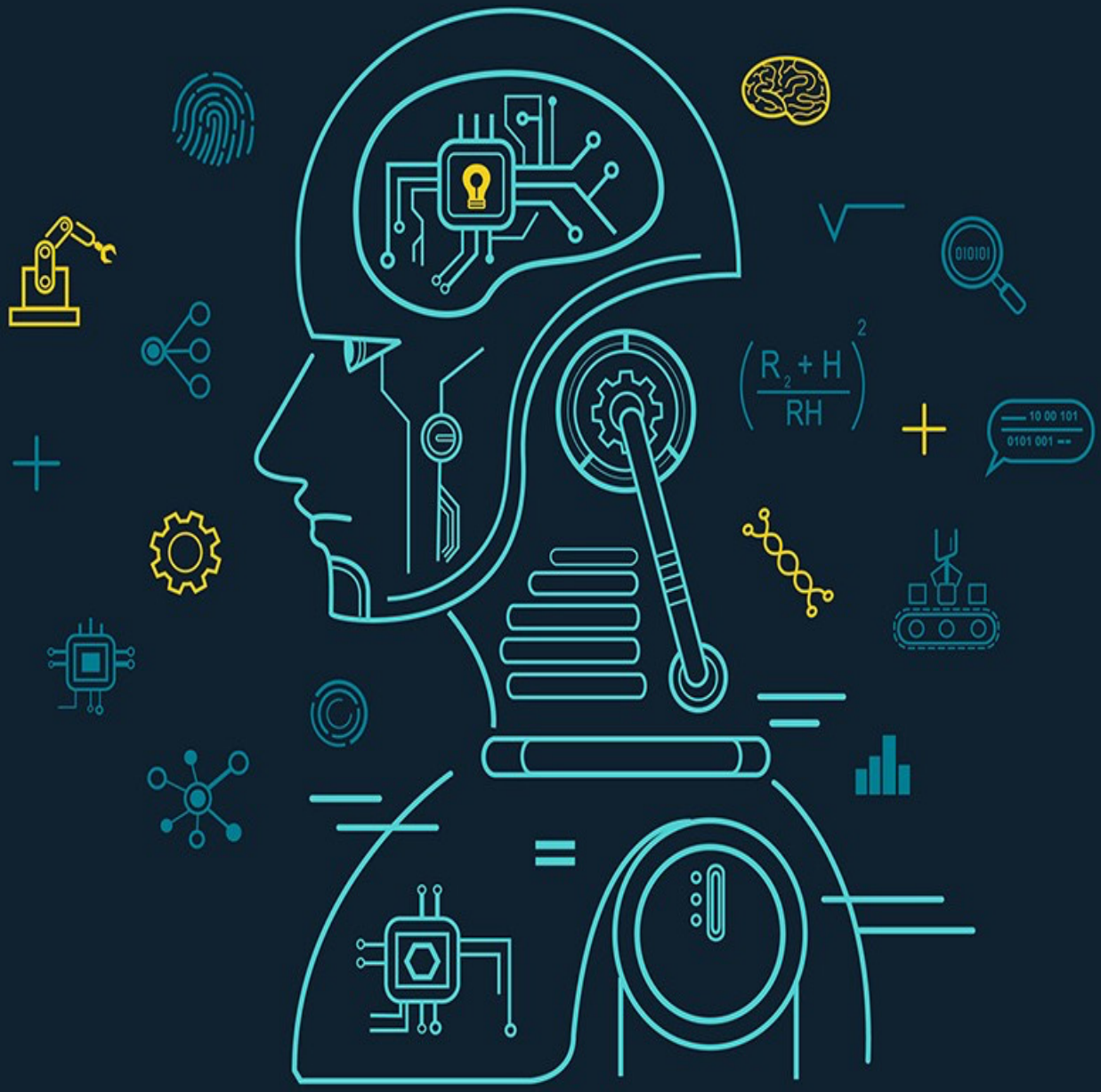
GUIDE TO



MACHINE LEARNING

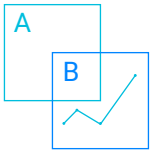
Source: https://mck.co/36jWuNk

Vinay Baburao

Type of Analytics

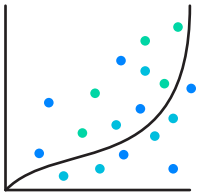
Machine Learning Focused

Prescriptive



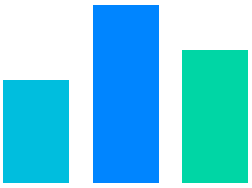
- Recommendation to achieve goals
- Used by leading data & Internet companies

