1. Introduction

Example:

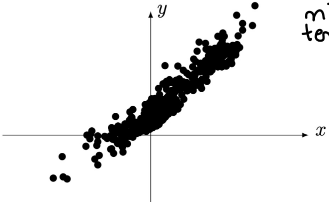
Use ML to learn a predictor for the maximum daytime temperature for a specific day's minimum temperature observed in the morning.

1. Download recordings of min & max daytime temperatures for the most recent days and denote the dataset by:

$$D = \{z^{(1)}, \dots, z^{(m)}\}$$

where $z^{(i)} = (x^{(i)}, y^{(i)}) \; \forall i \in \{1, \dots, m\}$

min temp, max temp, previous day for which $x^{(i)}$ & $y^{(i)}$ have been recorded.



2. ML will learn a hypothesis $h(x)$

$$h(x) = \hat{y} \leftarrow \text{output prediction/ hypothesis / approximation}$$

input min temp

hypothesis map:
- reads in data points (low level properties) and delivers a prediction
- ML methods learn a hypothesis map from a typically large set of candidate maps
  - ↳ candidate maps are hypothesis space or model underlying an ML method

Based on visualization of $(x^{(i)}, y^{(i)})$ we can approximate $[y \sim w_1 x + w_0]$ with weights $w_1 \in \mathbb{R}^+$ & $w_0 \in \mathbb{R}$.

We will restrict the ML method to only consider linear maps

$h(x) := w_1 x + w_0 \quad w_1 \in \mathbb{R}^+ \quad w_0 \in \mathbb{R}$

$\Rightarrow$ monotonically increasing

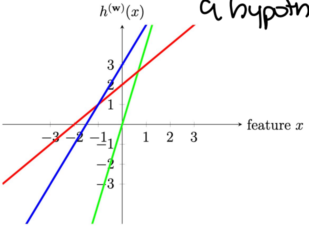
↳ this is a hypothesis space

↳ parameterized by $\vec{w} = \begin{bmatrix} w_1 \\ w_0 \end{bmatrix}$

indication: $h^{(w)}(x) = w_1 x + w_0$

3. Choose between a space of hypothesis to try and find the best one.

loss functions: quantify the quality of a hypothesis map.



} 3 hypothesis maps

- Starting from one initial guess, ML method will continually improve based on new observed data

1.2 Flavours in Machine Learning

- Features: properties that we measure or compute easily in an automated fashion.
- Labels. properties that cannot be measured easily

1.1.1 Supervised Learning

- using training sets w/ labeled data points for which we know the correct label values

1.1.2 Unsupervised Learning

- using training sets w/ labeled data points for which we don't know the correct label values

1.1.3 Unsupervised Learning

- loss function to evaluate and compare different hypotheses
  - ↳ assigns a non-neg loss val to a pair of data point and a hypothesis, out of large hypothesis space