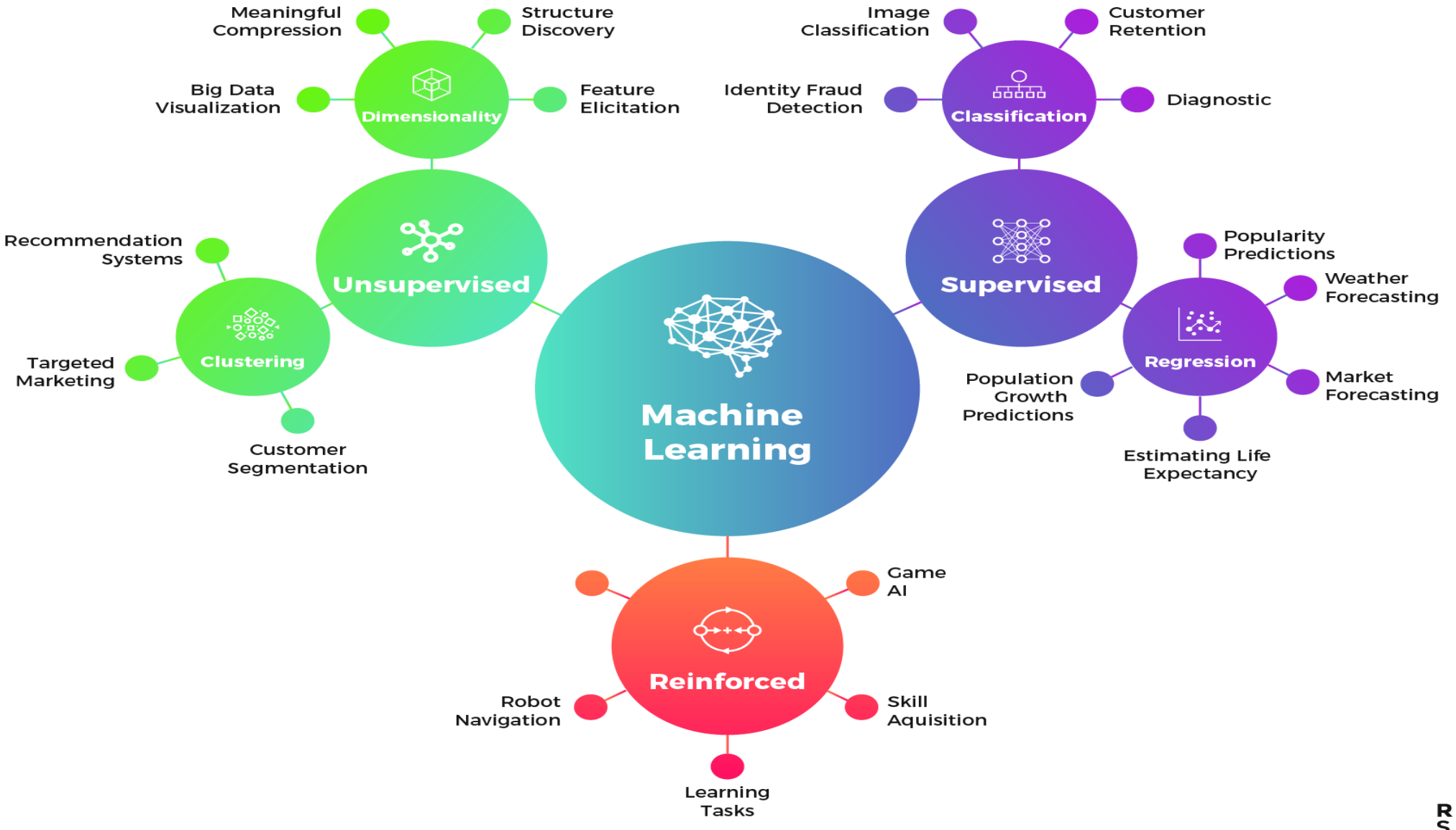
Artificial Intelligence

Zeeshan Abbas (Visiting Lecturer)
