Building Informatics and Smart Building

* Informatics [Definition]
  + Study of computational systems, especially for data storage and retrieval （特别对于数据存储和检索系统）
  + US: about Data science
  + Europe: information technology / computer science [central: transformation of information]
  + Other: health / medical domain try to define Informatics
    - Cross- training field of study
    - **Informatics** is a fundamental theorem / technique improves any task
* Building informatics [Definition]
  + leveraging information technology, computer science, and related data science technologies → to improve building operation efficiency & user comfort in buildings.
  + No clear consensus & definition
  + Most: dealing with collecting, managing, and analysing building data → improve building efficiency & user comfort.

* Smart Building [Definition]
  + **digitally connected structures**: combine optimized building & operational automation with intelligent space management → user experience, increase productivity, reduce costs, mitigate physical & cybersecurity risks
  + **converges** various **building-wide systems** (E.g. HVAC, lighting, alarms, electrical, fire safety, plumbing and drainage, security&access control) into a single IT managed network infrastructure.
  + **integrate & account** for intelligence, enterprise, control, and materials and construction as an entire building system
  + Core: **adaptability, not reactivity （核心为适应性，而不是反应性）**
  + → to meet the drivers for building progression: 1. energy and efficiency, 2. longevity,（寿命） 3. and comfort and satisfaction.
  + The increased amount of information → wider range of sources will allow these systems to become adaptable, & enable a Smart Building to prepare itself for context and change over all timescales.

* Smart Building V.S. Intelligent Building
  + Often used interchangeably
  + "Intelligent Building" was popular in 1980's/90's
  + Intelligent Building [Definition]

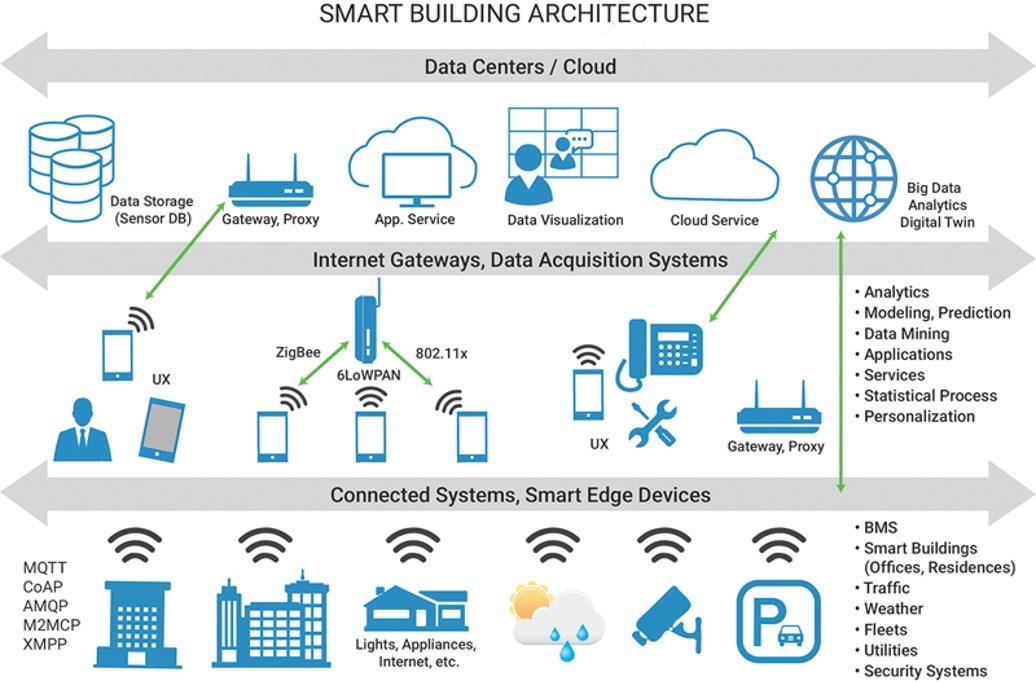
**creates an environment** that **maximizes** the efficiency of the occupants of the building & allowing effective management of resources with **minimum life-time costs**

* Similar:
  + Focus on creating **integrated system** interrelates various subsystems → single control frame work
  + Use information technology and data generated by building
* Difference:
  + Smart Building also **includes wider integration** with utilities and city infrastructure to realize smart city （Smart Building 有公共设施和城市基建，实现 Smart city）
  + Smart Building includes use of emerging **machine learning and AI** for advanced control and diagnostics （Smart Building有AI和ML）
  + Smart Building considers interaction of users with building and surrounding environment → improve comfort of users apart from building operations.
  + Smart Building encompasses the use of emerging IoT, wireless communication, and other related technologies (technologies were not developed when Intelligent building concept was proposed) （Smart Building 会用一些 Intelligent Building 时期未开发的技术）

* Smart Building - Characteristic （特征）
  + Climate Response: buildings to adapt to external climate conditions
  + Grid Response: buildings to adapt the information coming from grid （电网） → **maximize** the energy/economic efficiency at district/city scale
  + User Response: enable a real-time interaction between users and technologies → optimized performance & user comfort
  + Monitoring and Supervision: enable real-time monitoring of building operations and users’ behaviour.

* Smart Building - Benefits
  + **Lower Operational costs** – use emerge tech [Building System] → save time, space, money
  + **Lower Energy cost** – emission saving → achieve *net zero energy buildings*
  + **Greater Flexibility** – **dynamically adapt** the buildings according to external and internal changes by real-time monitoring using sensors
  + **Improved user comfort, efficiency, and wellbeing** – improve health, comfort, efficiency of users by adjusting the surrounding environment & providing collaboration opportunities （协作机会）

* **Smart Building - Framework**



* The bottom layer: systems and devices: HVAC, lighting, security and access control, water, elevator, drainage, etc.
* The middle layer: **gateways** to acquire & transmit data from different building systems
* The upper layer: services supported by cloud including storage, visualization, and data analytics.
* E.g. Deloitte: 1. Physical 2. Digital 3. Use case
* A smart building framework can be analysed from three aspects:
  + Different systems/devices within the smart buildings
  + Different technologies used within the smart buildings
  + Different use cases/features supported within the smart buildings

* Smart Building - Underlying Systems
  + Building Automation Systems (BAS) / BMS (Management): computerized, integrated, and intelligent network of hardware and software installed in a building → monitors and controls the building’s mechanical and electrical equipment.
    - Integrates advanced technologies: Information Technology (IT), Internet of Things (IoT), Big Data, Artificial Intelligence (AI), modern control, etc.
    - Major Components:
      * Network (wired / wireless)
      * Controllers (programmable)
      * Actuator （执行器）
      * Sensors
      * Computers / Servers
      * Software
    - Typical Functions:
      * Air conditioning and ventilation system monitoring and control

(on/off control of major equipment, control temperature and humidity, thermal comfort, etc.)

* Lighting control

(dimming （调光）, special lighting effects, motion sensor, daylight, etc.)

* Fire detection and alarm

(smoke detector, heat detector, fire alarm, etc.)

* Security and access control

(monitoring camera, password, key, finger print, iris, bio-features, etc.)

* Lift control

(destination control, up peak, down peak, etc. ) （目的地控制、上峰、下峰等）

* Utility management interface （设施管理界面）

(historical data analysis, cost management, etc.)

* Smart Building - Certification （认证）
  + One popular certification - provided by ***Smart Building Collective***
    - Smart Building Certification → to award buildings that *have the IT infrastructure* in place to measure a building and behavior.
    - Includes measures: optimization & innovation in buildings.
    - Imp. **Considers how the building data is being utilized.**
    - **4 - level** of certifications:
      * Bronze
      * Silver
      * Gold
      * Platinum
  + Certification is provided for both build phase and operation phase of buildings （在建造和运营阶段认证）
    - In Design | Core & Shell: Certify the technical infrastructure in place → give a clear view on what is possible with the building
    - In Operation:
      * Goes **beyond** only the technical infrastructure and looks（超越技术基础设施）
      * How the technical infrastructure is being used to optimize, improve, and manage the efficiency of the building
  + 6 - themes （六个主题）
    - Building Usage: effective usage（有效使用）of building for users.
    - Building Performance: optimization of investment, improved efficiencies, consumption, minimization of environmental impact.
    - Building Environment: measures the health of the physical environment
    - Health, Safety, and Security: if compliant with health and safety standards.
    - User Behavior & Collaboration: explores how the building empowers user behavior and collaboration, supporting the maximization of value and purpose of the space（空间价值最大化）.
    - Interactive Design & Connectivity: measures how well integrated with one another and the building
  + Other Certifications
    - Also provided by some government organizations and other third-party organizations
    - Government level certification E.g. *Smart Readiness Indicator* (SRI), [European Union (EU)]
    - SRI rates 3 key functionalities:
      * optimise energy efficiency and overall in-use performance
      * adapt their operation to the needs of the occupant
      * adapt to signals from the grid (for example energy flexibility)
    - **Beam Plus [HK]**

comprehensive set of **performance criteria** for a wide range of sustainability issues relating to the planning, design, construction, commissioning, fitting out, management, operation and maintenance of a building.

* Categories: new buildings (NB), existing buildings (EB), data centers (DC), Interiors (BI), Neighbourhood (ND), and existing schools (ES).
* Smart Building Acceleration Act [USE]
* Third-party organisation E.g. ***SmartScore provided by WiredScore company***
* LEED, BREEAM etc.

