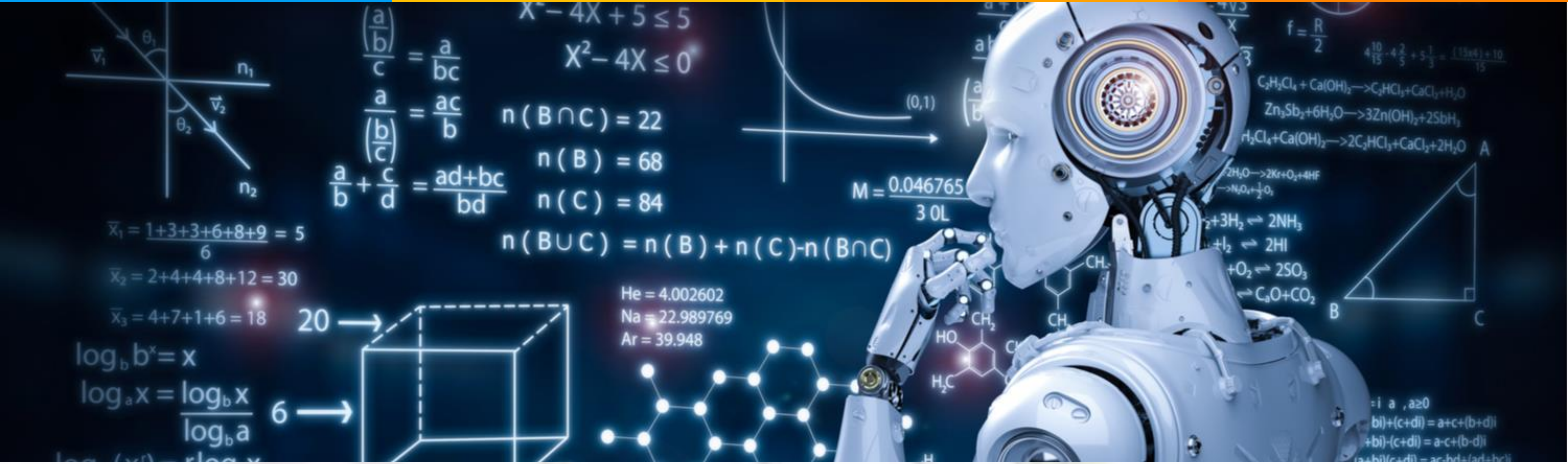
Duration: 60 min | Srikanth Shenoy | AML 5102 | M.E. AI & ML



Lecture 01 - Introduction to Machine Learning

1. Course plan & logistics

Logistics

- Course Name
  - Applied Machine Learning – AML 5102
- Hours per week: 3 lectures + 3 hours lab
- Total 36 hours theory + 36 hours lab

Course Objectives

- Understand/compare/choose between ML paradigm
- Supervised & Unsupervised ML, Feature Engg.
- My unofficial objectives -
  - Make you develop a passion for all things ML
  - Help you excel in theoretical & practical aspects of ML
  - Aid you in gaining employable skills
- Not comprehensive intro to popular software frameworks like numpy, pandas, sklearn, tensorflow, pytorch
  - Numerous videos available on youtube

Slight changes in the order

- Decision Trees goes towards the end (after LR)
- Added Perceptron – model based classification without optimization
- SVM moves into place after Linear Regression & Feature Engineering
- Ensemble methods will be covered in slightly more depth

Your objectives

Evaluation

- Memorization not important for exams
- Exam will test your
  - Core understanding & interdisciplinary thinking
  - Capability to apply Linear algebra to
    - Various scenarios & Machine Learning problems
- Objective type, True/False, Problems, 2-3 sentence
  - Negative marking for objective type, True/False
  - True/False will also need a 1-2 sentence reasoning.
  - Both have to be correct.

