AI & DS

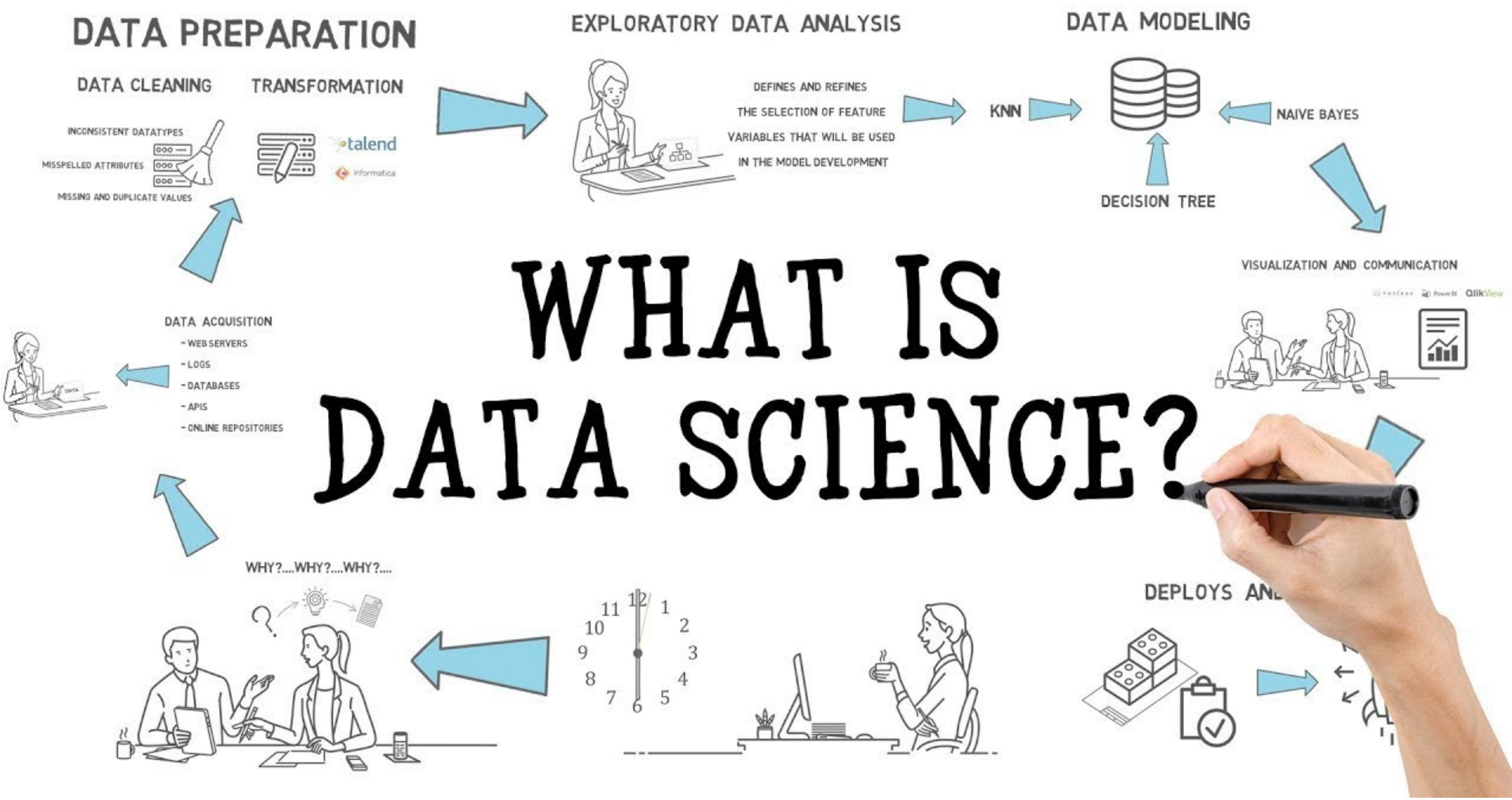
-CURRENT TRENDS

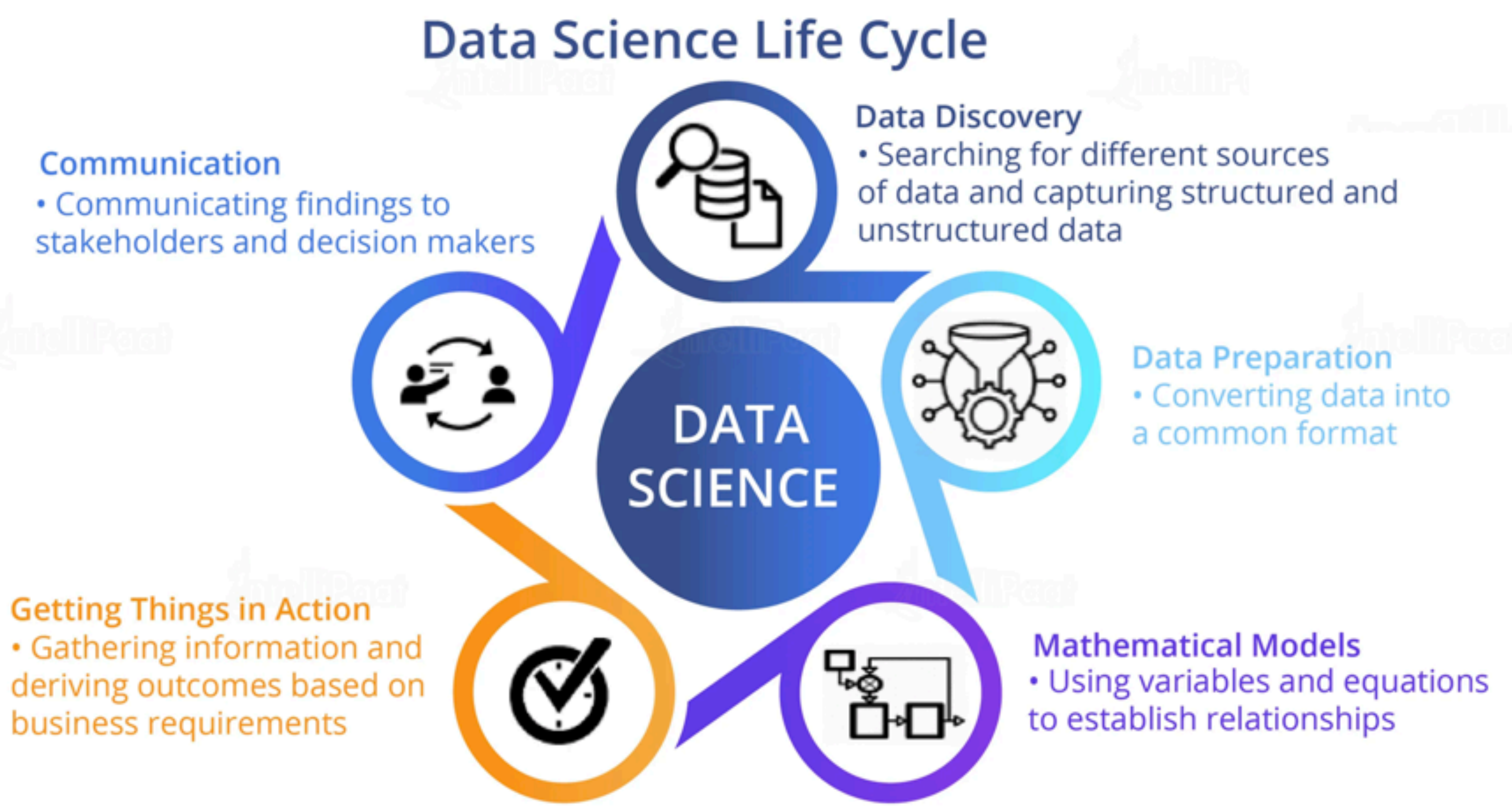
MR. PRANAV SHASTRI

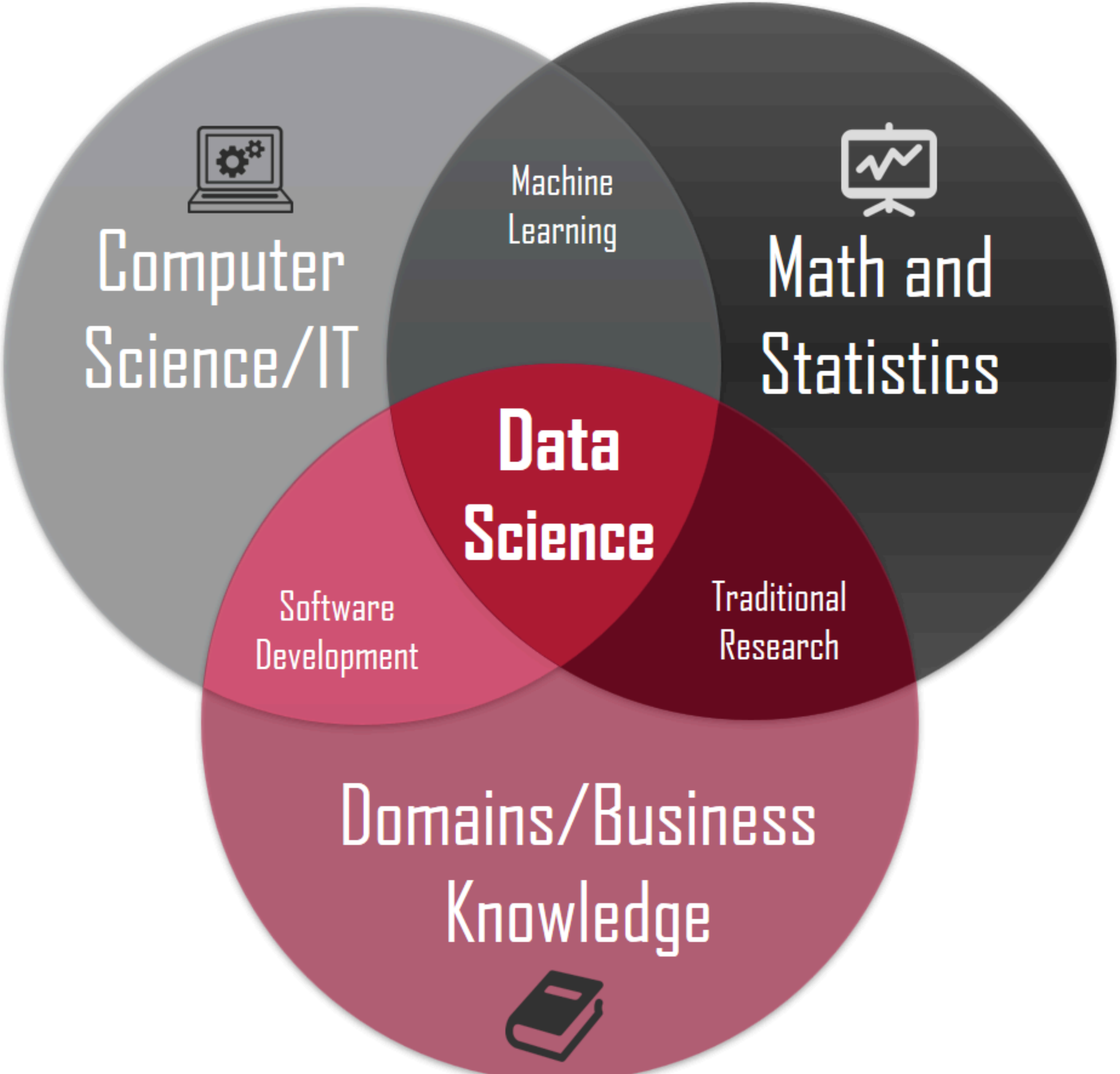
MASTER TRAINER,

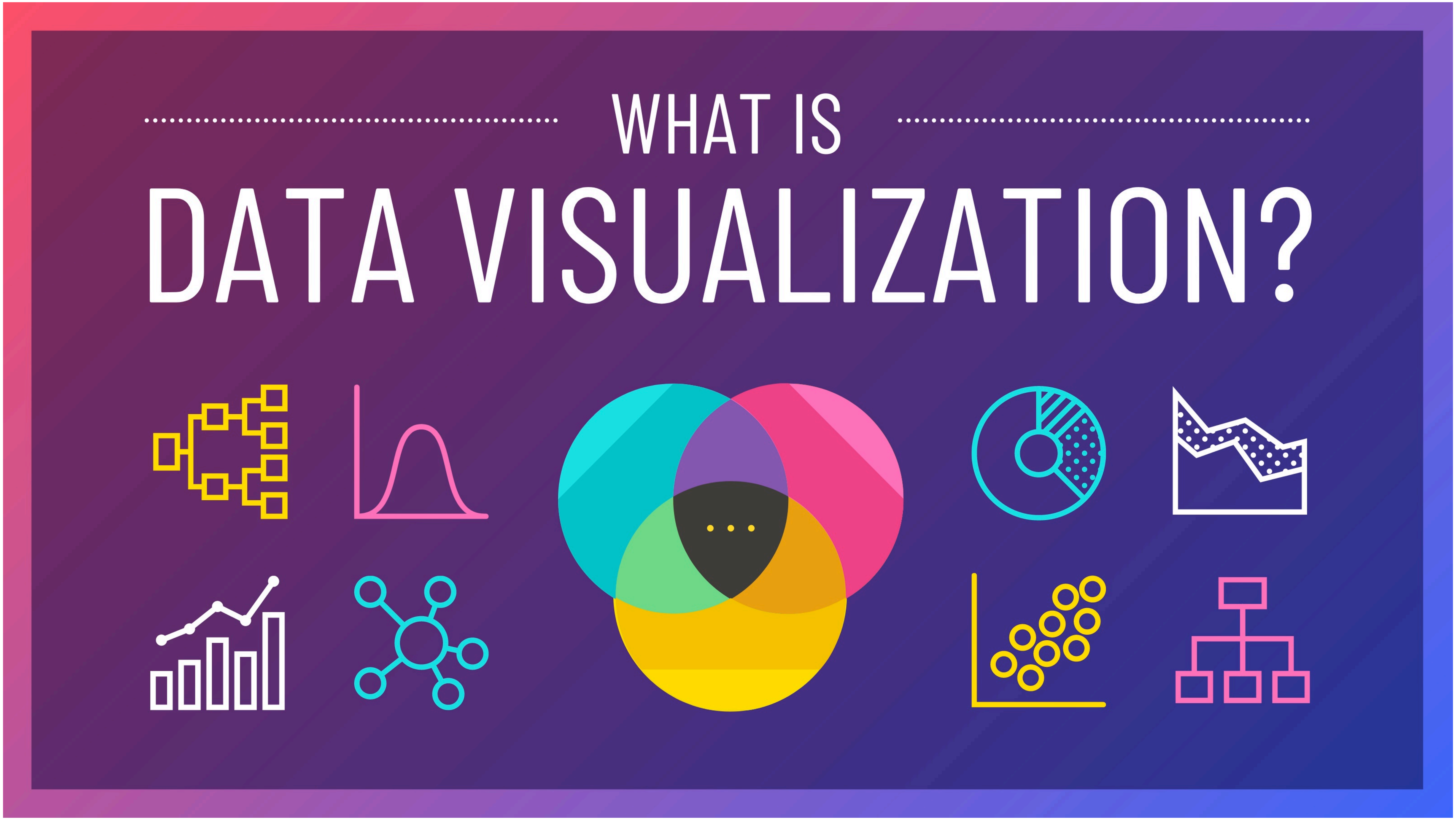
WILEYS

DATA SCIENCE

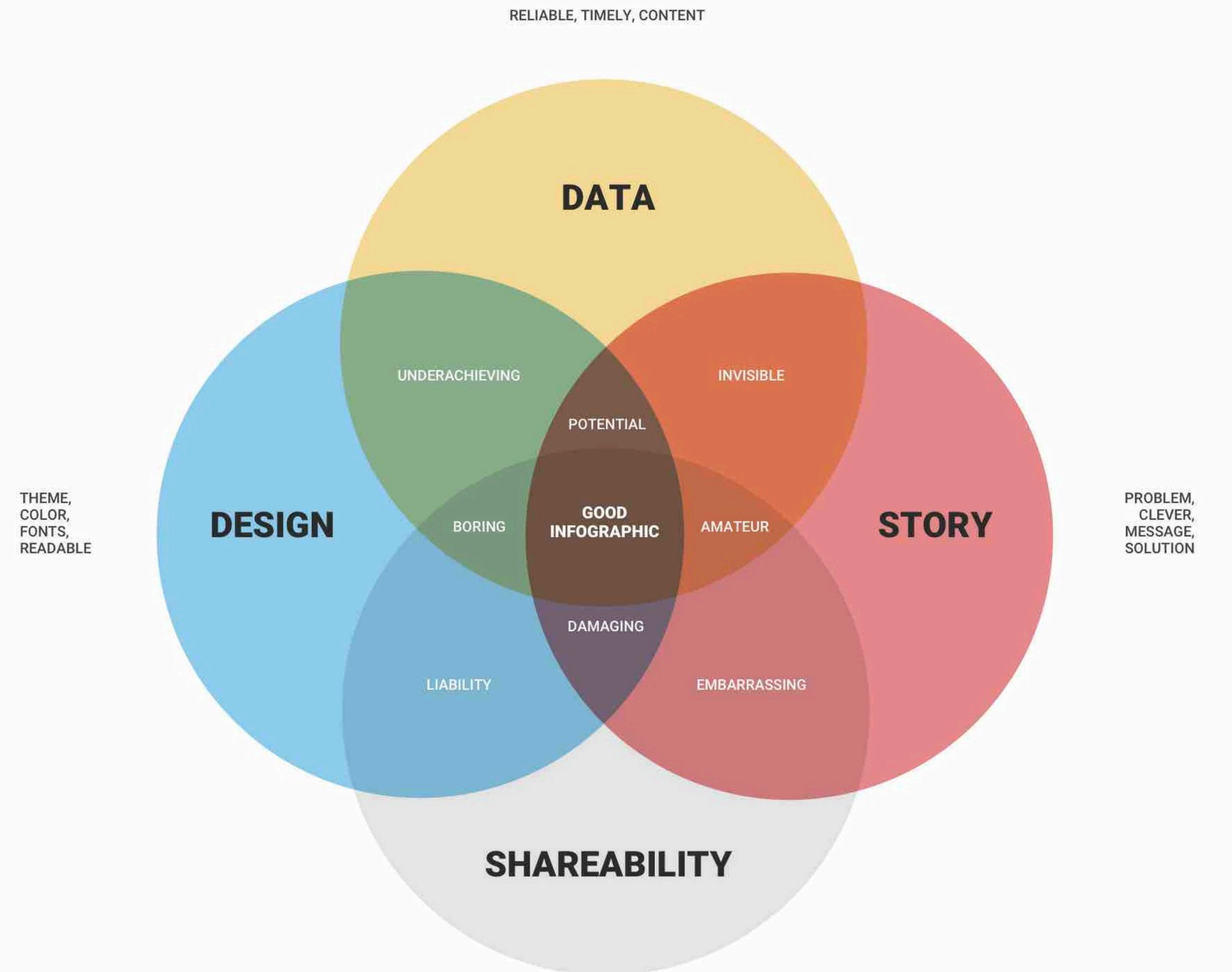

DATA PREPARATION
DATA CLEANING
TRANSFORMATION
INCONSISTENT DATATYPES
MISSPELLED ATTRIBUTES
MISSING AND DUPLICATE VALUES
talend
informatica
EXPLORATORY DATA ANALYSIS
DEFINES AND REFINES
THE SELECTION OF FEATURE
VARIABLES THAT WILL BE USED
IN THE MODEL DEVELOPMENT
DATA MODELING
KNN
NAIVE BAYES
DECISION TREE
VISUALIZATION AND COMMUNICATION
QlikView
WHAT IS
DATA SCIENCE?
DATA ACQUISITION
- WEB SERVERS
- LOGS
- DATABASES
- APIS
- ONLINE REPOSITORIES
WHY?....WHY?....WHY?....
DEPLOYS AN

