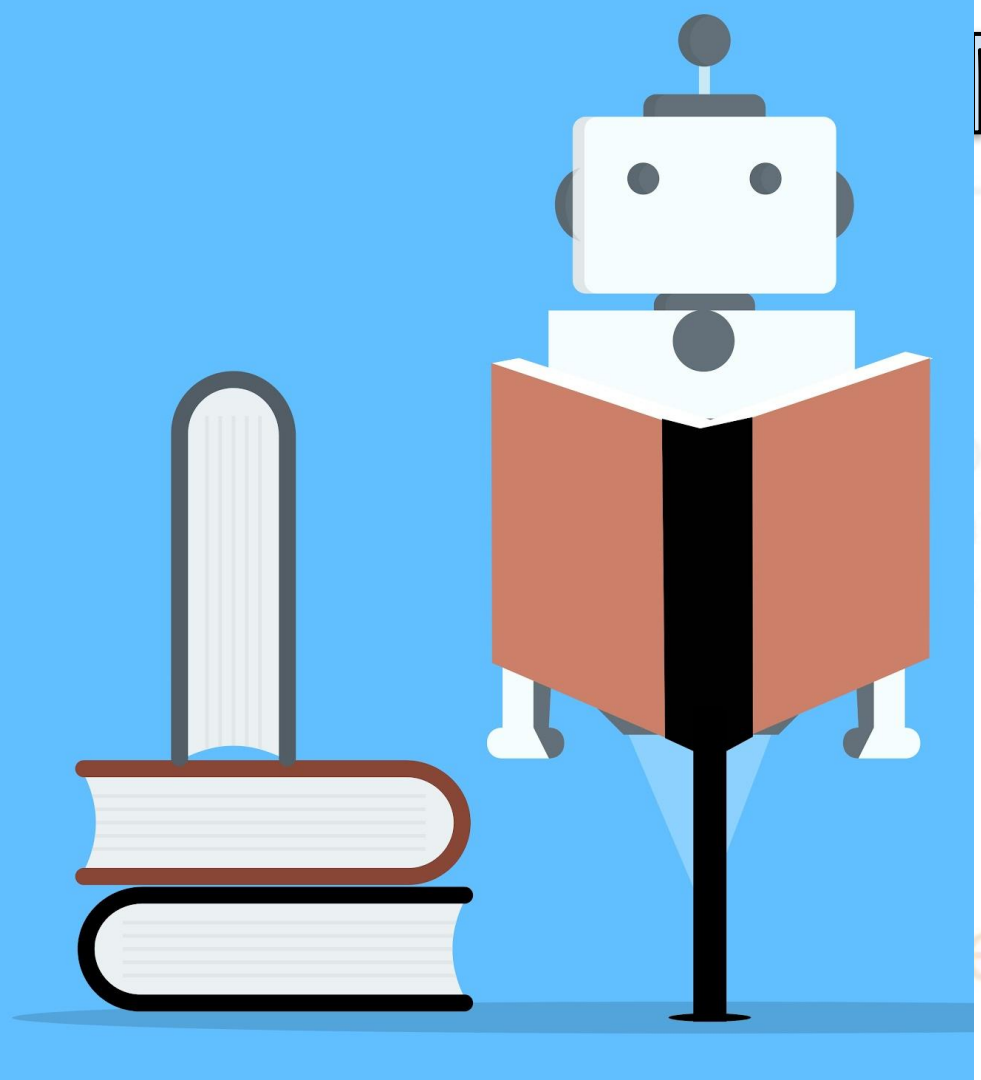


Mastering Machine Learning with Python

Ritvij Bharat Pvt. Ltd.

in association with

iHUB DivyaSampark, IIT Roorkee

(27th Aug 2024 - 18th Oct 2024)

Delivered by:

Shreyas Shukla

Overview

"Mastering Machine Learning with Python" is designed to provide you with a solid foundation in the field of machine learning

Machine Learning is an evolving field. This course will prove to be your necessary first step.

By the end of this course, you should be able to:

1. Learn python for Data Domain.
2. Use various Data libraries like Numpy, Pandas, MatplotLib, Seaborn, SkLearn etc.
3. Analyze data using Python through various real-life projects.
4. Understand the fundamental concepts of machine learning, such as supervised and unsupervised learning.
5. Apply different algorithms and techniques to solve real-world problems using Machine Learning.
6. Utilize popular Machine Learning libraries and frameworks.
7. Evaluate and interpret the performance of various Machine Learning models.

Structure

Python

1. Anaconda and Jupyter Installation and Familiarization
2. Python Basics
3. Statements
4. Functions
5. Object Oriented Programming

Data Analytics

1. Numpy
2. Pandas
3. MatplotLib
4. Seaborn
5. Data Preprocessing
6. Real World Projects

Machine Learning

1. Machine Learning Fundamentals
2. Supervised Learning:
   A. Linear Regression (Simple, Multiple, Polynomial etc.)
   B. Logistic Regression (KNNs, SVMs, Decision Trees etc.)
3. Unsupervised Learning:
   A. Clustering (K-Means, Hierarchical)
   B. DBSCAN
4. Dimensionality Reduction

Expectations

1. This course is designed to be very hands-on and interactive. You'll learn by doing, and we encourage you to actively participate in the class.

2. Don't be afraid to make mistakes and learn from them.

3. Throughout the course, you will have access to notes and code snippets.

4. Recordings of all the lectures will be made available to you.

How to approach the course :

1. For Slides: Listen Carefully. Do not focus too much in making notes. You can always go back and watch the recordings.
2. For Coding Session: Follow Along. Ask whenever you are stuck.
3. Be Consistent.

Practice. Practice. Practice.

- All the study materials such as video recordings, Code files, Slides etc. will be made available in the Telegram Group Only.
- Interact as much as you can, both in Telegram Group as well as during live sessions.

