ENSEMBLE LEARNING AS A PEER PROCESS

ABSTRACT

Ensemble learning, in its simplest form, entails the training of multiple models with the same training set. In a standard supervised setting, the training set can be viewed as a 'teacher' with an unbounded capacity of interactions with a single group of 'trainee' models. One can then ask the following broad question: *How can we train an ensemble if the teacher has a bounded capacity of interactions with the trainees?*

Towards answering this question we consider how humans learn in peer groups. The problem of how to group individuals in order to maximize outcomes via cooperative learning has been debated for a long time by social scientists and policymakers. More recently, it has attracted research attention from an algorithmic standpoint which led to the design of grouping policies that appear to result in better aggregate learning in experiments with human subjects.

Inspired by human peer learning, we hypothesize that using partially trained models as teachers to other less accurate models, i.e. viewing ensemble learning as a peer process, can provide a solution to our central question. We further hypothesize that grouping policies, that match trainer models with learner models play a significant role in the overall learning outcome of the ensemble. We present a formalization and through extensive experiments with different types of classifiers, we demonstrate that: (i) an ensemble can reach surprising levels of performance with little interaction with the training set (ii) grouping policies definitely have an impact on the ensemble performance, in agreement with previous intuition and observations in human peer learning.

1 INTRODUCTION

Humans learn in peer groups. That is necessitated by resource constraints, as teachers are relatively scarce and they have a bounded individual teaching capacity. The problem of how to group students in order to maximize outcomes via cooperative learning has been debated for a long time by social scientists and policymakers and it remains a sensitive issue (Esposito, 1973; Richer, 1976; Boaler et al., 2000). However, broadly speaking, it is clear that higher-skilled teachers are usually matched with higher-skilled students, reflecting perhaps a collective intuition that such groupings maximize an overall 'educational welfare' of the human society, under practical resource constraints.

From a machine learning point of view, human educational systems can be viewed as mechanisms for training ensembles of individuals. This analogy between ensemble learning and human education has inspired previous works on ensemble methods (see section 5) and is central to the present work.

1.1 MAIN MOTIVATION AND FOCUS: PEER ENSEMBLE LEARNING

We are inspired by the following human peer learning scenario:

N individuals undergo a learning process in T rounds. In each round, the individuals are divided into K groups, and in each group, the highest-skilled individual becomes the teacher for that group. The question is then how different grouping decisions affect the aggregate knowledge collected by the N individuals in T rounds and the goal is to design grouping policies that maximize aggregate learning.

This scenario was studied recently from an algorithm perspective which led to the design of grouping policies that appear to also lead to better aggregate learning in experiments with human subjects Wei et al. (2021).

Motivated by these recent findings, we study an analogous problem in the context of machine learning. Human individuals are replaced by identical-architecture classifiers. One of the N models is replaced by the 'environment' that holds the training set X. Then, in each of the T rounds of learning, the models get grouped, and the 'best' model M of each group acts as the tutor by providing to the rest of the group M's own partially accurate labeling $M(X)$ of the training set X.

1.2 Motivation #2: Ensemble Training under Training Capacity Constraints

A natural objective when training an ensemble of N classifiers is to maximize test accuracy. Therefore there has been significant research in ensemble training algorithms, like Adaboost (Freund & Schapire, 1995; 1999) and Gradient Boost (Breiman, 1997; Friedman, 2001). These algorithms are sequential by nature. On the other hand, maximizing accuracy under natural training-time constraints leads to parallel algorithms, with bagging being a prominent example (Efron, 1979; Breiman, 1996). An interesting characteristic of bagging is that -by design- each classifier has restricted access to the dataset, both to its points and their attributes. It is precisely that restriction that makes bagging more powerful relative to a basic ensemble.

In this paper we view the number of accesses to the labels as an additional resource, and we impose a *constraint* on how many learners can interact in the standard forward-backpropagation manner with the true labels in each parallel round of learning[1].

Learning an ensemble of models under constraints on the frequency of querying the training set, is a question that – to our knowledge– has not been considered in the literature. While the question is not motivated by current practical considerations, it is not hard to imagine ways for it to acquire practical importance. More importantly though, it is a potentially fundamental learning-theoretic question whose study can lead to novel insights in machine and human learning.

1.3 Contributions

There are conceivably many ways one can specify and further study our main question. We approach it via the analogy between human education and ensemble learning. We hypothesize that using partially trained models as 'teachers' to other less accurate models, i.e. viewing ensemble learning as a peer process, can provide a solution to our central question. We further hypothesize that grouping policies should play an important role in maximizing the performance of the ensemble.

Towards addressing our hypotheses, we present a concrete formulation of the problem and we conduct extensive experiments with various types of neural network classifiers. We demonstrate that:

a. Grouping policies have a significant impact on the ensemble performance, in agreement with previous theoretical findings and observations in human peer learning.

b. A peer learning-based ensemble can reach surprising levels of performance even under training resource constraints.

2 Peer Ensemble Learning: Formulation and Questions

We are given a training set X with categorical labels y. We want to train an ensemble of N-1 identical-architecture classifiers $h_1, \ldots, h_{N-1}$, in T *training rounds*. In each round, each classifier C_i can either either act as a learner, or as a trainer by providing its (partially correct) predictions $h_i(X)$ for the training set X. We view the training set as the Nth "classifier" h_N, which will always be used as a trainer providing $y = h_N(X)$ to some of the learners.

