https://ttpoll.com/p/967596

COSC 325: Introduction to Machine Learning

Dr. Hector Santos-Villalobos

Dr. Santos

THE UNIVERSITY OF
TENNESSEE
KNOXVILLE

Lecture 01: Course Overview

THE UNIVERSITY OF TENNESSEE KNOXVILLE

Are you in the right place?

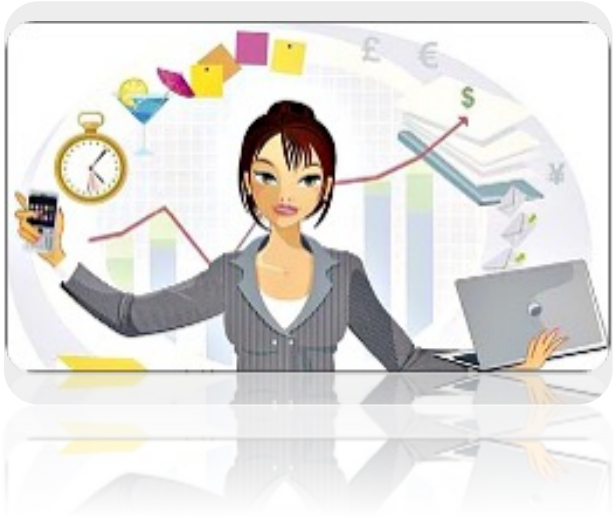
- Course
  - Course: COSC 325, Introduction to Machine Learning
  - Pre-Requisites: ECE 313 or ECE 317 or MATH 323; and MATH 251 or MATH 257 with a grade of C or better.
  - Recommended Background: Python Programming Language, Numpy, SciKit Learn
- Logistics
  - Location: MKB 524
  - Lectures: Tuesday/Thursday: 11:20 am to 12:35 pm
  - Canvas Link: https://utk.instructure.com/courses/206990

https://ttpoll.com/p/967596

Today's Topics

Team Introduction

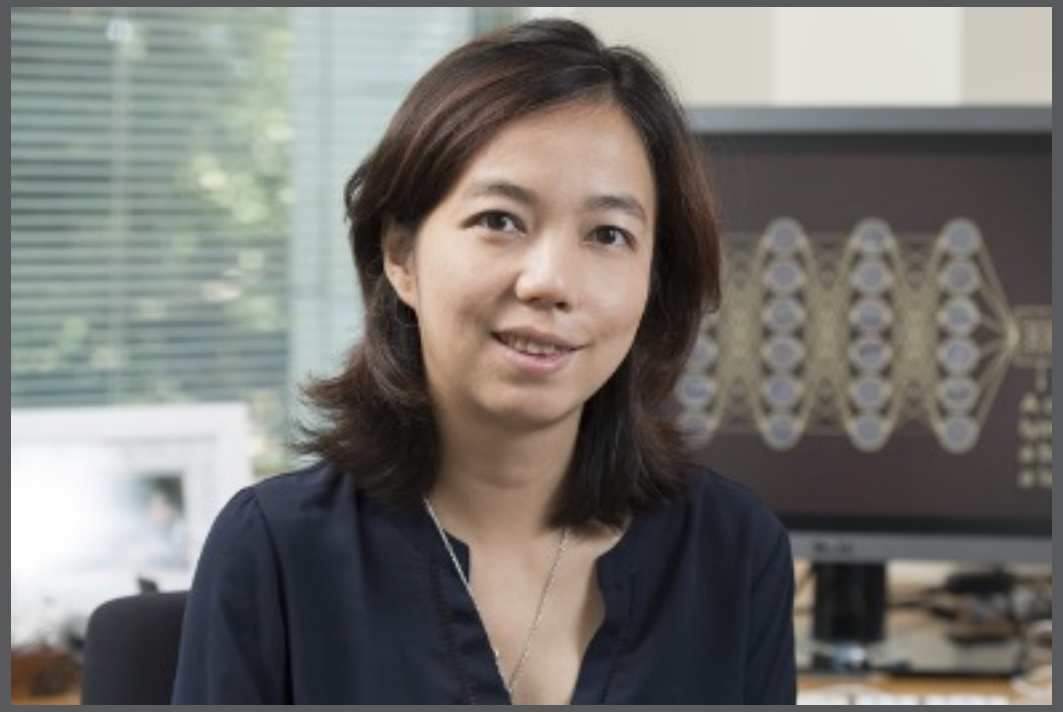
Course Goals, Logistics, and Policies



https://ttpoll.com/p/967596

“Artificial intelligence, deep learning, machine learning — whatever you’re doing if you don’t understand it — learn it. Because otherwise, you’re going to be a dinosaur within three years.” - **Mark Cuban (2020).**

