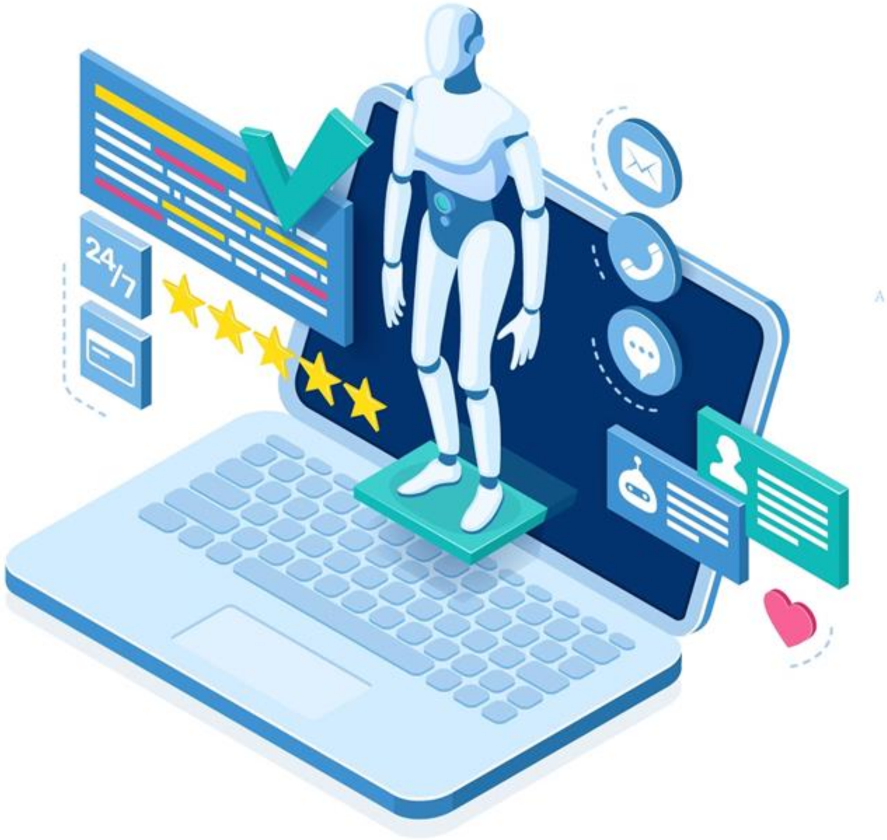
UNIVERSITAS INDONESIA
Veritas, Probitas, Iustitia

FAKULTAS
ILMU
KOMPUTER

Machine Learning
in Brief



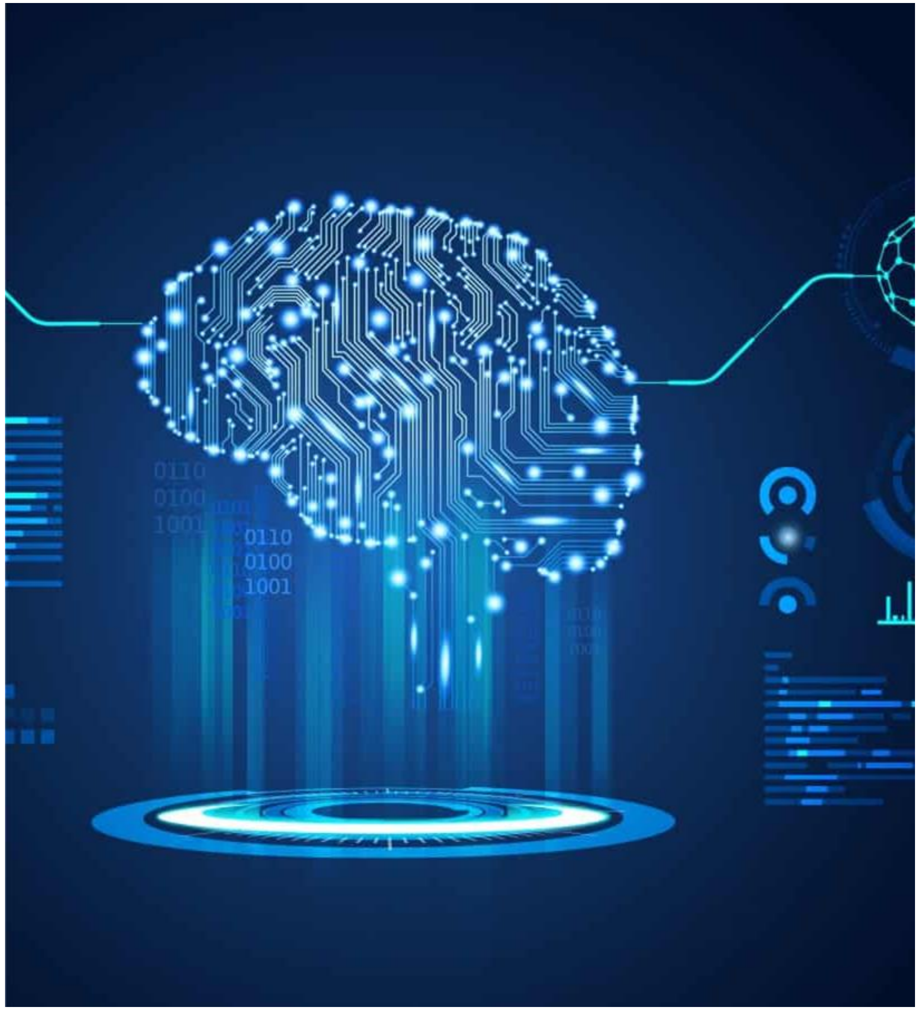
Prof. Dr. Eng. Wisnu Jatmiko
Guru Besar Fakultas Ilmu Komputer Universitas Indonesia

PENS, 21 November 2020

Outline of The Presentation

1. Overview of **Artificial Intelligence**
2. Definition and History of **Machine Learning**
3. **Types of Learning** (Application in Lab1231)
4. Machine Learning **Evaluation**
5. Summary - **Five Tribes** of Machine Learning

Artificial Intelligence (AI)

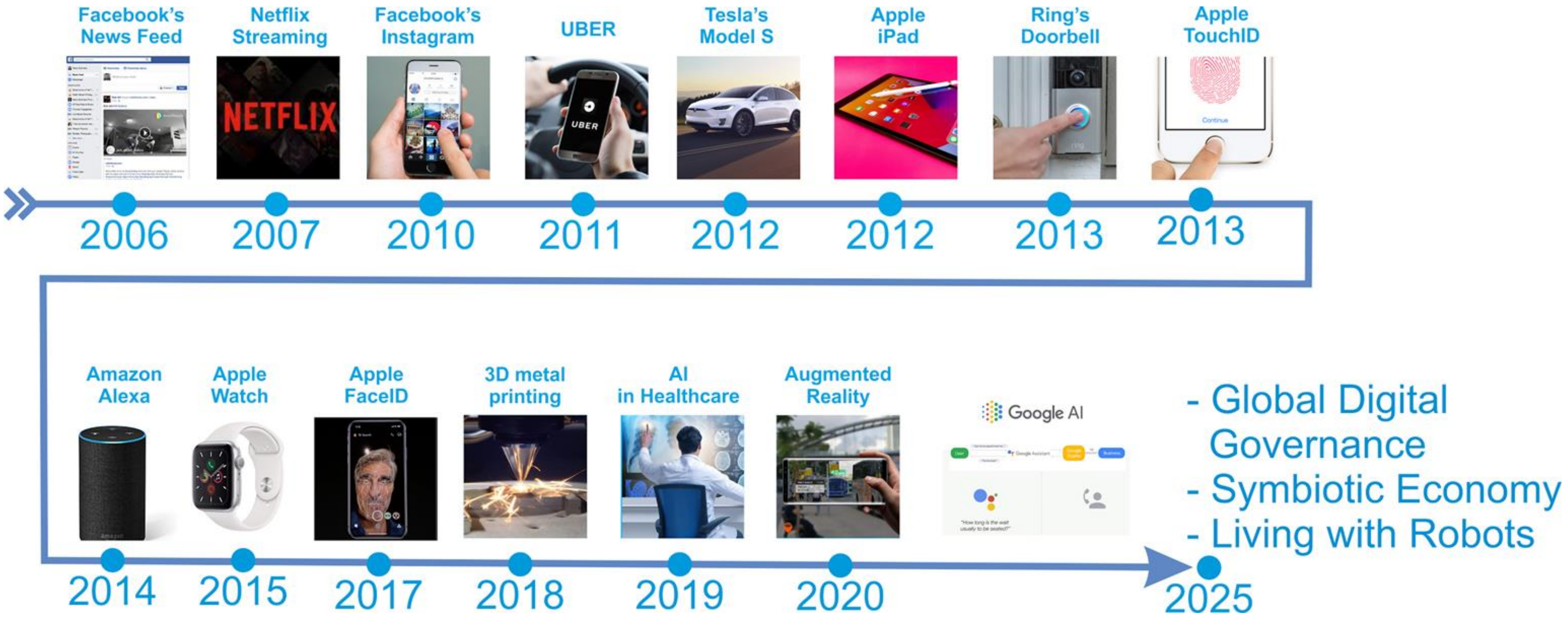


- …. making a machine behave in ways that would be called intelligent if a human were so behaving. McCarthy, Minsky, Rochester & Shannon, 1956.

- Turing test (1951) , “imitation game”, tests if a computer can successfully pretend to be a human in a dialogue via screen & keyboard. Dictionary.com

A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955, AI Magazine, Vol. 27(4), 2006
Adapted from Slide “Pattern Recognition: Statistics to Deep Networks” Anil K. Jain, Michigan State University, ICACSIS-IWBIS 2020