zeeshanabbas5@hotmail.com

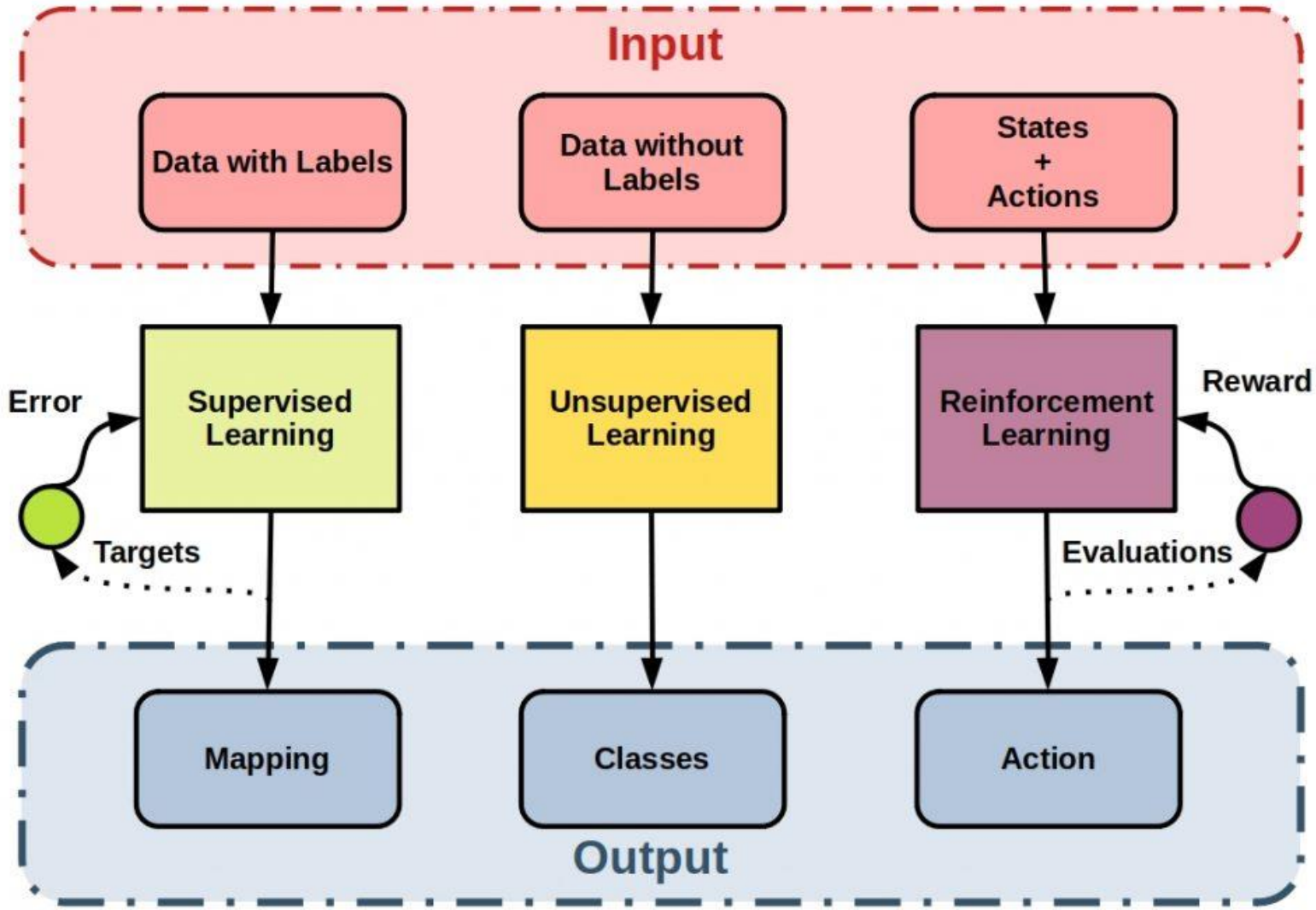
TR/AI at main · ZeeshanAbbas/TR (github.com)