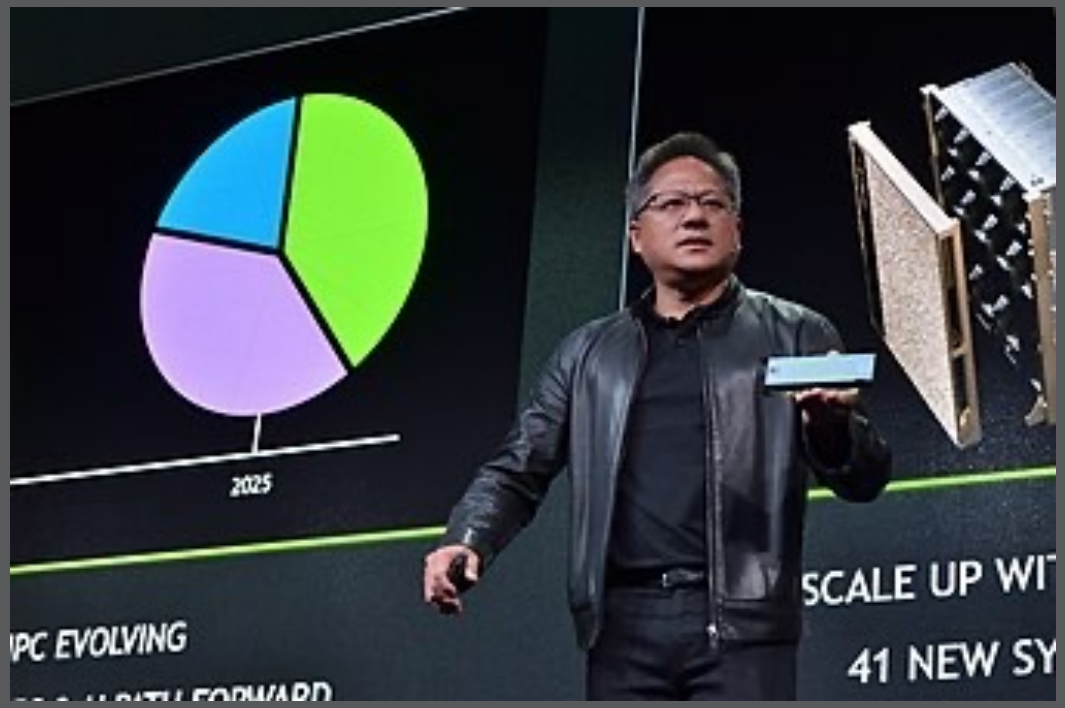
https://ttpoll.com/p/967596



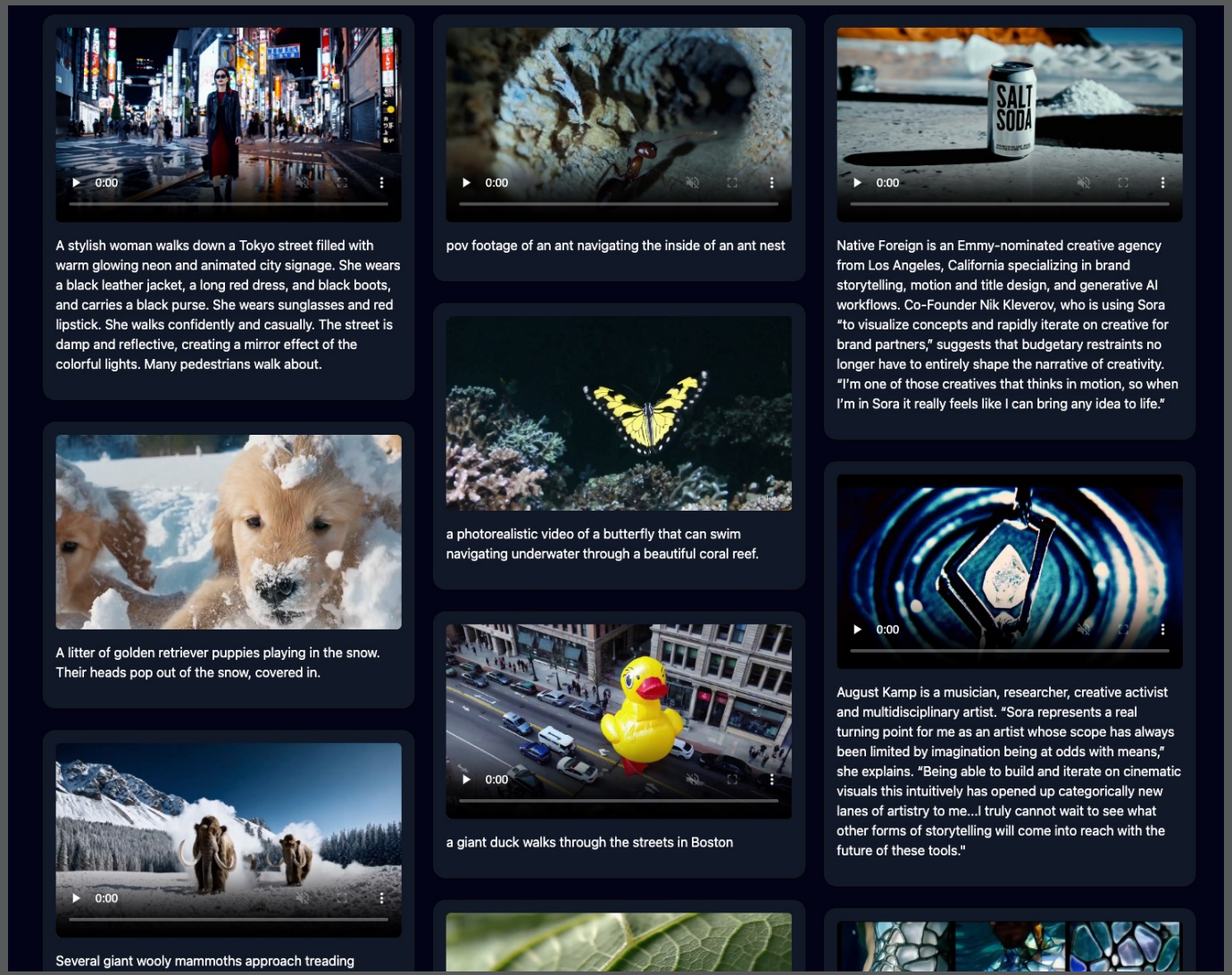
“We will not only use the machines for their intelligence. We will also collaborate with them in ways that we cannot even imagine. **– Fei Fei Li, Director of Stanford’s Artificial Intelligence Lab**.”

https://ttpoll.com/p/967596

"AI will be the most transformative technology of the 21st century. It will affect every industry and aspect of our lives." – **Jensen Huang (2021), CEO at NVIDIA.**



https://ttpoll.com/p/967596

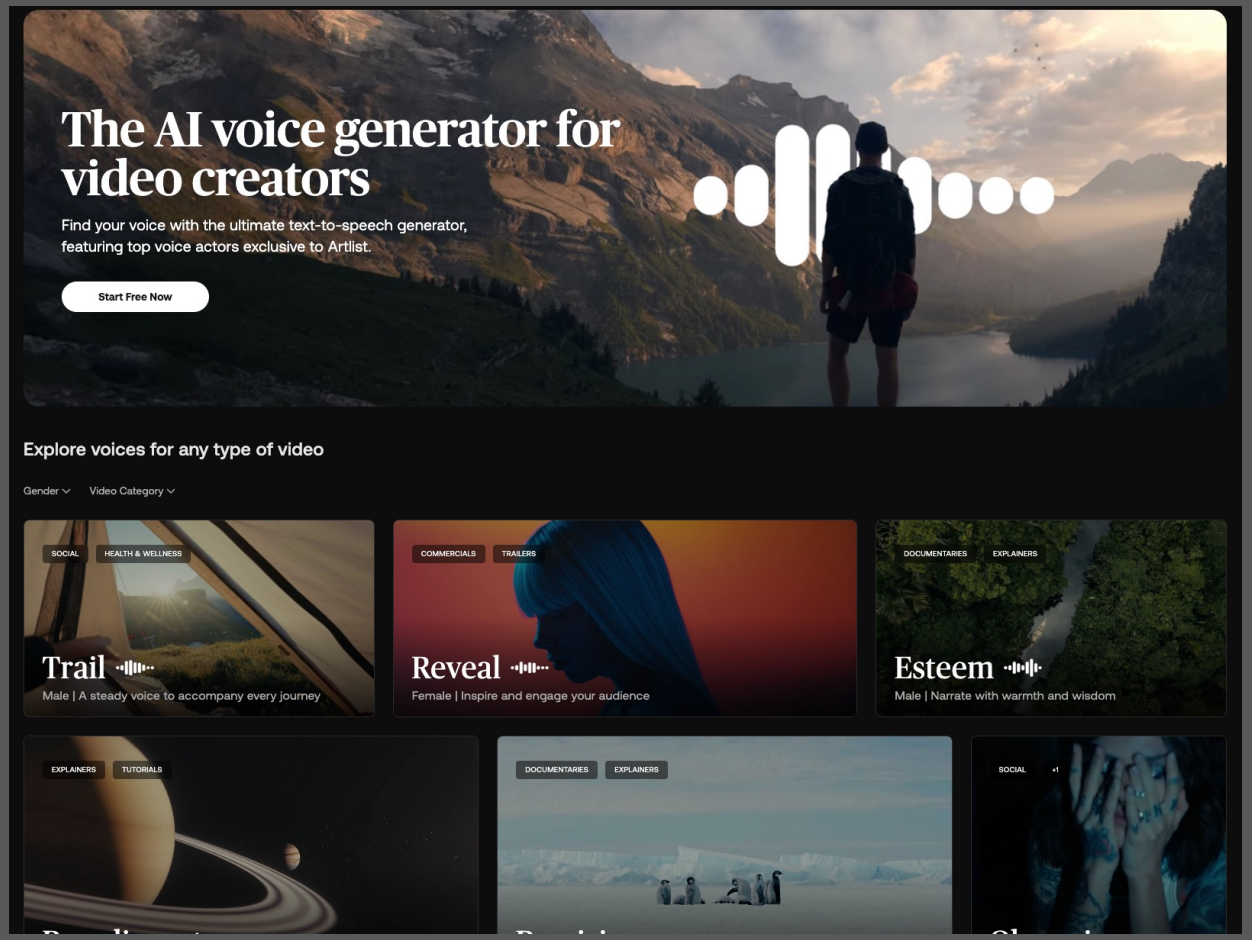


https://sora.aitubo.ai/videos/featured

Sora

https://ttpoll.com/p/967596

Artlist.io



https://artlist.io/voice-over

https://ttpoll.com/p/967596

Course team

- Instructor
  - Dr. Hector Santos
    - Email: hsantosv@utk.edu
    - Office: Min H. Kao Building, Suite 352
    - Office Hours:
      - Thursday: 10:10 am to 11:10 am
    - Telephone: 865-974-3076

