

Machine Learning

By

Mrs. Anjali Purohit

“Computers are able to **see, hear** and **learn**.



“Welcome to the future.”

“A baby learns to crawl, walk and then run. We are at the **crawling stage** when it comes to applying machine learning.”

*“Machine learning will **automate jobs** that most people thought could only be done by people.”*

~Dave Waters



Lets hear from Mr.Sundar Pichai,CEO/Google
about ML

!!!!!!

Human Jobs replaced by Machine Learning

1. **Translator – replaced by Google Translate**
2. **Driver - Driverless car**
3. **Cashiers, Bank tellers and clerks – Online banking/ATM**
4. **Stewards in Hotel – Replaced by Robots**
5. **Food delivery – Replaced by swiggy/Zomato/Uber eats**
6. **Factory daily wage labours (Assembly-line and factory workers) – Automation**
7. **Data entry operators**
8. **Travel Guides/Local guide in Tourism spot – Google Maps**
9. **Travel agents**
10. **Home Security guard – Home CCTV camera/Home surveillance system**
11. **Provision Stores – E-Provision store (BigBasket.com)**
12. **Journalists and reporters – replaced by computer-assisted writing**
13. **Pilots - The United States Military is already using Auto-pilot**
14. **Postal workers**
15. **Film projectionists**

Robotic Kitchen – Moley Robotics



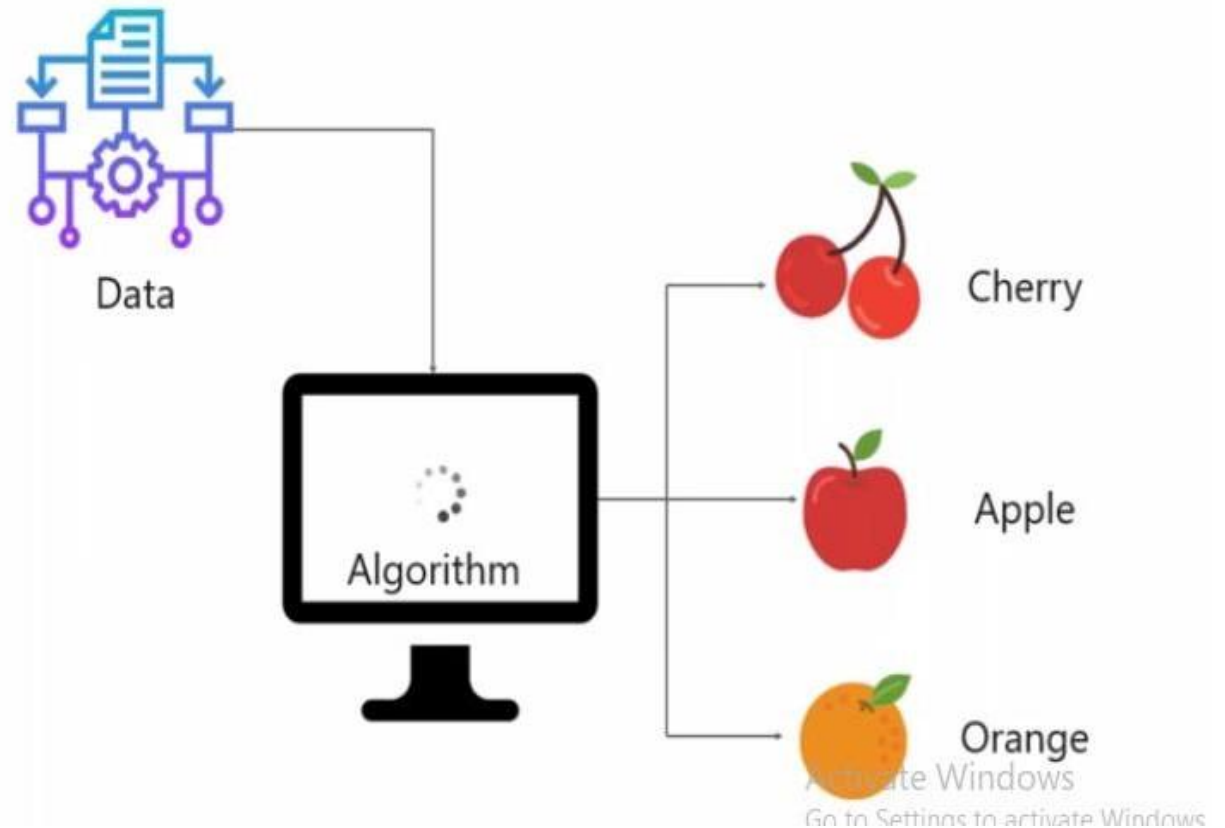
AGEN DA

1. Introduction to ML
2. Applications
3. Terminology in ML
4. ML Process
5. Top Python Machine Learning Libraries
6. ML types & Algorithms
7. List of IDEs
8. Installation of Anaconda Software



What Is Machine Learning?

Machine learning is a subset of artificial intelligence (AI) which provides machines the ability to learn automatically & improve from experience without being explicitly programmed.



1. What is machine learning?

- Machine learning is the study of computer algorithms that improve automatically through experience. It is seen as a subset of artificial intelligence.
- As intelligence requires knowledge, it is necessary for the
- computers to acquire knowledge.

Machine Learning - Mitchell (1997) definition

“A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P , if its performance at tasks in T , as measured by P , improves with experience E .”

In simple language machine learning is a field in which human made algorithms have an ability learn by itself or predict future for unseen data.

What is machine learning?

“Learning is any process by which a system improves performance from experience.”

- Herbert Simon

Definition by Tom Mitchell (1998):

Machine Learning is the study of algorithms that

- improve their performance P
- at some task T
- with experience E .

A well-defined learning task is given by $\langle P, T, E \rangle$.

Traditional Programming



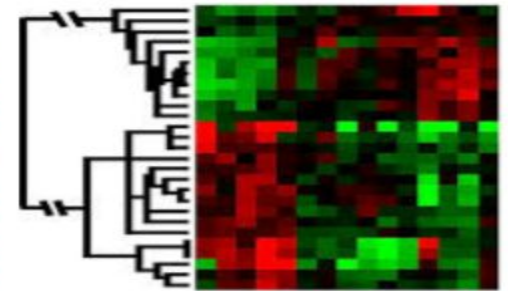
Machine Learning



Why Do We Use ML

ML is used when:

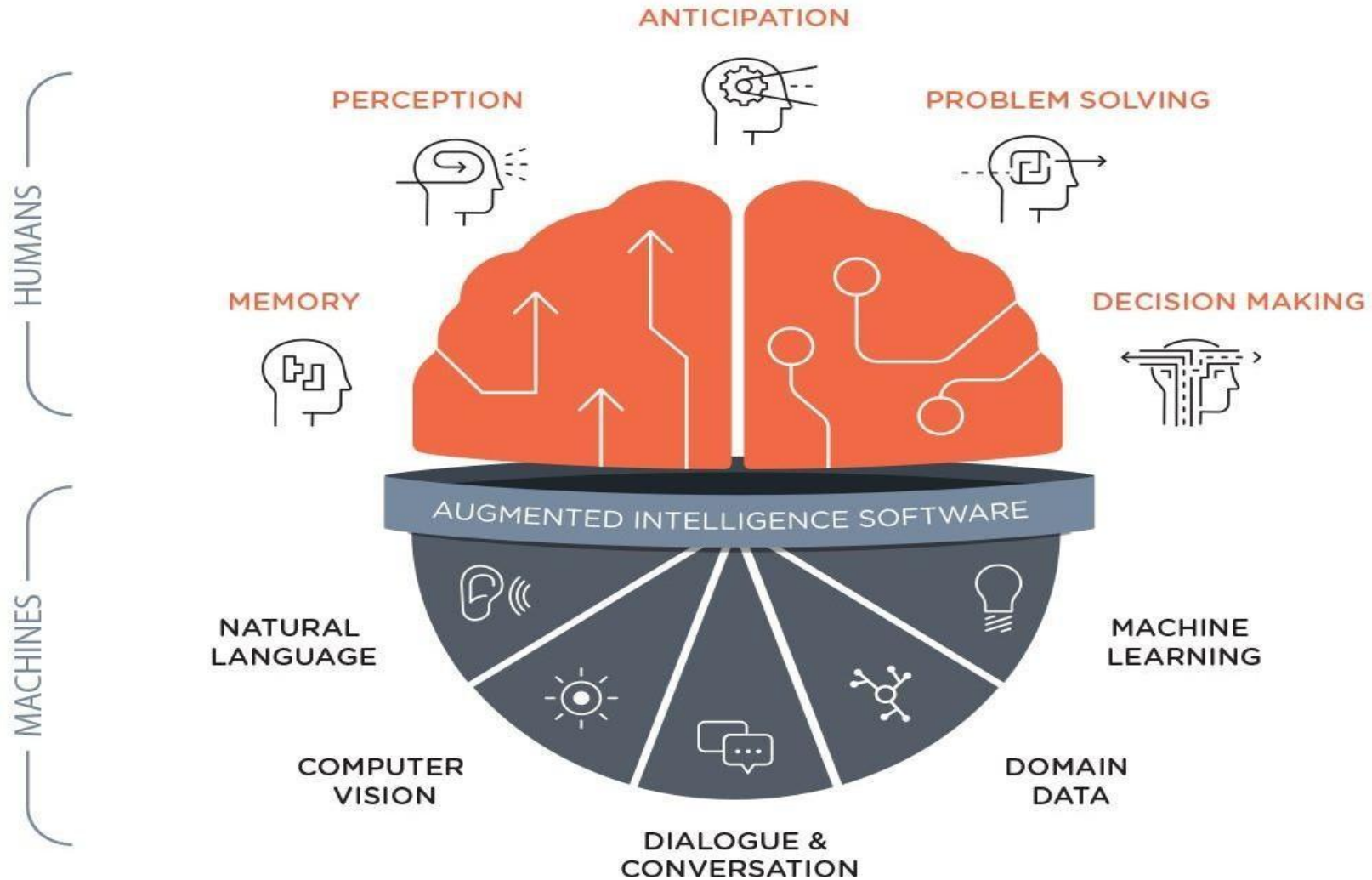
- Human expertise does not exist (navigating on Mars)
- Humans can't explain their expertise (speech recognition)
- Models must be customized (personalized medicine)
- Models are based on huge amounts of data (genomics)



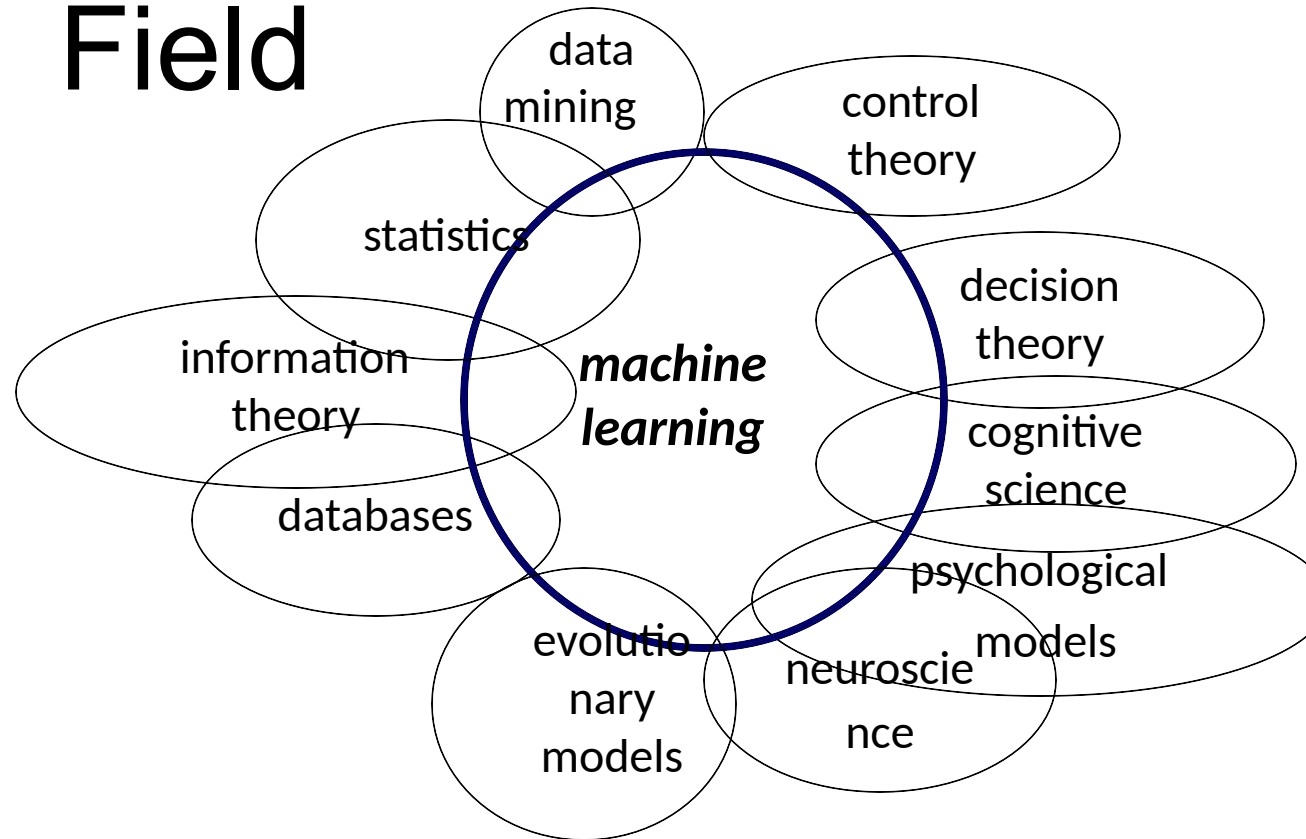
Learning isn't always useful:

- There is no need to “learn” to calculate payroll

Artificial Intelligence



Multidisciplinary Field



Machine learning is primarily concerned with the accuracy and effectiveness of the *computer system*.

