UNIVERSITY OF MACAU
FACULTY OF SCIENCE AND TECHNOLOGY
DEPARTMENT OF COMPUTER AND INFORMATION SCIENCE
2024-2-CISC7401-001 ADVANCED MACHINE LEARNING
Syllabus
2nd Semester 2024/2025
Part A – Course Outline

Course Description:
How can a machine learn from experience, to become better at a given task? How can we automatically extract knowledge or make sense of massive quantities of data? These are the fundamental questions of machine learning. Machine learning algorithms use techniques from statistics, optimization, and computer science to create automated systems which can sift through large volumes of data at high speed to make predictions or decisions without human intervention. Machine learning as a field is now incredibly pervasive, with applications from the web (search, advertisements, and suggestions) to national security, from analyzing biochemical interactions to traffic and emissions to astrophysics. This is an Advanced Machine Learning course, taking 3 credits and lasting for 14 weeks.

Course type:
Theoretical

Prerequisites:
- Proficiency in Python programming.
- Linear algebra (e.g., matrix computations, eigenvalues and eigenvectors, singular value decomposition).
- Basic statistics and probability
- Multivariate calculus
- Algorithms

Preferably access to GPUs (e.g., Google Colab).

Recommended background
- **Proficiency in Python**. All class assignments will be in Python. For those who aren't as familiar with Python, please follow this tutorial. In the tutorial, go over all topics under "Python Tutorial", "Python NumPy". It will also be beneficial to go over all topics under "Machine Learning".
- **Familiar with Jupyter Notebook**. We will use Jupyter Notebook (https://jupyter.org) to implement algorithms, demonstrate experiment qualitative and quantitative results, and write up assignment report. Converting a Jupyter Notebook file to pdf / html / python script is trivial (https://nbconvert.readthedocs.io/en/latest/usage.html).
- **Familiar with PyTorch**. Some coding assignments involving deep learning will use PyTorch, which is an excellent python-based toolbox for machine learning and deep learning. For those who haven’t used it before, please refer to its official tutorial. In the tutorial, go over “Introduction” topics, “Learning PyTorch”, and “Image and Video”.
- **College Level Linear Algebra**. You should be comfortable taking derivatives and understanding matrix vector operations and notation. Go over the “Essence of linear algebra” playlists by “3Blue1Brown” at https://www.youtube.com/c/3blue1brown/playlists

- **Basic Probability and Statistics.** You should be familiar with basics of probabilities, Gaussian distributions, mean, standard deviation, etc. Go over the "Probabilities of probabilities" playlist by "3Blue1Brown" at https://www.youtube.com/c/3blue1brown/playlists

Textbook(s) and other required material:
There are no required textbooks. The following textbooks are useful references and free online:
- Daume, A Course in Machine Learning
- Barber, Bayesian Reasoning and Machine Learning
- Hastie, Tibshirani, and Friedman, The Elements of Statistical Learning
- MacKay, Information Theory, Inference, and Learning Algorithms

Course objectives:
The objective of this course is to teach fundamental algorithms and problems related to Machine Learning, and practical implementations and theories of classic and contemporary Machine Learning algorithms. Upon completion of the course, students will:
1. Have a good understanding of fundamental problems in Machine Learning.
2. Have a good understanding of classic and contemporary Machine Learning algorithms to solve various problems.
3. Be aware of assumptions, strengths, and weaknesses of Machine Learning algorithms.
4. Be able to develop Machine Learning methods to solve real-world problems.

The learning outcomes will be assessed based on course projects.

Topics covered:
1. Model Complexity;
2. Bayes classifiers, naïve Bayes;
3. Nearest Neighbor, K Nearest Neighbor (KNN);
4. Decision tree, entropy, information gain;
5. Linear classification, linear regression;
6. Support Vector Machine (SVM), kernel trick;
7. Neural Network, backpropagation, convolutional neural network;
8. Unsupervised learning, K-means clustering, agglomerative clustering, Gaussian Mixtures, EM;
9. Ensemble methods, bagging, boosting;
10. Latent spaces, Principal Component Analysis (PCA), SVD, eigen-face;
11. Reinforcement learning, Markov Process, Markov Reward Process, Markov Decision Process.