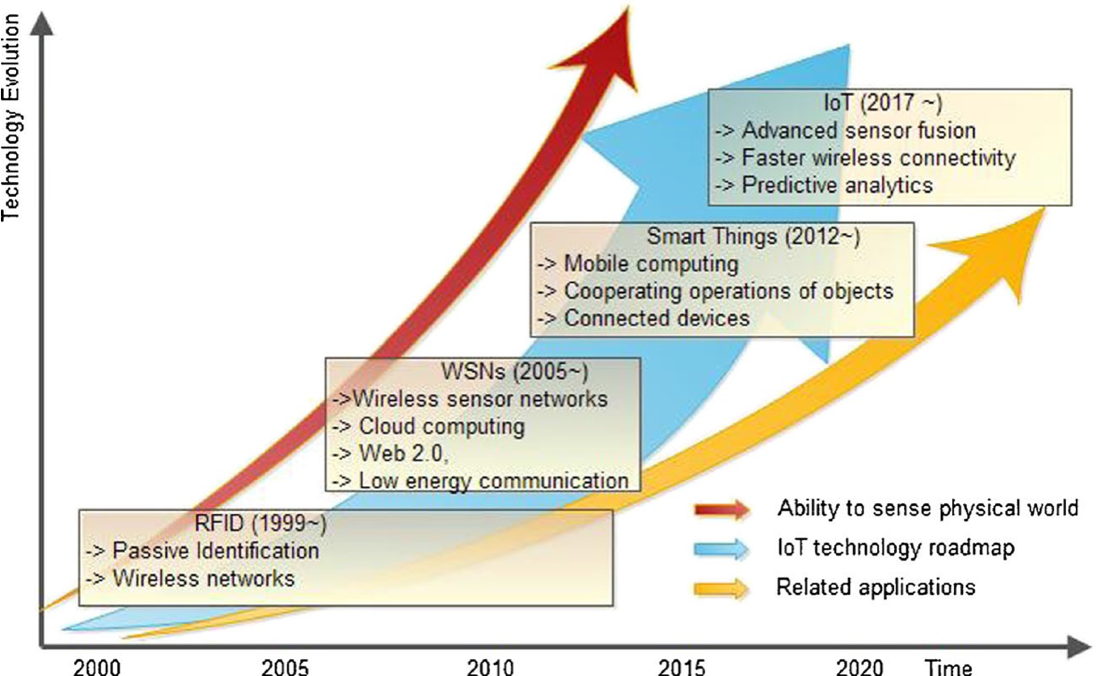
* Smart Building - Challenges
  + Limited interoperability due to heterogeneity of sensors, devices, and data （传感器，设备和数据的异构性→互操作性有限）
  + Complex dependency of different systems
  + Large scale of devices and data
  + Limited integration across systems and technologies （跨系统和技术的集成有限）
  + Lack of standardized technologies and solutions
  + Security & Privacy concerns
* Concept:
  + **Integration** can be defined as the ability to group intelligent, standalone systems into a cohesive building management system.
  + **Interoperability** is defined as the ability to link multiple standalone building services systems from variety of manufacturers (such as those for air conditioning, fire safety, lighting and access control) into a comprehensive system.

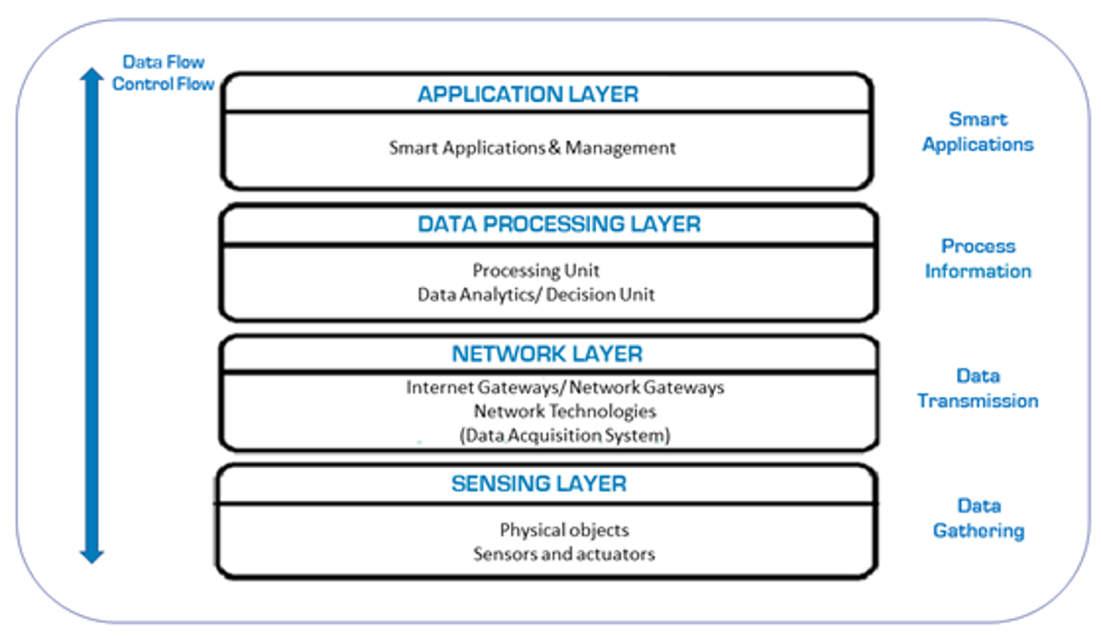
* Smart Building - Future
  + Will plug into **“smart” urban infrastructure** & cater to an increasingly dense and technology-savvy urban population （应合日益密集和精通技术的城市人口）
  + Skyscrapers will be continuously adapting & made from materials that from filtering air to producing energy

Internet of Things (IoT) in Smart Building

* Brief History of IoT
  + Kevin Ashton, co-founder of Auto-ID lab at MIT, in a presentation given to Proctor & Gamble in **1999** [RFID at first]

* IoT Definition
  + Physical objects (or groups of such objects) with sensors, processing ability, software and other technologies → connect and exchange data with other devices and systems → over the Internet or other communications networks.
  + Connecting any device (so long as it has an on/off switch) → Internet and to other connected devices.
  + The IoT is a giant network of connected things and people – all of which collect and share data about the way they are used and about the environment around them.



* IoT Architecture
* 
  + Sensing Layer (/Perception Layer/Device Layer) - [collecting, accepting, and processing data over the network]
    - Include:
      * IoT devices - sensors & actuators
    - Do:
      * Sensors: sense surrounding & transmit data
      * Actuators: receive the control signal & perform action
    - Each IoT device is also associated with an identity → enable **monitoring & management**.
    - IoT devices at *sensing layer* can have different form factors (ranging nanoscale low-power devices to autonomous vehicles and satellites) （在Sensing Layer的IoT设备可以有**不同的外形尺寸**）
    - In recent years due to improvement in technology → IoT devices have begun to **have computation and storage resources** → enabling local data processing
  + Network Layer (/Communication Layer) - [transmitting and processing the data collected by sensor devices]
    - Include:
      * Gateway
      * Routers

→ **transmit & process** data collected by sensors

* Data acquiring systems (DAS):

performing data aggregation and conversion functions (**collecting and aggregating** data from sensors, then **converting analog data to digital data**, etc.)

* Data Processing Layer (/Middleware Layer) - [**processing, analyzing, storing data**]
  + Processing: done by cloud / edge devices
  + Include:
    - Different types of *data processing functionalities* （数据处理功能）
      * Data **accumulation**
      * Data **abstraction**
      * Etc.
    - Privacy & Security functions etc.
* Application Layer
  + cloud / data centers, data is managed & is used by end-user applications from various sectors
  + Include:
    - **Different interface**
      * Graphical user interface (GUIs)
      * Command-line user interfaces (GUIs)

→ enable user interaction and provide application services to users

* Management services:
  + Remote device monitoring & management
  + Security control…

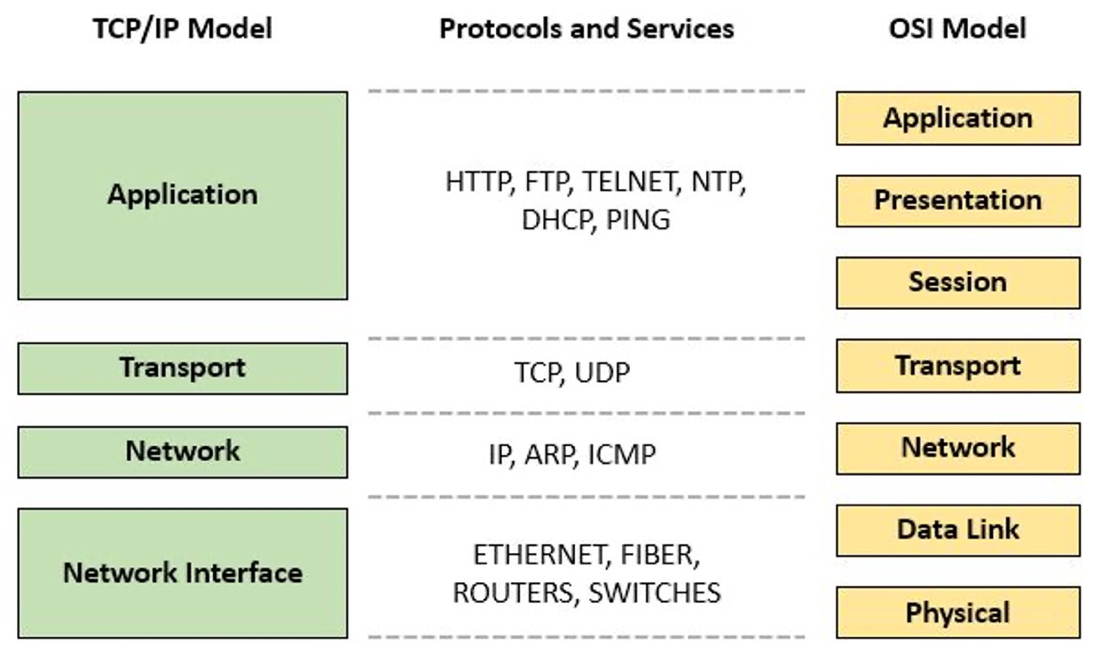
* IoT Benefits
  + Efficient resource monitoring and utilization
  + Minimize human effort by enabling automation
  + Save time and money
  + Enhance real-time data collection
  + Improve security and efficiency
  + Enable new analytic insights by connecting different systems
  + Add value by creating new business opportunities

