# Introduction to Machine Learning (CSCI-UA.473)

# (Fall 2021)

# Overview of the Course

The goal of this course is to introduce undergraduate students to the field of machine learning. No prior knowledge of machine learning is required. However a good understanding of concepts of probability, linear algebra, and some form of calculus is a must in order to understand the contents of the course. The course introduces students with a diverse set of topics from the field, including linear parametric models, non-linear parametric models including deep learning models, supervised learning framework, unsupervised learning framework, non-parametric models, ensemble learning, and reinforcement learning, among others.

# Target Audience

This course is aimed at 3rd- or 4th-year undergraduate students in computer science. Please contact either Leeann Longi <longi \_at\_ cs.nyu.edu> or Romeo Kumar <kumar \_at\_ cs.nyu.edu>, the student advisors at the CS Department, directly for any questions pertaining to enrollment for the course.

# Expectation from Students

The course is composed of a mix of theoretical lectures and practical labs. The assessment will be based on a set of homework assignments and a final exam.

* Students will use their NYU credentials to complete all tasks and communications for this course (sending emails etc)
* Students are expected to attend and actively participate in all the lectures and during the lab sessions. Active participation particularly in the lab sessions is necessary to successfully learn machine learning concepts
* Students are required to bring their laptops to the lab sessions for it will involve hands on coding
* Machine learning is a vast and growing field and it is unreasonable to expect that a semester-long course can cover all the topics in depth. While there is no official book for the course, the students are strongly encouraged to read the supplementary material provided (or pointed) to within the class
* Students should **consult the course syllabus first** if they have any questions about assignment parameters/due dates. If the answer is not on the syllabus, then students should reach out to the instructor
* While students are encouraged to discuss/brainstorm assignments with their peers, each student should write its own assignment. Copying assignments is forbidden and NYU has a strict policy towards academic integrity (see at the end of this document). Furthermore, the goal of this course is for you to learn machine learning concepts and not to merely get good grades. Cheating will certainly not help you accomplish the goals
* If you are incorporating any text/figures from an external source, you have the obligation to make it clear. Without doing so is considered plagiarism. The standard way of explicitly calling out the external source is by adding an explicit statement within your work, which calls out (cites) which external source you used for your material. This not just includes the text in your assignment but also includes the descriptions of your baseline model and your data for example. Failure to do so will result in a zero grade for the submitted work and, except in cases of obvious mistakes, a University investigation.

# General Information

### Lectures

* Schedule: 2:00 pm - 3:15 pm on Tuesdays and Thursdays
* All the lectures will be in person
* Lectures will be a mix of theory and labs
* The lectures will be delivered through slides with some parts being on the white board
* The slides will be made available online after the class
* Most lectures on Thursdays will be lab sessions
* The lab sessions will be led by the tutor

### Instructor

* Name: [Sumit Chopra](http://www.spchopra.org/)
* Email: [sumit@cs.nyu.edu](mailto:sumit@cs.nyu.edu)
* Office: 60 5th Avenue, Room 501
* Preferred Communication Method: Email

### Assistants

* Grader and Tutor Name: Umang Sharma
* Grader and Tutor email: us453@nyu.edu

### Office Hours

* **Chopra:** 4:00 pm - 5:00 pm, Thursdays (or by appointment)
* **Sharma:** 3:00 pm - 4:00 pm, Fridays

### Grading and Assignments

* Homeworks (70%) + Final Exam (30%)
* You **must pass** the final exam to pass the course
* There will be six (7) homework assignments through the semester
* All homework assignments should be written in typeset in Latex and submitted as pdfs. Handwritten notes or MS Word documents will not be accepted
* The due date for the homeworks will be one week from the date assigned AND before the lecture. For example, assignment given on 09/07 after the class will be due on 09/14 **before** the class
* No late assignments will be accepted. This is a strict policy.
* **Final exam**: TBD

### Course Website

* [Brightspace](https://brightspace.nyu.edu/d2l/home/133519)

### Books

The following resources will be useful but do not need to be purchased

* [The Elements of Statistical Learning](https://www.amazon.com/Elements-Statistical-Learning-Prediction-Statistics/dp/0387848576/): Trevor Hastie, Robert Tibshirani, and Jerome Friedman
* [Pattern Recognition and Machine Learning](https://www.microsoft.com/en-us/research/uploads/prod/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf): Christopher Bishop
* [Learning from Data](https://www.amazon.com/Learning-Data-Yaser-S-Abu-Mostafa/dp/1600490069/): Yaser S. Abu-Mostafa, Makil Magnon-Ismail, and Hsuan-Tien Lin
* <http://www.inference.org.uk/mackay/itprnn/book.html>: David J.C. MacKay
* [Introduction to Machine Learning with Python](http://shop.oreilly.com/product/0636920030515.do): Andreas C. Müller & Sarah Guido

### Other Material for Review

* [Lecture notes](https://github.com/nyu-dl/Intro_to_ML_Lecture_Note/blob/master/lecture_note.pdf) by[Kyunghyun Cho](http://www.kyunghyuncho.me/)
* [Linear Algebra and Vector Calculus](http://cs229.stanford.edu/section/cs229-linalg.pdf)
* [Probability Theory](http://cs229.stanford.edu/section/cs229-prob.pdf)

### Lab Sessions

* Most lectures on Thursdays will be lab sessions led by the tutor
* Lab sessions will use Python 3 and PyTorch
  + [Installation instructions for classic Jupyter](https://jupyter.readthedocs.io/en/latest/install/notebook-classic.html) or [Jupyter Lab](https://jupyterlab.readthedocs.io/en/stable/getting_started/installation.html) (Necessary for Labs)
  + Installation instructions for [PyTorch](https://pytorch.org/get-started/locally/) (Necessary for Labs)
  + [Installation of python packages](https://docs.python.org/3/installing/index.html) (Necessary for Labs)
  + [Anaconda](https://www.anaconda.com/download/) to install necessary packages (Highly recommended)
  + [Consider using a virtual environment](https://realpython.com/python-virtual-environments-a-primer/) (Advanced and not necessary)
* The lab materials will be distributed via Brightspace before the lab session

