HYPERPARAMETER OPTIMIZATION

MODEL UNCERTAINTY IN MACHINE LEARNING

Todd Young
University of Tennessee

01

INDUCTIVE BIAS

ASSUMPTIONS OF OUR MODELS

02

NO FREE LUNCH

HOW WELL YOU DO IS DETERMINED BY HOW 'ALIGNED' YOUR LEARNING ALGORITHM $P(h|d)$ IS WITH THE ACTUAL POSTERIOR, $P(f|d)$.
-DAVID H. WOLPERT

03

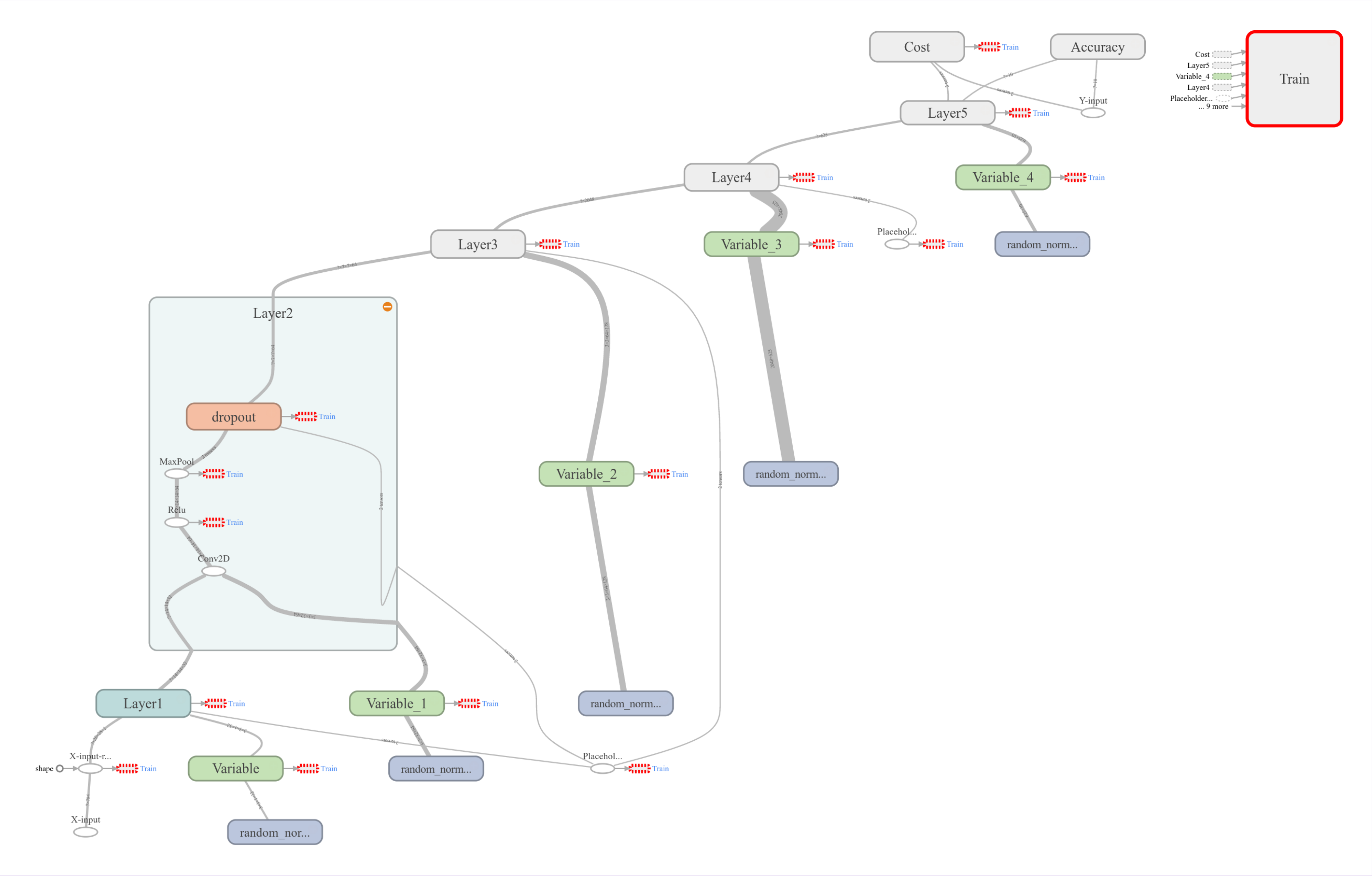
DEEP LEARNING?