Python Overview

Let’s have a brief overview of:

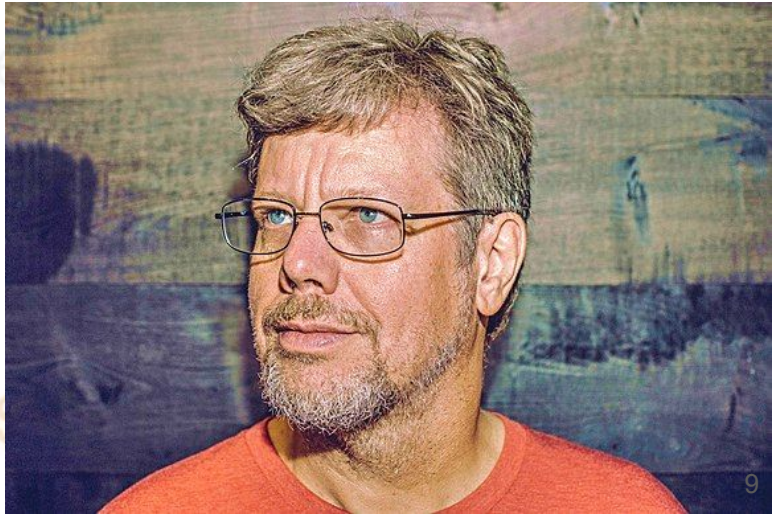
1. what Python is
2. why choose Python for programming
3. what you can do with Python.

Let’s have a brief overview of Python

This section in particular is geared towards people new to programming.

Brief History of Python

- Created in 1990 by Guido van Rossum
- Python 3 released in 2008
- Specifically designed as an easy to use language
- High focus on readability of code



Why Python?

- Clear, logical code that is easy to read and learn.
- Lots of existing libraries and frameworks allow users to apply Python to a wide variety of tasks.
- Focuses on optimizing developer time, rather than a computer's processing time.

Data Science and Machine Learning

- Analyze large data files
- Create visualizations
- Perform machine learning tasks
- Create and run predictive algorithms

Once you understand base Python and begin working with a few libraries, you'll quickly begin to see the vast potential Python has for your own Data Analytics projects

Python 2 vs Python 3

Choosing between Python 2 vs 3 used to be a very difficult decision for newcomers to the Python programming language.

Many companies still had legacy Python 2 code to be maintained.

The versions were similar enough that it was easy to learn both simultaneously.

Now every major external python package has been updated to support Python 3!

Python 3 is the future of Python.

We use Python 3 for this course.

Installing Python

Install Anaconda Distribution for Python.

- Anaconda installs Python **and** an easy to use development environment and navigator launch tool.

Briefly run Jupyter Notebook.

Explore “no install” online options.

To install Python we will use the free Individual Anaconda distribution.

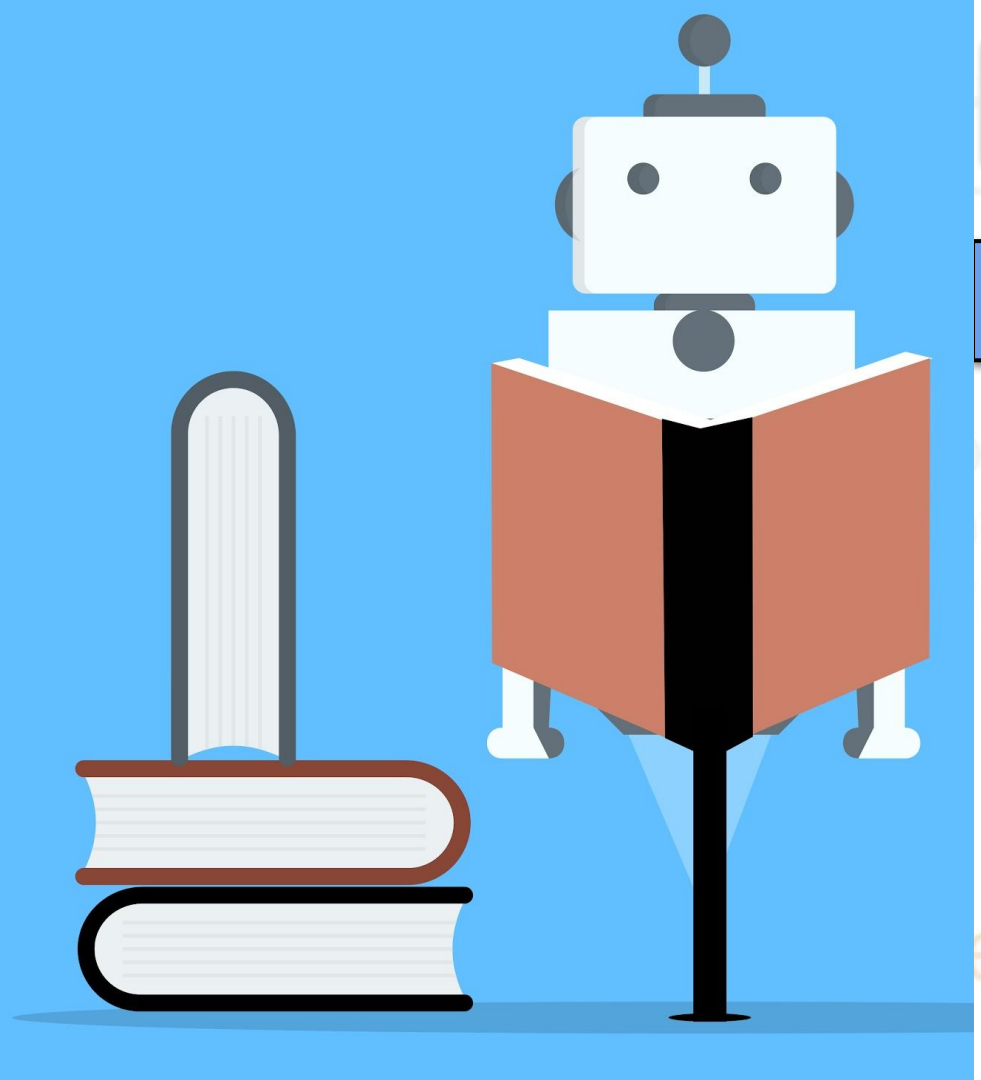
This distribution includes Python as well as many other useful libraries.

