

# Introduction to Machine Learning

Amol Mavuduru





## How this event is going to work...

- **First part:** a brief slide presentation about what **machine learning** really is.
- **Second part:** a code-based workshop where we will learn how to train ML algorithms in **Python** using the famous Titanic dataset.



# For the Coding Section

Github: <https://github.com/AmolMavuduru/IntroToMachineLearning>

Necessary software/libraries (Anaconda should come with most of the packages listed below it):

- **Python and Anaconda** (recommended for managing these packages): <https://www.anaconda.com/download/#macos>
- **Numpy** (to install: **conda install numpy**)
- **Pandas** (to install: **conda install pandas**)
- **SciKit-Learn** (to install: **conda install sklearn**)
- **Matplotlib** (optional) (**conda install matplotlib**)
- **Seaborn** (optional) (**conda install seaborn**)



## For the Coding Section (if not using Anaconda):

- I would recommend using Anaconda but if you would prefer not to, you can install the libraries with pip.
- Jupyter: <http://jupyter.org/>
- Numpy (to install: **pip install numpy**)
- Pandas (to install: **pip install pandas**)
- SciKit-Learn (to install: **pip install sklearn**)
- Matplotlib (optional) (**pip install matplotlib**)
- Seaborn (optional) (**pip install seaborn**)



# What is Machine Learning??

- Machine learning is an area of **artificial intelligence** that is focused on using **mathematical** and **statistical** techniques to “**give machines the ability to learn**” without being explicitly programmed.
- Usually involves **some form of data** with **examples** to learn from.
- Commonly applied to problems where there is **no deterministic algorithm** or **100% accurate solution**.



# Applications of Machine Learning

- Self-driving cars
- Image recognition
- Speech recognition
- Natural language processing
- Medical diagnosis
- Predicting the weather
- Predicting stock prices
- Almost anything that involves making predictions or automating human tasks!



# How Machine Learning Works

- It all starts with **relevant data**
- Then we **train** an **algorithm** using **examples** from that data
- Next, we use this algorithm to **generate predictions** on **unseen data**



# Types of Machine Learning Problems

- Supervised Learning (labeled data)
  - Classification (predicting a discrete class)
  - Regression (predicting a continuous value)
- Unsupervised Learning (unlabeled data)
  - Clustering
  - Anomaly or outlier detection





# What problem are we focusing on?

- Titanic dataset
- Predict survival on the Titanic using passenger information
- Can also use our algorithms to predict survival chances