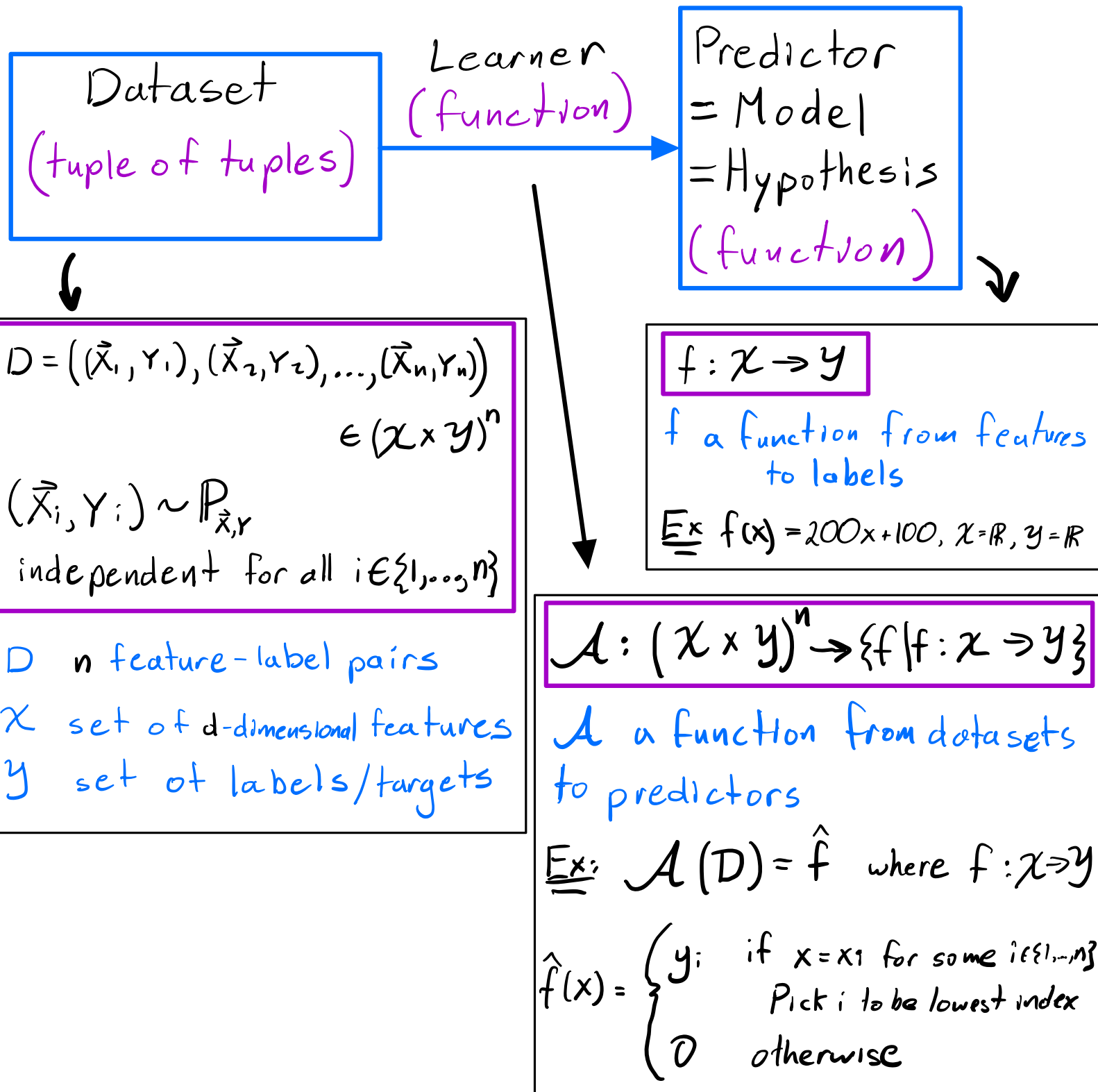


Motivation

Supervised Learning: Learning from a randomly sampled batch of labeled data



Setting:

We are given a random dataset of size n

$$D = ((\vec{X}_1, Y), \dots, (\vec{X}_n, Y_n)) \in (\mathcal{X} \times \mathcal{Y})^n$$

where $(\vec{X}_i, Y_i) \sim P_{\vec{X}, Y}$ are independent for all $i \in \{1, \dots, n\}$

$\vec{X}_i$ feature vector

Y_i label or target

We will always assume the features are vectors.

Ex (of features and labels/targets):

$\vec{X}_i \in \mathbb{R}^3$ # of rooms, # of floors, age of a house

$Y_i \in \mathbb{R}$ price

$\vec{X}_i \in \mathbb{R}^2$ amount of chemical 1, amount of chemical 2
in a wine

$Y_i \in \{0, 1\}$ type of wine

$\vec{X}_i \in \mathbb{R}^{400}$ pixel value of a $20 \times 20 = 400$ pixel image

$Y_i \in \{\text{cat}, \text{dog}, \text{bird}\}$ type of animal

What is a feature and what is a label is a design choice. Usually a feature is info that is easy to gather. And the label is hard, which is why you want to predict it