* IoT Challenges
  + Heterogeneity （异质性）
  + Scalability （可扩展性）
  + Integration of devices from different systems and vendors （集成来自不同系统/供应商设备）
  + Dynamic and intermittent（间歇性） wireless networks
  + Energy concerns - due to low-cost and low-power devices
  + Reliability concerns due to device failures
  + Remote device management and diagnosis
  + Lack of standardized protocols （缺少标准化协议）
  + Privacy and security concerns

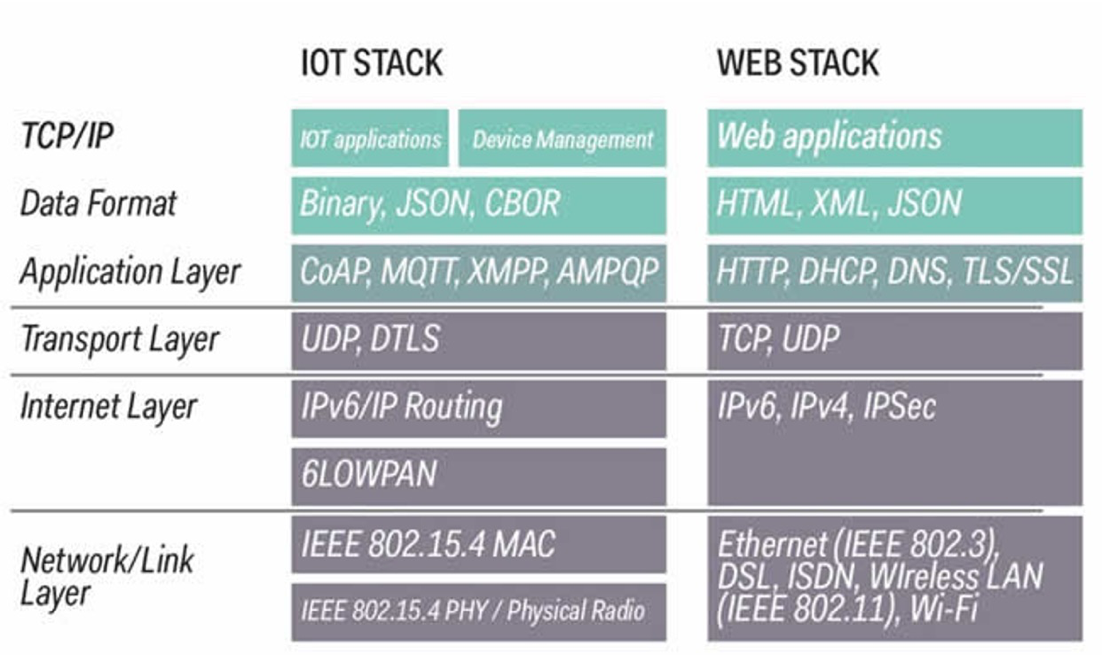
* Enabling Technologies for IoT
  + Low-cost and low-power sensing
  + Low-power and high data rate networking
  + Interoperability of different protocols （不同协议的互操作性）
  + Cloud/Edge computing platforms
  + AI/ML analytics enabling valuable insights and business solutions
  + Novel security and privacy solutions （新颖的安全和隐私解决方案）

* OSI Model
  + Conceptual model → **coordination of [ISO] standards** （协调ISO标准） → systems interconnection
  + **Layer:**
    1. Physical Layer - Transmits raw bit stream over physical medium （在device之间传递raw bit stream）
    2. Datalink Layer - Defines the *format of data* on the network
    3. Network Layer - Decides which *physical path* the data take
    4. Transport Layer - Transmits data *using transmission protocols* [TCP & UDP]
    5. Session Layer - Maintains connections & controlling ports and sessions
    6. Presentation Layer - Ensures data is in *usable format* & *data encryption* occurs （数据加密）
    7. Application Layer - *Human - computer interaction* layer

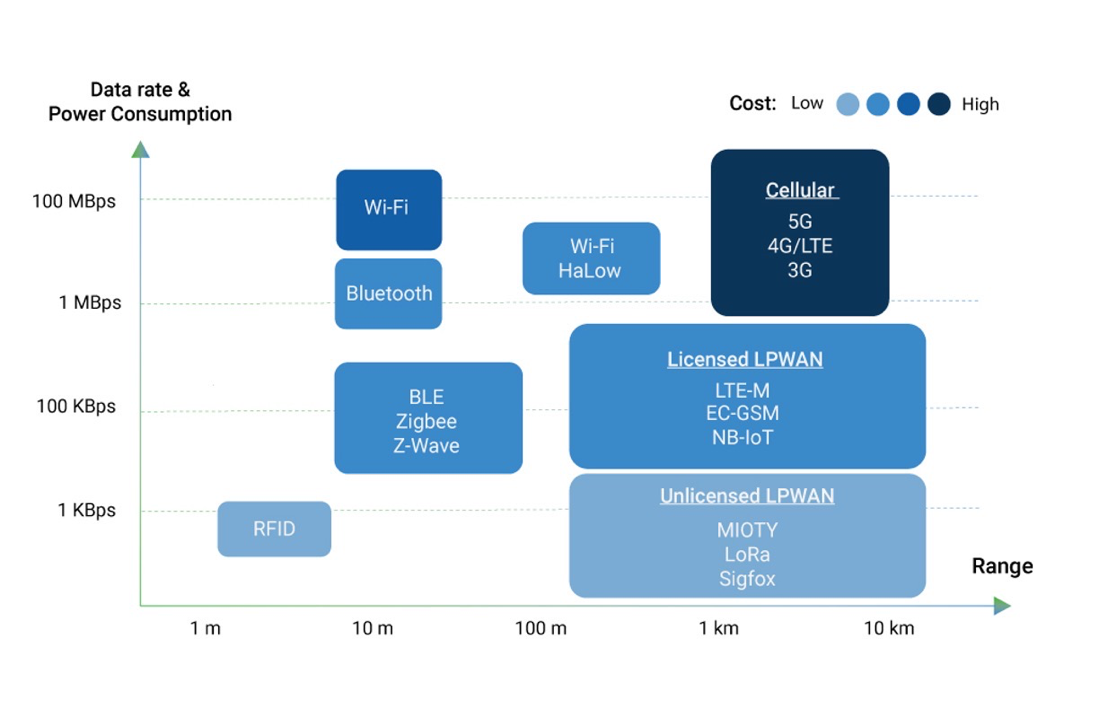
* **TCP/IP Model (Very IMP!!!)**



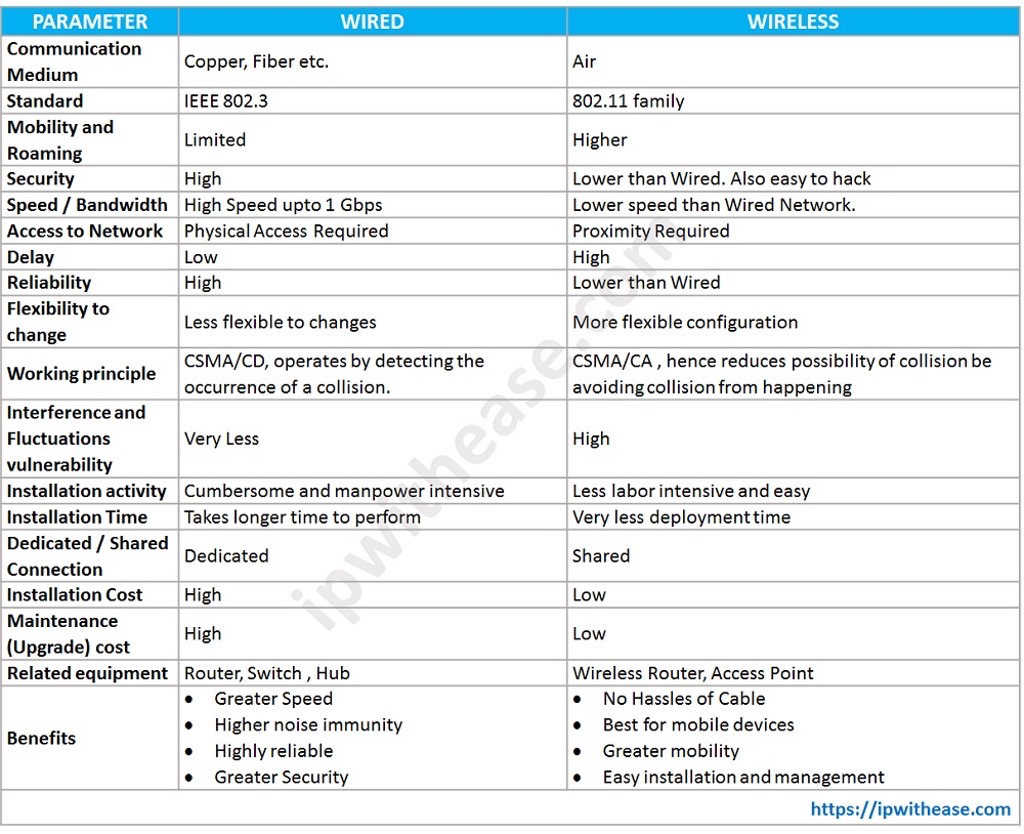
* IoT Stack



* Network Protocols for Buildings （网络协议） [communication mudiums used in IoT - Wire & Wireless]
  + Wired Protocols: BACnet, KNX, LonWorks, DALI, Ethernet
  + Wireless Protocols: WiFi, Bluetooth, ZigBee, Z-wave, 6LoWPAN, NB-IoT, 3G/4G/5G



* Wired V.S. Wireless Protocols:



* IoT Solutions for Smart Building
  + **Localization for occupants & resources tracking** （本地化&资源追求）
    - Enable navigation inside the building for new occupants
    - Enable new applications E.g. personalized experience in museums
    - Automate/Predict user requirements based on location information
    - Enable building managers to locate equipment or facilities needing repair
    - Understand occupant’s behaviour and predict unique events inside buildings by dynamically localizing occupants.
  + **Energy management**
    - Use smart meters/sensors to monitor and predict the *building environment and energy usage*
    - Provide a context-aware solution to control HVAC, lighting and other systems inside building based on **dynamic occupant behaviour**
    - Enable **dynamic adjustment** of supply and demand of electricity/energy by leveraging real-time information from smart meters
  + **Facility management**
    - Enable **timely preventive maintenance and malfunction detection** of building equipment to ensure the facility's optimal condition.
    - Provide adaptive and real-time access to building facilities for relevant personnel
    - Enable integration and remote monitoring and management of different system
    - Enable predictive maintenance and fault diagnosis by using large amount of collected data
  + **Indoor comfort enhancement**
    - Enable robust monitoring and control of the indoor built environment to maximize comfort
    - Enable automatic adjustment of indoor setting based on real-time data
    - **Joint control** of HVAC, lighting, and other systems
    - Enable learning of occupant’s preference based on collected historical data
  + **Occupant safety and health security**
    - Use real-time data from wearable sensors, smartphone, and other sensor → detect safety & health-related emergency situations
  + **Efficient resource management for convenience**
    - Use IoT for smart parking → optimize space usage & reduce time wastage
    - Use real-time occupancy data → optimal space utilization in office
  + **Building health control**
    - Integrate data from multiple sensors → structural health monitoring, fire hazard…

Big Building Data and Cloud/Edge Computing

* Big Data [Definition]
  + **Primarily**（主要是）refers to data sets that are too large / complex to be deal with by traditional data-processing application software

* Big Data Features (5V's / 4V's)
  + Value （不必须有 not all Big Data have value）
  + Velocity (data inflow high velocity E.g. IoT have high velocity)
  + Veracity
  + Variety
  + Volume

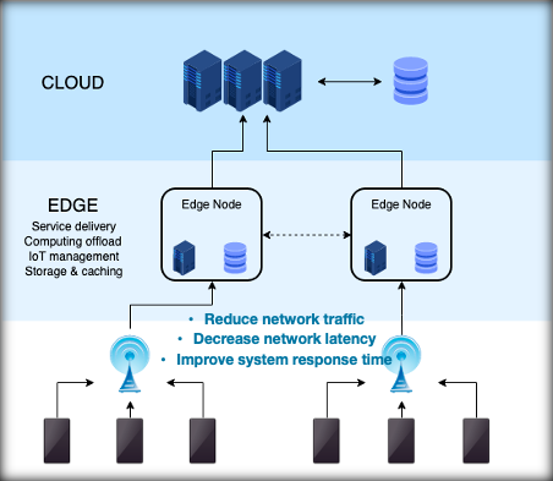
* Big Data Source in Building （一切和Building相关的量化数据都可以被称作是Big Data）
  + Energy consumption/production
  + building management systems
  + asset management data
  + building stock auditing
  + weather and climate data
  + social media data
  + socioeconomic data
  + energy end-user’s characteristics and comfort levels
  + open source data
  + etc.

* Big Data Challenges:
  + Data Challenges: Volume, variety, veracity, velocity, variability, visualization, value, etc.
  + Data **Processing** Challenges:
    - **Data Acquisition and Warehousing:** acquiring data from diverse sources and storing for value generation purpose （数据采集和仓储）
    - **Data Mining and Cleansing:** extracting and cleaning data from a collected pool of large scale structured/unstructured data （数据挖掘和清理）
    - **Data Aggregation and Integration:** aggregating and integrating clean data mined *from large* ***unstructured*** *data* （数据聚合和集成）
    - **Data Analysis and Modelling:** centers around solving the intricacy of relationships between different data and providing valuable insights （数据分析和建模）
    - **Data Interpretation:** visualising data and making data understandable for users （数据解释）
  + Data **Management** challenge
    - Privacy
    - Security
    - Data Governance
    - Data and Information sharing
    - Cost/Operational Expenditures （成本/运营支出）
    - Data Ownership （数据所有权）

* Power of Cloud
  + Cloud has become mainstream for business & many organizations (in last decade)
  + Cloud Computing includes providing on-demand services (database, servers, storage, applications, etc.) for users *via internet*

* The Great Cloud Bottleneck （巨大的cloud的瓶颈）
  + Cloud has many applications but not enough for IoT application
  + Drawbacks:
    - High Cost
    - High Response time (Large delay - sending and receiving data from remote cloud servers that are *located far from sensors/actuators*)
    - Bandwidth Congestion （带宽拥挤）
    - Limited Scalability （可扩展性有限）
    - Privacy Leakage

* Edge Computing
  + Edge Computing pushes the computation, storage, and other services *close to the data sources*



* Component
  + Local Devices
  + Localized Data Centre
  + Regional Data Centre
* Functionality
  + Caching （缓存）
  + Storage
  + Processing
  + Decision making
  + Security
  + …
* Benefits
  + Faster response time
  + Reliable Operations with intermittent Connectivity （具有间歇性链接的可靠操作）
  + Security & Compliance （安全和服从性）
  + Cost - Effective solutions
  + Interoperability between Legacy & Modern Devices （新老设备间的互操作性）
* Platforms
  + Extend cloud services to edge for data preprocessing: （将云服务扩展到Edge进行预处理）
    - Intelligent EdgeFabric
    - Azure IoT Edge
    - Cloud IoT Edge
    - AWS IoT Greengrass
  + Extend native Kubernetes to support edge computing seamlessly: （扩展原生Kubernetes来无缝支持Edge computing）
    - KubeEdge
    - Baetyl
    - Openyurt
  + **Features**
    - on-demand services
    - broad network access
    - resource pooling
    - rapid elasticity
    - measured service

* Challenges
  + Resource-constraint devices with limited resources （资源有限）
  + Heterogeneity of network, compute, and storage resources at edge devices
  + Dynamic network with intermittent network connectivity and device/link failures （间歇性动态网络连接故障）
  + Sharing and management of resources in distributed network （分布式网络的资源共享&管理）
  + Extreme scale of devices
  + Security and privacy concerns

AI and ML

* AI:
  + **processes and algorithms** that are able to simulate human intelligence, including mimicking cognitive functions such as perception, learning and problem solving.
  + Goal
    - To **Create Expert Systems** − exhibit intelligent behavior, learn, demonstrate, explain, and advice its users.
    - To **Implement Human Intelligence in Machines** − Creating systems that understand, think, learn, and behave like humans.
* Narrow V.S. General AI

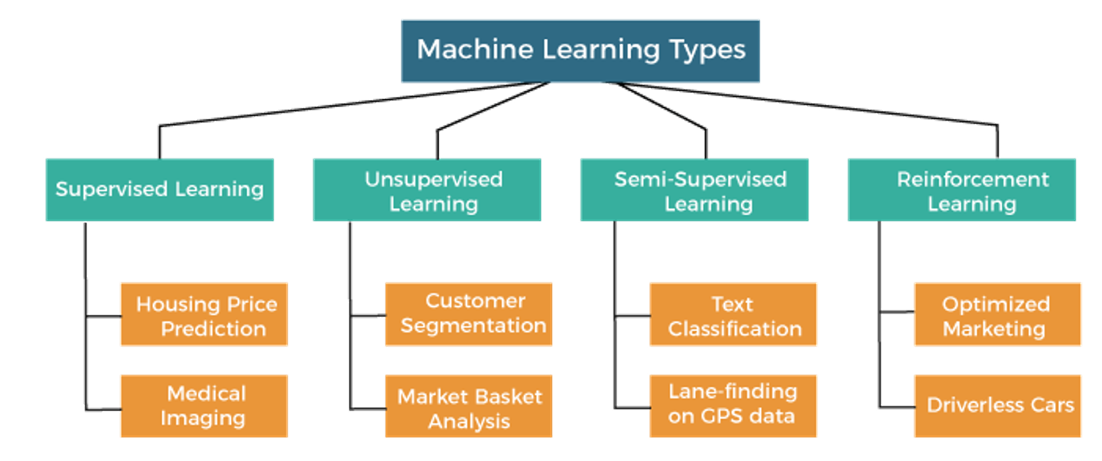


* Benefits of AI
  + High Accuracy with less errors
  + High-Speed
  + High reliability
  + Useful for risky areas （适用于危险区域）
  + Digital Assistant
  + Useful as a public utility

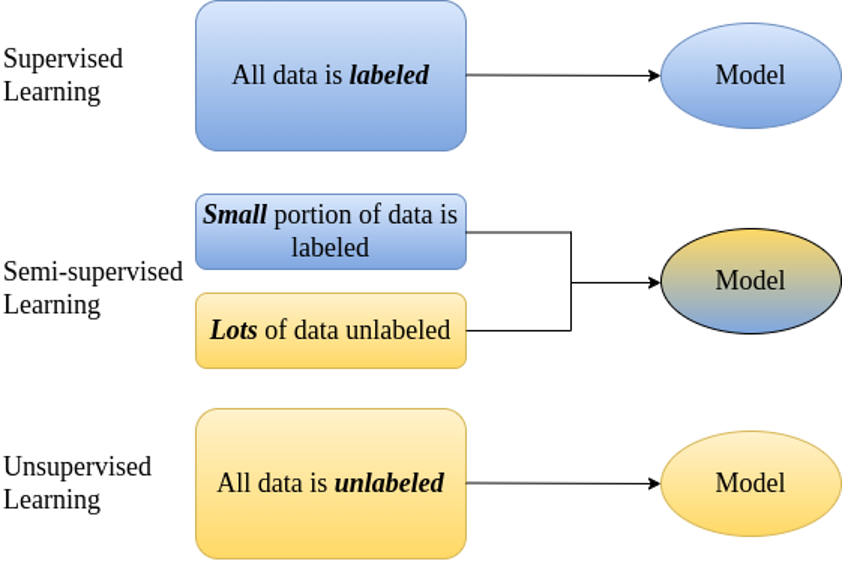
* Applications of AI
  + Data Security
  + Finance
  + Gaming
  + Astronomy
  + Healthcare
  + Transport
  + Agriculture
  + Etc.

