20211201

Wednesday, December 1, 2021 4:26 PM

- Today:
  - Review last time
  - Neural Network Workshop

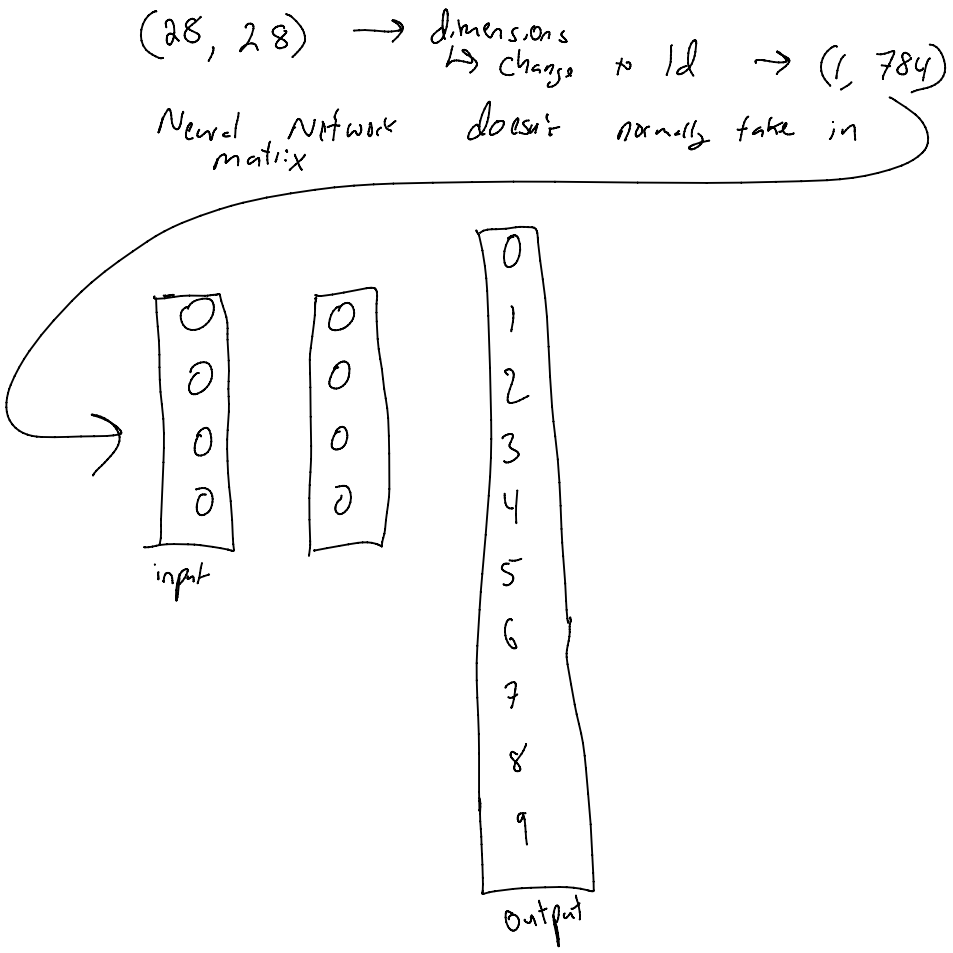
Why Data Privacy Matters:

- TLDR online
- There are ways to get involved with his research
  - How to incentivize for companies to be privacy conscious
  - Tools for users to regain control of their data
    - Today users have no choice or power
      - No bargaining power or a seat at the table
- Differential Privacy
  - A Mathematical definition of privacy
    - This definition is rising in popularity currently
  - How it works
    - Decide randomly whether the true or a random value is stored
    - You don't have data about any particular individual, but you still have global statistical estimates
  - Missing in fields:
    - Easy examples
  - How to maintain relationship between variables?
  - Technical limitations
    - You can tune the coin flip
      - More or less privacy

        - Lower probability of heads the more data you need
            - Today, there normally is a lot of data
    - A company can say they are using differential privacy, but not tell how the coin (how randomized the data is) is tune
    - Student looks at Apple's privacy
        - The more time data stayed there the more likely you can deanonymize the data
- He's looking for students to be part of this group
    - You are totally competent now, he wants to make this accessible to undergrads. He will help you get up to speed

- Privacy tools for average user
- Adversarial examples
    - Perturb the image and get something that looks totally the same to a human, and yet commonly computer vision struggles with it
        - Want to understand why computers are vulnerable to this
            - You don't want someone to mess with a self-driving car
            - Is this partially because the computer is only paying attention to some pixels?

Now Neural Network Workshop



Multi-class logistic regression

Need to normalize your data
↳ able to learn faster

Output of 1st layer is input of 2nd layer.
↳ for hidden layers it's best to be centered at 0 (like hyperbolic tangent instead of sigmoid)

https://playground.tensorflow.org/

Can try different neurons and different layers

Project:

- Will present IN CLASS
- Will need to send slides ahead of time

- Partners in zoom
- If it goes badly on Monday, then Wednesday will move online
- There is also demo day on spark

Fill out course evaluations as soon as you can

- Any feedback you have he wants to hear it
- He is really committed to making this course awesome
- Try to be as constructive as possible
- Make suggestions even if it sounds crazy
- There is now a data science Master's, data science center
- Stay in touch, he wants to know what people are up to
- Hopefully you can come back in a year for day of a life of a data scientist or guest lecture