A NEAT-based Multiclass Classification Method with Class Binarization

Zhenyu Gao
Vrije Universiteit Amsterdam
Amsterdam, the Netherlands
vu.gaozhy@gmail.com

Gongjin Lan*
Southern University of Science and Technology
Shenzhen, China
langj@sustech.edu.cn

ABSTRACT

Multiclass classification is a fundamental and challenging task in machine learning. Class binarization is a popular method to achieve multiclass classification by converting multiclass classification to multiple binary classifications. NeuroEvolution, such as NeuroEvolution of Augmenting Topologies (NEAT), is broadly used to generate Artificial Neural Networks by applying evolutionary algorithms. In this paper, we propose a new method, ECOC-NEAT, which applies Error-Correcting Output Codes (ECOC) to improve the multiclass classification of NEAT. The experimental results illustrate that ECOC-NEAT with a considerable number of binary classifiers is highly likely to perform well. ECOC-NEAT also shows significant benefits in a flexible number of binary classifiers and strong robustness against errors.

CCS CONCEPTS

• **Computing methodologies** → **Ensemble methods**; **Genetic algorithms**; • **Computer systems organization** → *Neural networks*;

KEYWORDS

Class binarization, Multiclass classification, Binary classification, Error Correcting Output Codes, NEAT, One-vs-One, One-vs-All.

ACM Reference Format:
Zhenyu Gao and Gongjin Lan*. 2021. A NEAT-based Multiclass Classification Method with Class Binarization. In *2021 Genetic and Evolutionary Computation Conference Companion (GECCO '21 Companion), July 10–14, 2021, Lille, France.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3449726.3459509

1 INTRODUCTION

Multiclass classification is a crucial branch of machine learning (ML) with many applications. One way to solve multiclass classification is the *class binarization techniques*, which convert multiclass classification to multiple binary classifications.

Several classifiers like SVM have been studied well under the framework of class binarization. However, studies on other base classifiers like NeuralEvolution (NE) remain an open issue. NeuroEvolution of Augmenting Topologies (NEAT), which is a popular NE algorithm [6], is capable of generating Artificial Neural Networks (ANNs) for multiclass classification directly (also referred to as standard-NEAT). However, its performance degrades rapidly as the number of classes increases [4]. We consider that one popular solution is to take advantage of class binarization techniques.

There are three well-known class binarization approaches: One-vs-One (OvO), One-vs-All (OvA), and Error-Correcting Output Codes (ECOC) [1]. In particular, OvO and OvA have been applied to the multiclass classification of NEAT [4]. Compared to OvO and OvA, ECOC has unique advantages [2]. This paper presents a novel method that combines ECOC with NEAT for multiclass classification.

2 OUR METHOD: ECOC-NEAT

In NEAT-based multiclass classification methods with class binarization, NEAT works as base classifiers (the NEAT algorithm for binary classification is referred to as *B-NEAT* by us). After employing classifiers as B-NEAT in ECOC, we have the new method, ECOC-NEAT.

In training stage, we first design a random code matrix with the size of l (l also means the number of binary classifiers. For k-class classification, there is $\lceil log_2 k \rceil \leq l \leq 2^{k-1} - 1, l \in \mathbb{Z}$). We can see that compared to a fixed number of classifiers in OvO or OvA, the classifier number in ECOC is flexible. In some cases of ECOC, only $O(\log_2 k)$ classifiers are required, which is much fewer than OvO (with $O(k^2)$ classifiers) and OvA (with $O(k)$ classifiers). Then we train l classifiers using B-NEAT to constitute the set of classifiers. In the decoding stage, we determine the final prediction based on the hamming distance. The ensemble decision is to assign a new sample to the label whose codeword is nearest to the output code.

3 EXPERIMENTAL SETUP

We perform experiments on three datasets, such as the Digit dataset from the ski-learn package [5], Satellite and Ecoli. datasets from the ML Repository of the University of California, Irvine [3].

A fixed hyper-parameter configuration is used throughout all experiments. For different datasets, we only adjust the number of attributes. The number of outputs is always equal to 2 for B-NEAT. In standard NEAT, the number of outputs equals classes number. The *Fitness* of each genome (i.e., ANN) is *Accuracy*, which is a straightforward metric for training. To make the comparison as fair as possible, we set the same number of generations ($G = 3000$) for each experiment. If there are l classifiers, each classifier is allowed to evolve g generations ($g = G/l$). All approaches require the implementation of NEAT. An open-source Python implementation of NEAT, NEAT-Python [1], is used in all experiments.

[1] https://github.com/CodeReclaimers/neat-python

GECCO '21 Companion, July 10–14, 2021, Lille, France

ACM ISBN 978-1-4503-8351-6/21/07.
https://doi.org/10.1145/3449726.3459509

To make persuasive conclusions, we repeat each experiment ten times (for Ecoli., we repeat 10-fold cross-validation ten times), and we average the metrics. Wilcoxon paired signed-rank test is used to test whether there is a significant difference between pairwise methods. The significant level $\alpha = 0.10$ for all the comparisons.

4 RESULTS

Besides NEAT, OvO-NEAT, and OvA-NEAT, we test ECOC-NEAT with different code sizes including the lower bound ($l = \lceil log_2 k \rceil$) and the upper bound ($l = 2^{k-1} - 1$). Table 1 displays the performance of different methods on the three datasets.

