UNIT-I :-

Introduction to Machine learning with python:

Introduction to machine learning, basic leaminology, Types of Machine learning and Applications, using python for Machine learning : Installing python and packages from the python package Index, Introduction to Numpy, Scipy, matplotlib and scikit-learn, Tiny application for Machine learning.

UNIT-II :-

Supervised learning :- Types of Supervised learning, Supervised Machine learning Algorithms : K-Nearest Neighbours, Regression models, Naive Bayes classifiers, Decision Trees, Ensembles of Decision Trees, Kernelised Support vector Machine, uncertainity estimates from classifiers.

UNIT-III :-

Building good training data sets : Dealing with missng data set into Separate training and test datasets, bringing features onto the same scale, selecting meaningful features, assesing feature importance with random forests. Compressing data via dimensionality reduction unsupervised dimensionality reduction via PCA, Supervised data compression via linear discriminant analysis.

UNIT-IV

Learning best practices for Model Evaluation and hyperparameter tuning :- Streamlining workflows with pipelines using K-fold cross validation to asses a model performance, debugging algorithms with learning and validation curves, fine tuning machine learning models via grid Search, looking at different performance evaluation metrics.

Combining different model for Ensemble learning: learning with ensembles, combining classifiers via majority vote, bagging - building an ensemble of classifiers from bootstrap samples, leveraging weak learners via adaptive boosting.

UNIT-V

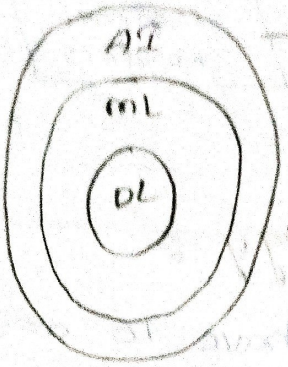
Working with Text Data (Data visualization):

Types of Data Represented as strings, Example Application Sentiment Analysis of Movie reviews, Representing Text Data as a Bag of words, stop words, Rescaling the data with tf-idf, Investigating Model Coefficients, Approaching a Machine ~~ensemble of classifiers~~ learning p

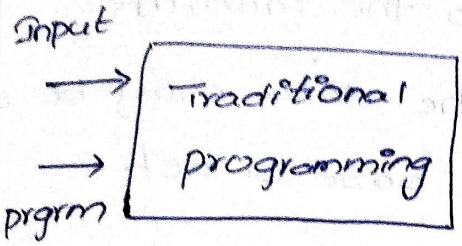
1/11/10.

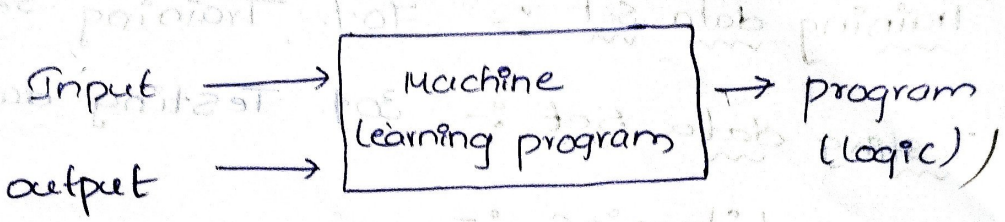
Machine Learning With Python

(1) Introduction to Machine learning :-



- Artificial Intelligence
- Machine Learning
- Deep Learning.

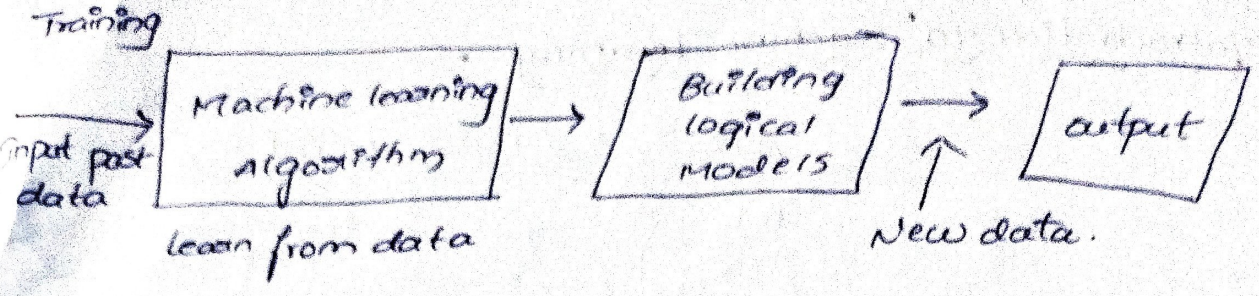




(1) Machine Learning :-

Sir Arthur Samuel an expert in AI, coined the term in Machine learning "it is a field of study that gives computers the capability to learn without being explicitly program."