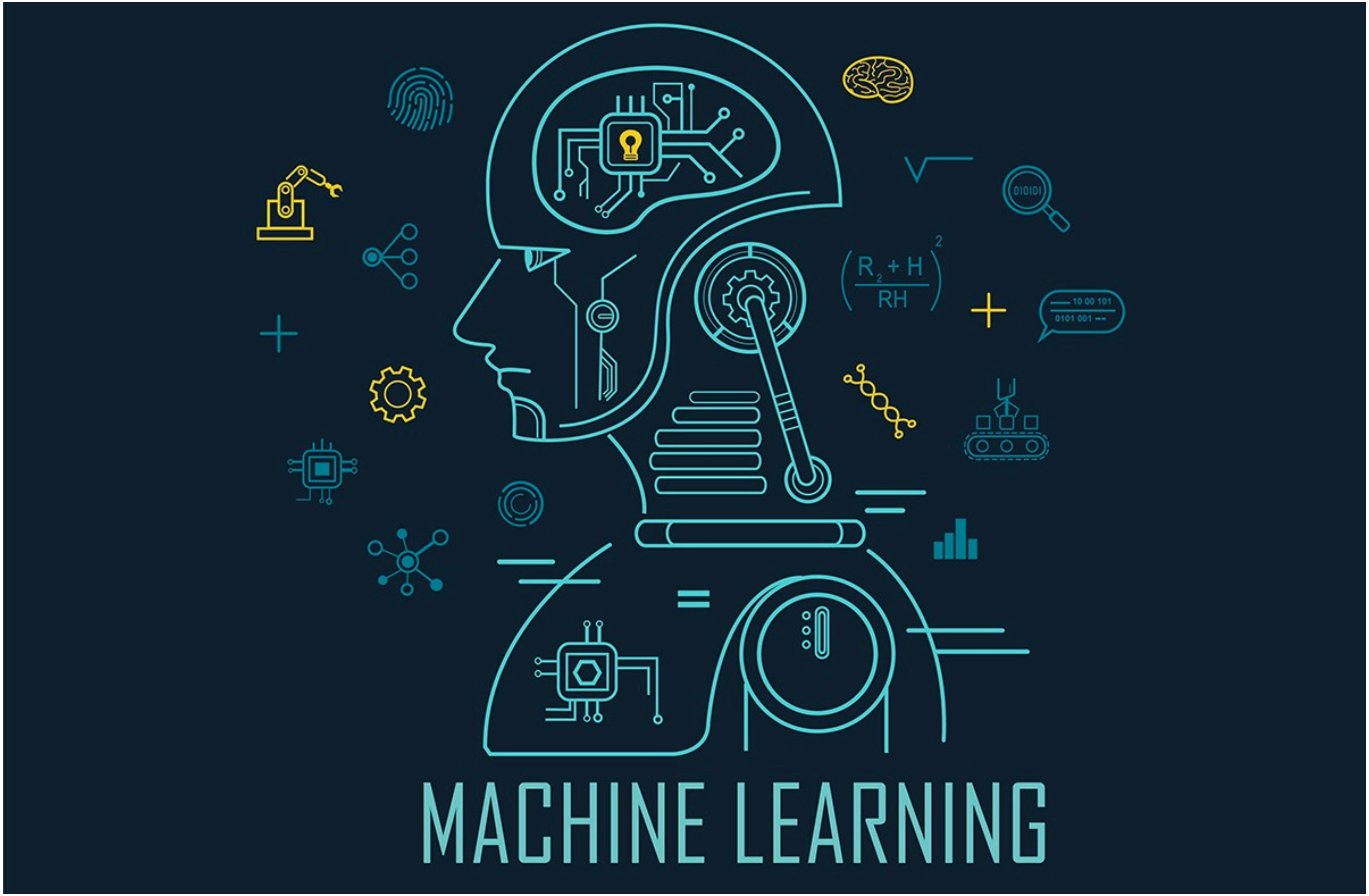

Data Science Life Cycle
Data Discovery
• Searching for different sources of data and capturing structured and unstructured data
Data Preparation
• Converting data into a common format
Mathematical Models
• Using variables and equations to establish relationships
Getting Things in Action
• Gathering information and deriving outcomes based on business requirements
Communication
• Communicating findings to stakeholders and decision makers
DATA SCIENCE


Computer Science/IT
Machine Learning
Math and Statistics
Data Science
Software Development
Traditional Research
Domains/Business Knowledge


WHAT IS
DATA VISUALIZATION?

DATA VISUALIZATION

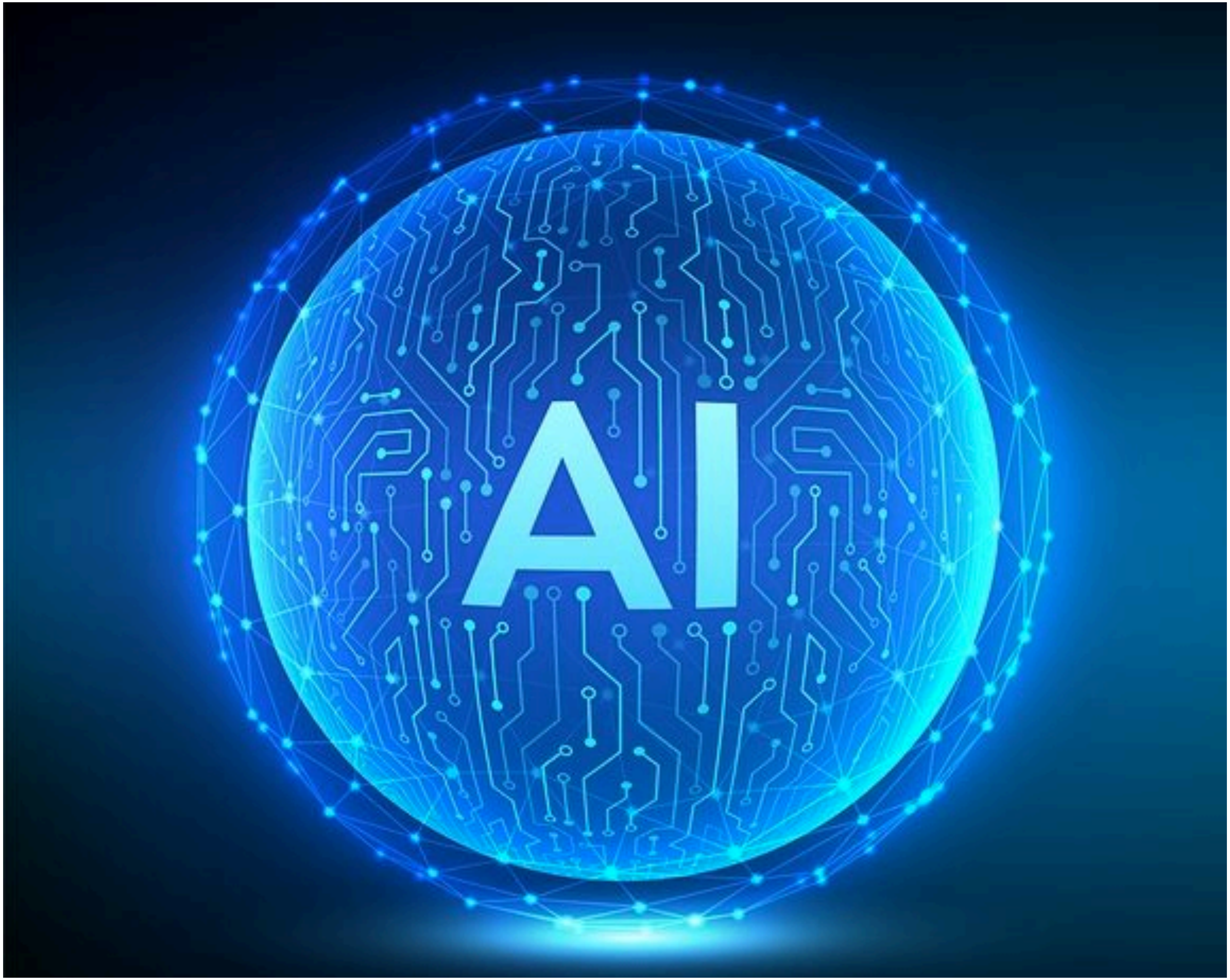


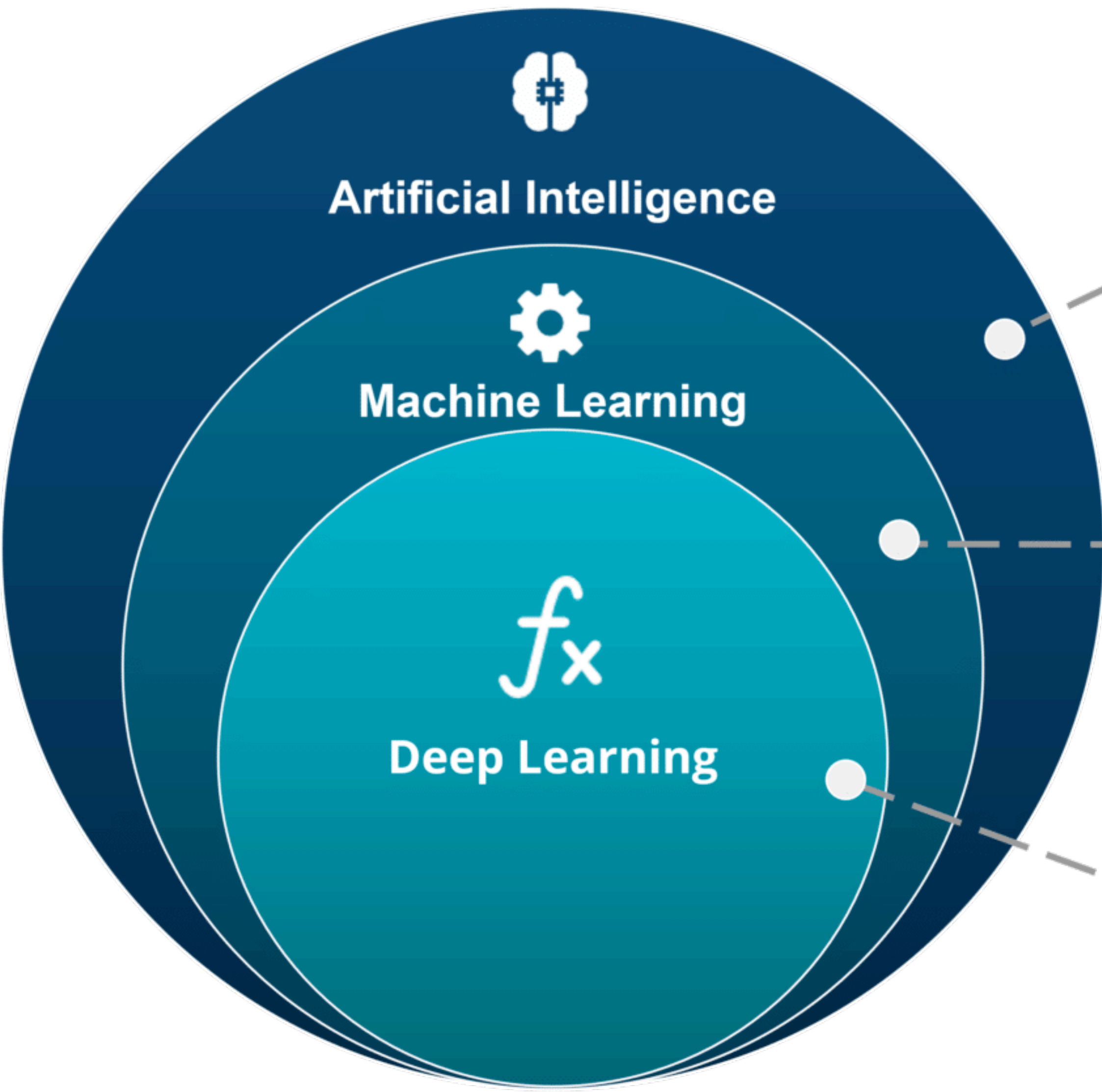

MACHINE LEARNING

WHAT IS ARTIFICIAL INTELLIGENCE?

- Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves.

