels. In addition, the experiments are repeated 10 times for each dataset.

TABLE I
THE DETAIL OF SYNTHETIC DATASETS

dataset	instances	features	classes
Sonar	208	60	2
Ionosphere	351	34	2
Wine	178	13	3
RandomRBF	200	10	2

B. The performance affected by the unlabeled rate

Figure 1 shows the complete trend of the average accuracy rate affected by the unlabeled rate μ varying from 0.1 to 0.9 on the RandomRBF dataset. The original SCSA is only learned from the labeled set L in the set T, while SCSA uses both the labeled set L and the unlabeled set U. From the figure, we observe that the curve of the SCSA is obviously higher than that of the original SCSA in most cases. The curve of the SCSA almost coincides with that of the original SCSA when the unlabeled rate decreases from 0.9 to 0.8. It shows that the accuracy of the SCSA can be increased by using the unlabeled samples except for the high unlabeled rate. The reason is that when the labeled set L is too small, there is not enough labeled samples to build reliable classifiers. These classifiers are more likely to assign the wrong labels to unlabeled samples, thereby degrading the performance of the refined classifiers.

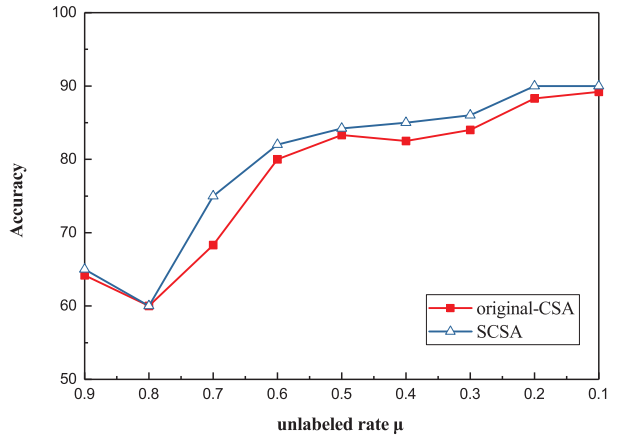

Fig. 2. The accuracy of SCSA and original SCSA over different unlabeled rates μ in the RandomRBF dataset

C. Comparison with semi-supervised methods

The average accuracy rates of the semi-supervised algorithms with different unlabeled rates are shown in Table II. We can see from the tables that the average performance of the compared algorithms is improved as the ratio of labeled data increases. For instances, the accuracy of SCSA is improved more than 5% when the unlabeled rate decreases from 0.4 to 0.2 on the RandomRBF dataset. Moreover, SCSA achieves remarkable or comparable performance to MPSVM and KNN-WDM on most databases. For example, when 40% samples are unlabeled, SCSA outperforms KNN-WDM on three out of four datasets. When 20% samples are unlabeled, SCSA has higher accuracy than KNN-WDM and reaches higher or comparable accuracy to MPSVM on all the datasets. This is because that SCSA adopts the ensemble learning technique to keep the diversity and labels samples by using the concomitant ensemble to avoid bias and overfit.

D. Comparison with supervised methods

Figure 3 and Figure 4 show the accuracy of the SCSA compared with popular supervised algorithms. We observe

TABLE II
THE AVERAGE ACCURACY RATES OF SEMI-SUPERVISED ALGORITHMS WITH DIFFERENT UNLABELED RATES μ

dataset	μ	MPSVM (%)	KNN-WDM (%)	SCSA (%)
Sonar	0.8	74.08	63.69	67.14
	0.6	79.35	72.39	79.05
	0.4	82.68	82.68	82.54
	0.2	82.23	83.19	83.33
Ionosphere	0.8	86.06	87.19	86.52
	0.6	88.19	89.30	88.90
	0.4	93.58	89.44	91.87
	0.2	93.87	91.73	92.56
Wine	0.8	94.92	88.65	97.22
	0.6	94.10	92.12	95.37
	0.4	95.77	91.83	97.22
	0.2	93.55	91.00	98.15
RandomRBF	0.8	81.50	69.67	60.00
	0.6	82.00	80.33	82.00
	0.4	84.00	80.92	85.00
	0.2	84.50	80.42	90.00

that the performance of SCSA with certain unlabeled rates is better than some of the supervised algorithms using the whole training labeled set T. For instances, when 40% samples are unlabeled, the average accuracy of SCSA is higher than that of CLONALG and KNN on the Ionosphere dataset. When 20% samples are unlabeled, SCSA outperforms CSCA on both Ionosphere and Sonar datasets, and reaches comparable performance with AIRS1 and AIRS2 on the Sonar dataset.