- Tom M. Mitchell professor of Machine learning depaetement, carneige university, has defined ML has "A computer program is said to learn from experience E with respective to tasks T, and performance measure P, if its performance of tasks in T as measured by P, improves with experience E."



* Basic terms / Terminology :-

1. Label data :- We have to assign a correct information (or) data to the machine.

2. unlabel data :- In this the data may be either in pattern base or not.

3. Training data set :- 70% Training set

4. Testing data set :- 30% Testing data set.

Standard Libraries :- brief

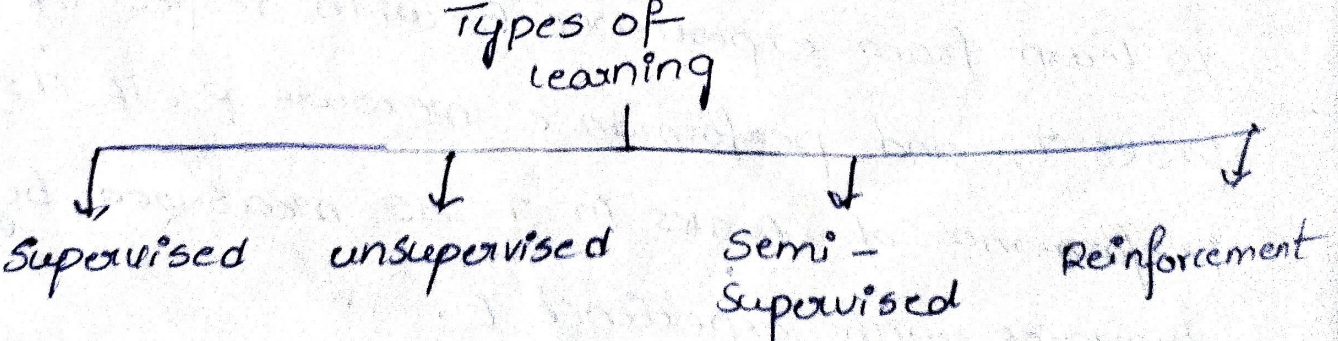
a) Numpy :- Mathematical operation

scipy :- Advanced calculation (or) scientific calculations.

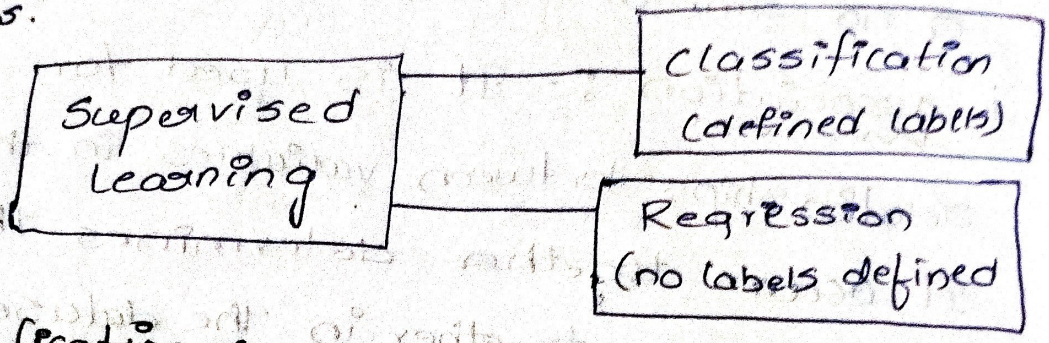
Matplotlib :- plots visualization.

Pandas :- Reading data set (.csv file)
csv - comma Separated value.

* Types of Machine learning :-



Supervised learning :- It is when the model is getting trained on a labelled dataset. A labelled dataset is one that has both input & output parameters.



