

Midterm Review

CSCI 4360/6360 Data Science II

Non-midterm

- **You only need to do EITHER the midterm OR the final project**
- Final project proposals
 - 1 page, max
 - 1 per team
 - Include team members (2 to 4 total), data science question to be answered, approach to answer it, measure of success
 - **Due Thursday by 11:59pm to Autolab**
- Final projects should be ~2x scope of a homework assignment
 - Kaggle competition
 - Sarcasm detection in Reddit comments
 - Identification of cancer in pathology images
 - Training a cryptocurrency trading bot

Midterm Format

- 7 true/false
 - No explanations required, but you're welcome to provide them
- 4 short answer
 - 2-3 sentences max
- 2 programming questions
- 2 bonus questions

Midterm Timing

- Appears on Autolab at midnight tonight/tomorrow morning
- Download the PDF, answer it at your own pace
- Write answers by hand, on MS Word, in Latex, annotate the PDF
- **Submit a PDF to Autolab by 11:59pm Tuesday** (unless we discussed a different time)

Midterm Q&A

- On Discord
- Regular office hours time 10:00 – 11:00am
- Regular lecture time 2:20 – 3:35pm
- Will answer questions asynchronously whenever possible (but cannot guarantee)

Midterm Topics

- Probability, Statistics, Linear Algebra
- Linear and ensemble models
- Optical flow, dynamic textures, linear dynamical systems
- Evolutionary computation
- Ethics in data science
- Reinforcement and PAC learning
- Graph theory
- Spectral clustering

Q&A

- Questions?

