

# Course Admin

---

EE-UY 4563/EL-GY 6143: INTRODUCTION TO MACHINE LEARNING  
PROF. SUNDEEP RANGAN

# People

---

❑ **Professor:** Sundeep Rangan, [srangan@nyu.edu](mailto:srangan@nyu.edu)

- 370 Jay St, 9<sup>th</sup> floor
- Office hours: Wednesdays 16:00 to 17:00 (in my office or on zoom)

❑ **Lecture:** Tuesdays 11:00-13:30, 370 Jay St, Room 202

- All lectures will be on zoom
- Attendance (in person or online) is OK. Attendance is not mandatory
- Bring your laptop. We will do in class exercises

❑ **Course Graders:**

- Ask for all questions regarding homeworks and labs

# Graders for Fall 2025

Grader name	Email	Allocation
Tejdeep Chippa	<a href="mailto:tejdeep.chippa@nyu.edu">tejdeep.chippa@nyu.edu</a>	Full
Roshan Nayak	<a href="mailto:rn2588@nyu.edu">rn2588@nyu.edu</a>	Full
Akshat Singh	<a href="mailto:akshat.singh@nyu.edu">akshat.singh@nyu.edu</a>	Full
Aadit Anand Fadia	<a href="mailto:aaf9407@nyu.edu">aaf9407@nyu.edu</a>	Full
Sheetal Prasad	<a href="mailto:sheetal.prasad@nyu.edu">sheetal.prasad@nyu.edu</a>	Half
Raunak Choudhary	<a href="mailto:rc5553@nyu.edu">rc5553@nyu.edu</a>	Half

- ☐ Graders will grade all labs, homeworks, and exams
- ☐ Please contact them if there are any questions
- ☐ Full graders will each have one hour of office hours per week

# Course Learning Objectives

---

- ❑ Formulate a task as a machine learning problem
  - Identify learning objectives, source of data, models, ...
- ❑ Load, pre-process and extract features from common data sources
  - images, text, audio, ...
- ❑ Mathematically describe simple models of the data
- ❑ Fit the models to data and use models for prediction and estimation
  - Use common tools
- ❑ Evaluate goodness of fit and refine models
- ❑ Evaluate the performance of methods using statistical techniques

# Grad vs Undergrad

---

- ❑ Class is simultaneously offered at the graduate and undergraduate level
- ❑ Undergrad EE-UY/CSE-UY 4563: Intro to Machine Learning
  - Covers fundamental algorithms and some analysis
  - In depth coverage of software tools including python, Google Cloud, Tensorflow
  - Python-based lab exercises + mandatory project
- ❑ Grad EL-GY 6143: Intro to Machine Learning
  - More algorithms and more mathematical analysis. Faster paced.
  - Software tools must be learned at home. Less coverage in class
  - Python-based lab exercises + optional project
- ❑ Lecture notes are mostly common with supplementary material for grad students indicated
- ❑ Many labs are common

# Texts and Other Resources

---

- ❑ Undergrad: James, Witten, Hastie and Tibshirani, “An Introduction to Statistical Learning”,
  - <http://www-bcf.usc.edu/~gareth/ISL/ISLR%20First%20Printing.pdf>
  - Very clear explanation of concepts.
  - But examples are in R. And there is no review of probability
- ❑ Grad: Hastie, Tibshirani, Friedman, “Elements of Statistical Learning”
  - <https://web.stanford.edu/~hastie/Papers/ESLII.pdf>
  - More advanced text with more analysis
- ❑ Tivandar Danka, Mathematics of Machine Learning
  - Nice background of all the mathematics you need
- ❑ Bishop, “Pattern Recognition and Machine Learning” (more advanced)

# More Resources

---

- ❑ Many classes online
- ❑ [Udacity](#)
- ❑ [Cornell CS4780](#)
  - YouTube videos
- ❑ [Berkeley Deep Learning](#)
  - Deep learning specifically
- ❑ Andrew Ng's machine learning class:
  - <https://www.coursera.org/learn/machine-learning>
  - A little less mathematical than this class
- ❑ Many, many others online...

