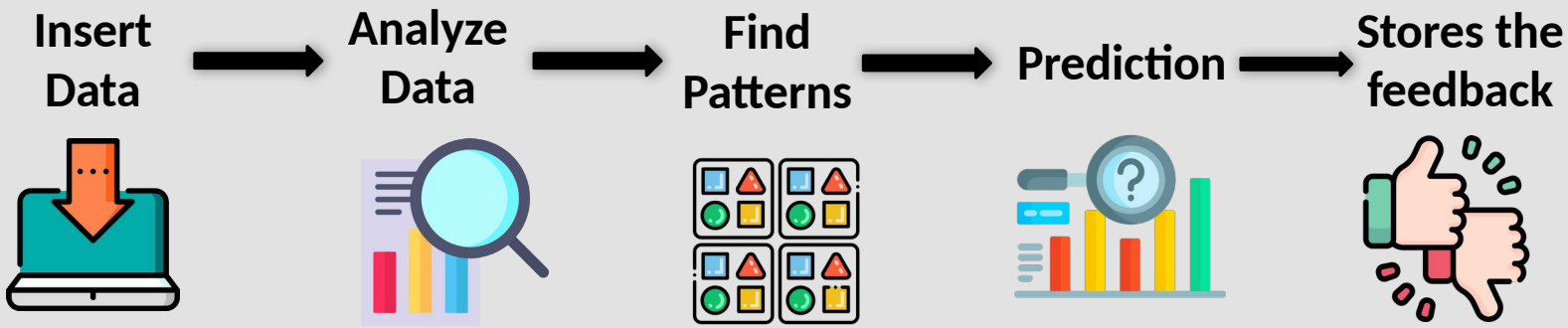
DAY 7

Instructor: Balu Mohandas Menon

Christian B. Wiberg
Philip Jess Teining

HOW DOES ML WORK?



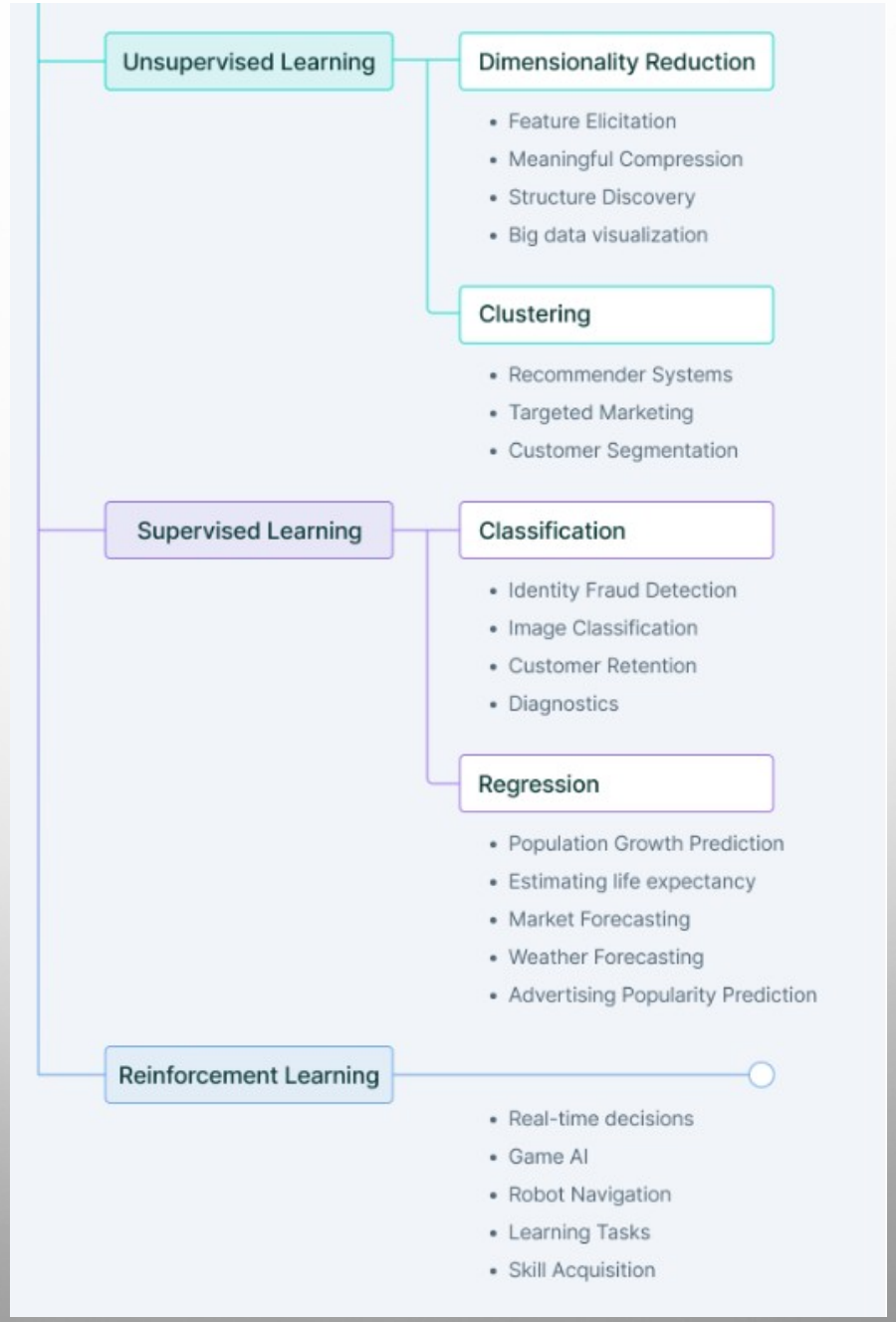
https://data-flair.training/blogs/machine-learning-tutorial/