* Challenges / Disadvantages of AI
  + High Cost
  + Can't think out of the box
  + No feelings and emotions
  + Increase dependency on machines
  + No Original Creativity
  + Ethical concerns

* Machine Leaning (ML) - **Part of AI**
  + **studying** how computer agents can improve their perception, knowledge, thinking, or actions based on experience or data.
  + Type



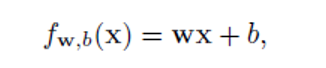
* Supervised Learning
  + [input data & l**abelled** output data]
  + Main Goal: map the input variable (x) with the output variable (y) （将输入与输出变量映射）
  + E.g. Risk Assessment, Fraud Detection, Spam filtering, etc.
  + **Classified into two types problem**
    - Classification
    - Regression
* Unsupervised Learning
  + [**not include labelled data**]
    - Trained with data *neither classified nor labelled* - model acts on data without any supervision
  + Main Goal: group / categories the unsorted dataset according to the *similarities, patterns and differences* - find the **hidden patterns** from the input dataset
  + Discover similarities & differences in information. E.g.exploratory data analysis, cross-selling strategies, customer segmentation, and image and pattern recognition.
  + **Classified into:**
    - Clustering
    - Dimensionality reduction
    - Association
* Semi-Supervised Leaning
  + [both **labelled & unlabelled** data]
  + Solve: not having enough labeled data for a supervised learning algorithm & helps if it’s too costly to label enough data.
  + Conjunction with a small amount of labeled data & unlabelled data → improvement in learning accuracy.
  + Variety of problems from ***classification and regression to clustering and association*** for different application domains such as medical, smart buildings, etc.



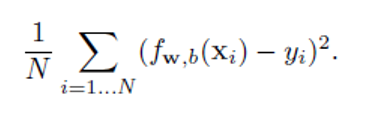
* Reinforcement Learning
  + Involves agents that learn to take suitable action → **maximize reward** in a particular situation.
  + **Not aware** of the correct output/label instead the agent learns to take optimal action based on experience of which action results in maximum reward.
  + **Similar to human being**
  + Employ: Game theory, Operation Research, Information theory, multi-agent systems.

* ML Algorithms

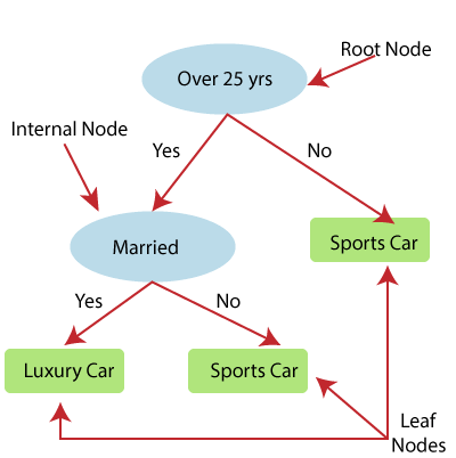
1. Regression Problem: model the relationship between a dependent (target) & independent (predictor) variables with one/more independent variables → **predicting** continuous/real values of dependent variable
   * Linear regression: linear relationship between the independent variable (X-axis) and the dependent variable (Y-axis) **[Linear problem]**
     + Simple linear regression: if there is only one input variable (x)
     + Multiple linear regression: more than one input variable
     + Objective: Learn value of parameters **w** and **b** → hyperplane in linear regression is as **close to all training examples as possible** （尽可能接近若有训练样本）



* Formal definition: To *minimize the squared error loss* （最小化平方误差）



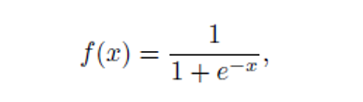
* Decision Trees [**supervised learning algorithm**]
  + Solve:
    - Both Classification & Regression
    - Both categorical & numerical （解决分类和数据问题）
  + Build a tree-like structure
    - Each internal node represents the "test" for attribute （每个内部节点表示属性的“测试”）
    - Each branch represent the **result** of test （每个分支表示测试结果）
    - Each leaf node represents the **final decision / result** （每个叶节点表示最终决策/结果）



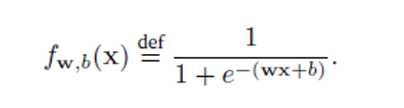
1. Classification Problem: program learns from the given dataset or observations → classifies new observation into a number of classes or groups
   * Binary Classifier: classification problem has only two possible outcomes
   * Multi-class Classifier: classification problem have more than two outcomes
   * Logistic Regression: predicts the output of a *categorical dependent variable*.（分类因变量的输出）
     + Outcome: categorical / discrete value

E.g. 0 or 1; true or false; yes or no

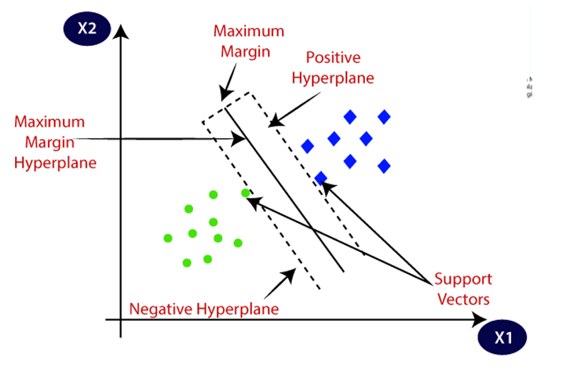
* Imp. It is not giving the exact value as 0 and 1 → give the probabilistic values which **lie between 0 and 1**.
* Sigmoid function (to get probability value between 0-1)



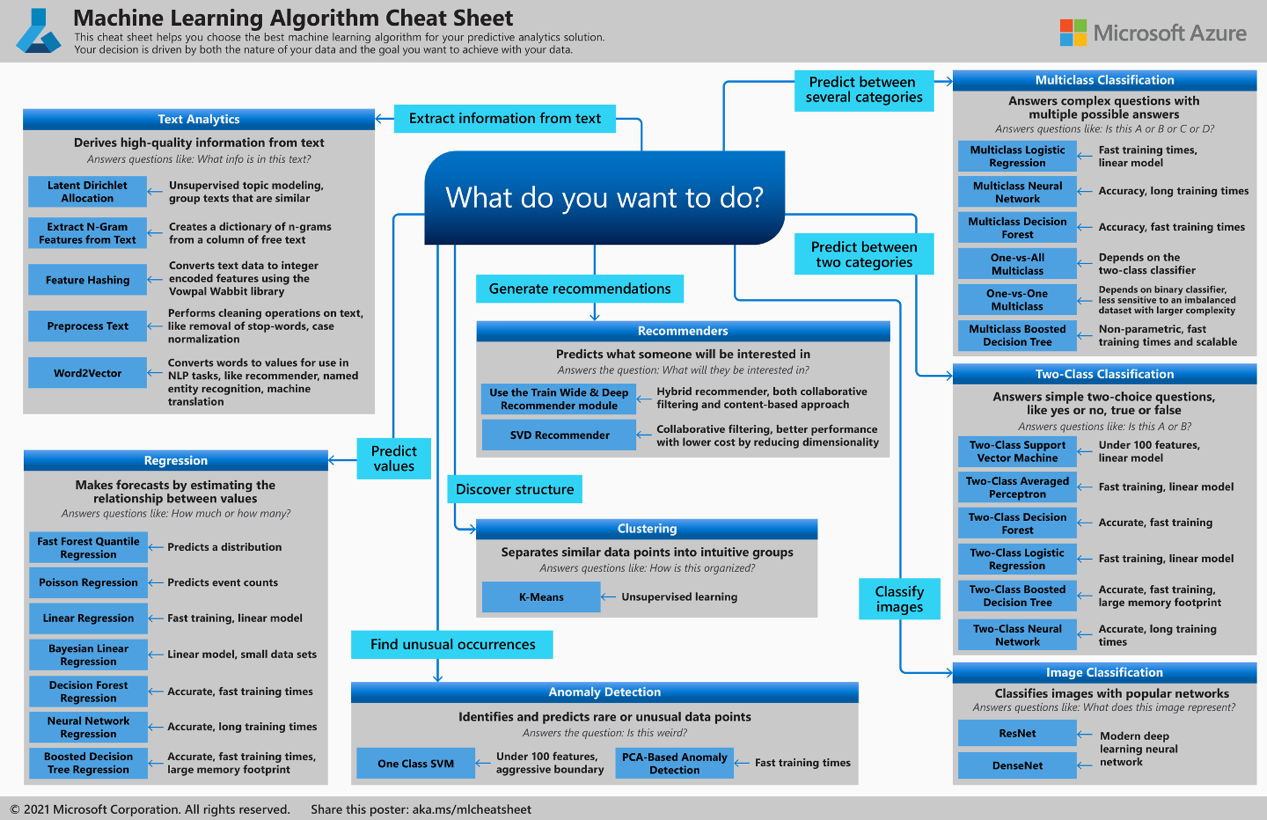
Plug in linear regression:



* Objective: **maximize the log-likelihood** of out training set
* Support Vector Machines (SVM)
  + Objective: create best line / decision boundary → can segregate（分开） n-dimensional space into classes → we can easily put the new data point in the correct category in future
  + **Hyperplane**: Best decision boundary
  + Maximum Margin （最大的距离）→ Maximum Margin Hyperplane
  + Support Vectors - 距离 Positive Hyperplane / Negative Hyperplane 最近的一个点