Types of Machine Learning



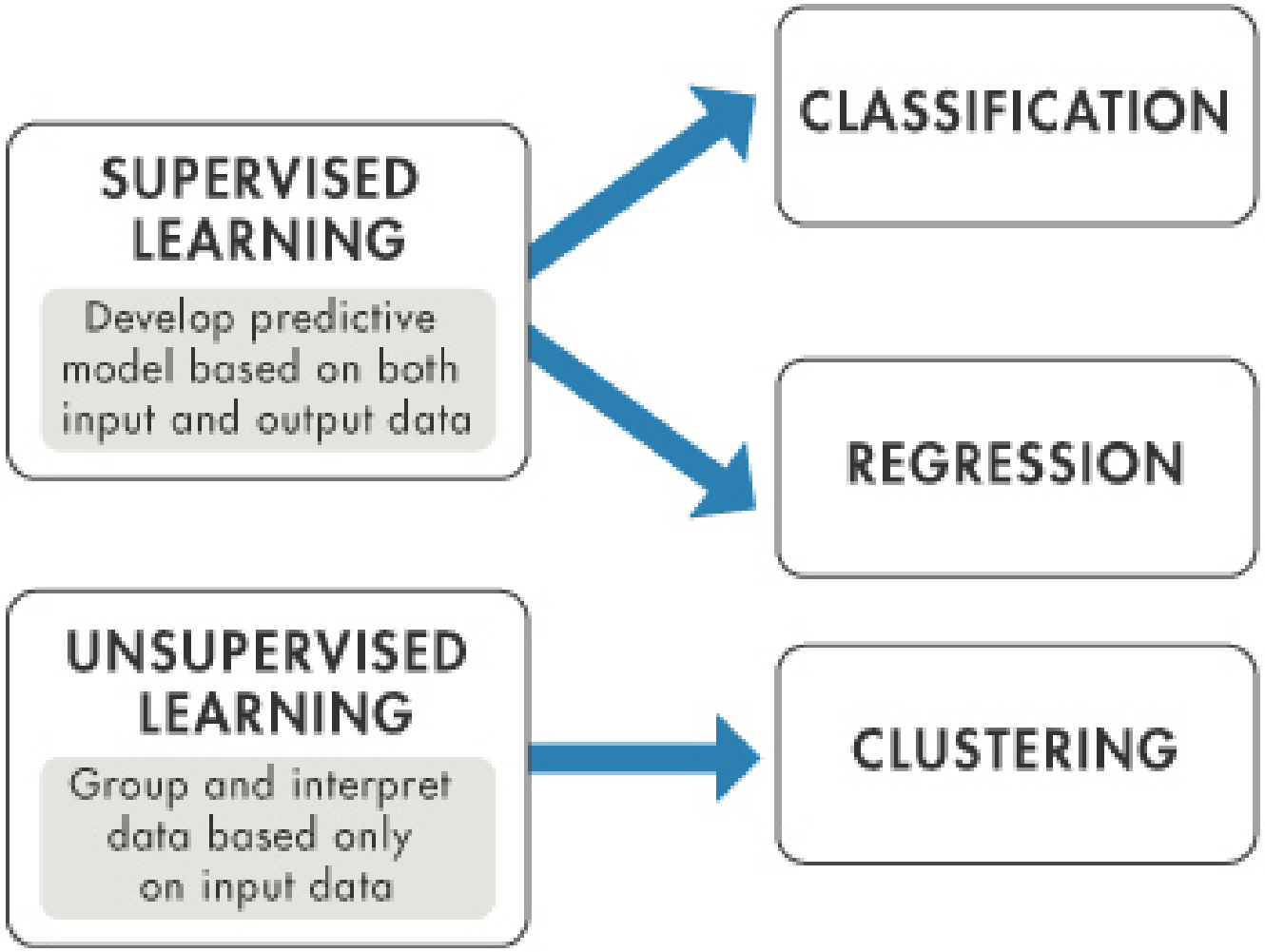
ROCKET SOURCE

Theory behind types of Machine Learning

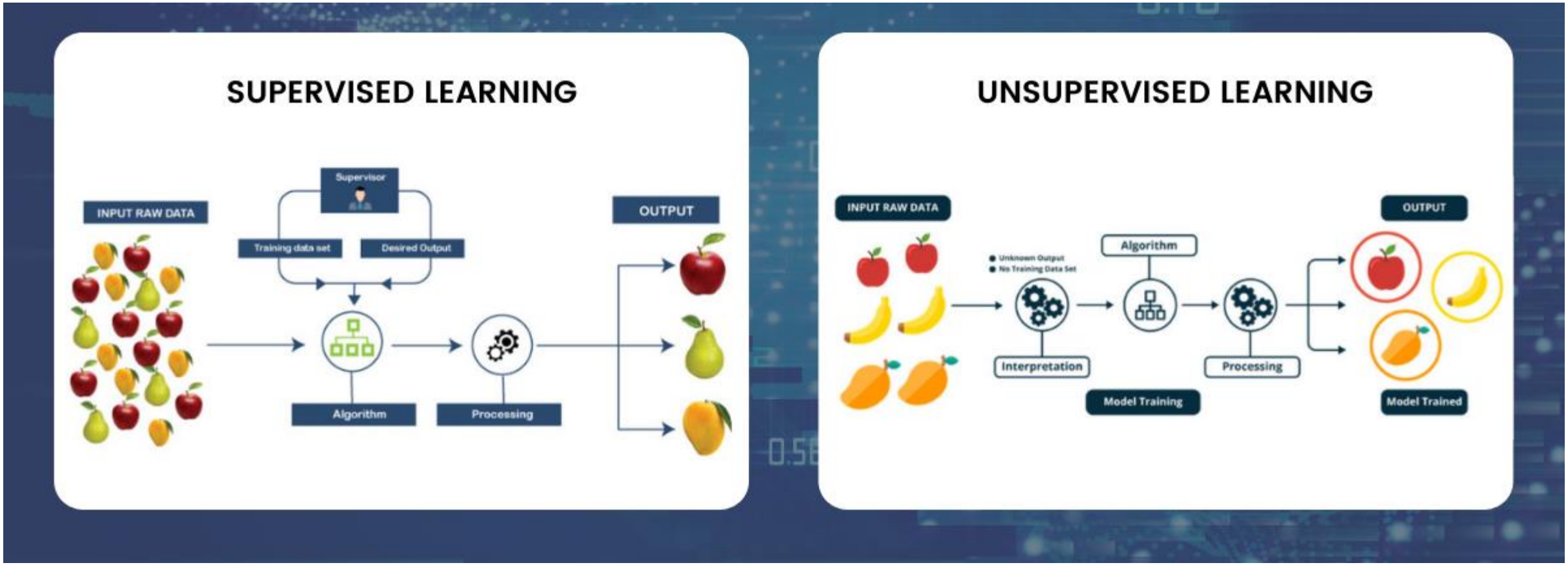


Supervised versus unsupervised learning

- The world of machine learning is broadly divided into supervised and unsupervised learning.
- There are other divisions too, but we'll discuss those later.