Table 1: Performance Results. l-bit ECOC means the size of code is l. Bold values with gray background means the maximal values. Underline means the result is significantly better than standard NEAT, double underline means it is better than OvO-NEAT.

Dataset Method	#Classifiers	Test Accuracy	Variance
Digit (10 classes)			
Standard NEAT	1	0.4489	9.56×10^{-4}
OvO-NEAT	45	0.8664	4.94×10^{-4}
OvA-NEAT	10	0.7400	10.46×10^{-4}
4-bit ECOC	4	0.5350	72.01×10^{-4}
10-bit ECOC	10	0.6506	15.78×10^{-4}
45-bit ECOC	45	0.8189	9.04×10^{-4}
511-bit ECOC	511	**0.8989**	**0.95**$\times10^{-4}$
Satellite (6 classes)			
Standard NEAT	1	0.6377	19.05×10^{-4}
OvO-NEAT	15	0.7790	1.95×10^{-4}
OvA-NEAT	6	0.6832	32.19×10^{-4}
3-bit ECOC	3	0.7039	9.79×10^{-4}
6-bit ECOC	6	0.7140	21.49×10^{-4}
15-bit ECOC	15	0.7745	3.61×10^{-4}
31-bit ECOC	31	**0.7903**	**0.46**$\times10^{-4}$
Ecoli. (8 classes)			
Standard NEAT	1	0.7540	0.99×10^{-4}
OvO-NEAT	28	0.8421	0.79×10^{-4}
OvA-NEAT	8	0.7868	1.53×10^{-4}
3-bit ECOC	3	0.7649	2.28×10^{-4}
8-bit ECOC	8	0.7896	3.24×10^{-4}
28-bit ECOC	28	**0.8489**	**0.79**$\times10^{-4}$
Exhaustive ECOC	127	0.8365	0.88×10^{-4}

The above results illustrate that class binarization techniques can incredibly improve the performance of multiclass classification compared to the standard-NEAT [4]). Then, the size of ECOC has a significant influence on multiclass performance. A larger ECOC (i.e., with more binary classifiers) usually results in better performance because more extended codes have more bits to correct errors.

Compared to OvO, one of ECOC methods' benefits is its robustness against errors is significantly stronger. Figure 1 leads us to conclude that if several classifiers in OvO-NEAT are missing, the test accuracy declines almost linearly. However, the accuracy of ECOC-NEAT methods decreases slightly.

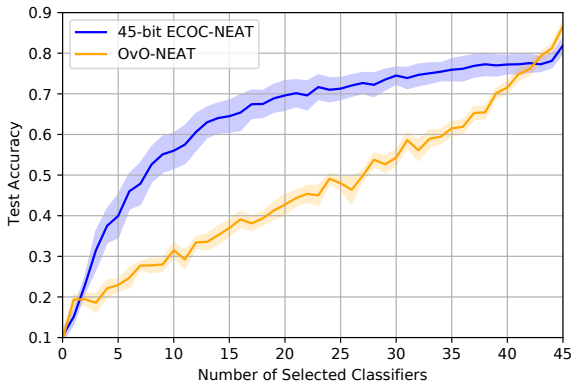

Figure 1: Test Accuracy vs. Number of Selected Classifiers on the Digit Dataset. Both methods contain 45 binary classifiers. Translucent bands indicate 95% confidence intervals.

This finding illustrates that if the classifiers are easily lost or contaminated in some scenarios, ECOC-NEAT has a more acceptable performance than OvO-NEAT.

It is unsurprisingly that ECOC is more robust against errors than OvO. Each base classifier in OvO provides independent and equally important information. If several classifiers are missing, the corresponding information is lost; nevertheless, for ECOC, the information between different classifiers can complement each other well. When several classifiers are missing, other classifiers can compensate for the lack of information to some extent.

5 CONCLUSIONS

We propose the ECOC-NEAT method that applies ECOC to the multiclass classification of NEAT. Our experimental results show that ECOC-NEAT with a considerable size performs consistently well; ECOC-NEAT also offers significant advantages in a flexible classifier number and strong robustness.

REFERENCES

[1] Erin L Allwein, Robert E Schapire, and Yoram Singer. 2000. Reducing multiclass to binary: A unifying approach for margin classifiers. *Journal of machine learning research* 1, Dec (2000), 113–141.
[2] Thomas G Dietterich and Ghulum Bakiri. 1994. Solving multiclass learning problems via error-correcting output codes. *Journal of artificial intelligence research* 2 (1994), 263–286.
[3] Dheeru Dua and Casey Graff. 2017. UCI Machine Learning Repository. (2017). http://archive.ics.uci.edu/ml
[4] Tyler McDonnell, Sari Andoni, Elmira Bonab, Sheila Cheng, Jun-Hwan Choi, Jimmie Goode, Keith Moore, Gavin Sellers, and Jacob Schrum. 2018. Divide and conquer: neuroevolution for multiclass classification. In *Proceedings of the Genetic and Evolutionary Computation Conference*. 474–481.
[5] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12 (2011), 2825–2830.
[6] Kenneth O Stanley and Risto Miikkulainen. 2002. Evolving neural networks through augmenting topologies. *Evolutionary computation* 10, 2 (2002), 99–127.