Class/laboratory schedule:

Timetabled work in hours per week			No of teaching weeks	Total hours	Total credits	No/Duration of exam papers
Lecture	Tutorial	Practice				
3	0	0	14	42	3	n/a

Student study effort required:

Class contact:	
Lecture	42 hours
Tutorial	0 hours
Other study Effort	
Self-study	20 hours
Course project	48 hours
Total student study effort	110 hours

Student assessment:
Final assessment will be determined on the basis of
* Three course projects: 60%=3*20%
* Final Project: 40%

Course Outline:

Weeks	Topics	Course Work	Homework / project
1 Jan 6-10	Introduction	python tutorial pytorch tutorial	
2 Jan 13-17	Model Complexity; Nearest Neighbor	Nearest neighbor	Project 1 out: install PyTorch and Jupyter; use CIFAR10 and MNIST to implement Nearest Neighbor for recognition
3 Feb 6-7	Bayes Classifiers; Naïve Bayes; Bayes Error	Naïve Bayes classifier	
4 Feb 10-14	Linear Regression, Gradient Descent, Cross Validation, Regularization	Linear regression, cross validation	
5 Feb 17-21	Linear Classification, Perceptron, Logistic Regression, Multi-Class, Cross-Entropy	Linear classifier, cross-entropy	Project 1 due Project 2 out: train a linear classifier and an SVM for recognition, tune hyperparameteres through validation
6 Feb 24-28	VC dimension, Structural risk minimization, AIC, BIC	VC-dimension,	
7 Mar 3-7	SVM, Lagrangian and Dual, Kernel Trick	SVM, kernel trick	
8 Mar 10-14	Decision Tree, Entropy, Information Gain	Decision tree, information gain	Project 2 due Project 3 out: train a deep neural network and convolutional neural network for recognition.
9 Mar 17-21	Neural Networks, Back Propagation, Convolutional Neural Networks	Neural networks, CNN	
10 Mar 24-28	Ensemble of learners, Bagging, Boosting	Ensemble method	
11 Apr 7-11	Clustering, K-means, Agglomerative clustering, Gaussian Mixture, EM	k-means, Gaussian mixture models	Project 3 due Final project out: propose a research project worthy of one month to work on; do literature

			review; design experiments; prepare datasets; design methods; benchmark methods; write project report
12 Apr 22-25	Latent space models, PCA, SVD, Eigen-face, recommendation systems	Latent space, PCA	
13 Apr 28-May 2	Reinforcement Learning, MDP, Monte Carlo	Reinforcement Learning	
14	Presentation of course project		Final project report due

Student Disabilities Support Service:
The University of Macau is committed to providing an equal opportunity in education to persons with disabilities. If you are a student with a physical, visual, hearing, speech, learning or psychological impairment(s) which substantially limit your learning and/or activities of daily living, you are encouraged to communicate with your instructors about your impairment(s) and the accommodations you need in your studies. You are also encouraged to contact the Student Disability Support Service of the Student Counselling and Development Section (SCD), which provides appropriate resources and accommodations to allow each student with a disability to have an equal opportunity in education, university life activities and services at the University of Macau. To learn more about the service, please contact SCD at scd.disability@umac.mo, or 8397 4901 or visit the following website: https://sao.um.edu.mo/

Coordinator:
Shu Kong, Assistant Professor of CIS
https://aimerykong.github.io

Persons who prepared this description:
Shu Kong, December 24, 2024

Part B General Course Information and Policies

2nd Semester 2024/2025
Instructor: Prof. Shu Kong Office: E11-4025
Office Hour: Tuesday and Wednesday at 14:00-15:00, or by appointment
Email: skong@um.edu.mo

Time/Venue

Lecture	Friday 19:00-22:00	E11-1012

Grading Distribution:

Percent. Grade	Final Grade	Percent. Grade	Final Grade	Percent. Grade	Final Grade
100 – 93	A	77 – 73	B-	57 – 53	D+
92 – 88	A-	72 – 68	C+	52 – 50	D
87 – 83	B+	67 – 63	C	below 50	F
82 – 78	B	62 – 58	C-		

Homework Policy:
The completion and correction of homework are helpful for learning. As a result,

- There are four course projects.
- Projects are due two weeks after assignment unless otherwise noted.
- There is no late penalty but is a reward (+5 points) if submitting >24 hours earlier by the deadline.
- Revisions submitted after the deadline will not be rewarded for +5 points.
- Possible revision of grades may be discussed with the instructor within one week after deadline.
- The course grade will be based on the weighted average of the homework and final project grades.

Other Important Notes:

- Check the ummoodle web pages for announcements, and lectures.
- Cheating is absolutely prohibited by the university.
- UM students can 'use ChatGPT or other generative-AI systems to ***enhance*** their learning' and that they 'should be aware that they ***must be authors of their own work***' (email 'Notes on the use of generative-AI systems', 11 April 2023). See details in FAQs for Students: Using Generative AI Tools in Graded Assignments (https://ctle.um.edu.mo/resource/faqs-for-students-using-generative-ai-tools-in-graded-assignments/).