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 Induction

Supervised Probabilistic Data Models

Independent and Identically Distributed Random Variables

Independent Random Variables

Predictors

Independent Sigma Algebras

Random Variables

Dependent Events

Probability Measures

Measurable Functions

Topological Sigma Algebra

Approximation

Conditional Event Probabilities

Measures

Topologies

Similarity Functions

Event Probabilities

Sigma Algebras

Metrics

Real Series

Extended Real Numbers

Distance

Real Limits

Space Distance

Absolute Value

Plane Distance

Probability Distributions

Interval Length

Real Space

Intervals

Real Plane

Real Sequences

Real Line

Subset Algebras

Real Summation

Sequences

Integral Line

Real Optimizers

Real Order

Cardinality

Real Numbers

Rational Order

Lists

Geometry

Rational Numbers

Set Numbers

Optimizers

Integer Arithmetic

Finite Sets

Direct Products

Natural Summation

Natural Additive Identity

Natural Multiplicative Identity

Integer Products

Equivalent Sets

Family Unions and Intersections

Family Operations

Identity Elements

Arithmetic

Integer Sums

Function Inverses

Integer Order

Comparisons

Families

Operations

Natural Powers

Function Images

Function Composites

Integer Numbers

Order

Converse Relations

Natural Products

Equivalence Relations

Functions

Natural Order

Natural Sums

Relations

Recursion Theorem

Ordered Pair Projections

Subset Systems

Peano Axioms

Set Products

Uncertain Outcomes

Ordered Pairs

Set Powers

Natural Induction

Unordered Triples

Natural Numbers

Generalized Set Dualities

Set Symmetric Differences

Successor Sets

Set Dualities

Intersection of Empty Set

Partitions

Set Unions and Intersections

Set Complements

Pair Unions

Set Intersection

Set Differences

Set Unions

Pair Intersection

Unordered Pairs

Empty Set

Set Specification

Set Equality

Set Inclusion

Deduction

Quantified Statements

Logical Statements

Statements

Identities

Sets

Names

Letters

Objects