Jan Macdonald
Technische Universität Berlin
macdonald@math.tu-berlin.de

AFG Oberseminar SS'17
July 13, 2017

A Generalization Error Bound For Image Classifiers

Abstract

The goal of statistical learning is to fit a prediction model to a set of training data that were drawn i.i.d. according to an unknown joint probability distribution on the space of inputs and outputs. The trained model is then used to estimate the conditional probability distribution on the space of outputs given some previously unseen input data.

A major question is how well such a trained model generalizes to new data that was not used during the training. Or in other words, how well can we control the error that was made by approximating the underlying unknown distribution by using only a finite sample of training data. And how does this error behave when we increase the number of samples?

The answers to these questions depend heavily on the capacity of the prediction model. Here we consider the case of a multi-class classification problem with multinomial logistic regression classifier model and derive a bound on the generalization error using covering numbers.

As an example we study the problem of digit recognition and present results for the MNIST database of handwritten digits.

What digits can be seen?

And how certain can we be about it?