In a standard parallel ensemble training, the training set is used to train N-1 classifiers per round, as shown in Figure 1(a). Motivated by peer learning we generalize the parallel round to a setting

[1] In this paper we consider an epoch as one round of learning.

where the N classifiers are split into k groups of size $C = N/k$ and each group has one trainer h_i that provides labels $h_i(X)$ to its group. The groups and their trainers can be updated in each round.

The above scenario implies a **training capacity constraint** $C = N/k - 1$ for h_N.

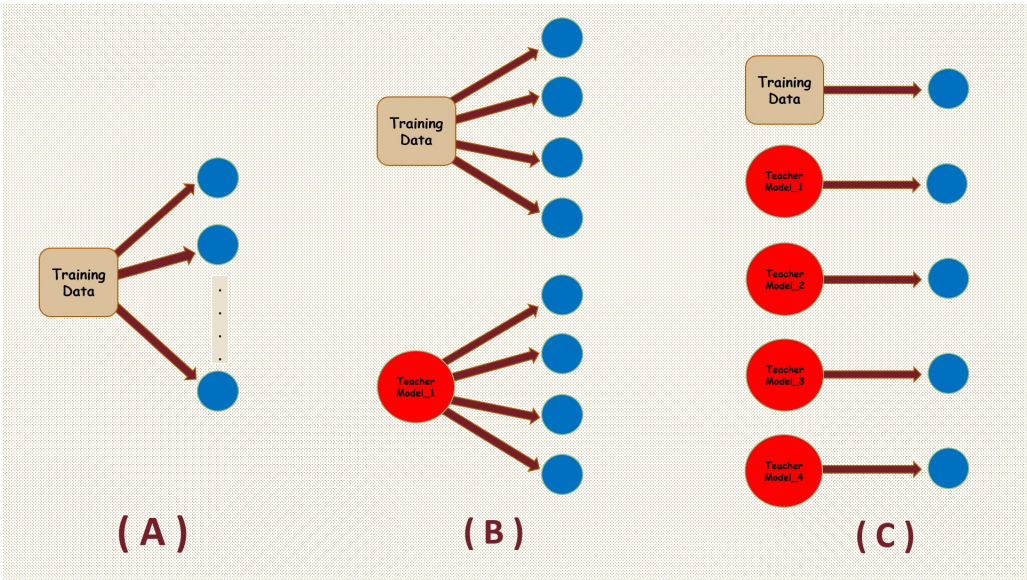


Figure 1: **(a)** One round of training in basic ensemble learning. **(b,c)** Peer ensemble learning with groups of size 2 and $N/2$. The trainer models and their mappings to learner models can change in every round.

Under this general framework we pursue two basic directions of research.

a. Grouping policies and their effect under a fixed training capacity. Assuming a fixed value for the training capacity $C = N/k$, there are $M = N!/C!$ different possible groupings in each round, and M^T different *policies*, i.e. sequences of groupings. One basic question is to what extent the choice of policy affects various metrics of ensemble performance.

We can identify a natural baseline: a random split into k groups in each round. We can then seek to design policies that will outperform a baseline random policy.

Looking at the ensemble from a social perspective, we are interested in studying metrics beyond the standard validation/test accuracy of the ensemble, but also the aggregate accuracy, the median accuracy, or a random sub-ensemble accuracy. We are more generally interested in properties of the distribution of the models $h_1, \ldots, h_{N-1}$ after T rounds.

b. The effect of the training capacity restriction. It is probably expected that imposing the training capacity restriction C may result in reduced ensemble performance metrics, e.g. reduced accuracy for a given number of epochs/rounds. On the other hand it is a priori unclear if a capacity-restricted ensemble can make a more efficient use of the training set, i.e. reach a *higher* accuracy for a fixed number of access to the true labels $h_N(X)$. Thus our second goal is to examine the effect of C on key ensemble metrics.

3 Experimental Study

Architectures and Datasets. We do experiments with three different types of architectures and two different types of datasets. More concretely, we use 'toy' versions of LeNet and Resnet (LeCun et al., 1998; He et al., 2016). These networks are used on Fashion-MNIST dataset which is consist of 60,000 training images and 10,000 test images of fashion and clothing items, taken from 10 classes, where each image is a standardized 28×28 size in grayscale (784 total pixels). We also use a standard GCN (Kipf & Welling, 2017; Zhou et al., 2020) on ogbn-arxiv dataset which is a un-directed graph, representing the citation network between all Computer Science (CS) arXiv papers, where each node

is an arXiv paper with a 128-dimensional feature vector. For each of these datasets we use a fixed split to training, validation and test datasets.

Pre-training. Motivated by the analogy to human peer learning, in addition to starting the entire ensemble training from scratch (i.e. with N-1 randomly initialized models), we also consider the effect pre-training, where the initial models undergo different degrees of training with the true labels, before they enter the ensemble training. More concretely, learner i undergoes i rounds of pre-training.

Number of learners and groups. We set $N = 10$ throughout our experiments. We consider the cases shown in Figure 1, i.e. $k = 1$, $k = 2$ and $k = N/2$. We refer to these cases as 'Split-in-Two' and 'Split-in-Five' groupings.

Policies. Grouping policies are based on measuring the validation accuracy of the learners $h_1, \ldots, h_{N-1}, h_N$ before **each round**. Concretely, let v_i be the validation accuracy of h_i, and let m_i be the models whose validation accuracy is the ith lowest in the list $\{v_1, \ldots, v_N\}$.

