EAS 509 Course Syllabus
Statistical Learning and Data Mining-II
Fall 2022

Instructor Dr. Nikolay Simakov
Email: nikolays@buffalo.edu
Classes:

- Wednesday, Friday, 4:40PM - 5:50PM, Knox 14
- November 16, no in-person class (recorded version will be available)

Office Hour: Monday 5.30-6.30 PM

- In zoom: available on request
  - Meeting URL: https://buffalo.zoom.us/j/91738672279?pwd=bEFGSFhEYnBqQ2Z5MEYya3dWbVhlUT09
  - Meeting ID: 917 3867 2279
  - Passcode: 163614
- In person: available on request
  - Furnas 805

Communication with Instructional Team:

- In addition to Zoom and in-person meetings, students can communicate with the Instructional Team via email.
- Email: The subject line of all emails to anyone on the instructional team must begin with [EAS 509] followed by description of the topic (e.g. [EAS 509] Question Regarding HW 2 is a valid subject line, [EAS 509] or Question Regarding HW 2 are not valid subject lines). Any emails which do not contain a properly formatted subject line may not be answered.

Teaching Assistant

Raktim Mukhopadhyay
Email: raktimmu@buffalo.edu

Course Description:

- An introduction to the mathematical theory and computational methodology at the heart of statistical learning and data mining.
- Mostly discuss unsupervised learning, time series analysis, Survival Analysis, Change Point
- R programming will be used.
- Theoretical Underpinnings, Mathematical fundamentals, and hands on application will be covered this course
- Will conflate the traditional course design and applicable research

Prerequisites: EAS 508: Statistical Learning and Data Mining-I, Understanding of Calculus, Introductory Linear Algebra, Probability Theory, Familiarity with programming

Textbook (Available On-line):

- **An Introduction to Statistical Learning** 2nd Edition, by Gareth James, Daniela Witten, Trevor Hastie. https://www.statlearning.com/
- (Advanced) **The Elements of Statistical Learning** Data Mining, Inference, and Prediction, by Trevor Hastie, Robert Tibshirani, Jerome Friedman. https://hastie.su.domains/ElemStatLearn/
- **Forecasting: Principles and Practice** (3rd ed). Rob J Hyndman and George Athanasopoulos. https://otexts.com/fpp3/

UB School Calendar:

Monday, Aug. 29, 2022	Classes Begin
Monday, Sept. 5, 2022	Labor Day Observed
Wednesday, Nov. 23-Saturday, Nov. 26, 2022	Fall Recess
Monday, Nov. 28, 2022	Classes Resume
Friday, Dec. 9, 2022	Last Day of Classes
Saturday, Dec. 10-Sunday, Dec. 11, 2022	Reading Days
Monday, Dec. 12-Monday, Dec. 19, 2022	Semester Final Examinations
Tuesday, Dec. 20, 2022	Winter Recess Begins

Exams:

Exam	Date	Format
Mid Term	October 14 or 19	Closed book, in class (Clustering + PCA + Some Time Series)
Final Exam	Somewhere between, Dec. 10, 2022 and Dec. 20, 2022	Either in class or whole day at home

Tentative Course Design:

EAS 509: Statistical Learning and Data Mining-II	
Week	**Topic**
Week 01- 05	• Syllabus and Course Design • Interactive Data Visualization using R • Basic Introduction to Shiny • Clustering • K-Means Clustering • Hierarchical Clustering • Clustering Techniques Comparison • PCA • Kernel PCA

Week 06- 10	• Time Series Analysis • Working with Dates and Time using R • Time Series (Data Preprocessing and Visualization) • Time Series Forecasting • ARIMA Models • Multivariate Time Series Analysis
Week 12-15	• Survival Analysis • Survival and Censoring Times, Kaplan-Meier Survival Curve, Log-Rank Test, Regression Models with a Survival Response • Change Point Detection • Association Rule Learning • Group 3 Presentation

Homeworks:

In general homework is due in a week after the topic is complete in lectures.