Predictive



- Anticipate what will happen
- Data driven, key source of insight

Descriptive



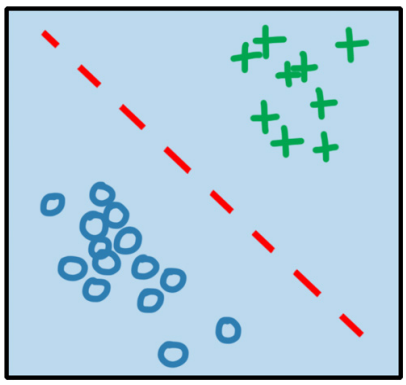
- Describe what happened
- Used across industries

Complexity

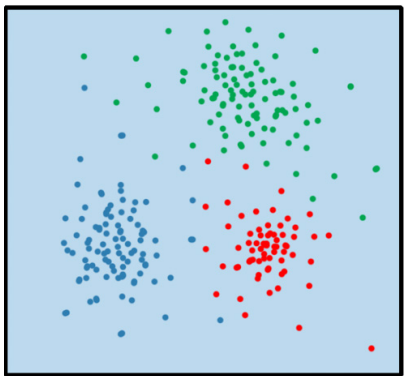
Source: https://mck.co/36jWuNk

Vinay Baburao

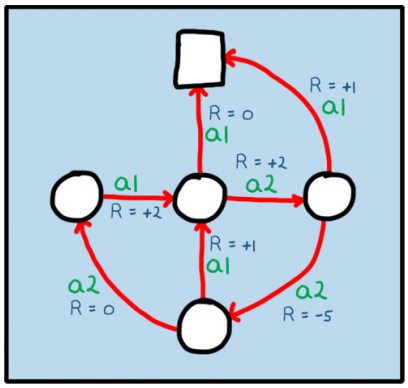
Major Types of ML



Supervised Learning



Unsupervised Learning

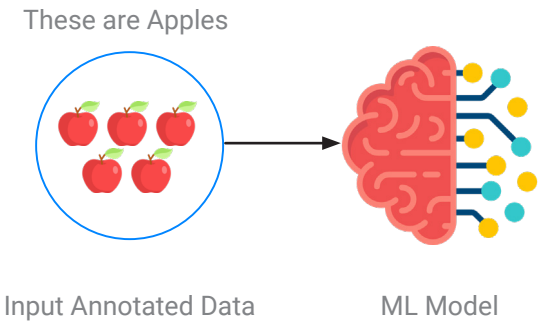


Reinforcement Learning

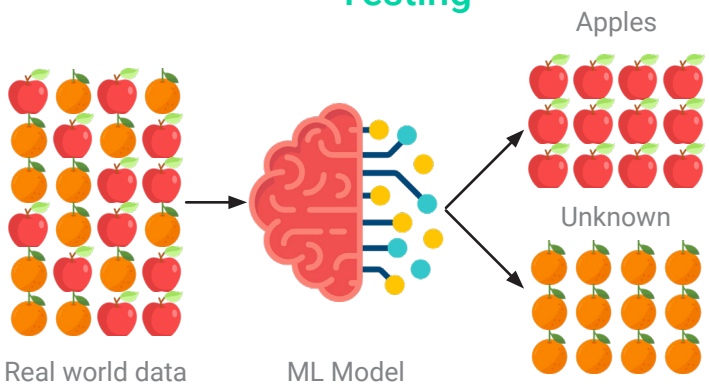
Source: Mathworks

Supervised

Training



Testing



What it is

An algorithm uses training data and feedback from humans to learn the relationship of given inputs to a given output

(Example: how the inputs “time of year” and “interest rates” predict housing prices)

When to use

You know how to classify the input data and the type of behavior you want to predict, but you need the algorithm to calculate it for you on new data

How it works

1.A human labels the input data and defines the output variable

(Example: housing prices)