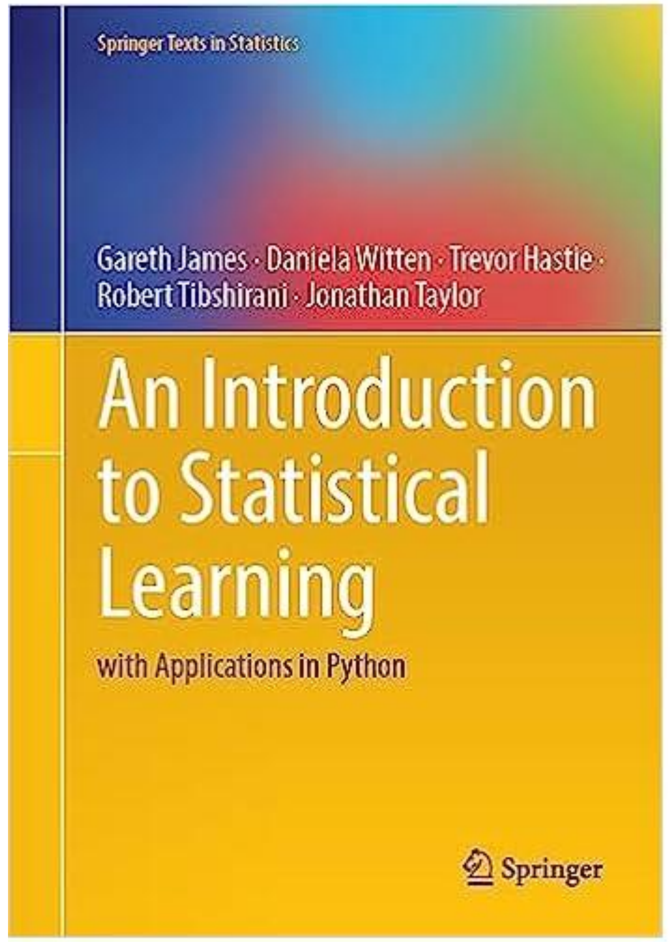
Grading policy

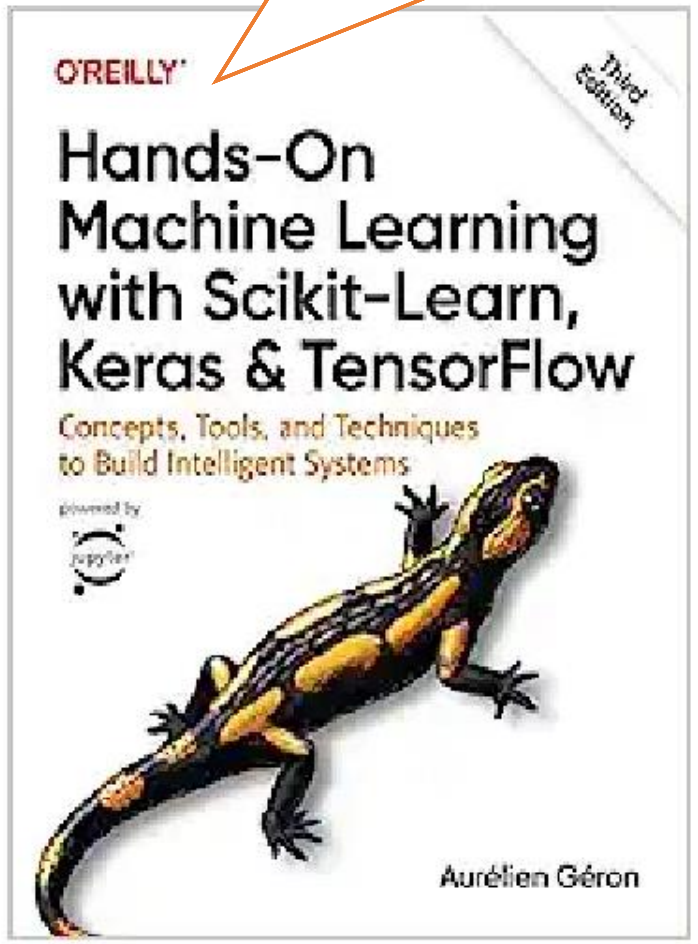
- 2 sessional + 1 end semester exam
- One big implication of chosen approach ☹
- Continuous assessment
  - 2-3 problem sets per semester
  - Many surprise quizzes
  - Daily summarization by randomly chosen student