Free

## Intro to Machine Learning

Course

This class will teach you the end-to-end process of investigating data through a machine learning lens, and you'll apply what you've learned to a real-world data set.

# Pre-Requisites

---

- ❑ Undergrad probability required for both UG and Grad version:
  - Basics of random variables, densities, Gaussian distributions, correlation, expectation, conditional densities, Bayes' theorem
  - Will provide a short review
  - NYU classes: Data analysis or Intro Probability are sufficient
- ❑ Undergraduate calculus and linear algebra
  - Vectors, matrices, partial derivatives, gradients.
  - Again, we will provide a brief review
- ❑ No machine learning experience is necessary
  - If you have ML experience, do NOT take this class.
  - Take Graduate probability (Fall) then Advanced machine learning (Spring)



# Pre-Requisites Programming

---

## ❑ Python

- All labs are in python, similar to object-oriented MATLAB, but many more libraries.
- And free!

## ❑ What you need to know

- You do not need to know python before class. But we will go over it quickly.
- You should have experience in some programming language (eg. MATLAB).
- You should know or being willing to learn object-oriented programming

## ❑ Resources:

- Installing python and ipython notebook (make sure you install Version 3.12)  
<http://jupyter-notebook-beginner-guide.readthedocs.io/en/latest/index.html>
- Python tutorial: <https://docs.python.org/3/tutorial/>
- Numpy: <http://cs231n.github.io/python-numpy-tutorial/>

# Grading: Undergraduate

---

- ❑ Midterm 1: 25%, Midterm 2: 25%, Labs, HW: 25%, Final project: 25%
- ❑ Labs: Simple python exercises
  - Given as jupyter notebook that you complete.
- ❑ Midterms
  - Each over approx. 3-4 weeks of material
  - Closed book with cheat sheet.
  - Follows homework and quiz problems + some very basic python questions
- ❑ Final project:
  - Use machine learning in some interesting way.
  - Must use data and python analysis.
  - Provide final report.

# Grading: Graduate

---

- ❑ Midterm 35%, Final 35%, Labs / HW 30%
  - Optional project: Up to 20%
- ❑ Labs: Simple python exercises
  - Given as jupyter notebook that you complete.
- ❑ Midterms & final
  - Each over approx. 6-7 weeks
  - Open book but no electronic aids.
  - Follows homework and quiz problems + some very basic python questions
- ❑ Optional final project:
  - Use machine learning in some interesting way.
  - Must use data and python analysis.
  - Provide final report.

# Lecture Sequence

Date	Prior Years		Fall 2025	
	Unit	Description	Unit	Description
9/2/2025	1, 2	Course Admin. What is ML? Multiple Linear Regression	1, 2	Course Admin. What is ML? Multiple Linear Regression
9/9/2025	3	Model Selection	3, 4	Model Selection, Regularization and LASSO
9/16/2025	4	Regularization and LASSO	5	Logistic Regression
9/23/2025	5	Logistic Regression	6	Nonlinear Optimization
9/30/2025	6	Nonlinear Optimization	8	KNN, Kernel methods [New]
10/7/2025		Midterm review	13	Decision Trees
10/14/2025		No class: Spring break		No class: Spring break
10/21/2025		Midterm		Midterm
10/28/2025	7	SVMs	8	Neural Networks
11/4/2025	8	Neural Networks	9	Convolutional neural networks
11/11/2025	9	Convolutional neural networks	10	PCA
11/18/2025	10	PCA	11	Word Embeddings and tokenization [New]
11/25/2025	11	Clustering and K-means	12	Clustering and K-means [Modified]
12/2/2025		Final Exam Review / Make Up		Final Exam Review / Make Up
12/9/2025		Final Exam		Final Exam

# Machine Learning Project

---

- ❑ Perform an interesting machine learning task of your choice
- ❑ Many possible areas:
  - Machine vision, brain-computer interfaces, natural language processing, sentiment analysis, ...
  - Anything that interests you
- ❑ Groups of 2 necessary
  - Submit all material as that group
- ❑ Use real data
  - write code
- ❑ Place all material in a github repo (including documentation) and submit only github repo

# Project Grading

---

## ☐ Formulation

- How well did you formulate the problem? Was it clear? Was that tied to the right objective?