Understanding Machine Learning

Machine Learning vs Statistical Inference vs Pattern Recognition vs Data Mining

Perspective:

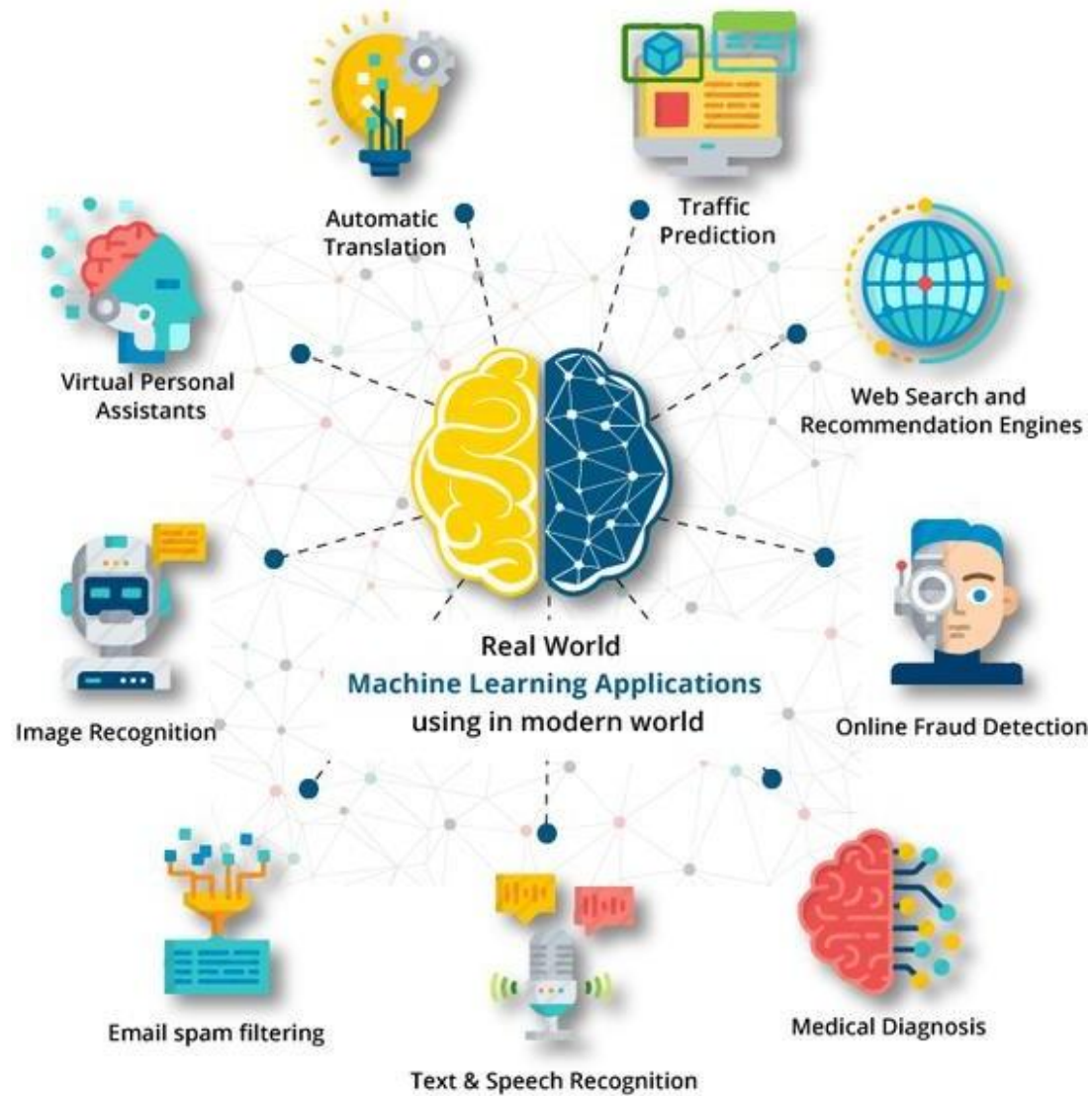
same concepts evolving in different scientific traditions

- **Statistical Inference (SI)**: field of Applied Mathematics
- **Machine Learning (ML)**: field of AI
- **Pattern Recognition (PR)**: branch of Computer Science focused on perception problems (image processing, speech recognition, etc.)
- **Data Mining (DM)**: field of Database Engineering

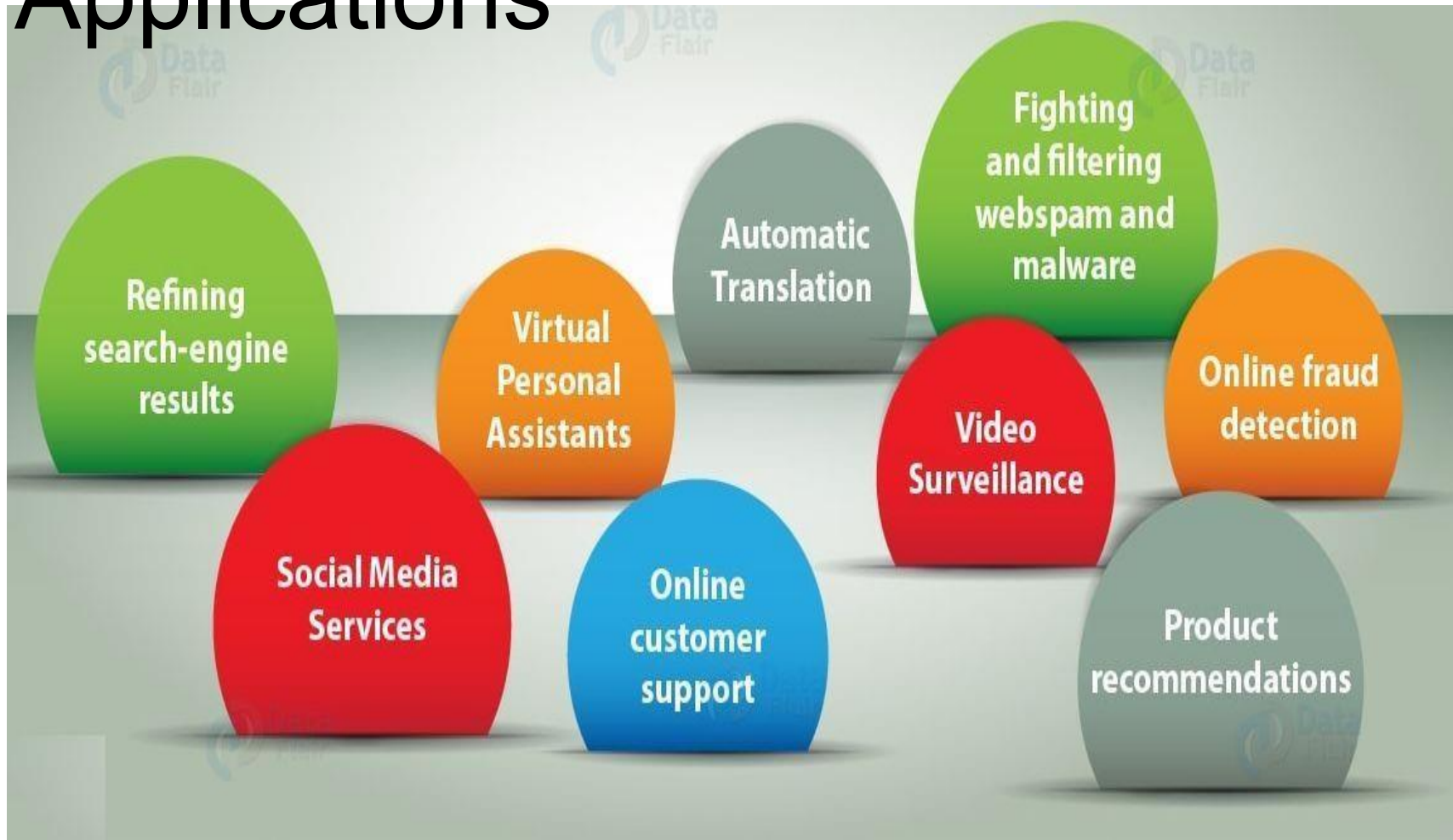
Why Machine

- **Learning?**
 - industrial/manufacturing control
 - mass spectrometer analysis, drug design, astronomic discovery
- **Black-box human expertise**
 - face/handwriting/speech recognition
 - driving a car, flying a plane
- **Rapidly changing phenomena**
 - credit scoring, financial modeling
 - diagnosis, fraud detection
- **Need for customization/personalization**
 - personalized news reader
 - movie/book recommendation

2. Applications of



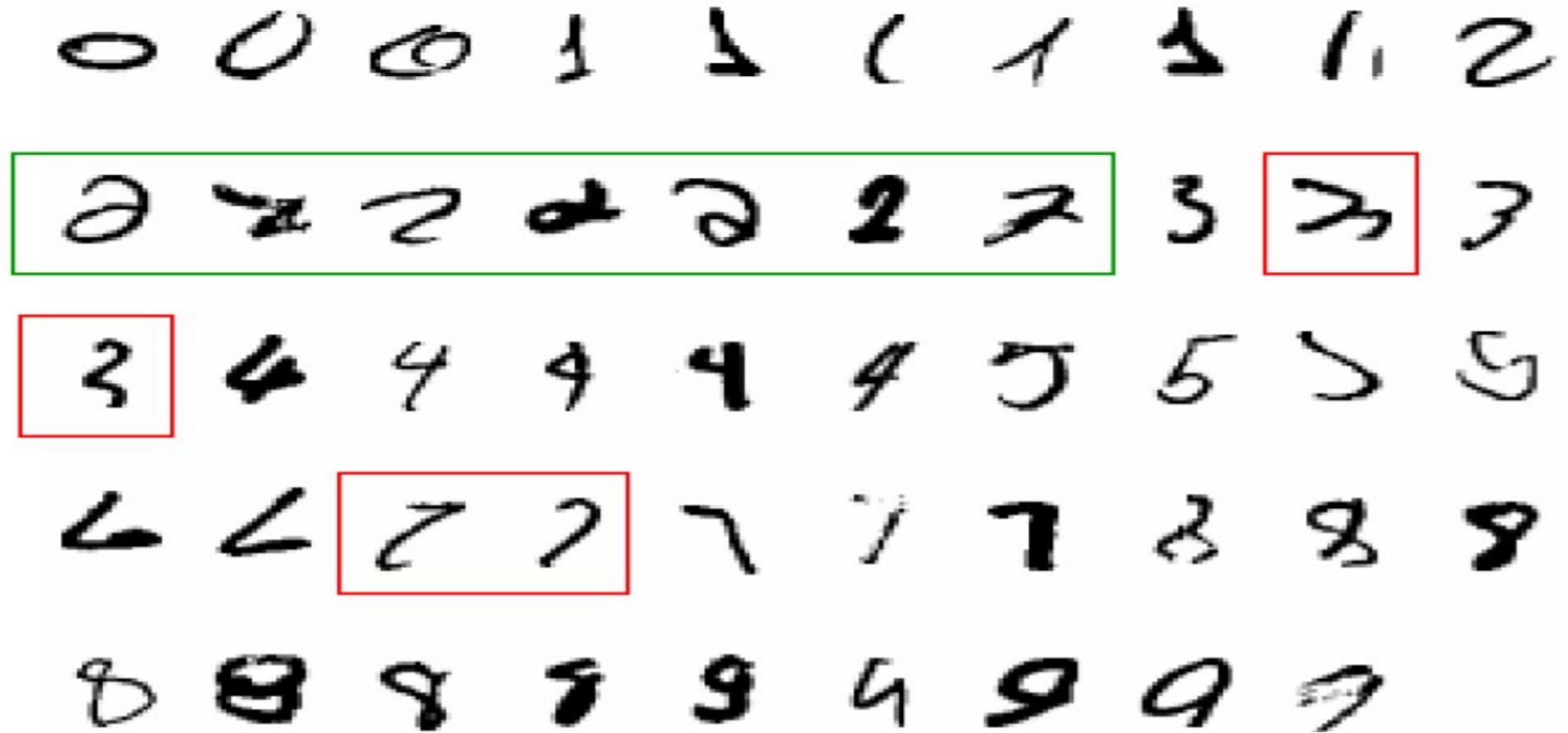
Machine Learning Applications



Some more examples of tasks that are best solved by using a learning algorithm

- **Recognizing patterns:**
 - Facial identities or facial expressions
 - Handwritten or spoken words
 - Medical images
- **Generating patterns:**
 - Generating images or motion sequences
- **Recognizing anomalies:**
 - Unusual credit card transactions
 - Unusual patterns of sensor readings in a nuclear power plant
- **Prediction:**
 - Future stock prices or currency exchange rates

A classic example of a task that requires machine learning: It is very hard to say what makes 2