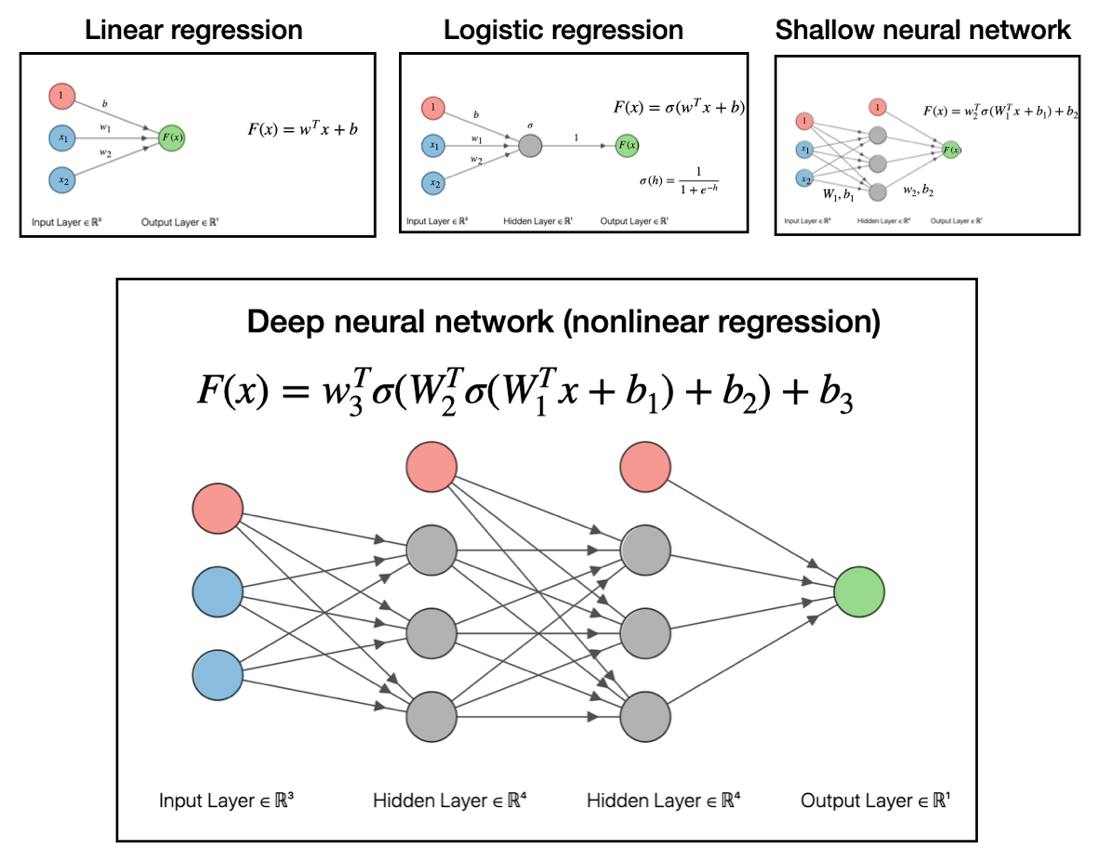
1. Clustering Problem: divide the population / data points into a number of groups [more similar to point in same group & dissimilar to point in other group]
   * Collection of object - similarity & dissimilarity between them
   * K-means Clustering: partitions n observations into k clusters [each observation belongs to the cluster]
     + Prototype: observation with nearest cluster mean
     + Steps for K-means clustering:
       - First, we initialize random k points, called means / cluster centroids
       - We categorize each item to its closest mean and we update the mean’s coordinates（更新均值坐标） [averages of the items *categorized in that cluste*r so far]
       - We **repeat the process** for a given number of iterations and at the end → we have our clusters. （不断迭代重复）



Deep Neural Networks

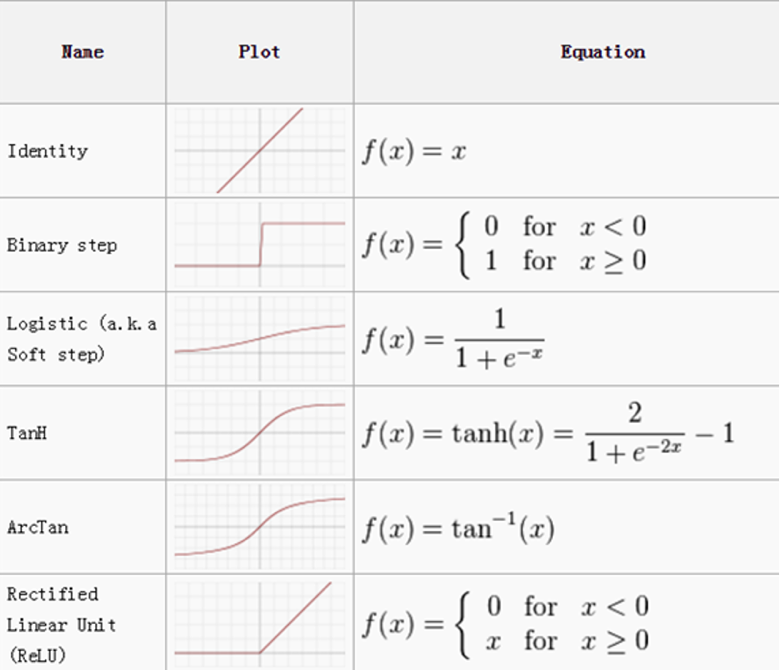
* Artificial Neural Network (ANNs) / neural networks (NNs)
  + computing systems inspired by the biological *neural networks that constitute animal brains*.

* Deep Neural Network (DNN)
  + [**a class of machine learning algorithms**] - uses multiple layers to *progressively extract higher-level features* from the raw input. （使用多个层 - 逐步提取较高级别的特征）
    - E.g. in image processing
      * Lower layers: identify edges
      * Higher layers: identify concepts relevant to human (digits/letters/faces)
  + Hidden Layer:
    - 1 - in shallow neural network
    - Many - in deep neural network
  + DNN is an ANN with multiple hidden layers [between **input and output layers**]

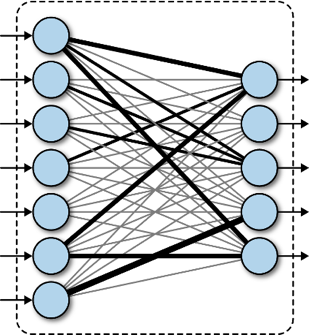


* Multilayer Perceptron (MLP)
  + Fully connected class of feedforward artificial neural network (ANN) （**完全连接的前馈人工神经系统**）（向前的 和Recurrent相对）
  + MLP consists **at least 3 layers of nodes**:
    - Input layer
    - Hidden layer
    - Output layer
  + Each node is a **neuron** [**use nonlinear activation function**]

* **Activation Functions in Neural Networks**



* Fully Connected Neural Network (FCNN)
  + Consist of a series of *fully connected layers* [Connect every neuron in one layer to other layer]

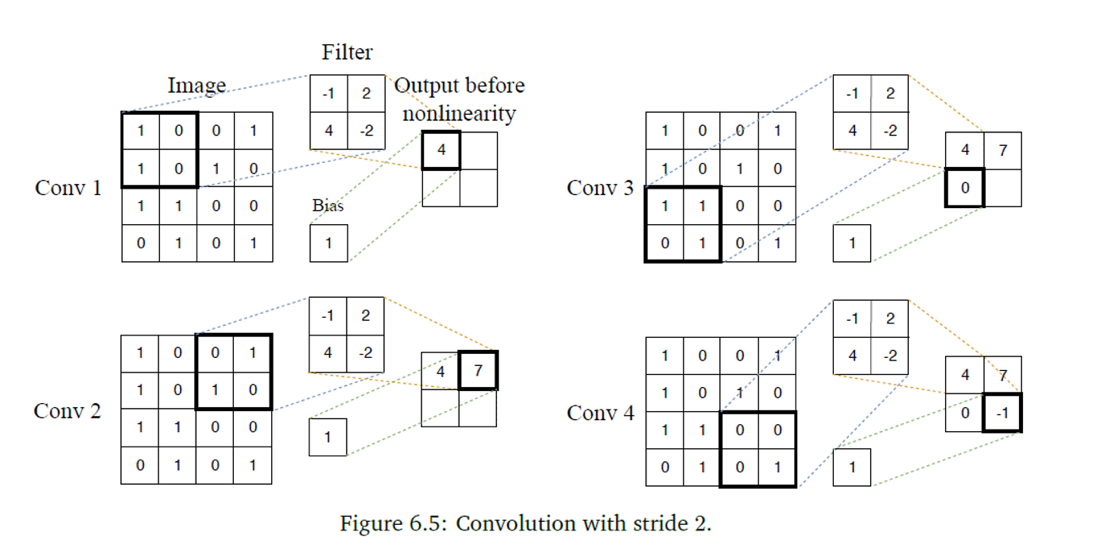


* Major advantages: "structure agnostic"（与结构无关）- there are **no special assumptions** needed to be made about the input
* Drawback: tend to have weaker performance than *special-purpose networks* tuned to the structure of a problem space
* There are large number of weight parameters to be learned in Fully Connected Neural Network (Due to fully connected layers)
* **Fully connected neural networks cannot be applied for image data**

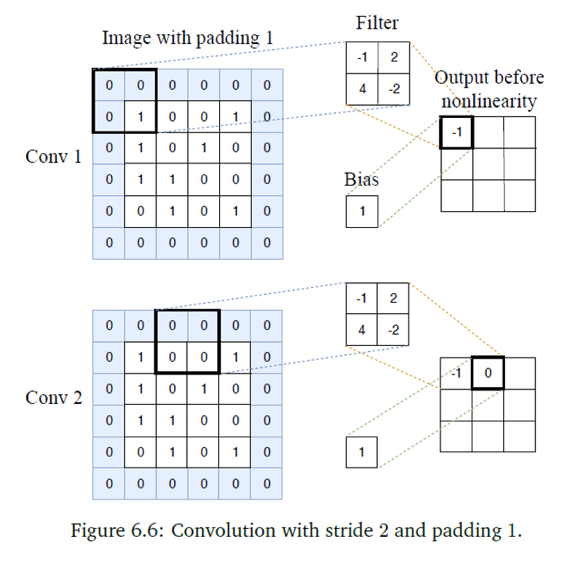
E.g. for image of size 64\*64\*3, need 12288 weights（直接将size相乘，就是一个neuron的first hidden layer的weights） for just one neuron in first hidden layer

* Convolutional Neural Network （卷轴神经） (CNN)
  + Specialized **type of ANN** - use a mathematical operation called convolution in place of general matrix multiplication in at least one of their layers （在它们的至少一层中代替一般的矩阵乘法）
  + Not all nodes in neuron are connected to output nodes → ***number of weight*** << FCNN
  + 3 main types of layers
    - Convolutional layer