## ☐ Approach

- Does your approach properly solve your problem? Was that made clear?

## ☐ Evaluation and Interpretation

- Did you comprehensively test the results? How well did you select / create the data?
- Did you test against alternative approaches?

## ☐ Presentation

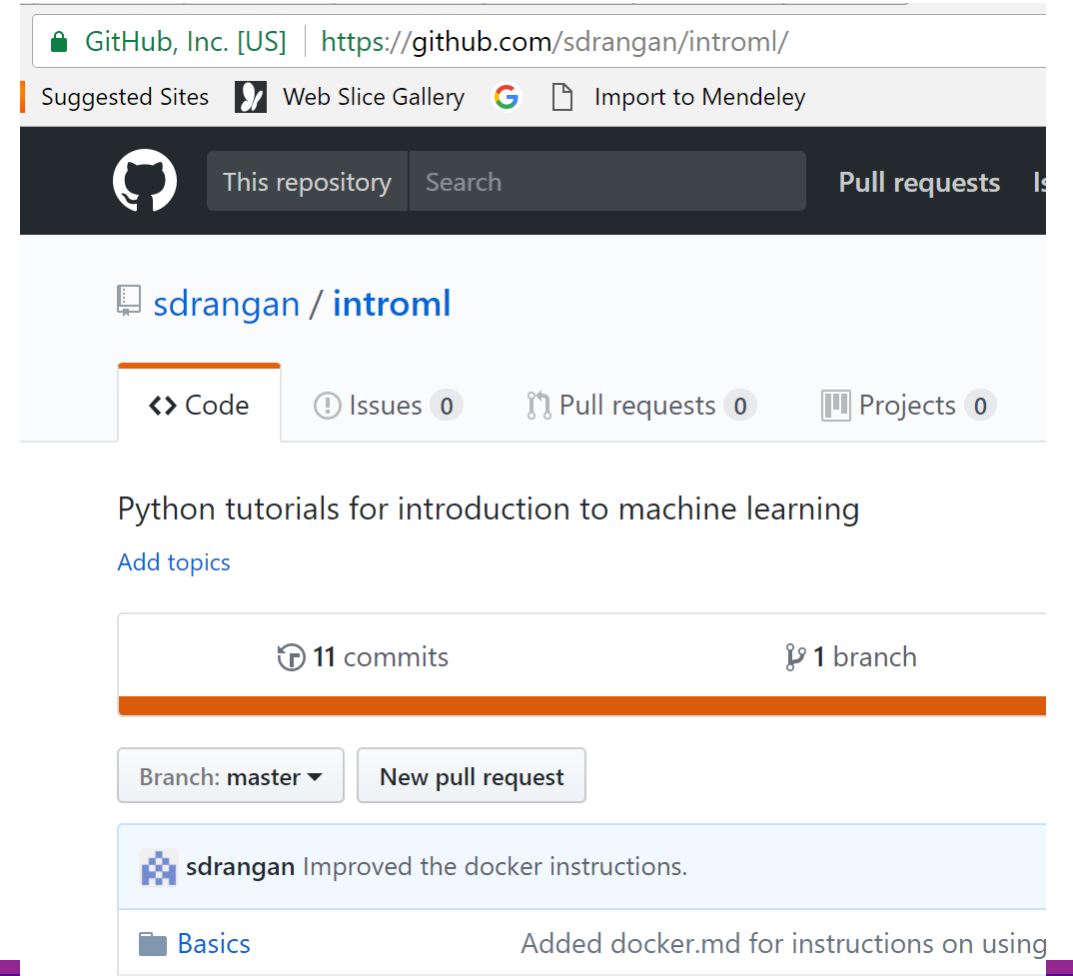
- Were the ideas clear? Were all the details conveyed. Did you highlight the main points?
- You can select a number of formats. Whatever makes sense. A github page