Supervised versus unsupervised learning

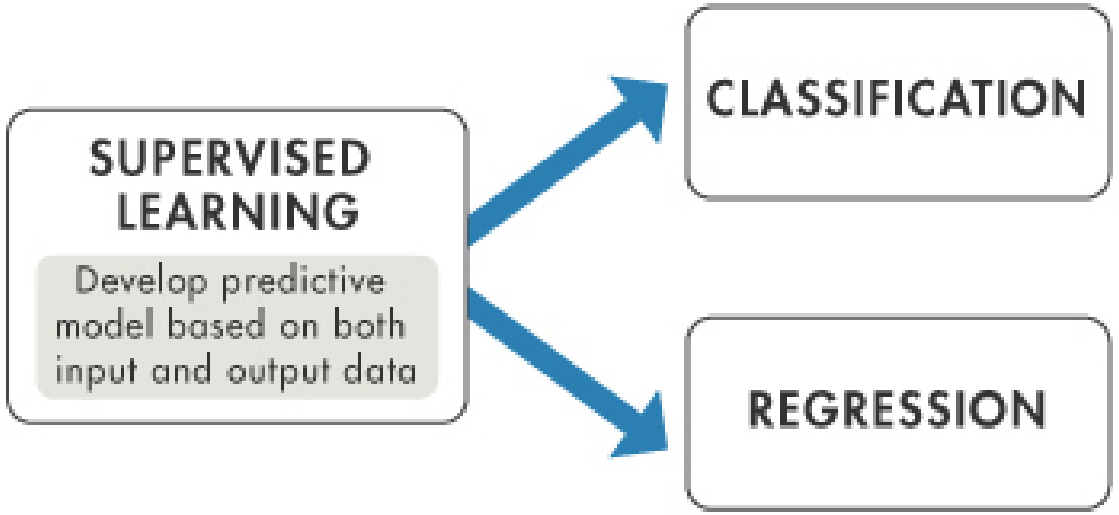


Supervised versus unsupervised learning

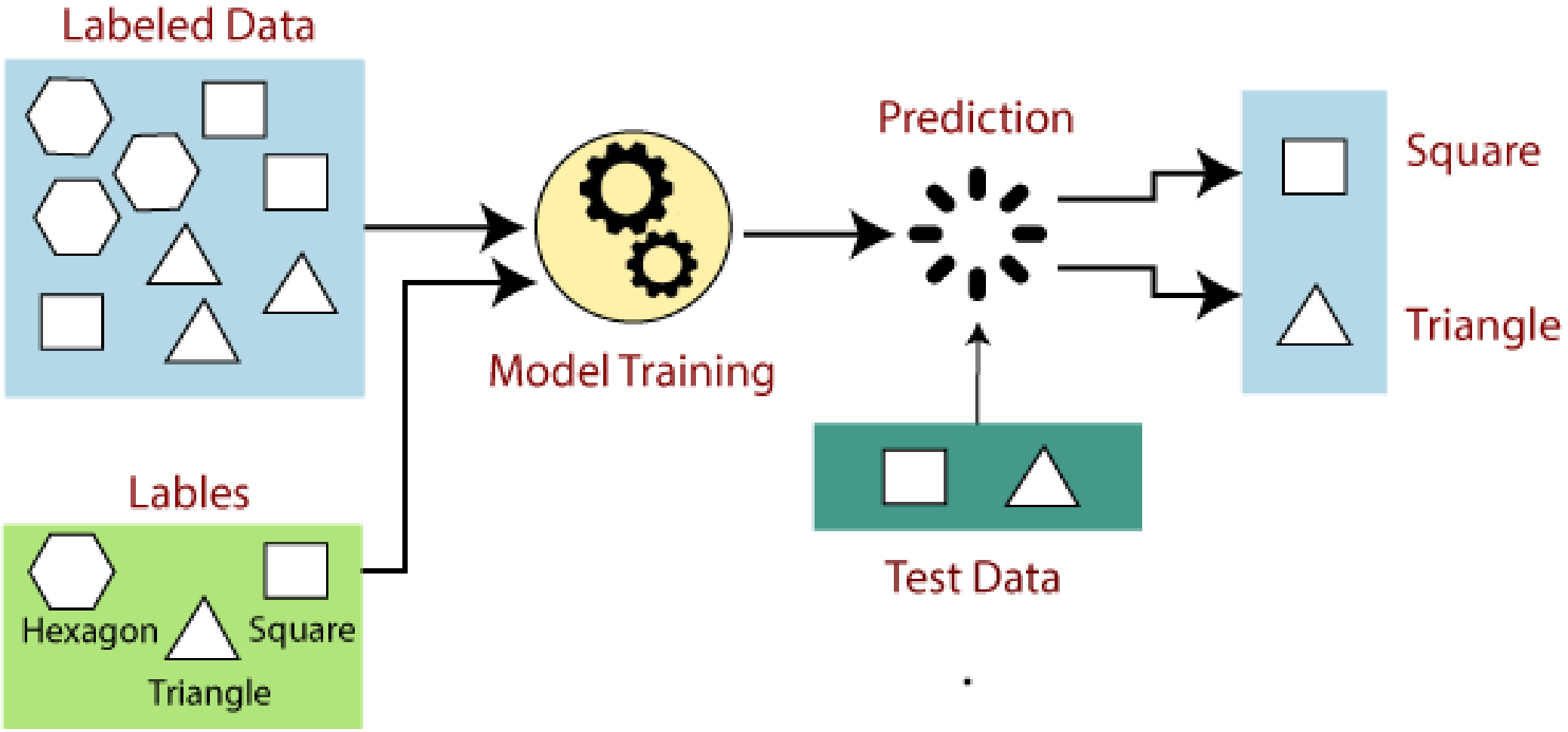
CLASSIFICATION	REGRESSION	CLUSTERING
Support Vector Machines	Linear Regression, GLM	K-Means, K-Medoids Fuzzy C-Means
Discriminant Analysis	SVR, GPR	Hierarchical
Naive Bayes	Ensemble Methods	Gaussian Mixture
Nearest Neighbor	Decision Trees	Hidden Markov Model
Neural Networks	Neural Networks	Neural Networks