Performing convolution operation to input matrix (with stride 2)

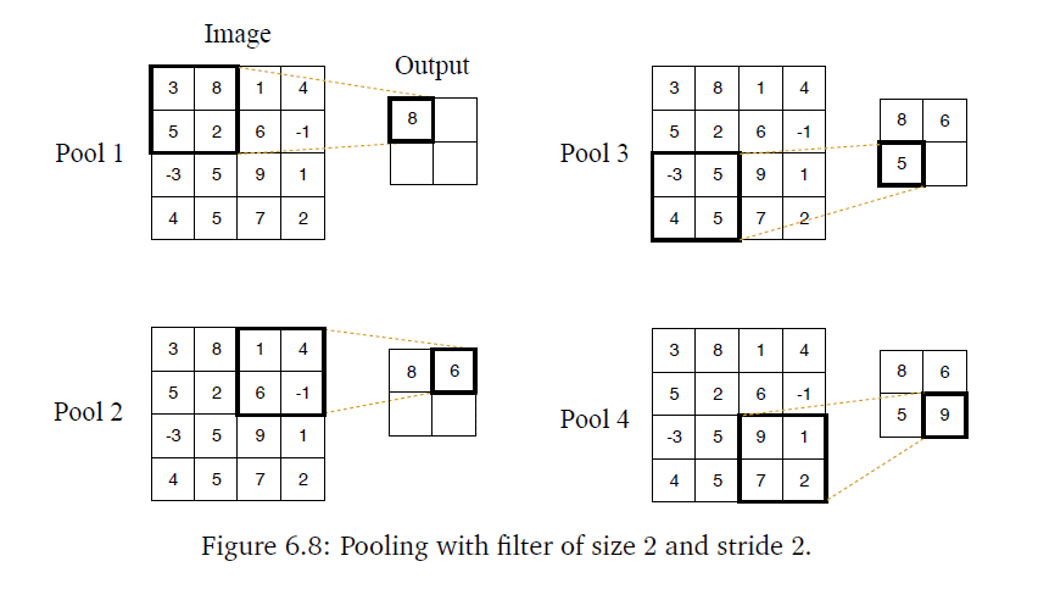


* 图中的计算就是将Filter和Image中的格一一对应相乘得数 → +bias → output中的数字
* Convolution operation using padding (stride 2 and padding1)

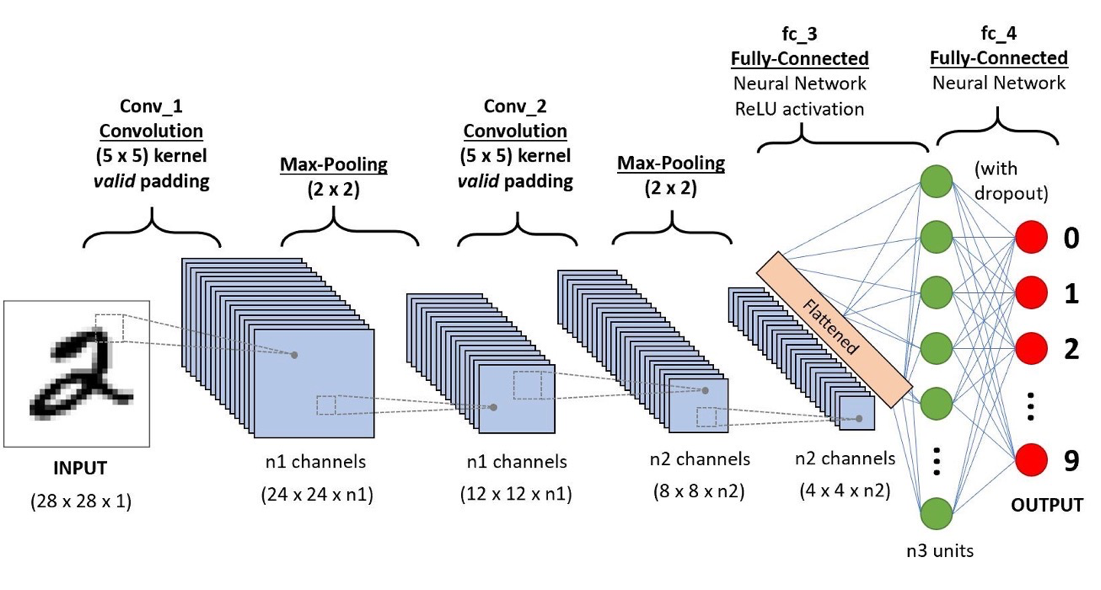


* Polling layer

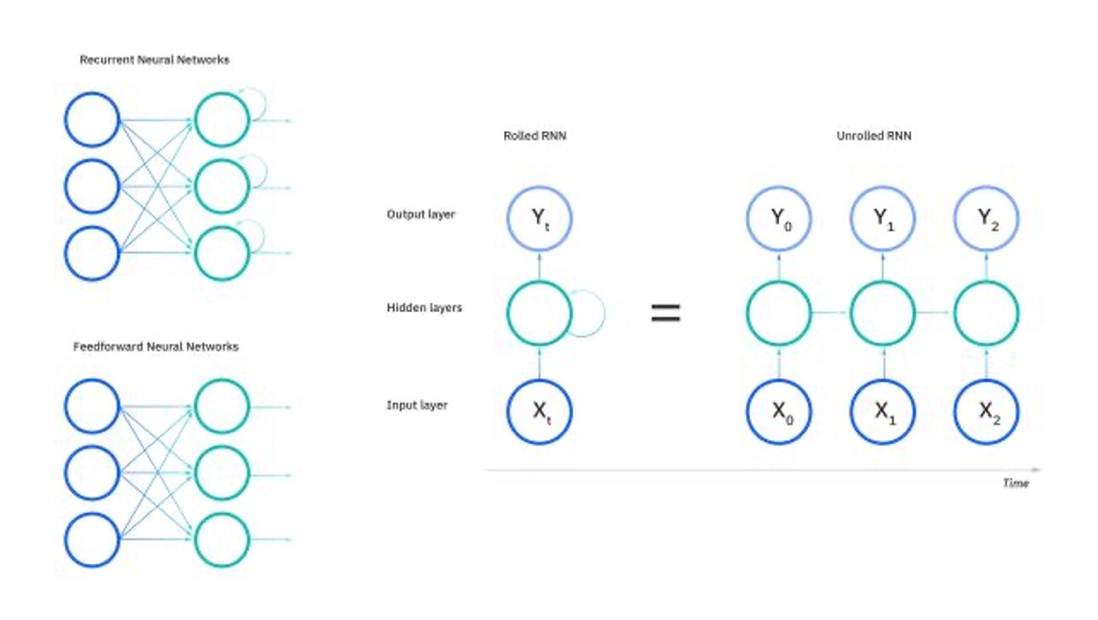
Apply a fixed operator - **usually max / average** （下图中就是用的max）



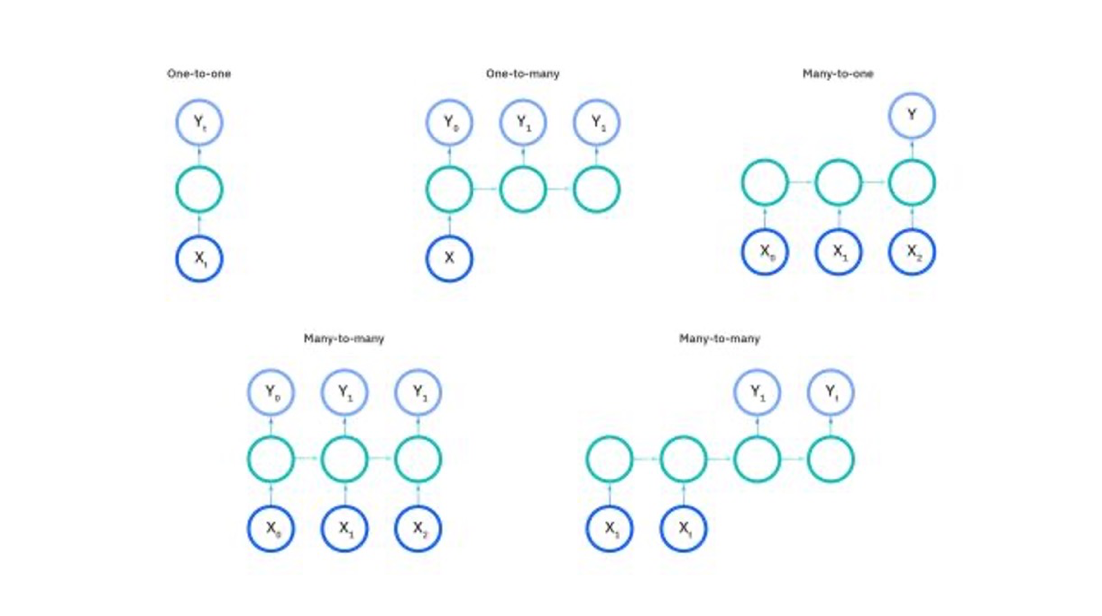
* Fully-connected (FC) layer
* CNN architectures make the explicit assumption that the inputs are images - CNN architectures are able to successfully **capture the spatial and temporal（时间与空间）dependencies in the input image data**.