Machine Learning touching our Daily Life



Walmart use Robots in their stores for inventory management, packing, pricing checks

Restaurants have Robot chefs and Waiters



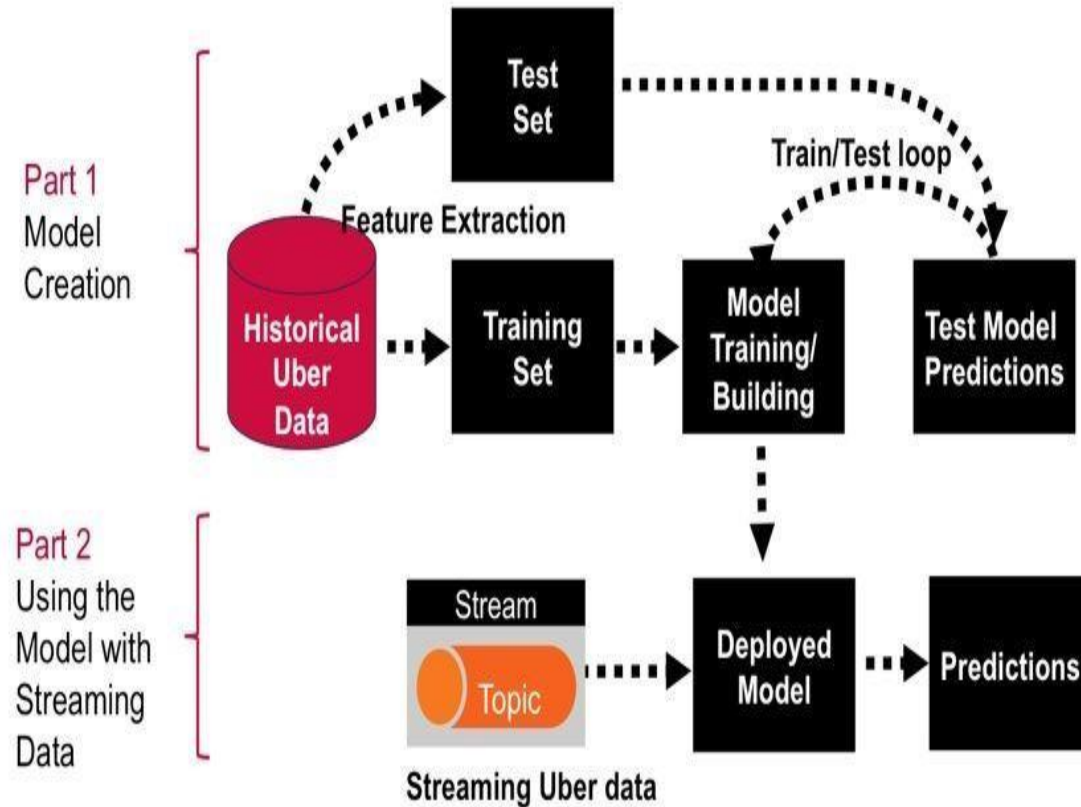
ML in Shopping Malls

Target knows when you are **pregnant**. A Minneapolis **man** discovered that his **daughter** was **pregnant** after **Target** started sending her coupons for baby clothes and cribs, according to an eye-opening New York Times story. "My **daughter** got this in the mail!" he said. ... Feb 20, 2012

 [www.citypages.com › news › target-coupons-inform-father-of-daught...](http://www.citypages.com/news/target-coupons-inform-father-of-daught...) ▼

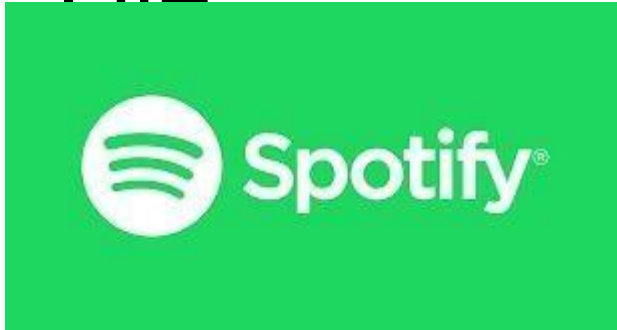
[Target coupons inform father of daughter's pregnancy | City ...](#)

Machine Learning touching our Daily Life



Michelangelo, an internal ML-as-a-service platform that democratizes machine learning and makes scaling AI to meet the needs of business as easy as requesting a ride.