Most-Influential Technology

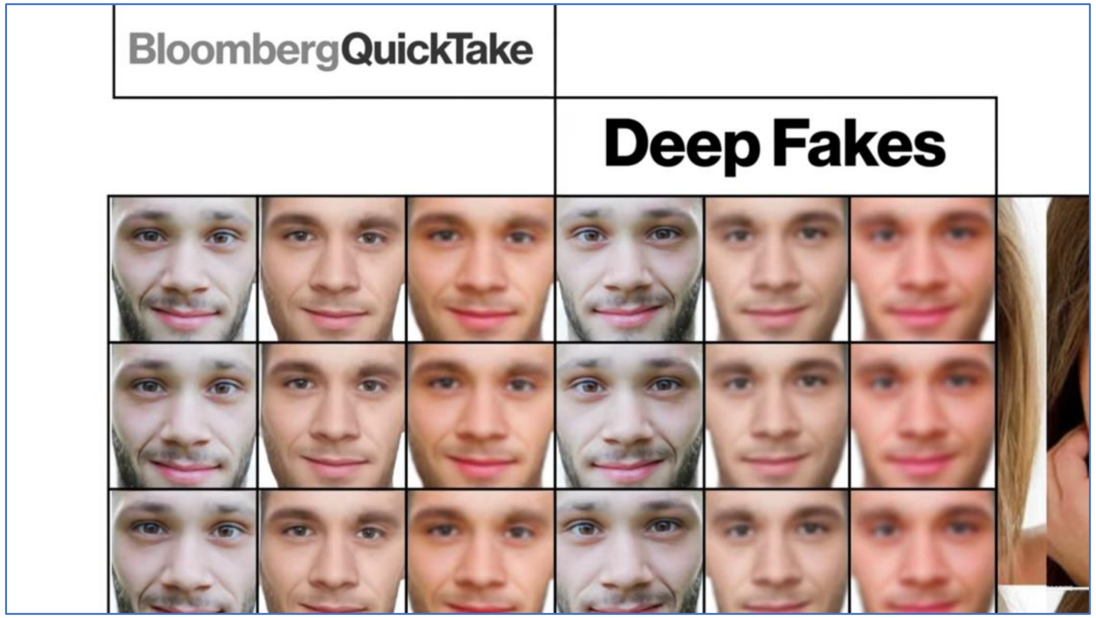


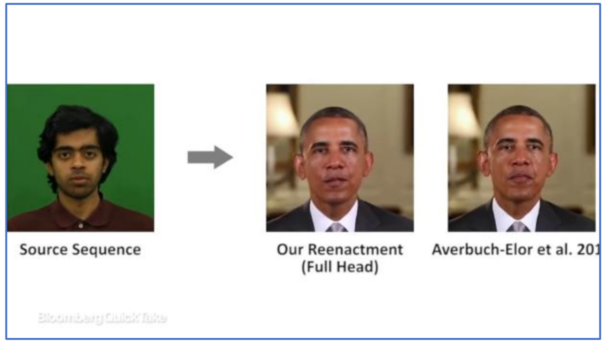
Source: https://www.washingtonpost.com/technology/2019/12/26/we-picked-most-influential-technologies-decade-it-isnt-all-bad/

Self-driving Car

Sumber: https://www.youtube.com/watch?v=aaOB-ErYq6Y&t=49s

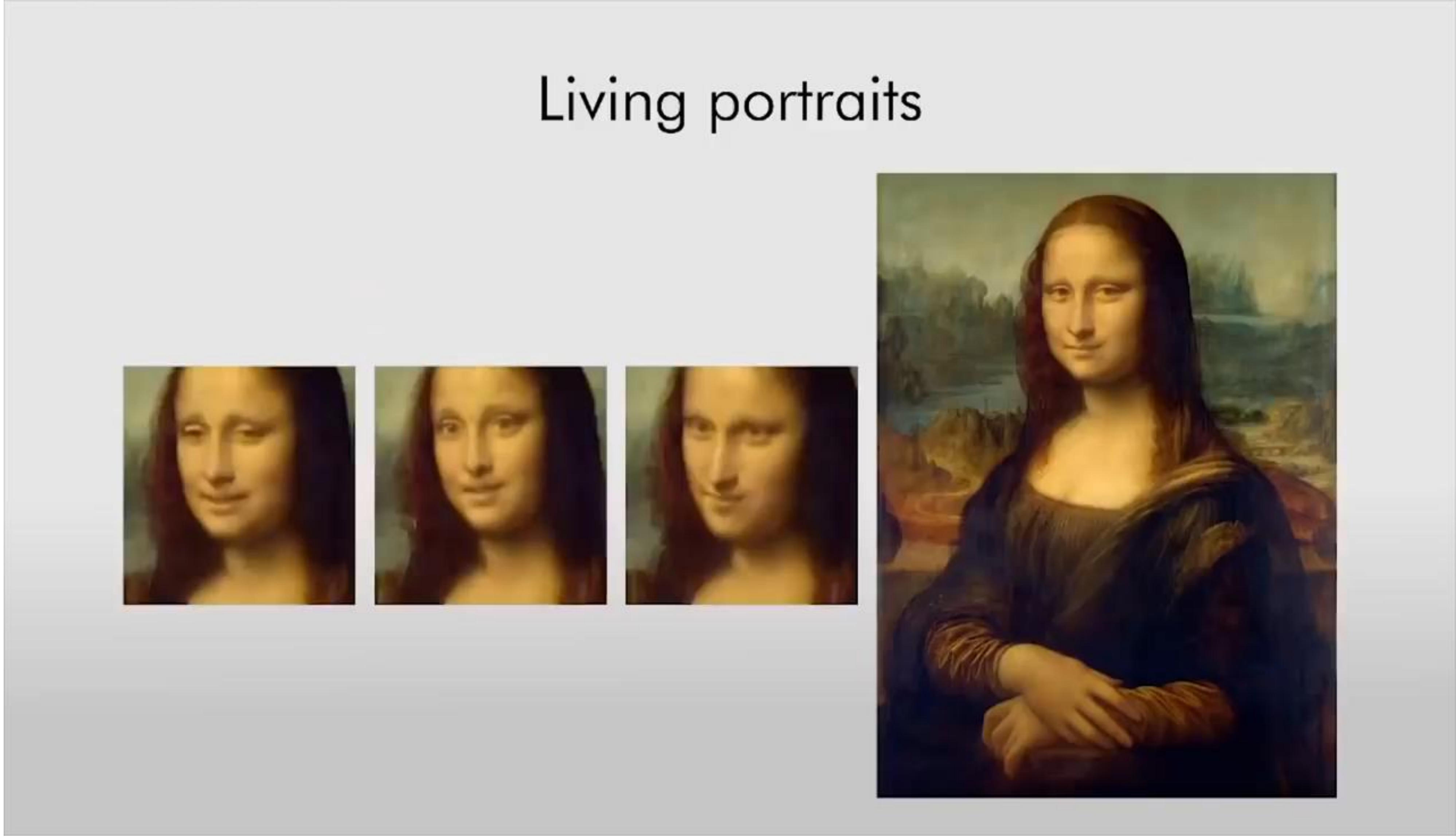
Deep Fake





Source: https://www.youtube.com/watch?v=gLoI9hAX9dw

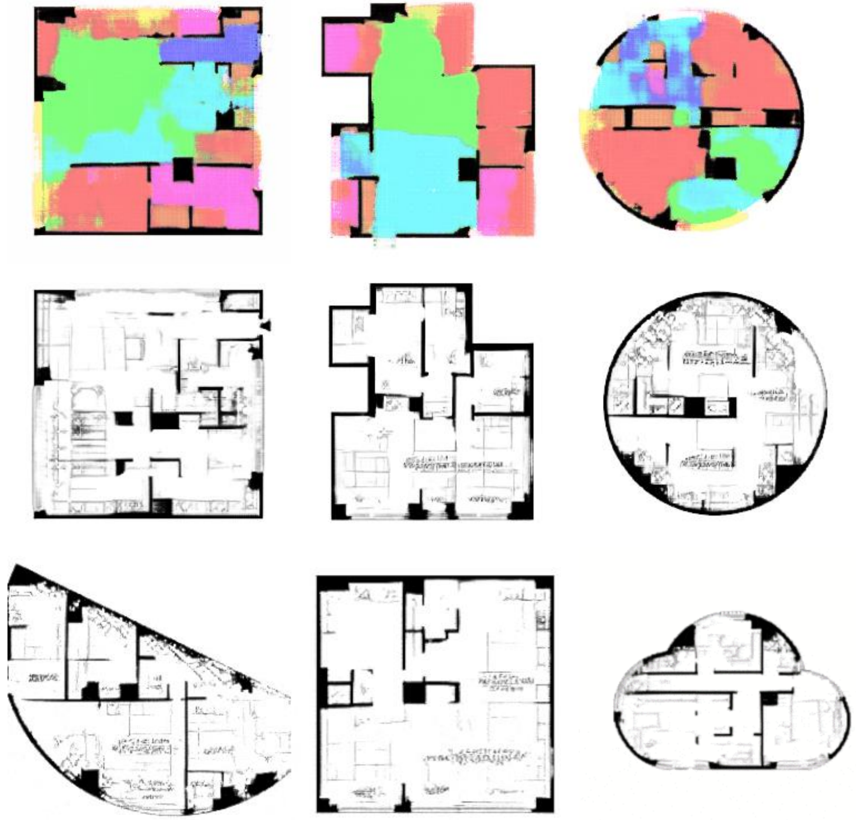
Deep Fake - Monalisa



Source: https://www.youtube.com/watch?v=P2uZF-5F1wI

ArchiGAN:
a Generative Stack for Apartment Building Design

Model II interface

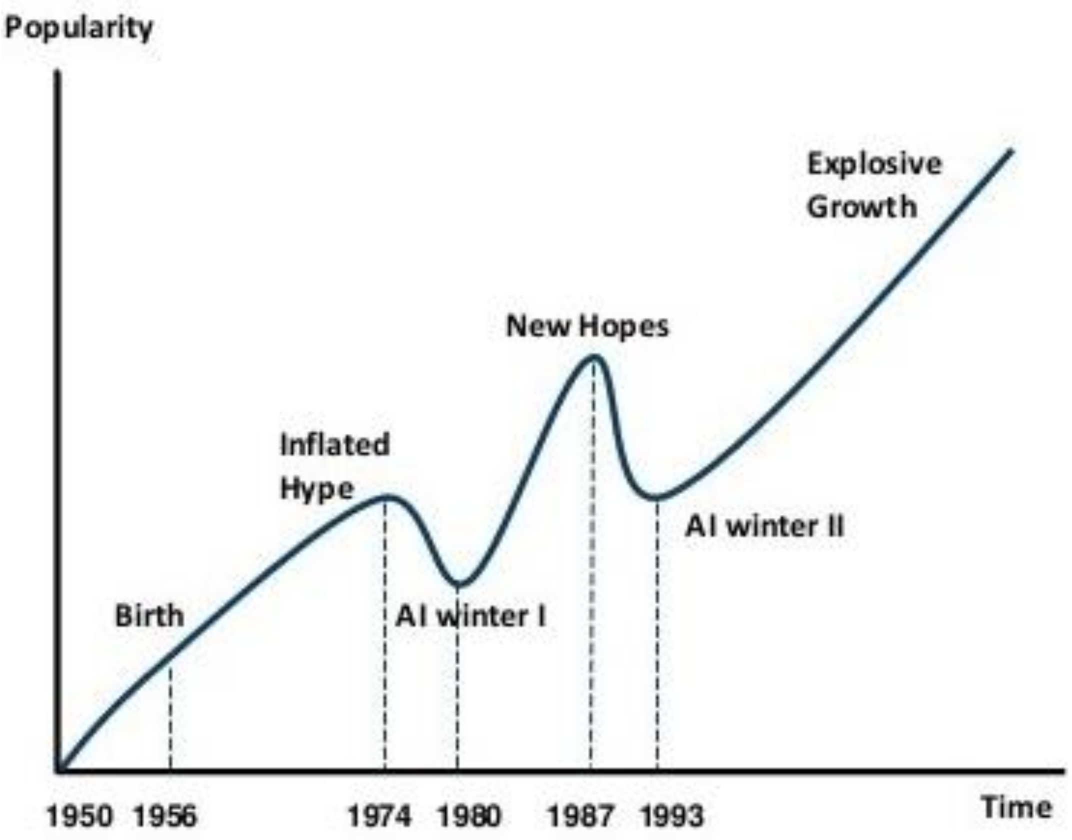


GAN-enabled Space Layout under Morphing Footprint

Source: https://developer.nvidia.com/blog/archigan-generative-stack-apartment-building-design/

AI **Winter**

AI HAS A LONG HISTORY OF BEING "THE NEXT BIG THING"...



