Lahore University of Management Sciences

EE 514/CS 535 Machine Learning

Spring 2020-21

Instructor	Zubair Khalid
Room No.	SBASSE 9-251
Office Hours	TBA
Email	zubair.khalid@lums.edu.pk
Telephone	042 3560 8477
Secretary/TA	Rabeeya Hamid, TBA, TBA
TA Office Hours	TBA
Course URL (if any)	https://www.zubairkhalid.org/ee514_2020.html

Course Teaching Methodology (Please mention following details in plain text)

- Teaching Methodology: We will follow hybrid (synchronous/asynchronous approach) for content delivery. Assessment includes timed Quizzes on LMS, Homeworks/Assignments and take-home (timed) Exams.
- Lecture details: Recorded lectures would be made available over the course YouTube channel or LMS. We will have one recitation/tutorial almost every week during the regular time slot.

Course basics				
Credit Hours	3 hours			
Lecture(s)	Nbr of Lec(s) Per Week	2	Duration	75 minutes
Recitation/Lab (per week)	Nbr of Lec(s) Per Week		Duration	
Tutorial (per week)	Nbr of Lec(s) Per Week	1 (optional)	Duration	50 minutes

Course distribution	
Elective	This is an elective course.
Open for Student Category	Juniors, seniors and graduates.
Close for Student Category	Please see prerequisites below.

Course description

Machine learning (ML) studies the design and development of algorithms that learn from the data and improve their performance through experience. ML refers to a set of methods and that help computers to learn, optimize and adapt on their own. ML has been employed to devise algorithms for diverse applications including object detection or identification in computer vision, sentiment analysis of speaker or writer, detection of disease and planning of therapy in healthcare, product recommendation in e-commerce, learning strategies for playing games, recommending movies to customers, speech recognition systems, fraudulent transaction detection or loan application approval in banking sector, to name a few.

This course provides a thorough introduction to the theoretical foundations and practical applications of ML. We will learn fundamental algorithms in supervised learning and unsupervised learning. We will not only learn how to use ML methods and algorithms but will also try to explain the underlying theory building on mathematical foundations. While reviewing the several problems and algorithms to carry out classification, regression, clustering, dimensionality reduction, we will focus on the core fundamentals which unifies all the algorithms. The theory discussed in class will be tested in assignments, quizzes and exams.

Course prerequisites

- Undergrads (Seniors/Juniors) must have passed:
  - An Ugrad/Grad course in Probability (MATH230 (Probability) OR DISC203 (Probability & Statistics) OR CS501 (Applied Probability))
  - And, a programming course (CS200/EE201 (Intro. to Programming))
  - And, a course on Linear Algebra (MATH120 (LA with Diff. Equations))
- Grads are strongly advised to brush up their programming skills and take CS501 (Applied Probability), may be in parallel with ML
- All students must possess strong programming skills and proficiency in algorithm implementation in JAVA/C/Python/MATLAB

Lahore University of Management Sciences

Course objectives
The goal of this course is to get the students excited about Machine Learning and to enable them to: • To provide a thorough introduction to ML methods • To build mathematical foundations of ML and provide an appreciation for its applications • To provide experience in the implementation and evaluation of ML algorithms • To develop research interest in the theory and application of ML

Learning outcomes
By the end of the course, students should be able to: • Understand and recognize a machine learning problem • Formulate and execute solutions to the machine learning problems • Understand the core theoretical concepts serve as foundations of ML algorithms • Understand the trade-offs among model complexity, data size, and model performance for different algorithms • Apply and interpret information theoretic and probabilistic machine learning methods on real-world datasets

Grading Breakup and Policy (remote)	
Online timed quizzes (1-2 per week):	20%
Project:	10%
Programming Assignments or homework(s)/	25%
Mid examination:	20%
Final examination:	25%

Examination detail		
Midterm Exam	Yes/No:	Yes
	Duration:	2 hours
	Exam Specifications:	Timed (take-home) exam
Final Exam	Yes/No:	Yes
	Duration:	2.5 – 3 hours
	Exam Specifications:	Timed (take-home) exam.

Textbook(s)/Supplementary Readings
Textbooks • Pattern Recognition and Machine Learning, Christopher M. Bishop – CB • Machine Learning: a Probabilistic Perspective, Kevin Murphy – KM • Machine Learning, Tom Mitchell – TM • The Elements of Statistical Learning: Data mining, Inference, and Prediction, by Hastie, Tibshirani, Friedman – HTF • Information Theory, Inference, and Learning Algorithms, David Mackay – DM

Course Policies
Plagiarism: All work MUST be done independently. In certain assignments students will be allowed to have discussions with peers, in which case they must mention the name and roll number of the student with whom the discussion took place and the nature of the discussion. Even in those assignments, all implementations need to be done independently. Any plagiarism or cheating of work from others or the internet will be immediately referred to the DC. If you are confused about what constitutes plagiarism, it is YOUR responsibility to consult with the instructor or the TA in a timely manner. No "after the fact" negotiations will be possible.

