- Chairs have requested users to enter domain conflicts. Please click here (/ICML2019 /User/ConflictDomains) to enter your conflict domains.

Submission Summary

Conference Name	36th International Conference on Machine Learning
Paper ID	2366
Paper Title	Heavy-Tailed Universality Predicts Trends in Test Accuracies for Very Large Pre-Trained Deep Neural Networks
Abstract	Given two or more Deep Neural Networks (DNNs) with the same or similar architectures, and trained on the same dataset, but trained with different solvers, parameters, hyper-parameters, regularization, etc., can we predict which DNN will have the best test accuracy, and can we do so without peeking at the test data? In this paper, we show how to use a new Theory of Heavy-Tailed Self-Regularization (HT-SR) to answer this. The HT-SR theory suggests, among other things, that modern DNNs exhibit what we call Heavy Tailed Mechanistic Universality (HT-MU), meaning that the correlations in the layer weight matrices can be fit to a power law with exponents that lie in common Universality classes from Heavy Tailed Random Matrix Theory (HT-RMT). From this, we develop a Universal capacity control metric that is a weighted average of these heavy tailed power law exponents, Rather than considering small toy NNs, we examine over 50 different, large-scale pre-trained DNNs, ranging over 15 different architectures, trained on ImagetNet, each of which has been reported to have different test accuracies. We show that this new capacity metrics correlates very well with the reported test accuracies of these DNNs, looking across each architecture (VGG16/.../VGG19, ResNet10/... /ResNet152, etc.). Moreover, we show how to approximate the metric by the more familiar Product Norm capacity measure, as the average of the log Frobenius norm of the layer weight matrices. Our approach requires no changes to the underlying DNN or its loss function, it does not require us to train a model (although it could be used to monitor training), and it does not even require access to the ImageNet data.
Created on	1/17/2019, 10:54:38 PM
Last Modified	1/17/2019, 10:54:38 PM
Authors	**Michael Mahoney** ("University of California, Berkeley") <mmahoney@stat.berkeley.edu> Charles Martin (Calculation Consulting) <charles@CalculationConsulting.com>
Primary Subject Area	Deep Learning -> Algorithms
Secondary Subject Areas	Deep Learning -> Architectures Deep Learning -> Deep Learning Theory Deep Learning -> Optimization

Submission Files	icml_test_accuracy_SUBMv1_MAIN.pdf (527.5 Kb, 1/22/2019, 10:22:19 PM)
Supplementary Files	icml_test_accuracy_SUBMv1_SUPP.pdf (783.2 Kb, 1/22/2019, 10:23:48 PM)
Submission Questions Response	**1. Student author** No **2. Agreement on anonymity** Agreement accepted **3. Agreement on Author List** Agreement accepted **4. Previous or Concurrent Submissions** Agreement accepted **5. Previous or Concurrent Submissions** No. **6. Agreement on Code Submission** Agreement accepted