NC STATE UNIVERSITY

College of Engineering
Department of Nuclear Engineering

https://www.ne.ncsu.edu/

Campus Box 7909
Burlington Engineering Labs
2500 Stinson Drive
Raleigh, NC 27695-7909

NE 795 Advanced Topics in Nuclear Engineering
Scientific Machine Learning

Fall 2023, 3 Credit Hours

1 Course Information

- **Schedule**
  - *Time:* Tuesdays and Thursdays, 10:15 AM - 11:30 AM
  - *Location:* Room 461, Riddick Hall
  - *Course website:* **Moodle** will be used to post lecture notes, homework, computer projects and other materials.
    https://moodle-courses2324.wolfware.ncsu.edu/course/view.php?id=3900
  - *Lecture recordings:* **Panopto** will be used to host the lecture recordings. The link will be posted once available.
    https://ncsu.hosted.panopto.com/Panopto/Pages/Sessions/List.aspx#folderID=%22663613de-f60d-4073-b7b9-b056013ec30a%22
- **Instructor**
  - *Instructor:* Dr. Xu Wu, Assistant Professor of Nuclear Engineering
  - *Office:* Burlington Laboratory 2110
  - *Office Hours:* Mondays, 1:00 - 3:00 PM, in-person, or Zoom (by appointment)
  - *Phone:* 919-515-6570
  - *Email:* xwu27@ncsu.edu
  - *Website:* https://www.ne.ncsu.edu/people/xwu27
- **Teaching Assistant**
  - *TA:* Ziyu Xie
  - *Email:* zxie22@ncsu.edu
  - *Office Hours:* TBD

2 Course Description

- **Artificial Intelligence (AI)** tries to build smart machines to automate intellectual tasks normally performed by humans, in a way that we consider intelligent. **Machine Learning (ML)** is a subset of AI that studies computer algorithms which improve automatically through experience (data). ML algorithms typically build a mathematical model based on training data and then make predictions without being explicitly programmed to do so. Its performance increases with experience (data), in other words, the machine learns. **Scientific machine learning (SciML)** is a burgeoning discipline in AI/ML that blends scientific computing and ML. SciML is a core component of ML. It consists of computational technologies that can be trained with scientific data to augment or automate human skills.

- This course aims at augmenting the applications of AI/ML in scientific computing especially in the nuclear engineering area, and preparing the students for transformative solutions across various DOE missions, and for driving the next wave of data-driven scientific discovery in nuclear engineering.
- After this course, the students will understand most of the fundamentals of supervised ML, including neural networks and other algorithms for regression/classification, as well as dimensionality reduction for unsupervised ML. The students will be able to implement ML algorithms in nuclear engineering scientific computing applications, with a heavy focus on uncertainty quantification, sensitivity analysis, calibration and inverse problems.

3 Student Learning Outcomes

- The students completing this course will be able to:
  - analyze common regression/classification problems using classical supervised ML techniques;
  - demonstrate understanding of the fundamentals of dimensionality reduction for unsupervised ML;
  - evaluate the performance of ML models;
  - build ML models and write numerical codes to solve regression, classification and dimensionality reduction problems;
  - use open source ML libraries such as scikit-learn, keras, TensorFlow and PyTorch.
  - apply SciML methods on uncertainty quantification, sensitivity analysis, calibration and inverse problems;
  - implement advanced techniques for UQ of neural networks;
  - solve research problems using the SciML techniques learned in this class;
  - develop research skills that combine the complementary perspectives of computational science and computer science to craft a new generation of ML methods for complex applications across nuclear engineering.

4 Prerequisites

- Background in calculus, linear algebra, scientific computing. Undergraduate level courses in these areas are sufficient.
- Background in Probability and Statistics, equivalent to ST 311 (Introduction to Statistics) and ST 371 (Introduction to Probability and Distribution Theory).
- Prior experience in Machine Learning is **not** required.
- **Python** is the recommended programming language.
- In the first lecture, instructions and resources to learn Python will be provided, as well as Python ecosystems and libraries that will be useful for this course.

