Course Plan

Data Privacy

1. April 3 – Data Privacy
2. April 10 – Data Privacy – Algorithms
3. April 15 – Data Privacy - Algorithms, Homework
   - April 15 - Project Ideas submission
4. April 17 – PyDeequ, Projects, + data privacy
5. April 22 – Data Privacy - Algorithms
6. April 24 - Data Privacy - Algorithms

- Finals – April 29, 2025 – 9.30 AM – 12.30 PM

- **April 24 - Data Privacy**
- **May 11 – Presentation – 10.30 AM – Zoom**

Project Ideas

- **Data Quality**

    1) **The High-Interest Credit Card of Technical Debt**

        https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/43146.pdf

    2) **Clean Lab**

    o **https://archive.ics.uci.edu/dataset/124/cmu+face+images**
    o **Kaggle Imagenet dataset which does not use CleanLab**

Project Ideas

- **Data Privacy**

    1) Comparison of K-Annonymity, L-diversity and Differential privacy for a dataset

    2) US Broadband data – Differential Privacy

    https://archive.ics.uci.edu/dataset/126/internet+usage+data