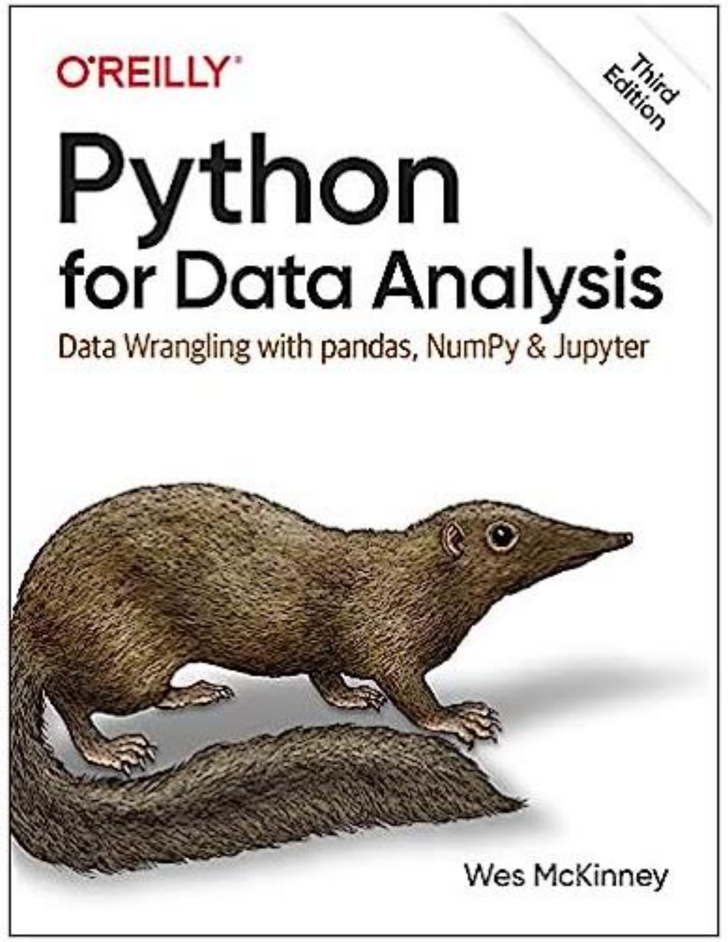
Textbook & References

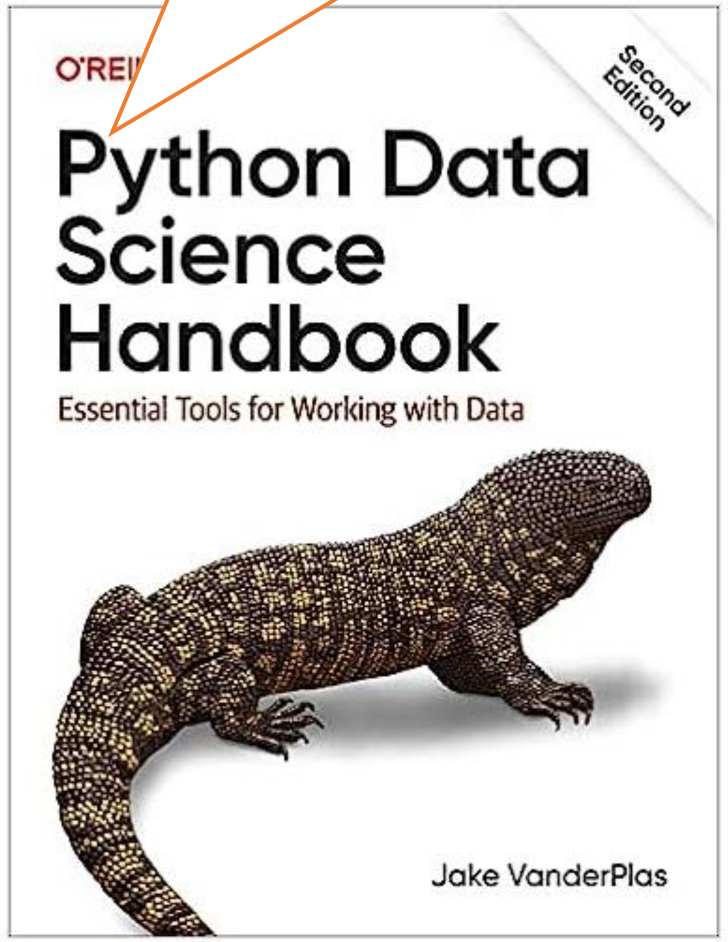
Helps in placements

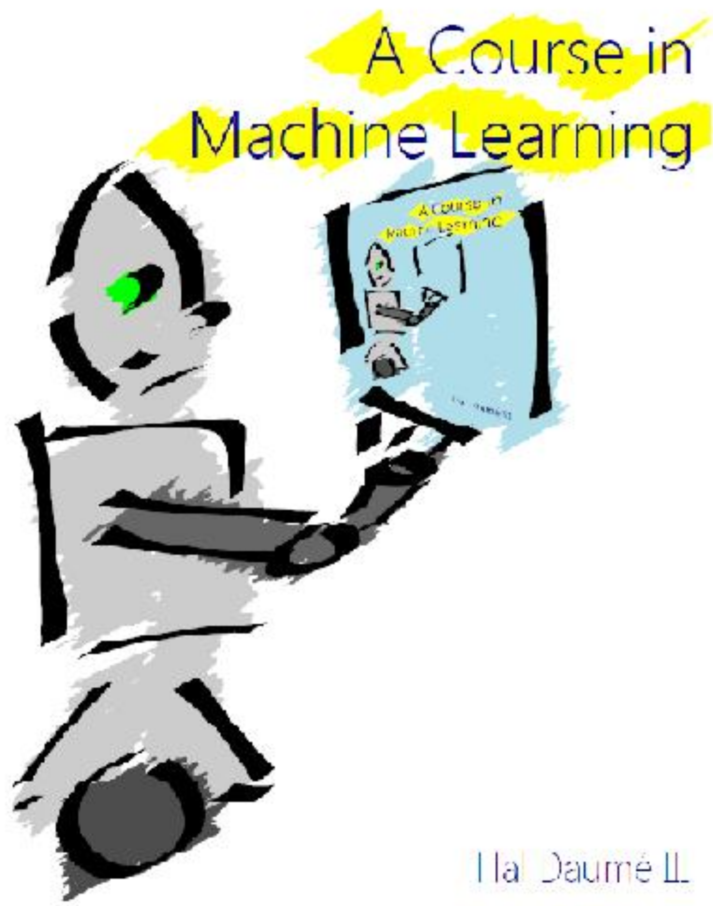
Good book for hands-on numpy, pandas & more







