

# MACHINE LEARNING

SIX WEEKS SUMMER TRAINING

BY

TECHVANTO ACADEMY

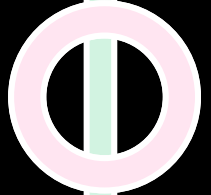
COURSE CODE : CSE443

SUBMITTED BY

NAME: MAMIDIPAKA VENKATA SAI NIKHIL

REGISTRATION NO: 12013998

SECTION: K20MD



## CERTIFICATE OF COMPLETION

This certifies that

**MAMIDIPAKA VENKATA SAI NIKHIL**

has efficiently completed **6 Weeks** Live Training in **Machine Learning** conducted by  
**Techvanto Academy, New Delhi**  
with an **A** grade on basis of overall performance and evaluation.

**We wish a great success for his/her future endeavours..!!**



**MR. SHEKHAR SAINI**  
Director

**#startupindia**



**REG. ID**

**TA30ML22063**

**Date: 25th May'22-10th July'22**






# About Summer Training :

We followed a structured methodology for our 6 Weeks Summer training Program by Techvanto Academy for Machine Learning using Python.



Program starts from Basics of the Python to the implementation of various Machine Learning algorithms like Linear Regression, Decision Tree, SVM and many more.



At the end prepared a Well- planned mini Project Loan Prediction(To determine whether the bank will approve the loan or not ) in which we will see how the machine makes predictions based on different columns of a dataset.



We have used various Machine Learning Algorithms for prediction and taken out which gives more accuracy.





# Why this course :

This course is a complete package that helped me learn Machine Learning and its Algorithms from basic to an advanced level. The course curriculum has been divided into 6 weeks where one can practice questions & attempt the assessment tests and clearing their doubts to their associated mentor.



The course offers me a wealth of programming challenges that will help me to prepare for interviews with top-notch companies like Microsoft, Amazon, Cisco etc.



Science is about learning machine through previous experiences. It is no less surprising that science progressed to the limit where a computer learned over time, but to mark your existence in the world as such, you know how to teach your machine.





# Machine Learning & it's importance:

- Machine learning (ML) is a type of artificial intelligence (AI) that allows software applications to become more accurate at predicting outcomes without being explicitly programmed to do so.
- ML applications are fed with new data, and they can independently learn, grow, develop, and adapt.
- Machine learning has become a significant competitive differentiator for many companies. It involves data exploration and pattern matching with minimal human intervention. Many of today's leading companies, such as Facebook, Google and Uber, make machine learning a central part of their operations.





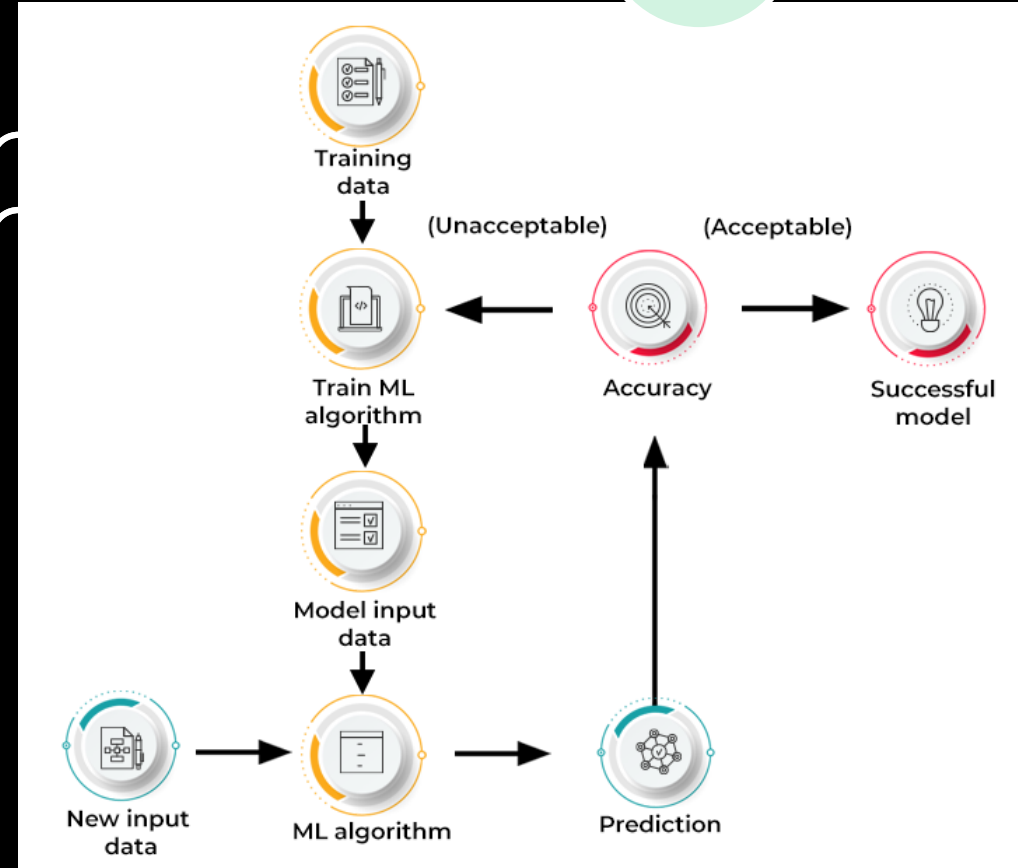
# How does Machine Learning work?