* Recurrent Nerual Network （递归神经） (RNN)
  + Type of NN - output from the previous step are fed as input to current step （上一步的输出作为当前步骤的输入馈送到当前步骤）
  + Use: *sequential data / time series data*



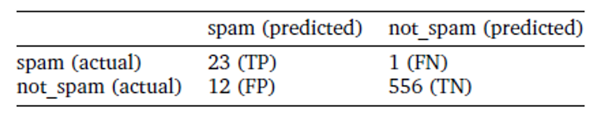
* Output in RNN can vary in length
* Different types of RNNs → Different use cases



* Performance Metrics （性能指标）
  + Model is trained → evaluate on *test data*
  + Regression:
    - Mean squared error
    - Root mean
    - Square error
    - Mean absolute error
  + Widely metrics & tools for **classification model**:
    - Confusion matrix - a table summarizes successful classification model [*predicting examples - various class*]



* **Binary** Classification



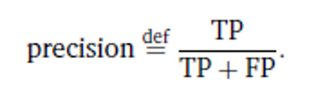
* **Multiclass** Classification



Confusion matric → further calculate precision & recall↓

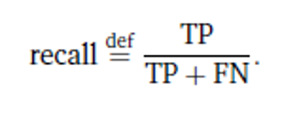
* Precision/Recall
  + **Precision:** ratio of correct positive predictions (TP) / overall number of positive prediction (TP+FP) [fraction of predictions as a positive class were actually positive.]

E.g. precision is the proportion of **relevant documents** in the list of *all returned documents*

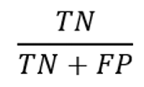


* **Recall:** ratio of correct positive predictions (TP) / overall number of positive examples (TP+FN) [fraction of all positive samples were correctly predicted as positive by the classifier.]

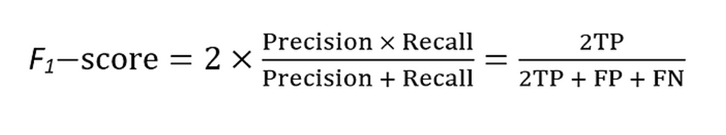
E.g. recall is the ratio of the **relevant documents** returned by the search engine to the *total number of the relevant documents* that could have been returned



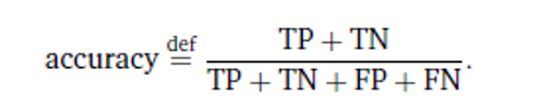
* In the case of the spam（垃圾邮件）detection problem, we want to have **high precision** (we want to *avoid making mistakes by detecting that a legitimate message is spam*) and we are ready to **tolerate lower recall** (we tolerate some spam messages in our inbox). （precision与recall是相反的，如果想要更高的precision就会带来较低的recall）
* Specificity
  + proportion of actual negatives correctly identified (TN) / all negatives (TN+FP) [跟recall相反]



* F1-score
  + Calculate **harmonic mean** → find balance between precision & recall



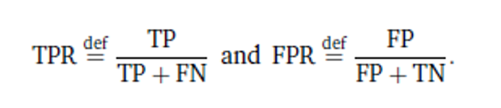
* Accuracy
  + correctly classified examples (TP+TN) / total number of classified examples (TP+TN+FP+FN)



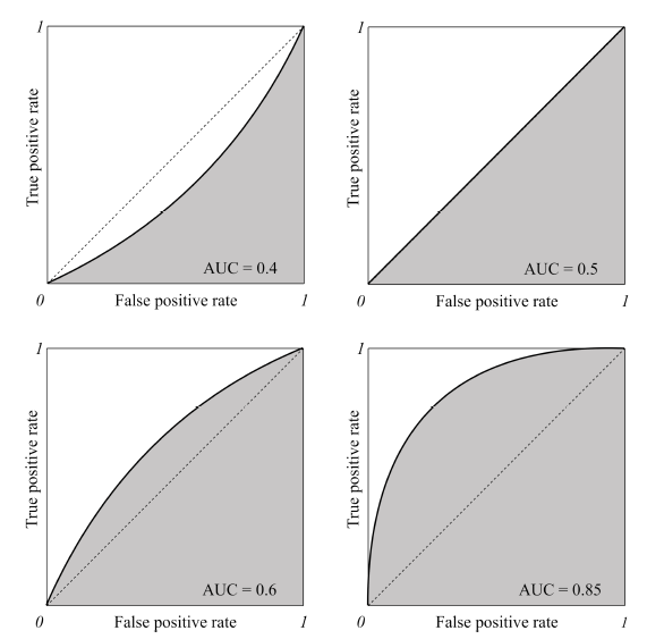
* When errors in predicting all classes are **equally important** → Accuracy is useful

E.g. spam/not spam is not this case ← this is not equal (false positives less than false negative)

* Area under the ROC curve
  + ROC stands for operating characteristic; graph is plotted against TPR (TP Recall) & FPR (FP Recall) [various threshold values]
  + ROC curves can **only assess** classifiers that *return some confidence score / a probability* of prediction（只能用于评估返回预测的置信分数/概率的classifier)



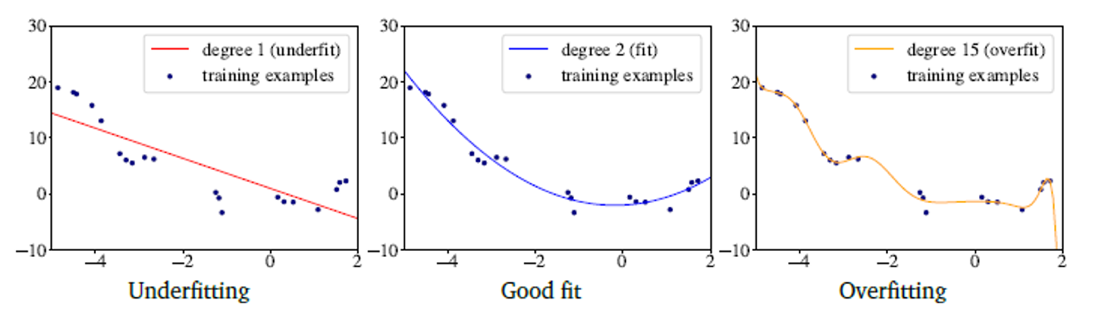
* ROC AUS is area under the curve [higher its number → better]

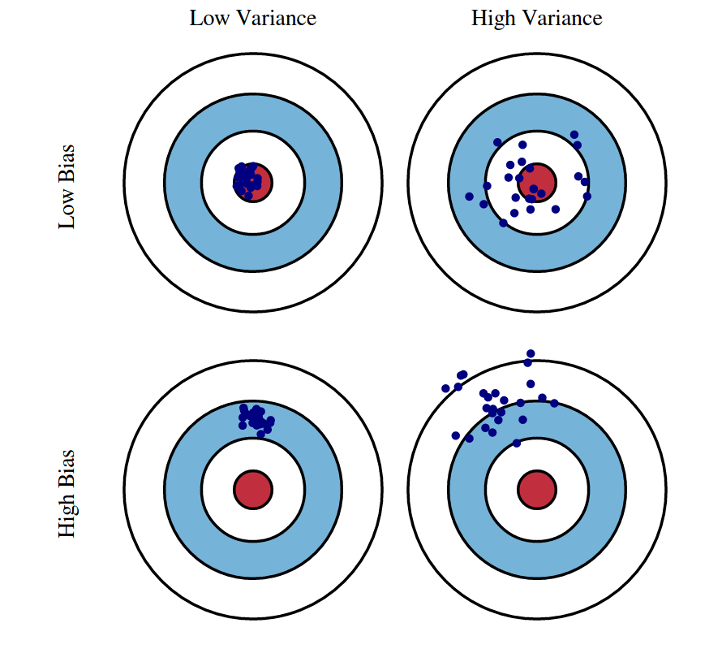


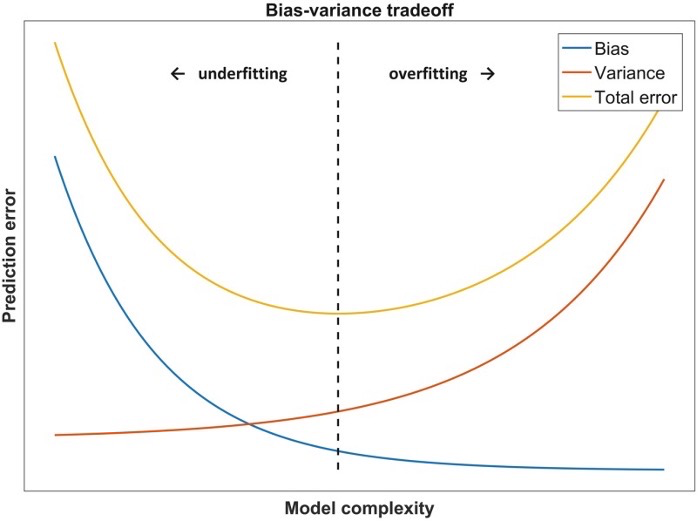
* Overfitting & Underfitting

If model is good ML model → generalizes any new input data from problem domain in proper way

* Bias:
  + Assumptions function actually the *error rate of training data*
  + Error rate: high - High Bias; low - Low Bias
* Variance:
  + Difference between the error rate of training data & testing data
  + **Difference of errors high** - high variance
  + **Difference of errors low** - low variance
  + Goal: usually want to make **low variance** for generalized out model
* Underfitting: model/algorithm cannot capture the underlying trend of data （无法捕捉数据基本趋势）
  + Model/algorithm **not fit** the data well enough - have fewer data to build accurate model & we want to build a linear model with fewer non-linear data （Model 和 algorithm都有）
  + Not perform well **both** on training data & new data
  + **Reasons for Underfitting:**
    - High bias & low variance
    - Size of training dataset is not enough
    - Model is too simple
    - Training data not clean - contain noise in it
* Overfitting: model does not make accurate predictions on **testing data （只有model）**
  + When model trained with so much data ← model start to learn from noise & inaccurate data entries
  + Evaluation of ML algorithms on training data is different from unseen data (new data)
  + Reasons for Overfitting:
    - High variance & low bias
    - The model is too complex
    - The size of training data is too small （容易被一些noise影响）







* Address Overfitting / Underfitting
  + Reduce underfitting:
    - Increase model complexity
    - Increase the number of features, performing feature engineering
    - Remove noise from data
    - Increase the number of epochs（周期）/ increase the duration of training（增加训练时间） → get better result
  + Reduce overfitting
    - Increase training data
    - Reduce model complexity
    - Early stopping during the training phase (loss begins increase → stop training)
    - **Ridge Regularization and Lasso Regularization**
    - Use dropout for neural networks