PH451 PH551

January 8, 2025

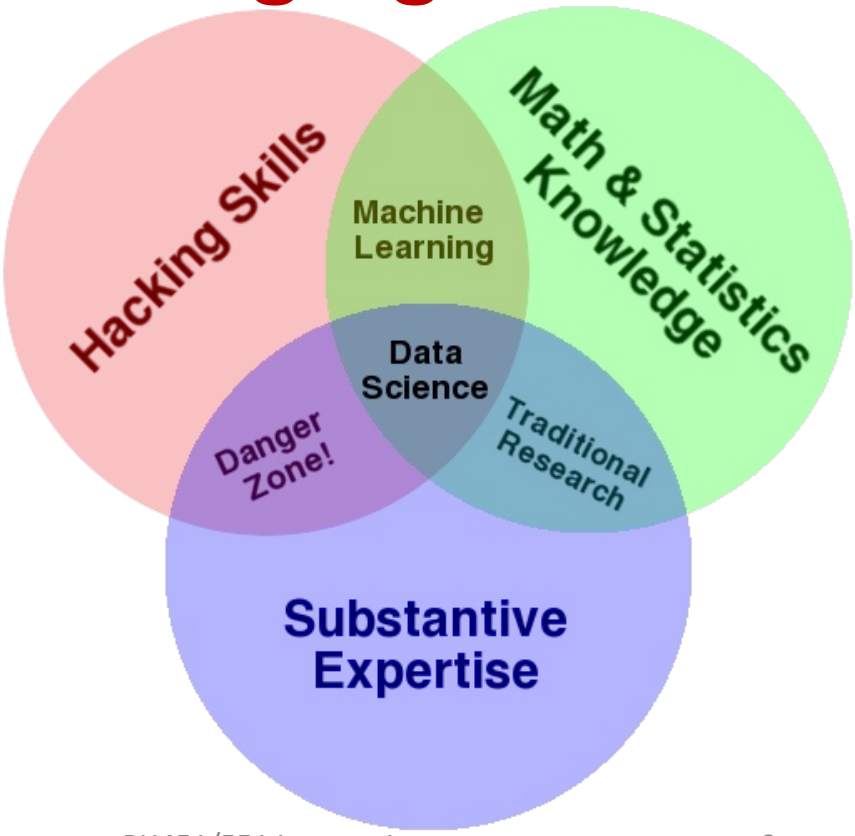
Today's Outline

- **Introduction and Overview**
- **Syllabus**

PH 451/551 Machine Learning

Machine Learning (ML) /Data Science (DS) is an emerging discipline

- Forms the basis of many scientific and engineering applications



Course covers

Fundamentals of machine learning and classification theory

- **Theory and practice**

Popular algorithms

- **decision trees, neural networks…**

Types of learning

- **supervised, unsupervised…**

Machine Learning Applications

- **classification, regression, clustering, anomaly detection…**

Pre-Requisites

No formal pre-requisites

However, good to have a familiarity with

- **basic calculus, linear algebra, statistics (probability). We will briefly cover key concepts during the course**
- **basic programming (python) for practical exercises**
- **No knowledge of physics is required**

Lecture and Hands-on

- **The course is taught with two weekly lectures:**
  - **Tue and Thu 1:00 – 2:30 pm**
- **Practical (hands-on) sessions** will be incorporated into the regular lecture times
  - **Make sure to attend all lectures as they will feature in-class activities and practical sessions**

Practical Sessions

During hands-on/practical sessions: you will see **demos** of some of the **important algorithms** and techniques that you will learn about in the lectures

- **You will have a chance to learn the experimental work that underlines data science as a science**
  - **Hands-on sessions will be graded and attendance is required.**

Course Website

- We will use Blackboard Learn to manage the course
- I will send course announcements through the Blackboard messaging system
- I will also post course materials (lecture notes, exams/quizzes) there
- Best way to contact me is through Blackboard Learn – I'll check messages there daily.