5 Outline of Topics

Part 1: Artificial Neural Network (ANN)

1. Introduction
2. Perceptrons, Sigmoid Neurons and ANNs
3. Classifying Handwritten Digits and Gradient Descent
4. Backpropagation
5. Cross-entropy, Softmax and Negative Log-likelihood Cost
6. Overfitting and Regularization
7. Bias, Variance, Hyperparameters and Weight Initialization
8. Vanishing Gradient, Activation Functions and SGD Variations

Part 2: Fundamental Techniques of Machine Learning

9. Generalized Linear Regression, Subset Selection and Shrinkage
10. Logistic Regression
11. K-Nearest Neighbor (KNN) and Feature Scaling
12. Naive Bayes (NB)
13. Decision Tree (DT)
14. Ensemble Learning - Bagging, Boosting and Stacking
15. Support Vector Machine (SVM)
16. Dimensionality Reduction with Principal Component Analysis (PCA)
17. Gaussian Processes (GP) Modeling (2 lectures)

Part 3: Uncertainty Quantification (UQ) and Sensitivity Analysis (SA) with ML

18. Forward UQ Methods (3 lectures)
19. SA Methods
20. Inverse UQ Methods
21. UQ of Deep Neural Networks (DNNs)

6 Course Schedule

- Detailed course schedule will be available in a separate document, including dates for each section/topic, homework/project assignment dates and deadlines, midterm exam date, etc.

Table 1: List of homework and topics.

	Topics	Weights	Deadlines
HW1	Artificial Neural Networks (ANNs)	10%	10/05/2023
HW2	Other Fundamental ML Techniques	10%	10/xx/2023
HW3	Gaussian Process (GP) and Principal Component Analysis (PCA)	10%	11/xx/2023
HW4	Uncertainty Quantification (UQ) and Sensitivity Analysis (SA)	10%	11/xx/2023

7 Assignments and Grading Policy

(1) **Homework problems (40%)**, see Table 1.

- Homework will be assigned periodically throughout the semester. The last homework assignment may be due during the last week of classes.
- Homework will be assigned and submitted via Moodle.

(2) **Computational projects (60%)**

- The computer project will be evaluated by a **proposal**, a **progress report**, a **final report** and a **final presentation**. Each of the milestones will be submitted through Moodle. See Table 2 for the schedule.
- You may work by yourself or in teams (max. two students) for the project. There will be two options.
- **Option A (recommended)**: each group needs to propose a computer project based on the research of the group members. It is expected to be in the areas of SciML applications in prediction, uncertainty, sensitivity, calibration and validation, etc.
- **Option B**: if a team is unable to determine a proper computer project, the instructor will assign a published paper to the team. The team is expected to reproduce the results in the paper. If the codes are already available online from the original authors, the students will be asked to solve a different problem that is pertinent to Nuclear Engineering applications.

Table 2: Project schedules

Milestones	Weights of total	Deadlines
Proposal	6%	09/30/2023
Progress report	12%	10/29/2023
Final presentation	24%	11/xx/2023
Final report	18%	12/xx/2023

- The **proposal** should be 1-2 pages. It should contain the following information: (1) project title, (2) team member(s), (3) description of the problem to be solved with ML, (4) scope of the research to be performed and ML methodologies to be used, (5) anticipated results, and (6) reading list (important papers you will need to read).

- The **progress report** should include (1) a high quality introduction (including literature survey on similar work published in the literature), (2) what have you done so far, (3) what remains to be done and (4) a clear description of the division of work among teammates, if applicable.
- The **final report** should be 10 - 15 pages, including tables/figures/references/appendix. You should submit a PDF file electronically. The codes should be submitted separately in a Zip file. The report should use the format similar to a research paper. It should include (1) introduction, including the motivation, the issue to be addressed, the objective, the literature review and a quick summary of the problem/research/findings, (2) problem definition, (3) the methodologies, (4) ML implementation and major results, (5) discussions and conclusions, including comments on the results and open questions, and (6) major references.
- The **final presentation** will be in person or online via Zoom. Depending on how many teams we have, we will use the last two lectures' time for the presentation. Each project will have **15 minutes** to present including Q/A.

(3) **Grading**

- The course will be graded on the letter grading scale listed in Table 3 and will count toward your GPA. More information can be found at https://studentservices.ncsu.edu/your-grades/general-info/.
- *Requirements for Credit-Only (S/U) Grading*: In order to receive a grade of S, students are required to take all exams and quizzes, complete all assignments, and earn a grade of C- or better. Conversion from letter grading to credit only (S/U) grading is subject to university deadlines. Refer to the Registration and Records calendar for deadlines related to grading. For more details refer to http://policies.ncsu.edu/regulation/reg-02-20-15.
- *Requirements for Auditors (AU)*: Information about and requirements for auditing a course can be found at http://policies.ncsu.edu/regulation/reg-02-20-04.
- *Policies on Incomplete Grades*: If an extended deadline is not authorized by the instructor or department, an unfinished incomplete grade will automatically change to an F after either (a) the end of the next regular semester in which the student is enrolled (not including summer sessions), or (b) the end of 12 months if the student is not enrolled, whichever is shorter. Incompletes that change to F will count as an attempted course on transcripts. The burden of fulfilling an incomplete grade is the responsibility of the student. The university policy on incomplete grades is located at http://policies.ncsu.edu/regulation/reg-02-50-3.

