COMP740/COMP840
Machine Learning Applications and Tools
Course Syllabus *

Course Information

Credits:	COMP740 4 credits / COMP840 3 credits
Term:	Fall 2020
Time:	Thursday 1-3:50pm
Location:	Remote through Zoom URL: https://unh.zoom.us/j/96229512239

Instructor Information

Name:	Karen Jin
Office:	Rm 139, Pandora Mill building
Email:	karen.jin@unh.edu

To make an appointment with me in my virtual Zoom office: https://calendly.com/karenajin You may also email to arrange a time to meet other than the slots listed.

Course Description

Introduces students to practical approaches of machine learning. The course is an exploration of creative applications of artificial intelligence using modern machine learning components and tools, including deep learning techniques. Different application domains are considered, such as computer vision, natural language processing, and cyber security. Students learn to evaluate the effectiveness of machine learning systems as well as their potential prediction problems.

Course Objectives

Upon completion of this course, students should be able to:

- Learn how to promote yourself using tools such as Github, Juypter Notebook and Kaggle.
- Use Python's essential libraries/tools including pandas, NumPy, scikit-learn and Matplotlib.
- Learn how to use tools such as Tensorflow to build Machine Learning models and deploy them on Cloud ML Engine.
- Learn to apply Machine Learning techniques to both well-known data sets and new datasets with supervised learning and unsupervised learning.
- Create a reproducible portfolio that highlights your ability to perform as a Machine Learning practitioner.

Recommended Text:

"Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow", 2nd Edition, by Sebastian Raschka and Vahid Mirjalili

Github: https://github.com/rasbt/python-machine-learning-book-2nd-edition

Kindle version: https://www.amazon.com/dp/B0742K7HYF?ref_=kcr_store_sample

"Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems" by Aurélien Géron, O'Reilly, 2017
Notebooks available from https://github.com/ageron

"Python Data Science Handbook: Essential Tools for Working with Data" by Jake VanderPlas, O'Reilly, 2016
Text and notebooks available from https://github.com/jakevdp/PythonDataScienceHandbook

Weekly Schedule:

Week 1 (August 31): Introduction to ML. Supervised v.s. Unsupervised learning
Week 2 (September 7): Training simple ML classification model
Week 3 (September 14): ML Classifiers
Week 4 (September 21): Training Datasets
Week 5 (September 28): Model Evaluation
Week 6 (October 5): Text Processing and Sentiment Analysis
Week 7 (October 12): Embedding ML Model into a Web Application
Week 8 (October 19): Introduction to Artificial Neural Network
Week 9 (October 26): Google Cloud Platform Auto ML
Week 10 (November 2): Image Classification with Tensorflow I
Week 11 (November 9): Image Classification with Tensorflow II
Week 12 (November 16): Individual Project Presentation
Week 13 (November 23): Thanksgiving No Class
Week 14 (November 30): Group Project Check-in
Week 15 (December 7): Final Exam
Week 16 (December 14): Group Project Presentation

Grading:

- **Homework: 25%**
  - Weekly assigned (approximately)
  - All homework assignments are individual assignments.
  - Each homework is due at the date and time listed. It is your responsibility to allow enough time during the week to complete the activities despite obstacles (technology driven or personal) that may arise. Late work is anything submitted after the announced due date and time. You will lose 10% per day for work submitted late for up to 48 hours after the due day. Submissions must be completed through MyCourses in order to receive credits.

- **Class Participation: 15%**

The participation is calculated based on 1) both the quantity and quality of the discussion posts in the Slack discussion forum, and 2) individual check-in meetings with the instructor. Each check-in meeting is 15-20 mins long.
To earn full points, you must fulfill the discussion requirements for the discussion topic assigned, and complete at least three check-in meetings during the semester.
Desired attributes for a discussion posting.

- Ability to synthesize the main concepts from instructor, course content, external resources and class community.
- Use of proper grammar.
- Ideas are organized, persuasive and elevate the overall dialogue.
- Opinions are substantiated.
- Demonstration of critical or creative thinking.
- Evidence of preparation.

- **Individual Implementation Project: 20%**

  Machine Learning
  - You may use either classification or regression technique
  - Must use a real data set (not fake data set, i.e. randomly generated numbers)

  Juypter Notebook Documentation
  - Must explain in detail the data sets and the participation of training and testing sets
  - Include data exploration and visualization
  - Explain the metric used to evaluate the accuracy of model
  - Discuss the tradeoffs you made

  Presentation: Each student will give a 10-minute presentation in class. Presentations are scheduled on November 19**.**

- **Group Implementation Project: 25%**

  In this group project, you will practice with your concepts and skills in supervised learning tasks. You are asked to implement a web application that utilized machine learning models, either pre-trained or your own. Teams of two people will be formed. To receive full credit for the Group Project you must fulfill these requirements:

  Code base: Both members of the team must commit code on the project Github codebase and use the project management tool to keep track your progress.
  Machine Learning: You may use a pre-trained model or build a model on your own.
  Deployment:
  Build a web application on a simple framework such as Flask. Minimum requirements:
  For COMP740 students, the application should be deployed as a web app and hosted externally; for COMP840 students, the application should be containerized and deployed on a cloud service of your choice.
  Detailed Project Requirements TBA

Presentation: Each group will give a 10-15 minute presentation in class. Presentations are scheduled on during last week of class.