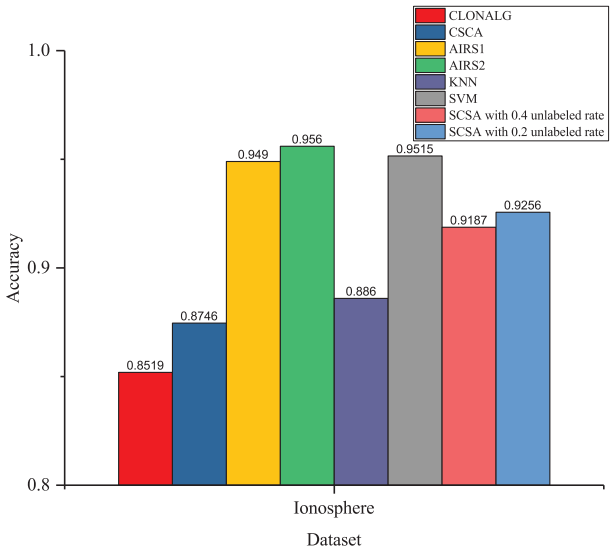

Fig. 3. The accuracy rates in the Ionosphere dataset

V. Conclusions

Most artificial immune-based algorithms are devised for supervised learning tasks, which require completely labeled data to train. To address this issue, we propose a semi-supervised algorithm based on CSA, named Semi-Clonal Selection algorithm (SCSA). Unlike most conventional artificial immune

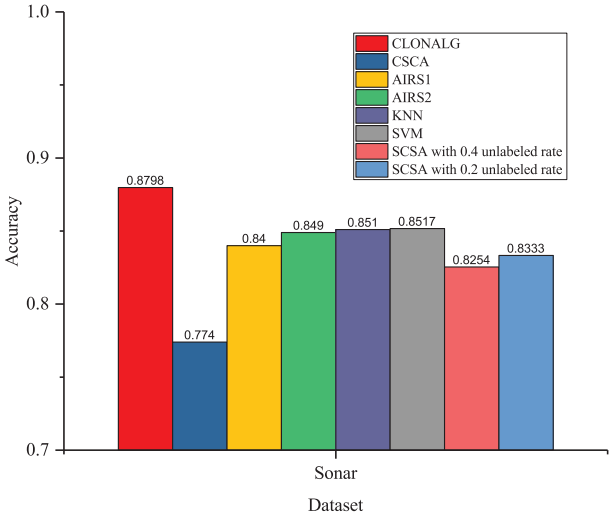

Fig. 4. The accuracy rates in the Sonar dataset

algorithms, it adopts the ensemble learning technique to further develop the co-training method and estimate the confidence of unlabeled data. Then, the classifiers are refined by using the newly labeled samples. Finally, we introduce the theory of learning from noisy examples to decide whether there are enough newly labeled samples to ensure performance. Experiments indicate that unlabeled samples can substantially improve the performance of SCSA. Meanwhile, compared with the state-of-the-art semi-supervised and supervised classification algorithms, it achieves remarkable or comparable performance in most datasets. In the next phase of work, we will investigate the influence of the varying levels of noise on the algorithm's robustness and use incremental learning to extend our method into time-varying data streams. We also explore how to apply our proposed framework into anomaly detection over data streams.

REFERENCES

[1] A. Watkins and L. Boggess, "A new classifier based on resource limited artificial immune systems," in *Proc. Congr. Evol. Comput*. IEEE, 2002, pp. 1546–1551.

[2] G. Dudek, "An artificial immune system for classification with local feature selection," *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 6, pp. 847–860, 2012.

[3] G. Magna, P. Casti, S. V. Jayaraman, M. Salmeri, A. Mencattini, E. Martinelli, and C. Di Natale, "Identification of mammography anomalies for breast cancer detection by an ensemble of classification models based on artificial immune system," *Knowledge-Based Systems*, vol. 101, pp. 60–70, 2016.

[4] K. D. Gupta and D. Dasgupta, "Negative selection algorithm research and applications in the last decade: A review," *IEEE Transactions on Artificial Intelligence*, vol. 3, no. 2, pp. 110–128, 2021.

[5] C. Chen, Y. Wan, A. Ma, L. Zhang, and Y. Zhong, "A decomposition-based multiobjective clonal selection algorithm for hyperspectral image feature selection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–16, 2022.

[6] B. Mumtaz, S. Kanwal, S. Alamri, and F. Khan, "Feature selection using artificial immune network: An approach for software defect prediction." *Intelligent Automation & Soft Computing*, vol. 29, no. 3, 2021.

[7] L. E. Jim, N. Islam, and M. A. Gregory, "Enhanced manet security using artificial immune system based danger theory to detect selfish nodes," *Computers & Security*, vol. 113, p. 102538, 2022.

[8] Y. Wang, T. Li, X. Liu, and J. Yao, "An adaptive clonal selection algorithm with multiple differential evolution strategies," *Information Sciences*, vol. 604, pp. 142–169, 2022.

[9] X. Sun, H. Wang, S. Liu, D. Li, and H. Xiao, "Self-updating continual learning classification method based on artificial immune system," *Applied Intelligence*, vol. 52, no. 11, pp. 12 817–12 843, 2022.

