AI-ML Training

Prerequisite:

Hardware

Computer/Laptop intel Core i5 and above, with Hard Disk min 500gb and above and min Ram 8gb and above

Internet connect without any proxy settings, as it would be required to install certain Data Science packages during the training (e.g., nltk, keras, TensorFlow, Pytorch)

Software

1) Web Browser

2) Anaconda installed from the below link and the below versions.

Anaconda Packages: https://repo.anaconda.com/archive/

Anaconda for WinOS: Anaconda3-2020.11-Windows-x86_64.exe

Anaconda for MacOS: Anaconda3-2020.11-MacOSX-x86_64.pkg

Required skills: Python

Introduction to Data Science

It consists of 3 fields: AI, ML and DL.

Artificial intelligence:

- Mimic human behavior
- Incorporate the human behavior to machines
- The effort to automate intellectual tasks normally performed by humans.
- Example: Playing Chess, image classification, language translations and speech recognition.

Machine learning:

- Systematic study of algorithms and systems that improve their knowledge or performance with experience.
- Programming paradigm
- System is trained rather than explicitly programmed.
- Presented with many examples relevant to a task and it finds statistical structure in these examples which eventually allows the system to come up with rules for automating the task.
- Implicitly Rule learning/ Pattern learning
- Error is identified by calculating the difference between actual and predicted value.
- What is Pattern? Calculation done on input to get the output.

Example:

Input number   Output number

2= 4

3=6

4=8

5=?

Deep learning:

- Deep learning is a key technology behind driverless cars, enabling them to recognize a stop sign, or to distinguish a pedestrian from a lamppost.
-

What is Learning?

- We learn something in certain ways.
- How to humans learn: Remember, generalize and keep adapting to changing things
- We will incorporate these things into machines

Use cases:

- Adaptive behavior is incorporated by Artificial intelligence
- Remember and Generalize – ML
- Remember and Generalize – DL

Difference between ML and DL?

Machine learning is about computers being able to think and act with less human intervention; deep learning is about computers learning to think using structures modeled on the human brain. Machine learning requires less computing power; deep learning typically needs less ongoing human intervention.

ML Use cases:

- Spam filters
- Computer games
- Voice recognition
- Algorithm to decide whether bank will give loan or not

The need for ML?

- Making Data-Driven Decisions.
- Efficiency and Scale.
- Learning specific patterns from the data.

Application development:

- Traditional approaches.
- We need to explicitly mention the logic/ Rules to get the output.

Why do we need the ML for it?

Situation where the logic or rules are complex. If we increase the complexity of the problem, then we require ML to solve it.
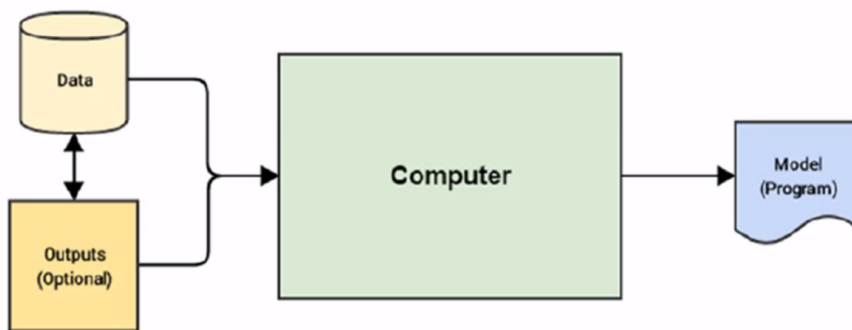
Input number   Output number

21234=435565

23454=435455

234=?

The simpler the better.

ML Paradigm:

Data + Output -> Computer -> Model (Program)

Machine Learning Paradigm

Traditional programming Paradigm:

Data -> Computer - Model (Program) -> Output

Why make machine learn?

Real world problem and business cases – Complex

- Lack of human expertise
- Scenarios and behavior can keep changing over time. Example: Infrastructe, network connectivity.
- Extremely difficult to explain or translate expertise into computational tasks.
- Address domain specific cases huge volumes of data with too many complex conditions and constraints.

Typical Machine learning tasks:

- Classification or categorization. Example: Identify dog, cat or any other classes

- Regression – Prediction of any number. Example: Salary prediction, housing prices prediction.
- Anomaly detection – identify unusual patterns. Examples: Fraud detection of credit card
- Translation – Feed data belonging to a specific language and translating it to other language.
- Clustering or grouping – To build categories. Example: solution to rubric cube
- Transcriptions – unstructured data like media, image, audio. Examples: Image to text translation.