Supervised learning

- **Supervised learning** refers to the process of building a machine learning model that is based on **labeled** training data.
- In Supervised Learning, the model learns by example.
- Along with our input variable, we also give our model the corresponding correct labels (correct output).
- During training, the model gets to look at which label corresponds to our data and hence can find patterns between our data and those labels.



Supervised learning



Supervised learning

Steps Involved in Supervised Learning:

- First Determine the type of training dataset
- Collect/Gather the labeled training data.
- Split the training dataset into a training dataset, test dataset, and validation dataset.
- Determine the input features of the training dataset, which should have enough knowledge so that the model can accurately predict the output.
- Determine the suitable algorithm for the model, such as support vector machine, decision tree, etc.
- Execute the algorithm on the training dataset. Sometimes we need validation sets as the control parameters, which are the subset of training datasets.
- Evaluate the accuracy of the model by providing the test set. If the model predicts the correct output, which means our model is accurate.

Supervised learning

Advantages of Supervised learning:

- With the help of supervised learning, the model can predict the output on the basis of prior experiences.
- In supervised learning, we can have an exact idea about the classes of objects.
- A supervised learning model helps us to solve various real-world problems such as fraud detection, spam filtering, etc.

Disadvantages of Supervised learning:

- Supervised learning cannot predict the correct output if the test data is different from the training dataset.
- Training required lots of computation time.
- In supervised learning, we need enough knowledge about the classes of objects.
- Supervised learning models are not suitable for handling complex tasks.

Supervised learning

I- Classification:

- The classification is one such technique of machine learning where we classify data into a given number of classes.
- In machine learning, classification solves the problem of identifying the category to which a new data point belongs.
- During classification, we arrange data into a fixed number of categories so that it can be used most effectively and efficiently**.**
- With the help of these pre-categorized training datasets, classification in machine learning programs leverage a wide range of algorithms to classify future datasets into respective and relevant categories.

Supervised learning

I- Classification algorithms:

You can apply many different classification methods based on the dataset you are working with. It is so because the study of classification in statistics is extensive.

The top five machine learning algorithms are listed below:

1. Logistic Regression
2. Naive Byes
3. K-Nearest Neighbors
4. Decision Tree
5. Support Vector Machine

Supervised learning

I-(i) Logistic Regression

- Logistic Regression was used in the biological sciences in the early twentieth century. It was then used in many social science applications.
- Logistic Regression is used when the dependent variable(target) is categorical.
- Logistic Regression is much similar to Linear Regression except that how they are used. Linear Regression is used for solving Regression problems, whereas Logistic regression is used for solving the classification problems.