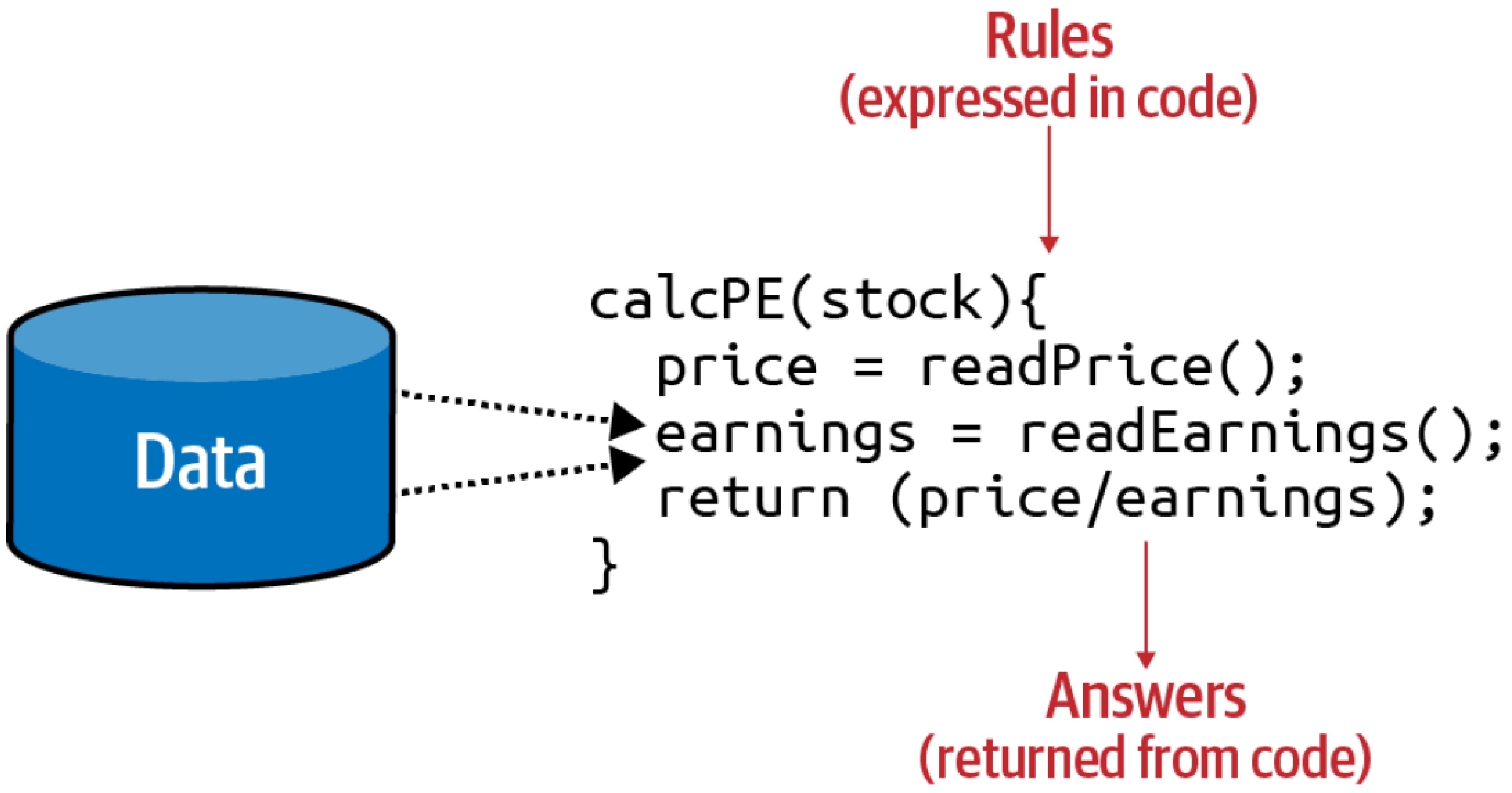




- Recommended, but none required
- Reading materials from various sources
- Mandatory and recommended videos
- For every lecture – Slides, Lecture Notes (for select topics)