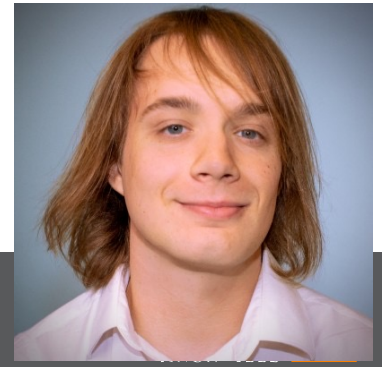
https://ttpoll.com/p/967596

Course team

GTA: Kyle

- Teaching Assistants
  - Kyle Musgrove
    - Email: kmusgro1@vols.utk.edu
    - Office: 350 Min H. Kao Bldg.
    - Office Hours: Monday 2:00 pm - 3:00 pm and Wednesday 10:00 am - 11:00 am
  - Zifan "Fred" Yu
    - Email: zyu20@uthsc.edu
    - Office: 350 Min H. Kao Bldg.
    - Office Hours: Tuesday 10:00 - 11:00 am
  - Mikolaj Jakowski
    - Email: mjakowsk@vols.utk.edu
    - Office: 350 Min H. Kao Bldg.
    - Office Hours: TBD

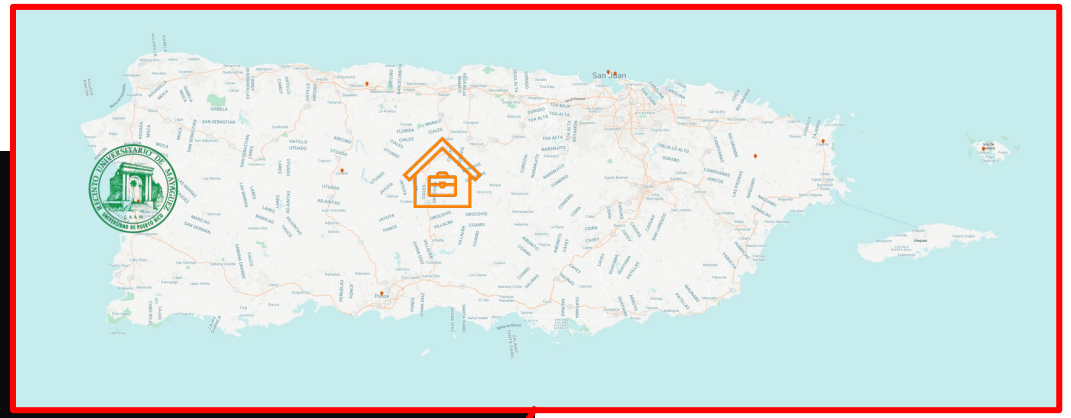
GTA: Fred



UTA: Miko

https://ttpoll.com/p/967596

Who am I?



https://ttpoll.com/p/967596

A little about me…

Assistant Professor

- Applied Science Manager, Amazon Prime Video and Studios
- Group Leader, Oak Ridge National Laboratory
- Ph.D. Purdue University
- Founder of Cyber Identity and Biometrics group @ ORNL

Multi-Modal Analytics, Reasoning, and Computational Imaging Laboratory

North Star:

Advance Self-Supervised Multi-Modal Machine Learning to achieve timely, effective, efficient, and personalized precision health.

https://ttpoll.com/p/967596

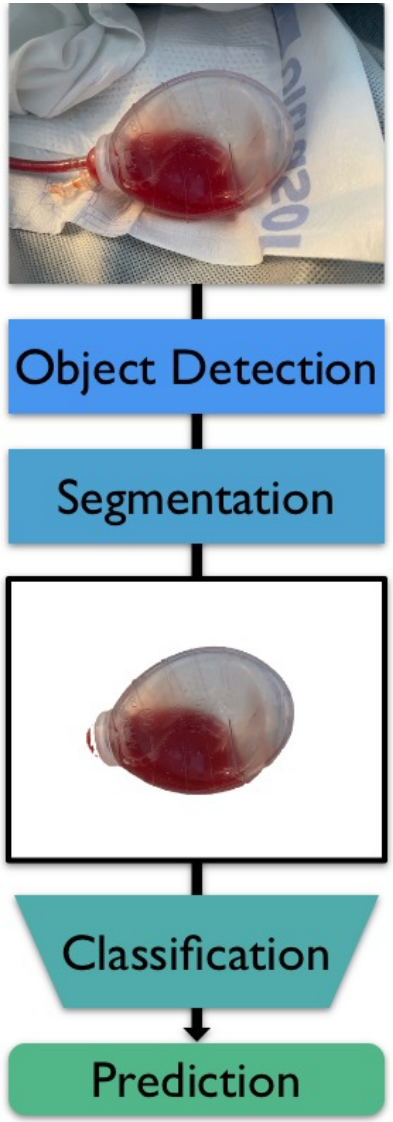
MARCI Active Projects

THE UNIVERSITY OF TENNESSEE KNOXVILLE

https://ttpoll.com/p/967596

Surgical Drain Classification

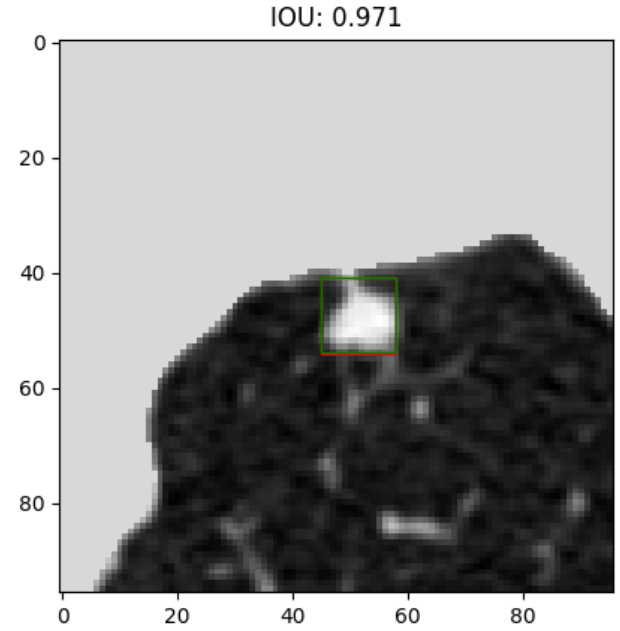
- Collect a novel dataset of thousands of drains.
  - Label each drain with fluid type, fluid turbidity, doctor's diagnosis, diagnosis confidence, etc.
- Pass drains through object detection (YOLO) and segmentation (SAM) ML models to mask them out.
- Train classifier/predictor on masked drains to predict one or multiple labels for the drain images.
- Use the trained predictor in a mobile device on new patients' drains to speed up doctors' workflow.