Machine learning algorithms are molded on a training dataset to create a model. As new input data is introduced to the trained ML algorithm, it uses the developed model to make a prediction.

Further, the prediction is checked for accuracy. Based on its accuracy, the ML algorithm is either deployed or trained repeatedly with an augmented training dataset until the desired accuracy is achieved.



# HOW MACHINE LEARNING WORKS



# TYPES OF MACHINE LEARNING



**Supervised  
Machine Learning**

**Unsupervised  
Machine Learning**

**Semi-Supervised  
Learning**

**Reinforcement  
Learning**



## 1. Supervised Learning:

In supervised learning technique, we train the machines using the "labelled" dataset, and based on the training, the machine predicts the output. Here, the labelled data specifies that some of the inputs are already mapped to the output.

- The main goal of the supervised learning technique is to map the input variable( $x$ ) with the output variable( $y$ ).
- Supervised machine learning can be classified into two types. They are:
  - i. Classification
  - ii. Regression

Applications: Image segmentation, Medical diagnosis, Fraud detection

## 2. Unsupervised Learning:

In unsupervised learning technique, we train the machine using the "unlabeled" dataset and the machine predicts the output without any supervision.

- The main aim of the unsupervised learning algorithm is to group or categories the unsorted dataset according to the similarities, patterns, and differences.
- Unsupervised learning can be further classified into two types. They are:
  - i. Clustering
  - ii. Association

Applications: Network Analysis, Recommendation Systems



### 3. Semi-Supervised Learning:

Semi-Supervised learning is a type of Machine Learning algorithm that lies between Supervised and Unsupervised machine learning and uses the combination of labelled and unlabeled datasets during the training period.

- The main aim of semi-supervised learning is to effectively use all the available data, rather than only labelled data like in supervised learning.

### 4. Reinforcement Learning:

Reinforcement learning works on a feedback-based process, in which an AI agent (A software component) automatically explore its surrounding by hitting & trail, taking action, learning from experiences, and improving its performance. Agent gets rewarded for each good action and get punished for each bad action; hence the goal of reinforcement learning agent is to maximize the rewards.

- Reinforcement learning is classified into two types. They are:
  - i. Positive Reinforcement Learning
  - ii. Negative Reinforcement Learning

Applications: Video games, Resource management, Robotics, Text mining.