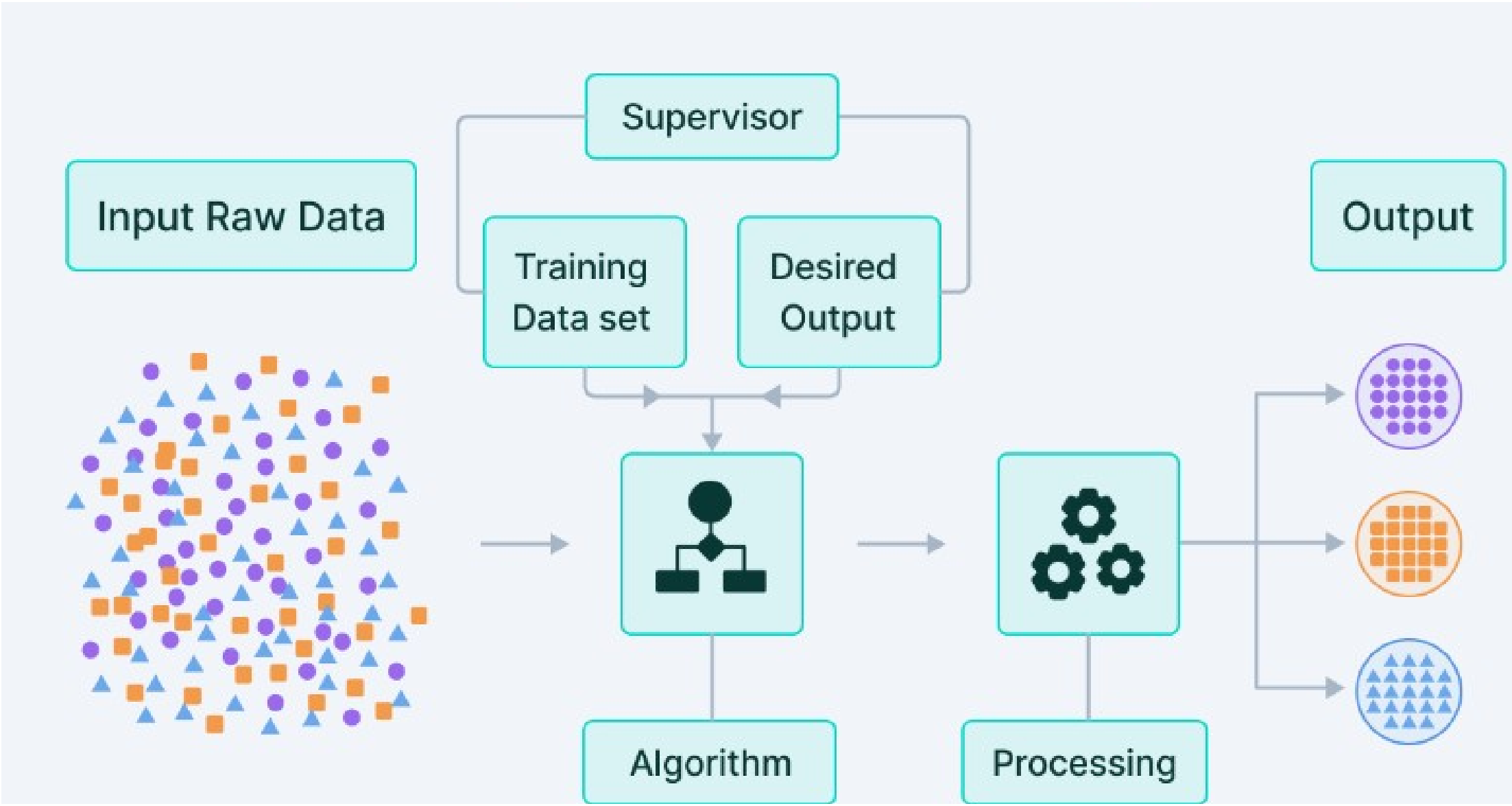
TYPES OF ML

SUPERVISED LEARNING

UNSUPERVISED LEARNING

REINFORCEMENT LEARNING


Unsupervised Learning
Dimensionality Reduction
Feature Elicitation
Meaningful Compression
Structure Discovery
Big data visualization
Clustering
Recommender Systems
Targeted Marketing
Customer Segmentation
Supervised Learning
Classification
Identity Fraud Detection
Image Classification
Customer Retention
Diagnostics
Regression
Population Growth Prediction
Estimating life expectancy
Market Forecasting
Weather Forecasting
Advertising Popularity Prediction
Reinforcement Learning
Real-time decisions
Game AI
Robot Navigation
Learning Tasks
Skill Acquisition


Supervisor
Input Raw Data
Training Data set
Desired Output
Output
Algorithm
Processing

SUPERVISED MACHINE LEARNING METHODS

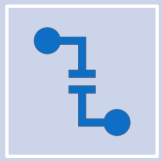
Classification : Classification refers to taking an input value and mapping it to a discrete value.

Output typically consists of categories, ex: whether it is going to rain today or not.

Regression : predicted output values are real numbers.

Predicting the price of a house or the trend in the stock price at a given time, etc.

