Course Outline - Machine Learning (CSE 445)
Section 1
Department of ELectrical and Computer Engineering
North South University

Summer 2021

Instructor:	Dr. Sifat Momen	**Time:**	ST 9:40 - 11:10
Email:	sifat.momen@northsouth.edu	**Place:**	virtual

Textbook:

1. Machine Learning by *Tom M. Mitchell*, McGraw-Hill Education, 1^{st} edition, 1997.
2. Hands-on Machine Learning with Scikit-Learn, Keras & Tensorflow: Concepts, Tools and Techniques to Build Intelligent Systems by *Aurélien Géron*, O'REILLY, 2019.

References:

1. Introduction to Machine Learning with Python: A guide for Data Scientists by *Andreas C. Müller & Sarah Guido*, O'Reilly Media, 1^{st} edition, 2016.
2. Pattern recognition and Machine Learning by *Christopher M. Bishop*, Springer, 2006.

Objectives: This is a foundation course on Machine learning. After completing this course successfully, a student should have deep insights on machine learning concepts and will have the knowledge of applying machine learning in understanding and forecasting data. Students will gain practical knowledge of how to collect, pre-process and apply ML algorithms. Students will have profound understanding on the classification of learning: Unsupervised and supervised learning, Connectionist learning and Reinforcement learning. Students will learn different classification and regression algorithms including decision tree induction, KNN, naïve bayes algorithm, support vector machine, linear regression, logistic regression and artificial neural network. Ensemble learning techniques will also be covered in substantial details. Unsupervised learning algorithms for clustering such as K-means and dbscan will be explored in this course. Feature selection and feature engineering will also be covered extensively in this course. Students will get pragmatic tips on machine learning issues including dealing with imabalance data, tackling overfitting of model on training data, tackling curse of dimensionality and understanding the bias variance tradeoff. Finally, students will learn about computaitonal learning theory including PAC learning and VC dimension.

Prerequisites: It is expected that students pursuing this course have a mature level of understanding in algorithms, linear algebra and probability & statistics.

Tentative Lecture Sequence:

- Introduction to Machine learning and Concept learning
- Data Visualization Techniques and Data Pre-processing
- Introduction to classification problem, zeroR and oneR classifier, confusion matrix
- Decision tree induction and tackling overfitting of the model on the dataset
- Strategies of testing & training, K-fold cross validation and Performance evaluation of classifiers
- Hyperparameter Tuning
- Instance based learning
- Bayesian learning
- Bias variance tradeoff
- Feature selection and engineering
- Linear regression & regularization
- Logistic regression
- Support vector machine
- Artificial Neural Network
- Computational learning theory, PAC learning, VC dimension
- Ensemble learning & Random Forest
- Dimensionality Reduction
- Clustering algorithms using unsupervised learning
- Mixture models and EM (optional)

Assessment weights: Assignment (5%), Quiz (10%), Viva (5%), Midterm (25%), Final (30%), Project (25%).

N.B.: One of the deliverables of the project is a paper. More details regarding paper template and submission will be given later. The paper has to written using LaTeX.

Course Policy:

- In no way, there will be any makeup of final exam.
- In case of an absolute emergency, if a student misses the midterm exam, a proper application, backed up by supporting documents, has to be provided. After careful inspection, if I feel that this can be approved, then an arrangement of makeup exam will be made. However, a student will incur a 20% penalty for this.
- Failure to submit the project will result in an **F** grade.
- Lack of knowledge of academic honesty policy is not a reasonable explanation for a violation. Exhibiting any kind of academic dishonesty will not be condoned and will be dealt strictly.