Anaconda can also easily be installed on to any major OS, Windows, MacOS, or Linux.

www.anaconda.com/downloads

Free “No Install” Options:

- jupyter.org/try
- Google Collab Online Notebooks
- Repl.it
  - Google Search:
    - “Python Interpreter Online”



Machine Learning

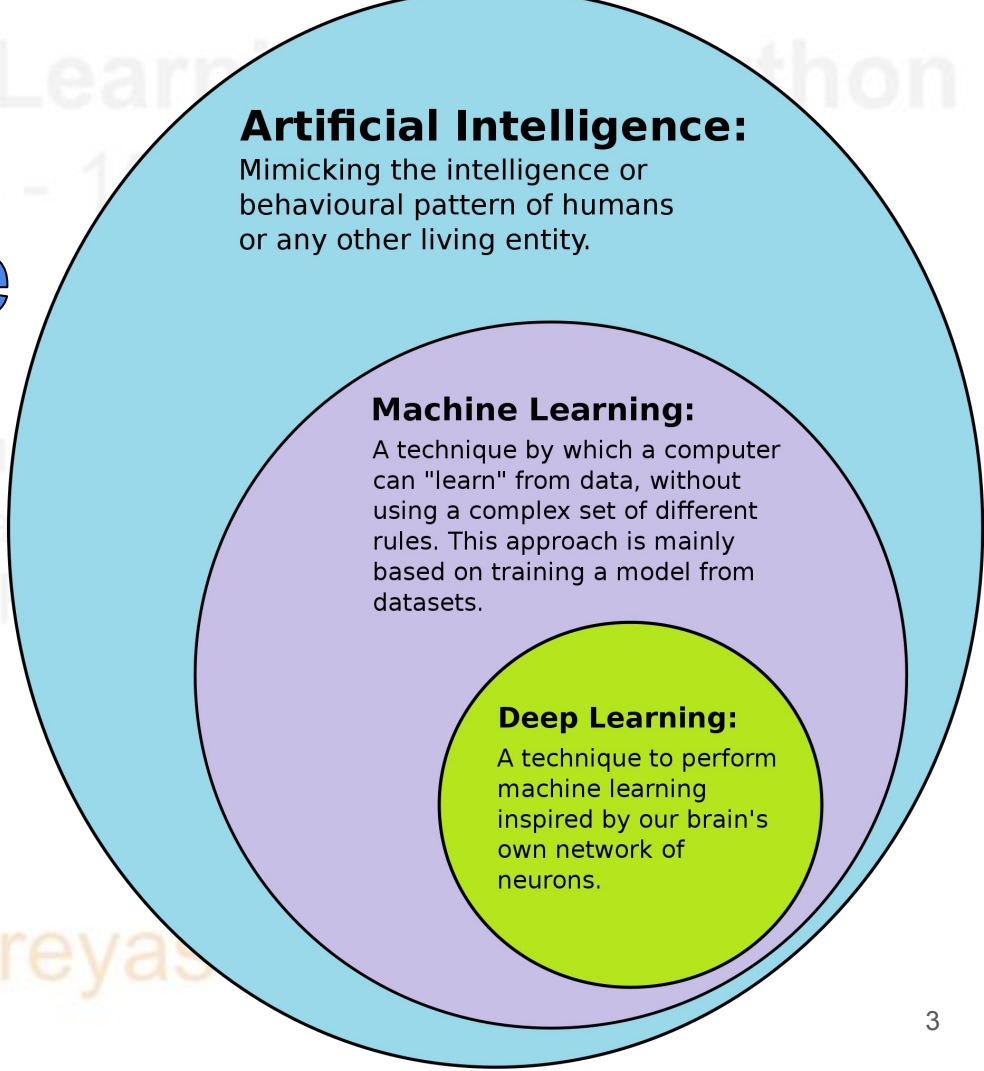
WELCOME!

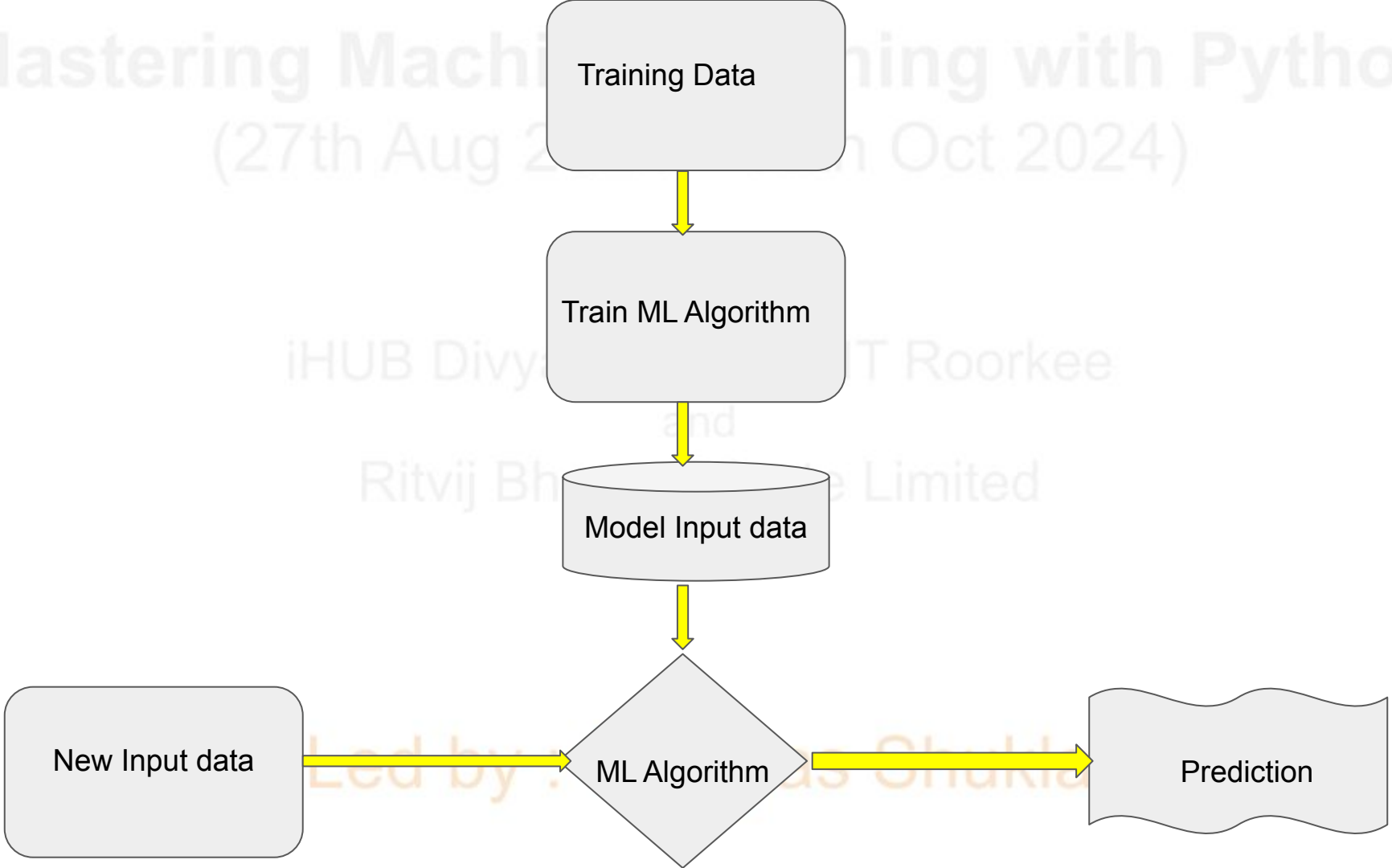
What is Machine Learning?