Following Wei et al. (2021) we define two policies for selecting the groups:

Dynamic-A: [Best-Trains-Best]. The trainers of the k groups are models $m_N, \ldots, m_{N-k+1}$, i.e. the models with highest validation accuracy. The rest of the ordered list $m_{N-k}, \ldots, m_1$ is split into k contiguous buckets that are assigned in that order to $m_N, \ldots, m_{N-k+1}$. In other words, the best trainers train the best learners.

Dynamic-B: [Equitable] The trainers of the k groups are models $m_N, \ldots, m_{N-k+1}$. The rest of the models in the ordered list $m_{N-k}, \ldots, m_1$ are assigned in a round-robin fashion to $m_N, \ldots, m_{N-k+1}$. In this grouping each trainer selects 'equitably' its learners.

We also consider the following two baselines.

Random:. The models are randomly split into k groups, and in each group the model with highest validation accuracy becomes the trainer for that group.

Static: [Best-Trains-Worst] The trainers of the k groups are models $m_N, \ldots, m_{N-k+1}$, i.e. the models with highest validation accuracy. The rest of the ordered list $m_{N-k}, \ldots, m_1$ is split into k contiguous buckets that are assigned in the reverse order to $m_N, \ldots, m_{N-k+1}$. In other words, the best trainers train the worst learners. This grouping is decided once before the first round and stays the same throughout the T rounds.

Number of random experiments. For each setting for the tuple (*pre-training*,k, *Policy*,*Dataset*) we train $N_{dataset}$ experiments, where $N_{lenet} = 10$, $N_{resnet} = 20$, $N_{GCN} = 10$. We report a number of metrics for each round/epoch, taking the average of $N_{dataset}$ values for each metric.

Metrics. In each round we record and report: (i) The ensemble accuracy on the test set. To compute the ensemble prediction on a test point x, we compute N-1 probability distributions output by the softmax layer of the corresponding classifiers, we multiply these probabilities pointwise, and we select label corresponding the maximum product. (ii) The average accuracy of the N-1 classifiers on the dataset.

4 RESULTS

In this section we summarize our experimental results.

a. The effect of grouping policies. The choice of grouping policy appears to affect the outcome of the ensemble training. In Figure 2 we plot the ensemble accuracy per epoch and we comment on the findings. Overall the results show that *Dynamic-A* gives a higher ensemble accuracy and beats the *Random* baseline, although this is not entirely clear for Resnets. Similar observations hold for the average ensemble accuracy as shown and discussed in Figure 4.

b. The effect of no pre-training. The experiments of part (a) are repeated without using pre-training. The results are summarized in figures 7 and 6 of the Appendix. The absence of pre-training induces a smoother convergence behavior, reduces the difference in performance among policies, and possibly renders *Dynamic-B* better for average accuracy, at least for some architectures.

c. The effect of the Capacity constraint. In Figure 5 we plot the ensemble accuracy with respect to the total number of forward operations. Split-In-Five access the ground truth only once per 5 forward operations, i.e. a 20% frequency, whereas Split-In-Two access the ground truth 4 times per 8 forward operations, a 50% frequency. Not surprisingly, Split-In-Two outperforms Split-In-Five.

In Figure 3 we plot the ensemble accuracy by the number of true label accesses. Split-In-Two does 4 non-true label access for each true label access, while the corresponding Split-In-Five ratio is 1/1. Nevertheless it squeezes more accuracy out of its limited ground truth use.

d.Misc results. The experiments of part (c) above are repeated for other policies. While the results are the same for *Dynamic-B* and *Random*, an interesting phenomenon occurs for *Static* where Resnet Split-in-Five ensembles are not just unable to learn, but tend to behave worse for later rounds. Explaining this behavior is left open for future work.

5 Related Works

We view this paper as an addition to the broad literature on machine learning methods inspired by human education, including its social aspects. We have in particular drawn inspiration from curriculum learning Bengio et al. (2009); Wu et al. (2021); Soviany et al. (2021) and works on learning ensembles with diverse priors Jain et al. (2021).

6 Conclusion and Future Work

We presented a study of ensemble learning as a peer process. Our work is inspired by human peer group learning and it has been further motivated by a computational view of ensemble learning as a parallel process with constraints on the ground truth access.

We performed a large set of experiments that appear to confirm that grouping policies play a role in aggregate measures of ensemble learning, as we intuitively tend to believe about human peer learning. Interestingly, our experiments show that a specific grouping policy designed in the context of theoretical research on human peer learning Wei et al. (2021) appears to also result in better peer ensemble learning.

In this paper we have omitted studying properties of the accuracy distribution of the ensemble models, in particular with respect to inequality (or "skill diversity") measures that have been central in the theoretical study of Wei et al. (2021), or other ensemble properties like robustness to random or adversarial noise. We thus believe that this work is only the first step towards discovering and studying new machine learning phenomena whose theoretical explanation can be an interesting challenge.