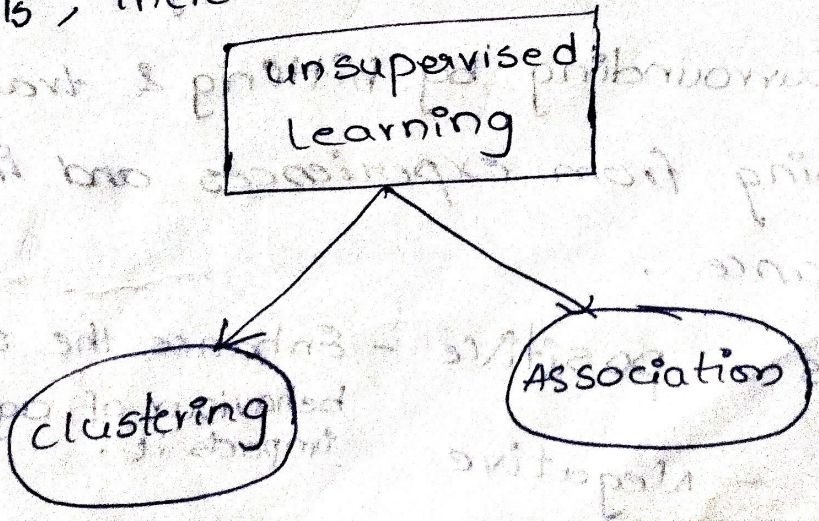
i, Classification :- It is a supervised learning task where output is having defined labels (discrete values). In binary classification, the model predicts either 0 or 1 ; yes or no but in case of multi-class classification the model predicts more than one class.

ii, Regression :- It is a Supervised Learning task where output having continuous value.

Example of SL :-

- Linear regression, Nearest Neighbor, SVM, etc.

2. unsupervised Learning :- It is different from the Supervised learning technique as its name suggests, there is no need for Supervision.



i. Clustering :- It is the method of grouping the objects into clusters such that objects with most similarities remains into a group and has less or no similarities with the objects of another group.

ii. Association :- It is used for finding the relationships between variables in the database. It occurs ~~together~~ determines the set of items that occurs together in the dataset.

3. Semi - Supervised learning :-

It is a type of Machine learning algorithm that lies between Supervised and unSupervised learning. It represents the intermediate ground between Supervised & unsupervised learning

To overcome the drawbacks of Supavised Learning and unsupervised learning algorithms, the concept of Semi-Supervised learning is introduced.

4. Reinforcement Learning :-

Reinforcement learning works on a feedback-based process, in which an AI agent automatically explore its Surrounding by hitting & trail, taking action, learning from experiences and improving its performance.

Categories – positive – Enhance the strength of behaviour of agent & positively impacts it.

– Negative.

↓

Exactly opposite to positive RL.

* Application of Machine Learning :-

29

1. Image Recognition :- It is one of the most common applications of machine learning. It is used to identify objects, persons, places etc. the popular use case of image recognition and face detection is Automatic friend tagging suggestion.

2. Speech Recognition :- It is a process of converting voice instruction into text, is also known as "speech to text" (or) computer search recognition.

3. Traffic prediction :- It predicts the traffic conditions such as whether traffic is cleared, slow-moving or heavily congested with 2 ways.

   Real Time location of the vehicle from google map app & sensors.

   Average time has taken on past days at the same time.

4. product recommendations :- It is widely used by various e-commerce & entertainment companies such as Amazon, Netflix etc. for product recommendation to the user.

5. Self-driving Cars :- Machine learning plays a significant role in self-driving cars. Tesla, the most popular car manufacturing company is working on self-driving car.

6. Email Spam and Malware Filtering :-

Whenever we receive a new email, it is filtered automatically as important, normal & spam. we always receive an important mail in our inbox with important symbols & spam emails in our spam box.

- content Filter
- Header Filter
- permission Filters
- Rule-based Filters.

7. Virtual personal Assistant :-

We have various personal assistants such as Google assistant, Alexa, cortana, Siri. As the name suggests, they help us in finding the information using our voice instruction.