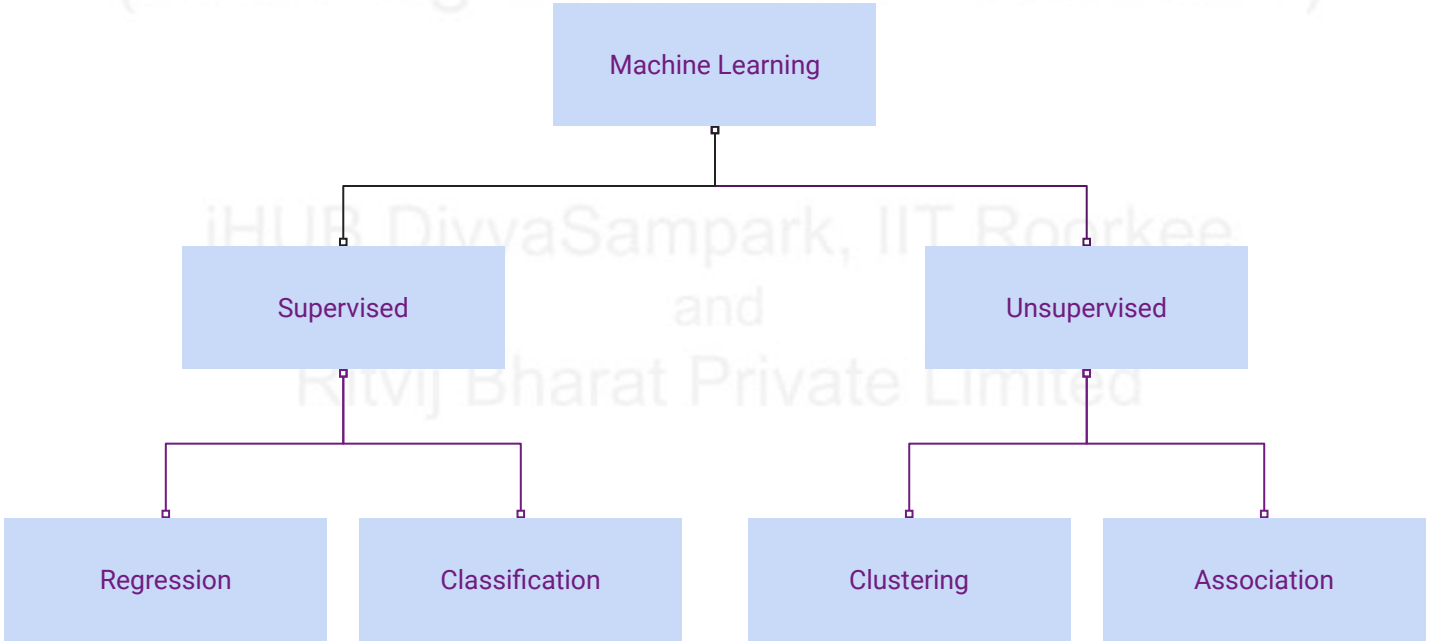
"Field of study that gives computers the ability to learn without being explicitly programmed"

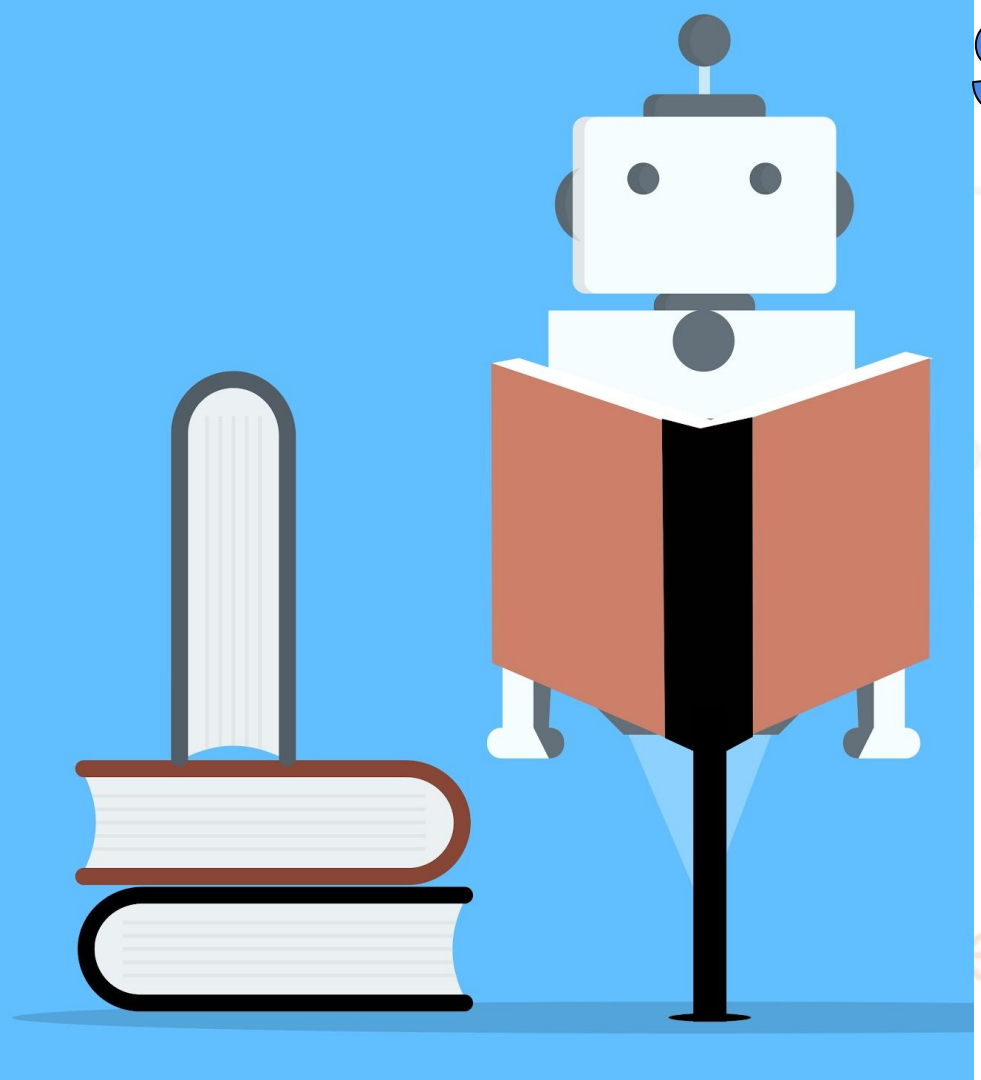
- Arthur Samuel (1959)