Timeline of AI Development

- **1950s-1960s**: First AI boom - the age of reasoning, prototype AI developed
- **1970s**: AI winter I
- **1980s-1990s**: Second AI boom: the age of Knowledge representation (appearance of expert systems capable of reproducing human decision-making)
- **1990s**: AI winter II
- **1997**: Deep Blue beats Gary Kasparov
- **2006**: University of Toronto develops Deep Learning
- **2011**: IBM's Watson won Jeopardy
- **2016**: Go software based on Deep Learning beats world's champions

Another AI Winter?
(1974–1980;
1987–1993)

The **Two Big Paradigms** of AI

Symbolic AI	Statistical AI
• Intelligence as manipulation of abstract symbols • Knowledge represented declaratively using sentences of symbols (e.g., in formal logic) • Knowledge derivation via logical reasoning • Top-down design: 1. knowledge specification (e.g., by domain experts) 2. knowledge representation 3. inferencing by symbol processing applications	• Intelligence as (numerical) signal manipulation • Knowledge represented as some structured collection of numerical values • Knowledge derivation via numerical operation • Bottom-up design: 1. data as examples (not high level knowledge) 2. build model representing the data 3. use model to solve intelligent tasks (e.g,. classification, regression)

Source: Frank van Harmelen, “On the nature of AI, and on the relation between symbolic and statistical AI”, https://www.slideshare.net/Frank.van.Harmelen/panel-on-the-future-of-ai-at-the-flamish-royal-society-of-sciences-and-arts

The two big paradigms of AI: **Framework examples**

Symbolic AI	Statistical AI
• production rules, • logic programming, • expert systems, • semantic network, • frames, • logic-based knowledge representation, • ontology, • classical search algorithms, • automated planning, • linguistic modeling	• machine learning, • neural networks, • deep learning, • Bayesian probabilistic reasoning, • evolutionary systems (e.g., genetic algorithms, swarm intelligence), • fuzzy logic, • clustering

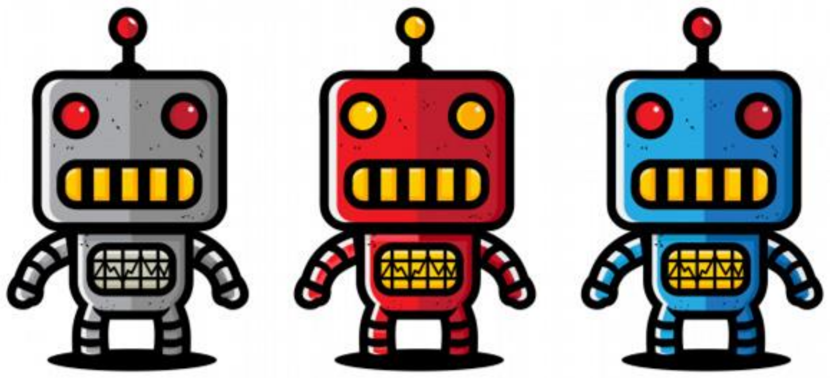
Source: Frank van Harmelen, "On the nature of AI, and on the relation between symbolic and statistical AI", https://www.slideshare.net/Frank.van.Harmelen/panel-on-the-future-of-ai-at-the-flamish-royal-society-of-sciences-and-arts

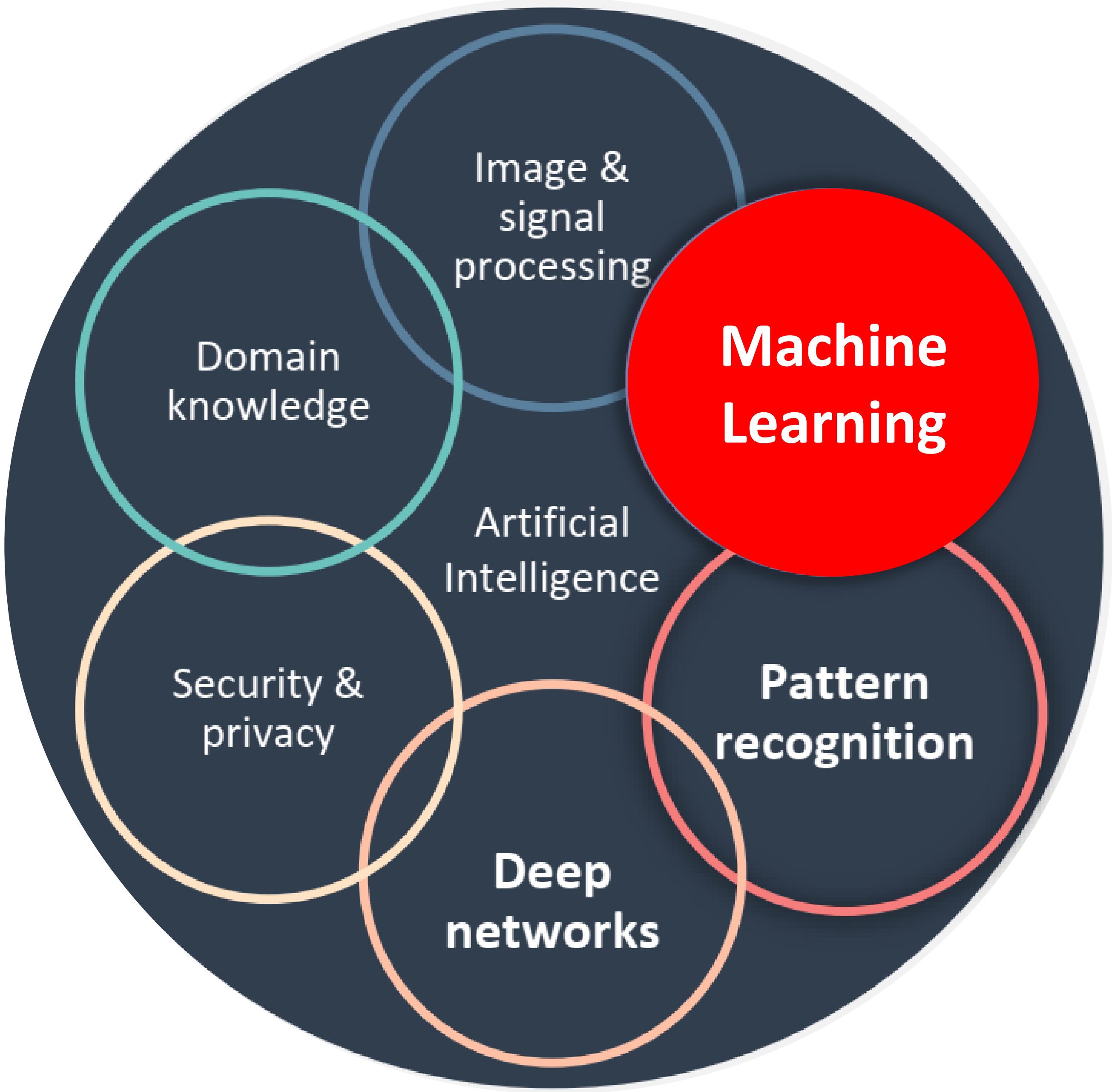
The two big paradigms of AI: **Application examples**

Symbolic AI	Statistical AI
• Planning (autonomous space mission) • Reasoning (diagnosis, design, decision support) • Language generation (conversations) • Search engines	• Pattern recognition (images, sound, shapes) • Speech generation (sound) • Motor skills (robots) • Search engines

Source: Frank van Harmelen, “On the nature of AI, and on the relation between symbolic and statistical AI”, https://www.slideshare.net/Frank.van.Harmelen/panel-on-the-future-of-ai-at-the-flamish-royal-society-of-sciences-and-arts

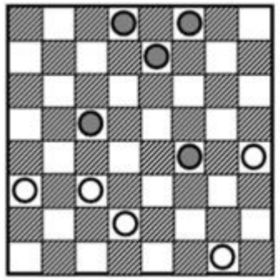
Artificial Intelligence:
Many Facets





Taken from Slide "Pattern Recognition: Statistics to Deep Networks" Anil K. Jain,Michigan State University, ICACSIS-IWBIS 2020

History of **Machine Learning**

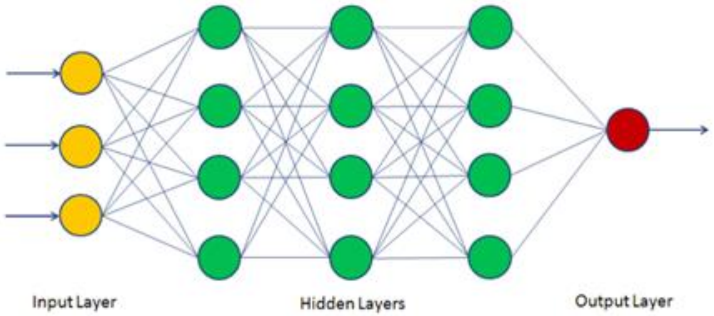


1950s

- Samuel's checker player
- Selfridge's Pandemonium

- Neural networks: Perceptron
- Pattern recognition
- Learning in the limit theory
- Minsky and Papert prove limitations of Perceptron

1960s

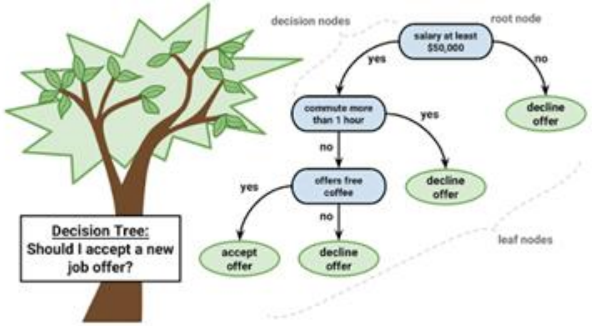


1970s

- Symbolic concept induction
- Winston's arch learner
- Expert systems and the knowledge acquisition bottleneck
- Quinlan's ID3
- Michalski's AQ and soybean diagnosis
- Scientific discovery with BACON
- Mathematical discovery with AM

Slide credit: Ray Mooney

History of **Machine Learning** (cont.)



1980s

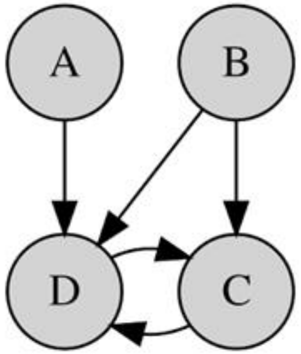
- Advanced decision tree and rule learning
- Explanation-based Learning (EBL)
- Learning and planning and problem solving
- Utility problem
- Analogy
- Cognitive architectures
- Resurgence of neural networks (connectionism, backpropagation)
- Valiant's PAC Learning Theory
- Focus on experimental methodology

1990s

- Data mining
- Adaptive software agents and web applications
- Text learning
- Reinforcement learning (RL)
- Inductive Logic Programming (ILP)
- Ensembles: Bagging, Boosting, and Stacking
- Bayes Net learning

Slide credit: Ray Mooney

History of **Machine Learning** (cont.)