Programming Languages/Tools

- We will use Python for the course
  - Practical exercises etc.
  - Exception: for some activities you will be able to use any language or framework of your choosing
- Machine learning frameworks
  - scikit-learn, Pytorch, C5.0
  - optional: Keras/Tensorflow

Course Diary

- We will use Notion to keep a weekly course diary (linked to BB)
- This is where you will find required weekly readings, in-class and homework assignments and their due dates
- Please check this page frequently! It will be regularly updated

Slack

- We will use Slack for technical questions and asynchronous discussions
- Please join via the link on Blackboard and regularly check that space

Pre-Course Survey

- Please fill the pre-course survey on Blackboard (required)
- It should take only a few minutes
  - Due before next class

How Your Final Grade is Calculated

- Participation/in-class activities: 15%
- Homework: 15%
- Quiz (February): 10%
- Exam (March): 20%
- Final Project: 40%

Final Projects

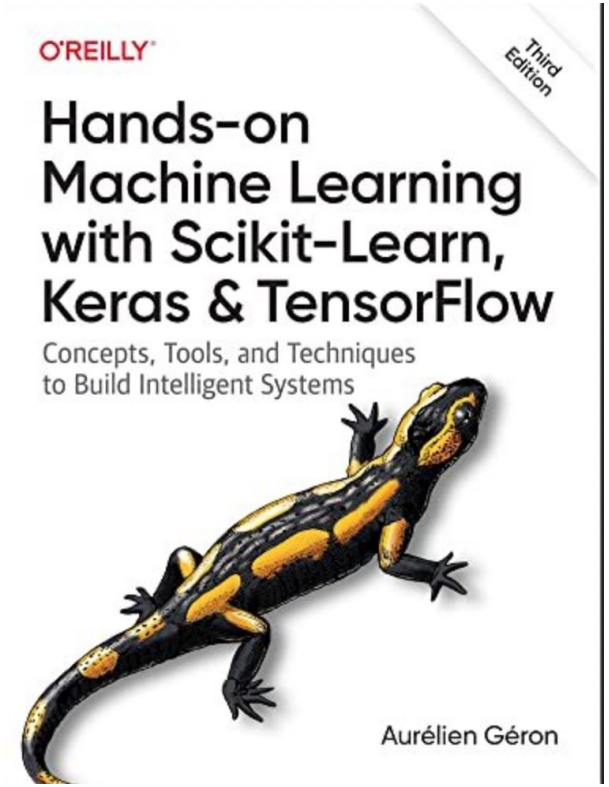
- Semester long activity (40% of grade)
- Team activity
  - Self-designed project in consultation with instructor
  - Has to include an ML component
  - Can pick any topic you feel passionate about
  - Graduate students – can pick something from your specific field or research area

Final Projects

- Pre-proposal 5%
- Proposal 10%
- Project Outline and Demo 25%
- Presentation 10%
- Peer Evaluation 10%
- Final Project Submission (write-up, code) 40%

Required Textbook

- **"Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow" by Aurelien Geron**



Additional Resources

- "Deep Learning" by Goodfellow, Bengio and Courville (free)
- "Python Data Science Handbook" by J. VanderPlas (free)
- "C4.5" by J. Ross Quinlan

General Attendance Policy

- Students are expected to attend all classes
- Performance in hands-on activities is graded and counts toward overall course grade

Late Submissions

- **Assignments are due on their specified due dates**
- **Late submissions for assignments/projects will be penalized:**
  - **0 - 24 hours late** **-10 points**
  - **24 - 48 hours late** **-20 points**
  - **> 48 hours late** **0 points**

Collaboration Policy

Individual homework assignments, quizzes and midterm are to be solved individually:

- **your own answers and code**
- **can collaborate with others in figuring out the answers to homework and discuss approaches on a high-level**
  - No code sharing or results

Collaboration Policy

For team-based activities:

- **in-class exercises**
- **mini-hackathon exercises (in-class or for homework) and final projects**