2. What is Machine Learning

Traditional Programming



Solving it traditional way

Explicit coding will not work

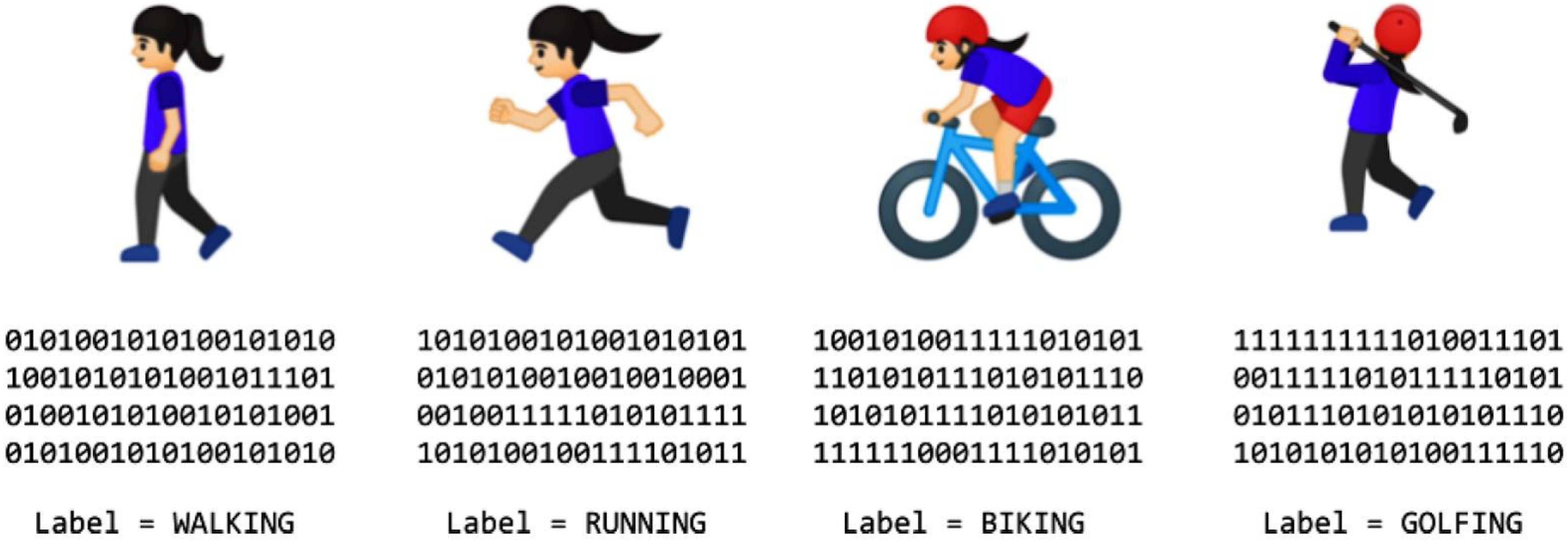
```java
public class Model {
    public static double score(double[] input) {
        double var0;
        if (input[4] <= 0.728433221578598) {
            if (input[0] <= 0.3851476162672043) {
                if (input[0] <= -0.8715016543865204) {
                    if (input[3] <= 0.629336342215538) {
                        if (input[0] <= -1.2120069861412048) {
                            if (input[2] <= -0.34459860622882843) {
                                if (input[1] <= -0.018694877624511172) {
                                    if (input[0] <= -1.438126027584076) {
                                        if (input[2] <= -1.5603920221328735) {
                                            if (input[0] <= -1.5781668424606323) {
                                                var0 = 1607.5101318359375;
                                            } else {
                                                if (input[2] <= -2.2775022983551025) {
                                                    var0 = 1727.7850341796875;
                                                } else {
                                                    var0 = 1728.89697265625;
```