Machine Learning touching our Daily Life

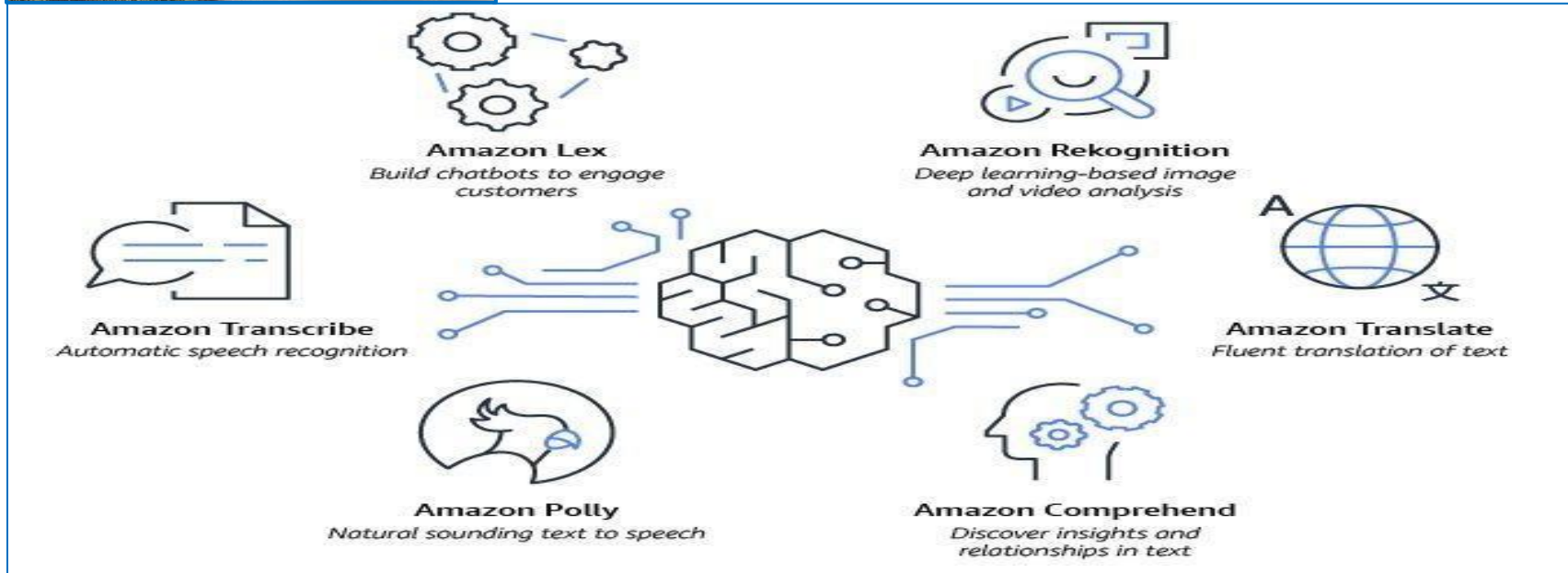


Song Recommendations based on mood and interest



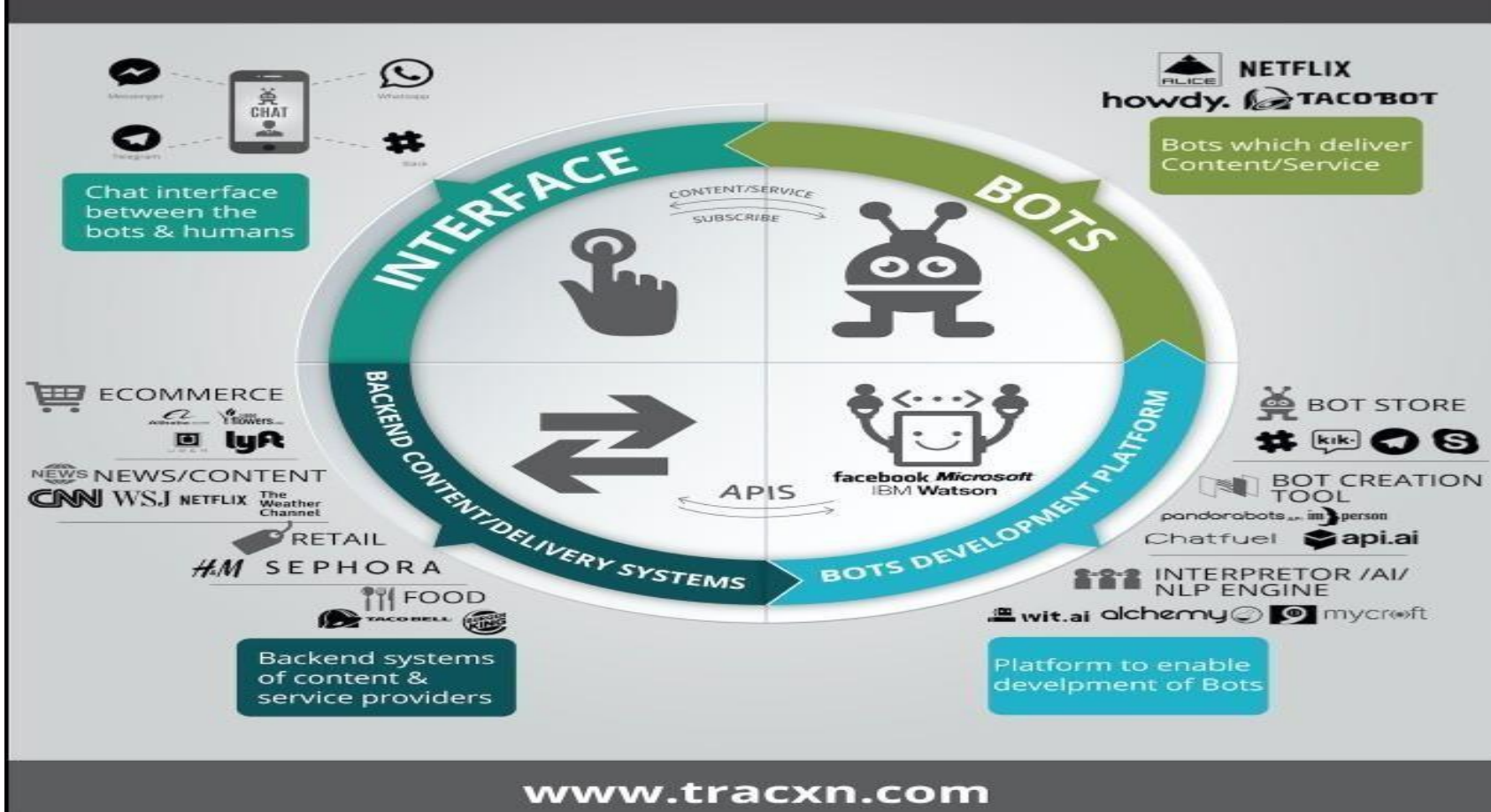
Content specific vaccines for Children

Amazon – Game Changer of the Decade



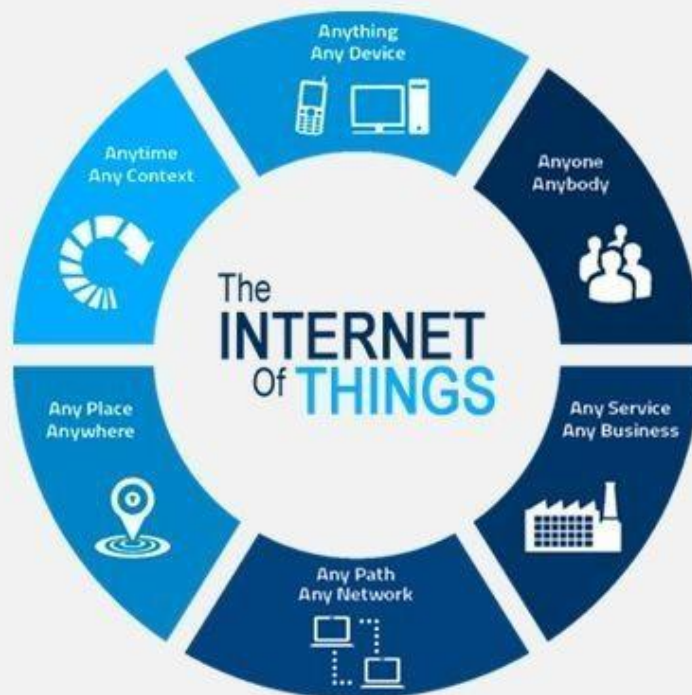
Chat bots

CHAT BOTS - VALUE CHAIN



Internet of Things

Global IoT spend
in manufacturing
to reach \$70
billion by 2020



20 billion
connected
devices by 2020

Personalization



Sophia
social humanoid



**Google
Assistant**



**Amazon
Alexa**
Virtual

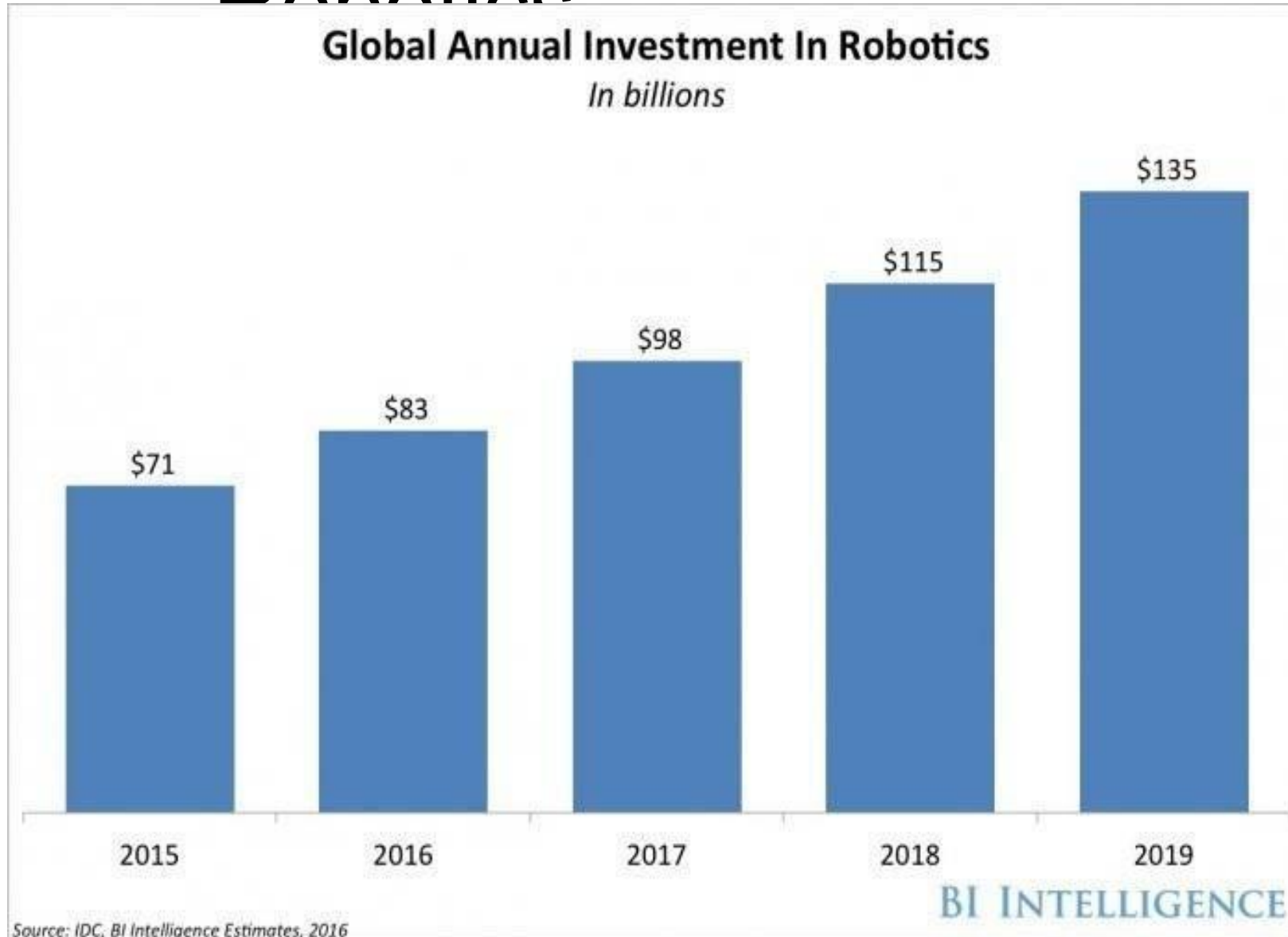
Personalization Platforms



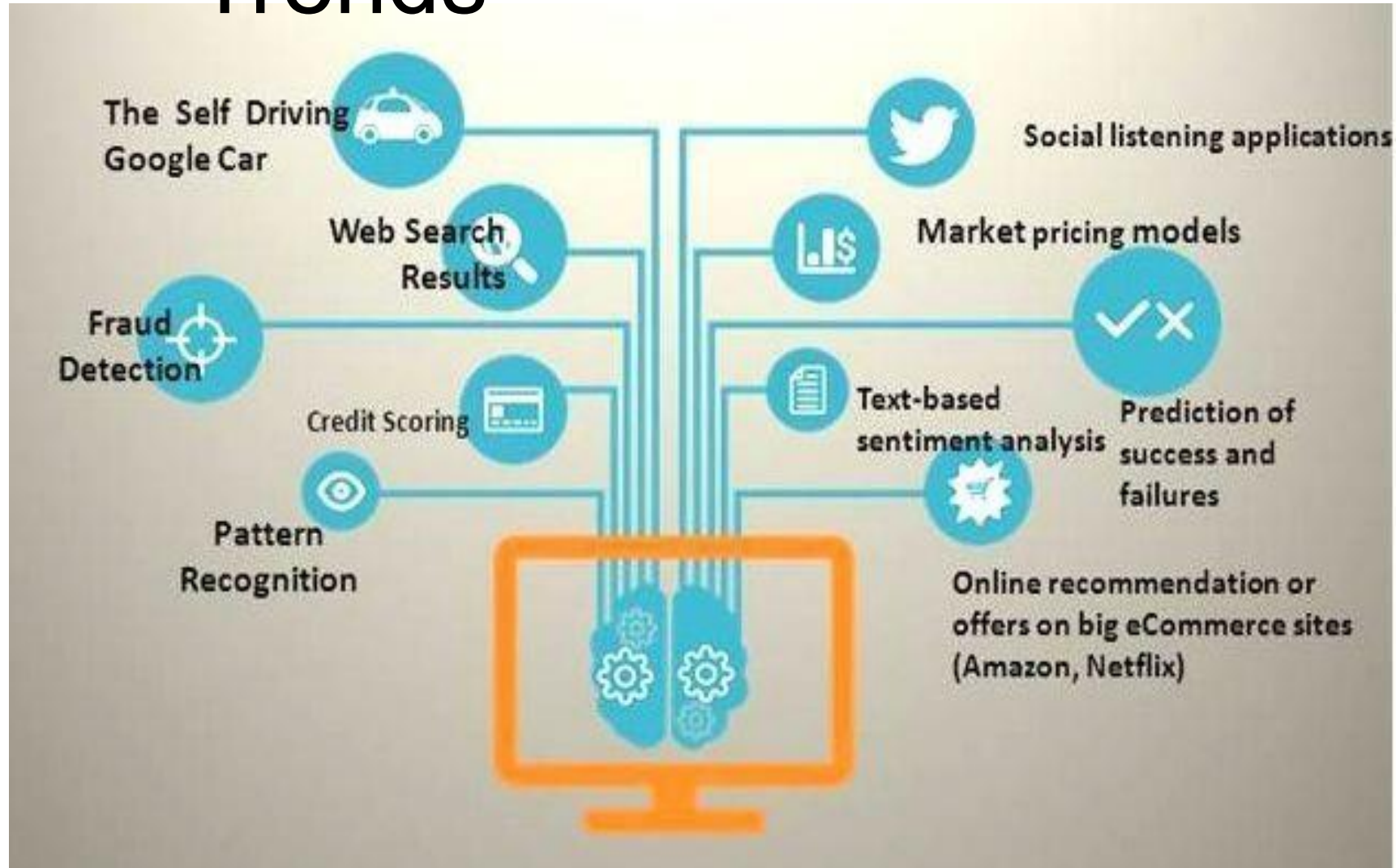
Machine Learning Applications across Industries



Investment in Robotics

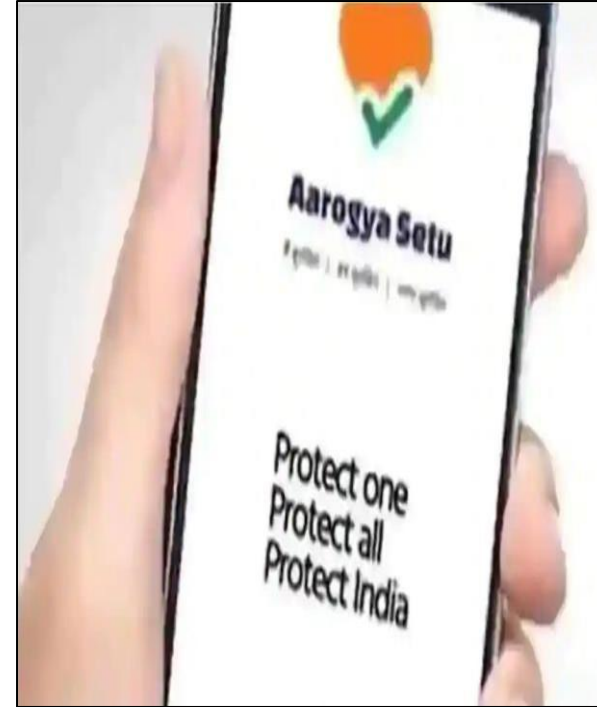


Machine Learning Trends



Machine Learning Trends

“Telephone took 75 years to reach 50 million users, radio 38 yrs, television 13 yrs, Internet 4 yrs, Facebook 19 months, Pokemon Go 19 days. AarogyaSetu, India’s app to fight COVID-19 has reached 50 mn users in **just 13 days-fastest** ever globally for an App,” Kant said in his tweet.



The app will calculate this based on their interaction with others, using cutting edge algorithms and artificial intelligence.

3. Terminology in ML

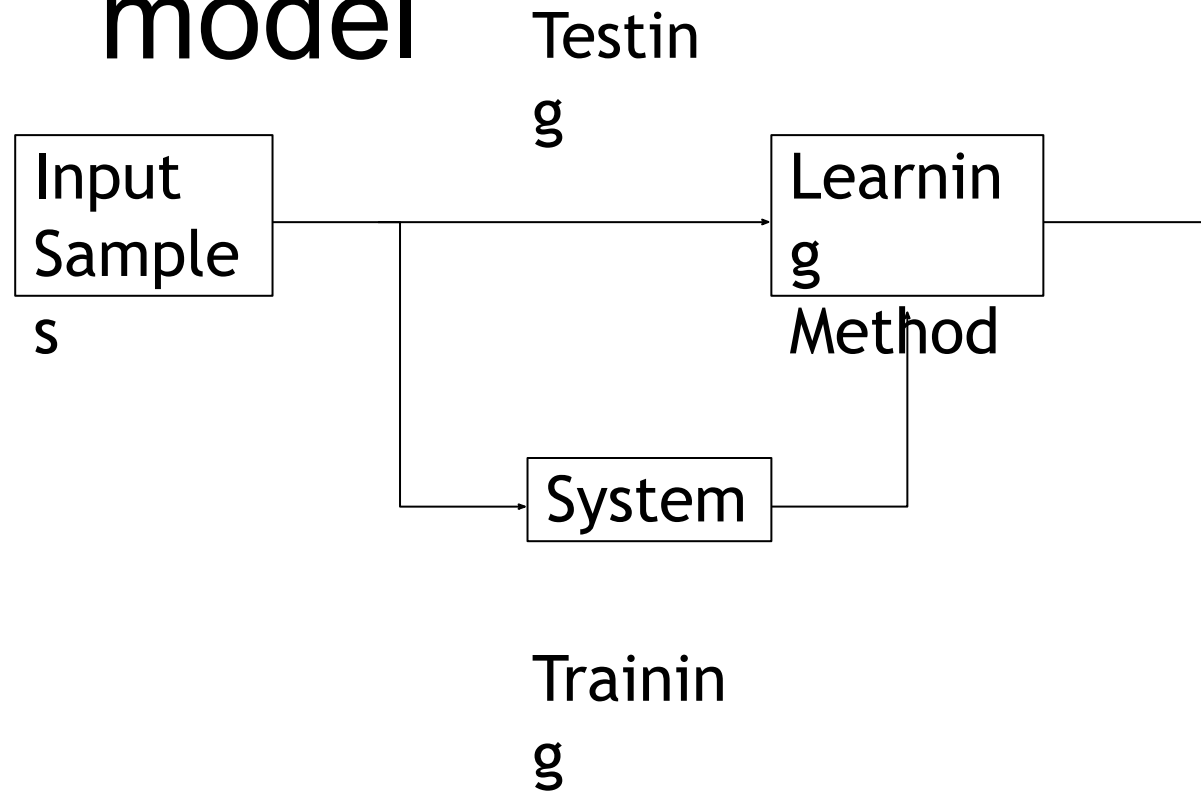
1. **Dataset** – A set of data examples, that contain important features to solve a problem
2. **Features** – How datas are interpreted for data extraction
3. **Model** – The rule which predicts the output with given input
4. **Attribute** - A quality describing an observation (e.g. color, size, weight). In Excel terms, these are column headers.