Collaboration is allowed and highly encouraged

Academic Misconduct Policy

- All students in attendance at University of Alabama are expected to be honorable and to observe standards of conduct appropriate to a community of scholars.
- The University of Alabama expects from its students a higher standard of conduct than the minimum required to avoid discipline.
- Please see syllabus for full policy

Exam Rules

- **Cellphones and other unapproved electronic devices must be turned completely off and placed with all other belongings on the floor. All watches must be put away. Do not put your phone or watch in your lap or on your chair or desk.**
- **Physically holding or concealing or otherwise using your cell phone, smart watch, or any other unapproved electronic device during the exam will be treated as academic misconduct.**

Disability Statement

- If you are registered with the Office of Disability Services, please make an appointment with me as soon as possible to discuss any course accommodations that may be necessary.

- If you have a disability, but have not contacted the Office of Disability Services, please do so as soon as possible. Students with disabilities must be registered with the Office of Disability Service before receiving academic adjustments.

UAct

The University of Alabama is committed to an ethical, inclusive community defined by respect and civility. The UAct website (www.ua.edu/uact) provides extensive information on how to report or obtain assistance with a variety of issues, including issues related to dating violence, domestic violence, stalking , sexual assault, sexual violence or other Title IX violations, illegal discrimination, harassment, child abuse or neglect, hazing, threat assessment, retaliation, and ethical violations or fraud.

Course Content

- Introduction; Probability and Statistics; Python; Classification Theory; Types of data and learning
- Constructing classifiers; Performance Metrics; Linear models; Kernels; Support Vector Machines
- Decision Trees; Decision Rules
- Ensembles; Boosting; Bagging; Random Forests; Regression Trees; Hierarchical models; Regularization; Multi-target Regression

Course Content

- **Feature Importance, Engineering, Pre-processing and Transformation; Dimensionality Reduction; Principal Components; TSNE; Locally-Linear Embedding;**
- **Unsupervised methods and Clustering; Genetic Algorithms**
- **Neural Networks; Gradient Descent and Back-propagation; Activation and Loss Functions**

Course Content

- **Deep Learning, Optimizers; Convolutional Neural Networks; Recurrent Neural Networks; LSTMs, Transformers, LLMs**
- **Generative Models: Generative Adversarial Networks; Autoencoders; Variational Methods; Energy Models; RBMs; Hopfield Networks; Bayesian Networks; Anomaly detection**
- **Geometric Deep Learning, Graph Neural Networks, Physics-Inspired Models, Symbolic Regression, Quantum ML**

Instructors

- **Primary Instructor:**
  - **Prof. Sergei Gleyzer**

- **Teaching and Learning Assistants**
  - **Eric Reinhardt (GTA)**
  - **Thanh Nguyen (Undergraduate LA)**
  - **Dinesh PR (Graduate LA)**

Sergei Gleyzer

Foundational and Applied Machine Learning

New machine learning algorithms; Physics-inspired models; Representational learning; Foundation Models for Science, Realtime Machine Learning

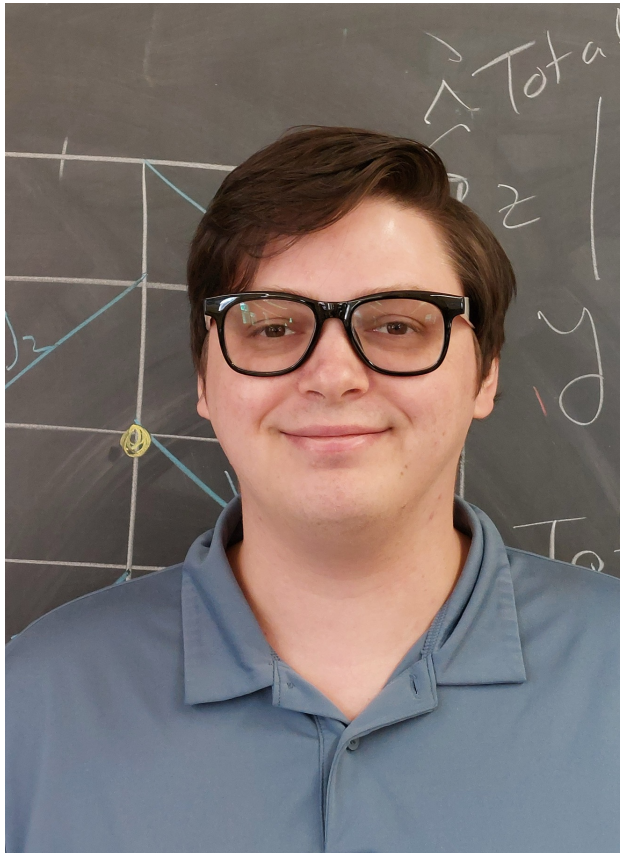
Applications: High-Energy Physics (Large Hadron Collider/CERN), astrophysics, quantum computing; end-to-end deep learning; discovery science with AI