2.The algorithm is trained on the data to find the connection between the input variables and the output

3.Once training is complete–typically when the algorithm is sufficiently accurate–the algorithm is applied to new data

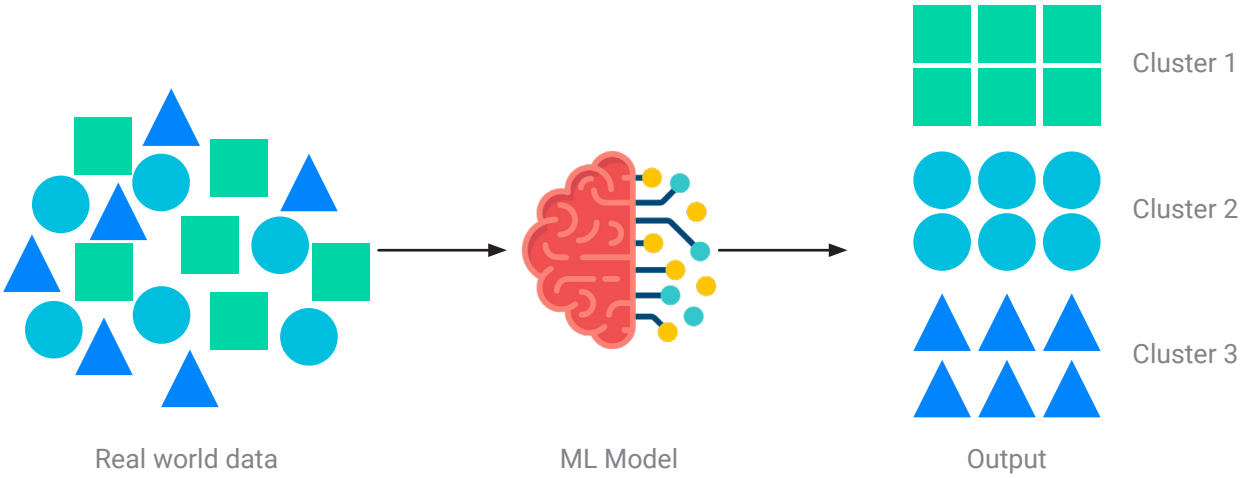
Algorithms

1. Linear Regression
2. Logistic regression
3. Linear discriminant analysis
4. Decision tree
5. Naive Bayes
6. Support vector machine
7. Random forest
8. AdaBoost
9. Gradient-boosting trees
10. Simple neural network

Source: https://mck.co/36jWuNk

Vinay Baburao

Unsupervised



What it is

An algorithm explores input data without being given an explicit output variable
(Example: explores customer demographic data to identify patterns)

When to use

You do not know how to classify the data, and you want the algorithm to find patterns and classify the data for you

How it works

1.The Algorithm receives unlabeled data
(Example: a set of data describing customer journeys on a website)
2.It Infers a structure from the data
3.The Algorithm identifies groups of data that exhibit similar behavior
(Example: forms clusters of customers that exhibit similar buying behaviors)