8. Online Fraud Detection :-

Machine learning is making our transcation safe and secure by detecting fraud transcation. Whenever we perform some online transcations, there may be various ways fraudulent transcation such as Fake accounts, fake ids. Feed Forward Neural Network helps us by checking whether

9. Stock Market trading :-

Machine learning is widely used in stock market trading. In stock market, there is always a risk of up and downs in shares. In ML long short term memory neural network is used for the prediction of stock market trends.

10. Medical Diagnosis :-

ML is used for diseases diagnoss. With this medical technology is growing very fast and able to build 3D models that can predict the exact possition.

11. Automatic language Translation :-

The technology behind the automatic translation is a sequence to sequence learning algorithm, which is used with image recognition and translates the text from one language to another language.)

* Installing python and package from the python package Index :-

Python is one of the most famous and powerful languages in the world. It is a dynamically typed, high-level interpreted language intended for general use.

Step 1 : Install python :

To install python we first, need to visit the website python.org and click on the downloads tab.

Step2 : Download python package from any Repository

once we are done with python installation we can download any python package of our choice.

Step 3 :- Extract the python package :

After we have zip file of python package we need to extract in the location of our choice.

Step 4 : Copy the package in the site package Folder.

The location where we need to place this package folder is the site packages folder which one would find in the Lib folder at the location.

Step 5 : Install the package :

After placing the package folder in the reach of python that is in the site package folder.

import "package name".

Ex:- import pyautogui.

Mid

** Introduction to Numpy :-

Numpy is a general-purpose array-processing package.

* It provides high-performance multi-dimensional array object and tools for working with these arrays.

* It is the fundamental package for scientific computing with python.

* It is open-source Software.

- A powerful N- dimensional array object.
- Sophisticated (broad casting) functions.
- Tools for integrating C / C++ and Fortran code.
- useful linear algebra and random number capabilities.

- Mac and Linux user can install Numpy via pip command.
- windows does not have any package manager analogous to that in linux or mac.
- Numpy's main object is the homogeneous multidimensional array.
- It is a table of elements, all of same type, indexed by a tuple of positive integers.
- In Numpy dimensions are called axes. the number of axes is rank.
- Numpy's array class is called nd array. It is also known as by the alias array.

2. Array Creation :- There are various way to create array in Numpy.
   - Arrange :- returns evenly spaced values within a given interval. step size is specified.
   - linspace :- returns evenly spaced values within a given interval. num no. of elements are returned
   - Flatten array :- Flatten method to get a copy of array collapsed into one dimension.

3. Array Indexing :- Numpy offers many ways to do array indexing.

- Slicing :- Numpy arrays can be sliced. Need to specify a slice for each dimension of the array.
- Integer array indexing :- one to one mapping of corresponding element is done to construct a new arbitrary array.
- Boolean array Indexing :- It is used to ~~pick~~ when we want to pick elements from array which satisfy some condition.

4. Basic Operations :- Built-in arithmetic functions are provided in Numpy.

- Operations on single array :- we can use overloaded arithmetic operators to do element-wise operation on array to create new array.
- unary operators :- operations are provided as method of ndarray class. This includes sum, min, max etc.
- Binary operators :- These operations apply on array elementwise and a new array is created.
- universal functions (ufunc) :-

Numpy provides familiar mathematical functions such as sin cos exp etc.

5. Sorting array :- There is a np.sort method for sorting Numpy arrays.

* Scipy :-

Mid

* It is an open-source scientific library of python that is distributed under a BSD license.
* It is used to solve the complex scientific and mathematical problems.
* The scipy is pronounced as sigh pi, it depends on Numpy, including the appropriate and fast N-dimension array manipulation.
* The scipy library supports integration, gradient optimization, special functions, ordinary differential equation solvers, parallel programming tools.
* The scipy is a data-processing & system prototyping enviironment as similar to MATLAB.
* It is easy to use and provides great flexibility to scientist & engineers.
* Travis oliphant, Eric Jones and pearu peterson merged code they had written and called the new package scipy.
* The newly created package provided a standard collection of common numerical operation on the top of Numpy.
* Numpy and scipy both are used for mathematical and numerical analysis.
* Scipy library contains full featured version of the linear algebra module as well many other numerical

Mid

Matplotlib :-

* Matplotlib is multi-platform data visualisation library built on Numpy arrays and designed to work with the broader SciPy stack.
* It was introduced by John Hunter in the year 2002

