Course Title	**Selected Topics in Information and Computational Systems - Machine Learning**
Course Number	**ENGR-UH 4560**

Course Description: Machine Learning is the basis for the most exciting careers in data analysis today. This course introduces students to the concepts of machine learning and deep learning. This course covers a broad introduction to machine learning techniques, which include both supervised learning and unsupervised learning techniques such as classification, support vector machines, decision trees, ensemble learning and random forests, dimensionality reduction, and neural networks and deep learning. In addition to learning about the most effective machine learning techniques, you will gain the practical implementation of applying these techniques to real engineering problems.

Course Structure	14 weeks course - 2 lectures per week (75 mins each) + 145 min recitation/laboratory per week, Spring semester.
Credits	4
Prerequisite Courses	ENGR-AD-202 Computer System Programming or equivalent programming skills
Instructor	Yi Fang, A1-191 , Email: yfang@nyu.edu
Teaching Assistant	Tadesse Mulugeta Email: tsm8@nyu.edu

Intended Learning Outcomes: Upon successful completion of this course, students will have obtained knowledge and comprehension of course topics as described below.

- Develop the programming skills using python to implement the basic machine learning algorithms [(1.a), (1.c),];
- Understand the fundamental concepts of machine learning algorithms for them: classification, clustering, regression, SVMs, decision trees, dimensionality reduction, and neural networks and deep learning [(1.a)];
- Develop a solution to an engineering problem based on learned machine learning algorithms [(1.c), (2.b)];

Textbook:	Hands-On Machine Learning with Scikit-Learning & TensorFlow, Aurelien Geron

Teaching and Learning Methodologies: Students are expected to arrive at class with an understanding of the basic definitions, concepts, and applications of relevant topics. Class time will be devoted to in-depth discussions of course topics. Lab project assignments are dedicated for reinforcing course topics via solving representative problem sets and holding class discussions.

Evaluation: The course will provide with two tracks of assignments and grades. Student may choose one of two tracks for the class. The first track will be based on six mini-projects, one large deep learning project, two presentations and the second track will be based on six mini-projects and three deep learning projects. The distribution of grades is subject to some revision at the discretion of the instructor. Typical weighting values are given below:

Track 1	Track 2
Mini Projects: 6*10%	Mini Projects: 6*10%
Large Deep Learning Project: 1*40%	Deep Learning Project One: 1*20%
	Deep Learning Project Two: 1*20%

Schedule: A typical schedule for course topics, projects, and exam dates is given in the table below (The actual schedule might be slightly different from this).

Week	Lecture Topics
1	▪ Introduction
2	▪ Regression
3	▪ Classification
4	▪ Support Vector Machine (SVM)
5	▪ Tree
6	▪ Ensemble Machine Learning
7	• K-Nearest Neighbors (KNN)
8	• Midterm Presentation
9	• Dimensionality Reduction
10	▪ Unsupervised Learning: Clustering
11	▪ Intro. To neural networks
12	▪ Deep Learning and its Application
13	▪ Deep Learning and its Application
14	▪ Final Presentation

Project Description:

Track-One: Large Deep Learning: Object Recognition and Detection by Deep Learning: In computer vision, object detection is an important task to identify the presence, location, and type of one or more objects in a given figure. It is a challenging problem that includes the development of the approaches for object recognition (e.g. where are they), object localization (e.g. what are their extent), and object classification (e.g. what are they). Recent deep learning techniques have demonstrated the state-of-the-art results for object detection. The "You Only Look Once," or YOLO was developed as the pioneering work for object detection using deep learning techniques. The YOLO achieves nearly state-of-the-art results with a single end-to-end model that can perform object detection in real-time. In this project, the student will be required to develop a YOLO model for object detection from the implementation from scratch using PyTorch library.

Deep Learning One: CNN for Object Detection and Classification: Image classification is the task of assigning an input image one label from a fixed set of categories. This is one of the core problems in Computer Vision that, despite its simplicity, has a large variety of practical applications. We would like you to establish a neural network involving advance DNN modules (i.e. convolution layers, RELU, pooling and fully connection layers and etc.) to distinguish the specific category of an input image. The implementation of CNN will be firstly trained with the training dataset. The trained model will be evaluated on the testing dataset.

Deep Learning Two: GAN-Human face synthesis: The Generative Adversarial Networks, or GANs, were introduced as an architecture for training generative models, such as deep convolutional neural networks for generating images. The GANs have made huge breakthroughs and can now produce highly convincing fake images of animals, landscapes, human faces, etc. The generative adversarial networks (GANs) are deep neural net architectures comprised of two nets, pitting one against the other (thus the "adversarial"). In this project, we will ask the student to understand the principles of the GANs, learn how generative algorithms work, and for that, contrasting them with discriminative algorithms is instructive, understand how discriminative algorithms try to classify input data; that is, given the features of an instance of data, they predict a label or category to which that data belongs.

Regression: In statistics, regression is a approach to modeling the relationship between a scalar dependent variable y and one or more explanatory variables (or independent variables) denoted X. The case of one explanatory variable is called simple linear regression. For more than one explanatory variable, the process is called multiple linear regression. Linear regression models are often fitted using the least squares approach, but they may also be fitted in other ways, such as by minimizing the "lack of fit" in some other norm (as with least absolute deviations regression), or by minimizing a penalized version of the least square loss function as in ridge regression (L2-norm penalty) and lasso (L1-norm penalty). Conversely, the least squares approach can be used to fit models that are not linear models. Thus, although the terms "least squares" and "linear model" are closely linked, they are not synonymous. This project will ask the student to implement the regression model to fit the data.