Lahore University of Management Sciences

SSE Council on Equity and Belonging

In addition to LUMS resources, SSE's **Council on Belonging and Equity** is committed to devising ways to provide a safe, inclusive and respectful learning, living, and working environment for students, faculty and staff. To seek counsel related to any issues, please feel free to approach either a member of the council or email at cbe.sse@lums.edu.pk.

Mental Health Support at LUMS

In addition to LUMS resources, SSE's **Council on Belonging and Equity** is committed to devising ways to provide a safe, inclusive and respectful learning, living, and working environment for students, faculty and staff. To seek counsel related to any issues, please feel free to approach either a member of the council or email at cbe.sse@lums.edu.pk.

Harassment Policy

SSE, LUMS and particularly this class, is a harassment free zone. Harassment of any kind is unacceptable, whether it be sexual harassment, online harassment, bullying, coercion, stalking, verbal or physical abuse of any kind. Harassment is a very broad term; it includes both direct and indirect behaviour, it may be physical or psychological in nature, it may be perpetrated online or offline, on campus and off campus. It may be one offense, or it may comprise of several incidents which together amount to sexual harassment. It may include overt requests for sexual favours but can also constitute verbal or written communication of a loaded nature. Further details of what may constitute harassment may be found in the LUMS Sexual Harassment Policy, which is available as part of the university code of conduct.

LUMS has a Sexual Harassment Policy and a Sexual Harassment Inquiry Committee (SHIC). Any member of the LUMS community can file a formal or informal complaint with the SHIC. If you are unsure about the process of filing a complaint, wish to discuss your options or have any questions, concerns, or complaints, please write to the Office of Accessibility and Inclusion (OAI, oai@lums.edu.pk) and SHIC (shic@lums.edu.pk) —both of them exist to help and support you and they will do their best to assist you in whatever way they can. You can find more details regarding the LUMS sexual harassment policy here.

To file a complaint, please write to harassment@lums.edu.pk.

Assessed Course Learning Outcomes

EE 514/ CS 535	By the end of the course, students should be able to:
CLO1:	Understand and recognize a machine learning problem
CLO2:	Formulate and execute solutions to the machine learning problems
CLO3:	Understand the core theoretical concepts serve as foundations of ML algorithms
CLO4:	Understand the trade-offs among model complexity, data size, and model performance for different algorithms
CLO5:	Apply and interpret information theoretic and probabilistic machine learning methods on real-world datasets

Relation to EE Program Outcomes

EE-514/CS 535 CLOs	Related PLOs	Levels of Learning	Teaching Methods	CLO Attainment checked in
CLO1	PLO1	Cog-1		
CLO2	PLO2	Cog-2		
CLO3	PLO1	Cog-1		
CLO4	PLO4	Cog-4		
CLO5	PLO2	Cog-3		

Lahore University of Management Sciences

Week No.	Topic	Assessment and Additional Remarks
1.	Course Introduction Machine Learning Overview (Notes 01) Supervised Learning: Formulation, Setup, Train-test split, Generalization (Notes 02)	Notes 01 Notes 02
2.	k-Nearest Neighbor (kNN) Algorithm Algorithm Formulation Distance Metric Choice of k Algorithm Convergence Storage Time Complexity Analysis Fast kNN	Notes 03 PA01 Due
3.	The Curse of Dimensionality and Connection with kNN Dimensionality Reduction: Feature Selection and Extraction Principal Component Analysis	Notes 03(b)
4.	Classifer Performance Evaluation: Confusion Matrix Sensitivity, Specificity, Precision Trade-offs, ROC, AUC, F1-Score and Matthew's Correlation Coefficient	Notes 04 HW01 Due
5.	Multi-class Classification, Evaluation, Micro, Macro Averaging Regression: Linear Regression, Polynomial Regression, Overfitting	Notes 05 PA02 Due
6.	Gradient Descent Algorithm Regularization	Notes 05
7.	Probability Review Bayesian Learning Framework, MAP and ML Hypothesis Linear Regression as ML estimation Naive Bayes Classifier	Notes 06 Notes 07 HW02 Due PA3 Due
8.	Naïve Bayes Classifier for Text Classification Bayesian Networks Introduction	Notes 08 Mid-Exam
9.	Logistic Regression: Mathematical Model, Decision Boundaries, Loss/Cost Function, Gradient Descent Multi-class Logistic Regression	Notes 09
10.	Perceptron and Perceptron Classifier, Perceptron Learning Algorithm and its Geometric Intuition Perceptron Learning Algorithm Convergence SVM Overview	Notes 10 PA04 Due
11.	Hard SVM, Soft SVM, Kernel Trick	Notes 11 HW03 Due
12.	Neural Networks Introduction, Model, Forward Pass	Notes 12 PA5 due
13.	Neural Networks: Back Propagation	Notes 12
14.	Unsupervised Learning, Clustering Overview K-means Clustering Agglomerative Clustering	Notes 13 Project Due

Prepared and Revised by:	**Zubair Khalid**
Revision Date:	**29-12-2020**