Supervised learning

I-(i) Logistic Regression (Types)

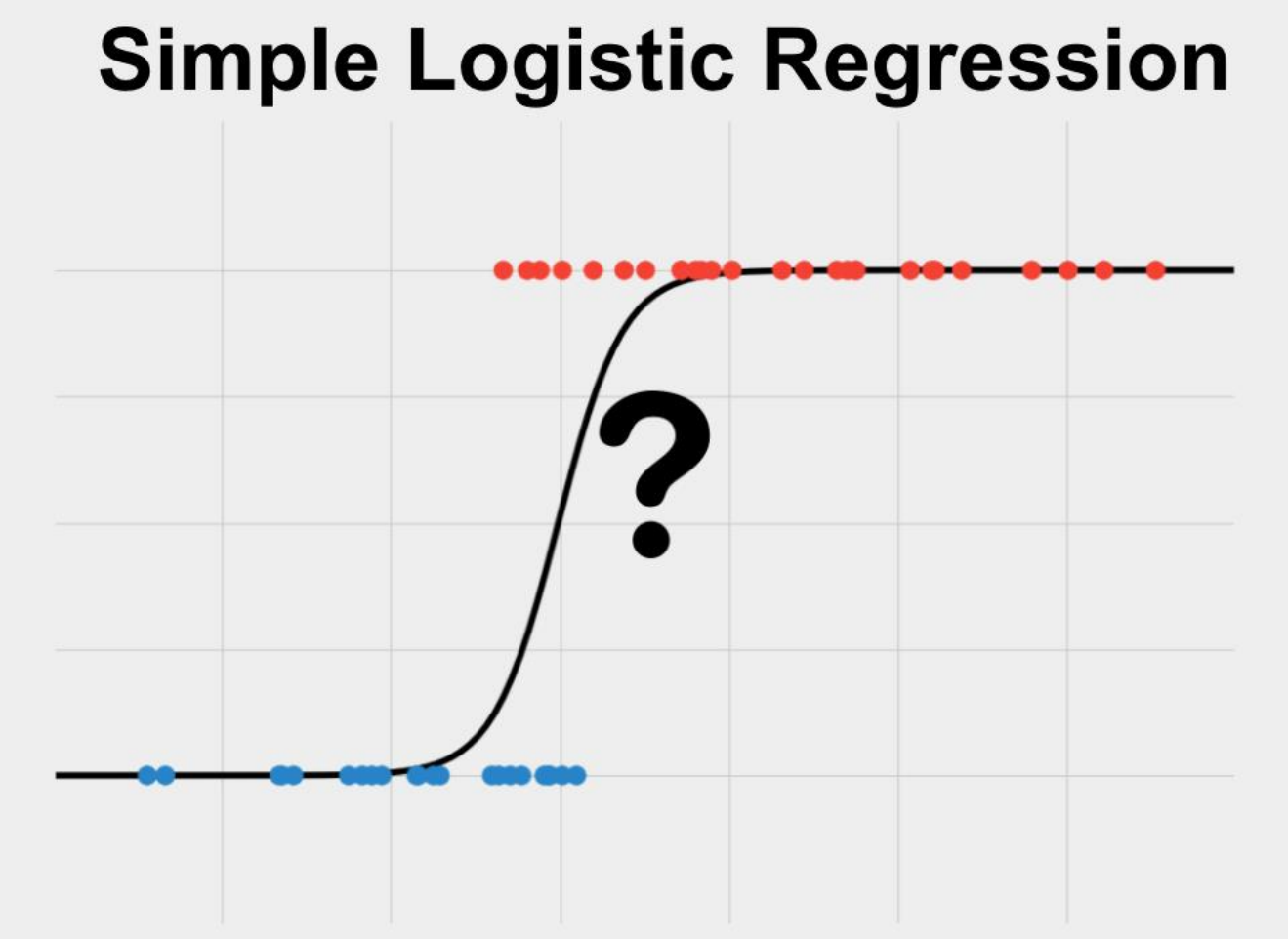
On the basis of the categories, Logistic Regression can be classified into three types:

- **Binomial**: In binomial Logistic regression, there can be only two possible types of the dependent variables, such **as 0 or 1, Pass or Fail**, etc.
- **Multinomial**: In multinomial Logistic regression, there can be 3 or more possible unordered types of the dependent variable, such as **"cat", "dogs"**, or **"sheep"**
- **Ordinal**: In ordinal Logistic regression, there can be 3 or more possible ordered types of dependent variables, such as **"low", "Medium",** or **"High".**

Supervised learning

I-(i) Logistic Regression (Simple/ Binomial)

Simple Logistic Regression is a statistical test used to predict a single binary variable using one other variable. It also is used to determine the numerical relationship between two such variables.