Artificial Intelligence vs Machine Learning vs Deep Learning



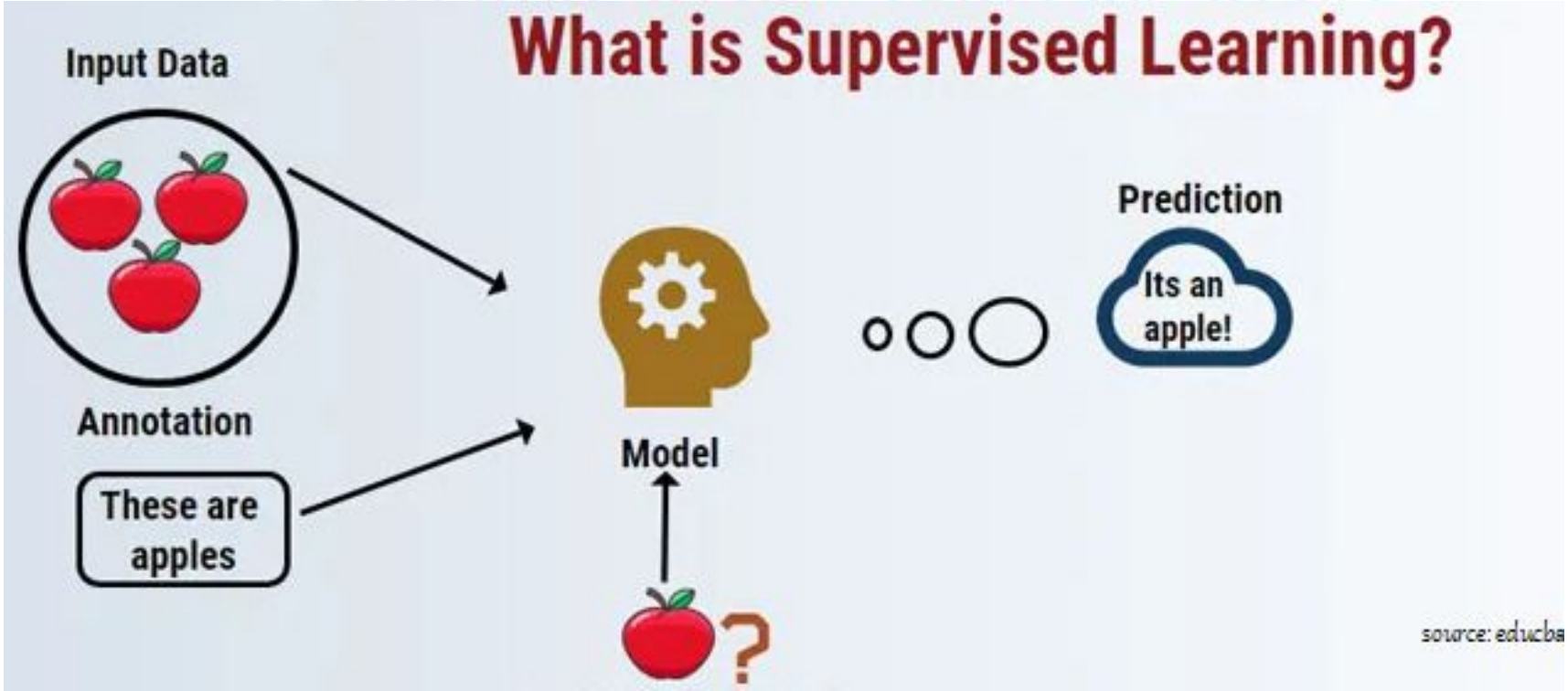






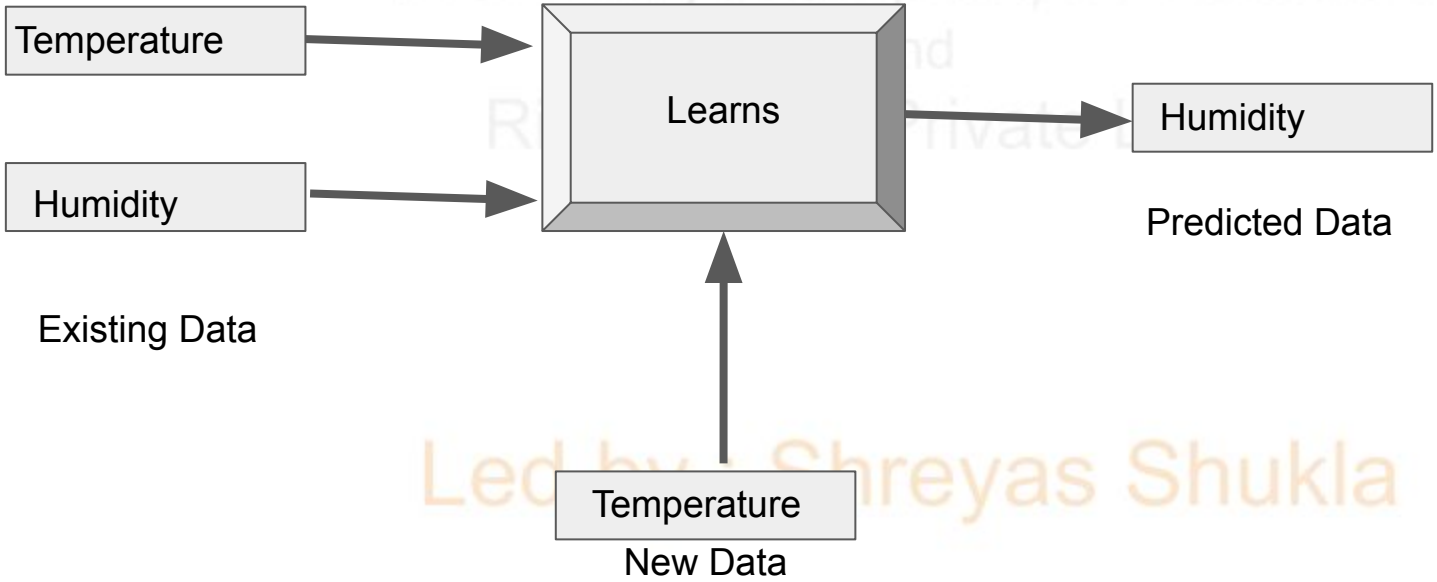
Supervised Learning

“Model able to predict with the help of labelled data”