Traditional approach can't scale

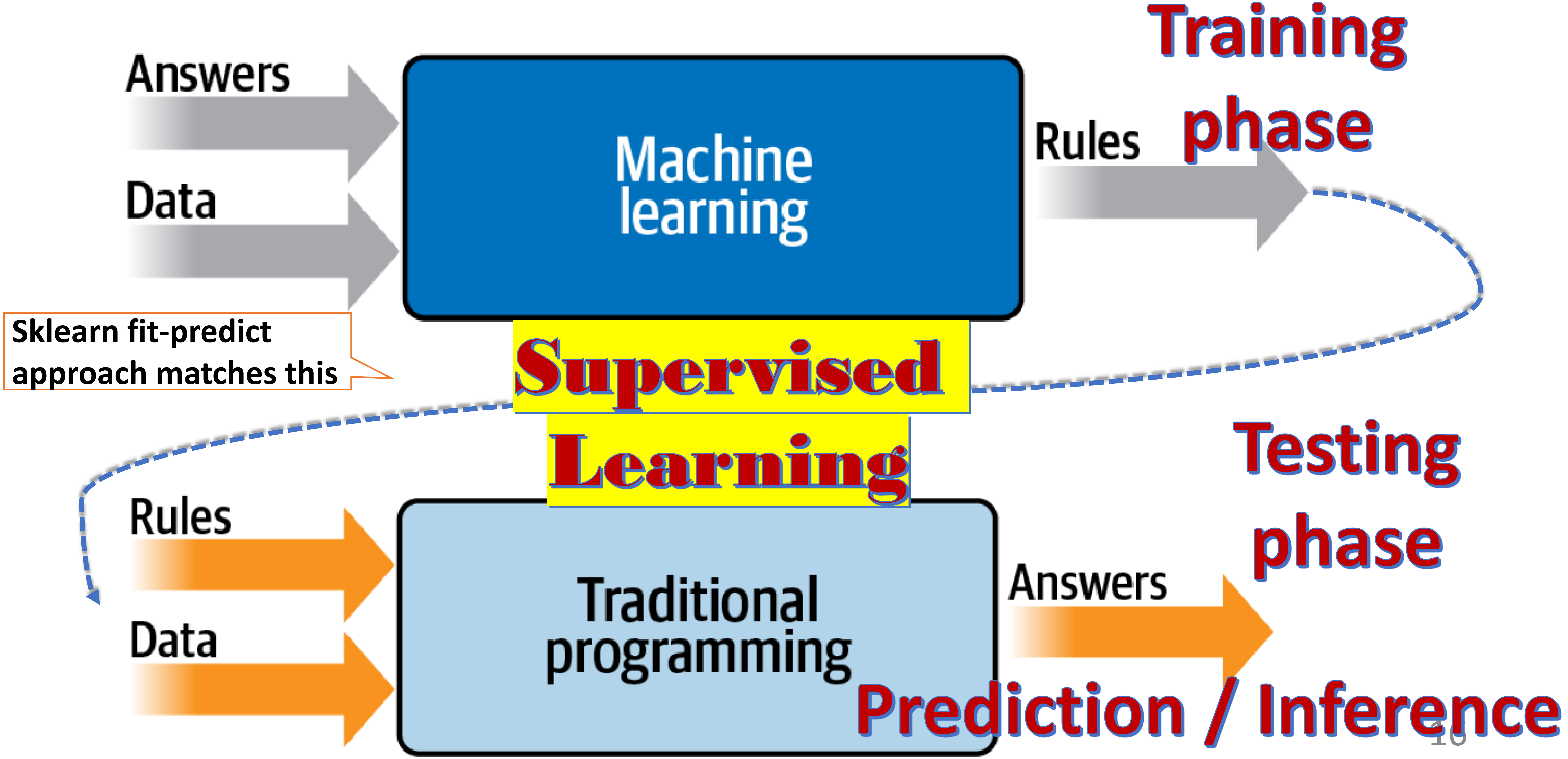
- Consider running activity. What variations exist?
  - Front pose, perspective, low lighting




