



CS 329P : Practical Machine Learning (2021 Fall)

# 9. Model Tuning

Qingqing Huang, Mu Li, Alex Smola

<https://c.d2l.ai/stanford-cs329p>

# Manual Hyperparameter Tuning



- Start with a good baseline, e.g. default settings in high-quality toolkits, values reported in papers
- Tune a value, retrain the model to see the changes
- Repeat multiple times to gain insights about
  - Which hyperparameters are important
  - How sensitive the model to hyperparameters
  - What are the good ranges



# Manual Hyperparameter Tuning



- Needs careful experiment management
- Save your training logs and hyperparameters to compare, share and reproduce later
  - The simplest way is saving logs in text and put key metrics in Excel
  - Better options exist, e.g. tensorboard and weights & bias
- Reproducing is hard, it relates to
  - Environment (hardware & library)
  - Code
  - Randomness (seed)



# Automated Hyperparameter Tuning



- Computation costs decrease exponentially, while human costs increase



- Cost per training for a typical ML task:
  - E.g. 1M user logs, 10K images
- Cost of a data scientist per day >\$500
- Use algorithms if it outperforms human after 1000 trials
- Typically beat 90% data scientists



	Time	Cost on cloud
Trees	10min on CPU	\$0.4
Neural networks	1hour on GPU	\$5

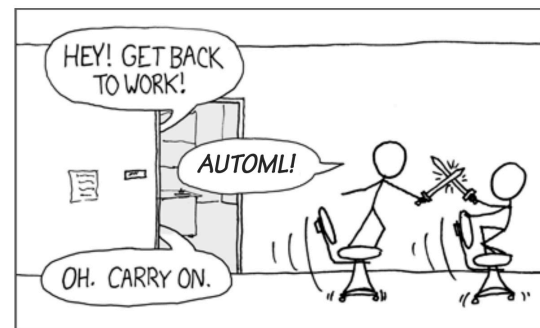


# Automated Machine Learning (AutoML)



- Automate every step in applying ML to solve real-world problems: data cleaning, feature extraction, model selection...
- **Hyperparameter optimization (HPO):** find a good set of hyperparameters through search algorithms
- **Neural architecture search (NAS):** construct a good neural network model

*THE DATA SCIENTIST'S #1 EXCUSE FOR  
LEGITIMATELY SLACKING OFF:  
"THE AUTOML TOOL IS OPTIMIZING MY MODELS!"*



# Summary



- Hyperparameter tuning aims to find a set of good values
- It's time consuming as data preprocessing
- There is a trend to use algorithm for tuning