

# Introduction to Python and Machine Learning

Mar 2019

Nima Mahmoudi

# Outline

- Introduction to Python
  - What is Anaconda?
  - What is PIP?
  - Jupyter Notebook
  - Basics
  - Operators and Expressions
  - Control Flow
  - Functions
  - Modules
  - Object Oriented Programming
- Introduction to Machine Learning
  - Loading the Dataset
  - Summarizing the Dataset
  - Visualizing the Dataset
  - Prediction using Machine Learning
  - Evaluating the results

# Introduction to Python

- Free and Open-Source
- Readable, very clear syntax
- On average, 3.5 times smaller code than Java/C++
- No declarations of arguments or variables
- Dynamic declaration and use of variables
- Python is interpreted (not compiled)
  - Compiler: a lot of time for processing the code, faster execution
  - Interpreter: almost no processing of the code, slower execution

# What is Anaconda?

- Free and Open-Source
- Package Manager and Distribution of Python
- Optimized for Data Science and Machine Learning
  - `conda -V`
  - `conda create -n NAME python=3.6`
  - `conda env list`
  - `conda activate NAME` (or `activate NAME`)
  - `python --version`

# What is PIP?

- It is a package manager for python
- PIP stands for PIP Installs Python
- A package is all the files you need for a module
- A module is code libraries you can include in your project
- Anaconda automatically installs pip for you
  - `pip install jupyter`
  - `pip install pandas`
  - `pip install sklearn`
  - `pip install -r requirements.txt`

# Jupyter Notebook

- `cd WORKSHOP_DIR`
- `jupyter notebook`



# The basics of python

- jupyter notebook
- If you don't have jupyter notebook installed:
  - [https://github.com/nimamahmoudi/python-ml-workshop-2019/blob/master/introduction to python.ipynb](https://github.com/nimamahmoudi/python-ml-workshop-2019/blob/master/introduction%20to%20python.ipynb)

# Useful Links

- <https://github.com/nimamahmoudi/python-ml-workshop-2019>
- <https://www.anaconda.com>
- <https://www.python.org>
- <https://www.jetbrains.com/pycharm/>
- <https://jupyter.org/>
- <https://pypi.org>
- <https://pypi.org/project/pip/>
- <https://python.swaroopch.com/>
- <https://www.slideshare.net/saketkc/python-workshop-11152935>



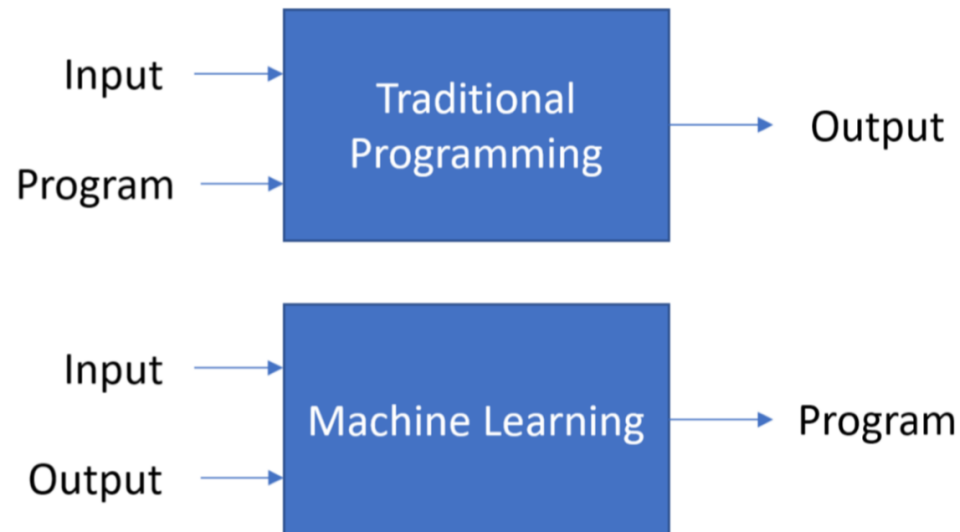
# Introduction to Machine Learning with Python

# Introduction to Machine Learning

- Machine learning is a field of artificial intelligence that uses statistical techniques to give computer systems the ability to “learn” (e.g., progressively improve performance on a specific task) from data, without being explicitly programmed.
  - Arthur Samuel, 1959
- A computer is said to learn from experience  $E$  with respect to some class of tasks  $T$  and performance measure  $P$ , if its performance at task  $T$ , as measured by  $P$ , improves with experience  $E$ .
  - Tom Mitchel, 1959

# Simple Definition

- We have a set of examples from which we want to extract regularities (patterns).



# Method Classification

- Supervised
  - The training data you feed to the algorithm includes the desired output.
  - E.g. Face recognition
- Unsupervised
  - The training data is unlabeled. Thus, the system tries to learn without a teacher.
  - E.g. Classify users into 10 groups
- Semi-Supervised
  - We have a lot of unlabeled data and a little bit of labeled data.
  - E.g. Google Photos

# Problem Types

- Classification

- In classification, the aim is to assign each input vector to one of finite number of discrete categories.
  - E.g. Email spam filtering (Spam or Ham)

- Regression

- If the desired output consists of one or more continuous variables, the task is called Regression.
  - E.g. Predict the price of a house

# Introduction to Machine Learning using Python

- jupyter notebook
- If you don't have jupyter notebook installed:
  - [https://github.com/nimamahmoudi/python-ml-workshop-2019/blob/master/introduction to ml.ipynb](https://github.com/nimamahmoudi/python-ml-workshop-2019/blob/master/introduction%20to%20ml.ipynb)

# Useful Links

- <https://github.com/nimamahmoudi/python-ml-workshop-2019>
- [https://en.wikipedia.org/wiki/Iris flower data set](https://en.wikipedia.org/wiki/Iris_flower_data_set)
- <https://machinelearningmastery.com/machine-learning-in-python-step-by-step/>
- <https://scikit-learn.org>
- <https://pandas.pydata.org>
- <https://matplotlib.org>
- <http://www.numpy.org>

# What's next?

- Check out the links
  - Look through each library
  - Try simple projects
- Take an online course!
- Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems
  - By Aurélien Géron
  - Great book, especially for beginners.