Machine Learning approach



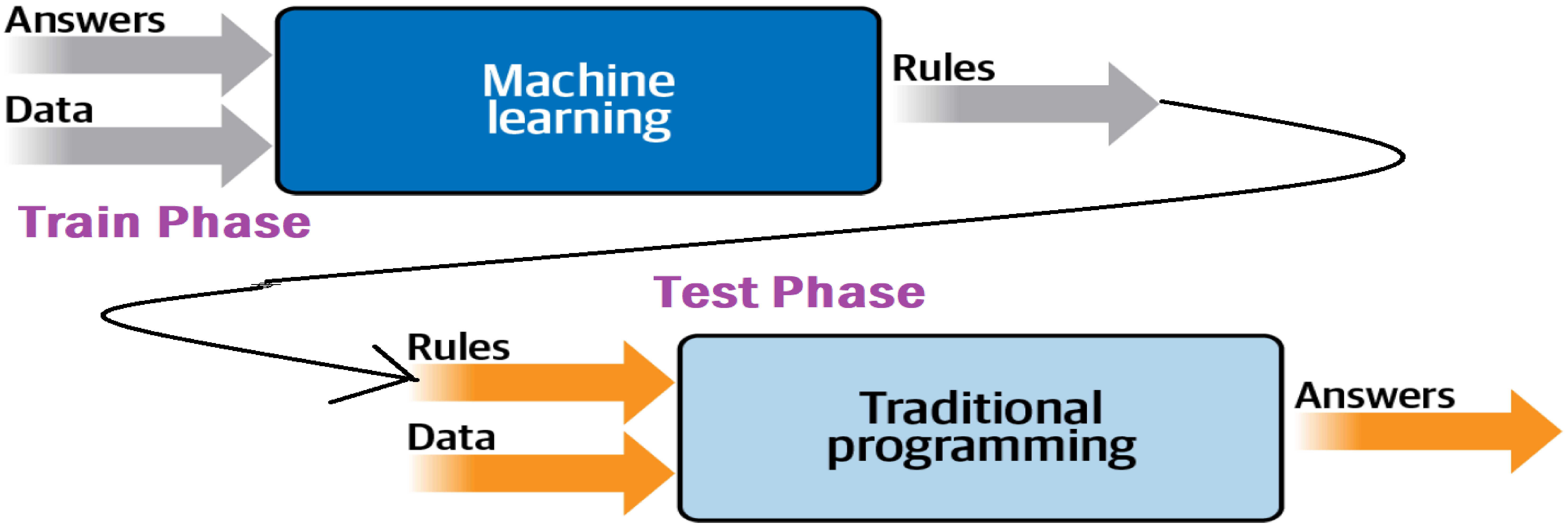
And then…


Training
phase
Answers
Data
Machine learning
Rules
Sklearn fit-predict approach matches this
Supervised
Learning
Testing
phase
Rules
Data
Traditional programming
Answers
Prediction / Inference

What format are the rules

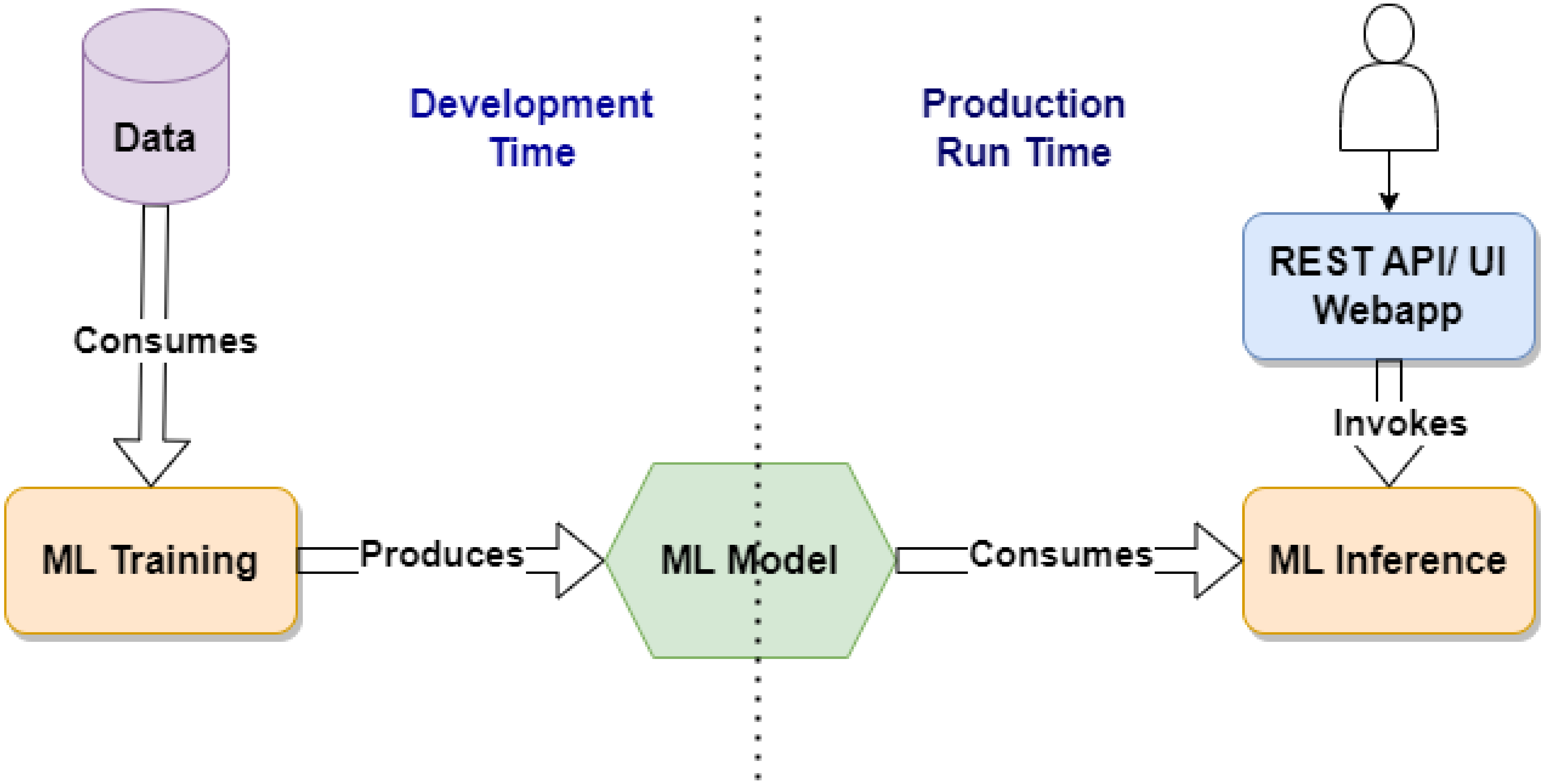
- In memory python object (tuned model)
- Pickle file
- Weights matrix stored as numpy, tf, pytorch
- NNEF
- ONNX – de-facto standard for exchanging deep learning models
- sklearn-onnx
- Mlflow
- Plain Ugly Code