Installation :-

```
python -mpip install -U matplotlib
```

Importing matplotlib :-

```
import matplotlib.pyplot as plt
```

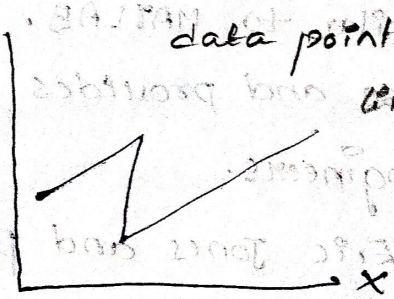
(or)

```
from matplotlib import pyplot as plt
```

Basic plots in Matplotlib :-

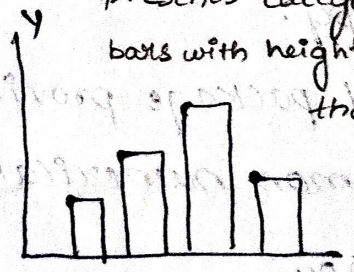
Matplotlib comes with variety of plots.

i) Line plot :- It is represented line format.



data points connected by straight line segments

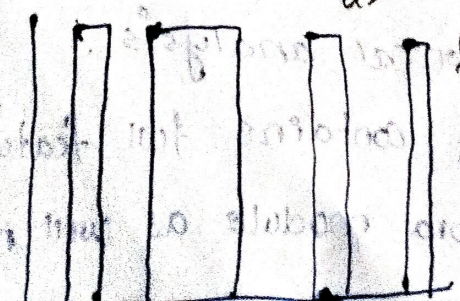
ii) Bar plot :-



presents categorical data with rectangular bars with heights & lengths proportional to the values that they represent.

iii) Histogram :-

* classifies data into various 'bins' or 'range groups'



used to represent the frequency distribution of a few data points of one variables.

iv. Scatter plot :-

Mathematical diagram using cartesian coordinates to display values for typically 2 variables for set of data.

* Scikit learn :-

It is a open-source python library that implements a range of machine learning, pre-processing, cross-validation and visualization algorithms using a unified interface. It used for data analysis & data-mining process.

Important features :-

- Simple and efficient tools for data mining and data analysis.
- It features various classification, regression and clustering algorithms including support vector machines, K-means etc.
- Accessible to everybody and reusable in various contexts.
- Built on the top of Numpy, Scipy and matplotlib.
- Open source, commercially usable - BSD license.

Installation :-

The latest version of scikit-learn is 1.1 and it requires python 3.8 or newer.

Scikit-learn requires :

- Numpy
- Scipy as its dependencies

```
pip install -U scikit-learn
```

* Tiny Machine Learning :-

- It is the field of study in ML and Embeded systems that explores the types of models you can run on small, low-powered devices like microcontrollers.

- It enables low-latency, low power and low-band width model at interference at edge devices.

Advantages of TinyML :-

i. Low latency :- The model runs on the edge, the data doesn't have to be sent to a server to run ~~interface~~ inference. This reduces latency of output.

ii. Low power Consumption :- Micro controllers consume very little power. This enables them to run without being charged for a really long time.

iii. Low Band width :- As the data doesn't have to be sent to the server constantly, less internet bandwidth is used.

iv. privacy :- The model is running on the edge, your data is not stored in any servers.

Applications :-

1. Industrial predictive Maintenance :-

* The predictive maintenance leads to significant cost savings.

* Machines are prone to fault.

* This device can alert the authorities according

potential issues even before it occurs.

2. Health care :-
   * It runs on solar power and can thus run indefinitely.
   * It works by detecting the mosquito breeding conditions and agitates the water to prevent it.

3. Agriculture :-
   * Machine learning models on the device using TensorFlow lite.
   * Since it works on the device, there is no need for an internet connection.

4. Ocean Life Conservation :-
   - Smart ML-powered devices are used to monitor whales in real-time in waterways around Settle and Vancouver to avoid whale strikes in busy shipping lanes.

Standard libraries :-

* pandas :- python library that is mainly used for data analysis. The users have to prepare the data set before using it for training the machine learning. pandas make it easy for the developers as it is developed specifically for data extraction. It has wide variety of tools for analysing data in detail, providing high-level data structures.