ML Terminology

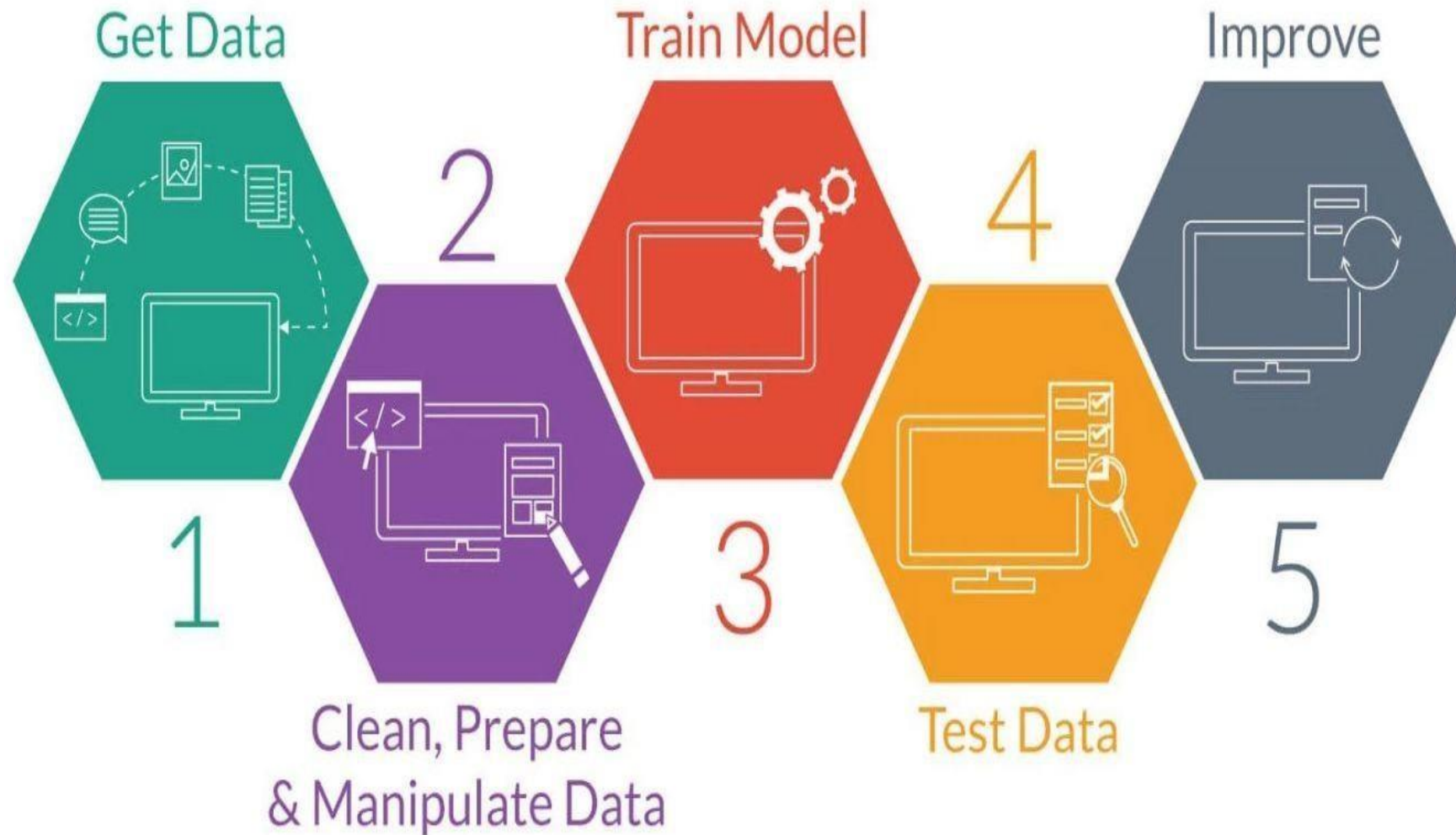
- **Examples:** Items or instances used for learning or evaluation.
- **Features:** Set of attributes represented as a vector associated with an example.
- **Labels:** Values or categories assigned to examples. In classification the labels are categories; in regression the labels are real numbers.
- **Target:** The correct label for a training example. This is extra data that is needed for supervised learning.
- **Output:** Prediction label from input set of features using a model of the machine learning algorithm.
- **Training sample:** Examples used to train a machine learning algorithm.
- **Validation sample:** Examples used to tune parameters of a learning algorithm.
- **Model:** Information that the machine learning algorithm stores after training. The model is used when predicting the output labels of new, unseen examples.
- **Test sample:** Examples used to evaluate the performance of a learning algorithm. The test sample is separate from the training and validation data and is not made available in the learning stage.
- **Loss function:** A function that measures the difference/loss between a predicted label and a true label. We will design the learning algorithms so that they minimize the error (cumulative loss across all training examples).
- **Hypothesis set:** A set of functions mapping features (feature vectors) to the set of labels. The learning algorithm chooses one function among those in the hypothesis set to return after training. Usually we pick a class of functions (e.g., linear functions) parameterized by a set of free parameters (e.g., coefficients of the linear function) and pinpoint the final hypothesis by identifying the parameters that minimize the error.
- **Model selection:** Process for selecting the free parameters of the algorithm (actually of the function in the hypothesis set).

4. ML Process

Learning system model

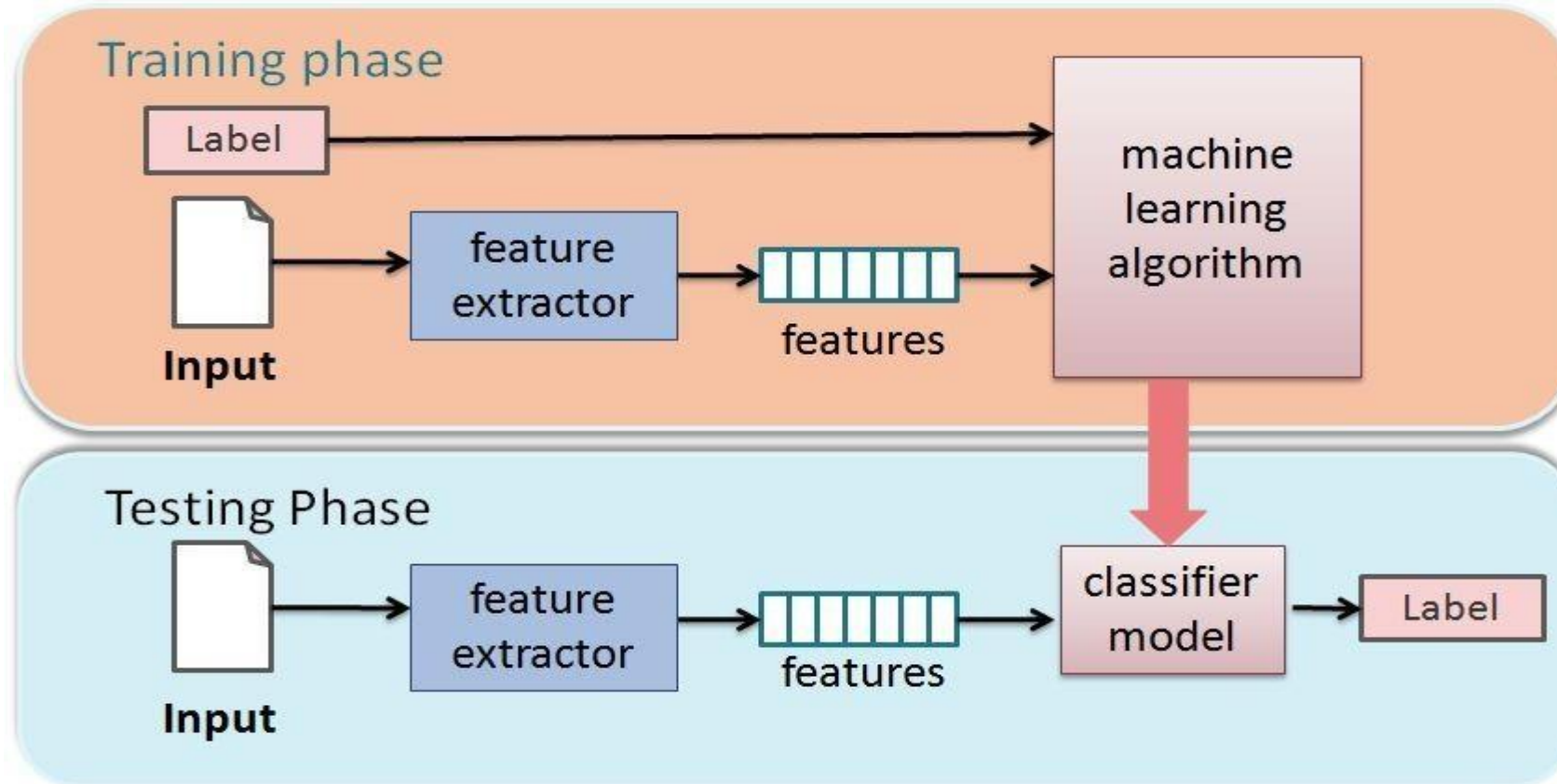


Machine Learning Workflow



Training and testing

- Training is the process of making the system able to learn.
- No free lunch rule:
 - Training set and testing set come from the same distribution
 - Need to make some assumptions or bias

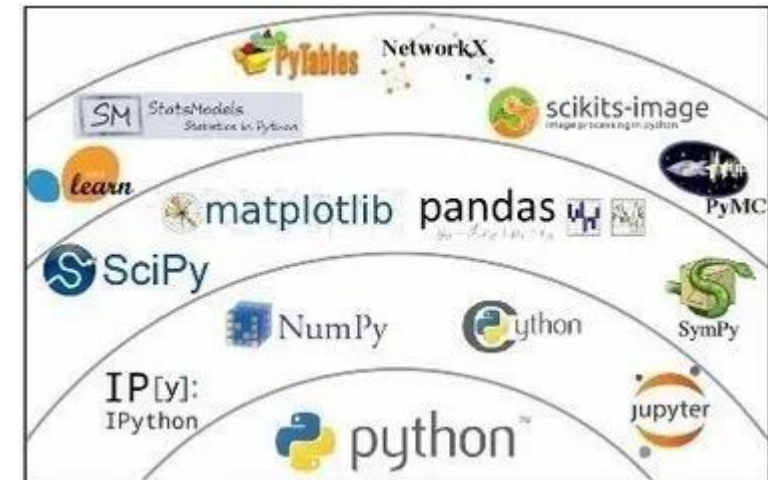


10 REASONS WHY PYTHON FOR MACHINE LEARNING

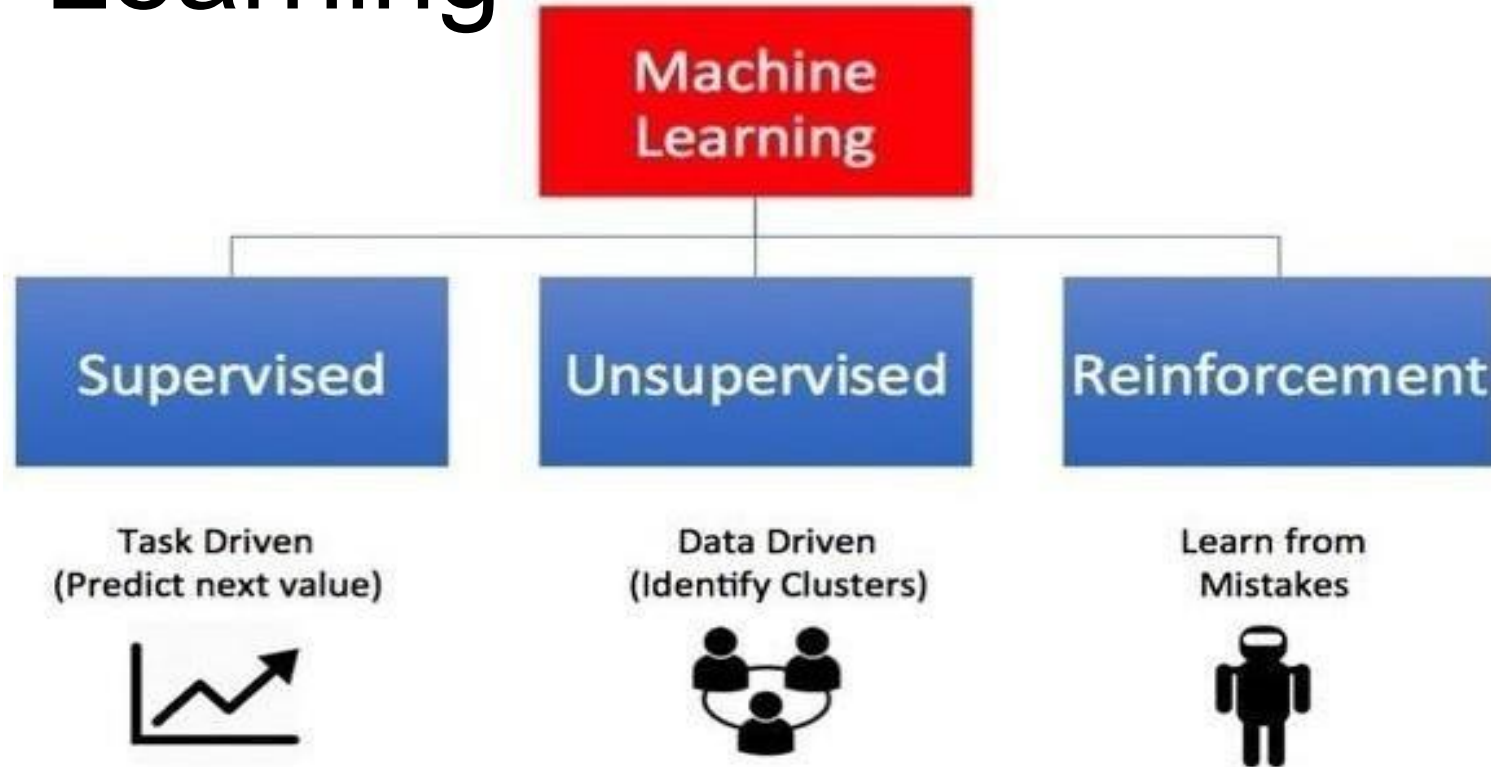


5. Top Python Machine Learning Libraries

- PANDAS
- NUMPY
- MATPLOTLIB
- SKLEARN
- SEABORN
- SciPy
- PyTorch
- Keras - an open-source library used for neural networks and machine learning.
- TensorFlow