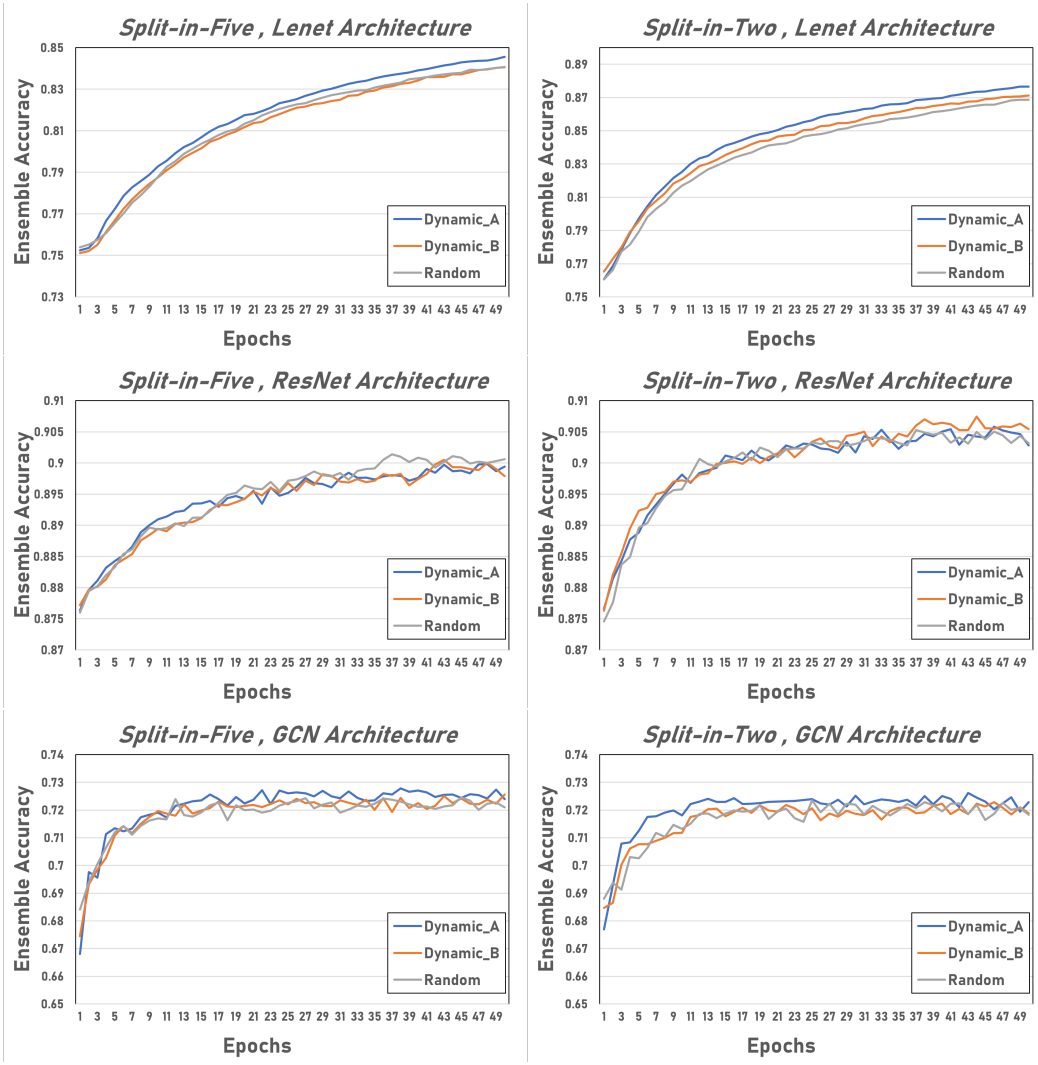
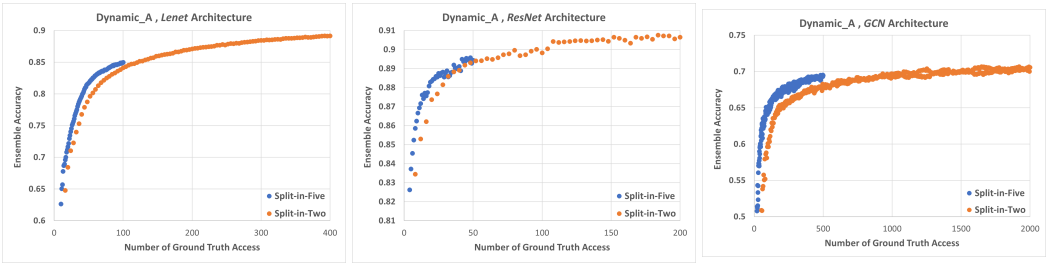


Figure 2: Ensemble Accuracy for **Split-in-Five** (left) and **Split-in-Two** (right) Grouping with Pre-Training. *Dynamic-A* is clearly better in Lenet, and GCN, but more random experiments are needed to stabilize the output for Resnet.



(a) Lenet (b) ResNet (c) GCN

Figure 3: **Dynamic A** Policy, Accuracy per true label accesses. Split-in-Two makes a more efficient use of ground truth relative to Split-in-Five.

