



Faculty of Engineering and Technology
Department of Electrical and Computer Engineering

Course Information	
Course Title	Machine Learning and Data Science
Course Number	ENCS5341
Prerequisites	ENCS3340
Semester	First Semester 2024/2025
Instructors	Dr. Ismail Khater Dr. Yazan Abu Farha

References
<ul style="list-style-type: none">• Data Mining Concepts and Techniques, Third Edition. Jiawei Han, Micheline Kamber, and Jian Pei. 2012• The Elements of Statistical Learning. Friedman, Tibshirani, and Hastie. 2012• Pattern Recognition and Machine Learning. Bishop, Christopher. 2007• Deep Learning, Ian Goodfellow, Yoshua Bengio and Aaron Courville, 2016• Probabilistic Machine Learning: An Introduction. Kevin P. Murphy. 2022• Lecture Notes

Course Description
<p>This course covers the theory and practical algorithms for machine learning from a variety of perspectives. Topics covered include: Regression (Linear and nonlinear), Classification (Logistic Regression, Artificial Neural Networks, Support Vector Machine, Naïve Bayes, Bayesian Network, HMM, Random Forests). Overfitting, Regularization, Evaluation Methods, Evaluation Measures, Model Selection. Clustering (Cluster Analysis, K-Means, EM and Mixture Models, Hierarchical Clustering, Cluster Evaluation). Data Set Analysis, Dimensionality Reduction, Feature Selections, Deep Learning, and Reinforcement Learning. In addition, short programming assignments include hands-on experiments with various learning algorithms, and a larger course project gives students a chance to dig into an area of their choice.</p>

Course Objectives
<p>The course aims to provide an introduction to the basic principles, techniques, and applications of Machine Learning. Programming assignments are used to help clarify basic concepts.</p> <p>The course covers the principles, design and implementation of learning programs that improve their performance on some set of tasks with experience. Upon successful completion of the course, students will have a broad understanding of machine learning algorithms and their use in data-driven knowledge discovery and program synthesis. Students will also be able to identify, formulate and solve machine learning problems that arise in practical applications. Students will have a knowledge of the strengths and weaknesses of different machine learning algorithms (relative to the characteristics of the application domain) and be able to adapt or combine some of the key elements of existing machine learning algorithms to design new algorithms as needed.</p>

Assessment Policy		
Assessment Type	Expected Due Date	Weight
Midterm Exam	TBA	30%
Final Exam	TBA	40%
Assignments	TBA	30%

Course Contents (Tentative)
Topics
1. Introduction <ul style="list-style-type: none"> Introduction and Motivation Probability and Linear Algebra review
2. Introduction to Data Science <ul style="list-style-type: none"> Data Science Life Cycle, Types of Data, Exploratory Data Analysis: Summary Statistics, Visualization. Data Preprocessing: Data Cleaning, Data Discretization, Data Normalization, Handling Missing Data, Encoding Categorical Data, Handling Unbalanced data. Important data science and machine learning libraries and tools.
3. Regression <ul style="list-style-type: none"> Linear and Polynomial Regression Multivariate Regression Overfitting and Regularization
4. Regularization and Model Selection <ul style="list-style-type: none"> Model Selection and Assessments: Bias/Variance Tradeoff, Evaluation Methods (Holdout set, Cross Validation, Leave-one-out cross-validation, Validation set). Ridge and lasso regression.
5. Classification <ul style="list-style-type: none"> K-NN Logistic Regression Support Vector Machines and Kernel Methods Ensembles Methods: Bagging, Boosting, Random Forests Hidden Markov Model.
6. Introduction to Deep Learning <ul style="list-style-type: none"> Artificial Neural Networks Convolutional Neural Networks Recurrent Neural Networks
7. Clustering <ul style="list-style-type: none"> Cluster Analysis K-Means, K-Medoids, Fuzzy C-Means Hierarchical Clustering Expectation Maximization and Gaussian Mixture Models – EM-GMM Cluster Validation Methods
8. Dimensionality Reduction <ul style="list-style-type: none"> Principal Component Analysis (PCA). Autoencoders.

Teaching and Learning Methods

Lectures, assignments, in-class activities and exams.

Additional Notes	
------------------	--

Assignments	No late assignments
-------------	---------------------

Exams	Comprehensive exams
-------	---------------------

Makeup Exams	<u>No makeup exam</u>
--------------	------------------------------

Attendance	Your attendances is very important
------------	------------------------------------

Key to a good grade	Reading the TEXTBOOK and HANDOUT + DOING the PROJECTS
---------------------	---