ONLY WITH THE PROPER PARAMETERS

PARAMETER CHOICES, A DARK ART

Harry Potter and the Half Blood Prince

FIVE LAYER CONVOLUTIONAL NEURAL NETWORK



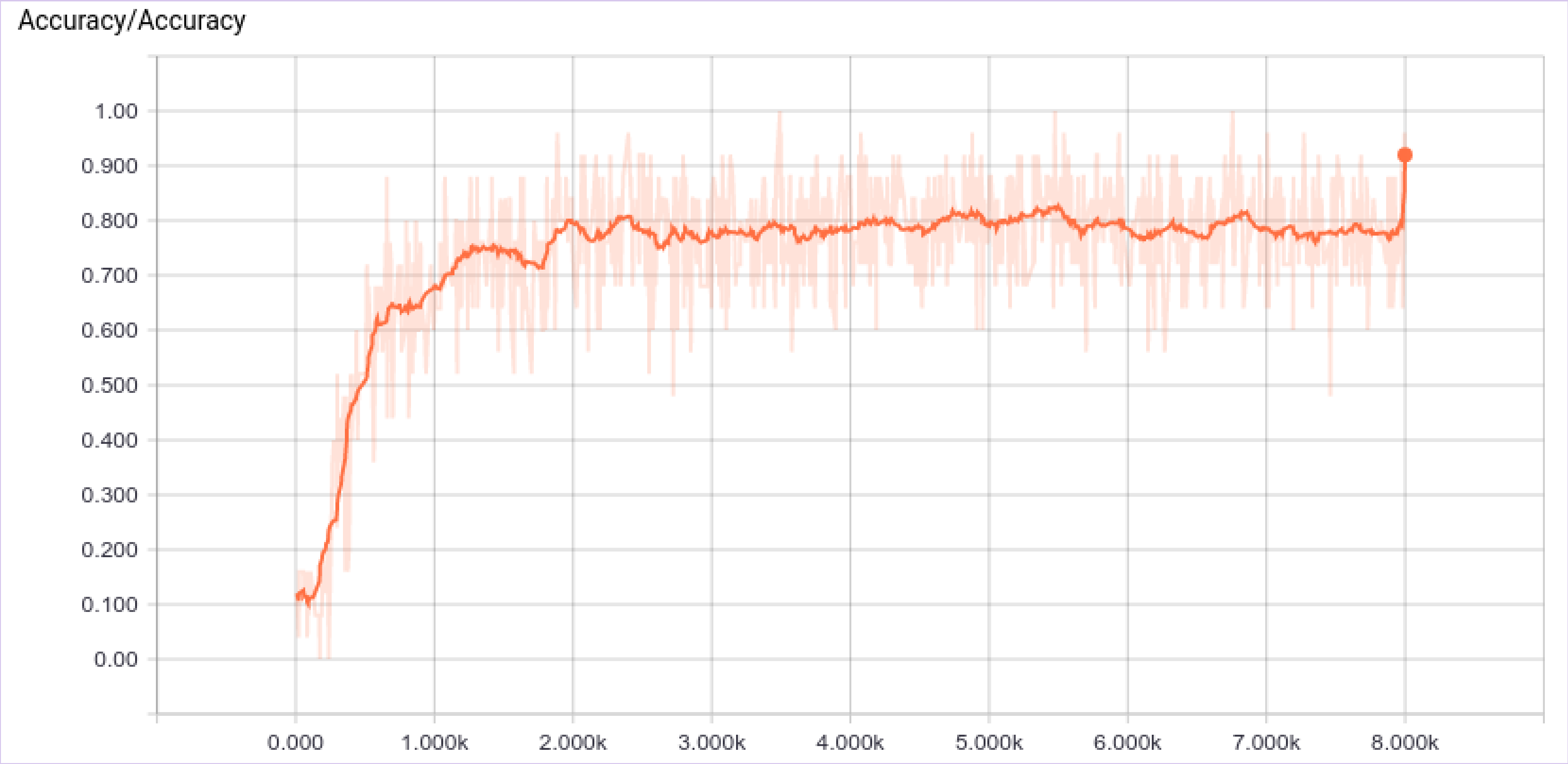
CNN ONE PARAMETERS:

Learning Rate: 0.1;

Number of Epochs: 8000; Batch Size: 25;

