Liquid Democracy for Low-Cost Ensemble Pruning

Extended Abstract

Ben Armstrong
University of Waterloo
Waterloo, Canada
ben.armstrong@uwaterloo.ca

Kate Larson
University of Waterloo
Waterloo, Canada
kate.larson@uwaterloo.ca

ABSTRACT

We show that there is a strong connection between ensemble learning and a delegative voting paradigm, *liquid democracy*, which can be leveraged to reduce ensemble training costs. We present an incremental training procedure that removes redundant classifiers from an ensemble via delegation. By carefully selecting the underlying delegation mechanism weight-centralization among classifiers is avoided, leading to higher accuracy than some boosting methods with a significantly lower cost than training a full ensemble. This work serves as an exemplar of how ideas from computational social choice can be applied to problems in nontraditional domains.

KEYWORDS

Liquid Democracy; Machine Learning; Ensembles; Pruning

ACM Reference Format:
Ben Armstrong and Kate Larson. 2024. Liquid Democracy for Low-Cost Ensemble Pruning: Extended Abstract. In *Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), Auckland, New Zealand, May 6 – 10, 2024*, IFAAMAS, 3 pages.

1 INTRODUCTION

Machine learning consumes increasingly large amounts of data and compute while performance often only improves incrementally [8]. This incentivizes mass data collection, increases model training time, and has significant environmental costs while restricting access to the most powerful ML models to groups large enough to support significant infrastructure. In this paper, we propose adapting an existing paradigm of delegative opinion aggregation to reduce compute requirements during classifier ensemble training.

Our model bridges two fields of research – machine learning and social choice. We introduce a new algorithm for incrementally pruning and re-weighting an ensemble during the training process to reduce computational costs of training while maintaining or improving test accuracy. This algorithm adds a new dimension to the long-standing link between ensemble learning and voting [5, 6] by showing similarities between pruning and delegation.

Historically, ensembles have been compared with groups of voters [10], allowing a theoretical link with early results in social choice which state that large groups of voters are more accurate than small groups [7]. Similarly, larger ensembles are frequently more accurate than smaller ensembles [11].

Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), N. Alechina, V. Dignum, M. Dastani, J.S. Sichman (eds.), May 6 – 10, 2024, Auckland, New Zealand.

Both ensemble learning and social choice have developed more complex methodologies which allow for smaller groups of classifiers or voters to outperform larger groups. Ensembles have benefited from both boosting and pruning methods [3] while social choice can be made more accurate via liquid democracy or sortition [1]. This paper establishes the similarity between pruning methods [12, 13] – where classifiers are removed from an ensemble to improve overall performance – and liquid democracy [4, 9]– a framework for voting in which some voters delegate to others rather than voting directly.

Our results show the benefits of this connection: Applying methods from liquid democracy to ensemble learning allow significant reductions in training cost. Analytically, we demonstrate that delegation between classifiers is highly unlikely to reduce group accuracy. Experimentally, our results show that a variety of delegation mechanisms can effectively reduce training cost without hurting accuracy. On some datasets, our methodology results in higher accuracy than common boosting methods

2 DELEGATIVE ENSEMBLE PRUNING

In this paper we develop a pruning process for incremental learning that parallels delegation within liquid democracy, wherein voters choose between participating directly or delegating their vote (and any delegations they have received) to another voter.

Our algorithm runs for some fixed number of steps, denoted T. Training data is divided into T equally sized disjoint sets. At time step $t \leq T$, 3 actions occur: (1) Each classifier that remains in the ensemble is partially trained on the t^{th} set of training data. This adds to their existing training, rather than replacing prior training, and requires a model that allows for partial training (such as an SVM or Neural Network). (2) A fixed, parameterized proportion of classifiers with the worst performance is chosen to delegate. (3) Using one of the delegation methods described below, each newly selected delegator chooses which classifier should receive their delegation. This removes the new delegator from the ensemble and increases the weight of the classifier receiving their delegation by an amount equal to the removed classifier's weight. This algorithm is described fully in the extended version of this paper [2].