### Prerequisites

### **A strong foundation in linear algebra, vector calculus, and introductory probability (standard probability distributions, continuous and discrete variables, expectations, and conditional expectation). Mathematical maturity and comfort with coding algorithms is required.**

### **Please assess your fit for the course by attempting Homework 0**. Homework 0 will be provided to you at the end of the first lecture. If most of the questions are not approachable, you may not have the right math background for this course and will likely struggle. **Note that Homework 0 will be considered a part of your final grade. It will have a 10% contribution towards it.**

* Required
* Data Structures (CSCI-UA.102)
* Linear Algebra (MATH-UA.140)
* Probability and Statistics (MATH-UA.235)
* Recommended
* CSCI-UA 310 Basic Algorithms
* DS-GA 1001 Introduction to Data Science
  + [Exercise materials](http://nbviewer.jupyter.org/github/briandalessandro/DataScienceCourse/tree/master/ipython/) are highly recommended.
* DS-GA 1002 Statistical and Mathematical Methods

# Schedule

Note that the schedule below is only a guideline. The contents of each lecture might change as the course progresses.

|  |  |  |
| --- | --- | --- |
| **Week** | **Tuesday** | **Thursday** |
| 1 (08/30) | -- | Course Introduction  Overview of Machine Learning |
| 2 (09/06) | Feasibility of learning   * Hypothesis space * Hoeffding Inequality * Bounding Out-of-sample errors * Two questions while learning * Learning in practice * Error measures * Noise | Linear Parametric Methods: Regression   * Loss function: Mean Squared Errors * Closed Form Solution * Overfitting * Subset selection and regularization   Overfitting   * Reasons for its existence * Regularization methods * Validation set |
| 3 (09/13) | Linear Parametric Methods: Classification   * Logistic Regression * Loss functions: for binary and multi-class classification * Iterative Optimization   + Batch gradient descent   + Stochastic gradient descent | (Lab 1) Introduction and Linear Models for Regression (mean squared error loss) |
| 4 (09/20) | Probabilistic Interpretation   * Maximum Likelihood Estimation * Maximum a Posteriori Estimation * Bayesian Statistics | (Lab 2) Linear Models for Classification: binary output (logistic regression), and multi-class output (softmax) |
| 5 (09/27) | More Linear Methods for Classification   * Support Vector Machines * Radial Basis Function Networks * Gaussian Processes | (Lab 3) Support Vector Machines and Radial Basis Function Networks |
| 6 (10/4) | Non-Linear Parametric Methods - 1 (Deep Learning Models)   * Perceptron * Multi-Layer Perceptron: Deep Neural Networks * Backpropagation Algorithm | (Lab 4) Fully Connected Deep Neural Networks |
| 7 (10/11) | \*\*\* Fall Break \*\*\* | (Lab 5) TBD |
| 8 (10/18) | Non-Linear Parametric Methods - 2 (Deep Learning Models)   * Convolutional Neural Networks | (Lab 6) Convolutional Neural Networks |
| 9 (10/25) | Non-Linear Parametric Methods - 3   * Decision Trees * Random Forests | (Lab 7) Random Forests |
| 10 (11/1) | Non-Parametric Models   * Nearest Neighbor Classification * Parzen Window Classification | (Lab 8) Nearest Neighbor Classifier |
| 11 (11/8) | Unsupervised Learning - 1   * Clustering: K-means * Mixture of Gaussians * The EM Algorithm | (Lab 9) Dimensionality Reduction and K-means Clustering |
| 12 (11/15) | Unsupervised Learning - 2   * Dimensionality Reduction * Principal Component Analysis * Matrix Factorization * Non-negative Matrix Factorization | (Lab 10) Matrix factorization |
| 13 (11/22) | Ensemble Learning   * Boosting * Gradient Boosted Decision Trees * Learning Ensembles | (Lab 11) Model Ensembles |
| 14 (11/29) | Reinforcement Learning | \*\*\* Thanksgiving \*\*\* |
| 15 (12/06) | Other Topics   * Self-Supervised Learning * Dealing with Sequential Data | (Lab 12) Self-supervised Learning using Sequential Data |
| 16 (12/13) | Pre-Exam Q&A | \*\*\* No Class \*\*\* |

# University Policies

### Academic Integrity

Work you submit should be your own. Please consult the CAS academic integrity policy for more information:<https://cas.nyu.edu/content/nyu-as/cas/academic-integrity.html>. Penalties for violations of academic integrity may include failure of the course, suspension from the University, or even expulsion.

### Religious Observance

As a nonsectarian, inclusive institution, NYU policy permits members of any religious group to absent themselves from classes without penalty when required for compliance with their religious obligations. The policy and principles to be followed by students and faculty may be found here: The University Calendar Policy on Religious Holidays (<http://www.nyu.edu/about/policies-guidelines-compliance/policies-and-guidelines/university-calendar-policy-on-religious-holidays.html>)

### Disability Disclosure Statement

Academic accommodations are available to any student with a chronic, psychological, visual, mobility, learning disability, or who is deaf or hard of hearing. Students should please register with the Moses Center for Students with Disabilities at 212-998-4980.

NYU's Henry and Lucy Moses Center for Students with Disabilities

726 Broadway, 2nd Floor

New York, NY 10003-6675

Telephone: 212-998-4980

Voice/TTY Fax: 212-995-4114

Web site:<http://www.nyu.edu/csd>