8 Recommended Texts

There is no required text but self-contained lecture notes will be posted. Some useful references are:

(1) **For Statistics**

- Robert, C., & Casella, G. (2013). Monte Carlo statistical methods. Springer Science & Business Media.
- Gelman, A., Carlin, J., Stern, H., and Rubin, D. (2014). Bayesian Data Analysis, 3rd Edition, Chapman & Hall.
  http://www.stat.columbia.edu/~gelman/book/

Table 3: Letter grades.

Letter grades	Percentages	GPAs
A+	96-100%	4.333
A	92-96%	4.0
A-	89-92%	3.667
B+	86-89%	3.333
B	82-86%	3.0
B-	79-82%	2.667
C+	76-79%	2.333
C	72-76%	2.0
C-	69-72%	1.667
D+	65-69%	1.333
D	61-65%	1.0
D-	56-61%	0.667
F	0-56%	0

- Sivia, D., & Skilling, J. (2006). Data analysis: a Bayesian tutorial. OUP Oxford. http://aprsa.villanova.edu/files/sivia.pdf

(2) **For Machine Learning and Deep Learning**

- Goodfellow, I., Bengio, Y. and Courville, A. (2016). Deep Learning, MIT Press. https://www.deeplearningbook.org/
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: data mining, inference, and prediction. Springer Science & Business Media. https://web.stanford.edu/~hastie/ElemStatLearn/
- Murphy, K. P. (2012). Machine learning: a probabilistic perspective. MIT press. https://www.cs.ubc.ca/~murphyk/MLbook/
- Aston Zhang, Zack C. Lipton, Mu Li, Alex J. Smola (2020), Dive into Deep Learning - an interactive deep learning book with code, math, and discussions https://d2l.ai/
- Eli Stevens and Luca Antiga (2019), Deep Learning with PyTorch - Essential Excerpts, Manning Publications https://www.manning.com/books/deep-learning-with-pytorch#toc

(3) **For Python progamming**

- Shaw, Z. A. (2017). Learn python 3 the hard way: A very simple introduction to the terrifyingly beautiful world of computers and code. Addison-Wesley Professional.
- Scipy Lecture Notes - One document to learn numerics, science, and data with Python https://scipy-lectures.org/
- McKinney, W. (2017). Python for data analysis: Data wrangling with Pandas, NumPy, and IPython. O'Reilly Media, Inc.

https://www.oreilly.com/library/view/python-for-data/9781491957653/
https://github.com/wesm/pydata-book

(4) **For scientific computing**

- Smith, R. C. (2013). Uncertainty quantification: theory, implementation, and applications (Vol. 12). SIAM.
  https://rsmith.math.ncsu.edu/UQ_TIA/
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., ... & Tarantola, S. (2008). Global sensitivity analysis: the primer. John Wiley & Sons.

9 Others

(1) **Late Assignments Policy**

- Unless stated otherwise, assignments are due by the end of day on the designated due date. Assignments turned in within 24 hours of this time are considered late and will be assessed a 25% penalty. Assignments turned in after 24 hours will be marked and returned to the student, but no credit will be assigned.
- To allow for unforeseen circumstances, each student will be granted a **one-time exemption**. The student should contact the course instructor **at least three days** before the original deadline and explain the situation in order to get an extension. The assignment must be turned in by the end of the new designated due date.

(2) **Course Attendance/Absence Policy**

- Required; Active class participation is strongly encouraged.
- NC State attendance policies can be found at: REG 02.20.03 - Attendance Regulations - Policies, Regulations & Rules (https://policies.ncsu.edu/regulation/reg-02-20-03-attendance-regulations/). Please refer to the course's attendance, absence, and deadline policies for additional details.
- *Absences Policy*: Personal Problems: We understand that sometimes life makes it difficult to focus on schoolwork. If you are having a personal problem that affects your participation in this course, please talk to us to create a plan. Please do not wait until the end of the semester to share any challenges that have negatively impacted your engagement and academic performance. The sooner we connect, the more options we will have available to us to support your overall academic success. If you are not comfortable speaking with us directly, please utilize the other student resources provided below in order to understand how to best approach success in this course given your personal needs as soon as possible.

