泛化理论

序章 泛化理论简介

§0.1.1 序言

@ 滕佳烨

Generalization

Measuring how model performs on *unseen* data.

Generally, our ultimate training goal is to minimize the **population loss**.
However, we only have finite training samples during the training process.
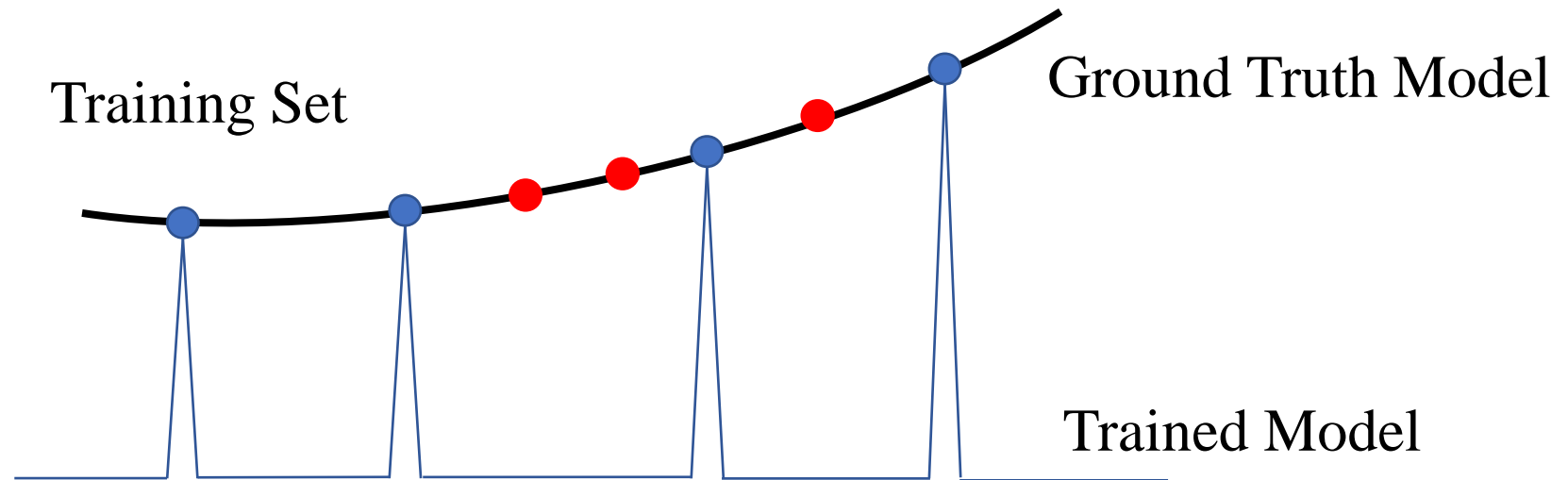We can only obtain and minimize the training loss (ERM).

Is the population loss minimized when the training loss is minimized?
Not always true.

Our Goal	GAP	Our Approach
Minimize population loss	←——————→	Minimize training loss

Generalization

Consider a failure example, where the model interpolate all training samples:

The training error is zero while the test error is large!
The model only learns/memorizes the information on dataset (small training error),
but cannot generalize to other data (large test error).

Generalization

So, when does the model generalize well?

Some intuition:
(a) Fitting noise may hurt generalization (overfitting)
(b) The optimization algorithm plays important roles
(c) Restricting the function class may help improve generalization
(d) More training samples should leads to better generalization

Schedule

C0: introduction to generalization
C1: Traditional Statistics
C2: Uniform Convergence
C3: Stability-based Bound
C4: PAC-Bayes and Information-based bound
C5: Implicit Bias
C6: Benign Overfitting
C7: Other approaches

Some knowledge about basic machine learning and probability theory may be useful!
(but not required)

Contact me…

I am planning to establish a group focusing on generalization topics.
In this group, we will
(a) Meet new friends (most importantly!)
(b) Discuss with each other on some interesting topics
(c) Read papers on Arxiv every day (hopefully…)
(d) Share elegant and novel papers with each other
(e) If possible, I will invite some friends to give talks on generalization every week (maybe not only generalization)

I am not sure I can finally establish such a group. If you are interested, always feel free to contact me! My WeChat ID is *adoutengjiaye*. (or, scan the QR code below)

Take-away messages

(a) Generalization plays an important role in machine learning
(b) It is still a mystery why neural networks generalize well in practice
(c) Contact me and join us for the WeChat Group!

All the slides will be available soon.

@ 滕佳烨

Thanks!