Hyperparameter Tuning Neural Networks

Brian Won

August 4, 2025

Contents

1 Lecture 1: Week1

1.1 Definitions

Data can be split into three parts:

- Training set: used to train the model
- HoldOut-Cross Validation set(Dev): used to tune the hyperparameters
- Test set: used to evaluate the performance of the model

Split the data into 70/30; training/test. Big Data depends on the algorithms decides which one does better.

Human errors are $\approx 0\% Optimal aka: Bayes Optimal Error Rate is much higher than human errors.$

High Variance: Memorizing all letters in the book for an exam, without understanding then even with small changes in the test, the model will perform poorly. Cannot generalize.

High Bias: Every four legs animals are dogs

$\left\|w^{[l]}\right\|^2 = \sum_{i=1}^{n^{[l]}} \sum_{j=1}^{n^{[l-1]}} \left(w_{i,j}^{[l]}\right)^2$

1.2 Examples

Example 1.1.