- The process of learning begins with observations or data, such as examples, direct experience, or instruction, in order to look for patterns in data and make better decisions in the future based on the examples that we provide. The primary aim is to allow the computers learn automatically without human intervention or assistance and adjust actions accordingly.


AI



ARTIFICIAL INTELLIGENCE

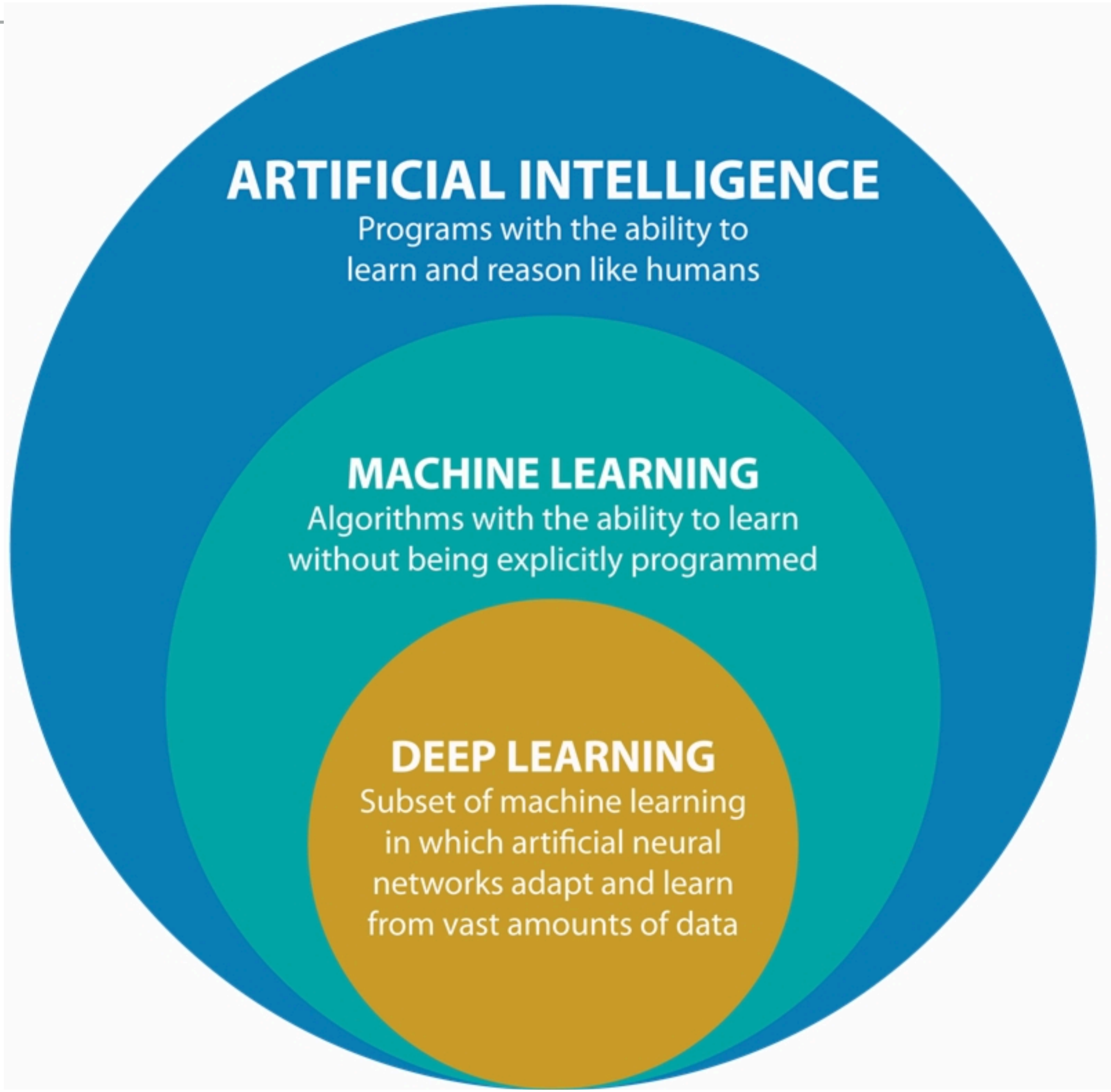
A technique which enables machines to mimic human behaviour

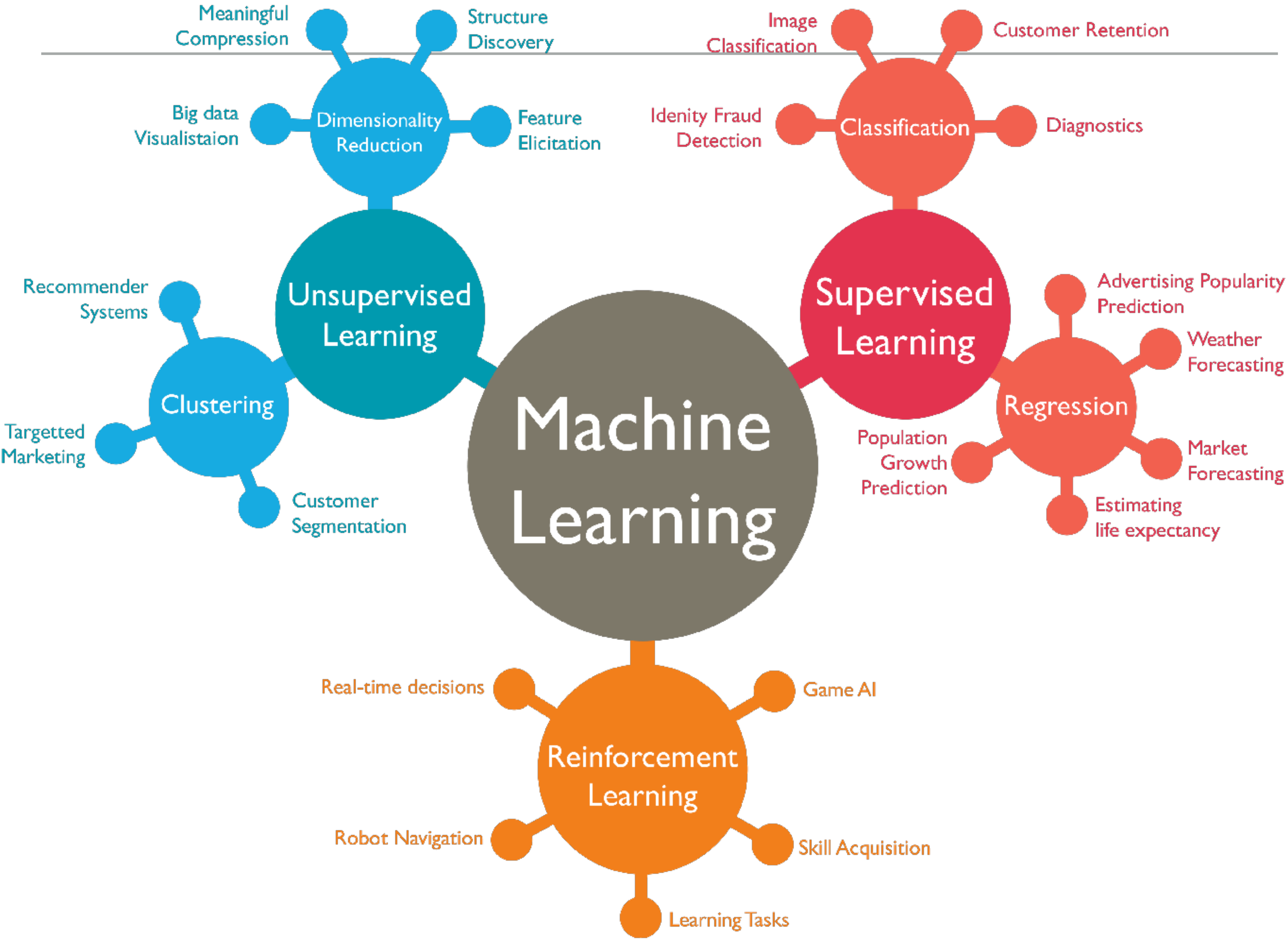
MACHINE LEARNING

Subset of AI technique which use statistical methods to enable machines to improve with experience