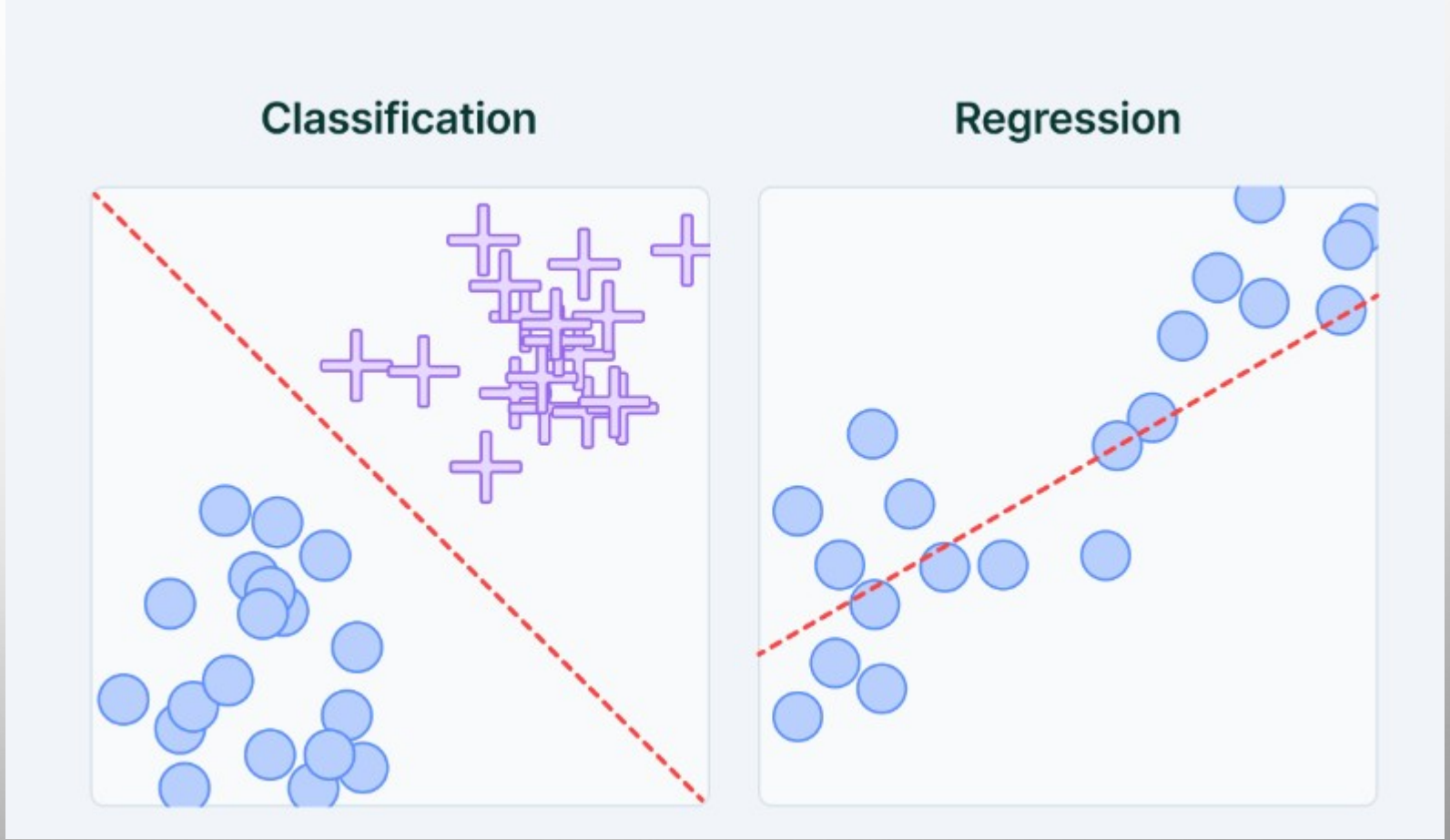
UNSUPERVISED MACHINE LEARNING METHODS



Clustering : Clustering is the type of Unsupervised Learning where we find hidden patterns in the data based on their similarities or difference.

Association : Association is the kind of Unsupervised Learning where we can find the relationship of one data item to another data item.

