



Why

We discuss inductors that produce relations consistent with their given datasets.

Definition

Let $(x_1, y_1), \dots, (x_n, y_n)$ be a dataset in $X \times Y$. Let \mathcal{R} be the set of all relations on $X \times Y$.

A *consistent inductor* $\{G_n : (X \times Y)^n \rightarrow \mathcal{R}\}_n$ is one for which, for all $n \in \mathbf{N}$, for all $D_n \in (X \times Y)^n$, D is consistent with $G_n(D_n)$. In other words, a consistent inductor always produces a relation with which the dataset is consistent.

The interpretation follows. Fix a relation R^* . And let every dataset “shown” to the algorithm G_n be constructed by selecting elements from R^* . In other words, every dataset is a sequence in R^* . In this case, a dataset $D_n \in (X \times Y)^n$ is always consistent with R^* and so a consistent inductor will never “eliminate” R^* . In other words, the inductor, in order to be consistent “must eliminate” every inconsistent relation.

We may “hope” to give the algorithm a “large and diverse” dataset, so that several of the elements of R^* are included. In this case, the algorithm can “eliminate” many smaller relations in \mathcal{R} which did not include records in the dataset.

Functionally consistent

The rub is that any dataset is consistent with the complete relation $X \times Y$. So we can often consider a set $\mathcal{H} \subset \mathcal{R}$ of relations. It is common to call this a *hypothesis class*, especially for the case in which it consists of functional relations.