Types of Machine learning:

- Supervised
- Unsupervised

Supervised learning:

Feed input and output data into the machine. Giving samples of data and then helping it to do the task.
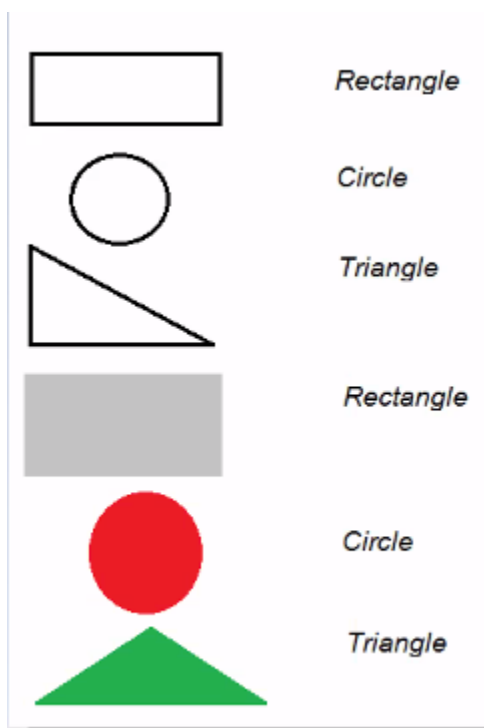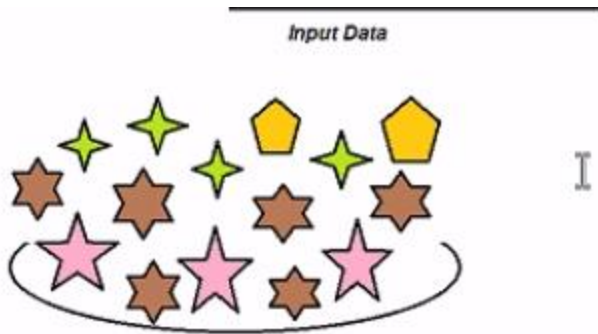

Unsupervised Learning:

Only based on the input data. Output data won't be provided.

Characteristic details of the data: Dimensions, size and color

Learning patterns among the data based on the similarities

Input Data

Check the anaconda python version:

It must be equal to or above 3.6 version.


```
Anaconda Prompt (Anaconda3) - python

(base) C:\Users\vsriniva>python
Python 3.8.8 (default, Apr 13 2021, 15:08:03) [MSC v.1916 64 bit (AMD64)] :: Anaconda, Inc. on win32
Type "help", "copyright", "credits" or "license" for more information.
>>>
```

Machine Learning with Python

Why Python?

- Its open source
- Easy to learn
- Interactive Environment
- Python has rich set of libraries that supports data science implementation.

Anaconda:

Python data science repository

IDE:

Supports the python packages

1) Jupyter Notebooks - Data exploration - Interactive Environment

2) Spyder – Algorithm Implementation – Scripting editor

3) Pycharm

4) Visual Code

5) Intellij

6) Eclipse

First two is the popular one. It has got interactive environment and scripting editor.

Jupyter Notebook:

Step 1: Create a Folder named "MLPractice" at any location of your choice

Step 2: WinOS

Start Anaconda Prompt

> **jupyter notebook --notebook-dir="C:\username\desktop\MLPractice"**

```
base) C:\Users\vsriniva>jupyter notebook --notebook-dir="C:\Users\vsriniva\Desktop\Data_science_training\ML_practice"
I 11:37:39.902 NotebookApp] The port 8888 is already in use, trying another port.
I 11:37:39.906 NotebookApp] The port 8889 is already in use, trying another port.
W 2022-03-07 11:37:43.587 LabApp] 'notebook_dir' has moved from NotebookApp to ServerApp. This config will be passed to ServerApp. Be sure to update your config before our next release
W 2022-03-07 11:37:43.587 LabApp] 'password' has moved from NotebookApp to ServerApp. This config will be passed to ServerApp. Be sure to update your config before our next release.
W 2022-03-07 11:37:43.589 LabApp] 'password' has moved from NotebookApp to ServerApp. This config will be passed to ServerApp. Be sure to update your config before our next release.
I 2022-03-07 11:37:43.603 LabApp] JupyterLab extension loaded from C:\Users\vsriniva\Anaconda3\lib\site-packages\jupyterlab
I 2022-03-07 11:37:43.603 LabApp] JupyterLab application directory is C:\Users\vsriniva\Anaconda3\share\jupyter\lab
I 11:37:43.614 NotebookApp] Serving notebooks from local directory: C:\Users\vsriniva\Desktop\Data_science_training\ML_practice
I 11:37:43.614 NotebookApp] Jupyter Notebook 6.3.0 is running at:
I 11:37:43.614 NotebookApp] http://localhost:8890/
I 11:37:43.615 NotebookApp] Use Control-C to stop this server and shut down all kernels (twice to skip confirmation).
_Lib: debug printing for files [.*] and level [100] is turned on
_Lib: debug printing for files [.*] and level [200] is turned on
_Lib: debug printing for files [.*] and level [300] is turned on
1752:.../engine/vf_shex/vf_shex.cpp(91): INFO: DllCanUnloadNow returned S_OK.
I 11:37:46.816 NotebookApp] 302 GET /tree (::1) 1.000000ms
I 11:37:50.112 NotebookApp] 302 POST /login?next=%2Ftree (::1) 122.000000ms
I 11:37:59.078 NotebookApp] Creating new notebook in
I 11:38:04.242 NotebookApp] Kernel started: 056d406e-741c-481b-b02b-65d6688d5b78, name: python3
```
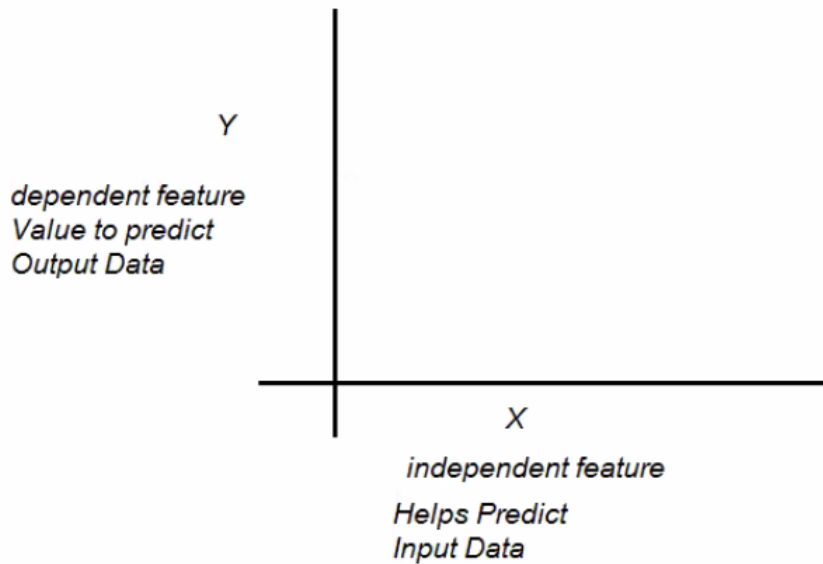
Practical session (Hands on):