e.g., understanding consumers' habits regarding our products can help us develop better cross-selling strategies.


Classification
Regression

SUPERVISED MACHINE LEARNING APPLICATIONS

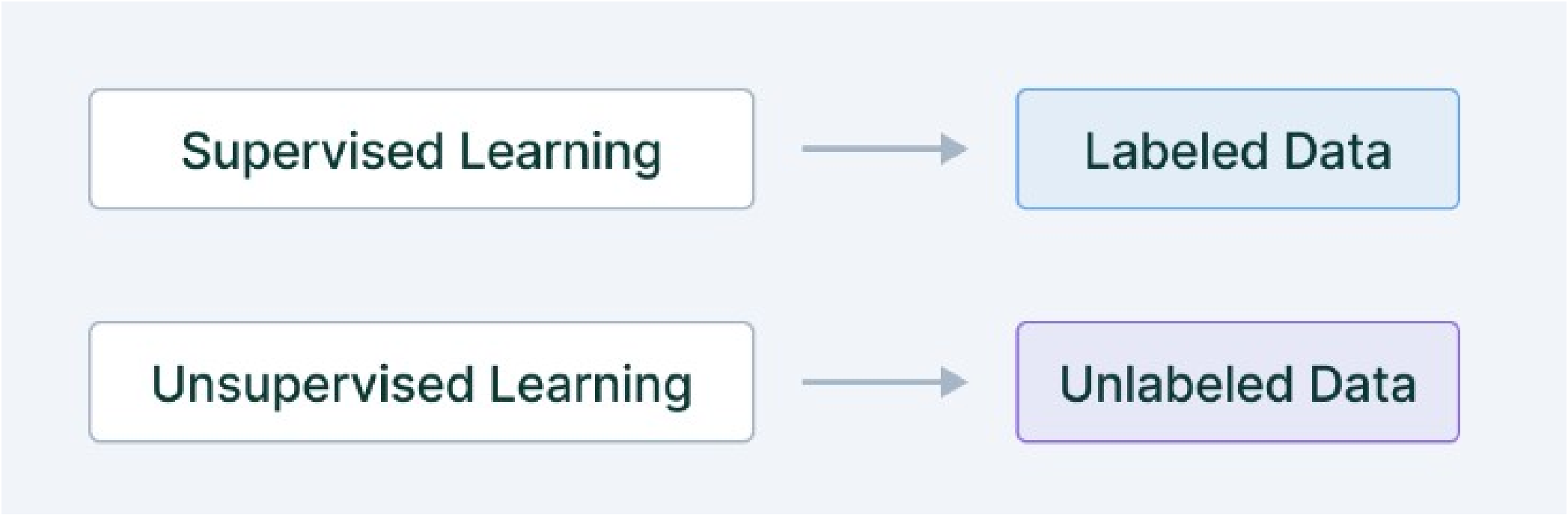
Predictive analytics (house prices, stock exchange prices, etc.)