(3) **Transportation**

- This course will not require students to provide their own transportation. Non-scheduled class time for field trips or out-of-class activities is NOT required for this class.

(4) **Safety & Risk Assumptions**

- N/A

(5) **Digital Course Components**

- This on campus course will be captured and distributed via the Internet and/or electronic media as part of the Engineering Online (EOL) program for the distance students.
- These video recordings may contain an image of you entering the classroom, asking a question or being a part of the studio class.
- Please notify Dr. Linda Krute, Director of EOL, at ldkrute@ncsu.edu if you do NOT want your image to be included in the lecture presentation. If EOL does not hear from you after the first week of the class, we will assume that you are in agreement with this procedure.
- Students may be required to disclose personally identifiable information to other students in the course, via digital tools, such as email or web-postings, where relevant to the course. Examples include online discussions of class topics, and posting of student coursework. All students are expected to respect the privacy of each other by not sharing or using such information outside the course.

(6) **Academic Integrity**

- Students are required to comply with the university policy on academic integrity found in the Code of Student Conduct found at https://policies.ncsu.edu/policy/pol-11-35-01/
- Absolutely no collaboration is permitted during closed-book tests. All the tests are closed book unless otherwise specified.
- Collaboration on homework assignments is allowed, but the submitted work must be your own individual work. Homework assignments must not be treated as group assignments. Zero grade will be assigned for particular homework for the first offense. Second offense will be reported to the Office of Student Conduct.
- Violations of academic integrity will be handled in accordance with the Student Discipline Procedures (NCSU REG 11.35.02) at https://policies.ncsu.edu/regulation/reg-11-35-02/. 11.35.02)

(7) **Additional NC State Rules and Regulations**

- Students are responsible for reviewing the NC State University Policies, Rules, and Regulations (PRRs) which pertain to their course rights and responsibilities, including those referenced both below and above in this syllabus:
  - Equal Opportunity and Non-Discrimination Policy Statement https://policies.ncsu.edu/policy/pol-04-25-05/ with additional references at https://oied.ncsu.edu/divweb/policies/.
  - Code of Student Conduct Policy https://policies.ncsu.edu/policy/pol-11-35-01/

(8) **Use of Electronic Devices in Class**

- Cell phones are to be turned OFF prior to entering the classroom/lab. No exceptions.
- Use of laptops/other electronic devices during class is permitted only for the purpose of following the posted lecture materials/taking electronic notes.

(9) **Accommodations for Students with Disabilities**

- Reasonable accommodations will be made for students with verifiable disabilities. In order to take advantage of available accommodations, students must register with Disability Services for Students at 1900 Student Health Center, Campus Box 7509, 515-7653.
- For more information on NC State's policy on working with students with disabilities, please see the "REG 02.20.01 - Academic Accommodations for Students with Disabilities" at https://policies.ncsu.edu/regulation/reg-02-20-01/.

(10) **Non-Discrimination Policy**

- NC State provides equal opportunity and affirmative action efforts, and prohibits all forms of unlawful discrimination, harassment, and retaliation ("Prohibited Conduct") that are based upon a person's race, color, religion, sex (including pregnancy), national origin, age (40 or older), disability, gender identity, genetic information, sexual orientation, or veteran status (individually and collectively, "Protected Status").
- Additional information as to each Protected Status is included in NCSU REG 04.25.02 (Discrimination, Harassment and Retaliation Complaint Procedure). NC State's policies and regulations covering discrimination, harassment, and retaliation may be accessed at http://policies.ncsu.edu/policy/pol-04-25-05 or https://oied.ncsu.edu/divweb/.
- Any person who feels that he or she has been the subject of prohibited discrimination, harassment, or retaliation should contact the Office for Equal Opportunity (OEO) at 919-515-3148.

(11) **Student Mental Health**

- As a student you may experience a range of personal issues that can impede learning, such as strained relationships, increased anxiety, alcohol/drug concerns, feeling down, difficulty concentrating and/or lack of motivation. These mental health concerns or stressful events may lead to diminished academic performance and may impact your ability to participate in daily activities. It is very important that you have a support system and that you ask for help when you are struggling. The Counseling Center at NC State offers confidential mental health services for full time NC State students, including same-day emergency services. Please visit https://counseling.dasa.ncsu.edu/ to get connected.