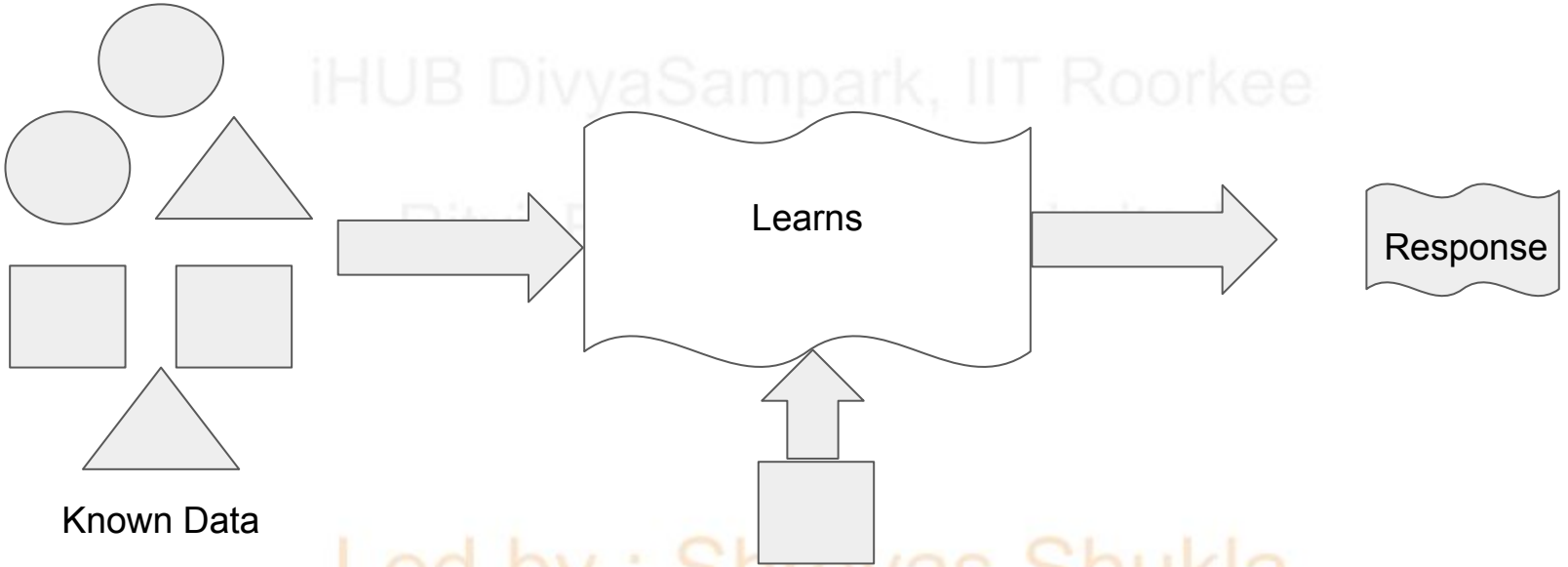
1. Regression

Relationship between two or more variables where a change in one variable leads to change in other