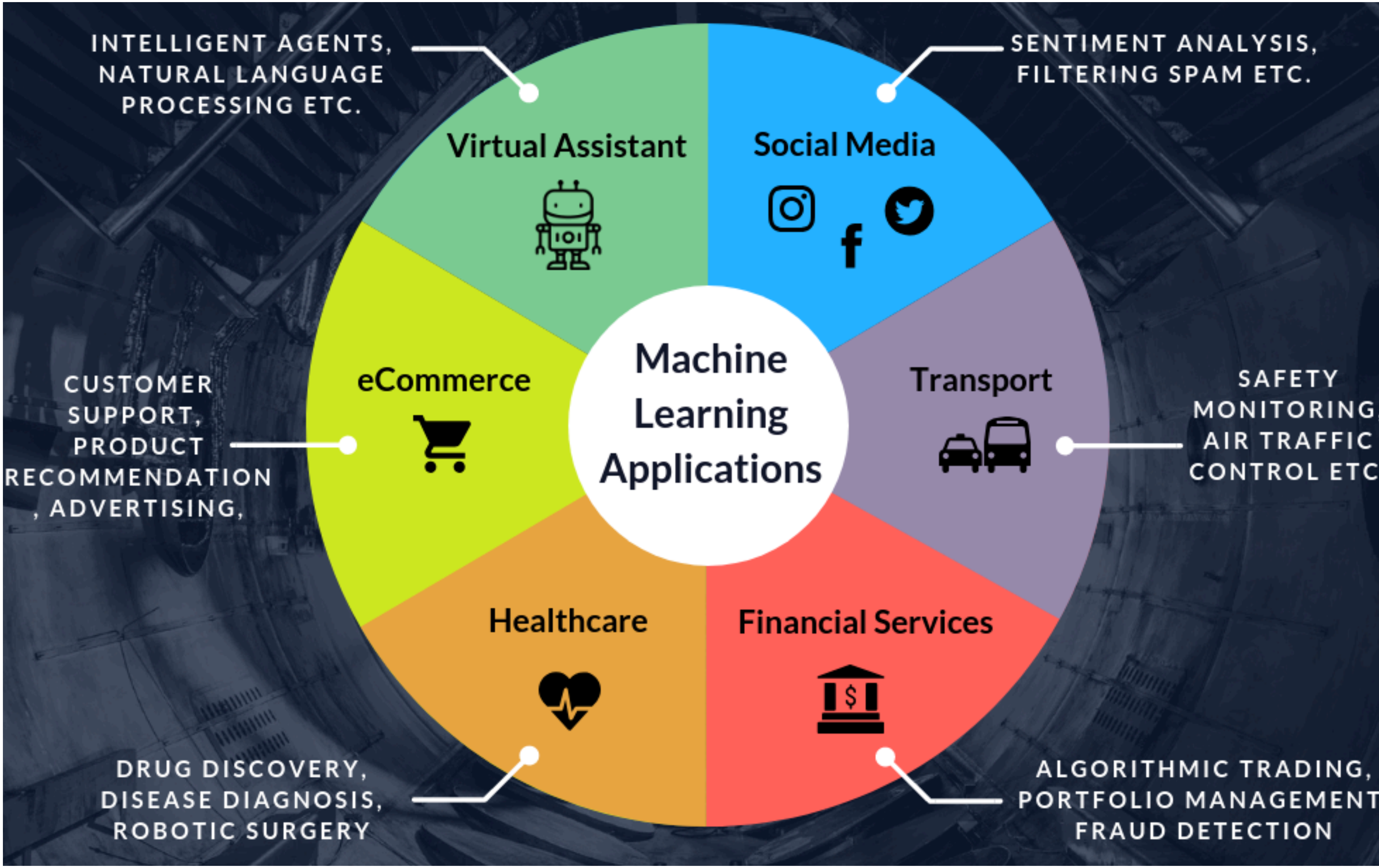
DEEP LEARNING

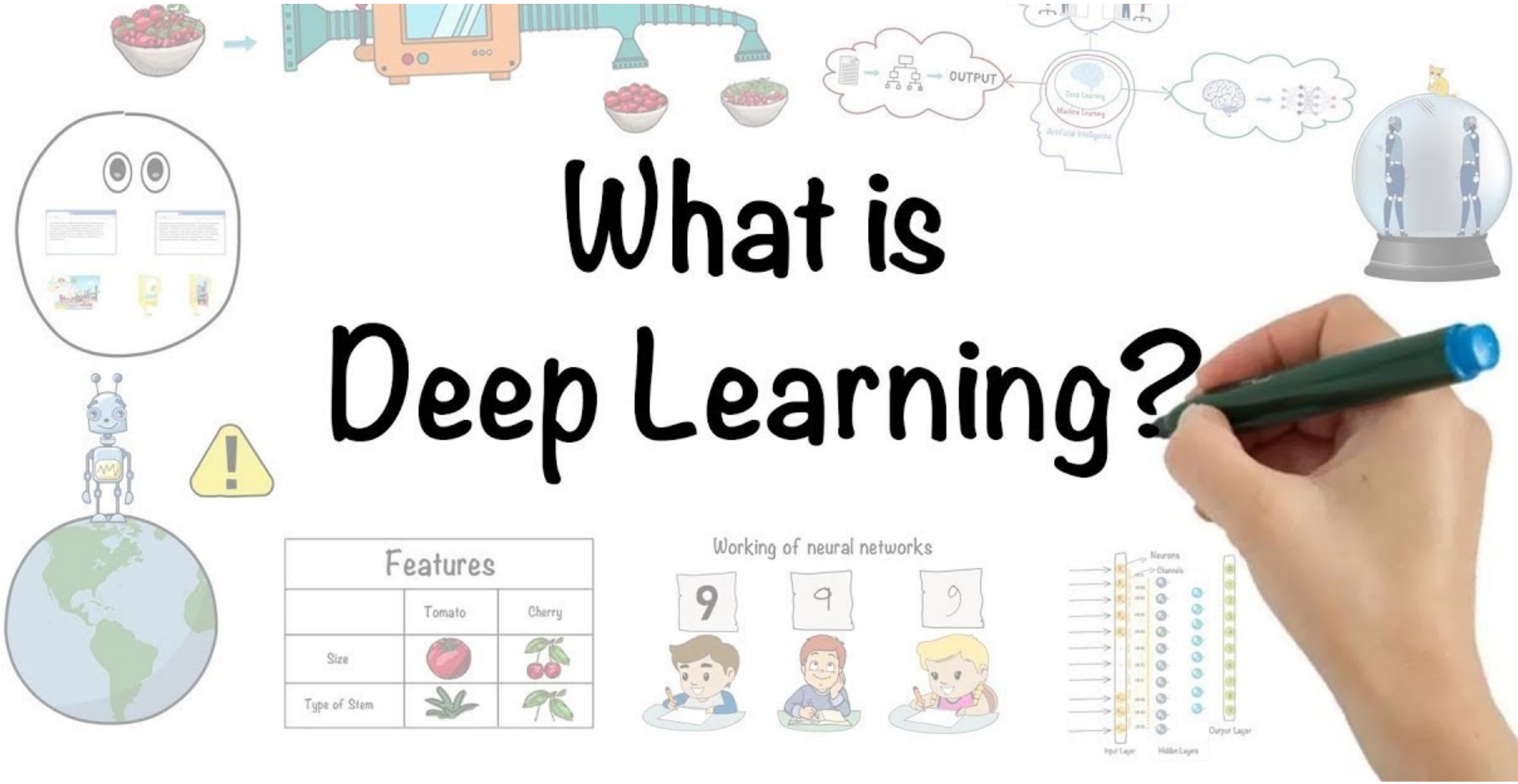
Subset of ML which make the computation of multi-layer neural network feasible

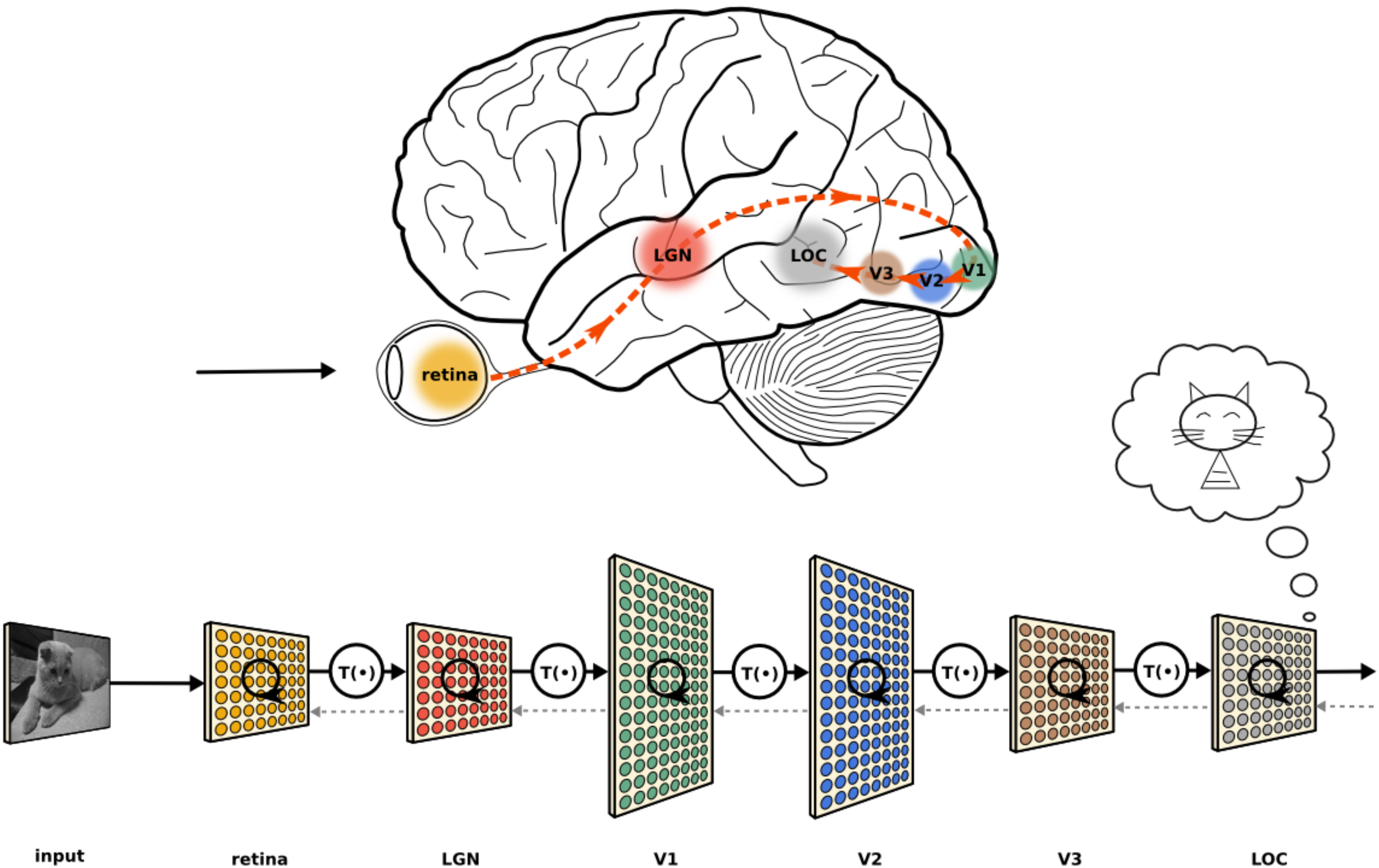

ARTIFICIAL INTELLIGENCE
Programs with the ability to learn and reason like humans
MACHINE LEARNING
Algorithms with the ability to learn without being explicitly programmed
DEEP LEARNING
Subset of machine learning in which artificial neural networks adapt and learn from vast amounts of data


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

APPLICATIONS




What is
Deep Learning?
OUTPUT
Features
Tomato
Cherry
Size
Type of Stem
Working of neural networks
9
9
9
Neurons
Channels
Input Layer
Hidden Layers
Output Layer



LGN
LOC
V3
V2
V1
retina
T(•)
input
retina
LGN
V1
V2
V3
LOC


BIG
DATA
0101
1010

Manufacturing

- Predictive maintenance or condition monitoring
- Warranty reserve estimation
- Propensity to buy
- Demand forecasting
- Process optimization
- Telematics

Retail

- Predictive inventory planning
- Recommendation engines
- Upsell and cross-channel marketing
- Market segmentation and targeting
- Customer ROI and lifetime value

Healthcare and Life Sciences

- Alerts and diagnostics from real-time patient data
- Disease identification and risk stratification
- Patient triage optimization
- Proactive health management
- Healthcare provider sentiment analysis

Travel and Hospitality

- Aircraft scheduling
- Dynamic pricing
- Social media – consumer feedback and interaction analysis
- Customer complaint resolution
- Traffic patterns and congestion management