2000s

- Support vector machines & kernel methods
- Graphical models
- Statistical relational learning
- Transfer learning
- Sequence labeling
- Collective classification and structured outputs
- Computer Systems Applications (Compilers, Debugging, Graphics, Security)
- E-mail management
- Personalized assistants that learn
- Learning in robotics and vision

2010s

- Deep learning systems
- Learning for big data
- Bayesian methods
- Multi-task & lifelong learning
- Applications to vision, speech, social networks, learning to read, etc.

Slide credit: Ray Mooney

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.

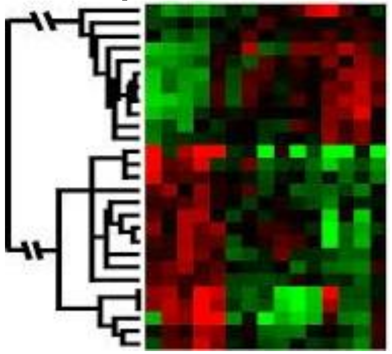
A well-defined learning task is given by <P, T, E>.

When Do We **Use Machine Learning**?

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)

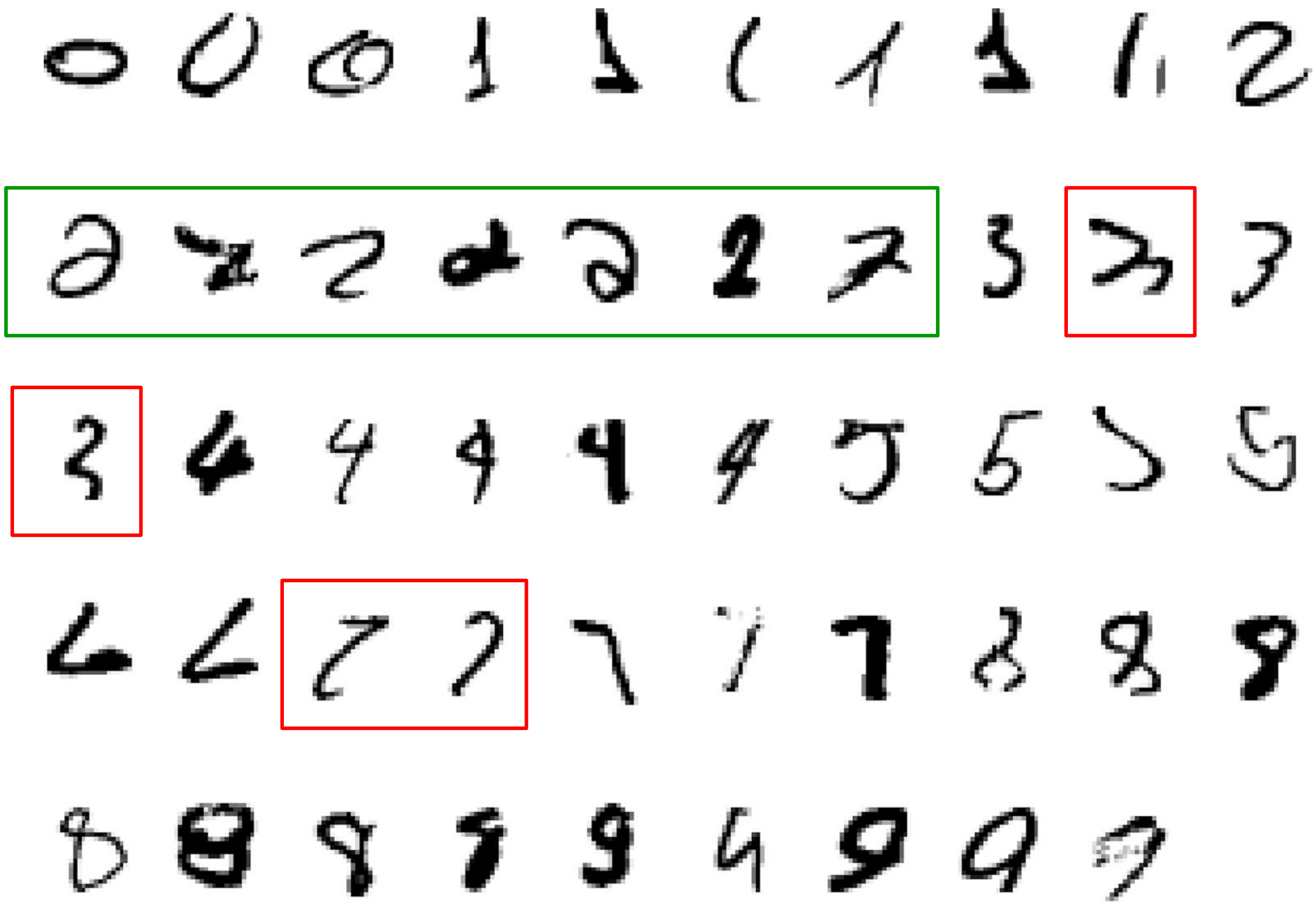


Learning isn’t always useful:

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

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



Slide credit: Geoffrey Hinton

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

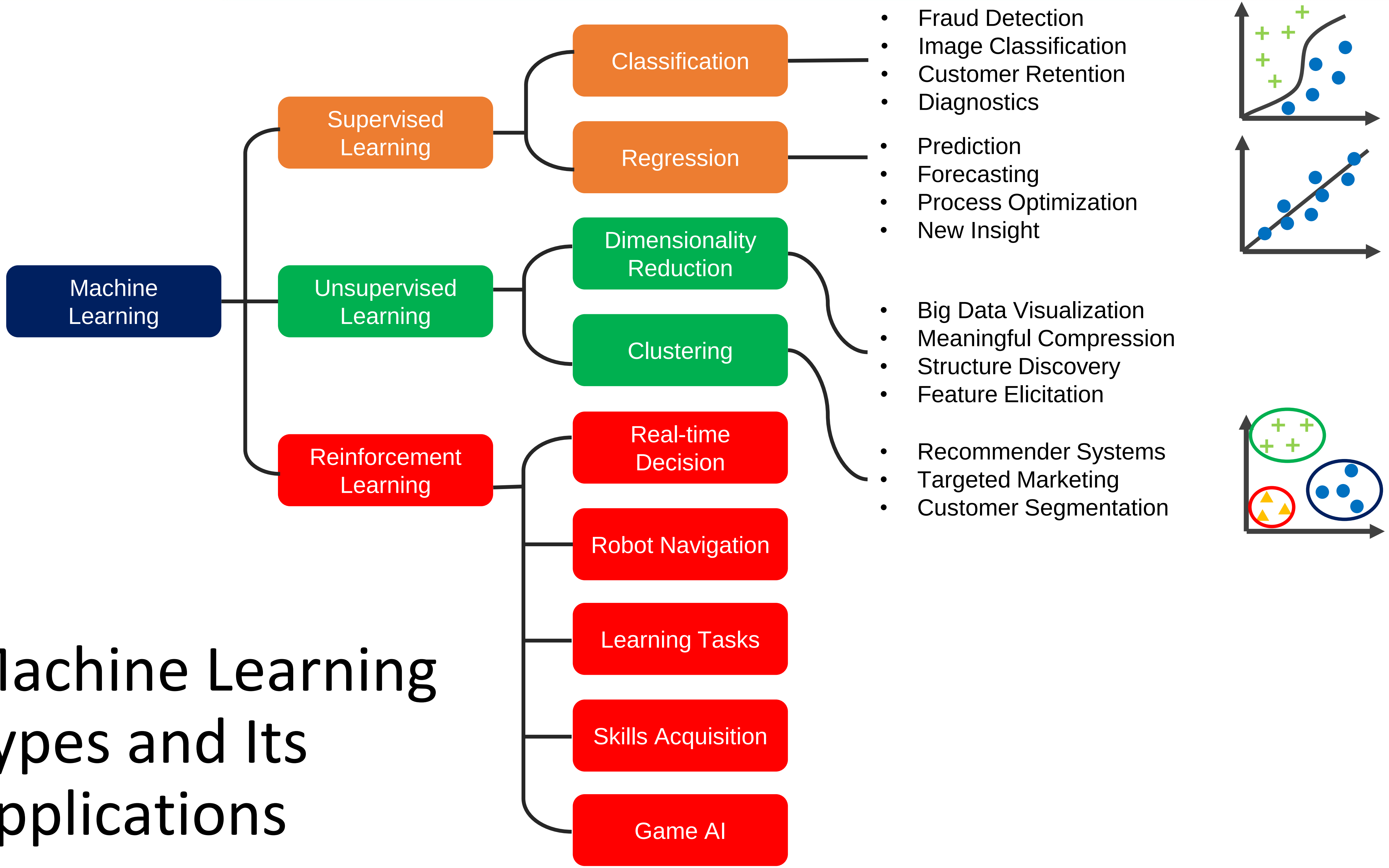
Slide credit: Geoffrey Hinton

Types of Learning

Supervised (Inductive) Learning	Given: Training data + desired outputs (labels)
Unsupervised Learning	Given: Training data (without desired outputs)
Semi-supervised Learning	Given: Training data + a few desired outputs (labels)
Reinforcement Learning	Rewards from sequence of actions

