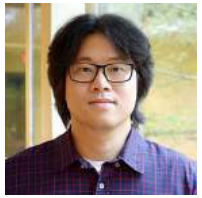
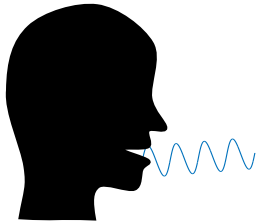
ENGR-E 511
“Machine Learning for Signal Processing”
Lecture 00: Course Logistics

Minje Kim
Department of Intelligent Systems Engineering
Email: minje@indiana.edu
Website: http://minjekim.com





INDIANA UNIVERSITY
SCHOOL OF INFORMATICS,
COMPUTING, AND ENGINEERING

Grade Distribution

- 6 assignments (90%)
- Q&A Community participation (10%)
- No final project

Six Assignments

- Languages allowed to use: Python3
  - Sorry, I can't allow R or MATLAB
  - No Tensorflow, PyTorch, or scikit-learn. Write your own code from scratch whenever possible.
- Late policy
  - Assignments will be released bi-weekly. Deadline is two weeks after the release date
  - Late submissions are accepted without any late penalty
    - If the sum of the late hours throughout the semester < seven days (168 hours)
  - Otherwise, you'll get only 80% of your **total late homework score you earned**
  - If you don't submit a homework, we add 168 hours to the late submission hours automatically
  - **The homework submission system is closed seven days after the deadline**
  - **We don't accept late submissions after the grace period. Don't send your homework to the AI's as an email attachment.**
- You can discuss about it with your friends, but your solution should be created by yourself
  - This is a graduate course, so I take plagiarism seriously
  - We will use Standford MOSS to check the plagiarism, and will report the misconduct to the registrar
- You're not supposed to use off-the-shelf toolbox functions
  - Numpy and scipy functions are usually fine, but no scikit-learn.
    - Okay, there are exceptions, like **scipy.cluster.vq.kmeans**. You're not supposed to use that.

Late Policy Examples

- Student A missed the original HW#1 deadline (2/6) and wasn't able to finish in 7 days after that
  - Student A has no way to submit HW#1 since the system is closed
  - Late penalty will start to apply to the other late homework submissions
- Student B was late for three times, and the sum is 167 hours (6 days + 23 hours)
  - B's good, too. No penalty. Don't send me an email about this.
- Student C was on time for only one assignment
  - C's total late hours is 250 hours for all other assignments
  - C got 18, 15, and 19 for the three late submissions
  - I'll apply late penalty to this, so C gets (18+15+19) X 0.8 = 41.6 for the three late submissions
- Student D got tested positive to Covid19
  - Send me an email

Class Participation

- InScribe
  - **All homework-related questions should be asked in Q&A Community**
  - I will distribute a point to a good answer or a good comment that are endorsed by the instructors
  - To a good question, too, but RARELY

Schedule

Expected to finish by	Module	Topic	Other Milestones
1/11	Module 00, Module 01	•Course Logistics •Probability	
1/13	Module 01	•Probability	
1/18	Module 01	•Linear Algebra	
1/20	Module 01	•Optimization	
1/25	Module 02	•Time-Frequency Transforms	
1/27	Module 03	•Clustering - Part 1	
2/1	Module 03	•Clustering - Part 2	
2/3	Module 03	•Clustering - Part 3	Homework #1 due by **2/6**
2/8	Module 04	•Dimension Reduction - Part 1	
2/10	Module 04	•Dimension Reduction - Part 2	
2/15	Module 05	•Bayesian Classification - Part 1	
2/17	Module 05	•Bayesian Classification - Part 2	Homework #2 due by **2/20**
2/22	Module 06	•Neural Networks - Part 1	
2/24	Module 06	•Neural Networks - Part 2	
3/1	Module 07	•Hidden Markov Models - Part 1	
3/3	Module 07	•Hidden Markov Models - Part 2	Homework #3 due by **3/6**
3/8	Module 08	•Nonlinear Methods - Part 1	
3/10	Module 08	•Nonlinear Methods - Part 2	
3/22	Module 09	•Support Vector Machines - Part 1	
3/24	Module 09	•Support Vector Machines - Part 2	Homework #4 due by **3/27**
3/29	Module 10	•Undirected Graphical Models - Part 1	
3/31	Module 10	•Undirected Graphical Models - Part 2	
4/5	Module 11	•Probabilistic Topic Modeling - Part 1	
4/7	Module 11	•Probabilistic Topic Modeling - Part 2	Homework #5 due by **4/10**
4/12	Module 12	•Adaptive Basis Function Models - Part 1	
4/14	Module 12	•Adaptive Basis Function Models - Part 2	
4/19	Module 13	•Kalman Filtering	
4/21	Module 14	•Array Processing	Homework #6 due by **4/24**
4/26	Module 15	•Advanced Audio Signal Processing Applications - Part 1	
4/28	Module 15	•Advanced Audio Signal Processing Applications - Part 2	

