

TTIC 31230, Fundamentals of Deep Learning

David McAllester, Winter 2020

Double Descent

Double Descent

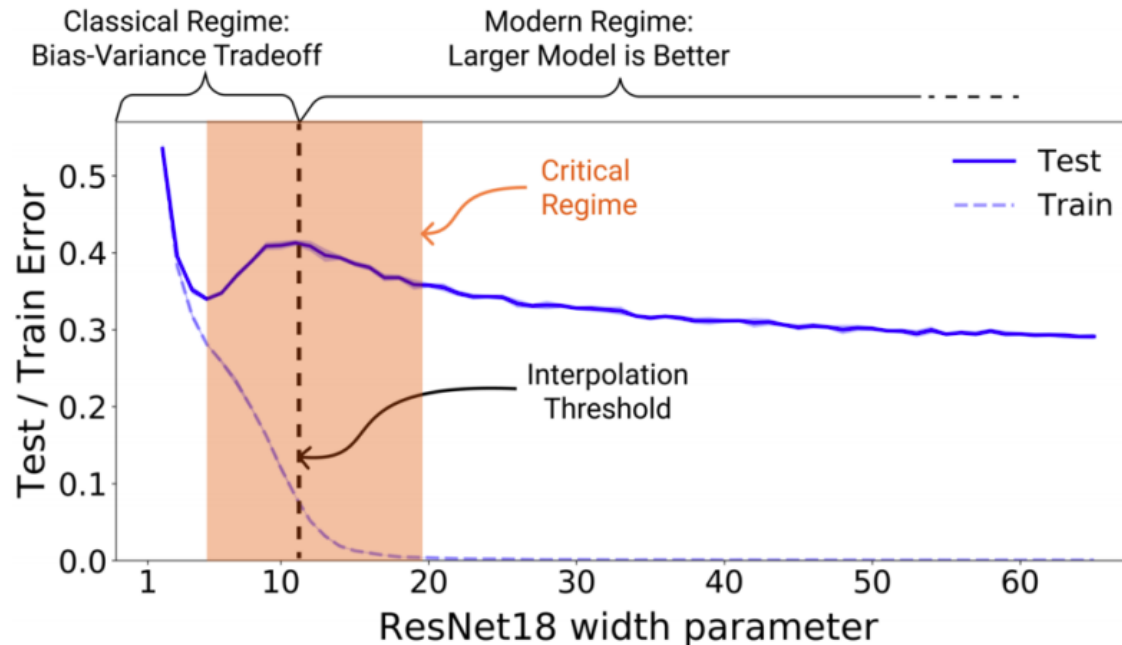
Reconciling modern machine learning practice and the bias-variance trade-off

Mikhail Belkin, Daniel Hsu, Siyuan Ma, Soumik Mandal, arXiv December 2018.

Deep Double Descent: Where Bigger Models and More Data Hurt

Preetum Nakkiran, Gal Kaplun, Yamini Bansal, Tristan Yang, Boaz Barak, Ilya Sutskever, ICLR 2020

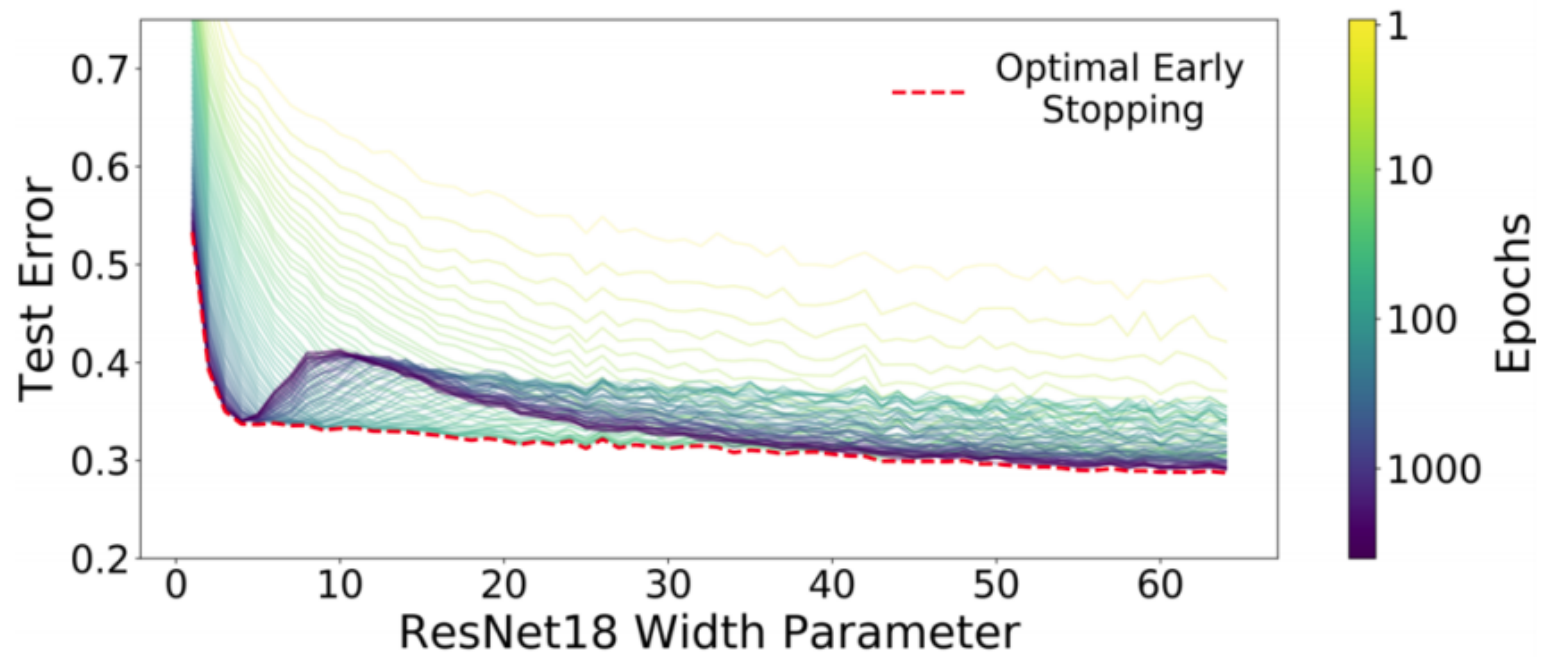
Double Descent



Deep Double Descent: Where Bigger Models and More Data Hurt

Preetum Nakkiran, Gal Kaplun, Yamini Bansal, Tristan Yang, Boaz Barak, Ilya Sutskever, ICLR 2020

Double Descent



Summary

There is never harm in doing early stopping — one should always do early stopping.

Regularization is any modification to the training algorithm motivated by reducing the training-validation gap.

While regularization modifications to training can be inspired by theory, the theory is weak.

Regularization proposals should be evaluated empirically.

END