THE OHIO STATE UNIVERSITY

STAT 8810
Advanced Topics in Statistics I (35715)
Autumn 2026

Syllabus - Graduate Instructor Information

Name: Subhadeep Paul

Professional Qualifications

Ph.D in Statistics, University of Illinois, Urbana-Champaign (2017)

Syllabus - Course Schedule

Statistical Foundations of AI

Autumn 2026

Week	Dates	Topics
1	10/20	Linear Neural Network (NN) for regression, Stochastic Gradient Descent (SGD)
1	10/22	Multilayer Perceptron (MLP) and feedforward NN

Week	Dates	Topics
2	10/27	NN for Classification, Regularization
2	10/29	Convolutional Neural Networks (CNN)
3	11/03	(University-mandated asynchronous online) Modern architecture for CNN, AlexNet, ResNet
3	11/05	Statistical Learning Theory, High-dimensional linear model
4	11/10	Statistical Theory for Deep Neural Networks (DNN)
4	11/12	Tabular data: Trees, Random Forests, Gradient-boosted trees, Project proposal due
5	11/17	Sequence data, recurrent neural networks, and language models
5	11/19	LSTM, Deep RNN, Bidirectional RNN
6	11/24	Attention, Transformer, and LLM

Week	Dates	Topics
7	12/01	Statistical Transfer Learning, Prediction-powered inference, Federated Learning
7	12/03	Student Project Presentation
8	12/08	Student Project Presentation, Final project report due

Course Details - Instructor Contact Details

Email address: paul.963@osu.edu

Course Details - Office Hours

Cockins Hall 231, Wednesdays 1:00PM – 2:00 PM

Course Details - Class Meeting Details

Class Meeting Details: 12:40 PM-02:30 PM - Enarson Classroom Bldg 258 - TR

Course Details - Course Description

A systematic advanced treatment of areas of current interest in Statistics. Topics will be announced each semester. Prereq: Grad standing in Statistics or Biostatistics, or permission of instructor. Repeatable to a maximum of 24 cr hrs or 8 completions.
Prerequisite(s) / Corequisites: Prereq: Grad standing in Statistics or Biostatistics, or permission of instructor.

Course Details - More About This Course

The course is intended to introduce methods and theory related to deep neural networks and artificial intelligence from a statistical perspective. We will learn the mathematics behind the modern deep learning methods, learn to implement those in data using python, and discuss statistical theory for these methods.

Module 1: Deep Neural Networks

1. Neural network for Linear regression
2. Mini batch Stochastic gradient descent
3. Implementation in Pytorch
4. Multilayer Perceptrons and feed forward neural networks
5. Backpropagation
6. Neural networks for classification

Module 2: Image data: Convolution Neural Networks.

Module 3: Sequence data, recurrent neural network, transformer, and LLM

Module 4: Statistical Learning Theory

1. Basics of decision theory
2. High dimensional linear regression model, lasso

Module 5: Statistical theory for deep learning

1. Approximation results
2. Statistical error rates

Module 6: Transfer Learning, Federated Learning, Prediction-powered inference

Course Details - Course Learning Outcomes

Upon successful completion of the course, students will be able to

1. Fit DNN models to numerical, image, and text datasets
2. Understand the workings and outputs of deep neural network and AI models such as Large language models
3. Understand the statistical theory behind DNN
4. Understand the concepts of transfer learning, federated learning, and prediction-powered inference.

Course Details - Grading Policy

Grades

Assignment or category	Percentage
Project proposal	20
Final Project Presentation	60
Final Project Report and Review	20
Total	100

Course Details - Grading Scale

Letter Grade	Percentage Range
A	100% to 93%
A-	< 93% to 90%
B+	< 90% to 87%
B	< 87% to 83%
B-	< 83% to 80%
C+	< 80% to 77%
C	< 77% to 73%
C-	< 73% to 70%

Letter Grade	Percentage Range
D+	< 70% to 67%
D	< 67% to 60%
E	< 60% to 0%

Course Details - Other

Course materials and technologies

Textbooks

Required

Zhang, A., Lipton, Z. C., Li, M., & Smola, A. J. (2023). Dive into deep learning. Cambridge University Press.