Financial Services

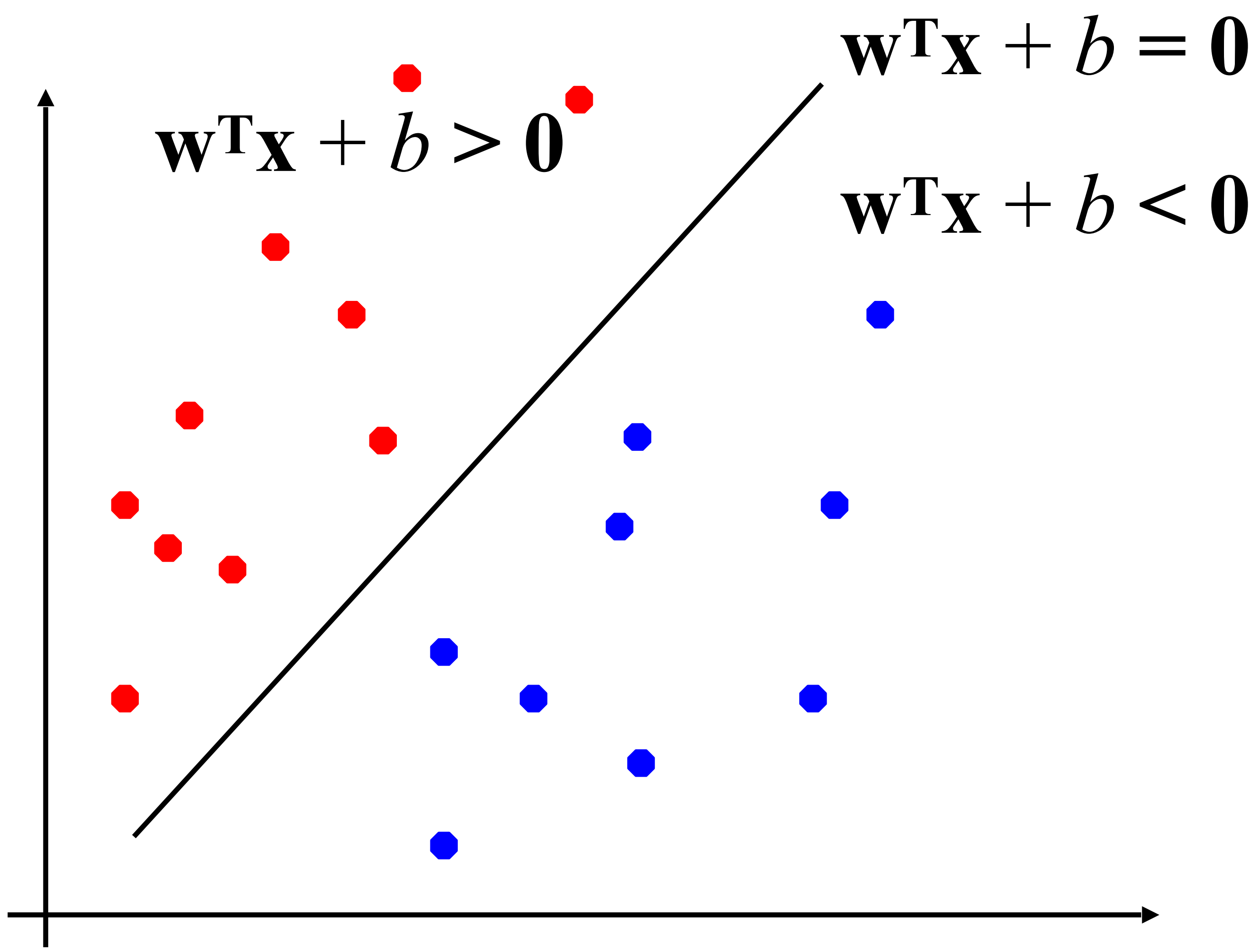
- Risk analytics and regulation
- Customer Segmentation
- Cross-selling and up-selling
- Sales and marketing campaign management
- Credit worthiness evaluation

Energy, Feedstock, and Utilities

- Power usage analytics
- Seismic data processing
- Carbon emissions and trading
- Customer-specific pricing
- Smart grid management
- Energy demand and supply optimization

Perceptron Revisited: Linear Separators

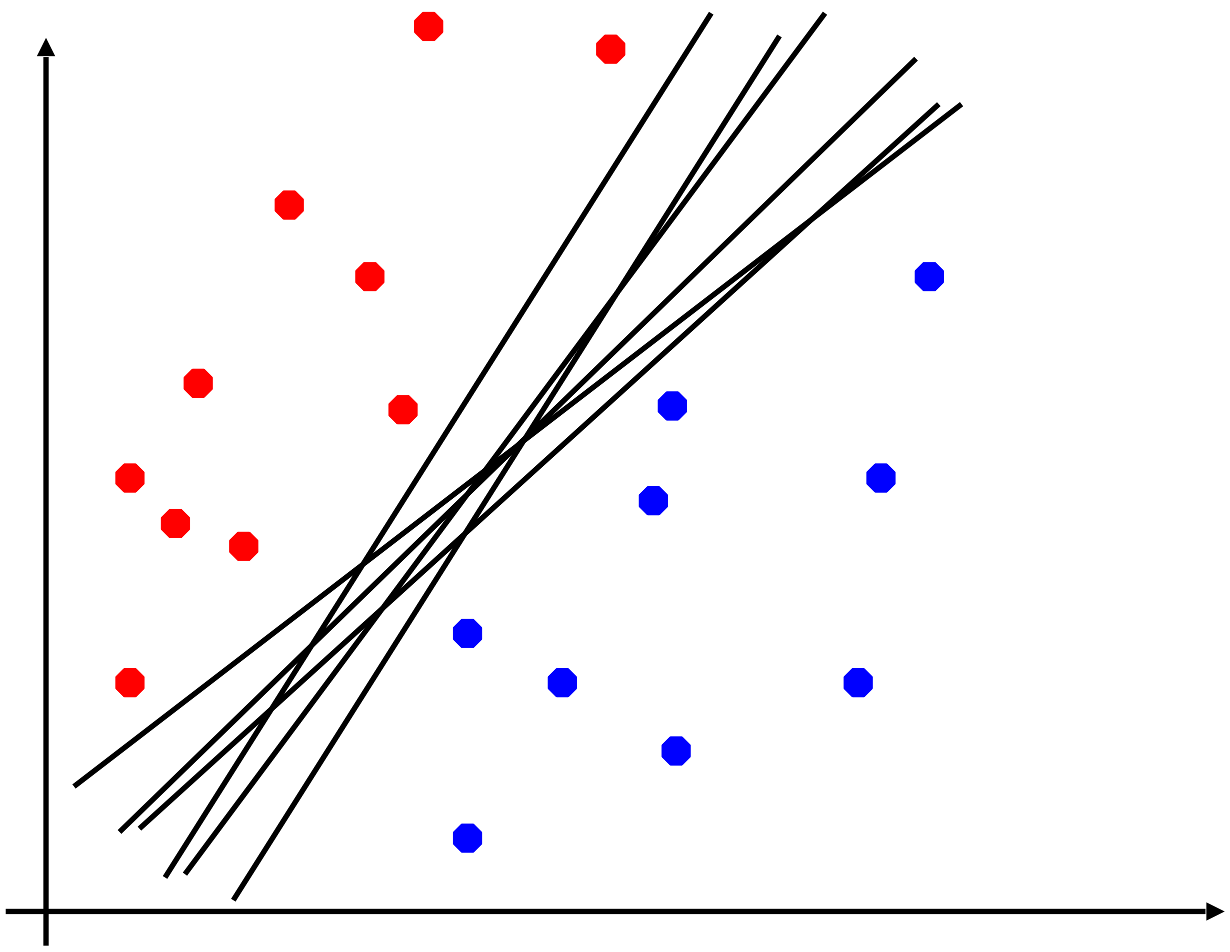
- Binary classification can be viewed as the task of separating classes in feature space:



$$f(\mathbf{x}) = \mathrm{sign}(\mathbf{w}^{\mathrm{T}}\mathbf{x} + b)$$

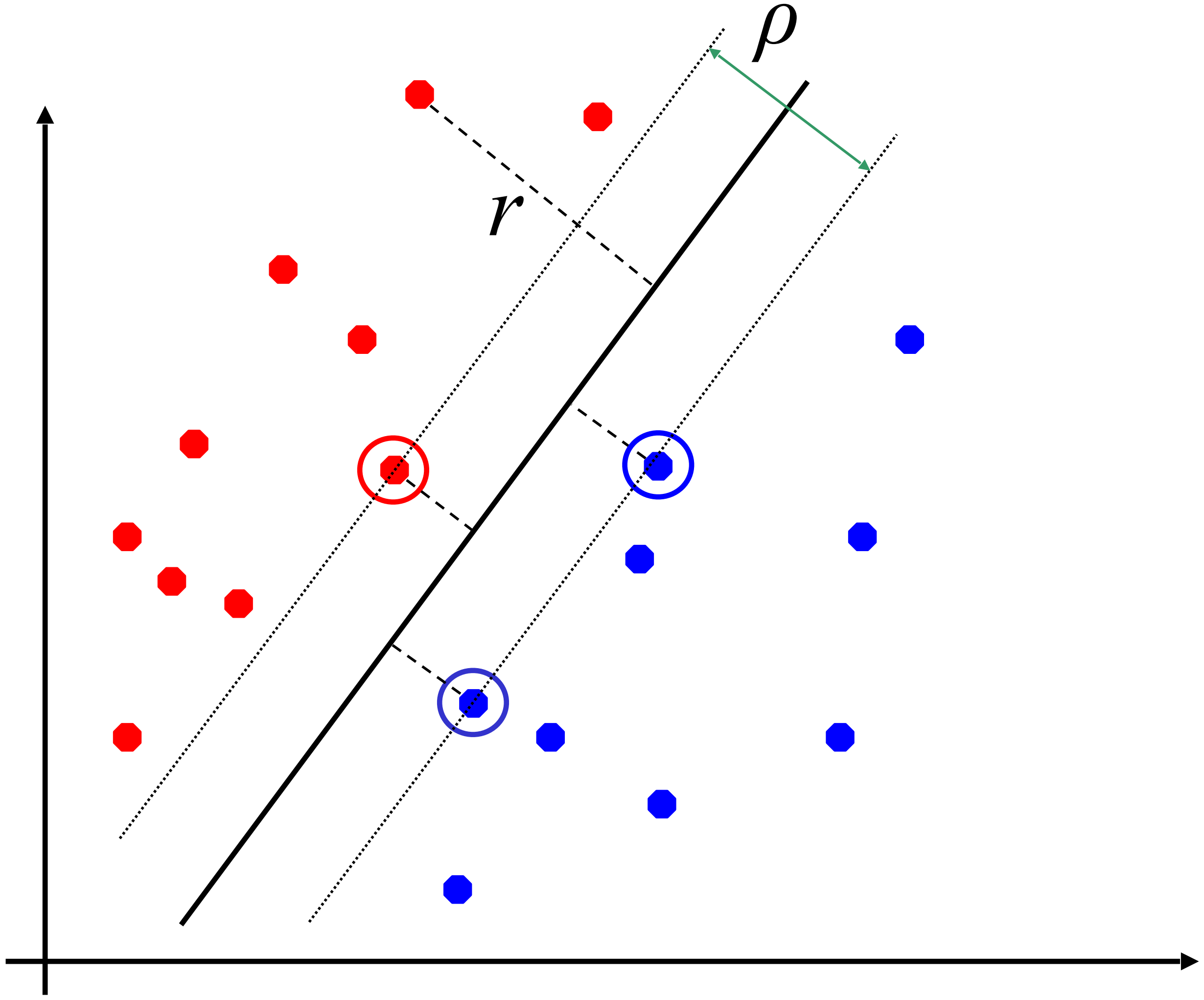
Linear Separators

- Which of the linear separators is optimal?