[10] L. N. De Castro and F. J. Von Zuben, "Learning and optimization using the clonal selection principle," *IEEE transactions on evolutionary computation*, vol. 6, no. 3, pp. 239–251, 2002.

[11] J. Brownlee, "Clonal selection theory & clonalg - the clonal selection classification algorithm (csca)," *Swinburne University of Technology*, 2005.

[12] A. Sharma and D. Sharma, "Clonal selection algorithm for classification," in *International Conference on Artificial Immune Systems*. Springer, 2011, pp. 361–370.

[13] Y. J. Kim, W. Nam, and J. Lee, "Multiclass anomaly detection for unsupervised and semi-supervised data based on a combination of negative selection and clonal selection algorithms," *Applied Soft Computing*, vol. 122, p. 108838, 2022.

[14] X. Yang, Z. Song, I. King, and Z. Xu, "A survey on deep semi-supervised learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 9, pp. 8934–8954, 2023.

[15] A. Fujino and N. Ueda, "A semi-supervised auc optimization method with generative models," in *Data Mining (ICDM), 2016 IEEE 16th International Conference on*. IEEE, 2016, pp. 883–888.

[16] C. Li, J. Zhu, and B. Zhang, "Max-margin deep generative models for (semi-) supervised learning," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 11, pp. 2762–2775, 2018.

[17] Z. Song, X. Yang, Z. Xu, and I. King, "Graph-based semi-supervised learning: A comprehensive review," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–21, 2022, doi:10.1109/TNNLS.2022.3155478.

[18] B. Jiang, H. Chen, B. Yuan, and X. Yao, "Scalable graph-based semi-supervised learning through sparse bayesian model," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 12, pp. 2758–2771, 2017.

[19] J. Lee, Y. Kim, and S. B. Kim, "Noise-robust graph-based semi-supervised learning with dynamic shaving label propagation," *Applied Soft Computing*, vol. 142, p. 110371, 2023.

[20] K. Wang, Y. Chen, Y. Zhang, X. Yang, and C. Hu, "Iterative self-training based domain adaptation for cross-user semg gesture recognition," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 2974–2987, 2023, doi:10.1109/TNSRE.2023.3293334.

[21] K. He, R. Mao, T. Gong, C. Li, and E. Cambria, "Meta-based self-training and re-weighting for aspect-based sentiment analysis," *IEEE Transactions on Affective Computing*, pp. 1–13, 2022, doi:10.1109/TAFFC.2022.3202831.

[22] T. S. Abdelgayed, W. G. Morsi, and T. S. Sidhu, "Fault detection and classification based on co-training of semisupervised machine learning," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 2, pp. 1595–1605, 2017.

[23] T. Joachims, "Transductive inference for text classification using support vector machines," in *ICML*, vol. 99, 1999, pp. 200–209.

[24] B. Gu, X.-T. Yuan, S. Chen, and H. Huang, "New incremental learning algorithm for semi-supervised support vector machine," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM, 2018, pp. 1475–1484.

[25] O. Chapelle, V. Sindhwani, and S. S. Keerthi, "Optimization techniques for semi-supervised support vector machines," *Journal of Machine Learning Research*, vol. 9, no. Feb, pp. 203–233, 2008.

[26] H. Chen, Y. Yu, Y. Jia, and B. Gu, "Incremental learning for transductive support vector machine," *Pattern Recognition*, vol. 133, p. 108982, 2023.

[27] M. Li and Z.-H. Zhou, "Improve computer-aided diagnosis with machine learning techniques using undiagnosed samples," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 37, no. 6, pp. 1088–1098, 2007.

[28] Y. Wang and T. Li, "Improving semi-supervised co-forest algorithm in evolving data streams," *Applied Intelligence*, pp. 1–15, 2018.

[29] S. Goldman and Y. Zhou, "Enhancing supervised learning with unlabeled data," in *ICML*, 2000, pp. 327–334.

[30] D. Angluin and P. Laird, "Learning from noisy examples," *Machine Learning*, vol. 2, no. 4, pp. 343–370, 1988.
[31] W.-J. Chen, Y.-H. Shao, D.-K. Xu, and Y.-F. Fu, "Manifold proximal support vector machine for semi-supervised classification," *Applied intelligence*, vol. 40, no. 4, pp. 623–638, 2014.
[32] H. Mohebbi, Y. Mu, and W. Ding, "Learning weighted distance metric from group level information and its parallel implementation," *Applied Intelligence*, vol. 46, no. 1, pp. 180–196, 2017.
[33] A. Watkins and J. Timmis, "Artificial immune recognition system (airs): Revisions and refinements," in *1st International Conference on Artificial Immune Systems (ICARIS2002)*, vol. 5. University of Kent at Canterbury Printing Unit, 2002, pp. 173–181.
[34] A. Bifet, G. Holmes, R. Kirkby, and B. Pfahringer, "Moa: Massive online analysis," *Journal of Machine Learning Research*, vol. 11, no. May, pp. 1601–1604, 2010.