Use the following criteria in your individual and group projects:
- Reproducibility of the Project: Is every step reproducible, e.g., in Juypter Notebook?
- Creativity: Did you create new insights that you shared to the class?
- Craftsmanship: Did you take care in crafting your presentation and put thought into little details like thoughtful data visualization?
- Difficulty: Was this a challenging problem you tackled?
- ML Technique: How did you approach the problem?
- Software Engineering Carpentry: Did you apply software engineering best practices?
- Resume Worthiness: Will this project get you a job?

- **Final Exam: 15%**
  The final exam consist of two components:
  - 30 minutes 1-on-1 oral assessment through Zoom
  - Take home exam.
  Both tests will be conducted during the week of December 7.

COURSE POLICIES REGARDING STUDENT BEHAVIOR

Attendance
Students are responsible for attending scheduled meetings and expected to abide by the University Policy on Attendance (as stated in the *UNH Student Rights, Rules, and Responsibilities*). Students who miss a scheduled meeting have the responsibility to email instructor about the circumstances for missing the meeting within a week of the absence.

Late submissions
Policy for late submissions is very strict and applies only in exceptional cases of student illness, accident, or emergencies that are properly documented. A late submission of an homework or project artifact may be granted only if the student:
- Emails prior to the deadline and
- Explains and provides evidence for the circumstances that have prevented the student from meeting class requirement.

Failing to comply with these rules may result in no credit for the course.

STATEMENT ON ACADEMIC HONESTY

Please see below the Academic Honesty policy for the Computing Program.

1. Graded work in this class should be entirely yours and shall not include work done by others or obtained from external sources.
2. Collaboration on work is allowed only upon explicit instruction from the course instructor and only within the constraints given for that specific work. If unclear, you must consult with me on what is allowable. It is your responsibility to get such clarification.
3. Collaboration on all the work in this class is encouraged, unless explicitly stated that the assigned work must be individual work with no collaboration permitted.
4. Sharing the products of your work or your direct participation in doing work for others is also an instance of academic misconduct. Refrain from doing work on behalf of somebody else or from sharing your work products with other students.
5. If the preparation of your artifacts benefit from collaborations with peers, tutors, tech assistants, course instructor (that's me), or any other person (friend, relative, etc.), in cases in which collaboration is allowed, submitted artifacts must include clear attribution to the kind of beneficial collaboration.
6. If the preparation of your artifacts benefit from online sources (forums, public GitHub repos, tutorials, MOOCs, etc.), in cases in which collaboration is allowed, submitted artifacts must include clear attribution to the source and source content you have used.

Failing to comply with these rules is considered a violation of the academic honesty policy.

See the University Academic Policy at https://www.unh.edu/student-life/09-academic-honesty.

for more information.

There are very serious repercussions if you deviate from the academic honesty policy:

1. The penalty for the first occurrence of an instance of academic dishonesty and plagiarism in this course is no credit for the graded work in question. Program Coordinator and Department Chair will be notified of your misconduct, and a letter from the Program Coordinator will be sent to you, course instructor, and faculty advisor.
2. The second instance of academic dishonesty in this course is penalized with failing the course. Program Coordinator, Department Chair, and Associate Dean of UNH Manchester or Graduate School (depending on your enrollment) will be notified. A letter from the Program Coordinator will be sent to you, course instructor, and faculty advisor. The Dean may decide on dismissal from the University.
3. If the first instance of academic dishonesty in this course is your second one, the course action described at #2 above will be followed.

Bottom line, do not plagiarize, do not share your work with others, and not collaborate, when allowed, without giving proper attribution.

STUDENTS WITH DISABILITIES

UNH Manchester is committed to providing students with disabilities with a learning experience which assures them of equal access to all programs and facilities of the University, which makes all reasonable academic aids and adjustments for their disabilities and provides them with maximum independence and the full range of participation in all areas of life at UNH Manchester. Students who need to document their disability and determine any accommodations, services, or referrals should schedule an appointment with the UNH Manchester Disability Services Coordinator by calling 641-4170. For more information, please see http://manchester.unh.edu/student/disability.

* Subject to changes and revisions