6.1 Types of Machine Learning



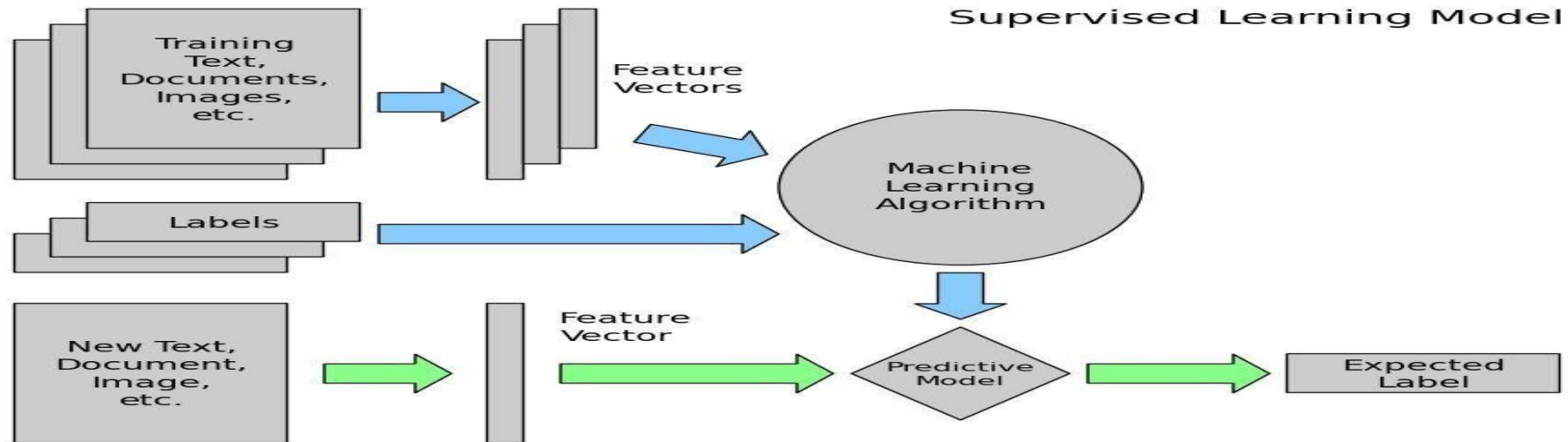
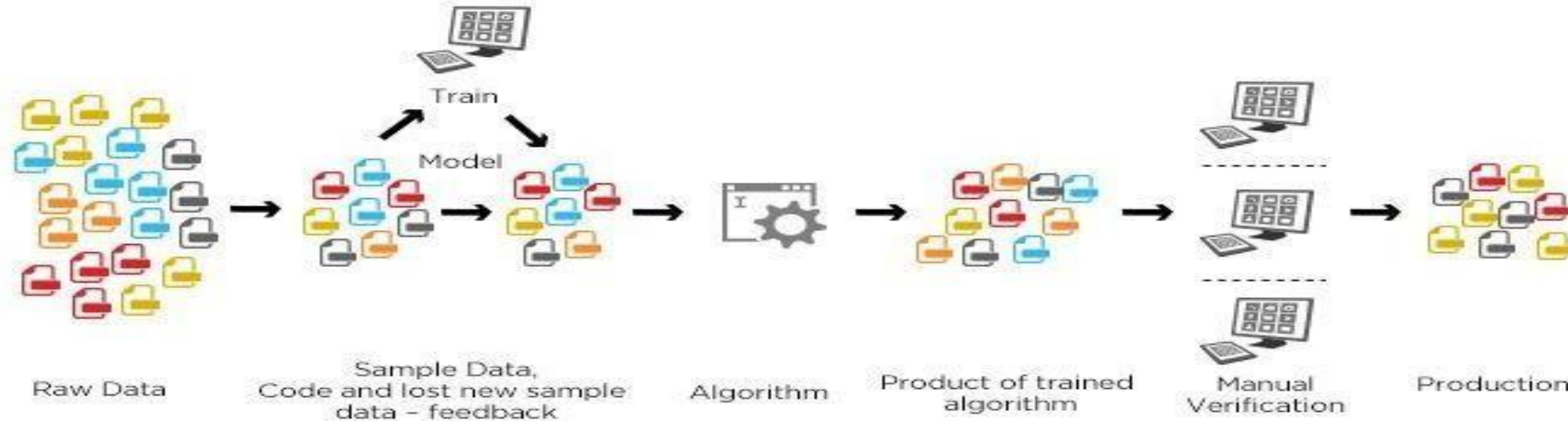
Learning Paradigms

Supervised learning

- Generates a function that maps inputs to desired outputs.
 - For example, in a classification problem, the learner approximates a function mapping a vector into classes by looking at input-output examples of the function
- Probably, the most common paradigm
- E.g., decision trees, support vector machines, Naïve Bayes, k-Nearest Neighbors, ...

Machine learning structure

- Supervised learning

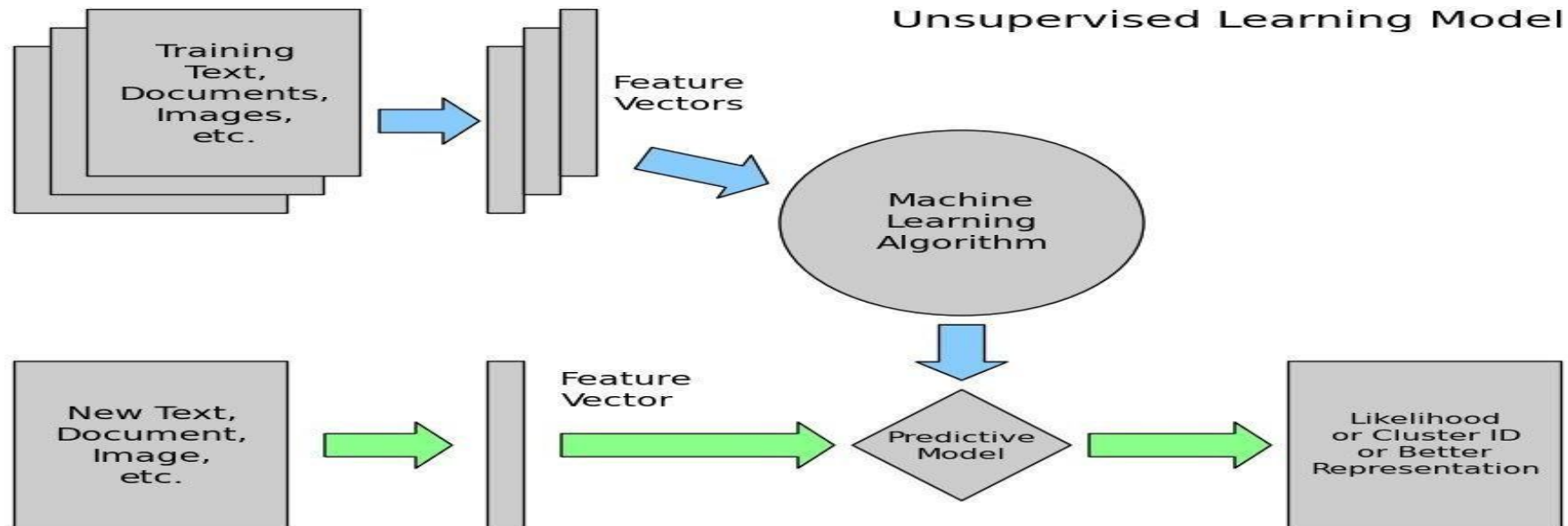
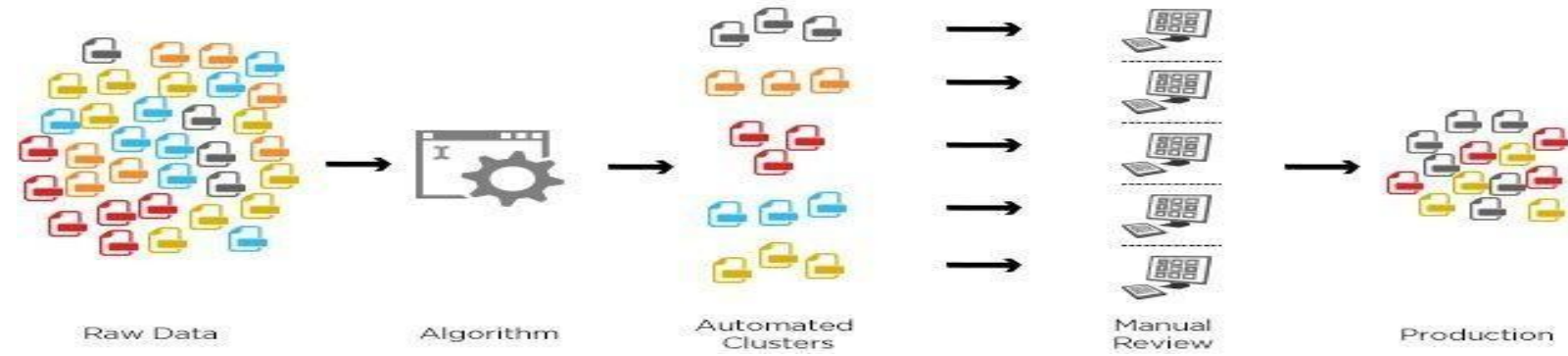


Learning Paradigms

- **Unsupervised learning**
 - Labels are not known during training
 - E.g., clustering, association learning
- **Semi-supervised learning**
 - Combines both labeled and unlabeled examples to generate an appropriate function or classifier
 - E.g., Transductive Support Vector Machine

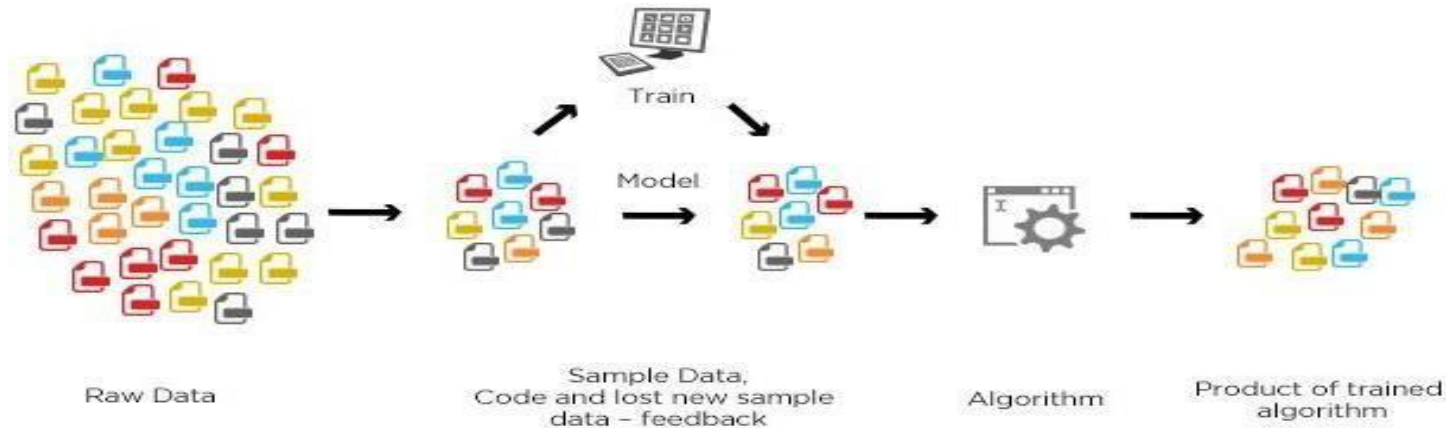
Machine learning structure

- Unsupervised learning



Machine learning structure

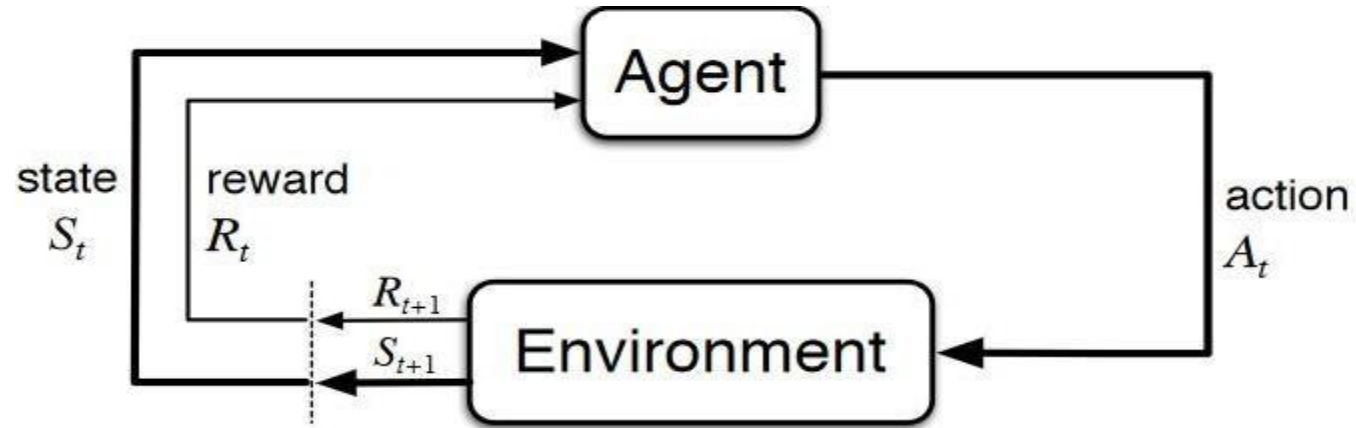
- semisupervised learning



The goal of a semi-supervised model is to classify some of the unlabeled data using the labeled information set.