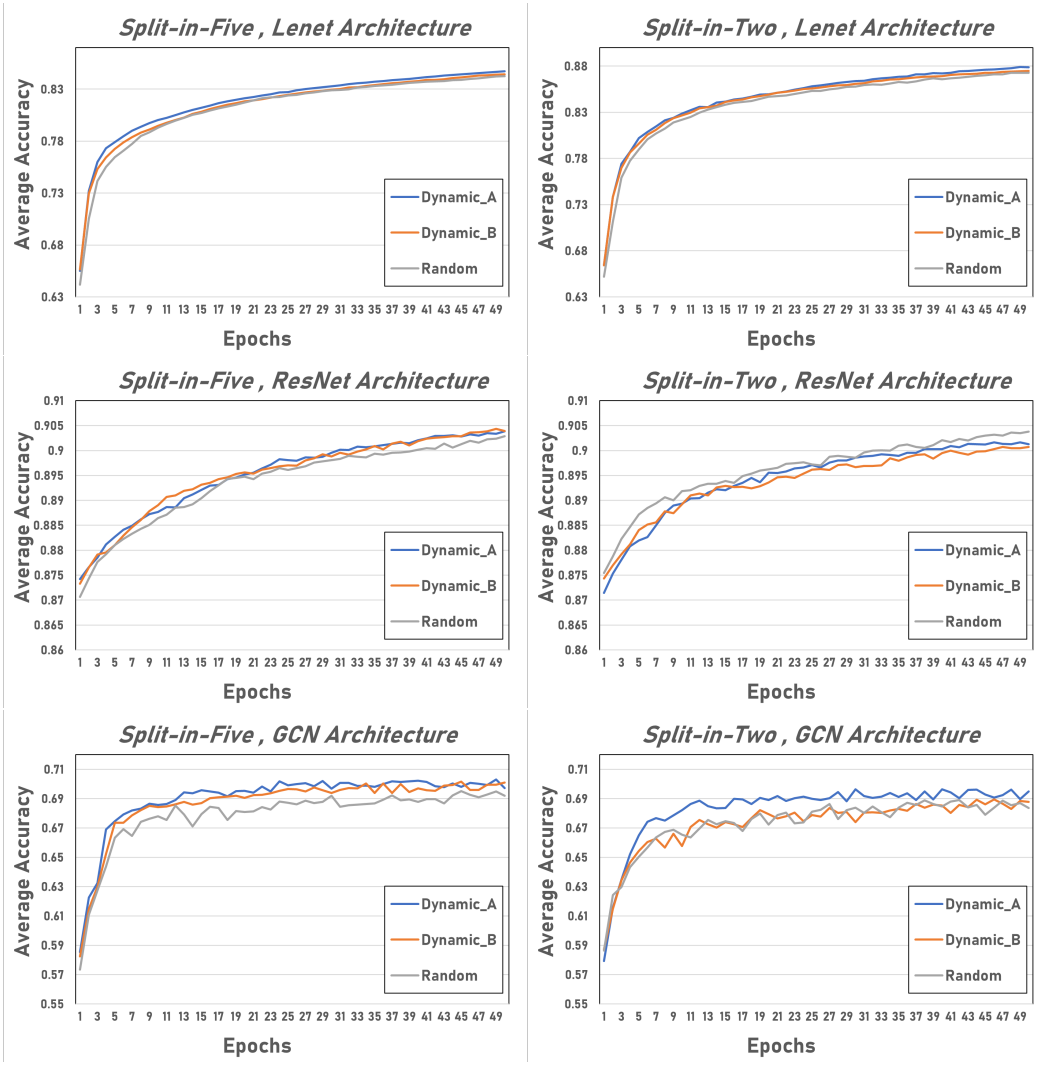
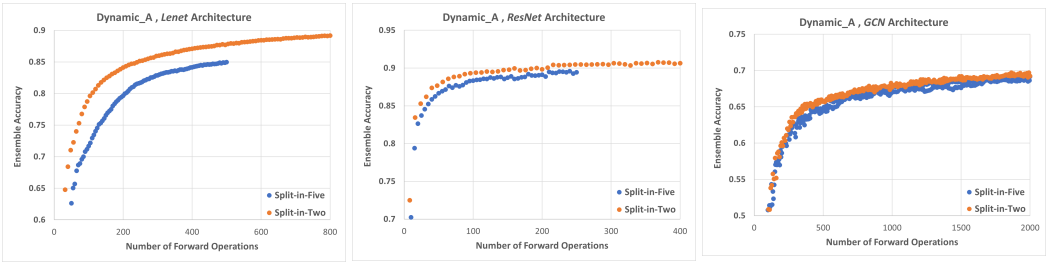


Figure 4: Average Accuracy for **Split-in-Five** (left) and **Split-in-Two** (right) Grouping without Pre-Training. *Dynamic-A* appears to outperform in Lenet and GCN although by smaller margins. The results are mixed for Resnet and more random experiments are needed to stabilize the output.



(a) Lenet (b) ResNet (c) GCN

Figure 5: *Dynamic-A* Policy, accuracy by number of forward operations without Pre-Training. Not surprisingly Split-In-Two ouperforms Split-In-Five, as it makes 4x higher use of the training set.

REFERENCES

Yoshua Bengio, Jérôme Louradour, Ronan Collobert, and Jason Weston. Curriculum learning. In *Proceedings of the 26th Annual International Conference on Machine Learning - ICML '09*, pp. 1–8, Quebec, Canada, 2009. ACM Press.

Jo Boaler, Dylan Wiliam, and Margaret Brown. Students experiences of ability grouping—disaffection, polarisation and the construction of failure1. *British Educational Research Journal*, 26(5):631–648, 2000.

Leo Breiman. Bagging predictors. *Machine Learning*, 24(2):123–140, August 1996.

Leo Breiman. Arcing the edge. Technical Report 486, Statistics Department, University of California at Berkeley, June 1997.

Bradley Efron. Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7(1): 1–26, 1979.

Dominick Esposito. Homogeneous and heterogeneous ability grouping: Principal findings and implications for evaluating and designing more effective educational environments. *Review of Educational Research*, 43(2):162–179, June 1973.

Yoav Freund and Robert E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. In *Computational Learning Theory, Second European Conference, EuroCOLT*, volume 904 of *Lecture Notes in Computer Science*, pp. 23–37, Barcelona, Spain, March 1995. Springer.

Yoav Freund and Robert E. Schapire. A short introduction to boosting. *Journal of Japanese Society for Artificial Intelligence*, 14(5):771–780, September 1999.