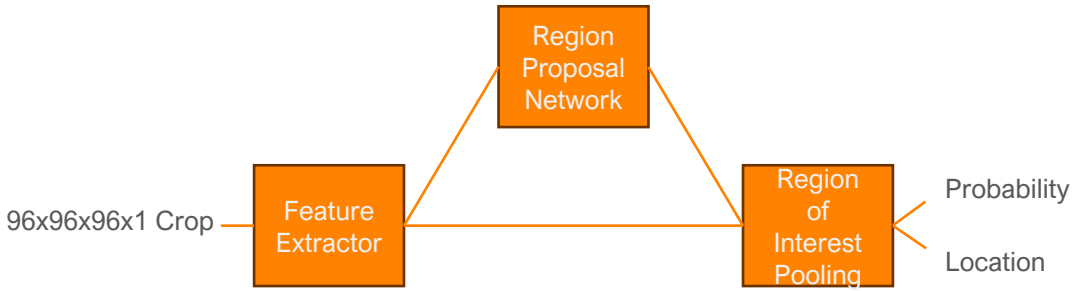


https://ttpoll.com/p/967596

Lung Cancer Detection

- 44% of cases are detected in a late stage witl a 5-year survival rate of 7%.
- Early detection 5-year survival rate of 61%
- Large organ to review.
- **Can we use ML to expedite X-ray CT review?**





FasterCNN Architecture

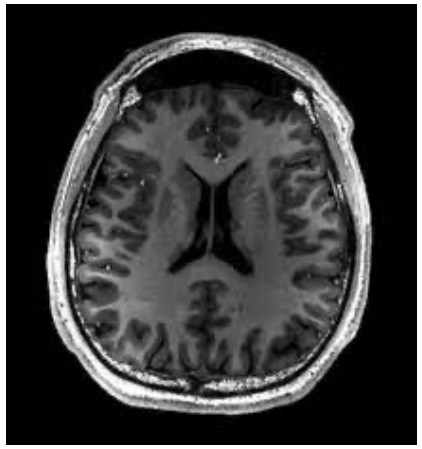
- Object *detection* and *classification* problem.
  - Nodule detection, followed by cancer classification.
- Medical images such as CT scans are 3D volumes, which is essentially a stack of 2D images
- **Developing deep learning 3D object detection methods.**

https://ttpoll.com/p/967596

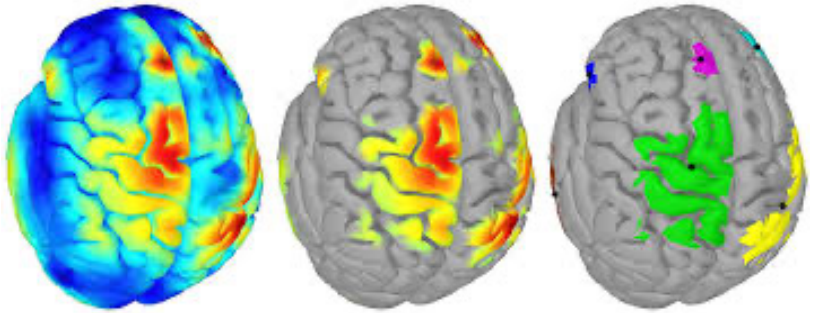
3D Multimodal Fusion

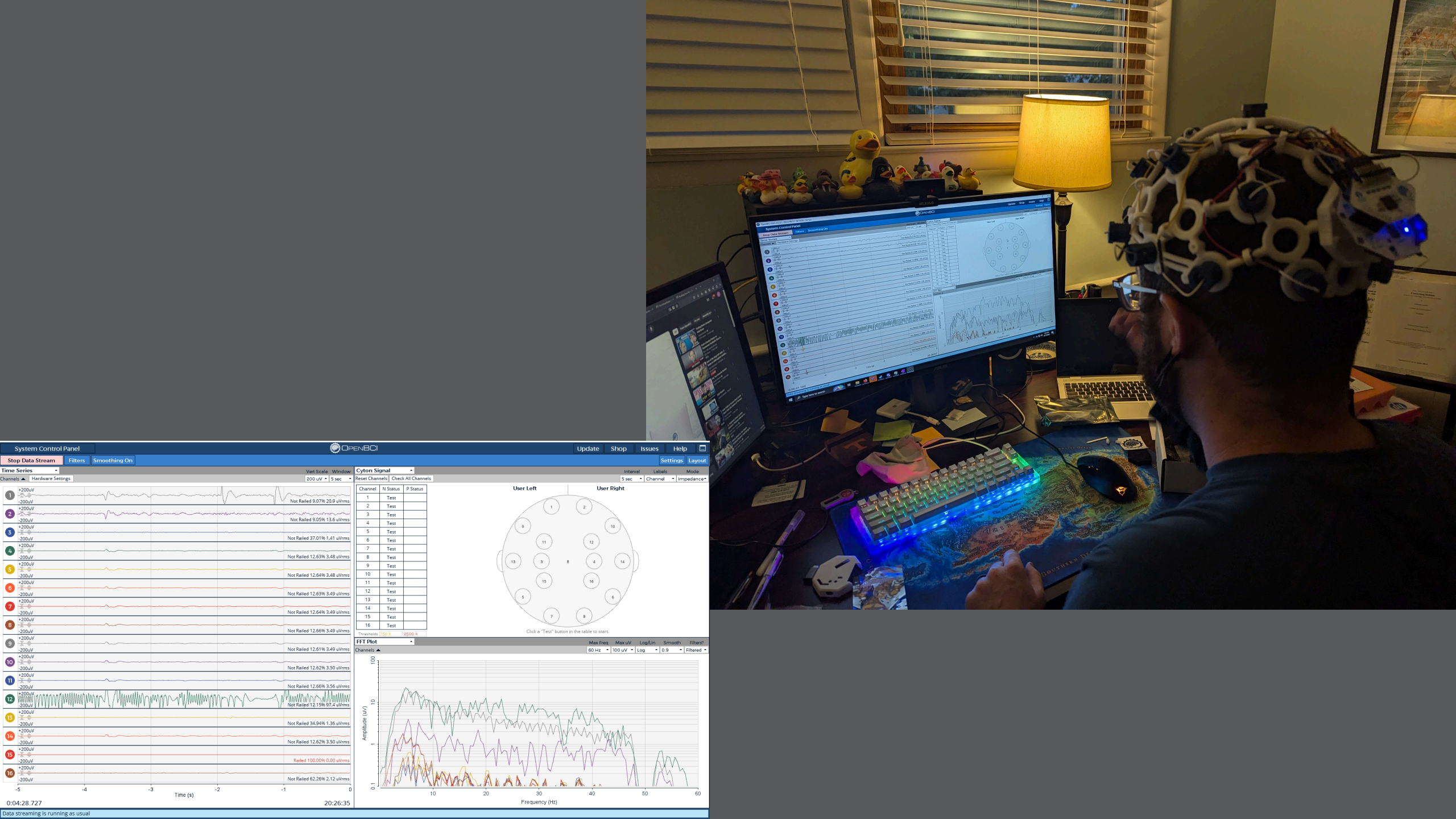
- The Problem
  - Alzheimer's Disease is a degenerative disease that can begin affecting a patient years before symptoms become obvious
- The idea
  - Take a series of 2D EEG signals, transfer them to a 3D space and map them to an MRI
    - Currently done with inverse operations, we want to explore a machine learning solution
  - From this combined mapping can we:
    1. Detect Alzheimer's Disease earlier than current methods using modern AI approaches
    2. Can we detect what signals and portions of the brain are contributing to the diagnosis

+



=




System Control Panel
OpenBCI
Update
Shop
Issues
Help
Stop Data Stream
Filters
Smoothing On
Settings
Layout
Time Series
Cyton Signal
User Left
User Right
Click a "Test" button in the table to start.
FFT Plot
Time (s)
Frequency (Hz)
Amplitude (uV)
0:04:28.727
20:26:35
Data streaming is running as usual

Course Logistics and Policies

THE UNIVERSITY OF TENNESSEE KNOXVILLE

https://ttpoll.com/p/967596

What is not included in this class?

- Technical depth on advanced ML methods
  - Deep learning backpropagation, CNNs, Diffusion, GANs, etc.
  - Data mining, human-computer-interaction, robotics
  - Reinforced learning
  - Biological-inspired artificial intelligence
- Operational deployment of ML models
  - Edge computing
  - HPC computing (Cloud computing)
  - Enterprise deployment

https://ttpoll.com/p/967596

Why take this course?