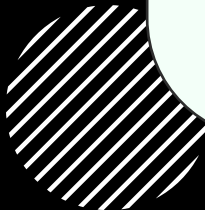


# **Project Description:**

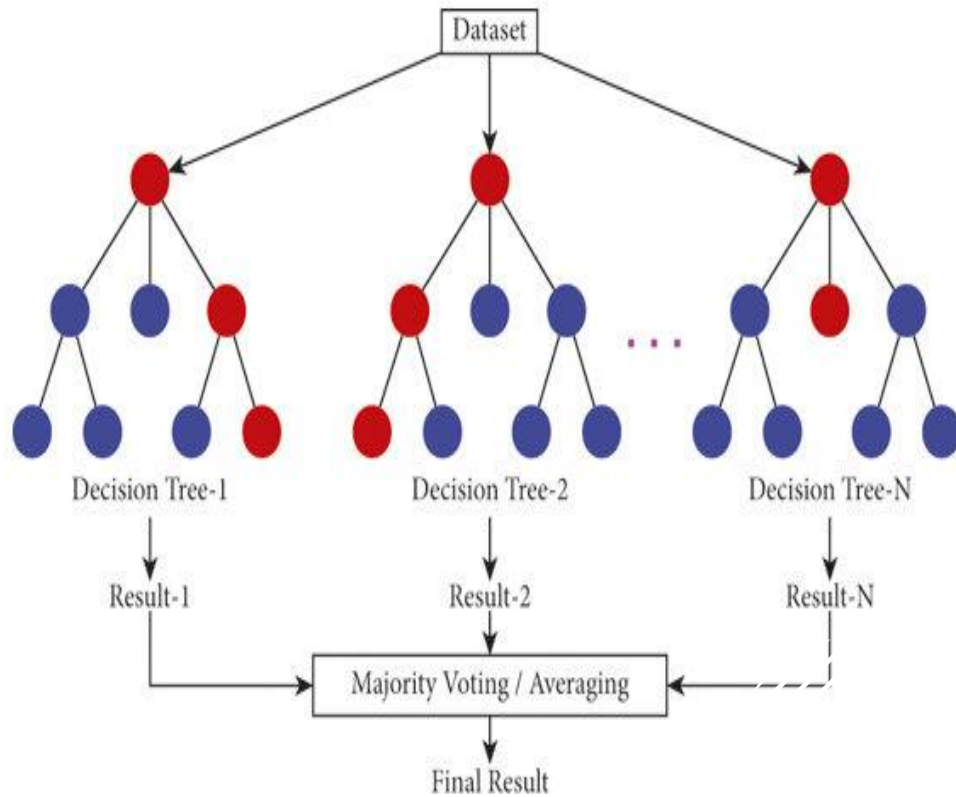
## Loan Repayment Prediction



Use loan\_data.csv to train the model. Predicts whether the bank should approve the loan of an applicant based on his profit using any two machine learning algorithms.

- I use two algorithms to predict the problem. They are:
    - Random Forest Algorithm
    - Decision Tree Algorithm
- 

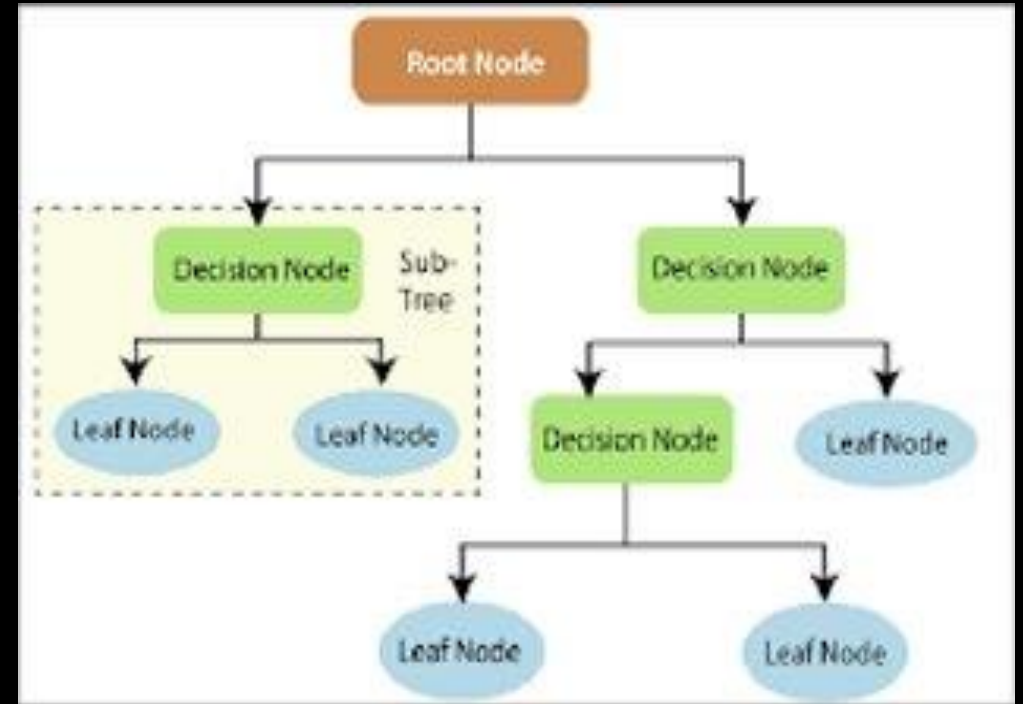
# Random Forest Algorithm:



- Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset.

# Decision Tree Algorithm:

- Decision Tree is the most powerful and popular tool for classification and prediction. Decision trees are able to generate understandable rules perform classification without requiring much computation. Decision trees are able to handle both continuous and categorical variables. Decision trees provide a clear indication of which fields are most important for prediction or classification.





# Approach used for the project:

At First, we need to import the libraries that we will use like NumPy, Pandas and Seaborn etc.

Then we need to read our csv file.

A dataset may contain duplicate values or missing values. So, we need to clean the data.

Visualize the data according to your research .

Find the Accuracy using Decision Tree Algorithm and Random Forest Algorithm.

At last, we need to find the best approach out of two.





## Decision Tree

<-----Classification Report----->

	precision	recall	f1-score	support
0	0.85	1.00	0.92	2431
1	0.00	0.00	0.00	443
accuracy			0.85	2874
macro avg	0.42	0.50	0.46	2874
weighted avg	0.72	0.85	0.78	2874

<-----Accuracy Scores----->

Train Accuracy score: 0.8374105011933174  
Test Accuracy score: 0.8458594293667363

## Random Forest

<-----Classification Report----->

	precision	recall	f1-score	support
0	0.85	1.00	0.92	2431
1	0.57	0.02	0.04	443
accuracy			0.85	2874
macro avg	0.71	0.51	0.48	2874
weighted avg	0.81	0.85	0.78	2874

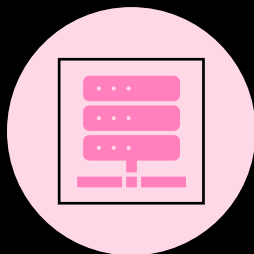
<-----Accuracy Scores----->

Train Accuracy score: 1.0  
Test Accuracy score: 0.8465553235908142

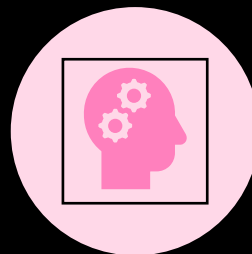




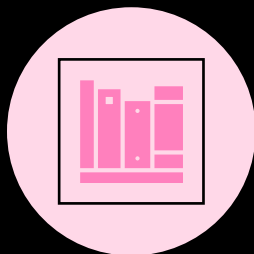
# Learning Outcomes:



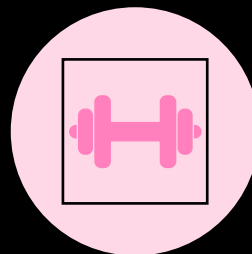
Have a good understanding of the fundamental issues and challenges of machine learning: data, model selection, model complexity, etc.



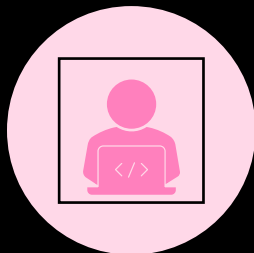
Appreciate the underlying mathematical relationships within and across Machine Learning algorithms and the paradigms of supervised and un-supervised learning.



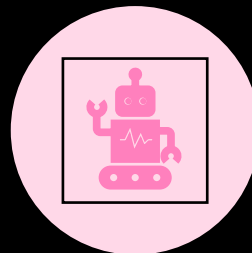
Have a good understanding of the libraries used in the project.



Have an understanding of the strengths and weaknesses of many popular machine learning approaches.



Be able to design and implement various machine learning algorithms in a range of real-world applications.



Ability to understand and apply scaling up machine learning techniques and associated computing techniques and technologies.





Thank  
You!