Text recognition

Spam detection

Customer sentiment analysis

Object detection (e.g. face detection)


Supervised Learning
Labeled Data
Unsupervised Learning
Unlabeled Data

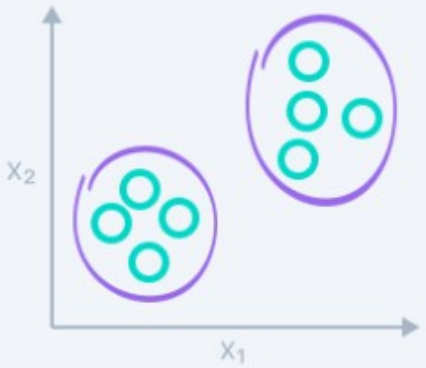
DATA

Supervised Learning learns from the training dataset by iteratively making predictions on the data and adjusting for the correct answer.

Supervised techniques deal with labeled data where the output data patterns are known to the system.

Unsupervised Learning models work on their own to discover the inherent structure of unlabeled data.

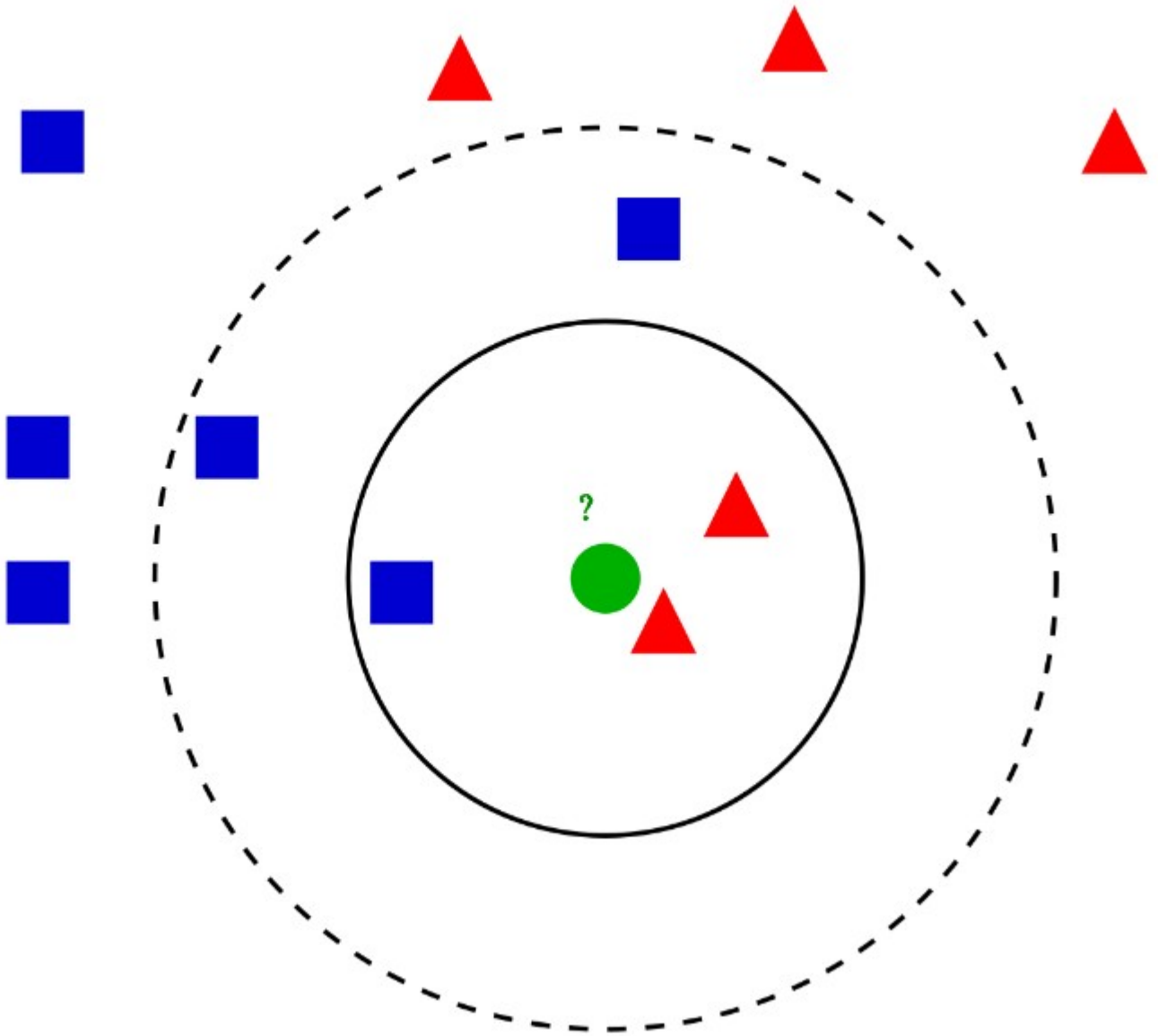
The unsupervised learning algorithm works with unlabeled data, in which the output is based solely on the collection of perceptions.