Supervised learning

I-(i) Logistic Regression (Simple/ Binomial)

Assumptions for Simple Logistic Regression:

Every statistical method has assumptions. Assumptions mean that your data must satisfy certain properties in order for statistical method results to be accurate.

The assumptions for Simple Logistic Regression include:

- Linearity
- No Outliers
- Independence

Supervised learning

I-(i) Logistic Regression (Simple/ Binomial)

When to use Simple Logistic Regression?

You should use Simple Logistic Regression in the following scenario:

- You want to use one variable in a **prediction** of another, or you want to quantify the numerical relationship between two variables
- The variable you want to predict (your dependent variable) is **binary**
- You have **one independent variable**, or one variable that you are using as a predictor

Supervised learning

I-(ii) Naïve Bayes

- Naïve Bayes is a technique used to build classifiers using Bayes theorem.
- Bayes theorem describes the probability of an event occurring based on different conditions that are related to this event.
- Naïve Bayes classifier assume that the value of any given feature is independent of the value of any other feature.
- This is called the independence assumption, which is the naïve part of a Naïve Bayes classifier.

$$P(A \mid B) = \frac{P(A)\, P(B \mid A)}{P(B)}$$

Supervised learning

I-(ii) Naïve Bayes

Naïve Bayes Classifier

$$P(A|B) = \frac{P(B|A)\ P(A)}{P(B)}$$

Thomas Bayes
1702 - 1761

Where :

- P(A | B) = how often **“A”** happens given that **“B”** happens
- P(A) = how likely A will happen
- P(B) = how likely B will happen
- P(B | A) = how often B happens given that A happens

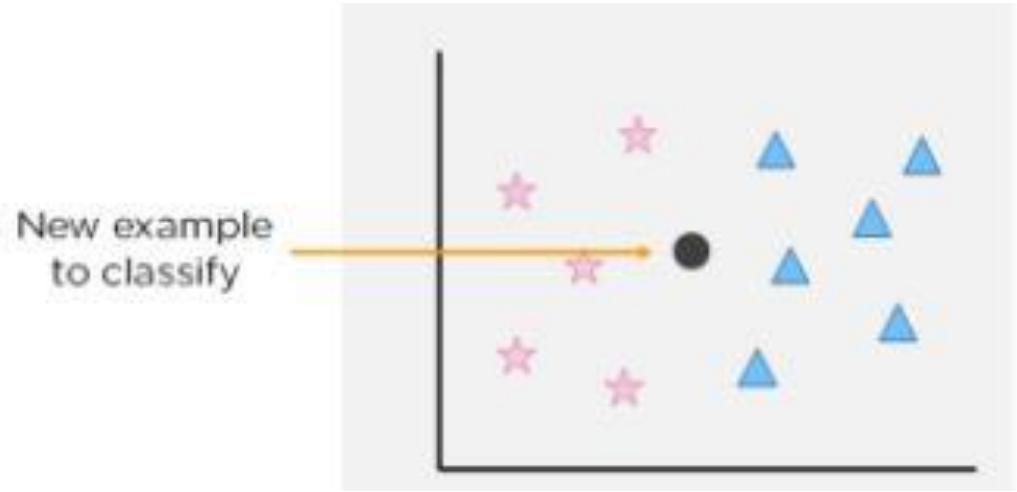
Supervised learning

I-(ii) K-Nearest Neighbor

- K-Nearest Neighbor is a classification and prediction algorithm that is used to divide data into classes based on the distance between the data points.
- K-Nearest Neighbor assumes that data points which are close to one another must be similar and hence, the data point to be classified will be grouped with the closest cluster.

Supervised learning

I-(ii) K-Nearest Neighbor



Supervised learning

I-(ii) K-Nearest Neighbor