Freely available here: https://d2l.ai/

Optional books:

Prince, Simon JD. Understanding deep learning. MIT press, 2023. Freely available here: https://udlbook.github.io/udlbook/

Bishop, Christopher M., and Hugh Bishop. Deep learning: Foundations and concepts. Vol. 1. Cham, Switzerland: Springer, 2024. Freely available here: https://link.springer.com/book/10.1007/978-3-031-45468-4

Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). The elements

of statistical learning: data mining, inference, and prediction (Vol. 2, pp. 1-758). New York: springer. Freely available here: https://link.springer.com/book/10.1007/978-0-387-21606-5

The books will be supplemented by papers and lecture notes.

Course Details - Course Policies

Project: There is no homework or exam for this course. The evaluation is based on a written project proposal and final project presentation and report is due at the end of the course (the tentative schedule of the presentation is given below).

There are three options for a project:

1. Micro-review: Write a micro-review of the literature on a topic (approximately 5-10 pages). I will provide you initial feedback on suitability of the topic in the project proposal, but most related topics are fine. The reviews are expected to be of very high quality. The topic should be broad enough to be able to write a substantial review on, yet not too broad that it becomes difficult to review in a short writeup.
2. Short Research Report: You may also perform original research on a topic related to deep learning, tree based methods, and AI, and submit a short research report (minimum 5 pages and no maximum). If you take the research route, you need to declare a stated research goal (even if vague) in the proposal document and show progress towards achieving this goal. I don't expect you to have completed the stated objective before the semester ends, however the written report should reflect a substantial good faith effort. You may use existing (unpublished) research that you may already be conducting related to the topic of the course as part of your project, but please get permission from other co-authors or advisors on presenting the research during our project presentation and

submitting the short research report.

3. Data analysis or system building project: You may also analyze a dataset using methods learnt in the class to obtain scientific insights, or build a software system. The submitted report should be minimum 5 pages with no maximum. You are free to choose any dataset you like. It is expected that you will choose a data and scientific problem such that the data analysis task will be substantial and worthy of a ML or AI project. You may also use data you have collected or analyzing as part of your research and you will not have to share the data with me. However, as in part (2), if you are using existing research for the project, be sure to take permission from co-authors and advisors for presentation and short written report.

You must work on in short groups of size 2-3 I expect the effort and quality of the documents to increase linearly with the number of authors. For all three options, 2 reports and a presentation are required: a project proposal (approx. 1 page), a final written report (see above for suggested lengths), and a final presentation on the last two days of scheduled class.

Tentative Dates for presentations:

Final project presentation: December 03, and December 08, during class time

Course Details - AI Use Policy

GenAI usage is allowed at level 4: Co-Creation and Drafting. See details below.

GenAI can fabricate information (for example, creating citations to sources that

do not exist) or omit key information. You are ultimately responsible for the accuracy and integrity of your submitted work. Unless explicitly permitted below, representing AI-generated content as your own work is a violation of the Code of Student Conduct. Please check with your instructor if you have questions.

These levels may apply to the entire course or specific assignments, as indicated by the instructor:

- Permitted: You may use GenAI to produce initial drafts, refine text, or co-create. You may use GenAI to translate materials.
- Prohibited: The materials in this course may be subject to copyright protection. Do not input or allow AI to access external or instructor-created course materials (e.g., assigned readings, slides) into AI systems without explicit instructor permission.
- Requirement: You must document and critically evaluate the accuracy of AI outputs. You are responsible for authoring the final text/work.

Syllabus - University Policies

Office of Undergraduate Education - Standard Syllabus Statements

- Academic Misconduct
- Religious Accommodations
- Accommodations for Disability and Illness
- Intellectual Diversity
- Grievances and Solving Problems
- Creating an Environment Free from Harassment, Discrimination, and Sexual Misconduct
- Copyrights
- Content Warnings
- Mental Health Support
- Resources for Military-Connected Students
- Campus-wide Class Cancellations