- There are a ton of machine learning quick-starts out in the world
- There are a ton of libraries that implement machine learning approaches for you
- To ***effectively*** leverage these tools, you need a fundamental understanding of how they work, how they're different, and when they're useful

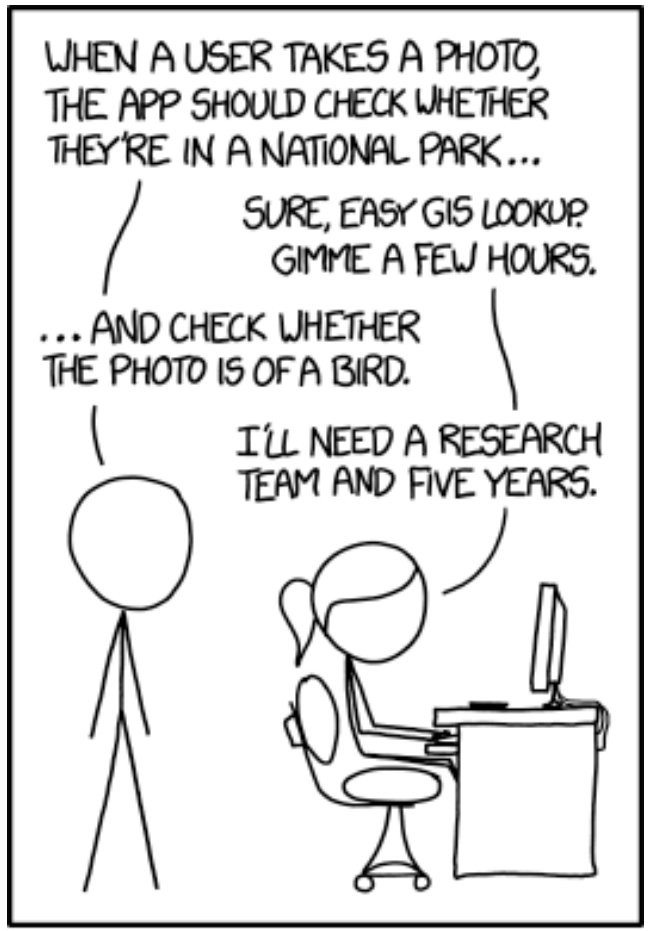
Source: https://xkcd.com/1838/

Slide credit: Dr. Schuman

https://ttpoll.com/p/967596

Will you be using ML in your job?

- Even if you're not implementing and applying machine learning techniques actively in your day-to-day job life:
  - You should be able to tell when machine learning is the appropriate tool for a particular problem and when it is NOT
- There may come a time when you are being asked to apply ML to something where it doesn't make sense to use it
- This class should help prepare you to justify when you shouldn't be using ML for a task as well as when you should



Source: https://xkcd.com/1425/

Slide credit: Dr. Schuman

https://ttpoll.com/p/967596

What do I want you to gain with this class?

- Understand the fundamentals of machine learning (ML).
- Understand how to train, explain, and assess ML models.
- Learn about different ML techniques and the applications they are suitable for.
- Apply ML techniques to real-world applications.
- Build an ML programming portfolio.
- Develop a strong ML foundation for future self-learning and advanced ML courses.

https://ttpoll.com/p/967596

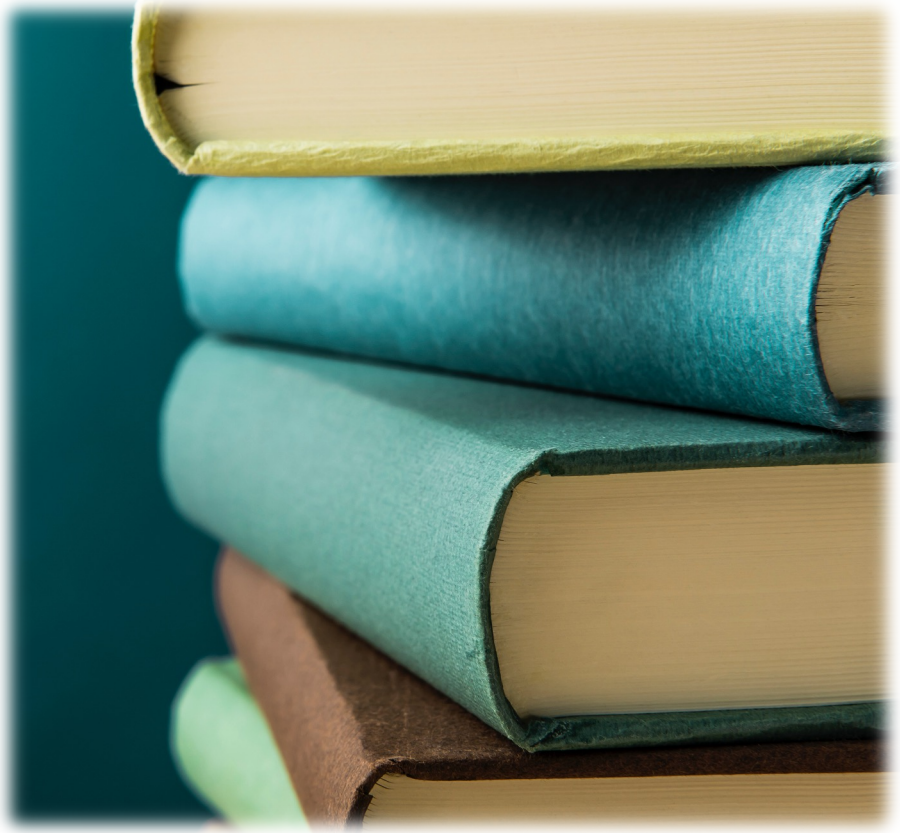
Inspiration for this course

- UTK, Dr. Schuman, COSC425, Introduction to Machine Learning
- Sebastian Raschka, Introduction to Deep Learning
- Coursera, Dr. Andrew Ng, Machine Learning Certification

https://ttpoll.com/p/967596

Recommended Books

- **The Hundred-Page Machine Learning Book,** Andriy Burkov, 2024, https://themlbook.com/wiki/doku.php
- **A Course in Machine Learning,** 2017, H. Daume III, http://ciml.info/
- **Probabilistic Machine Learning: An Introduction**, 2022, Kevin Murphy,
  - https://probml.github.io/pml-book
- **An Introduction to Statistical Learning**, 2023, James et. al.,
  - https://www.statlearning.com
- **Dive Deep Into Deep Learning**, 2023, Aston Zhang,
  - https://d2l.a
- **Understanding Machine Learning: From Theory to Algorithms**, 2014, Shai Shalev-Shwartz and Shai Ben-David,
  - http://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning



https://ttpoll.com/p/967596

Coding/Programming Tools

- Python 3
- Numpy
- Pandas
- SciKit-Learn
- Jupyter Notebooks
- Markdown
- Google Collab
- PyTorch

I assume you have novice to intermediate experience.