Jerome H. Friedman. Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5):1189–1232, October 2001.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, Las Vegas, Nevada, United States, June 2016. IEEE.

Saachi Jain, Dimitris Tsipras, and Aleksander Madry. Combining diverse feature priors. *CoRR*, arXiv:2110.08220, October 2021.

Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. In *5th International Conference on Learning Representations, ICLR 2017*, Toulon, France, April 2017.

Yann LeCun, Leon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, November 1998.

Stephen Richer. Reference-group theory and ability grouping: A convergence of sociological theory and educational research. *Sociology of Education*, 49(1):65–71, January 1976.

Petru Soviany, Radu Tudor Ionescu, Paolo Rota, and Nicu Sebe. Curriculum learning: A survey. *CoRR*, arXiv:2101.10382, January 2021.

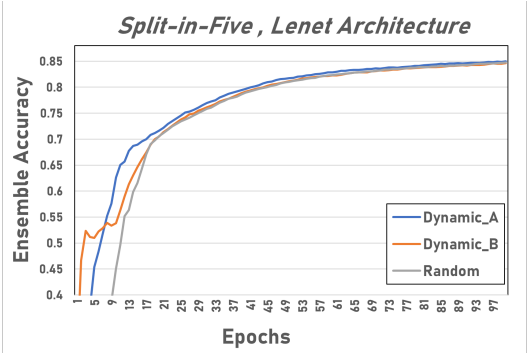
Dong Wei, Ioannis Koutis, and Senjuti Basu Roy. Peer learning through targeted dynamic groups formation. In *2021 IEEE 37th International Conference on Data Engineering (ICDE)*, pp. 121–132, Chania, Greece, April 2021. IEEE.

Xiaoxia Wu, Ethan Dyer, and Behnam Neyshabur. When do curricula work? In *9th International Conference on Learning Representations, ICLR 2021*, Austria, May 2021. OpenReview.net.

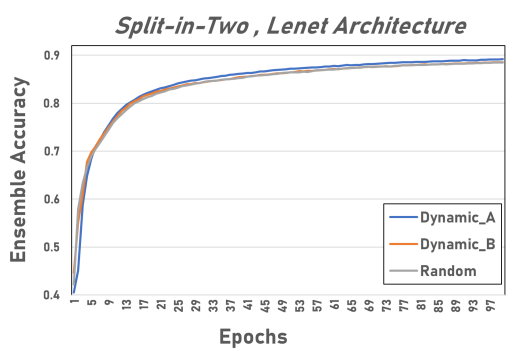
Jie Zhou, Ganqu Cui, Shengding Hu, Zhengyan Zhang, Cheng Yang, Zhiyuan Liu, Lifeng Wang, Changcheng Li, and Maosong Sun. Graph neural networks: A review of methods and applications. *AI Open*, 1:57–81, 2020.