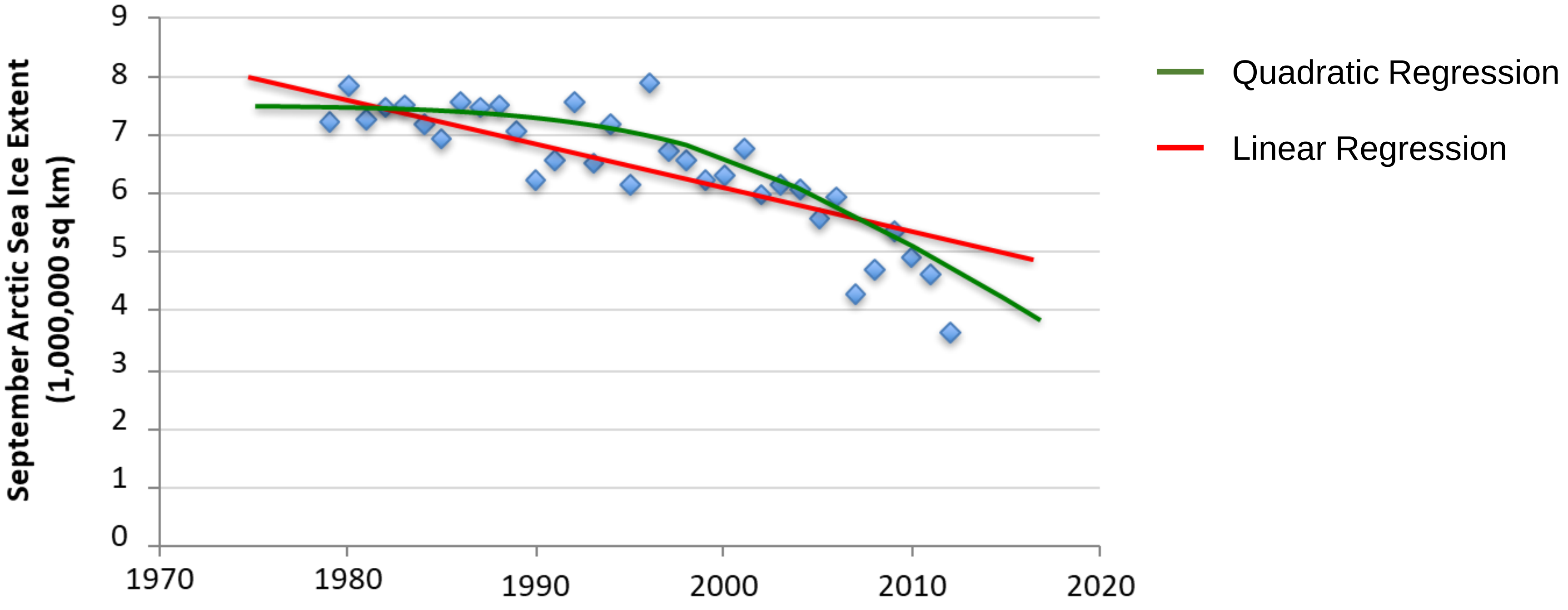
Based on slide by Pedro Domingos

Machine Learning Types and Its Applications



Supervised Learning: Regression

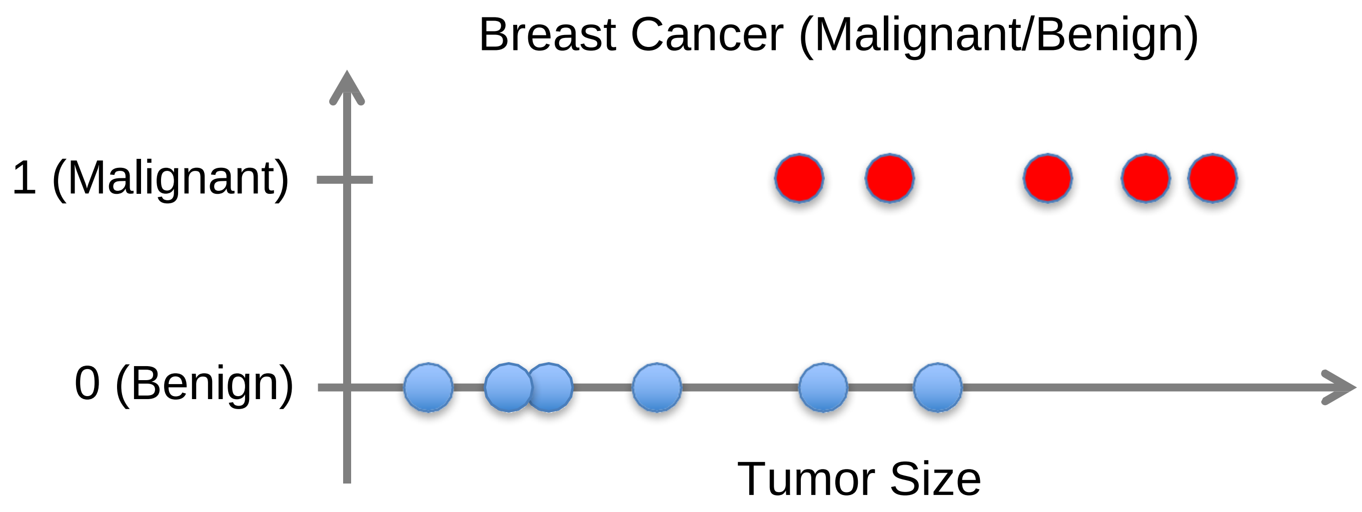
- Given $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$
- Learn a function $f(x)$ to predict y given x
  - y is real-valued == regression



Data from G. Witt. Journal of Statistics Education, Volume 21, Number 1 (2013)

Supervised Learning: Classification

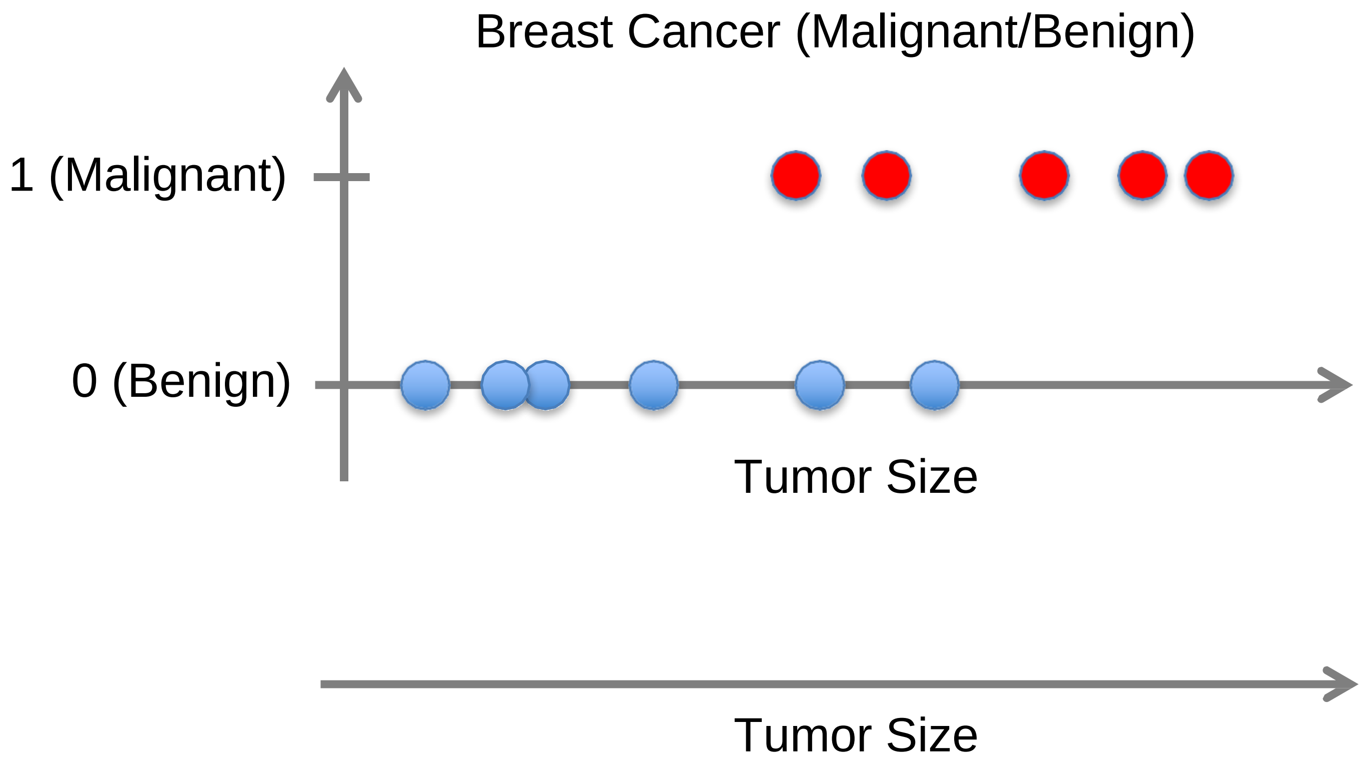
- Given $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$
- Learn a function $f(x)$ to predict y given x
  - y is categorical == classification



Based on example by Andrew Ng

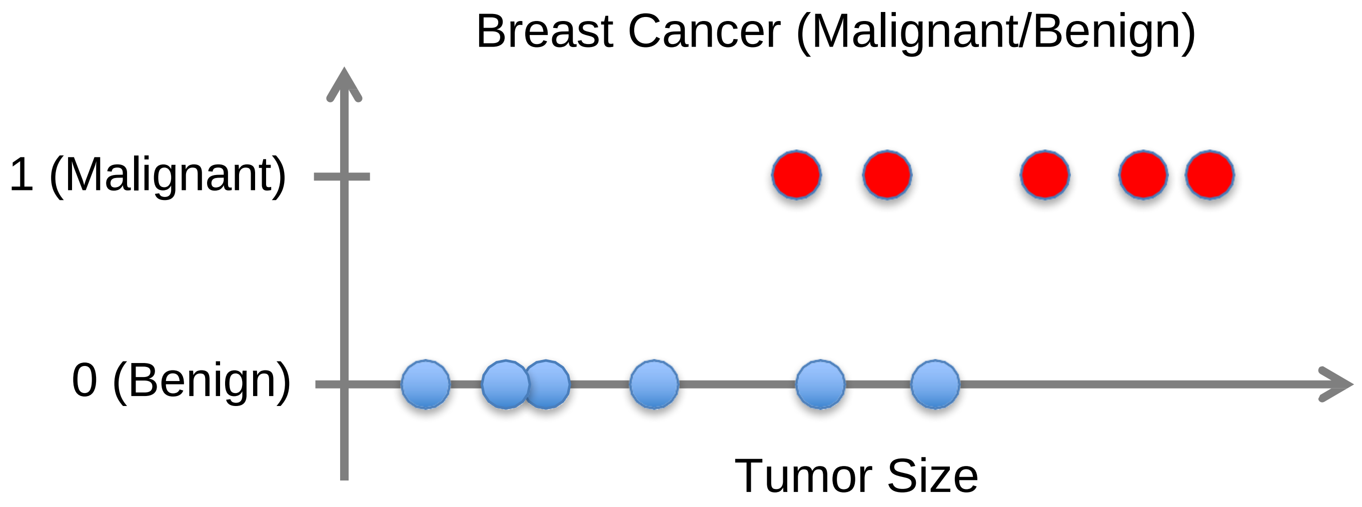
Supervised Learning: Classification

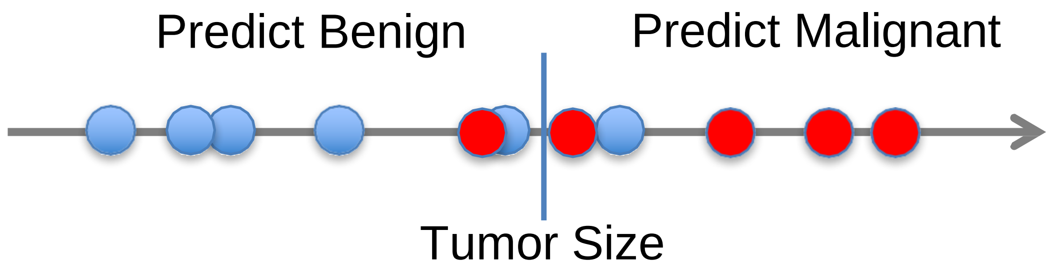
- Given $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$
- Learn a function $f(x)$ to predict y given x
  - y is categorical == classification