https://ttpoll.com/p/967596

Topic Schedule

- Programming Tools — Week 1-2
- How Do Machines Learn? — Week 2-4
  - Nearest Neighbors
  - Linear, Polynomial, Logistic Regression
- Decision Trees — Week 4-5
  - Ensemble Methods, Random Forest, Boosting
- Model Assessment and Improvements — Week 6-7
- Unsupervised Learning — Week 8-9
- Human-in-the-Loop Machine Learning — Week 10
- Advanced ML Techniques — Week 11-14
  - Support Vector Machines, Artificial Neural Networks, Deep Learning, Reinforced Learning

https://ttpoll.com/p/967596

Week	Lecture	Date	Topics	Artifacts
1	1	Tuesday, August 20, 2024	Course Overview	
	2	Thursday, August 22. 2024	Machine Learning Motivation	
2	3	Tuesday, August 27, 2024	Scientific Computing with Python	
	4	Thursday, August 29. 2024	SciKit-Learn and a Nearest Neighbor Classifier	HW#1
3	5	Tuesday, September 03, 2024	Linear Regression and Gradient Descent	
	6	Thursday, September 5. 2024	Polynomial and Logistic Regression	HW#2
4	7	Tuesday, September 10, 2024	Overfitting, Variance, Bias, and Regularization	
	8	Thursday, September 12. 2024	Decision Trees	HW#3
5	9	Tuesday, September 17, 2024	Ensemble Learning	
	10	Thursday, September 19. 2024	Random Forests and Boosting	CP#1
6	11	Tuesday, September 24, 2024	Dimensionality Reduction and Feature Selection	
	12	Thursday, September 26. 2024	Cross Validation and Model Confidence	HW#4
7	13	Tuesday, October 01, 2024	Model Evaluation, Explanation, and Selection	
	-	**Thursday, October 3. 2024**	**Midterm #1**	
8	-	Tuesday, October 08, 2024	FALL BREAK	
	14	Thursday, October 10. 2024	Introduction to Unsupervised Techniques	HW#5
9	15	Tuesday, October 15, 2024	Clustering with K-Means	
	16	Thursday, October 17. 2024	Anomaly Detection with Gaussian Mixtures	CP#2
10	17	Tuesday, October 22, 2024	Human-in-the-Loop Machine Learning 1	
	18	Thursday, October 24. 2024	Human-in-the-Loop Machine Learning 2	HW#6
11	19	Tuesday, October 29, 2024	Support Vector Machine (SVM) 1	
	20	Thursday, October 31. 2024	Support Vector Machine (SVM) 2	
12	-	Tuesday, November 05, 2024	ELECTION DAY	
	21	Thursday, November 7. 2024	Introduction to Artificial Neural Networks	HW#7
13	22	Tuesday, November 12, 2024	Deep Learning	
	23	Thursday, November 14. 2024	Reinforced Learning	
14	24	Tuesday, November 19, 2024	Ethical Artificial Intelligence	CP#3
	-	**Thursday, November 21. 2024**	**Course Project Presentations**	
15	-	**Tuesday, November 26, 2024**	**Course Project Presentations**	CP#4
	-	Thursday, November 28. 2024	THANKSGIVING RECESS	
16	-	**Tuesday, December 03, 2024**	**Midterm #2**	
	-	**Tuesday, December 10, 2024**	**Final Exam**	Bonus Assignment

Rough due date (Sunday)

HW: Homework

No quiz or homework due on exam week.

CP: Course Project

Spread workload.

https://ttpoll.com/p/967596

Expect Changes!

Change is good!

In-Class Experience

- Classes will be in person unless specified otherwise
- The class will combine PowerPoint lectures, whiteboard instruction, code examples, and in-class problems.
- Class materials (e.g., slides, code, recordings) will be posted to Canvas.
- Slides will be provided after each lecture.
- A notebook is recommended for in-class problem-solving and to capture additional notes.

Attendance Policy

- I will attempt to record all lectures and post them on Canvas
  - Technology can fail
  - Humans can forget to push the record button
  - Please do your best to attend
- Excused absences: Follow the Student Life's Website guidance (Link)
  - Missing a lecture: Communicate in advance (when possible) via email
  - Missing an exam: Excused absences due to injury, illness, death in the family, or distress will be granted upon the completion of the appropriate Student Life Absence Notifications form.

Disability Services

- Students who require accommodation(s) should make an appointment with the Office of Disability Services (https://sds.utk.edu/, 865-974-6087) to discuss their specific needs or with me during my office hours.

https://ttpoll.com/p/967596

In-Class Communication

- Engage!
  - Questions make the class fun.
  - An active classroom leads to more learning.
  - Raise your hand or use Discord for real-time questions during lectures
- Mutual respect
  - Please listen when someone is talking
  - Engage classmates before/after class and during group exercises

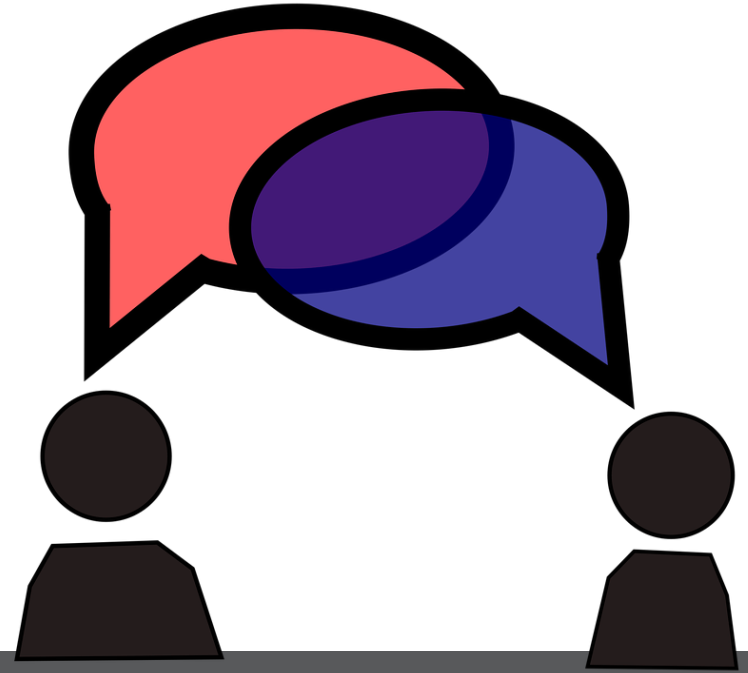
https://ttpoll.com/p/967596

Off-Class Communication

- My response time to emails and messages
  - 24 hours during weekdays
  - 48 hours during weekends or travel

- I receive many emails. Kindly send a reminder after the corresponding response time has elapsed.

- ***Discord*** is the preferred, class-wide communication tool.



https://ttpoll.com/p/967596

Discord Server
UTK Fall-24 COSC 325

Join @
https://discord.gg/DXpnvT9R

https://ttpoll.com/p/967596

Discord

- You can download the Discord app or use the browser version.
- General and homework-and project-specific questions
- Openly ask questions. You may be helping others.
- PM the TAs or me if you need ***privacy***.
  - We may ask to post the question if we believe it can help others.
  - If you don't want your name associated with the post, we can post it for you.
- All class announcements occur in Discord (Check channel daily).

https://ttpoll.com/p/967596

Again, mutual respect…

When posting to discussion boards, chatting through Zoom, and videoconferencing, please follow general professional guidelines; be civil and polite.

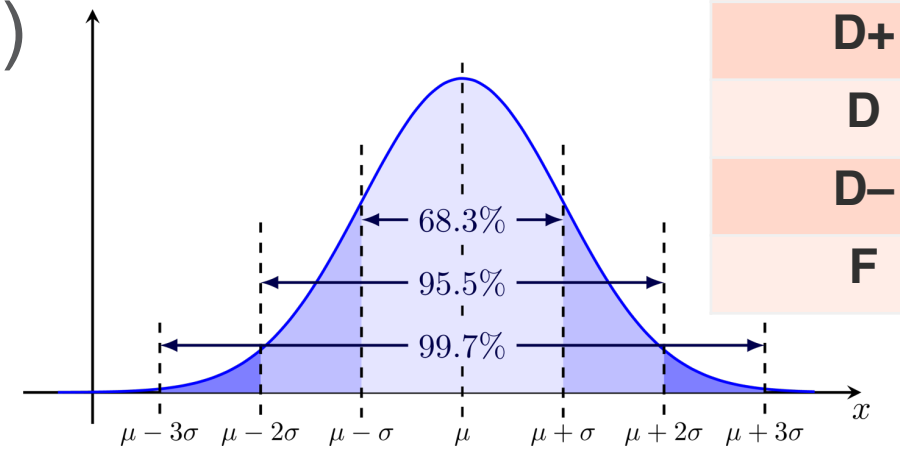
Course Work and Grading

THE UNIVERSITY OF TENNESSEE KNOXVILLE

https://ttpoll.com/p/967596

Grading Breakdown

- In-class exams (45%)
- Quizzes (10%)
- Programming homework (20%)
- Course project (20%).
- Attendance (5%)



