

# Classical Machine Learning

## Week 0

### Plan

- Setting up your learning and programming environment

### Getting started

- [Setting up your ML environment \(Setup\\_NYU.ipynb\)](#)
  - [Choosing an ML environment](#)  
[\(Choosing an ML Environment NYU.ipynb\)](#)
- [Quick intro to the tools \(Getting\\_Started.ipynb\)](#)

# Week 1

## Plan

- Motivate Machine Learning
- Introduce notation used throughout course
- Plan for initial lectures
  - *What*: Introduce, motivate a model
  - *How*: How to use a model: function signature, code (API)
  - *Why*: Mathematical basis -- enhance understanding and ability to improve results
- [Course Overview \(Course overview NYU.ipynb\)](#)
- [Machine Learning: Overview \(ML Overview.ipynb\)](#)
- [Intro to Classical ML \(Intro Classical ML.ipynb\)](#)

## Using an AI Assistant

AI Assistants are often very good at coding.

But using one to just "get the answer" deprives you of a valuable tool

- you can ask the Assistant *why* it chose to do something
- keep on asking
- treat it as a private tutor !

[Learning about KNN using an Assistant as a private tutor  
\(https://www.perplexity.ai/search/using-python-and-sklearn-please-407oe3uzTXu1i9xEHVR2MQ\)](https://www.perplexity.ai/search/using-python-and-sklearn-please-407oe3uzTXu1i9xEHVR2MQ)

# Week 2

## Recap of last week

- [Summary of Intro to Supervised Machine Learning \(Intro to Supervised Learning Summary.ipynb\)](#)

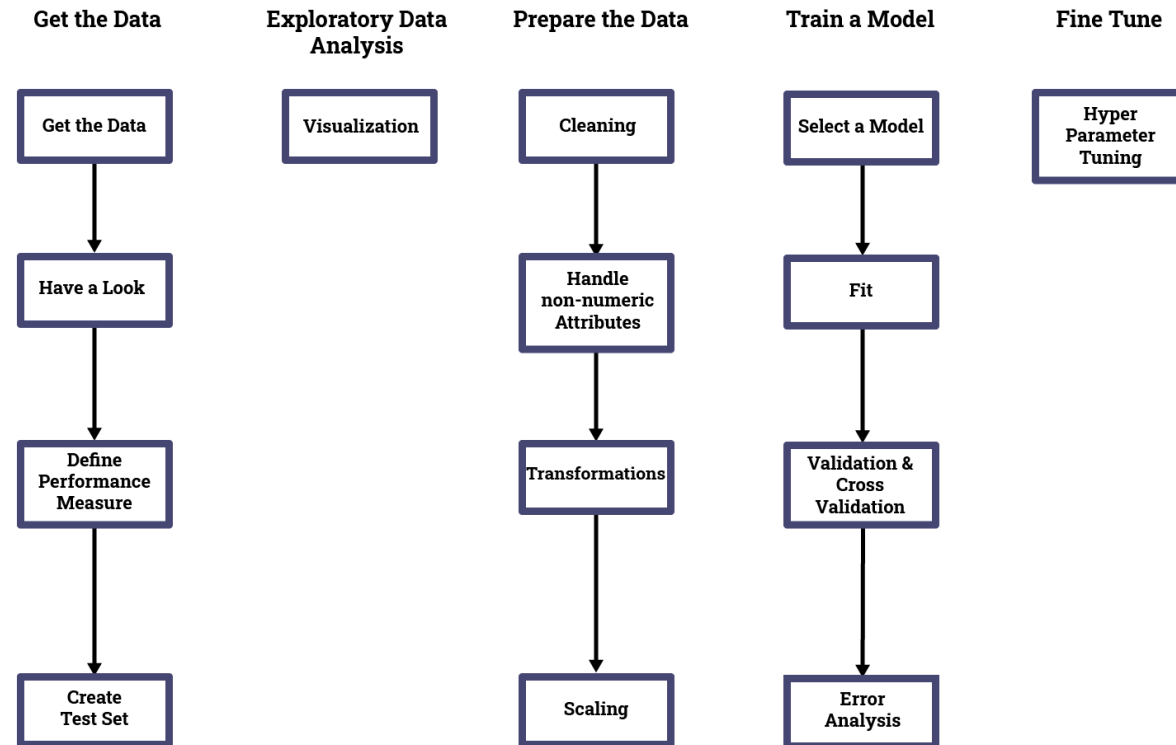
## Plan

We will learn the Recipe for Machine Learning, a disciplined approach to solving problems in Machine Learning.

We will illustrate the Recipe while, at the same time, introducing a model for the Regression task: Linear Regression.

Our coverage of the Recipe will be rapid and shallow (we use an extremely simple example for illustration).

I highly recommend reviewing and understanding this [Geron notebook \(external/handson-ml2/02\\_end\\_to\\_end\\_machine\\_learning\\_project.ipynb\)](#) in order to acquire a more in-depth appreciation of the Recipe.



