

CS 474/574 Machine Learning

1. HW1

Joshua Slagle
Dept. of Computer Science
Iowa State University
Ames, IA, USA

February 11, 2021

Warm Up

- ▶ $F = m \cdot a$
- ▶ $E = m \cdot c^2$
- ▶ $\mathcal{O}(n \cdot \log(n))$

1. Supervised
2. Unsupervised
3. Reinforcement

1x5

Representation of x

- ▶ x is usually not a simple (vector of) number(s). How to tell it to a computer?
- ▶ Example: bananas vs. apples
- ▶ **Feature engineering**: manually craft functions to **extract** features from raw data, e.g., SIFT, bag-of-words.
- ▶ Automated feature extraction in deep learning: E.g., filters in CNNs.
- ▶ If x involves categorical values (e.g., gender), there are usually two approaches: **One-hot encoding** and **embedding** (in DL context, to be discussed later).

Supervised ML

- ▶ Given many pairs of inputs and outputs:
 $\{(\mathbf{X}_1, \mathbf{y}_1), (\mathbf{X}_2, \mathbf{y}_2), \dots, (\mathbf{X}_N, \mathbf{y}_N)\},$
- ▶ that underline a “black-box” function $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ such that $\forall i \in [1..n], f(\mathbf{X}_i) = \mathbf{y}_i,$
- ▶ construct a function \hat{f} that approximates the function f .
- ▶ “approximate”: usually $\min ||\hat{f}(x) - f(x)||^p$ where p is usually 1 or 2. See ℓ_p -norm .
- ▶ The process of finding the approximation function \hat{f} is called **training** or **learning**.
- ▶ \hat{f} is called a **model** or an **estimator**.
- ▶ \mathbf{X}_i : an **input** (especially when raw data is used as the input) or **feature vector** (if using feature engineering).
- ▶ \mathbf{y}_i , often $\in \mathbb{R}^1$ a **label** (in classification) or **target** (used more generally and lately).
- ▶ Classification vs. Regression: When y is continuous or discrete. In modern DL context, such division is usually no mentioned, especially in generative tasks.