A APPENDIX

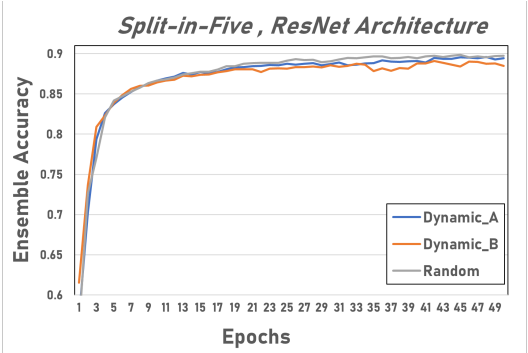
A.1 EFFECT OF GROUPING POLICY WITHOUT PRE-TRAINING



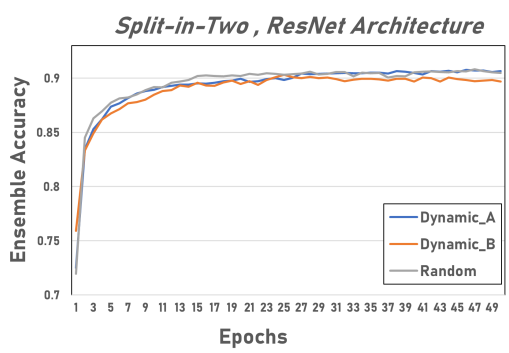
(a) Lenet



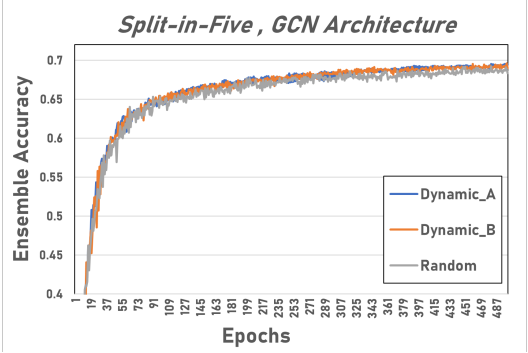
(b) Lenet



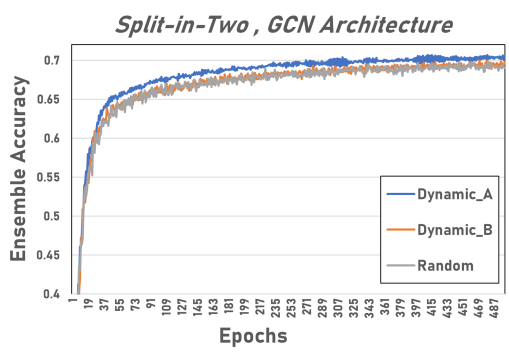
(c) ResNet



(d) ResNet

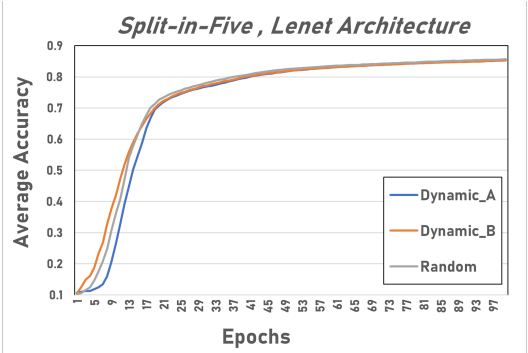


(e) GCN

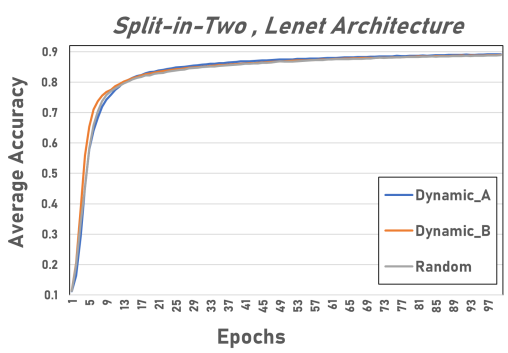


(f) GCN

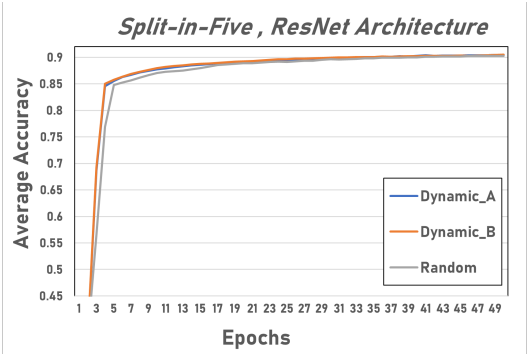
Figure 6: Ensemble Accuracy for **Split-in-Five** and **Split-in-Two** Grouping without Pre-Training. *Dynamic-A* still appears to outperform but by a smaller margin relative to using pre-training. An additional aspect is that the (average) convergence behavior of the ensemble performance becomes smoother.



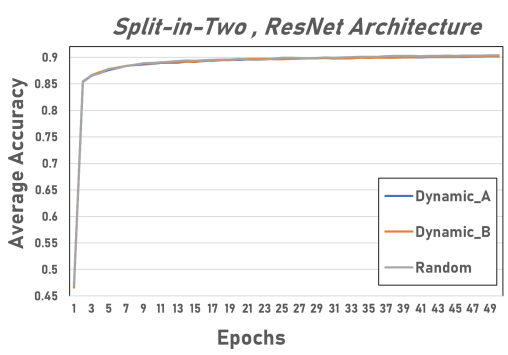
(a) Lenet



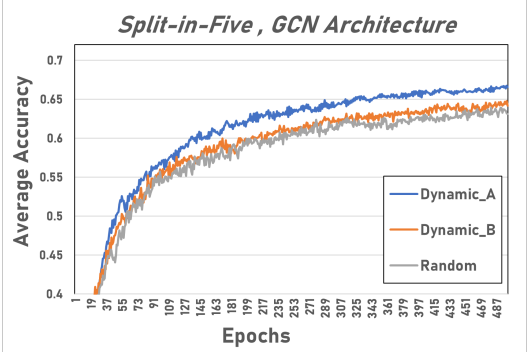
(b) Lenet



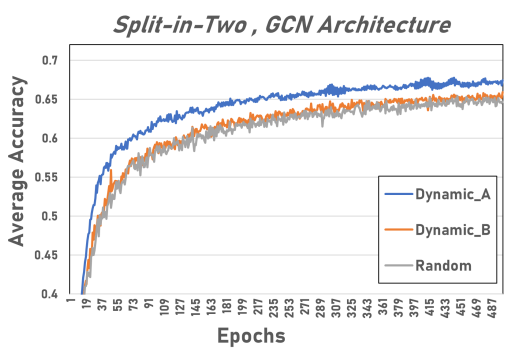
(c) ResNet



(d) ResNet



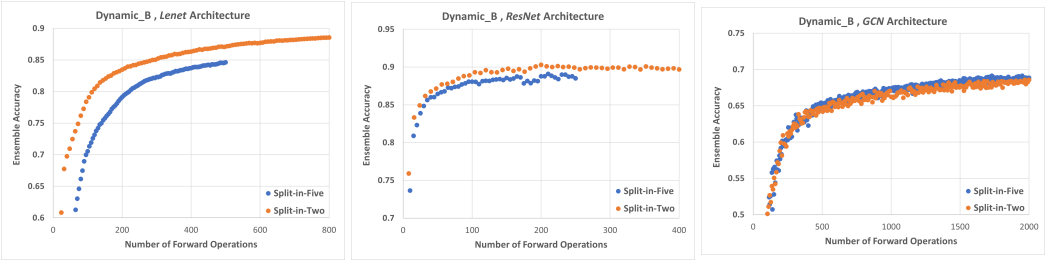
(e) GCN



(f) GCN

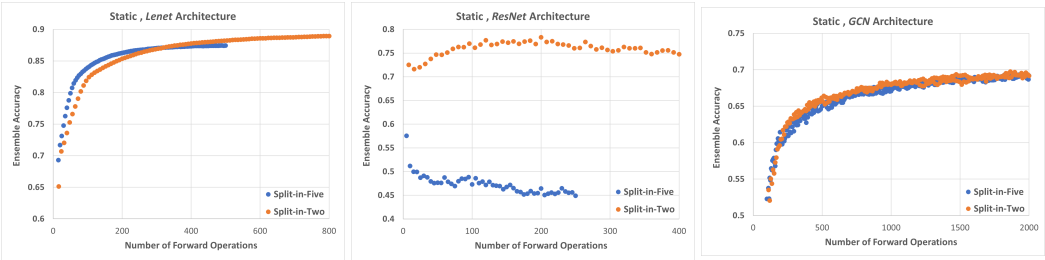
Figure 7: Average Accuracy for **Split-in-Five** (left) and **Split-in-Two** (right) Grouping without Pre-Training. *Dynamic-A* clearly outperforms in GCN, but the situation is reversed in Lenet.