Grade	Performance Level	Average
A	Superior	94-100
A–	Intermediate Grade	90-93
B+	Very Good	87-89
B	Good	84-86
B–	Intermediate Grade	80-83
C+	Fair	77-79
C	Satisfactory	74-76
C–	Unsatisfactory	70-73
D+	Unsatisfactory	67-69
D	Unsatisfactory	64-66
D–	Unsatisfactory	61-63
F	Failure	<60

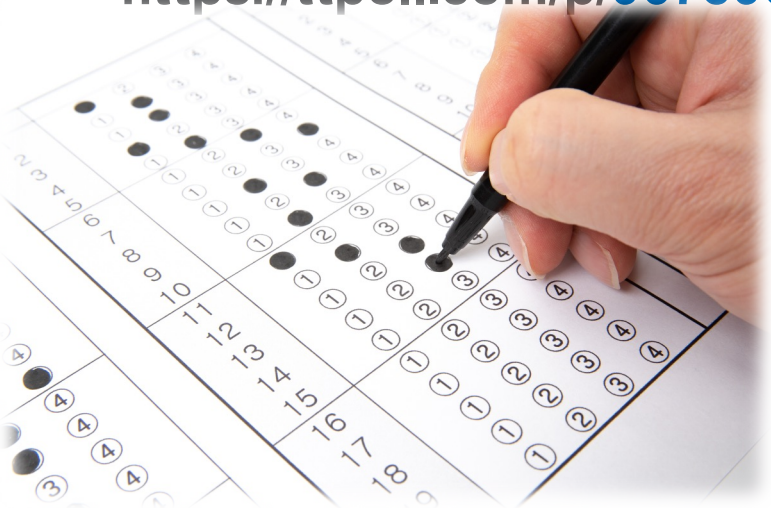
https://ttpoll.com/p/967596

In-Class Exams (45%)

- There will be three multiple-choice, in-class exams.
- Keep the two highest scores.
  - Exam #1: **Thursday, October 3rd, 2024**
  - Exam #2: **Tuesday, December 3rd, 2024**
  - Final Exam [Optional]: **Tuesday, December 10th, 2024**
- Excused absences due to injury, illness, death in the family, or distress will be granted upon completing the appropriate UTK Student Life Absence Notifications form (https://studentlife.utk.edu/caring-for-all-vols/absence/).

https://ttpoll.com/p/967596

Quizzes (10%)

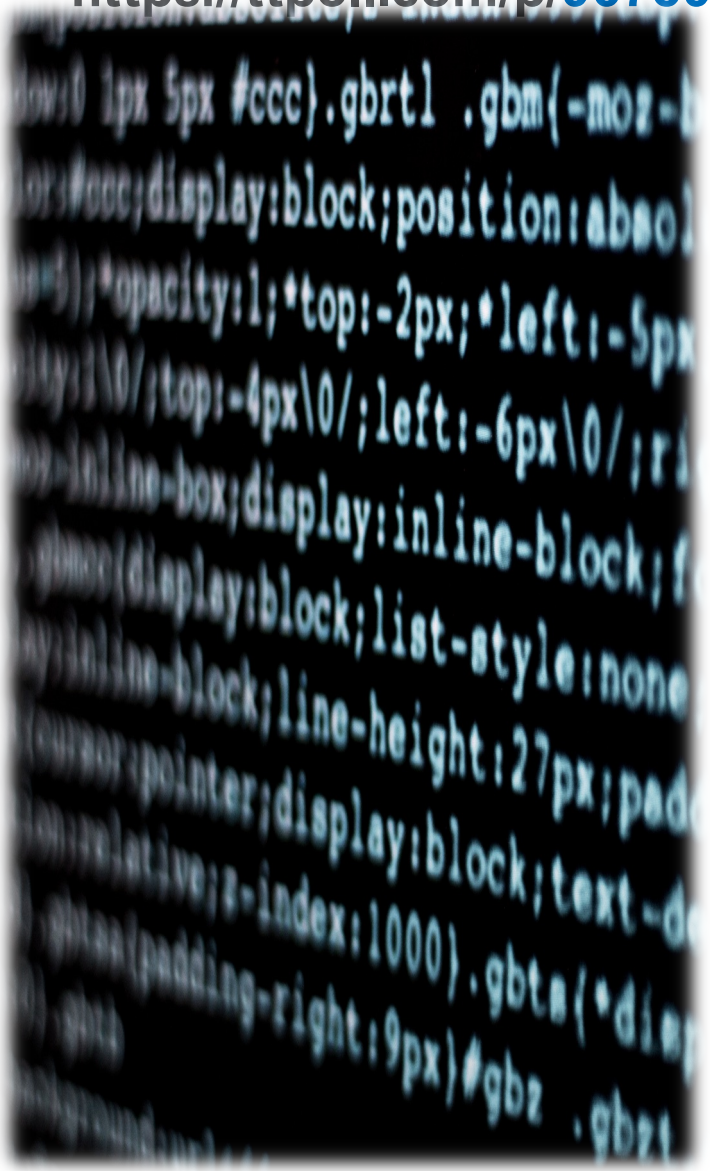


- About 5-10 multiple-choice questions
- Due every Sunday by midnight
- Retain your best 9 out of 12 quizzes (Top ~75%)
- Quizzes, solutions, and grading will be posted to Canvas.
- Note that missing a quiz will not be excused.

https://ttpoll.com/p/967596

Homework (20%)

- There will be **seven** homework assignments.
- **Homework is an individual task.**
- All submissions should be made on the Canvas website.
- All programming should be done in Python 3 and appropriately documented.
- Submit the following files:
  - A single PDF file for the report.
  - A single PDF file for each Python file
  - A zipped folder with the code and report



Homework (20%)

- **Late deadline:**
  - The late deadline is the “Available Until” date in Canvas (+2/3 days)
  - A ten-base point penalty will be applied to your score for each day after the submission deadline.
  - Automated zero score after late deadline.
- **Slack days:**
  - Everyone gets up to 5 slack days.
  - You can distribute slack days across homeworks.
  - Once you use all your slack days, you will fall into the late-submission category by default.

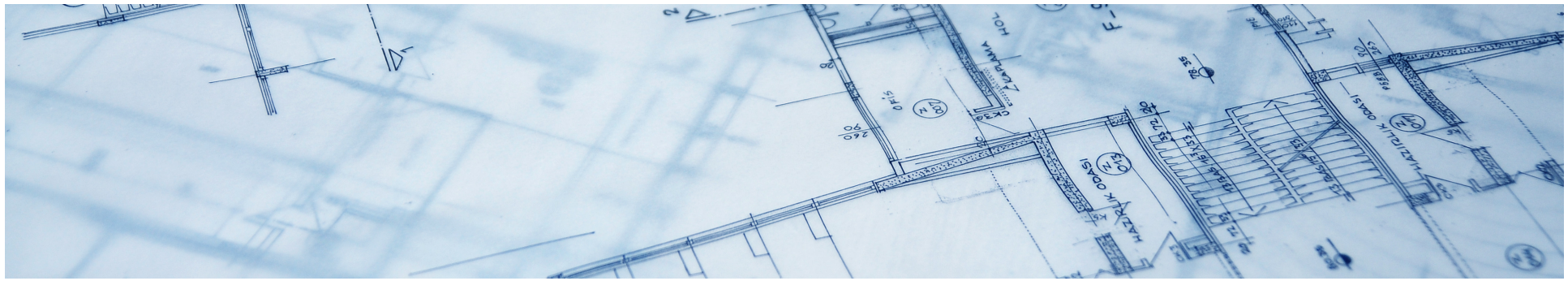
Course Project (20%)