## ☐ Bonus

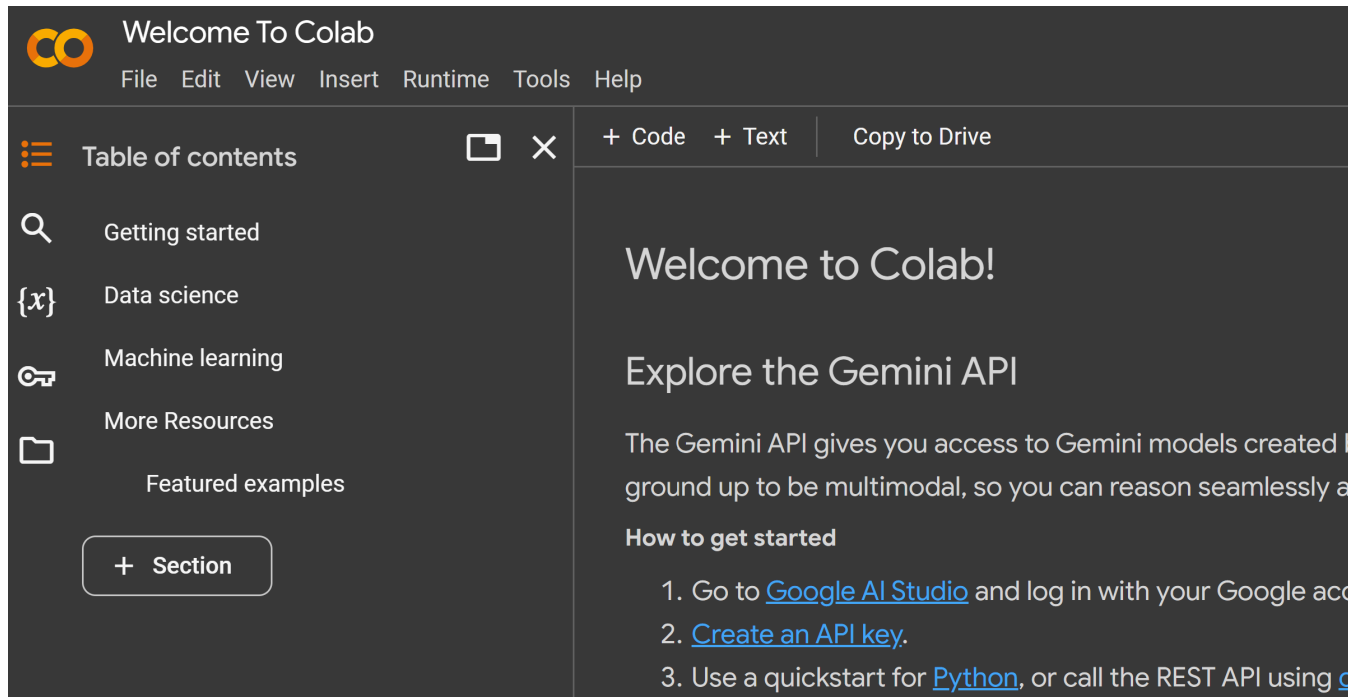
- Given for particularly hard / novel research

# Github

- ❑ Labs and demo posted on github
- ❑ <https://github.com/sdrangan/introml/>
- ❑ Also includes instruction for installing software
- ❑ Several tutorials of github on the web.
- ❑ Available on Windows, Mac and Unix.
- ❑ But, you can just clone the repo



# Google Colab



❑ Fully online python environment

❑ Many benefits

- Free to use
- Will work for all assignments and labs
- No machine to set up

❑ But it may be a bit slow

- Only matters for larger projects



# Software for Local Machine

---

- ❑ If you use your local machine, you will need to install several pieces of software:
- ❑ Python with various packages
  - Make sure you get 3.12
  - Anaconda
  - Jupyter notebook
  - See notes in NYU Classes
- ❑ Deep learning package
  - Either pytorch or Tensorflow
- ❑ Git hub
  - Guides: <https://guides.github.com/>
  - Available on Windows, Mac or Linux (including GCP instances)
  - All demos will be available on: <https://github.com/sdrangan/introml.git>

 PyTorch

 TensorFlow