- Speech Analysis
- Protein Sequencing
- Web content analysis

Reinforcement Learning

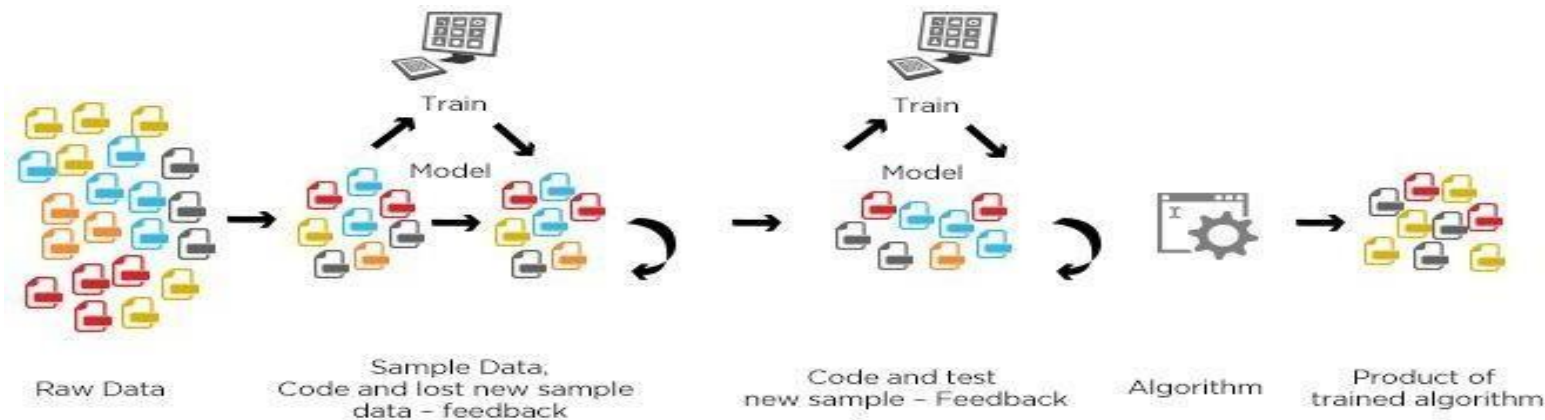


- In the absence of training dataset, it is bound to **learn** from its experience.
- We have an agent and a reward, with many hurdles in between. The agent is supposed to find the best possible path to reach the reward.

Types of Reinforcement:

positive, negative, punishment, and extinction.

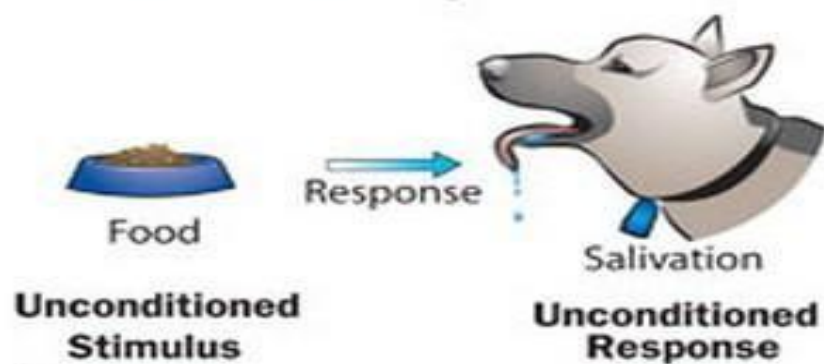
- Reinforcement learning Machine learning structure



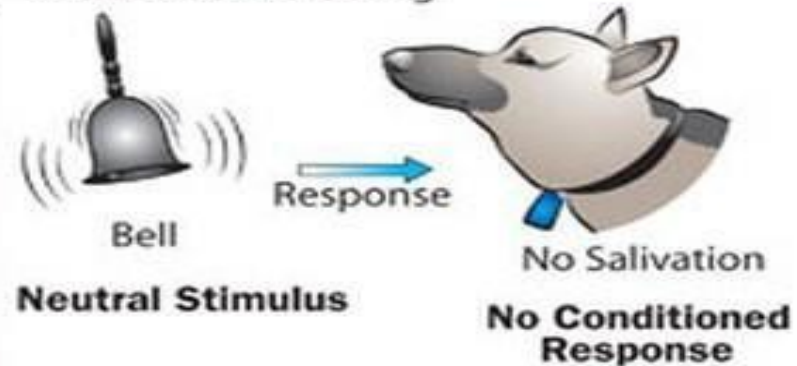
- It is concerned with how an agent should take actions in an environment so as to maximize some notion of cumulative reward.
 - Reward given if some evaluation metric improved
 - Punishment in the reverse case
- E.g., Q-learning, Sarsa

How Dog Training Works

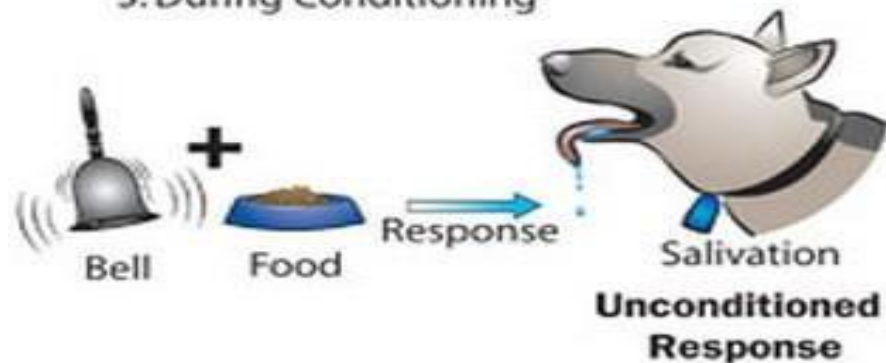
1. Before Conditioning



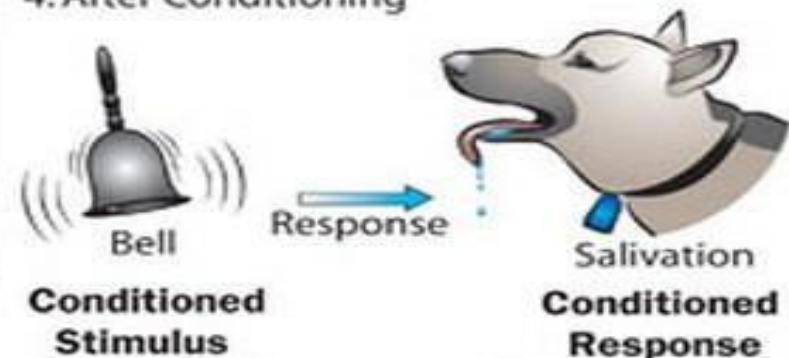
2. Before Conditioning



3. During Conditioning

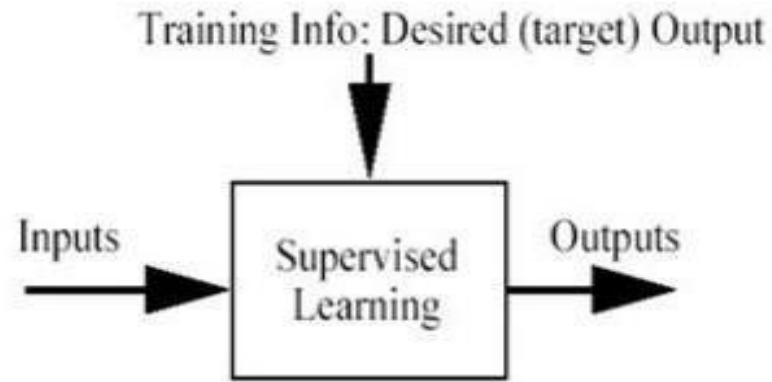


4. After Conditioning



©2006 HowStuffWorks

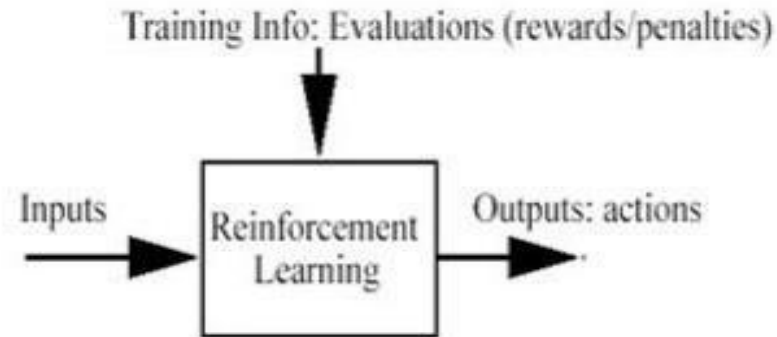
Supervised Learning



$$\text{Error} = (\text{target output} - \text{actual output})$$

Input is an instance, output is a classification of the instance

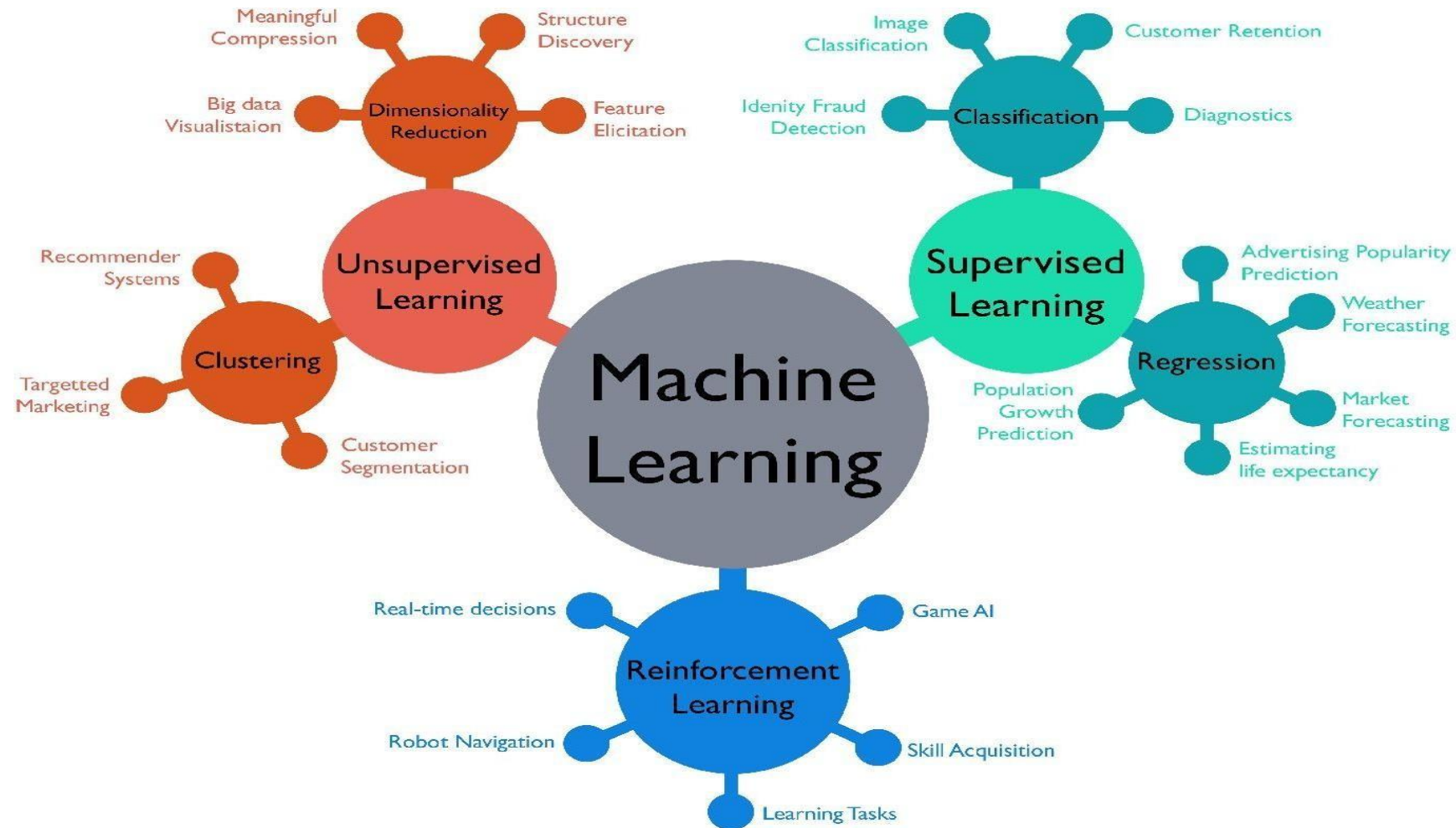
Reinforcement Learning



Objective: Get as much reward as possible

Input is some "goal" output
Is a sequence of actions to meet the goal

Machine Learning Perspectives



ML Algorithms

- **Supervised learning**
 - Prediction
 - Classification (discrete labels), Regression (real values)
- **Unsupervised learning**
 - Clustering
 - Probability distribution estimation
 - Finding association (in features)
 - Dimension reduction
- **Semi-supervised learning**
- **Reinforcement learning**
 - Decision making (robot, chess machine)