Textbooks

- Textbooks

- (IUCAT) Sergios Theodoridis and Konstantinos Koutroumbas, "Pattern Recognition," https://iucat.iu.edu/catalog/13489528
- **Students are encouraged to read the following books though not required:**
  - (IUCAT) Kevin Murphy, "Machine Learning: a Probabilistic Perspective," https://iucat.iu.edu/catalog/11676896
  - Chrostopher Bishop, "Pattern Recognition and Machine Learning,": https://www.amazon.com/Pattern-Recognition-Learning-Information-Statistics/dp/0387310738/ref=pd_bxgy_14_img_3?ie=UTF8&psc=1&refRID=8FAPG9N7R8GQ9AJ5EYHD
  - Pattern Classification: https://www.amazon.com/Pattern-Classification-Pt-1-Richard-Duda/dp/0471056693
  - Ian Goodfellow, Yoshua Bengio and Aaron Courville, "Deep Learning": http://deeplearningbook.org
    - Only Part 1

Communication

- Office hours and AI duties

- With me (by appointment)
  - Send me an email
  - **I won't accept homework-related questions because I have to encourage to use Q&A Community**
- With the AIs (times are in US Eastern)
  - 3-5pm on Tuesdays at Luddy 4147
    7-9pm on Wednesdays at iu.zoom.us/my/minje
    3-5pm on Thursdays at iu.zoom.us/my/minje
    3-5pm on Saturdays at iu.zoom.us/my/minje
- Each AI takes care of a homework assignment
  - For example, AI#3 will take care of
    - All HW#3 questions on Q&A (marking good answers)
    - Grading HW#3
    - Office hours both for residential and online students
  - But AI#3 will be off otherwise
  - AI#3 knows nothing about HW#2 or HW#4

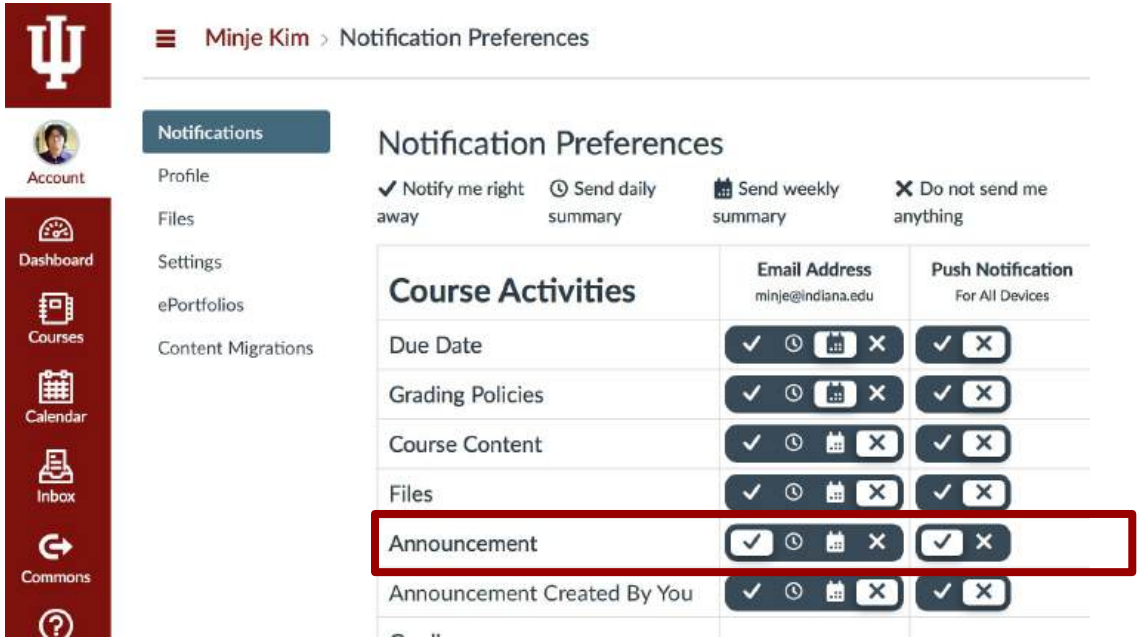
Dates	AI	Homework the AI covers
1/25, 1/26, 1/27, 1/29	Ambarish Gurjar	Homework 1
2/1, 2/2, 2/3, 2/5	Ambarish Gurjar	Homework 1
2/8, 2/9, 2/10, 2/12	Murali Parvataneni	Homework 2
2/15, 2/16, 2/17, 2/19	Murali Parvataneni	Homework 2 (Homework #2 due by **2/20**)
2/22, 2/23, 2/24, 2/26	Baekeun Park	Homework 3
3/1, 3/2, 3/3, 3/5	Baekeun Park	Homework 3 (Homework #3 due by **3/6**)
3/8, 3/9, 3/10, 3/12	Kaitlin Pet	Homework 4
3/15, 3/16, 3/17, 3/19	No Office Hours --- Spring Break	No Office Hours --- Spring Break
3/22, 3/23, 3/24, 3/26	Kaitlin Pet	Homework 4 (Homework #4 due by **3/27**)
3/29, 3/30, 3/31, 4/2	Grace Li	Homework 5
4/5, 4/6, 4/7, 4/9	Grace Li	Homework 5 (Homework #5 due by **4/10**)
4/12, 4/13, 4/14, 4/16	Zhengyi Li	Homework 6
4/19, 4/20, 4/21, 4/23	Zhengyi Li	Homework 6 (Homework #6 due by **4/24**)