Supervised Learning: Classification

- Given $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$
- Learn a function $f(x)$ to predict y given x
  - y is categorical == classification



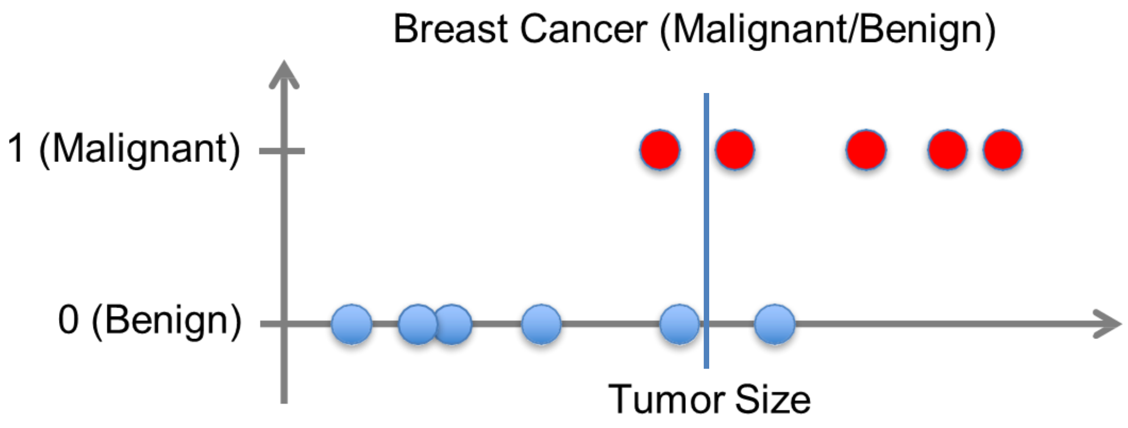


Based on example by Andrew Ng

Supervised Learning

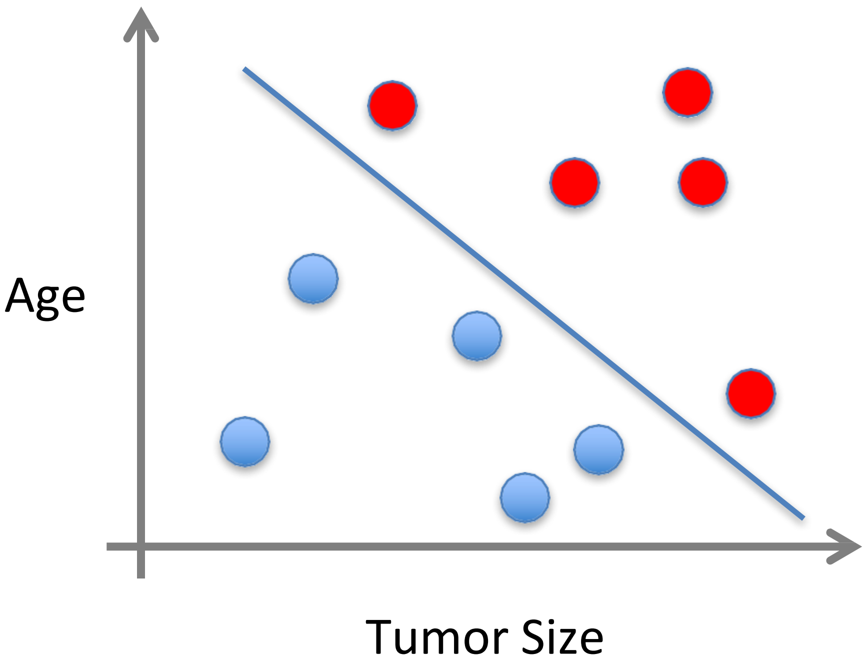
- x can be multi-dimensional
  - Each dimension corresponds to an attribute

1 Dimension



*there are some errors in classification

2 Dimension



*no errors in classification

Based on example by Andrew Ng

Supervised Learning: Example

Arrhytmia Classification using Fuzzy-Neuro Generalized Learning Vector Quantization

I Made Agus Setiawan[1], Elly M. Imah[2], and Wisnu Jatmiko[3]
[1]Computer Science Department, Udayana University, Bali, Indonesia
[2]Department of Mathematic, State University of Surabaya, Indonesia
[3]Faculty of Computer Science, Universitas Indonesia, Depok, Indonesia

E-mail: madeagus@unud.ac.id, wisnuj@cs.ui.ac.id

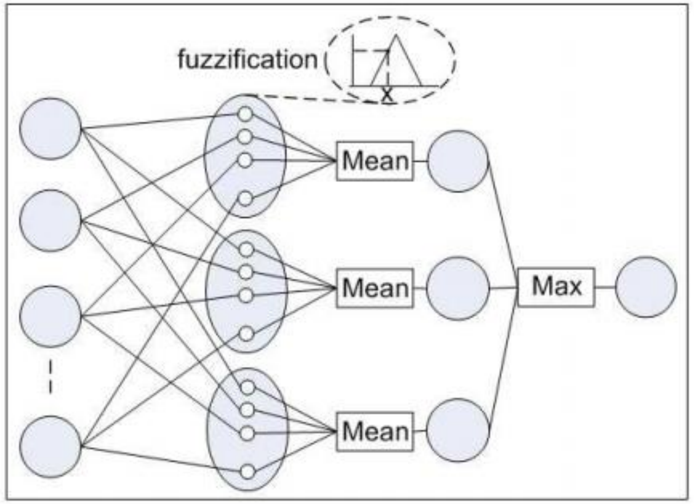
***Abstract*—Automatic heart beats classification has attracted much interest for research recently and we are interested to determine the type of arrhythmia from electrocardiogram (ECG) signal automatically. This paper will discuss a new extension of GLVQ that employ fuzzy logic concept as the discriminant function in order to develop a robust algorithm and improve the classification performance. The overall classification system is comprised of three components including data preprocessing, feature extraction and classification. Data preprocessing related to how the initial data prepared, in this case, we cut the signal beat by beat using R peak as pivot point, while for the feature extraction, we used wavelet algorithm. The ECG signals were obtained from MIT-BIH arrhythmia database. Our experiment showed that our proposed method, FN-GLVQ, was able to increase the accuracy of classifier compared with original GLVQ that used euclidean distance. By using *10-Fold Cross Validation*, the algorithm produced an average accuracy 93.36% and 95.52% , respectively for GLVQ and FNGLVQ.**

I. Introduction

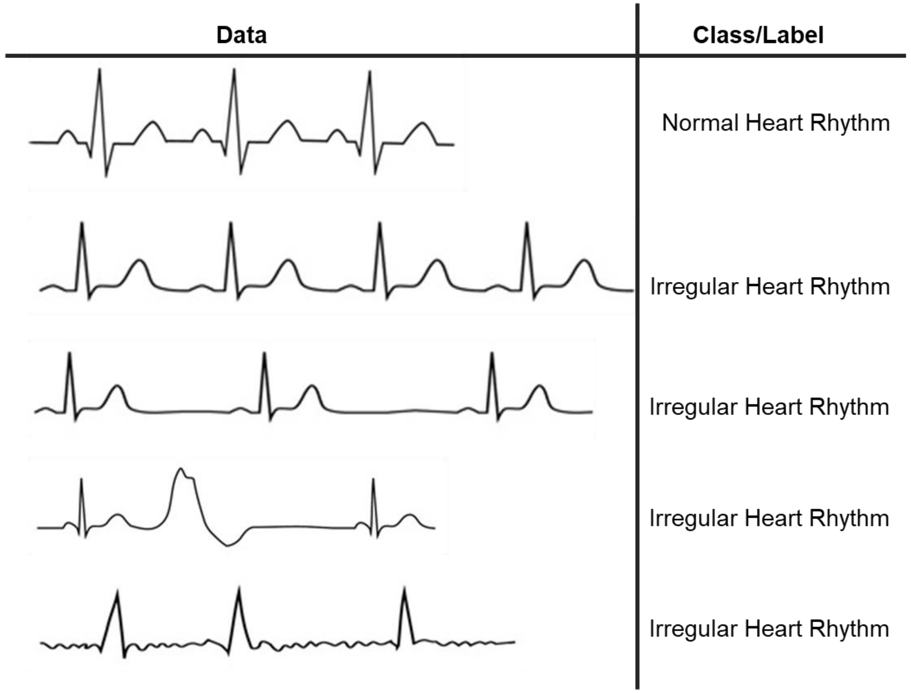
combining with Genetic Algorithm, like Nasiri doing [9] or combining with Particle Swarm Optimization (PSO) like Melgani works [10]. Ghongade et.al make a comparison for many feature extraction method like DFT, PCA, DWT, Morphological based and integrating it with ANN classifier [11].

So far, many method have been developed for arrhythmia detection, recognition and classification, however many method for arrhythmia beat classification have yet been able to handle noisy classification data. In our previous study we applied Fuzzy Neuro Learning Vector Quantization (FNLVQ) to solve unknown or un-categorical beat [12], [13], however the study was limited only six classes.

In this paper, we propose a new extension of Generalized LVQ by A.Sato [14] that employ fuzzy logic concept for the discriminant function in order to develop a robust algorithm and improve the classification performance. This paper is organized as follows. In section II, we describe pre-processing technique to re-



