Valiant's Theory

The PAC Model

Theory of Machine Learning – Spring 22

January 13, 2022

Last class

Logistics

- **Course webpage:** Canvas, can find lecture schedule, slides, scribe template, . . .
- TA: Chris Harker
- Scribe for today?

ML until the 1980s

- Many informal notions of learning: Rosenblatt and the “perceptron” algorithm, neural networks, . . .
- Limitations of perceptrons
- No formal theory to reason about, no clear definitions

Question: Can we formally define “learning”?

Theory of the Learnable

- Leslie Valiant 1983 – Theory of the Learnable (CACM)
- Drawing the boundaries of learnability – how to define it? what is possible?
- Really a theory of *supervised* learning
- I.e., deals with classification or *prediction* problems

Theory of the Learnable

- Input: “features” of input
- Hypothesis/model: function from input to prediction/label
- **Definition of a learning algorithm.** an algorithm that can find a *good* hypothesis without explictly being told what it is!

Most natural way. Give examples of inputs and outputs

Good hypothesis?

- Must do well on given inputs (hopefully perfectly)
- Must also do well on “unseen” inputs
- **How to formalize this?**

Valiant’s key assumption. Assume an “input distribution” (unknown to the learner)

Good hypothesis

assuming there is a *true* label for each input

Risk minimization is the goal

Definition of learnability

Complexity of ground truth label

Importance of *hypothesis class*

Learnability with finite hypothesis classes