Applied Machine Learning
CSCI 490-0J13/680-0K13

Spring 2021

Instructor:	Pratool Bharti	**Time:**	Tue/Thu 3:30 - 4:45p
Email:	pbharti@niu.edu	**Place:**	Online

Office Hours: Tue/Thu (4:45 - 5:45p)

Descriptions/ Objectives: This course offers a hands-on approach to machine learning and data science topics. In this class, we will discuss the application of machine learning methods like SVMs, Random Forests, Gradient Boosting and Neural Networks on real world dataset, including data preparation, model selection and evaluation. We will use Scikit-learn and Keras/Tensorflow tools and libraries in Python for all the implementations. Apart from applying models, we will also discuss software development tools and practices relevant to productionizing machine learning models.

Prerequisites: A good understanding of Probability, Calculus and Linear Algebra would be very helpful for this course. Familiarity with Python programming and basic use of NumPy, Pandas and Matplotlib is expected. Students enrolled for Python course in Spring 2021 may enroll for this course simultaneously. A good reference is the Python Data Science Handbook by Jake VanderPlas. It's online for free and available as a notebook at the link above. I highly recommend going through it.

Programming Platform: Students are required to create a free account in Google Colab (`https://colab.research.google.com`). All programming assignments are required to be done using Jupyter Notebook on Google Colab server. Details will be discussed in the first class.

Recommended References: Students need to consult following recommended book regularly. Online version is available free of cost. Other required materials will be posted on Blackboard.

- Muller C. A and Guido S, *Introduction to Machine Learning with Python: A Guide for Data Scientists*, An O'Reilly Media publication; 1st edition (October 21, 2016)

- Kuhn M and Johnson K, *Applied Predictive Modeling*, Springer; 1st ed. 2013, Corr. 2nd printing 2018 edition (March 30, 2018)

- Goodfellow I, Bengio Y and Courville A, *Deep Learning*, An MIT Press book, 2016,
  Online: - http://www.deeplearningbook.org.

Grading Policy: 5/6 Assignments (60%), Midterm 1 (20%), Midterm 2 (20%), Quizzes (Extra credits).

Grade	Average
A	≥ 90
B	≥ 80
C	≥ 70
D	≥ 60
F	< 60

Important Dates:

Midterm #1 . . . . March 09, 2021
Midterm #2 . . . . April 22, 2021

Course related Announcements: All course related materials and announcements will be posted on Blackboard. Students should login at least once everyday in the Blackboard.

Collaboration: You are welcome and encouraged to discuss course materials and reading assignments with other students. Please limit discussion of homework to general approaches. You are not allowed to share code between submissions or submission groups. For homeworks submitted individually, each individual is required to write their own solution. Collaboration is not permitted for any of the exams.

Assignments Policy: There will be approximately 5-6 individual assignments. Most or all of the assignments will be well-structured Python programs and written reports. Graduate students are required to do an additional assignment/ questions. There will be likely no extensions. The maximum possible points that will be awarded will decrease by 10% of the original maximum point value per day that they are late. Maximum late submission allowed is 48 hours. Assignments submitted later than 48 hours will be assigned 0 score.

Attendance Policy: Regular attendance is essential and expected.

Academic Integrity: Academic dishonesty is strictly prohibited. You are expected to do your own work on the assignments, programs and exams. Cheating includes, but is not limited to, copying work from other students, from other textbooks, from the Internet, or allowing others to do the same, without proper citation of source material. Please follow the link for more information https://www.niu.edu/academic-integrity/students/plagiarism/index.shtml. Lack of knowledge of the academic honesty policy is not a reasonable explanation for a violation.

Special circumstances: Students with special needs (disability accommodation, religious observances, required military service, major illness or other unexpected events) are encouraged to contact the instructor as soon as possible. Having a lot of work for your other classes, being busy at your job, and network problems are not special circumstances; they are normal circumstances that everyone has.

Disability accommodations: The instructor will provide all of the accommodations to which you are entitled by law. If you need an accommodation for this class, you must provide a notification letter from the Disability Resource Center. If you wish to take your exams at the DRC office, you must also follow DRC regulations with regard to exam scheduling. Therefore you should contact the DRC as soon as possible. They are located on the Campus Life Building, Suite 180, and can be reached at 815-753-1303 or drc@niu.edu. Also, please contact me privately as soon as possible to discuss possible accommodations. Please note that you will not be required to disclose your disability, only your accommodations. The sooner you let me know your needs, the sooner I can assist you in achieving your learning goals in this course.

Tentative Course Outline:

Week	Date	Class Topics
1	01/12	Introduction, Course Overview
	01/14	Math Essentials for Machine Learning
2	01/19	Matplotlib and Visualization
	01/21	Introduction to Supervised Learning

3	01/26	Introduction to Supervised Learning (continue)
	01/28	Data preprocessing
4	02/02	Linear Models for Regression
	02/04	Linear models for Classification, SVMs
5	02/09	Trees, Forests, Gradient Boosting & Ensembles
	02/11	University Closed
6	02/16	Model Evaluation
	02/18	Learning with Imbalanced Data
7	02/23	Model Interpretation and Feature Selection
	02/25	Model Interpretation and Feature Selection (continue)
8	03/02	Parameter tuning and Automatic Machine Learning
	03/04	Revision Day
9	03/09	First Mid-term Exam
	03/11	Mid-term Discussion
10	03/16	Dimensionality Reduction
	03/18	Clustering and mixture models
11	03/23	NMF and Outlier Detection
	03/25	Working with text data
12	03/30	Topic models for text data
	04/01	Word and document embeddings
13	04/06	Word and document embeddings (continue)
	04/08	Neural Networks
14	04/13	Keras and Convolutional Neural Nets
	04/15	Advanced Neural Networks
15	04/20	Revision Day
	04/22	Second Mid-term Exam