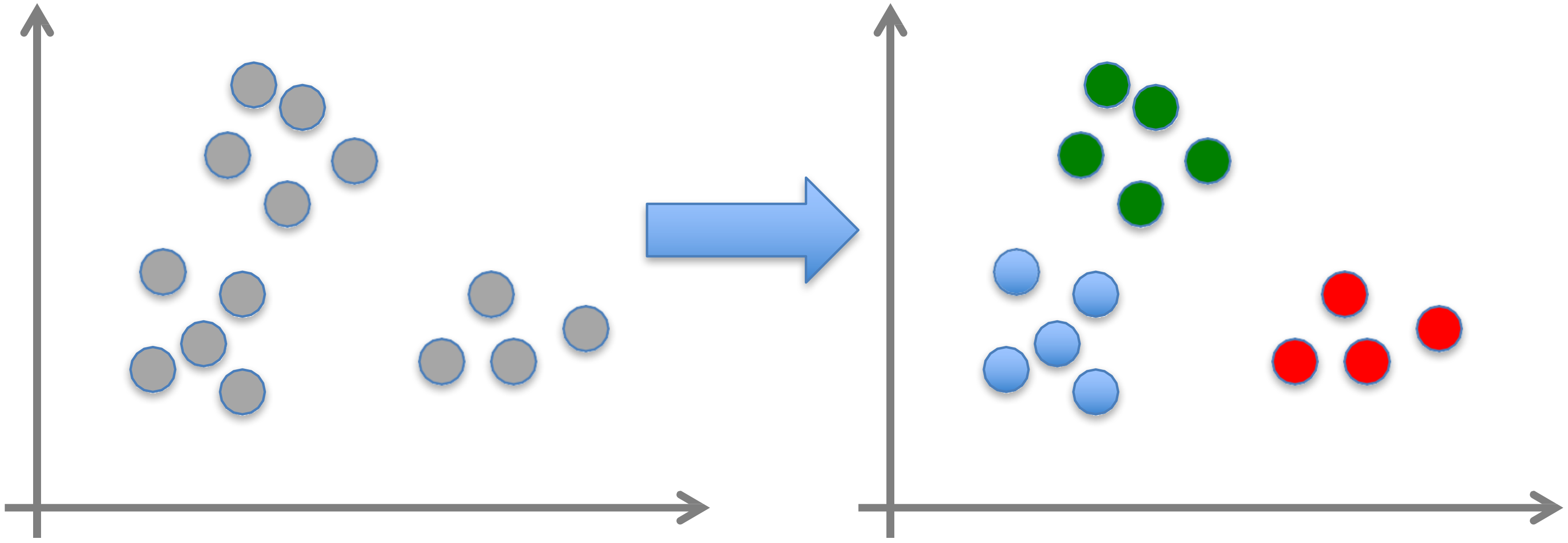
FNGLVQ Algorithm



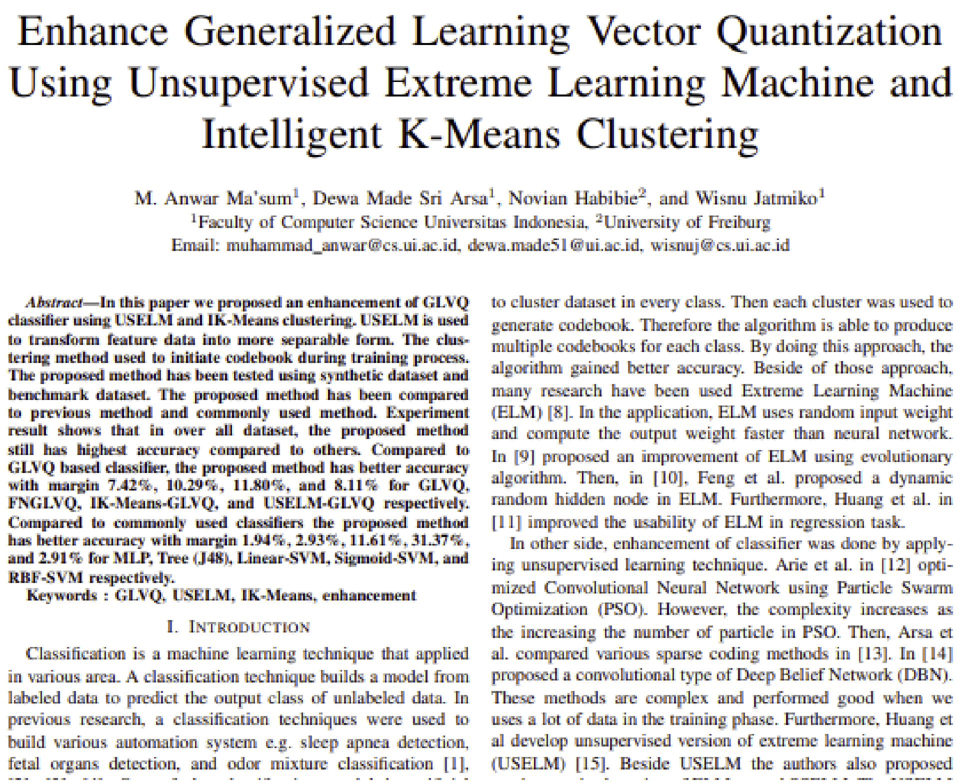
Classification

Unsupervised Learning

- Given $x_1, x_2, \ldots, x_\mathrm{n}$ (without labels)
- Output hidden structure behind the x's
  - E.g., clustering



Unsupervised Learning: Example



Enhance Generalized Learning Vector Quantization Using Unsupervised Extreme Learning Machine and Intelligent K-Means Clustering

M. Anwar Ma'sum[1], Dewa Made Sri Arsa[1], Novian Habibie[2], and Wisnu Jatmiko[1]
[1]Faculty of Computer Science Universitas Indonesia, [2]University of Freiburg
Email: muhammad_anwar@cs.ui.ac.id, dewa.made51@ui.ac.id, wisnuj@cs.ui.ac.id

***Abstract*—In this paper we proposed an enhancement of GLVQ classifier using USELM and IK-Means clustering. USELM is used to transform feature data into more separable form. The clustering method used to initiate codebook during training process. The proposed method has been tested using synthetic dataset and benchmark dataset. The proposed method has been compared to previous method and commonly used method. Experiment result shows that in over all dataset, the proposed method still has highest accuracy compared to others. Compared to GLVQ based classifier, the proposed method has better accuracy with margin 7.42%, 10.29%, 11.80%, and 8.11% for GLVQ, FNGLVQ, IK-Means-GLVQ, and USELM-GLVQ respectively. Compared to commonly used classifiers the proposed method has better accuracy with margin 1.94%, 2.93%, 11.61%, 31.37%, and 2.91% for MLP, Tree (J48), Linear-SVM, Sigmoid-SVM, and RBF-SVM respectively.**

Keywords : GLVQ, USELM, IK-Means, enhancement

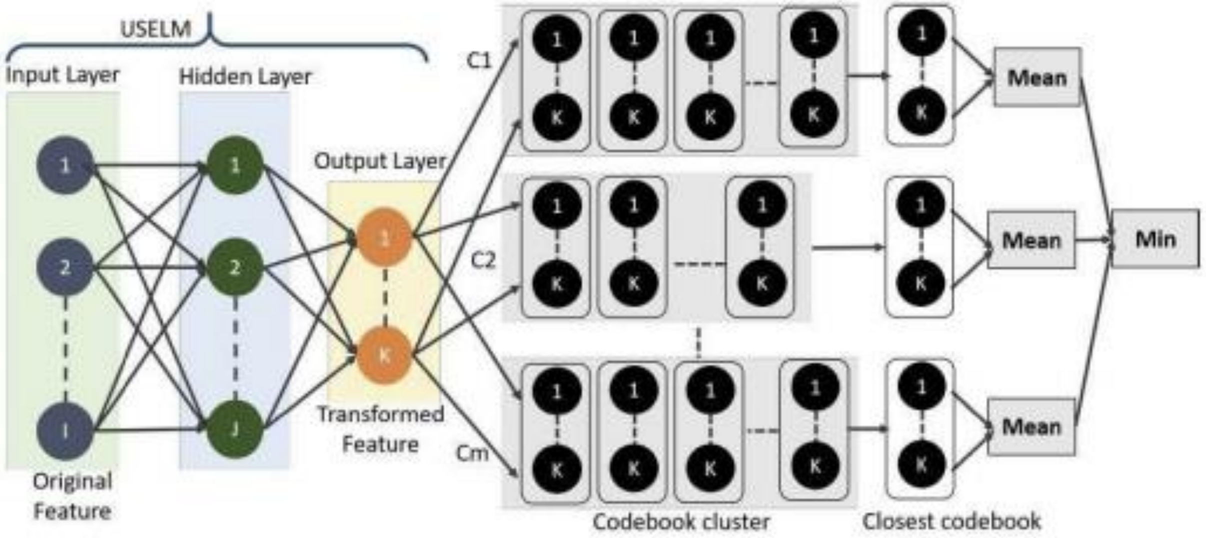
I. INTRODUCTION

Classification is a machine learning technique that applied in various area. A classification technique builds a model from labeled data to predict the output class of unlabeled data. In previous research, a classification techniques were used to build various automation system e.g. sleep apnea detection, fetal organs detection, and odor mixture classification [1],

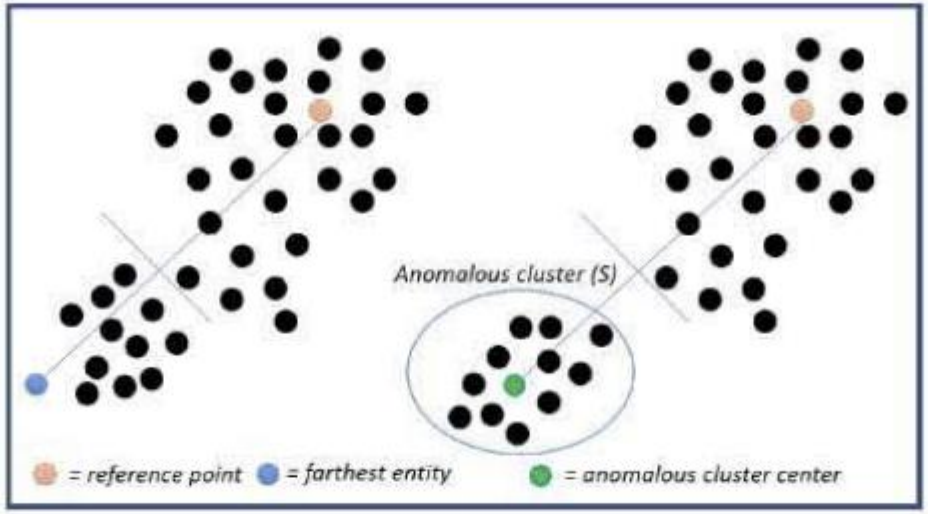
to cluster dataset in every class. Then each cluster was used to generate codebook. Therefore the algorithm is able to produce multiple codebooks for each class. By doing this approach, the algorithm gained better accuracy. Beside of those approach, many research have been used Extreme Learning Machine (ELM) [8]. In the application, ELM uses random input weight and compute the output weight faster than neural network. In [9] proposed an improvement of ELM using evolutionary algorithm. Then, in [10], Feng et al. proposed a dynamic random hidden node in ELM. Furthermore, Huang et al. in [11] improved the usability of ELM in regression task.