Python basics

Python Packages for Machine learning

Understanding the X and Y variables:

Columns   Feature / Variable
Rows   Observation

Y

dependent feature
Value to predict
Output Data

X

independent feature

Helps Predict
Input Data

independent feature	*dependent feature*
Helps Predict *Input Data*	*Value to predict* *Output Data*
X	Y
YrsExp	Salary
YrsExp, Skills, Edu, City	Salary
Investment, Revenue, tax, state	Profit
Age, Income, City	Customer Response

Packages and Modules:

Modules:

- Modules in Python are simply Python files with a .py extension
- The name of the module will be the name of the file.

- A Python module can have a set of functions, classes or variables defined and implemented.
- These are prebuilt.

Example:

Module color (color.py)

Function red()

Function blue()

Function green()

We will see how to import the module and use it.

import color

Color.red()

Color.green()

OR

from color import red, blue

From color import *

Packages:

- Packages are namespaces which contain multiple packages and module themselves. They are simply directories.
- We create a directory "drawing" include module in it: color, line, rectangle, square and circle
- To use line module from drawing package

Import drawing.line

From drawing import circle

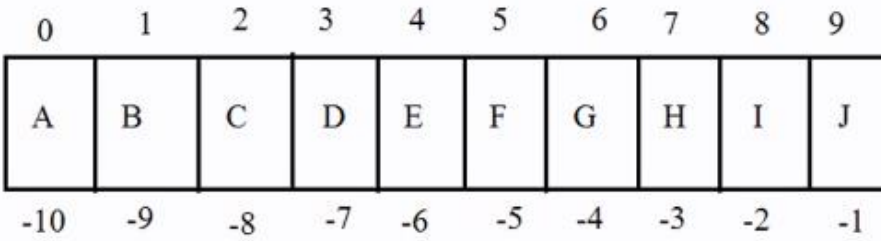Import matplotlib.pyplot as plt

From matplotlib import pyplot as plt2

To Install a new packages:

Conda install <package_name>

Or

pip install <package_name>

Slicing or traversing the array:

0	1	2	3	4	5	6	7	8	9
A	B	C	D	E	F	G	H	I	J
-10	-9	-8	-7	-6	-5	-4	-3	-2	-1

Syntax: letters[start: stop: step]

Positive direction: letters[start: stop: 1]

Negative direction: letters[start: stop: -1]

Difference between array and list:

Arrays	List
Homogeneous type storage	mixe datatype storage
Contigious mem. Single segment allocation in mem	non-Contigious mem. multiple segment allocation
Parallel execution	non-parallel execution
numy integrates C++ code in python and hence the c-compiler embedded interpretes it fast	no C++ code integrated
Uses vectorization algorithm no loops needed for oper.	Loop overheads for operations

2 Dimensional arrays:

0	1	2	3	4
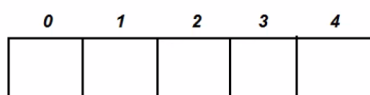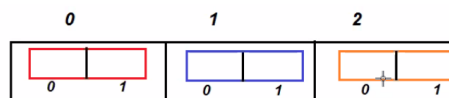
a   1d array

a.ndim   1
a.shape   (5,)

m   2d array

m.ndim   2
m.shape   (3,2)

0	1	2	3	4

a   **1d array**

a.ndim   1
a.shape   (5,)

0	1	2
0 1	0 1	0 1

m   **2d array**

m.ndim   2
m.shape   (3,2)

Column

     0   1

Row   0

      1

      2

matrix[rowIndex, colIndex]