CS 181AI
Lecture 3

Training (Part 1)

Arthi Padmanabhan

Jan 25, 2023

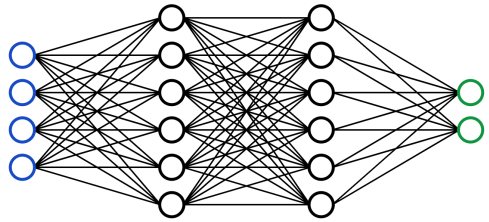
Logistics

- Reminder: Assignment 1 is due next **Monday, Jan 30**
- Reading group assignments were sent – let me know if you didn’t receive an email

Today's Plan

- Training a neural network (how the **algorithm** works)
- Get set up with training on the MNIST dataset

Neural Networks

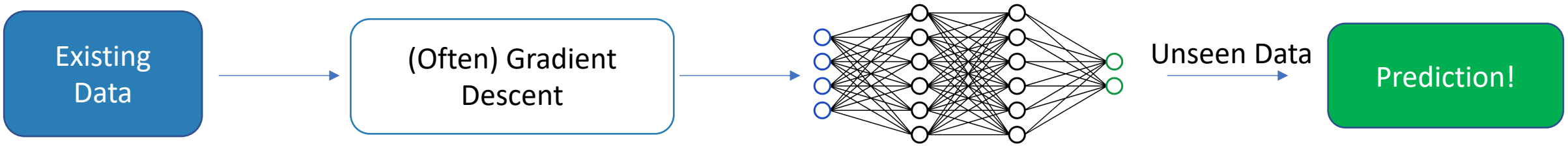


Unseen Data →

Prediction!

- Process of using data to create the model: called training
- Today, we'll look inside the black box!

Neural Networks



- Process of using data to create the model: called training
- Today, we'll look inside the black box!

Sample Model

- Total number of weights and biases = parameters
- Parameters are “tunable” – can be changed during training

Training Process

- We got some garbage result because all weights and biases were initialized randomly

0.01
0.3
0.1
0.09
0.17
0.14
0.05
0.03
0.09
0.01

Training Process

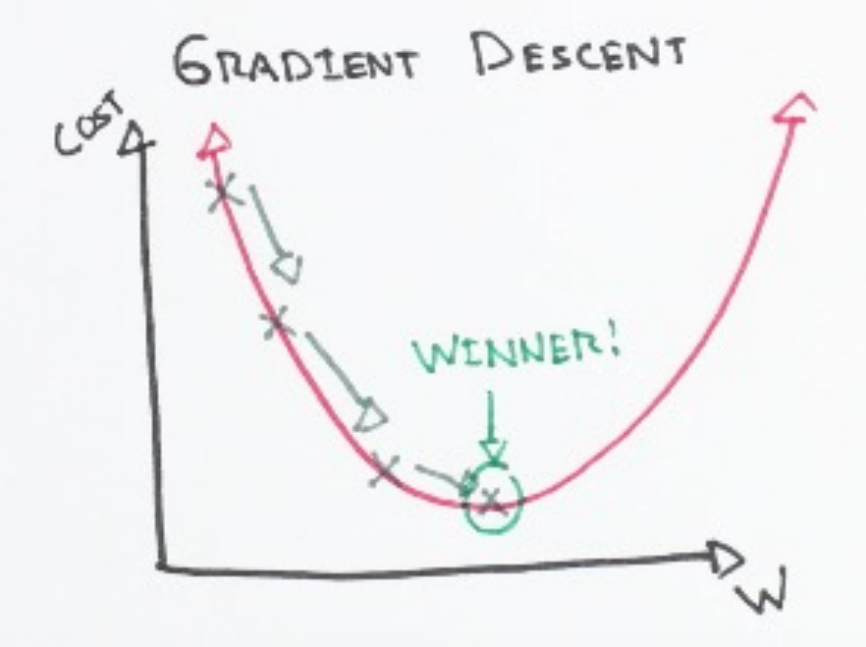
- We got some garbage result because all weights and biases were initialized randomly

0.01
0.3
0.1
0.09
0.17
0.14
0.05
0.03
0.09
0.01

0
0
0
1
0
0
0
0
0
0

Training Process

- There is some set of weights that will lead to the lowest cost
- There is a way to compute the direction to step in this n-dimensional space to get a lower cost
- How big to step depends on size of slope



Training Process

- We compute direction and steepness and might get something that looks like the right vector

weights

0.04
0.3
1.4
…
…
…
…
…
…
…
…
…

0.31
0.03
-1.25
…
…
…
…
…
…
…
…
…

w_0 should increase somewhat
w_1 should increase a little
w_2 should decrease a lot