Convolution Layer Dropout: 0.2; Fully Connected Layer Dropout: 0.2

TEST ACCURACY: 0.937



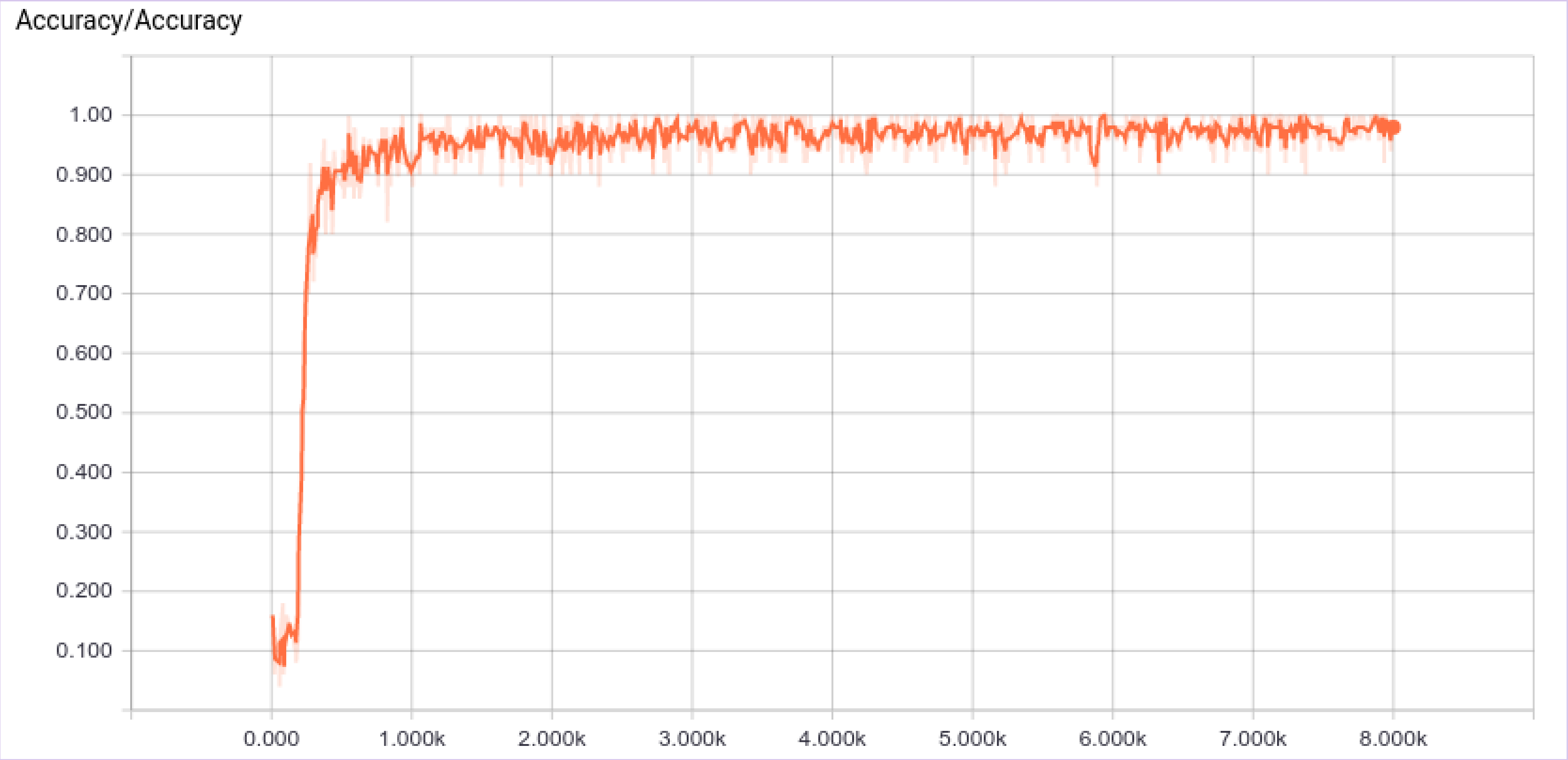
CNN TWO PARAMETERS:

Learning Rate: 0.01;

Number of Epochs: 8000; Batch Size: 50;