ML Algorithms

- **Classification**

- Learn a way to classify **unseen examples**, based on a set of labeled examples, e.g., classify songs by emotion categories. E.g., decision trees (e.g., C5.4)

- **Regression**

- Learn a way to **predict continuous output values**, based on a set of labeled examples, e.g., predict software development effort in person months
- Sometimes regarded as **numeric classification** (outputs are continuous instead of discrete)
 - E.g., Support Vector, Regression

ML Algorithms

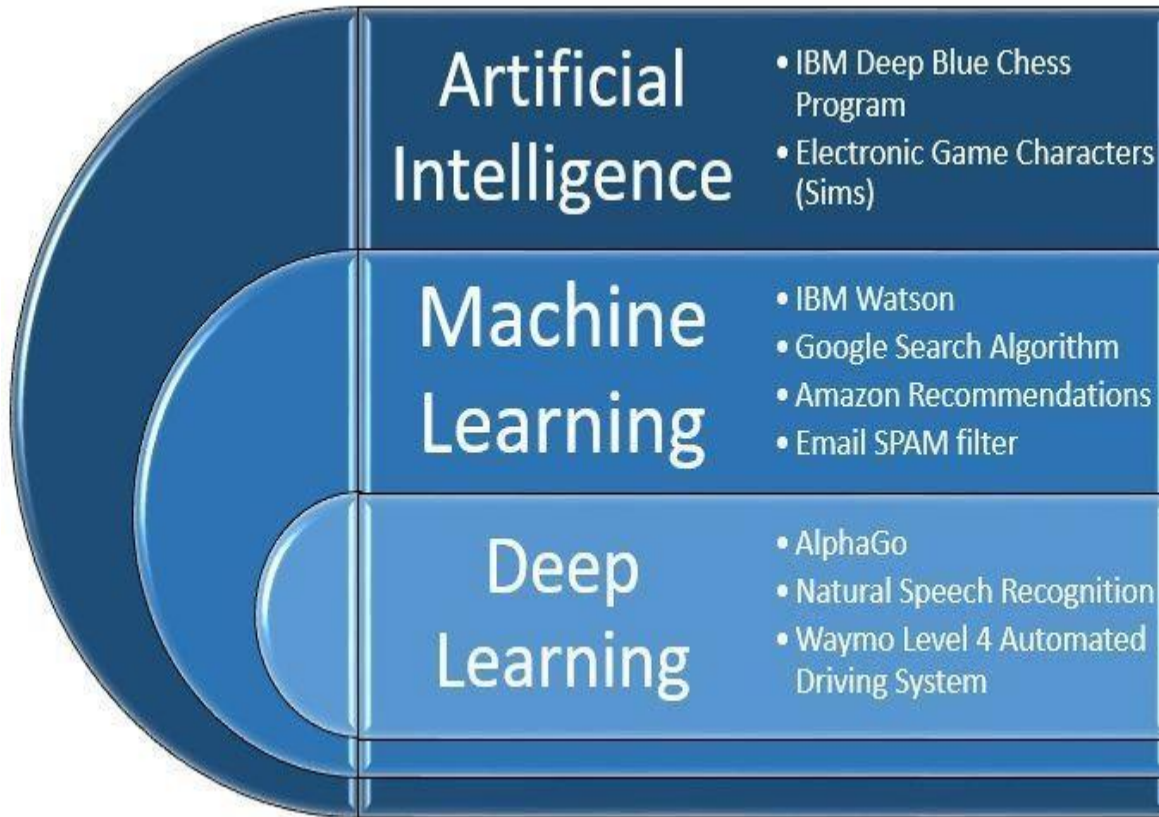
- **Association**

- Find any association among features, not just input-output associations (e.g., in a supermarket, find that clients who buys apples also buys cereals)
- E.g., Apriori

- **Clustering**

- Find natural grouping among data
- E.g., K-means clustering, DBSCAN, Heirarchial clustering

AI vs Machine Learning vs Deep Learning



Programs with the ability to learn and reason like humans

Algorithms with the ability to learn without being explicitly programmed

Subset of Machine Learning in which Artificial Neural Networks adapt and learn from vast amount of data

Deep Learning

Traditional Pattern Recognition: Fixed/Handcrafted Feature Extractor



Mainstream Modern Pattern Recognition: Unsupervised mid-level features



Deep Learning: Representations are hierarchical and trained



Future of machine learning

- Improved unsupervised algorithms
- Enhanced personalization
- Increased adoption of quantum computing
- Improved cognitive services
- Rise of robots

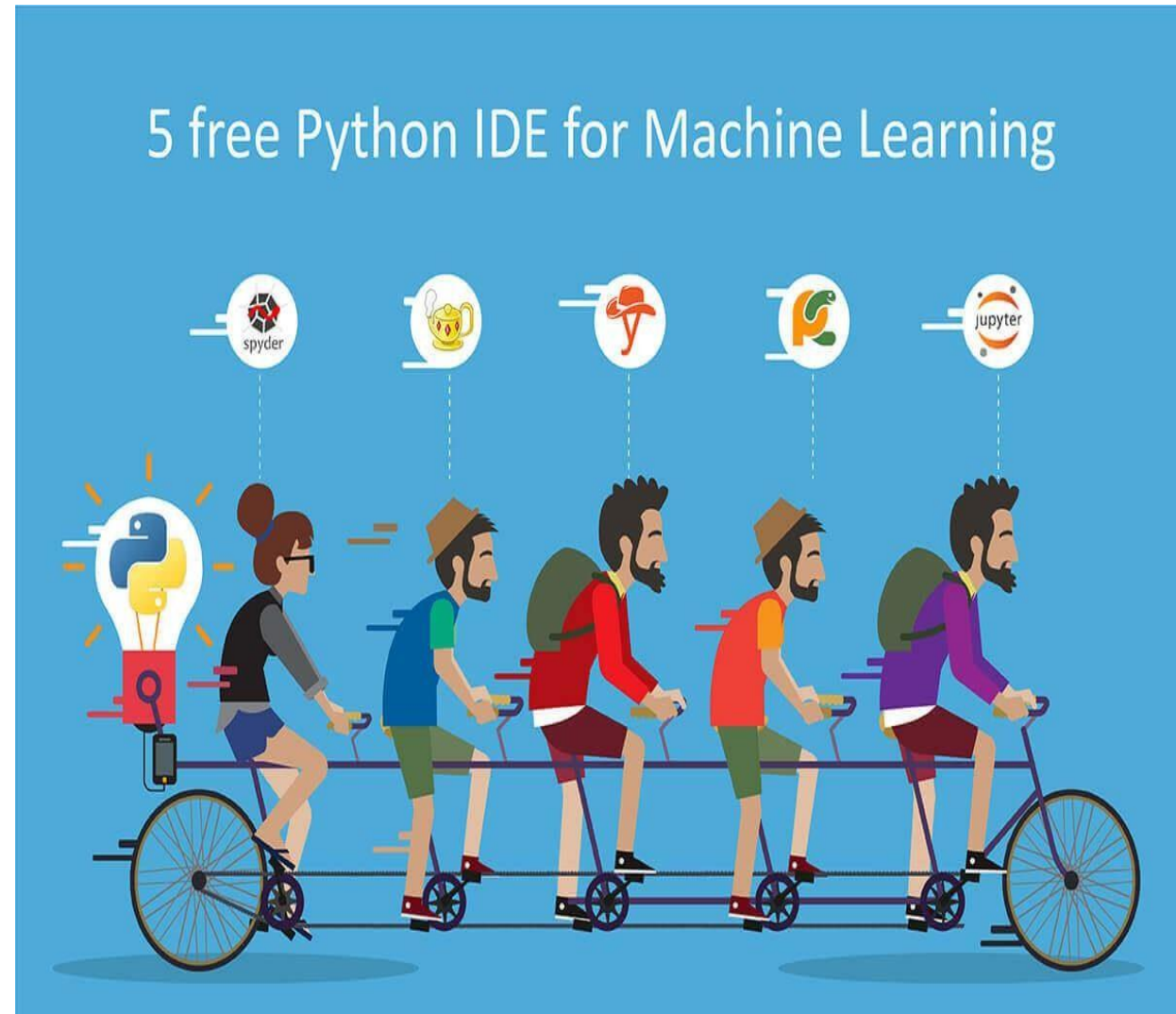
WORKING PROCESS WITH DATASET



1. Gathering data
2. Import Package
3. Data Analysis
4. **Data Visualization**
5. Data pre-processing
6. Model selection
7. Training and testing the model
8. Evaluation

List of IDEs

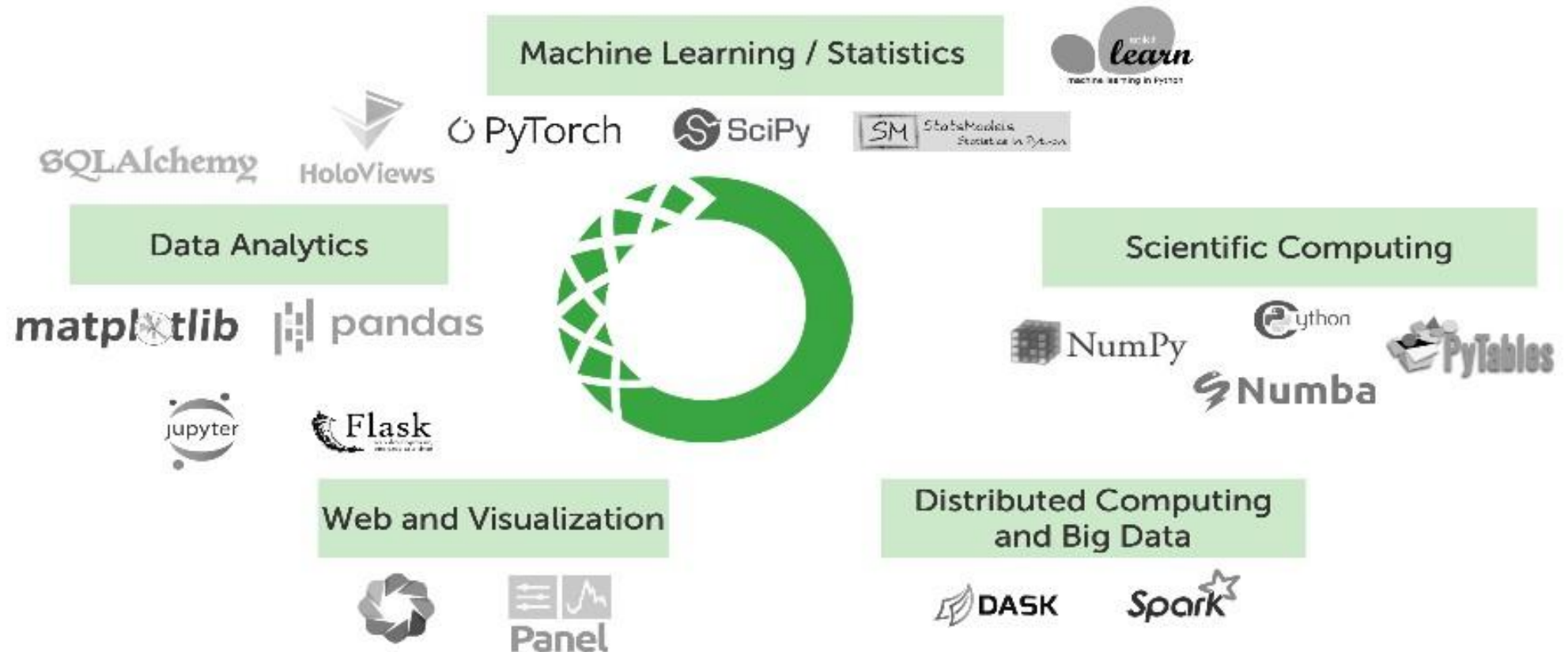
- **JuPyter/IPython Notebook**
- **PyCharm**
- **Spyder** - Spyder stands for Scientific **PY**thon **D**evelopment **E**nvi**R**onment.
- **Rodeo**
- **Geany**
- **Rstudio**
- **Scala IDE for Eclipse**
- **Visual Studio Code**
- **Atom**
- **Sublime Text 3**



8. Anaconda Installations



Why is Data Science so complicated?



Anaconda Distribution

- Thousands of curated packages
 - Analysis
 - Visualization
 - Modeling
- Mac OS, Linux, and Windows
 - 200+ packages pre-installed
 - It “just works”



ANACONDA.

Conda package manager

- Install, remove, and update packages
- Automatically installs dependencies
 - At the *correctly matched* versions
- Install packages written in any language
 - Not just for Python and R
- No compilers needed to install packages
- Works on Mac, Linux, and Windows
- Curate your own packages with
 - Anaconda Team Edition or Anaconda Enterprise Edition



Python's Jupyter Notebook

- Programming in the browser
- Code, Instructions and output are displayed "in-line"
- Useful for writing code that tells a story
- Used by scientists and researchers

Anaconda Installation

Link

- <https://www.anaconda.com/products/individual>

“Artificial Intelligence, Deep Learning,
Machine Learning — whatever you’re doing
if you don’t understand it — learn it.
Because otherwise you’re going to
be a dinosaur within 3 years.”