Supervised learning	Unsupervised learning
Input data is labeled	Input data is unlabeled
Has a feedback mechanism	Has no feedback mechanism
Data is classified based on the training dataset	Assigns properties of given data to classify it
Divided into Regression & Classification	Divided into Clustering & Association
Used for prediction	Used for analysis
Algorithms include: decision trees, logistic regressions, support vector machine	Algorithms include: k-means clustering, hierarchical clustering, apriori algorithm
A known number of classes	A unknown number of classes
	

CLASSIFICATION TECHNIQUES

- As stated earlier, classification is when the feature to be predicted contains categories of values. Each of these categories is considered as a class into which the predicted value falls. Classification algorithms include:
- Naive Bayes
- Logistic regression
- K-nearest neighbors
- (Kernel) SVM
- Decision tree
- Ensemble learning

K-Nearest Neighbors

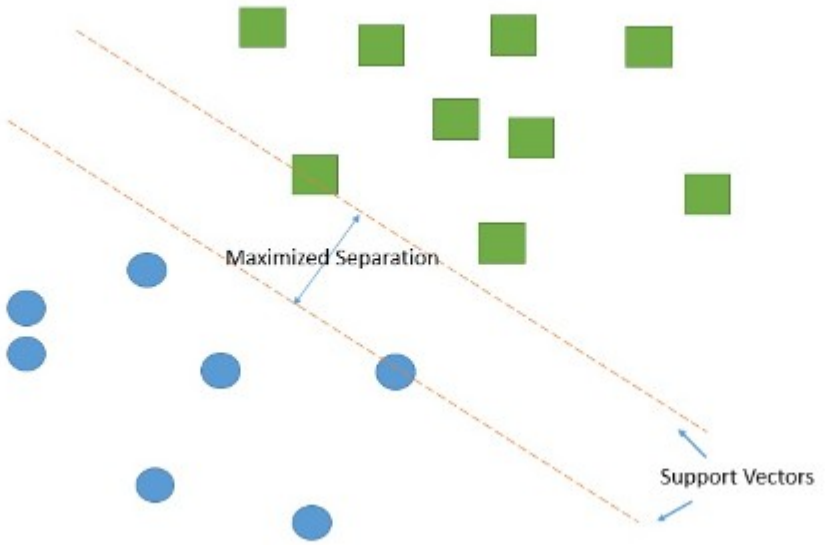


KNN & SVM

K-Nearest Neighbors operates by checking the distance. The k value in the k-NN algorithm defines how many neighbors will be checked to determine the classification of a specific query point

Support Vector Machines work by drawing a line between the different clusters of data points to group them into classes. Points on one side of the line will be one class and points on the other side belong to another class

Support Vector Machines



UNDERSTANDING TRAIN TEST SPLIT (SCIKIT-LEARN + PYTHON)