Recipe, as illustrated by Linear Regression

[The Recipe for Machine Learning: Solving a Regression task \(Recipe via Linear Regression.ipynb\)](#)

- A *process* for Machine Learning
  - Go through the methodical, multi-step process
    - Quick first pass, followed by Deeper Dives

## Fitting a model: details

Recall: fitting a model (finding optimal value for the parameters) is found by minimizing a Loss function.

Let's examine a typical Loss function for Regression

- [Regression: Loss Function \(Linear Regression Loss Function.ipynb\)](#)

## Iterative training: when to stop

Increasing the number of parameters of a model improves in-sample fit (reduces Loss) but may compromise out-of-sample prediction (generalization).

We examine the issues of having too many/too few parameters.

- [When to stop iterating: Bias and Variance \(Bias and Variance.ipynb\)](#)

## Get the data: Fundamental Assumption of Machine Learning

- [Getting good training examples \(Recipe Training\\_data.ipynb\)](#)

## Regression: final thoughts (for now)

- [Regression: coda \(Regression\\_coda.ipynb\)](#)

## Deeper dives

- [Fine tuning techniques \(Fine\\_tuning.ipynb\)](#)

## Using an AI Assistant

[Learning about Linear Regression using an Assistant as private tutor \(https://www.perplexity.ai/search/using-python-sklearn-and-matplotlibTYv7oGdRO6upR5L5OSirg\)](https://www.perplexity.ai/search/using-python-sklearn-and-matplotlibTYv7oGdRO6upR5L5OSirg)

# Week 3

## Recipe "Prepare the Data" step: Transformations

We recap the importance of adding *synthetic* features to our Linear Regression example

- and *preview* the *mechanical* process of creating these features via *Transformations*

### Transformations

- [Prepare Data: Intro to Transformations \(Prepare\\_data\\_Overview.ipynb\)](#)

## Classification Task

### Plan

- We introduce a model for the Classification task: Logistic Regression
- How to deal with Categorical (non-numeric) variables
  - classification target
  - features

### Classification intro



- [Classification: Overview \(Classification\\_Overview.ipynb\)](#)
- [Classification and Categorical Variables \(Classification\\_Notebook\\_Overview.ipynb\)](#)
  - [linked notebook \(Classification and Non Numerical Data.ipynb\)](#)

**Categorical variables** (contained as subsections of Classification and Categorical Variables)

We examine the proper treatment of categorical variables (target or feature).

Along the way, we run into a subtle difficulty: the Dummy Variable Trap.

- [Classification and Categorical Variables: Categorical Variables \(Classification\\_Notebook\\_Overview.ipynb#Categorical-variables\)](#)
  - [Categorical variables, One Hot Encoding \(OHE\) \(Categorical\\_Variables.ipynb\)](#)

**Deeper dives**

- [Log odds \(Classification\\_Log\\_Odds.ipynb\)](#)

# Additional Deep Learning resources

Here are some resources that I have found very useful.

Some of them are very nitty-gritty, deep-in-the-weeds (even the "introductory" courses)

- For example: let's make believe PyTorch (or Keras/TensorFlow) didn't exist; let's invent Deep Learning without it !
  - You will gain a deeper appreciation and understanding by re-inventing that which you take for granted

[Andrej Karpathy course: Neural Networks, Zero to Hero \(https://karpathy.ai/zero-to-hero.html\)](https://karpathy.ai/zero-to-hero.html)

- PyTorch
- Introductory, but at a very deep level of understanding
  - you will get very deep into the weeds (hand-coding gradients !) but develop a deeper appreciation

**fast.ai**

`fast.ai` is a web-site with free courses from Jeremy Howard.

- PyTorch
- Introductory and courses "for coders"
- Same courses offered every few years, but sufficiently different so as to make it worthwhile to repeat the course !
  - [Practical Deep Learning](https://course.fast.ai/) (<https://course.fast.ai/>)
  - [Stable diffusion](https://course.fast.ai/Lessons/part2.html) (<https://course.fast.ai/Lessons/part2.html>)
    - Very detailed, nitty-gritty details (like Karpathy) that will give you a deeper appreciation

## Stefan Jansen: Machine Learning for Trading (<https://github.com/stefan-jansen/machine-learning-for-trading>)

An excellent github repo with notebooks

- using Deep Learning for trading
- Keras
- many notebooks are cleaner implementations of published models

# Assignments

Your assignments should follow the [Assignment Guidelines](#)  
([assignments/Assignment\\_Guidelines.ipynb](#)).

# Regression

- Assignment notebook: [Using Machine Learning for Hedging \(assignments/Regression%20task/Using\\_Machine\\_Learning\\_for\\_Hedging.ipynb\)](#)
- Data
  - There is an archive file containing the data
  - You can find it
    - Under the course page: Content --> Data --> Assignments --> Regression task
    - You won't be able to view the file in the browser, but you **will** be able to Download it
  - You should unzip this archive into the *the same directory* as the assignment notebook
  - The end result is that the directory should contain
    - The assignment notebook and a helper file
    - A directory named Data

```
In [1]: print("Done")
```

Done

