APPLIED PHYSICAL SCIENCES

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

APPL 350
Data Science for Applied Science & Engineering
Spring 2021

Introduction	This course is intended as a first introduction to a variety of mathematical, statistical and computational methods that are of particular interest for analyzing data and performing machine learning across applied science and engineering, including visualization, transforms for time series and image analysis, dimensionality reduction, clustering, and various forms of classification (e.g., logistic regression, support vector machines, neural networks). Course activities will emphasize the ability to perform these data analyses in a computational environment and the written and oral communication of results.
Methods	This class will not be conducted like a traditional science lecture. Participation will be essential to success in this course. The approach of this class will require you to actively perform example analyses and discuss these during the class meeting period and/or in the asynchronous written collaboration environment. To do this effectively, you must engage fully with the assigned reading and/or videos prior to each class meeting. Homework and class presentations will be followed by discussions in which full class participation will be expected. Written computational assignments must be completed in the form of literate programming, combining extensive descriptive narrative text with embedded code and output (figures and tables) as produced in MATLAB Live Editor. (While similar functionality is also available in e.g., Jupyter and R Notebook, but this class will be conducted in MATLAB to be consistent with other courses in the APPL curriculum.) Paying attention to effective written and oral communication is essential to convey the value of any data analysis and, as such, grades will be as much about the quality of the communication as that of the computation.
Results	By the end of this course, students should be able to demonstrate greater expertise and experience performing a variety of data analyses, critically assess data analyses for appropriateness and possible overfitting, clearly communicate the processes and results of their data analyses.
Discussion	I ask that you come with a mindset and demonstrate a willingness to try new things, fail forward, learn from the experiences, refine your efforts, and communicate your experiences.
Engineering Student Outcomes	• Demonstrate constant curiosity about our changing world • Explore a contrarian view of accepted solutions • Integrate information from many sources to gain insight • Persist through and learn from failure • An ability to communicate effectively with a range of audiences • An ability to develop and conduct appropriate experimentation, analyze and interpret data, and use engineering judgement to draw conclusions • An ability to acquire and apply new knowledge as needed, using appropriate learning strategies

Class Essentials

CONTACT INFORMATION

Dr. (Insert name)	Teaching assistants
Office Location (Insert office #) **Email** (Insert email) **Phone** (Insert phone)	(Names and contact information about TAs or mark "not applicable")

LOGISTICS

Class meeting times
(Insert days/times)

Class meeting location
(Insert location)

Office Hours
(Insert days/times) or by appointment

Required Texts & Software

- Data-Driven Science and Engineering by Steven L. Brunton and J. Nathan Kutz (full text available online via UNC-Chapel Hill Libraries)

Pre-requisites

- APPL 101 or COMP 116 or permission of instructor

Course content

COURSE TOPICS

- Visualization
- Transforms for time series and image analysis
- Dimensionality reduction, including SVD and PCA
- Linear and logistic regression
- Unsupervised clustering
- Supervised classification
    - Support vector machines
    - Random forests
    - Neural networks

Course Schedule

Week 1: Mathematical preliminaries
Week 2: Singular Value Decomposition (SVD) and Principal Component Analysis (PCA)
Week 3: Truncation, Randomized SVD, Tensor Decompositions
Week 4: Fourier Series, Fourier Transforms, DFTs and FFTs
Week 5: Wavelet Transforms, *First Midterm Exam*
Week 6: Sparsity, Compression and Compressed Sensing
Week 7: Matrix completion and Robust PCA
Week 8: Curve fitting and Regression
Week 9: Logistic regression, Optimization, Model Selection
Week 10: Unsupervised Clustering
Week 11: Introduce Supervised Learning, *Second Midterm Exam*
Week 12: Linear Discriminants, SVMs and Random Forests
Week 13: Neural Networks
Week 14: Deep Convolutional Neural Networks

To help you succeed

Course Expectations and Policies

- Watch all online lectures and do all reading assignments before coming to class
- Participate in discussions (in-class and asynchronous) and problem-solving activities.
- During class time, do not use your phone or computer for something unrelated to class; research shows that this is distracting to other students in the class. If there is an urgent situation, then you can leave the classroom to use your phone or computer.
- Come to every scheduled class and lab session and let me know ahead of time if you cannot attend.
- Turn in assignments on time; if an assignment is up to 24 hours late, there is a 25% deduction, and if an assignment is beyond 24 hours late, you will get a zero. If you need an extension, you must ask at least 24 hours before the time that the assignment is due (you can avoid a grade deduction this way).

Student Resources

SEE, SAY, DO SOMETHING

We're happy you are here and eager to learn. Despite our best intentions to follow a plan, life may throw us a curve ball. If you or someone you know is experiencing some distress or you are concerned about the well-being of a student, please report it here: https://deanofstudents.unc.edu/carereport. It is important to support one another. If you see something, say, and do something.

ACCESSIBILITY RESOURCES

UNC-CH provides accommodations for any students with documented disabilities. If you have a disability and believe you require accommodations, please contact the Department of Accessibility Resources at http://accessibility.unc.edu. Please contact me early in the semester so we can make any necessary arrangements and discuss the learning checks.

Assignments & Evaluation

➢ Your Course Grade

40%: Homework and Projects	Weekly submission of .mlx files that must be executable in MATLAB Live Editor
20%: First midterm exam	Open-book, open-internet, timed challenge exercise to be completed and described in MATLAB Live Editor
20%: Second midterm exam	Open-book, open-internet, timed challenge exercise to be completed and described in MATLAB Live Editor
20%: Final exam	Open-book, open-internet, timed challenge exercise to be completed and described in MATLAB Live Editor, to be held during the final exam period
100%: total	

➢ Grade Interpretation & Honor Code

Your final course grade will be determined from a standard scale:

A	93+
A-	90.0 - 92.9
B+	87.0 - 89.9
B	83.0 - 86.9
B-	80.0 - 82.9
C+	77 - 79.9
C	73 - 76.9
C-	70 - 72.9
D+	67 - 69.9
D	60 - 66.9
F	<60

ACADEMIC HONESTY

Collaboration is encouraged on homework and project assignments, but ultimately the material submitted must represent your understanding of the material as developed through those collaborations. In particular, it is a violation of the honor code if a student duplicates work or obtains solutions from another student and submits it as their own.

Unlike the homework, collaboration is not allowed at all on any exams.

Please reference the honor code: http://honor.unc.edu.

➢ Major Course Due Dates

Exam 1	**(Date)**
Exam 2	**(Date)**
Final Project	**(Date)**

I reserve to right to make changes to the syllabus, including due dates and test dates (excluding the officially scheduled final examination), when unforeseen circumstances occur. These changes will be announced as early as possible so that students can adjust their schedules.