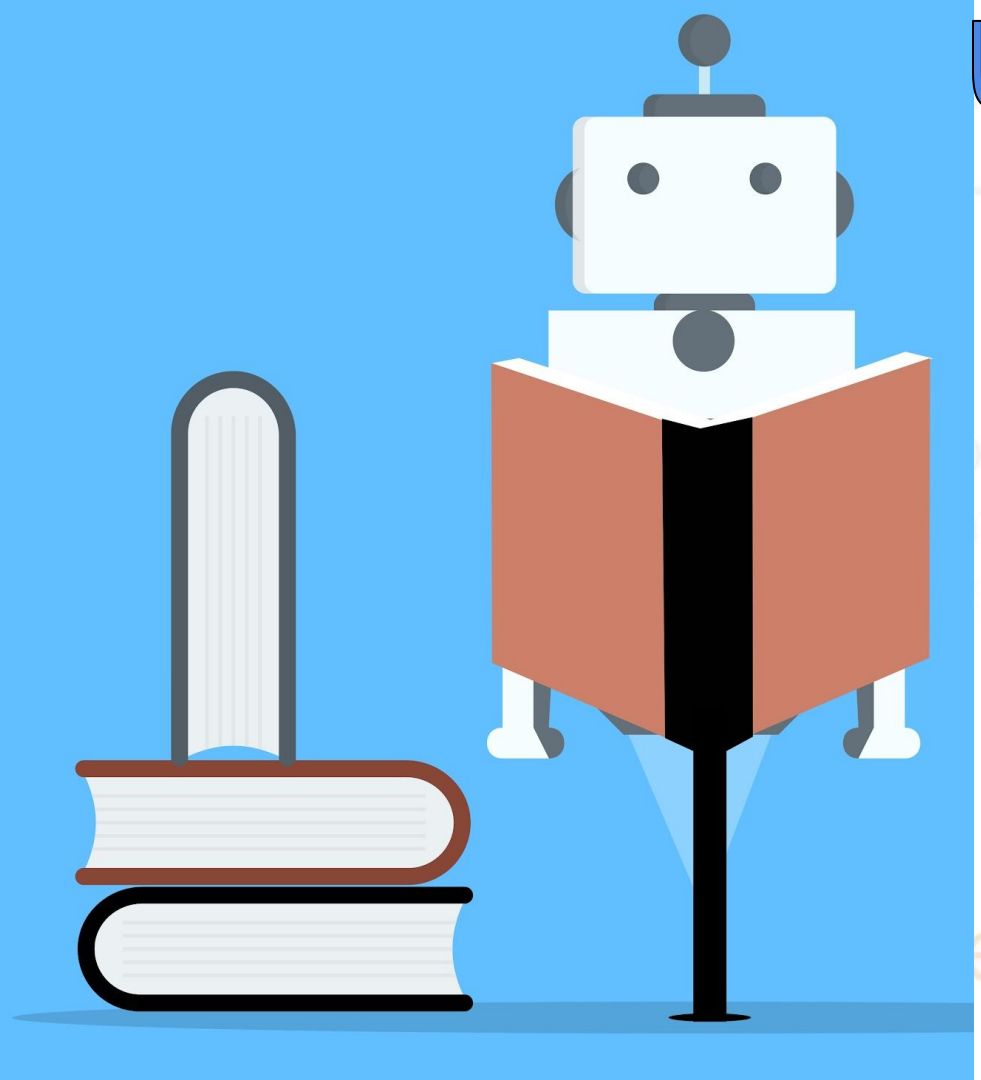
2. Classification

Output data is categorical i.e. with 2 or more classes (yes/no, win/lose/draw etc.)



Applications of Supervised Learning

1. Risk Assessment
2. Fraud Detection
3. Image Classification
4. Visual Recognition

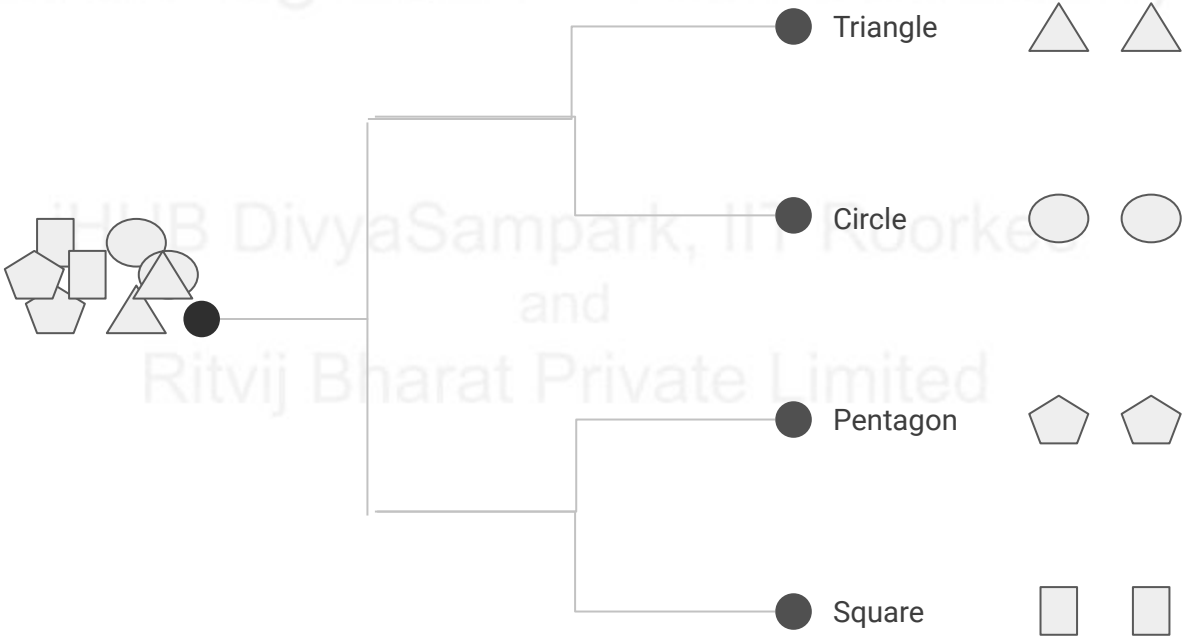


Unsupervised Learning

“Algorithm is trained using data that is unlabeled”


MUMBAI INDIANS HSSA
HSSA MUMBAI INDIANS
ACC

Unsupervised Learning

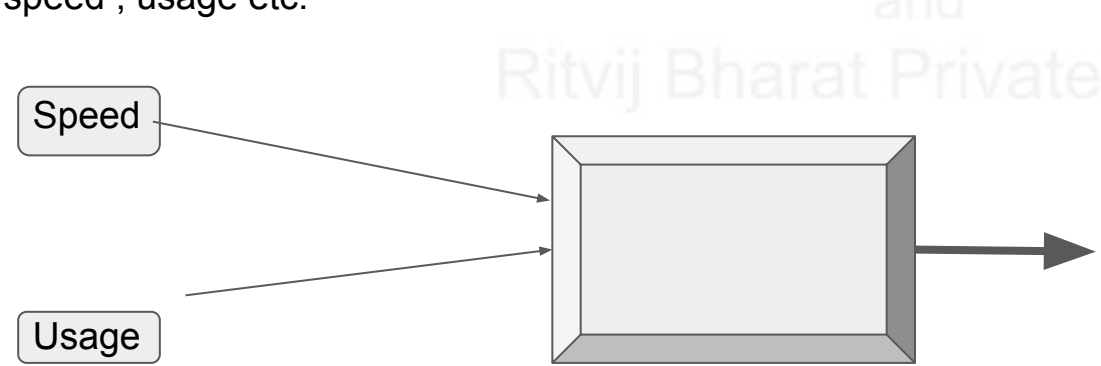


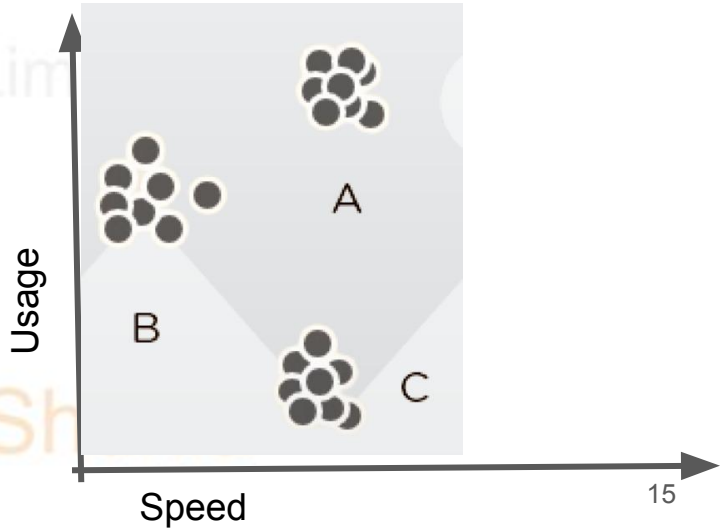
Group similar data together

1. Clustering

Method of dividing the objects into clusters which are similar between them and are dissimilar to the objects belonging to another cluster

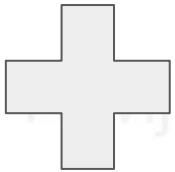
Suppose a Internet service provider wants to reduce its churn rate by providing personalized internet speed , usage etc.