Classification Margin

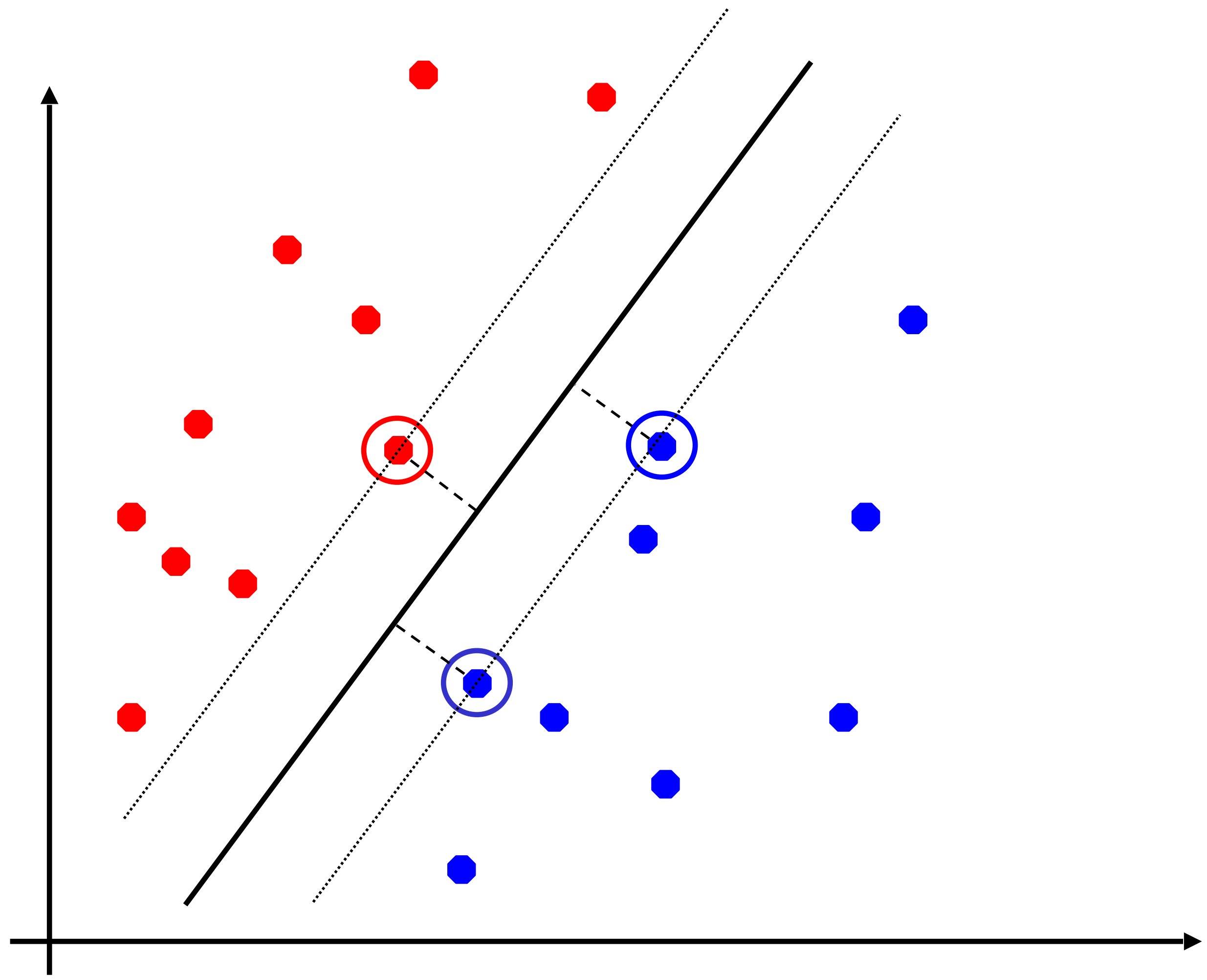
- Distance from example $\mathbf{x}_i$ to the separator is
- Examples closest to the hyperplane are ***support vectors***.
- ***Margin*** ρ of the separator is the distance between support vectors.



$$r = \frac{\mathbf{w}^T \mathbf{x}_i + b}{\|\mathbf{w}\|}$$

Maximum Margin Classification

- Maximizing the margin is good according to intuition and PAC theory.
- Implies that only support vectors matter; other training examples are ignorable.



Linear SVM Mathematically

- Let training set $\{(\mathbf{x}_i, y_i)\}_{i=1..n}$, $\mathbf{x}_i \in \mathbf{R}^d$, $y_i \in \{-1, 1\}$ be separated by a hyperplane with margin ρ. Then for each training example $(\mathbf{x}_i, y_i)$:

$$\begin{array}{ll} \mathbf{w}^{\mathbf{T}}\mathbf{x}_i + b \leq -\rho/2 & \text{if } y_i = -1 \\ \mathbf{w}^{\mathbf{T}}\mathbf{x}_i + b \geq \rho/2 & \text{if } y_i = 1 \end{array} \quad \Leftrightarrow \quad y_i(\mathbf{w}^{\mathbf{T}}\mathbf{x}_i + b) \geq \rho/2$$

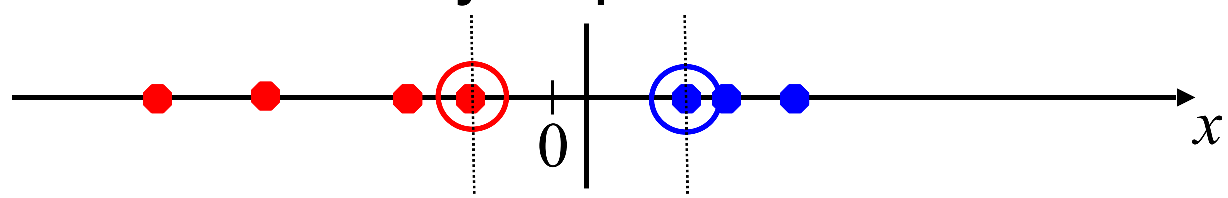
- For every support vector $\mathbf{x}_s$ the above inequality is an equality. After rescaling $\mathbf{w}$ and b by $\rho/2$ in the equality, we obtain that distance between each $\mathbf{x}_s$ and the hyperplane is $r = \frac{y_s(\mathbf{w}^T\mathbf{x}_s + b)}{\|\mathbf{w}\|} = \frac{1}{\|\mathbf{w}\|}$

- Then the margin can be expressed through (rescaled) $\mathbf{w}$ and b as:

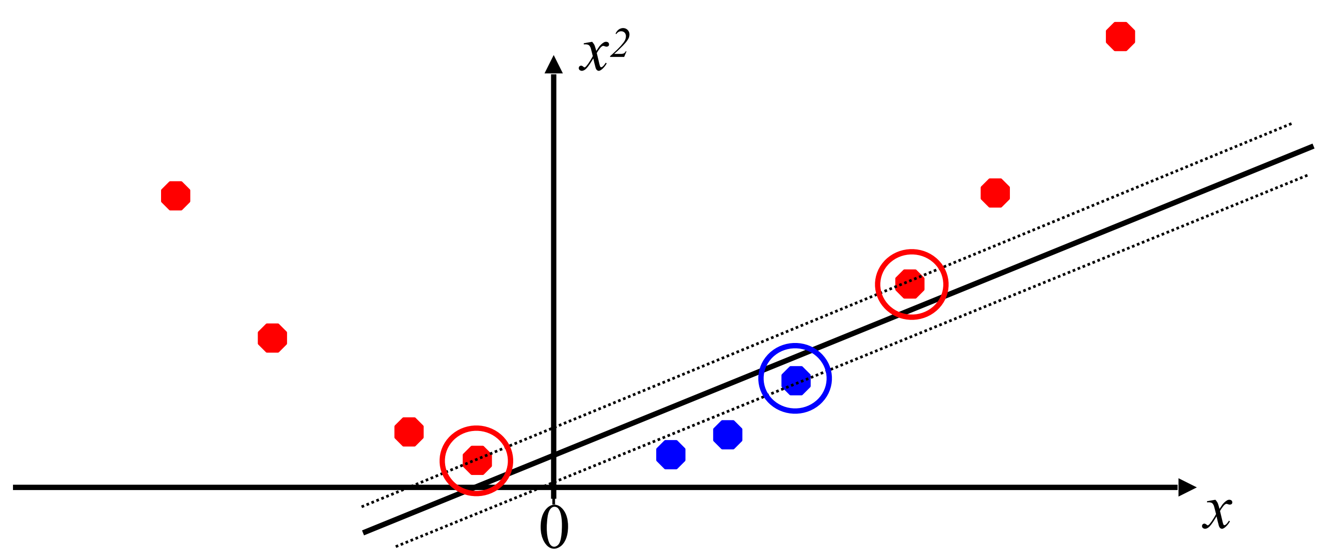
$$\rho = 2r = \frac{2}{\|\mathbf{w}\|}$$

Non-linear SVMs

- Datasets that are linearly separable with some noise work out great:



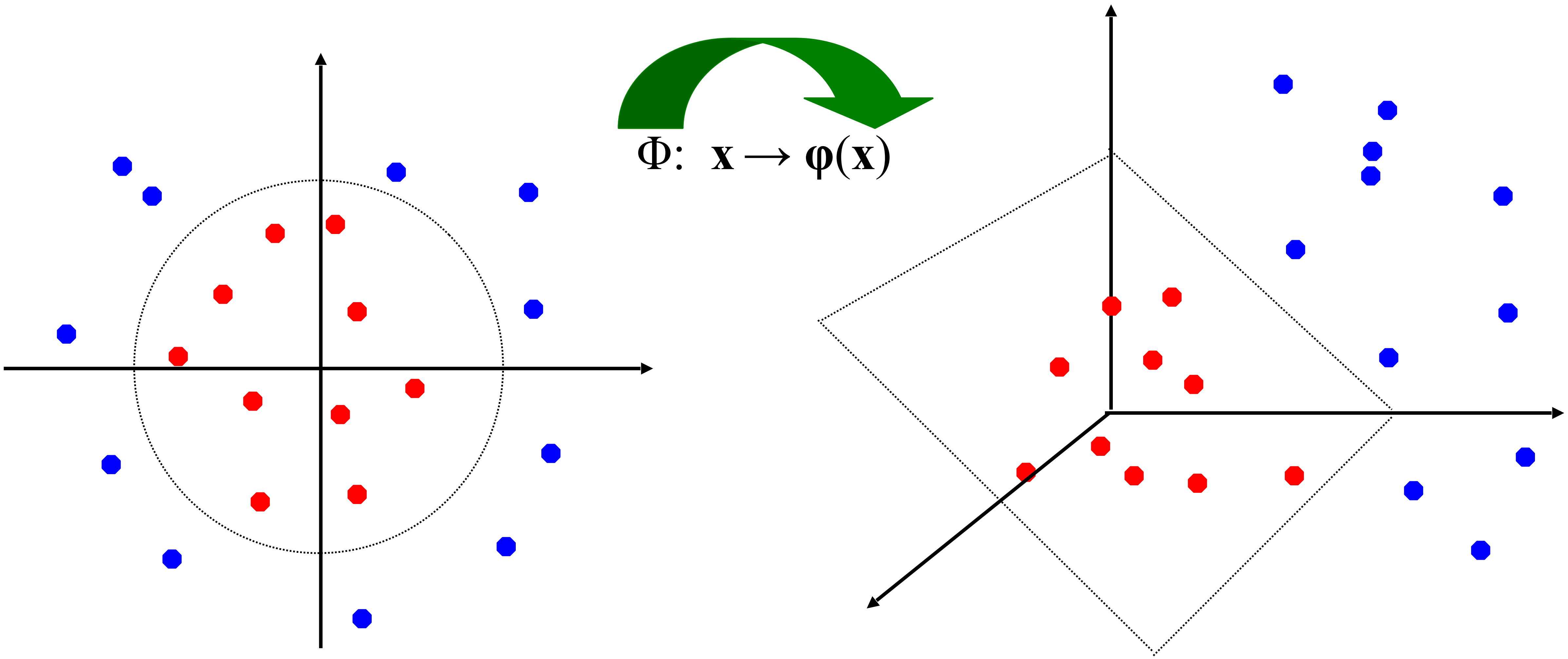
- But what are we going to do if the dataset is just too hard?



- How about… mapping data to a higher-dimensional space:

Non-linear SVMs: Feature spaces

- General idea: the original feature space can always be mapped to some higher-dimensional feature space where the training set is separable:



Examples of Kernel Functions

- Linear: $K(\mathbf{x}_i,\mathbf{x}_j)= \mathbf{x}_i^{\mathbf{T}}\mathbf{x}_j$
  - Mapping Φ: $\mathbf{x} \to \varphi(\mathbf{x})$, where $\varphi(\mathbf{x})$ is $\mathbf{x}$ itself

- Polynomial of power p: $K(\mathbf{x}_i,\mathbf{x}_j)= (1+ \mathbf{x}_i^{\mathbf{T}}\mathbf{x}_j)^p$
  - Mapping Φ: $\mathbf{x} \to \varphi(\mathbf{x})$, where $\varphi(\mathbf{x})$ has $\binom{d+p}{p}$ dimensions

- Gaussian (radial-basis function): $K(\mathbf{x}_i,\mathbf{x}_j) = e^{-\frac{\|\mathbf{x}_i-\mathbf{x}_j\|^2}{2\sigma^2}}$
  - Mapping Φ: $\mathbf{x} \to \varphi(\mathbf{x})$, where $\varphi(\mathbf{x})$ is *infinite-dimensional*: every point is mapped to *a function* (a Gaussian); combination of functions for support vectors is the separator.

- Higher-dimensional space still has *intrinsic* dimensionality d (the mapping is not *onto*), but linear separators in it correspond to *non-linear* separators in original space.

SVM applications

- SVMs were originally proposed by Boser, Guyon and Vapnik in 1992 and gained increasing popularity in late 1990s.

- SVMs are currently among the best performers for a number of classification tasks ranging from text to genomic data.

- SVMs can be applied to complex data types beyond feature vectors (e.g. graphs, sequences, relational data) by designing kernel functions for such data.

- SVM techniques have been extended to a number of tasks such as regression [Vapnik *et al.* '97], principal component analysis [Schölkopf *et al.* '99], etc.

- Most popular optimization algorithms for SVMs use *decomposition* to hill-climb over a subset of α_i's at a time, e.g. SMO [Platt '99] and [Joachims '99]

- Tuning SVMs remains a black art: selecting a specific kernel and parameters is usually done in a try-and-see manner.


100%
EVOLUTION
78%
STATUS
ARTIFICIAL INTELLIGENCE & MACHINE LEARNING TRENDS
2020

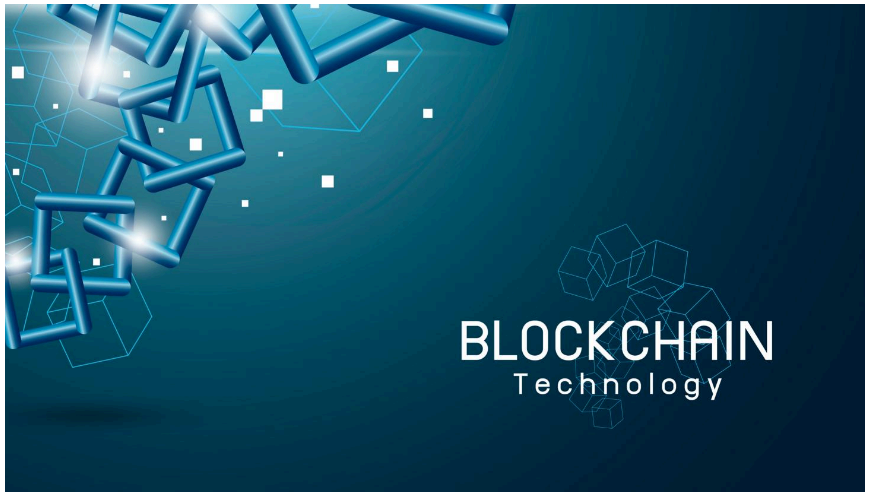
AI AND ML TRENDS TO LOOK FOR

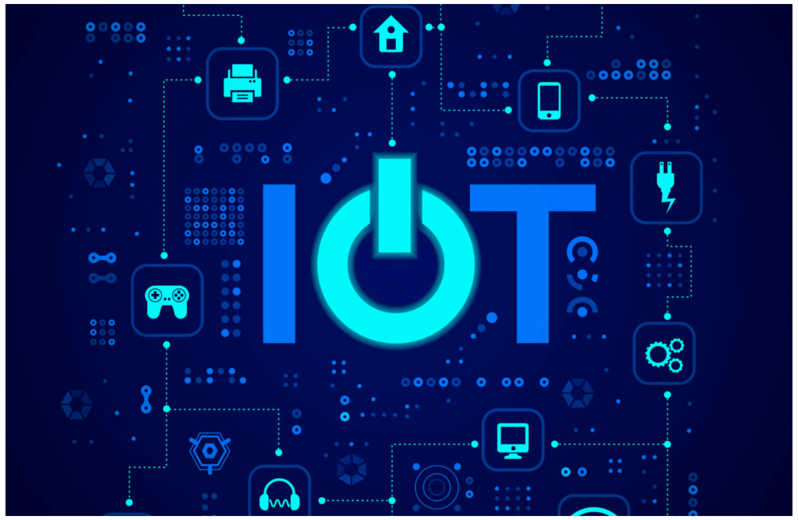
- According to a recent report by IDC, global spending for AI will reach $97.9 billion in 2023, up almost 3 times from $37.5 billion in 2019

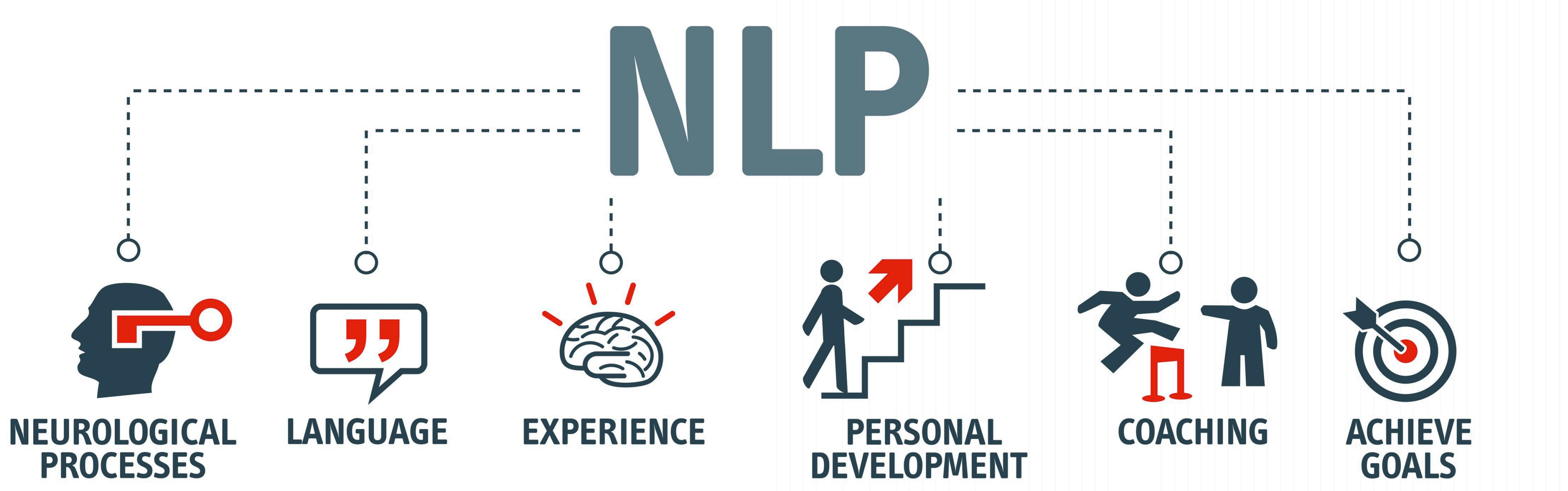
- By 2022, more than 3/4 of companies will use deep Neural Networks, rather than the classical machine learning

- The estimated value of the US Deep Learning market will be worth $935 Million by 2025, emphasising the popularity of ML.

TRENDS TO LOOK FOR








INDIA'S TOP 15 EMERGING JOBS FOR 2020, ACCORDING TO LINKEDIN
1. Blockchain Developer
2. Artificial Intelligence Specialist
3. JavaScript Developer
4. Robotic Process Automation Consultant
5. Back-end Developer
6. Growth Manager
7. Site Reliability Engineer
8. Customer Success Specialist
9. Full Stack Engineer
10. Robotics Engineer (Software)
11. Cybersecurity Specialist
12. Python Developer
13. Digital Marketing Specialist
14. Front-end Engineer
15. Lead Generation Specialist

ANY QUERIES???