As the algorithm progresses, fewer classifiers are trained at each increment – reducing computational requirements by a predictable and parameterized amount – and the most accurate classifiers are empowered based on their accuracy. Delegation mechanisms are functions that determine the flow of delegations within an ensemble. One mechanism is used across the entire ensemble. The mechanism is given a single classifier (as well as information about other delegations and classifier accuracies) and determines to which of the remaining classifiers the original classifier should delegate. We

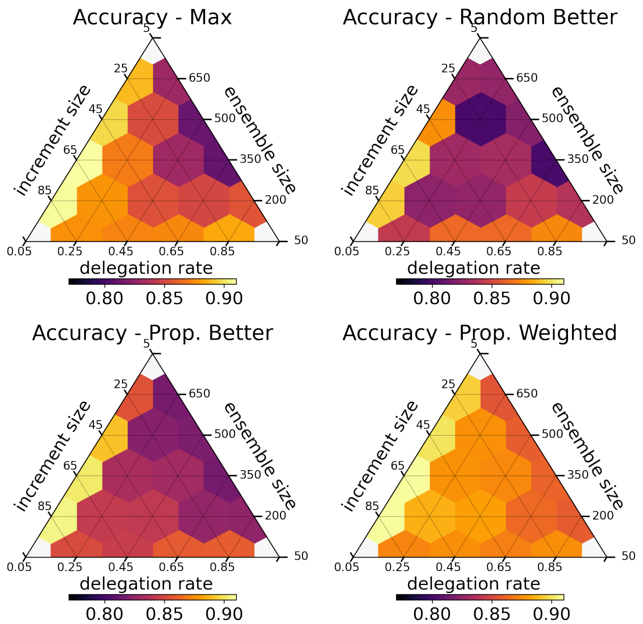

Figure 1: Test accuracy of fully trained ensemble across delegation mechanisms as parameters affecting accuracy are varied. Results are averaged over 50 trials using the spam base dataset. Hexagon colour corresponds to accuracy (lighter is better) and location corresponds to parameter values on the outside of each triangle. The Proportional Weighted mechanism has the highest accuracy across most parameter values.

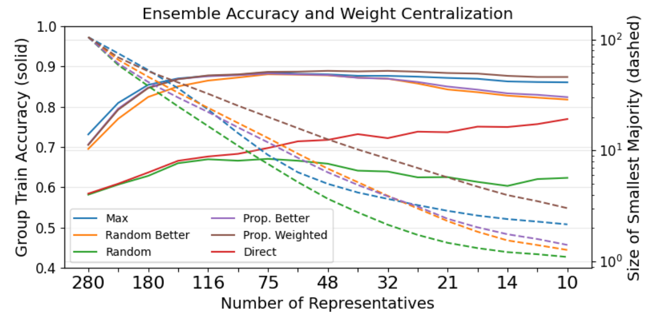

Figure 2: Ensemble behaviour as delegation proceeds, averaged over 500 randomly initialized trials on the spam base dataset. Prop. Weighted delegations consistently have higher accuracy and lower weight centralization. (solid) Ensemble accuracy on test data after each period of incremental training and delegation. (dashed) A minimum number of classifiers required to make a majority of weight.

explore the effects of the following delegation mechanisms on the accuracy, F1 score, and weight distribution of the final ensemble:

Max: Each delegating classifier delegates to the most competent classifier with less weight than the next most competent classifier.

Random Better: Each delegating classifier delegates to a classifier more competent than themselves chosen uniformly at random.

Proportional Better: Each delegating classifier delegates to a classifier more competent than themselves chosen with probability proportional to the competence of each potential representative.

Proportional Weighted: Each new delegator delegates to a more competent classifier with probability proportional to competence *and* weight of each potential representative, such that those with higher weight are less likely to receive delegations.

3 EXPERIMENTS

Our delegation procedure has several parameters which affect both the accuracy and computational training cost of ensembles. We perform a wide variety of experiments to explore the extent to which parameter values can maximize accuracy and minimize training cost. Specifically, we aim to minimize the total number of training examples learned across all classifiers.

Figure 1 shows ensemble test accuracy on a single dataset over many combinations of parameter values. Parameters being varied are *ensemble size*: number of classifiers in the initial ensemble; *increment size*: number of examples in each training increment; *delegation rate*: fraction of classifiers delegating at each increment. Each hexagon aligns with parameter values on the outside of the triangle and the colour of the hexagon indicates ensemble accuracy (lighter is more accurate). Ensembles of medium size with low delegation rate are found to be most accurate and the Proportional Weighted mechanism consistently performs best.

