Why

We want to talk about an inductor's error under a probabilistic supervised data model.

Definition

Let $(X, \mathcal{X}, \mu)$ and $f : X \to Y$ be a probabilistic supervised data model. Let $\mathcal{H} \subset (X \to Y)$ be a hypothesis class.

Let $\varepsilon, \delta \in (0, 1)$. An inductor $A : (X \times Y)^n \to \mathcal{H}$ is (ε, δ)-*probably approximately correct* for μ and f if

$$\mu^n \left[\operatorname*{error}_{\mu, f} \left(A((x_i, f(x_i))_{i=1}^n) \right) \leq \varepsilon \right] \geq 1 - \delta.$$

To show that an algorithm is probably approximately correct, we bound by δ the probability of the inductor producing a predictor which exceeds ε error.

A hypothesis class $\mathcal{H}$ is *probably approximately correct learnable* (or *PAC learnable*) if (a) for every underlying measure μ, labeling function $f : X \to Y$, error $1 - \varepsilon$ and confidence $1 - \delta$, (b) there exists $m_0 \in \mathbf{N}$ and a sequence of inductors $(A^m : (X \times Y)^m \to \mathcal{H})_{m=m_0}^\infty$ so that (c) A^m is (ε, δ)-probably approximately correct for all $m \geq m_0$. If the inductors are subfamilies of a family of inductors $(A^n)_{n \in \mathbf{N}}$ then we say that this family of inductors *PAC learns* $\mathcal{H}$.

We interpret these conditions as follows. No matter the underlying distribution and labeling function, given an accuracy and confidence, we can specify the number of samples required to make the algorithm "succeed." By which we mean that ("with high probability") the predictor's error will be "small."

Sample complexity

If a family of inductors $\mathcal{A} = (A_n)_n$ PAC learns $\mathcal{H}$, then condition (b) above is equivalent to the existence function mapping (ε, δ) to $\mathbf{N}$. Let

$m_0^\star$ be the smallest integer such that condition (b) holds. We call the function $(\varepsilon, \delta) \mapsto m_0^\star(\varepsilon, \delta)$ the *sample complexity* of the family $\mathcal{A}$ on the hypothesis class $\mathcal{H}$.

Other terminology

For (what seem to be) historical reasons, some authors refer to the definition above as *agnostic PAC learning* since it drops a common condition (developed further in these sheets) relating the hypothesis class to the correct labeling function f.

Probably Approximately Correct Inductors

Supervised Probabilistic Data Models

Independent and Identically Distributed Random Variables

Predictors

Independent Random Variables

Approximators

Independent Sigma Algebras

Random Variables

Similarity Functions

Probability Measures

Measurable Functions

Topological Sigma Algebra

Topologies

Independent Events

Measures

Metrics

Distance

Conditional Event Probabilities

Sigma Algebras

Extended Real Numbers

Space Distance

Absolute Value

Real Series

Real Limits

Plane Distance

Event Probabilities

Interval Length

Real Functions

Outcome Probabilities

Real Space

Intervals

Real Plane

Real Line

Real Optimizers

Integral Line

Real Order

Real Sequence

Size of Direct Product

Real Summation

Subset Algebras

Geometry

Rational Order

Real Numbers

Number of Set Products

Sequences

Natural Summation

Cardinality

Rational Numbers

Number of Disjoint Unions

Lists

Finite Set Examples

Optimizers

Integer Arithmetic

Natural Fractions

Direct Products

Set Numbers

Integer Products

Natural Multiplicative Identity

Natural Additive Identity

Finite Sets

Identity Elements

Natural Arithmetic

Equivalent Sets

Family Operations

Uncertain Events

Comparisons

Integer Order

Integer Sum

Natural Powers

Family Unions and Intersections

Function Inverses

Operations

Orders

Integer Numbers

Families

Function Composites

Function Images

Converse Relations

Natural Order

Equivalence Relations

Natural Sums

Functions

Relations

Recursion Theorem

Ordered Pair Projections

Peano Axioms

Uncertain Outcomes

Natural Induction

Set Products

Subset Systems

Natural Numbers

Generalized Set Dualities

Set Powers

Ordered Pairs

Partitions

Successor Sets

Intersection of Empty Set

Set Unions and Intersections

Unordered Triples

Set Symmetric Difference

Set Dualities

Set Intersections

Pair Unions

Set Complements

Pair Intersections

Set Unions

Set Differences

Empty Set

Unordered Pairs

Set Specification

Set Equality

Set Inclusion

Deductions

Quantified Statements

Logical Statements

Statements

Sets

Identities

Names

Letters

Objects

5