A.2 EFFECT OF TRAINING CAPACITY CONSTRAINT



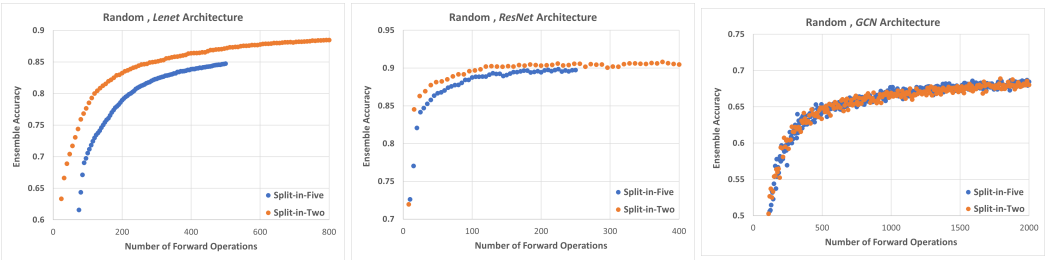
(a) Lenet (b) ResNet (c) GCN

Figure 8: **Dynamic B** Policy, Forward Operations without Pre-Training



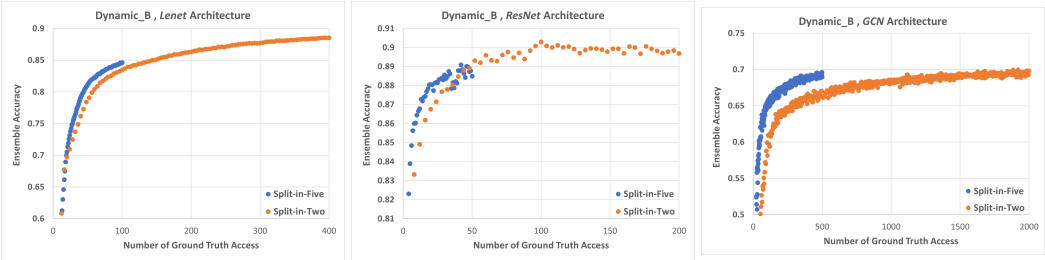
(a) Lenet (b) ResNet (c) GCN

Figure 9: **Static** Policy, Forward Operations without Pre-Training



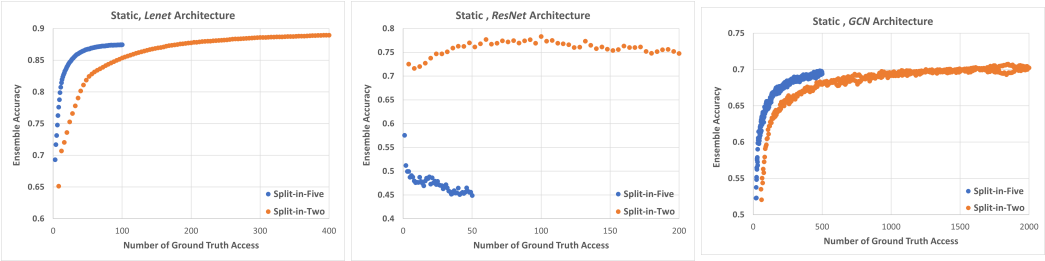
(a) Lenet (b) ResNet (c) GCN

Figure 10: **Random** Policy, Forward Operations without Pre-Training



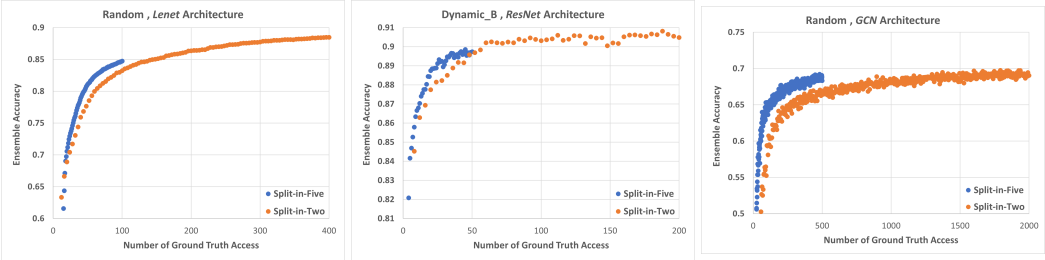
(a) Lenet (b) ResNet (c) GCN

Figure 11: **Dynamic B** Policy, Ground Truth Access without Pre-Training



(a) Lenet (b) ResNet (c) GCN

Figure 12: **Static** Policy, Ground Truth Access without Pre-Training



(a) Lenet (b) ResNet (c) GCN

Figure 13: **Random** Policy, Ground Truth Access without Pre-Training