Learning Objectives

- Target audience
  - An ambitious graduate student in his/her early years
  - A relaxed senior graduate students
  - A super ambitious undergrads (I generally don't recommend this course to undergrads)
  - Needs to have a good understanding about probability theories, linear algebra, and optimization
- Theory versus practice
  - Both
    - That's what engineering is about IMHO
  - But only for the topics that are important
- After taking this course
  - You can brag that you learned machine learning
    - To a limited amount of people around you
  - You know who to ask questions
  - You know what to do next

Canvas Announcement

- Canvas➔Account➔Notifications➔
  Set to receive emails (and push notifications if you use the app) right away for an announcement

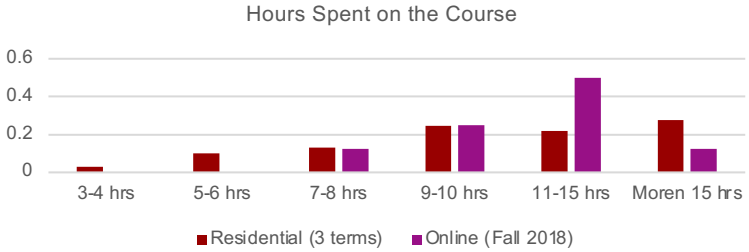


Submission Format

- IPython Notebook + HTML
  - Your notebook should be a comprehensive report, not just a code snippet.
    - Mark-ups are mandatory to answer the homework questions.
    - You need to use LaTeX equations in the markup if you're asked.
    - See an example here: https://nbviewer.jupyter.org/github/jupyter/notebook/blob/master/docs/source/examples/Notebook/Typesetting%20Equations.ipynb
  - Download your notebook as an .html version and submit it as well, so that the AIs can check out the plots and audio
    - Meaning you need to embed an audio player in there if you're asked to submit an audio file
    - Graders should be able to see your results without running the codes (although they will run it if they feel like)

Frequently Asked Questions

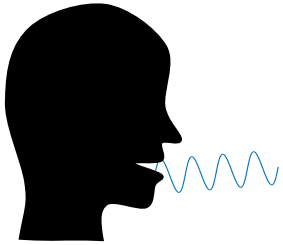
- This is my first machine learning course. Can I take it?
  - Please reconsider. I promise that the course will be fun, but it's quite demanding.
- You speak too slowly
  - I wouldn't be offended if you speed up by 1.5x
- Can I post my source code on Q&A Community to elaborate my question?
  - No, please don't post your code on Q&A Community
- Is there any template code that I can start from?
  - No, you need to start from scratch. It will be fun though.



Frequently Asked Questions

- Can I download the video lectures?
  - No, but you still have access to the video after the semester (until you graduate)
- Will you distribute the solution after the due date?
  - Unfortunately, no. But we will do our best to answer your questions.
- Can I upload my source code to a public Github repo
  - Unfortunately, no. Your code can be used in some plagiarism cases.
- I don't like the quality/resolution of the video. What should I do?
  - I don't recommend smartphones or tablets to watch the video lectures, because the high-resolution options (1080P and 1080P HQ) are only available in a desktop/laptop browser for now. Canvas/Kaltura recommend Chrome. The lectures, especially the equations, look pretty nice in this recommended setup.
- Where is the latest information about COVID19?
  - For international students: https://ois.iu.edu/coronavirus/index.html
  - In general: https://coronavirus.iu.edu

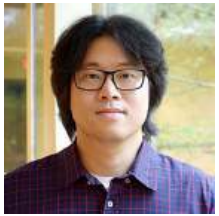
What Is a Signal?



(A) Speech

(B) EEG

(C) Timeline

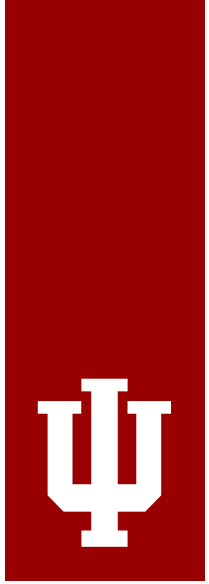


(D) Facial Image

(E) Stock Price

(F) All of them

- People have different definitions…
- My definition:
  “An ordered set of data samples.”



Thank You!