



## Why

Let  $X = \{a, b\}$  and  $Y = \{0, 1\}$ . The dataset  $((a, 0))$  is consistent, but it is not functionally complete. On the other hand, the dataset  $((a, 0), (b, 0), (a, 0), (a, 0), (a, 0), (a, 1))$  is complete but it is not *functionally* consistent.

In general, if  $y_i \neq y_j$  for some  $i$  and  $j$  where  $x_i = x_j$ , then the dataset is not functionally consistent. In the preceding example, both  $(a, 0)$  and  $(a, 1)$  appear.

If we emphasize the “predictive” aspect of a functional inductor, we interpret the input as an object we “see before” the output. And so treat  $y \in Y$  as an uncertain outcome which is the element associated to  $x \in X$ .

In this case, we may use the language of probability to discuss this uncertain outcome. If, for example,  $Y$  is finite, we can associate a distribution with each input  $x \in X$ .

## Definition

Let  $(X, \mathcal{X})$  and  $(Y, \mathcal{Y})$  be measurable spaces.

A *probabilistic functional inductor* (for a dataset of size  $n$  in  $X \times Y$ ) is a function mapping a dataset in  $(X \times Y)^n$  to a family of measures on  $(Y, \mathcal{Y})$ , indexed by  $X$ . We call a function from inputs to output measures a *probabilistic predictor*. We call the distribution a *probabilistic prediction*.

**Notation**

Let  $\mathcal{M}(Y, \mathcal{Y})$  be the set of measures on  $Y$ . Let  $D$  be a dataset in  $(X \times Y)^n$ . Let  $g : X \rightarrow \mathcal{M}(Y, \mathcal{Y})$  a probabilistic predictor. Let  $G_n(X \times Y)^n \rightarrow (X \rightarrow \mathcal{M}(Y, \mathcal{Y}))$  be a predictive probabilistic inductor. Then  $G_n(D)$  is a family of measures  $\{g_x : \mathcal{Y} \rightarrow [0, 1]\}_{x \in X}$ .