In other side, enhancement of classifier was done by applying unsupervised learning technique. Arie et al. in [12] optimized Convolutional Neural Network using Particle Swarm Optimization (PSO). However, the complexity increases as the increasing the number of particle in PSO. Then, Arsa et al. compared various sparse coding methods in [13]. In [14] proposed a convolutional type of Deep Belief Network (DBN). These methods are complex and performed good when we uses a lot of data in the training phase. Furthermore, Huang et al develop unsupervised version of extreme learning machine (USELM) [15]. Beside USELM the authors also proposed

The Algorithm



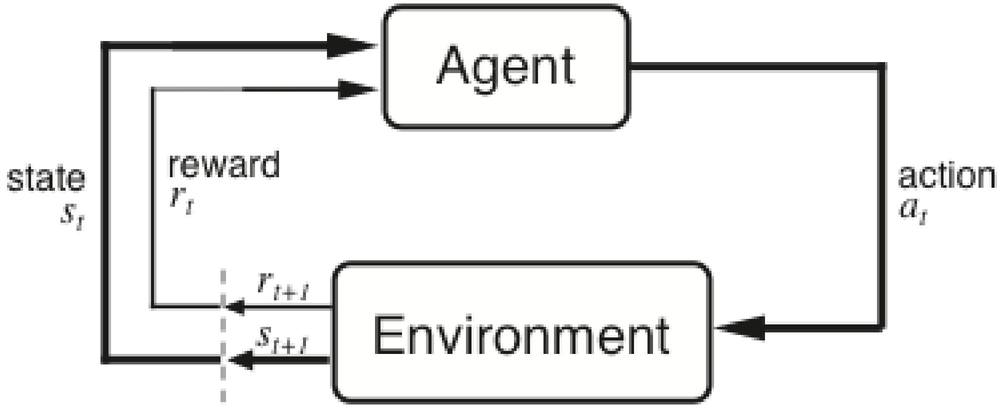
Clustering



Reinforcement Learning

- Given a sequence of states and actions with (delayed) rewards, output a policy
  - Policy is a mapping from states → actions that tells you what to do in a given state
- Examples:
  - Credit assignment problem
  - Game playing
  - Robot in a maze
  - Balance a pole on your hand

The Agent-Environment Interface



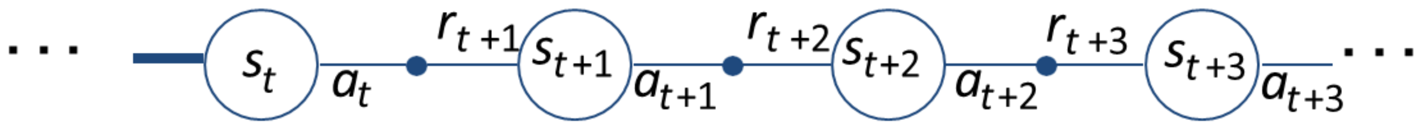
Agent and environment interact at discrete time steps : $t = 0, 1, 2, K$

Agent observes state at step t : $s_t \in S$

produces action at step t : $a_t \in A(s_t)$

gets resulting reward : $r_{t+1} \in \Re$

and resulting next state : s_{t+1}



Slide credit: Sutton & Barto

Reinforcement Learning

https://www.youtube.com/watch?v=4cgWya-wjgY

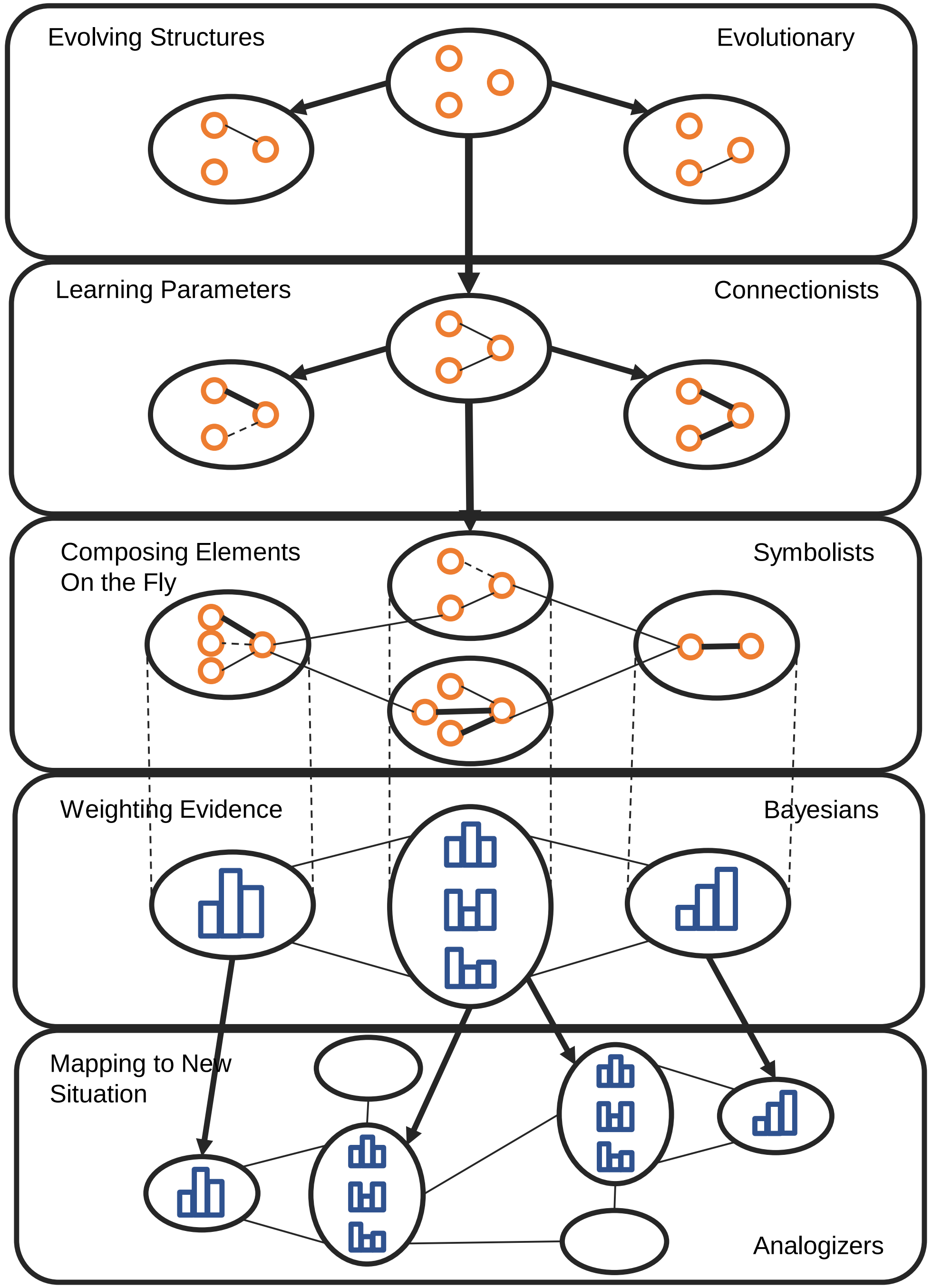
Machine Learning Evaluation

- Accuracy
- Precision and recall
- Squared error
- Likelihood
- Posterior
- probability
- Cost / Utility
- Margin
- Entropy
- K-L divergence
- etc.

Slide credit: Pedro Domingos

Five Tribes of Machine Learning

- **Evolutionary**
  - are interested in evolving structure. They are influenced by biology.
  - Signature Technology : **genetic algorithms, evolutionary programming, and evolutionary game theory**
- **Connectionists**
  - **use neural networks**. They are influenced by neuroscience.
  - Signature Technology : **deep learning technologies, including RNN, CNN, and deep reinforcement learning**
- **Symbolists**
  - **use formal systems**. They are influenced by computer science, linguistics, and analytic philosophy.
  - Signature Technology : **decision trees, production rule systems, and inductive logic programming**
- **Bayesians**
  - use **probabilistic inference**. They are influenced by statistics.
  - Signature Technology **:** Hidden Markov Models, graphical models, and causal inference
- **Analogizers**
  - are interested in mapping to new situations. They are influenced by psychology.
  - Signature Technology : **k-nearest neighbor, and support vector machines**

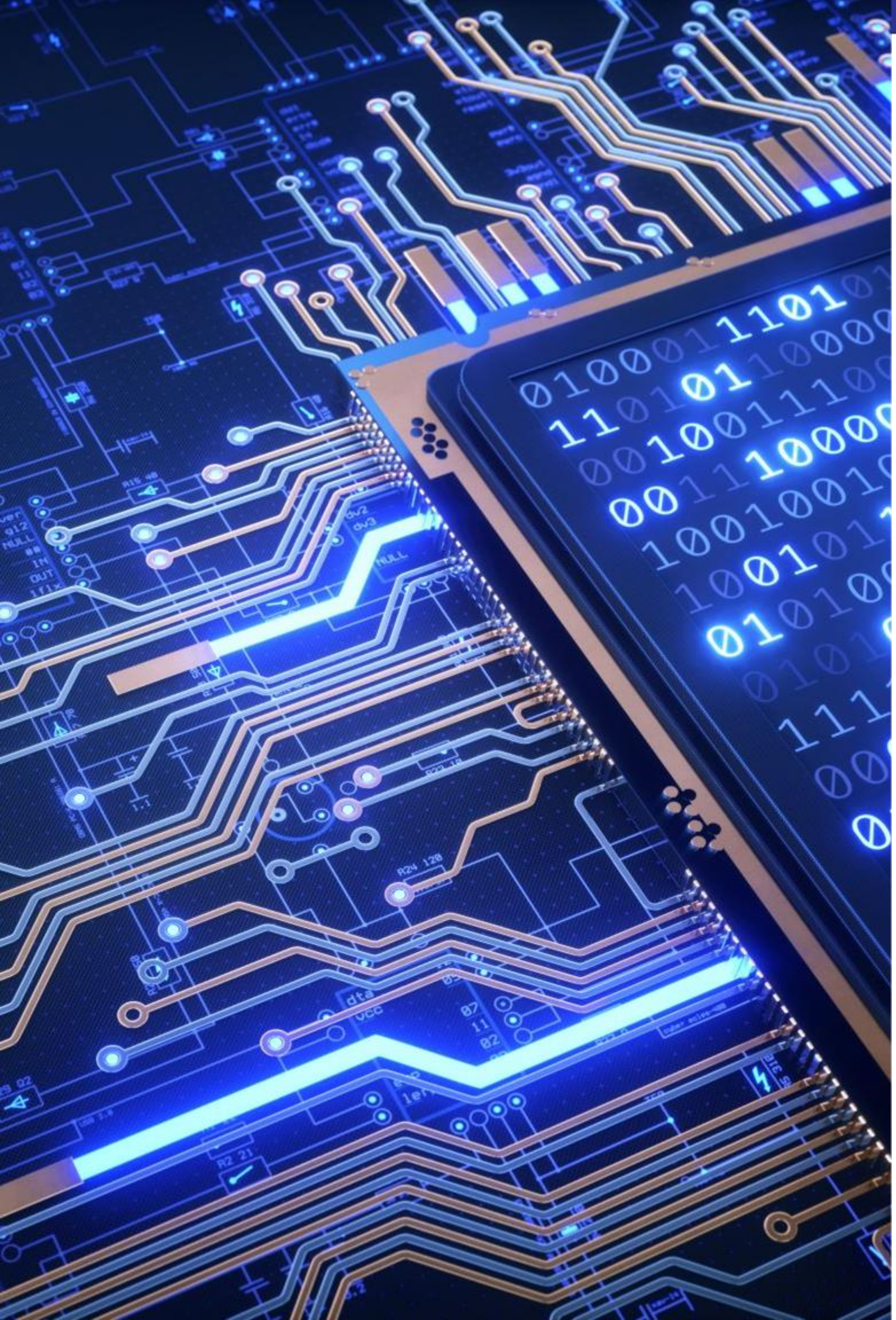


Domingos, Pedro. *The master algorithm: How the quest for the ultimate learning machine will remake our world*. Basic Books, 2015.

https://kevinbinz.com/2017/08/13/ml-five-tribes/

Reference

- AI – Digital Talent Schoolarship Kominfo 2019
- Domingos, Pedro. The master algorithm: How the quest for the ultimate learning machine will remake our world. Basic Books, 2015.
- Journal/ Conference Papers by Wisnu Jatmiko



The End of Presentation