Convolution Layer Dropout: 0.5; Fully Connected Layer Dropout: 0.8

TEST ACCURACY: 0.9896



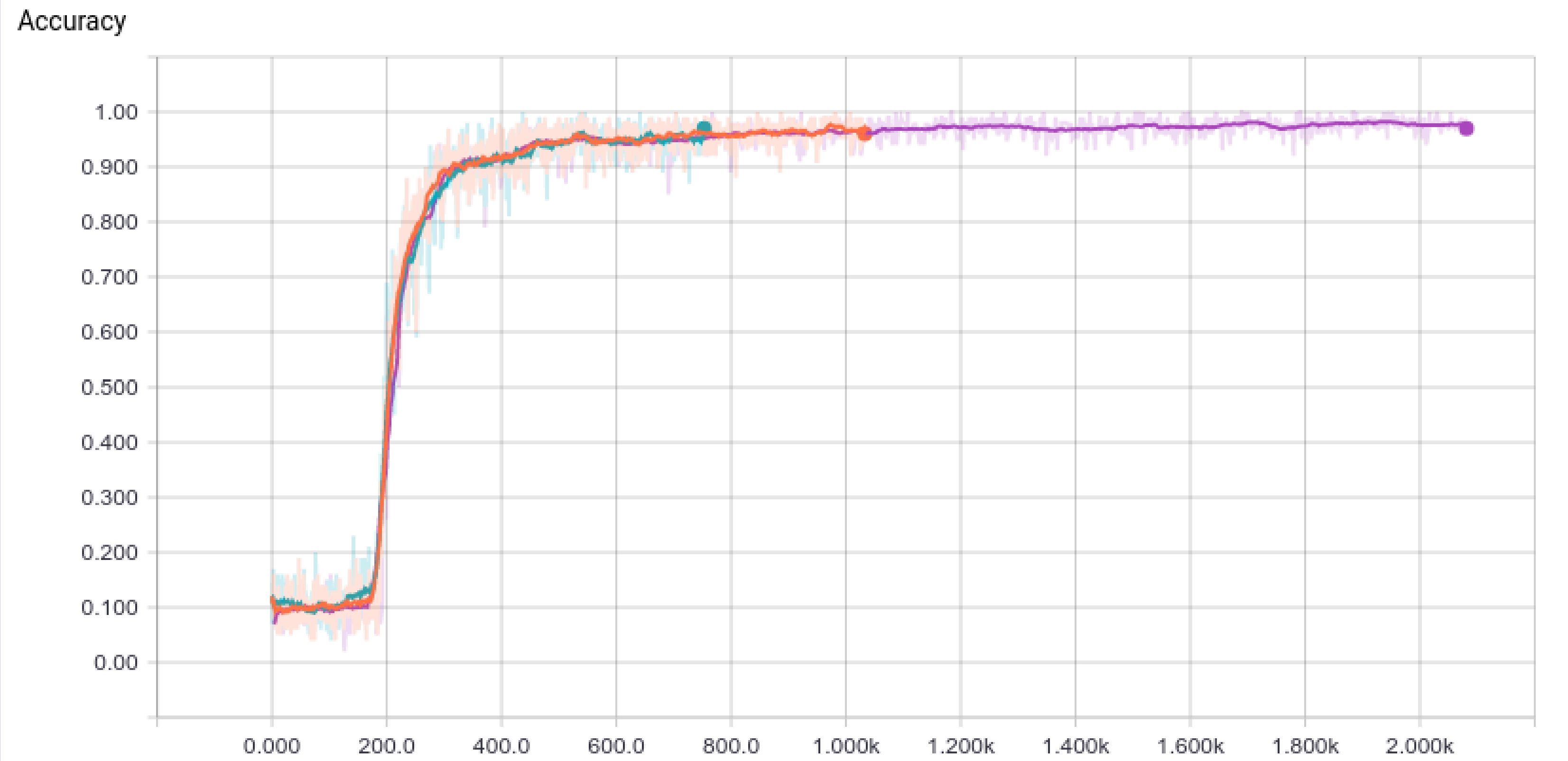
CNN THREE PARAMETERS:

Learning Rate: 0.001;

Number of Epochs: 8000; Batch Size: 100;

Convolution Layer Dropout: 0.7; Fully Connected Layer Dropout: 0.5

TEST ACCURACY: 0.9925



METHODS

FORMALIZING OUR HYPERPARAMETER SEARCH

01

RANDOM SEARCH

J BERGSTRA, Y BENGIO 'RANDOM SEARCH FOR HYPER-PARAMETER OPTIMIZATION' - JOURNAL OF MACHINE LEARNING RESEARCH, 2012

02

TREE-STRUCTURED PARZEN ESTIMATORS

HYPEROPT

BERGSTRA, JAMES S., ET AL. “ALGORITHMS FOR HYPER-PARAMETER OPTIMIZATION.” ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS. 2011.

03

GAUSSIAN PROCESSES

SPEARMINT

PRACTICAL BAYESIAN OPTIMIZATION OF MACHINE LEARNING ALGORITHMS
JASPER SNOEK, HUGO LAROCHELLE AND RYAN PRESCOTT ADAMS
ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, 2012

FUTURE WORK