SVM for classification: A Support vector machines (SVMs, also support vector networks) are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. Given a set of training examples, each marked as belonging to one or the other of two categories, an SVM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic binary linear classifier (although methods such as Platt scaling exist to use SVM in a probabilistic classification setting). An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall. This gap is also called maximum margin and the SVM classifier is called maximum margin classifier. This project will ask the student to implement the SVM model to classify the data.

K-NN: In pattern recognition, the k-nearest neighbors algorithm (k-NN) is a non-parametric method used for classification and regression. K-Nearest Neighbors is one of the most basic yet essential classification algorithms in Machine Learning. It is a supervised learning method which has intense applications in pattern recognition, data mining and intrusion detection. In general, KNN is a model that classifies data points based on the points that are most similar to it. It uses test data to make a guess on what an unclassified point should be classified as. This project will ask the student to train KNN for the data classification.

K-means and Hierarchical Clustering: K-means clustering is a method of vector quantization, originally from signal processing, that is popular for cluster analysis in data mining. k-means clustering aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster. This results in a partitioning of the data space into Voronoi cells. Hierarchical clustering involves creating clusters that have a predetermined ordering from top to bottom. For example, all files and folders on the hard disk are organized in a hierarchy. There are two types of hierarchical clustering, Divisive and Agglomerative.

PCA: In pattern recognition, high-dimensional data are very common which poses several challenges for classification and regression. The main purposes of a principal component analysis are the analysis of data to identify patterns and finding patterns to reduce the dimensions of the dataset with minimal loss of information. In general, Principal component analysis (PCA) is a procedure for reducing the dimensionality of the variable space by representing it with a few orthogonal (uncorrelated) variables that capture most of its variability. PCA is an unsupervised learning method which find patterns without reference to prior knowledge. This project deals with decision trees in data mining.

Classification with Neural Network: Artificial Neural Networks (ANN) or connectionist systems are computing systems vaguely inspired by the biological neural networks that constitute animal brains. The neural network itself is not an algorithm, but rather a framework for many different machine learning algorithms to work together and process complex data inputs. Such systems " learn " to perform tasks by considering examples, generally without being programmed with any task-specific rules. For example, in image recognition, they might learn to identify images that contain cats by analyzing example images that have been manually labeled as "cat" or "no cat" and using the results to identify cats in other images. They do this without any prior knowledge about cats, for example, that they have fur, tails, whiskers and cat-like faces. Instead, they automatically generate identifying characteristics from the learning material that they process.

Relationship to Outcomes:

	Shared Engineering Outcomes					Program Specific Criteria				
	(a)	(b)	(c)	(d)	(e)	CivE	CmpE	ElecE	MechE	GenE
Lectures		x					x	x		
Labs / Recitations		x					x	x		

	Student Learning Outcomes						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Lectures	x						
Labs / Recitations						x	

Shared Engineering Outcomes:

a) Apply techniques in the practice of leadership and innovation [(5)];
b) Identify social, economic, ethical and other factors that shape engineering solutions and incorporate them in conjunction with engineering principles in problem solving and designing systems, components, or processes to meet desired needs within realistic constraints [(1), (2), (4)];
c) Recognize and respond respectfully to cultural concerns and differences when solving problems both physical and ethical [(1), (2), (4)];
d) Exhibit guidance and organizational effectiveness in multidisciplinary teams as a participant and a leader [(5)];
e) Demonstrate competence in writing and speaking effectively, and in communicating significant technical information in a clear and concise manner [(3)].

Program Specific Criteria:

- CivE: Civil Engineering graduates will be able to work professionally in four of the technical areas of the civil engineering discipline (structural, geotechnical, transportation, and environmental), design systems, components, and processes in more than one civil engineering context, and apply the principles of project management.
- CompE: Computer Engineering graduates will be able to analyze and design complex computing and network devices and systems containing hardware and software components.

- ElecE: Electrical Engineering graduates will be able to analyze and design complex electrical, electronic, and communication devices and systems.
- MechE: Mechanical Engineering graduates will be able to analyze and design systems, components, and processes, and work professionally in both thermal and mechanical systems areas.
- GenE: General Engineering graduates will be able to analyze and design devices and systems in an interdisciplinary engineering area related to: Biomedical and Health Systems; Information, Communication, and Electronic Systems; or Urban Systems.

Student Learning Outcomes:

1. an ability to identify, formulate, and solve complex engineering problems by applying principles of engineering, science, and mathematics; 2. an ability to apply engineering design to produce solutions that meet specified needs with consideration of public health, safety, and welfare, as well as global, cultural, social, environmental, and economic factors;	3. an ability to communicate effectively with a range of audiences; 4. an ability to recognize ethical and professional responsibilities in engineering situations and make informed judgments, which must consider the impact of engineering solutions in global, economic, environmental, and societal contexts;	5. an ability to function effectively on a team whose members together provide leadership, create a collaborative and inclusive environment, establish goals, plan tasks, and meet objectives 6. an ability to develop and conduct appropriate experimentation, analyze and interpret data, and use engineering judgment to draw conclusions 7. an ability to acquire and apply new knowledge as needed, using appropriate learning strategies.