#	**Topic**	**Due Date**
0 (Optional-Extra Points)	Getting Familiar with UBLearn	09/06/2022
1	Interactive Visualization	~09/16/2022
2	Clustering	~09/28/2022
3	PCA	~10/7/2022
4	Time Series	~10/14/2022
5	Survival Analysis	~11/2/2022
6	Change Point Detection	~11/11/2022
7	Association Rules	~11/22/2022

Group Project:

- In depth topic of choice study using techniques from this course. Additionally, data descriptive statistics and visualization.
- 4 people in each group
    - Self-enroll to group in ublearn->tools->group
- No repetition of topics
    - first submission has priority.

Events	Due/Date
Group Formulation	09/30
Proposal Submission	10/28
Project progress report	10
Project Presentation	12/2 12/7 12/9
Project Document Submission	12/9/2022

Grade Policy:

Grades will be based on grades accumulated from the course requirements. Each graded entity is normalized to one by dividing received points by total possible points. Six top homework normalized grades are averaged. The final grade is calculated as weighted average using following weights:

Course Requirement	**Percent of Final Grade**
Homework 0 (extra point)	1%
Homework 1 to 7 (average of top 6)	40%
Midterm	20%
Final	25%
Group Project	15%

Exact cutoffs for specific grades will depend on the level of difficulty of exams and assignments. These cutoffs will be determined once Final Exam has been graded. However, the cutoffs are roughly the following:

Percentage	**Final Grade**
93	A
89	A-
85	B+
81	B
77	B-
73	C+
69	C
65	C-
60	D
< 60	F

In certain cases, students may be eligible to receive a temporary incomplete ("I") grade. Students may only be given an "I" grade if they have a passing average in course work that has been completed and establish well-defined parameters to complete course requirements. Prior to the end of the semester, students must initiate the request for an "I" grade and receive the instructor's approval.

Expectations of Students:

- Students are expected to act in a professional manner. A student's grade may be reduced due to unprofessional or disruptive behavior.
- Assignments are due at the assigned date and time.
- At the discretion of the instructor, late submission of assignments may receive a grade of zero (50% off on each unexcused delay day).
- Students are encouraged to discuss homework and share ideas, but each student must independently write and submit their own solution.
- Projects will be group projects. Four Students make a group. Group projects detail will be discussed in the class.

- Makeup exams will be given in the following circumstances only:
  1. You contact the instructor prior to the exam, and
  2. You have a valid and documented reason to miss the exam ○ Serious illness or family emergency are acceptable reasons
  ○ Sleeping in, lack of preparation, ennui, grogginess, etc. are not acceptable excuses

Accessibility Services and Special Needs:
If you have a disability and may require some type of instructional and/or examination accommodation, please inform your instructor early in the semester so that we can coordinate the accommodations you may need. If you have not already done so, please contact the Office of Accessibility Services (formerly the Office of Disability Services) University at Buffalo, 25 Capen Hall, Buffalo, NY 14260-1632; email: stuaccessibility@buffalo.edu Phone: 716-645-2608 (voice); 716-645-2616 (TTY); Fax: 716-645-3116; and on the web at http://www.buffalo.edu/accessibility/. All information and documentation is confidential. The
University at Buffalo and the School of Engineering and Applied Sciences are committed to ensuring equal opportunity for persons with special needs to participate in and benefit from all of its programs, services and activities.

Mental and Behavioral Health:

The personal safety and well-being of each and every student is vastly more important than any assignment or grade in a course. Students experiencing stress, depression, anxiety, fear, grief, relationship difficulties, eating problems, substance abuse, or other personal or social problems are strongly encouraged to contact the Student Life Counseling Services at 716-645-2720 for help. Counseling sessions are private, confidential, and free to all undergraduate and graduate students. In a crisis, contact the Erie County 24-hour crisis hotline at 716-834-3131 or text \GOT5" to 741-741. (Data usage while texting the Crisis Text Line is free and the number will not appear on a phone bill.)

Academic Integrity:
Academic integrity is a fundamental university value. Through the honest completion of academic work, students sustain the integrity of the university while facilitating the university's imperative for the transmission of knowledge and culture based upon the generation of new and innovative ideas. The UB graduate academic integrity policy is available at http://www.buffalo.edu/grad/succeed/currentstudents/policy-library.html. The sharing or posting of material related to EAS 508, and requests for help or answers on any electronic tool, such as Chegg and Course Hero, is considered a violation of the academic integrity policy. Consequences of violating the academic integrity policy may range from zero points to expulsion from the university, depending on the seriousness of the violation.