Figure 2 illustrates the training performance over time of an ensemble with an increment size of 25 and, delegation rate of 0.2 that began training with 350 classifiers. Solid lines show accuracy plateaus quickly during delegation and eventually begin to degrade under some delegation mechanisms. The number of classifiers required to form a majority of weight (dashed lines) reduces significantly, showing that weight becomes quite concentrated. However, some delegation mechanisms are more likely to avoid giving a majority of weight to a single classifier. Full experiments on several other datasets are shown in our full paper [2].

4 DISCUSSION

Our experiments demonstrate the benefits of delegation in multiple ways. We see that when compared with full ensembles, delegation improves accuracy while reducing training cost. The amount of improvement varies between datasets. The parameterization of our algorithm allows control over the extent of the training cost reduction. Additionally, the choice of delegation mechanism matters. The Proportional Weighted mechanism outperforms others by considering weight in addition to accuracy. Comparison with Adaboost shows that *on some datasets*, this procedure significantly outperforms boosting while providing similar performance on others.

This setting is very well suited for extension to other learning settings. In online learning or continual learning frameworks, ensembles may adapt to domain shifts in data by delegating to classifiers that have previously been pruned, if they are more suited for a new domain. This may also form a natural method of avoiding "catastrophic forgetting" by having classifiers delegate when out-of-domain and resume learning when useful.

REFERENCES

[1] Shiri Alouf-Heffetz, Ben Armstrong, Kate Larson, and Nimrod Talmon. 2022. How Should We Vote? A Comparison of Voting Systems within Social Networks. In *Thirty-First International Joint Conference on Artificial Intelligence (IJCAI 2022)*. 1734–1740.

[2] Ben Armstrong and Kate Larson. 2024. Liquid Democracy for Low-Cost Ensemble Pruning. arXiv:2401.17443

[3] Eric Bauer and Ron Kohavi. 1999. An Empirical Comparison of Voting Classification Algorithms: Bagging, Boosting, and Variants. *Machine Learning* 36 (1999), 105–139.

[4] Zoé Christoff and Davide Grossi. 2017. Binary voting with delegable proxy: An analysis of liquid democracy. *arXiv preprint arXiv:1707.08741* (2017).

[5] Cristina Cornelio, Michele Donini, Andrea Loreggia, Maria Silvia Pini, and Francesca Rossi. 2021. Voting with Random Classifiers (VORACE): Theoretical and Experimental Analysis. *Autonomous Agents and Multi-Agent Systems* 35, 2 (2021), 22.

[6] Florin Leon, Sabina-Adriana Floria, and Costin Bădică. 2017. Evaluating the Effect of Voting Methods on Ensemble-Based Classification. In *2017 IEEE International Conference on INnovations in Intelligent Systems and Applications*. IEEE, 1–6.

[7] Jean Antoine Nicolas et al. 1785. *Essai sur l'application de l'analyse à la probabilité des decisions rendues à la pluralité des voix. Par m. le marquis de Condorcet,...* de l'Imprimerie Royale.

[8] David Patterson, Joseph Gonzalez, Quoc Le, Chen Liang, Lluis-Miquel Munguia, Daniel Rothchild, David So, Maud Texier, and Jeff Dean. 2021. Carbon Emissions and Large Neural Network Training. *arXiv preprint arXiv:2104.10350* (2021).

[9] Alois Paulin. 2020. An overview of ten years of liquid democracy research. In *The 21st Annual International Conference on Digital Government Research*. 116–121.

[10] David M Pennock, Pedrito Maynard-Reid II, C Lee Giles, and Eric Horvitz. 2000. A Normative Examination of Ensemble Learning Algorithms.. In *ICML*. 735–742.

[11] Omer Sagi and Lior Rokach. 2018. Ensemble Learning: A Survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 8, 4 (2018), e1249.

[12] Christino Tamon and Jie Xiang. 2000. On the boosting pruning problem. In *Machine Learning: ECML 2000: 11th European Conference on Machine Learning Barcelona, Catalonia, Spain, May 31–June 2, 2000 Proceedings 11*. Springer, 404–412.

[13] Zhi-Hua Zhou, Jianxin Wu, and Wei Tang. 2002. Ensembling neural networks: many could be better than all. *Artificial intelligence* 137, 1-2 (2002), 239–263.