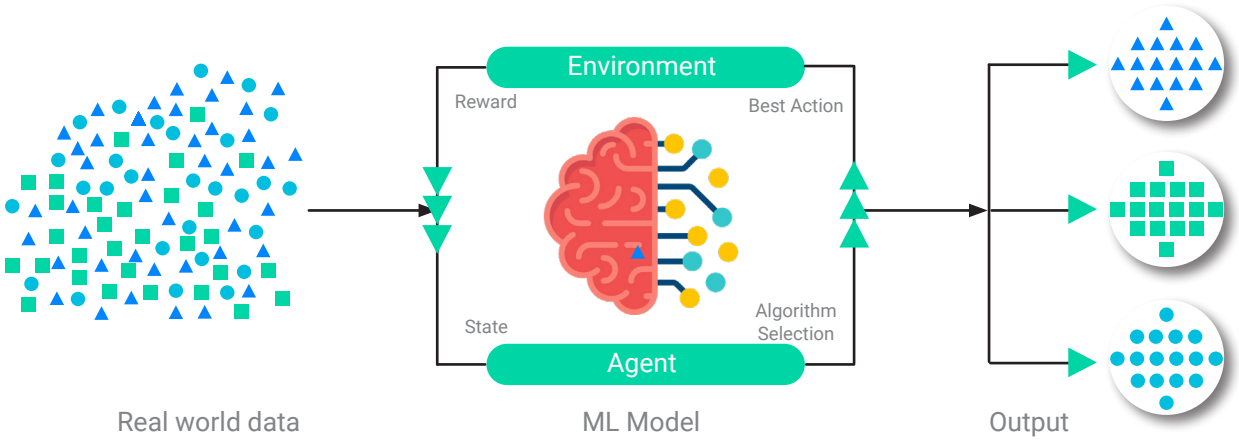
Algorithms

1. Auto encoder
2. K-means clustering
3. Gaussian mixture model
4. Hierarchical clustering
5. Recommender system
6. Manifold learning

Source: https://mck.co/36jWuNk

Vinay Baburao

Reinforcement



WHAT IT IS

An algorithm learns to perform a task simply by trying to maximize rewards it receives for its actions

(Example: maximizes points it receives for increasing returns of an investment portfolio)

WHEN TO USE

You don't have a lot of training data; you cannot clearly define the ideal end state; or the only way to learn about the environment is to interact with it

HOW IT WORKS

1.The algorithm takes an action on the environment

(Example: makes a trade in a financial portfolio)

2.It receives a reward if the action brings the machine a step closer to maximizing the total rewards available

(Example: the highest total return on the portfolio)

3.The algorithm optimizes for the best series of actions by correcting itself over time

Source: https://mck.co/36jWuNk

Use Cases

Supervised Learning

Linear regression

- Understand product-sales drivers such as competition prices, distribution, advertisement,etc
- Optimize price points and estimate product-price elasticities

Logistic regression

- Classify customers based on how likely they are to repay a loan
- Predict if a skin lesion is benign or malignant based on its characteristics (size, shape, color, etc)

Linear discriminant analysis

- Predict client churn
- Predict a sales lead's likelihood of closing

Decision Tree

- Understand product attributes that make a product most likely to be purchased
- Provide a decision framework for hiring new employees

Naive Bayes

- Analyze sentiment to assess product perception in the market
- Create classifiers to filter spam emails

Support Vector Machine

- Predict how many patients a hospital will need to serve in a time period
- Predict how likely someone is to click on an Online ad

Random forest

- Predict call volume in call centers for staffing decisions
- Predict power usage in an electrical-distribution grid

AdaBoost

- Detect fraudulent activity in credit-card transactions. Achieves lower accuracy than deep learning
- Simple, low-cost way to classify images. Achieves lower accuracy than deep learning

Gradient-boosting trees

- Forecast product demand and inventory levels
- Predict the price of cars based on their characteristics

Simple neural network

- Predict the probability that a patient joins a healthcare program
- Predict whether registered users will be willing or not to pay a particular price for a product

Source: https://mck.co/36jWuNk

Vinay Baburao

Use Cases

Unsupervised Learning

Auto Encoder

- Reduce noise in a medical image (eg, MRI) to analyze it more accurately
- Detect fake reviews and opinions on social media

K-means clustering

- Segment customers into groups by distinct characteristics (eg, age group)–for instance, to better assign marketing campaigns or prevent churn

Gaussian mixture model

- Segment employees based on likelihood of attrition
- Segment customers to better assign marketing campaigns using less-distinct customer characteristics

Hierarchical clustering

- Cluster loyalty-card customers into progressively more micro segmented groups
- Inform product usage/development by grouping customers mentioning keywords in social-media data

Recommender system

- Recommend what movies consumers should view based on preferences of other customers with similar attributes
- Recommend news articles a reader might want to read based on the article she or he is reading

Manifold learning

- Detect patterns in spread of a pandemic
- Group similar customers together and recommend next best product for them to buy
- Predict brain-tumor progression

Source: https://mck.co/36jWuNk

Use Cases

Reinforcement Learning

- Optimize the trading strategy for an options-trading portfolio
- Balance the load of electricity grids in varying demand cycles
- Stock and pick inventory using robots
- Optimize the driving behavior of self-driving cars
- Optimize pricing in real time for an Online auction of a product with limited supply

Source: https://mck.co/36jWuNk