2. Association

Customer 1
Bread
Milk
Butter
Wheat



Customer 2
Bread
Milk
Rice
Egg

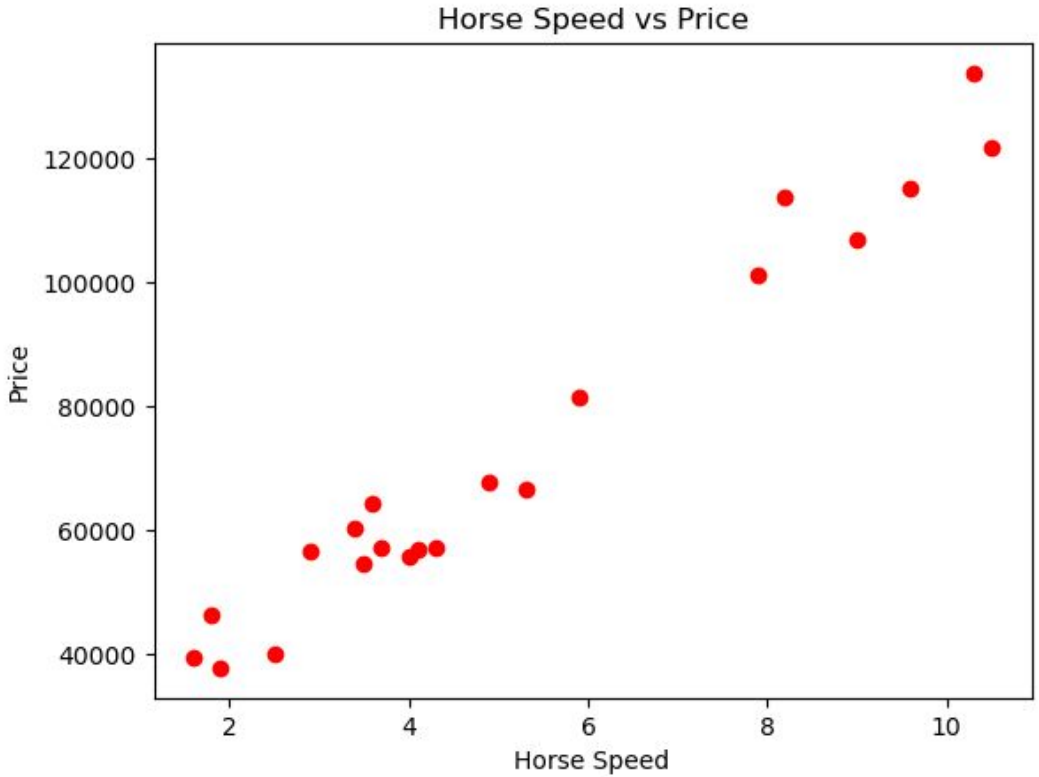
Customer 3

If a customer purchases milk, he is likely to purchase bread too?

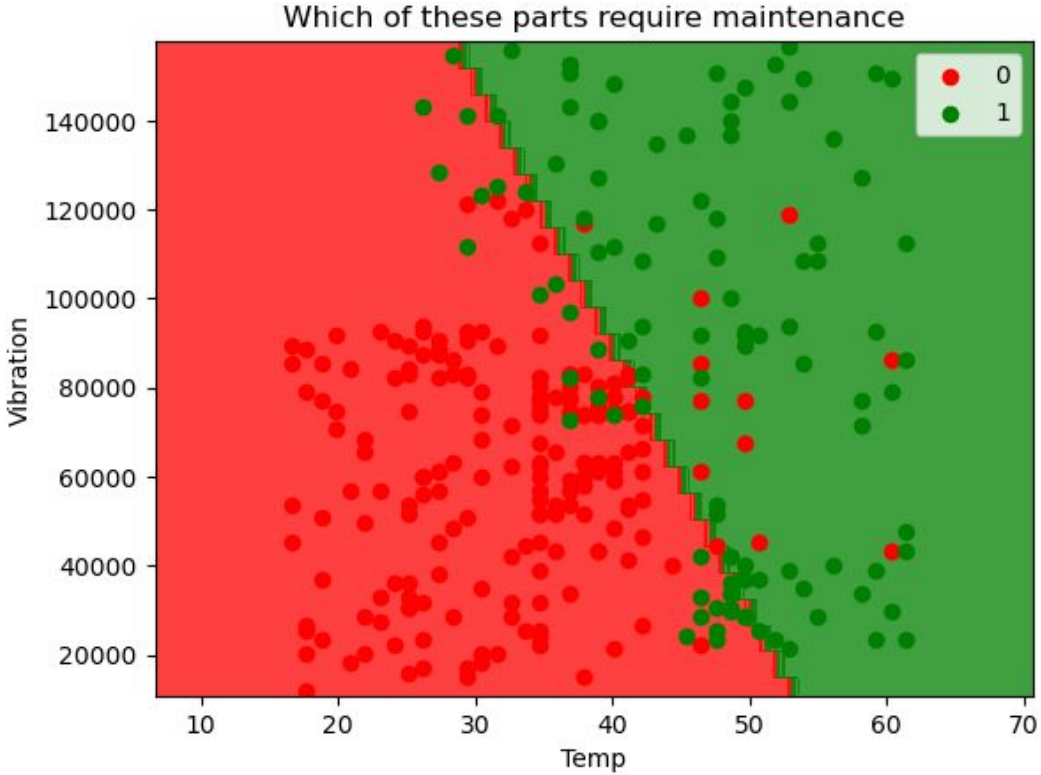
Applications of Unsupervised Learning

1. Market Basket Analysis
2. Semantic Clustering
3. Delivery Store Optimization

Regression



Classification



Real World

Collect & Store Data

Clean & Organize Data

Exploratory Data Analysis

Report

Visualization

Communication

Data Analysis