Sergei Gleyzer

- **Founder of Inter-experimental LHC Machine Learning WG, CMS Machine Learning Forum, Machine Learning for Science and HumanAI Organizations**
  - **APS Fellow 2024**
  - **Distinguished Teaching Fellow 2022-2026**
  - **UA ORED Senior Fellow 2025**
  - **Fermilab LPC Senior Distinguished Researcher Fellow 2025**

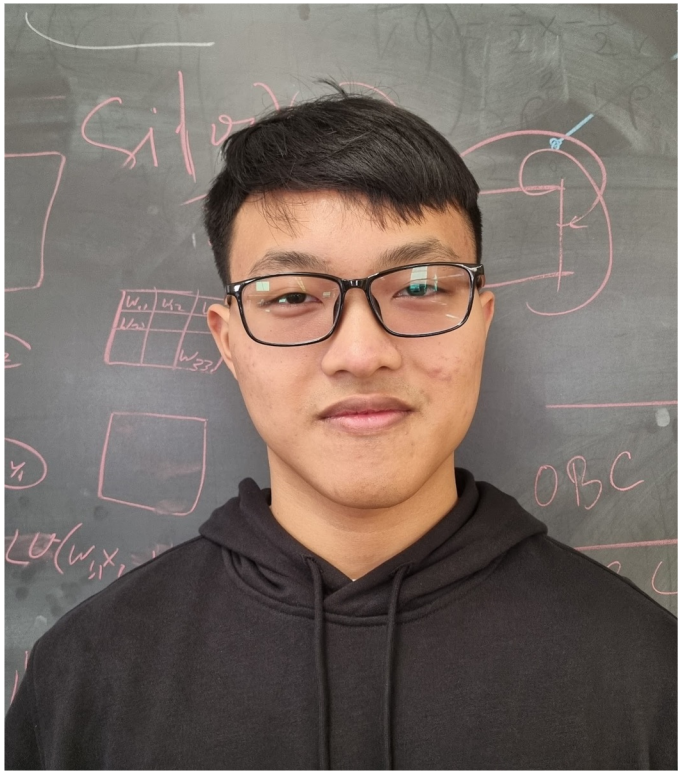
Eric Reinhardt

- 3rd year physics grad student
- 3rd time TA for the course
- Physics-informed models, foundational models, natural language processing
- Applications: theoretical and experimental high energy physics, condensed matter phys., mathematics



Thanh Nguyen

- Junior in Data Science,
- Interested in data analysis, ML applications, and learning quantum computing,
- Contributing to the visual inspection team for UA CMS,
- Research assistant in human trafficking at IDA,
- I play chess, video games, and read books in my free time.



Hi, I’m Dinesh

From India

2nd time (half) TA

Interests: Math, physics, generative ML, numerical simulations

Hobbies: athletics, treks, travel, history, gaming

Likes: Long arguments

Guest Speakers

During the semester we will have some invited speakers from academia and industry

Goal:

- **Real-world perspective for data science concepts and applications learned in class**
- **Showcase potential future career paths enabled by data science**

Class Introduction