- Teams comprised 2-3 students
- The course project breakdown is as follows:
  - PRFAQ (10%) - Project motivation, vision, impact
  - Midterm Report (30%): Document the motivation, goals, schedule, initial conceptual design, baseline, etc.
  - Presentation (20%): Oral summary and demonstration.
  - Final Report and Code (40%): Final and overall project documentation.
- I will provide instructions early on. So, you can start early.

https://ttpoll.com/p/967596

Course Project Schedule

- CP#1: PRFAQ: 09/22/2024 (Sunday)
- CP#2: Midterm Report: 10/20/2024 (Sunday)
- CP#3: Presentation: 11/19/2024 (Tuesday)
- CP#4: Final Project Report: 11/26/2024 (Tuesday)



https://ttpoll.com/p/967596

Attendance (5%)

- You are expected to attend in-person lectures.
- Attendance will be tracked via pop quizzes.
- Clicker: PointSolutions (Turning)
- Scoring based on answered questions
  - <= 50%, results in 0%
  - > 50% results in 2.5%
  - > 66%, results in 5% (Full attendance credit)
- Use UTK email to log in to PointSolutions
  - You should go through UTK SSO Authentication

pointsolutions

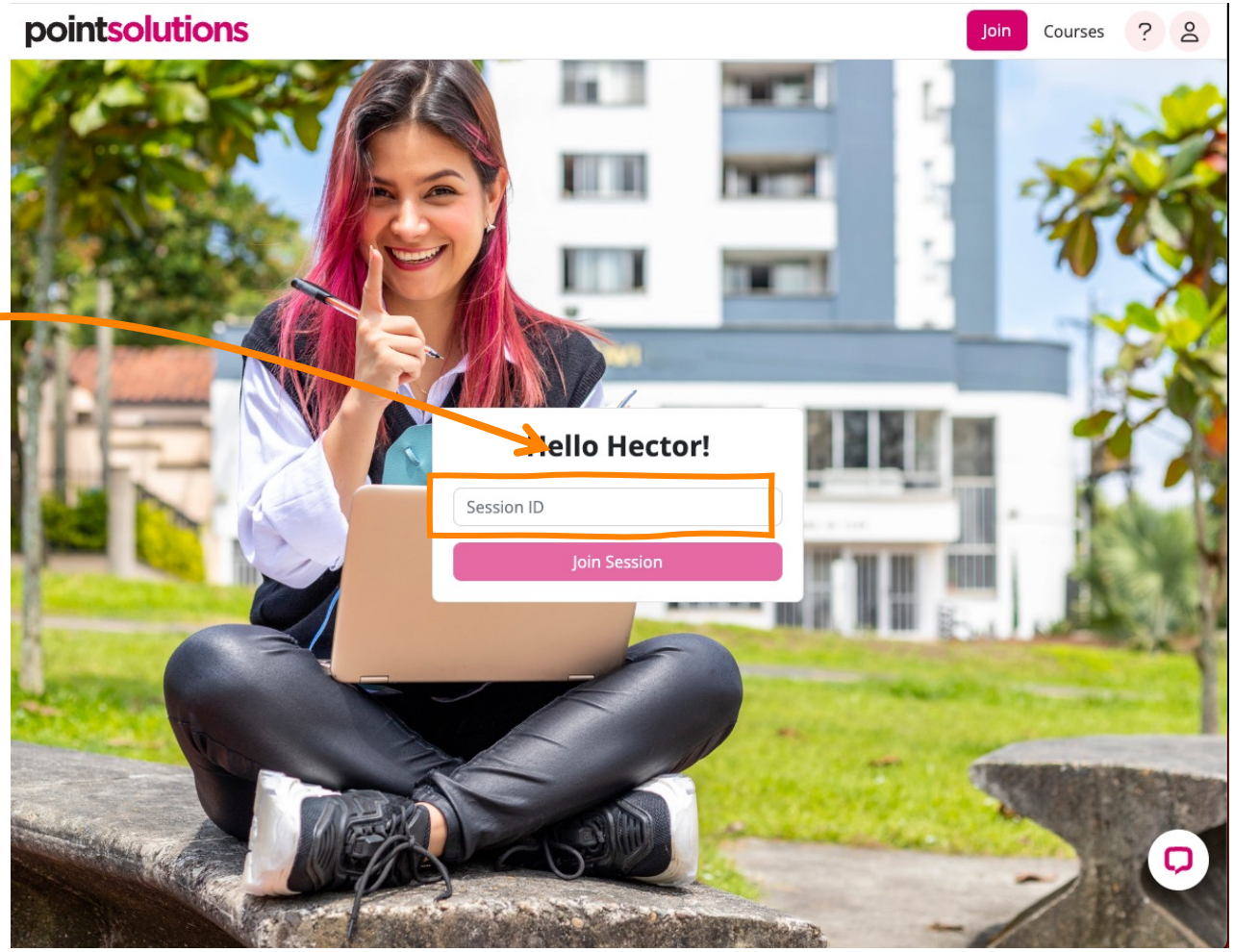
https://ttpoll.com/p/967596

Pop Quizzes

- You will go to a link like this.

  https://ttpoll.com/p/446768

- The session ID is the last 6 digits in the URL colored blue.
  - This ID is generated randomly before the session starts
  - The link will be available on each Pop Quiz slide.
  - You get a point for answering.



https://ttpoll.com/p/967596

Pop Quiz Questions

https://ttpoll.com/p/967596

Extra Credit

- Work distribution designed for flexibility
  - Drop lowest exam score (Final is optional)
  - Drop lowest three quiz scores.
  - Drop lowest homework score.
- Bonus assignment worth (3.33% of grade)
- No additional extra credit assignments

https://ttpoll.com/p/967596

Academic Integrity

- You are expected to comply with UTK policy on academic integrity and plagiarism.
- Collaboration and intellectual exchange on homework, exam study, and course project is encouraged.
- However, ***do not***
  - Copy and paste code from others.
  - Share solutions to problems in Discord.
  - Use other people's work as yours.
- I reserve the right to audit any course artifact.
- Please don't hesitate to reach out to me for clarification or help.

https://ttpoll.com/p/967596

Getting Help

- Try on your own… try hard!
- Search for resources online.
- Go to peers.
  - Look for hints/guidance and not full solutions.
- Go to teacher assistants.
  - Start with email.
  - They will not debug your code.
- Go to the instructor.

https://ttpoll.com/p/967596

AI use in class

- **I encourage the use of AI**
  - You must disclose (i.e., cite or acknowledge) the use of any Generative AI tool, such as ChatGPT
  - **I use ChatGPT** as a co-pilot in the creative process and illustration generation.
- ***Life is challenging…***
  - ***Take the easy path now and struggle later.***
  - ***Put the effort now and rip the benefits later.***
  - ***Choose your challenges.***
- Homework and projects are designed to help you learn.

https://ttpoll.com/p/967596

My Teaching Philosophy

- Teaching is a privilege
  - I know very little.
  - You help me learn a little bit more.
- We are all different.
  - I attempt to provide different learning approaches
    - Coding, exams, lectures, books, coding examples, etc.
- Repetition is king
- Search for the “why” and reinforce with the “how.”
- Pick your challenges.

https://ttpoll.com/p/967596

Next Lecture

- Why do we need this class?
- What is machine learning (ML)?
- ML Terminology
- ML Notation
- Basic programming tools