EEO 388: Foundations of Machine Learning

Fall 2024

Catalog Description:

This course provides an introduction to the fundamental concepts of machine learning. Statistical learning framework is utilized for clustering, classification, and prediction tasks. Concepts are reinforced through theoretical and programming assignments, with applications in computer vision, natural language processing and bioinformatics.

Course Prerequisites: EEO 224 and EEO 306

Course Credits: 3

Faculty: Vibha Mane
Email: vibha.mane@stonybrook.edu

Office Hours: Conducted online via Zoom; details posted on Brightspace.

Course Delivery: Online, asynchronous; course material posted on Brightspace.

Final Exam: Online, synchronous, via Zoom; details posted on Brightspace.

Textbooks: Gareth James et al., An Introduction to Statistical Learning, Springer, 2017.
EMC Education Services, Data Science and Big Data Analytics, Wiley, 2015

Course Dates and Duration: August 26 - December 19.

Grading

Assignments (Theory & Programming)	60
Exercise Sets	12
DB Participation*	12
Final Exam **(Synchronous)**	16

*Refer to the document “DBRubric.pdf”.

Course Schedule

- The schedule for Learning Modules, exercise sets, assignment submission and exam dates are posted in a file shared with the class.

Additional Notes

- There will be a synchronous Final Exam. All students are expected to attend.
- In the event you miss an exam due to significant illness or any other personal emergency, alternative accommodation will be made. However, it is your responsibility to provide documentation to support such a request.
- There are no makeup exams to boost scores.
- There are no extra credits.
- No resubmission of assignments is allowed.
- Any grade dispute should be brought to the attention of the instructor within one week of posting the grades.
- Stony Brook University makes religious accommodations. Please check the university religious holiday calendar for a list of major holidays and policies.

Course Learning Outcomes: Upon completion of the course, students will

- Develop an understanding of the fundamental concepts of machine learning and problem formulation.
- Acquire knowledge of statistical learning framework for performing tasks such as clustering, classification, and prediction.
- Reinforce the above-mentioned concepts through theoretical and programming assignments.

Topics Covered:
The course has 6 Learning Modules, with topics as described below. The course entails extensive programming in Python. Tutorials and example codes are provided in each module.

Module 1a	**Python Tutorial:** Python data structures - Array, Series, DataFrame; data visualization with Seaborn. **ML Overview and Exploratory Data Analysis:** Types of machine learning tasks and models; structured and unstructured data; scatter, histogram & density plots.
Module 1b	**Unsupervised Learning and Clustering:** k-means, hierarchical and Dbscan clustering; distance metrics.
Module 2a	**Probability Review:** Probability basics; discrete and continuous distributions; joint random variables; multivariate normal distribution. **Bayesian Learning:** Bayes Theorem, Naïve Bayes classifier.

Module 2b	**Resampling Methods:** Cross-validation, underfitting & overfitting, bias-variance tradeoff, bootstrap method. **Principal Component Analysis:** Feature reduction with principal component analysis.
Module 3a	**Tree-based Methods for Classification:** Regression & decision trees, information gain, bagging, boosting & random forests.
Module 3b	**Support Vector Machine for Classification:** Support vector classifier, slack variables, kernels & support vector machine.

University Policies

Academic Integrity Statement

Each student must pursue their academic goals honestly and be personally accountable for all submitted work. Representing another person's work as your own is always wrong. Faculty is required to report any suspected instances of academic dishonesty to the Academic Judiciary. Faculty in the Health Sciences Center (School of Health Technology & Management, Nursing, Social Welfare, Dental Medicine) and School of Medicine are required to follow their school-specific procedures. For more comprehensive information on academic integrity, including categories of academic dishonesty please refer to the academic judiciary website at http://www.stonybrook.edu/commcms/academic_integrity/index.html

Student Accessibility Support Statement

If you have a physical, psychological, medical, or learning disability that may impact your course work, please contact the Student Accessibility Support Center, Stony Brook Union Suite 107, (631) 632-6748, or at sasc@stonybrook.edu. They will determine with you what accommodation is necessary and appropriate. All information and documentation are confidential. Students who require assistance during emergency evacuation are encouraged to discuss their needs with their professors and the Student Accessibility Support Center. For procedures and information go to the following website: https://ehs.stonybrook.edu//programs/fire-safety/emergency-evacuation/evacuation-guide-disabilities and search Fire Safety and Evacuation and Disabilities.

Critical Incident Management Statement

Stony Brook University expects students to respect the rights, privileges, and property of other people. Faculty are required to report to the Office of Student Conduct and Community Standards any disruptive behavior that interrupts their ability to teach, compromises the safety of the learning environment, or inhibits students' ability to learn. Faculty in the HSC Schools and the School of Medicine are required to follow their school-specific procedures. Further information about most academic matters can be found in the Undergraduate Bulletin, the Undergraduate Class Schedule, and the Faculty-Employee Handbook.

Electronic Communication

Email and especially email sent via Brightspace is one of the ways the faculty officially communicates with you for this course. It is your responsibility to make sure that you read your email in your official University email account. For most students that is Google Apps for education (http://www.stonybrook.edu/mycloud).

If you choose to forward your official University email to your other off-campus account, faculty are not responsible for any undeliverable messages to your alternative personal accounts. You can setup Google mail forwarding using these DoIT-provided Instructions found at http://it.stonybrook.edu/help/kb/seeting-up-mail-forwarding-in-google-mail.

If you need technical assistance, please contact Client Support at 631-632-9800 or supportteam@stonybrook.edu.

Course Materials and Copyright Statement

Course material accessed from Brightspace, Zoom, Echo 360, VoiceThread, etc. is for the exclusive use of students who are currently enrolled in the course. Content from these systems cannot be reused or distributed without written permission of the instructor and/or the copyright holder. Duplication of materials protected by copyright, without permission of the copyright holder is a violation of the Federal copyright law, as well as a violation of Stony Brook's Academic Integrity.

Additional Resources

- To access mental health services, call Counseling and Psychological Services (CAPS) at 631-632-6720; Counselors are available to speak with 24/7.
- For updated information on the Academic Success and Tutoring Center (ASTC), please check www.stonybrook.edu/tutoring for the most up-to-date information.
- For IT Support: Students can visit the Keep Learning website at https://sites.google.com/stonybrook.edu/keeplearning for information on the tools you need for alternative and online learning.
- Need help? Report technical issues at https://it.stonybrook.edu/services/itsm or call 631-632-2358.
- For information on Library services and resources, please visit https://library.stonybrook.edu