Machine Learning is like exam



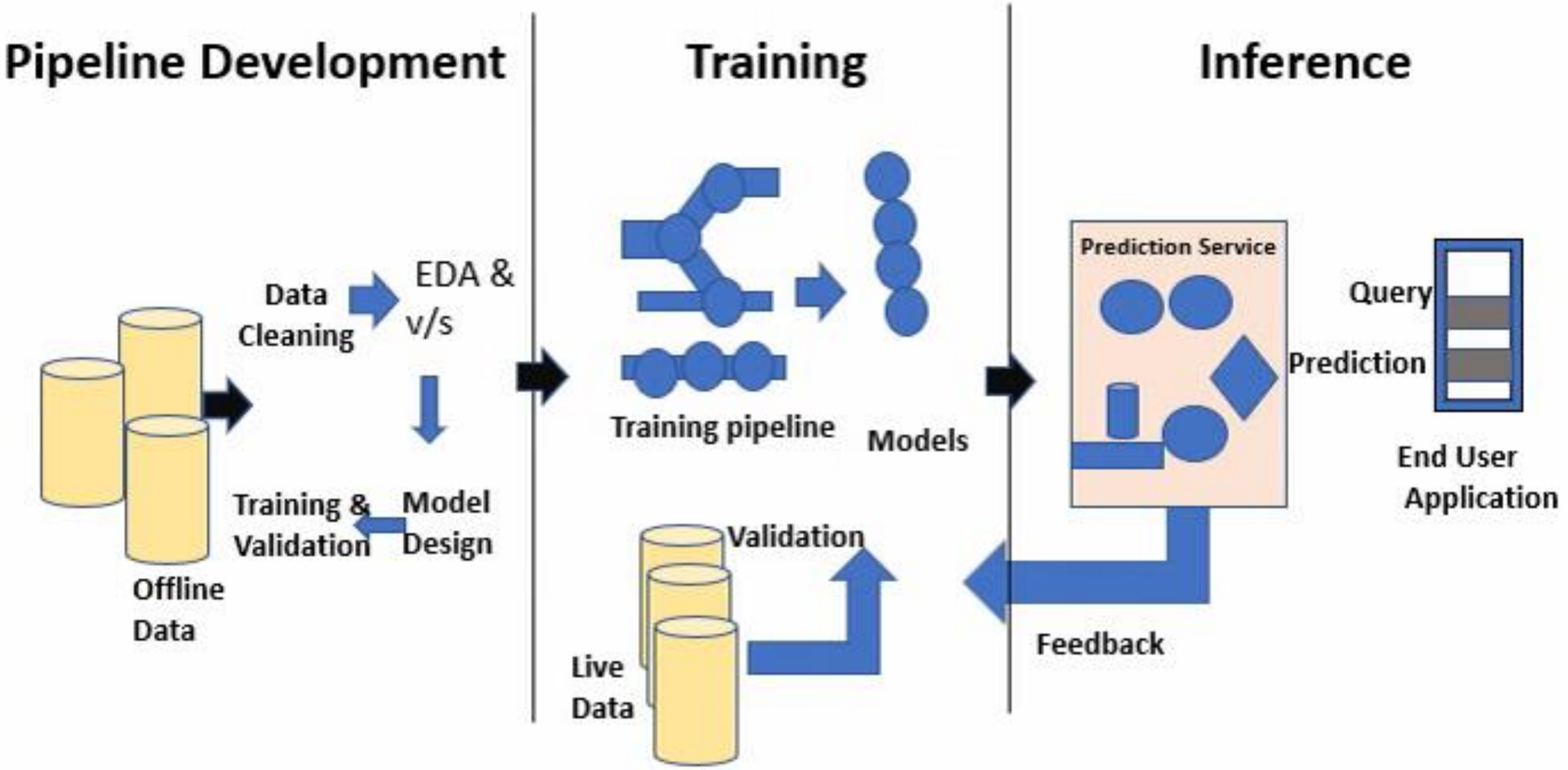
- Success is measured on
  - How one performs in test
  - Not how one practices

From development to production



Unsupervised Learning

Machine Learning lifecycle

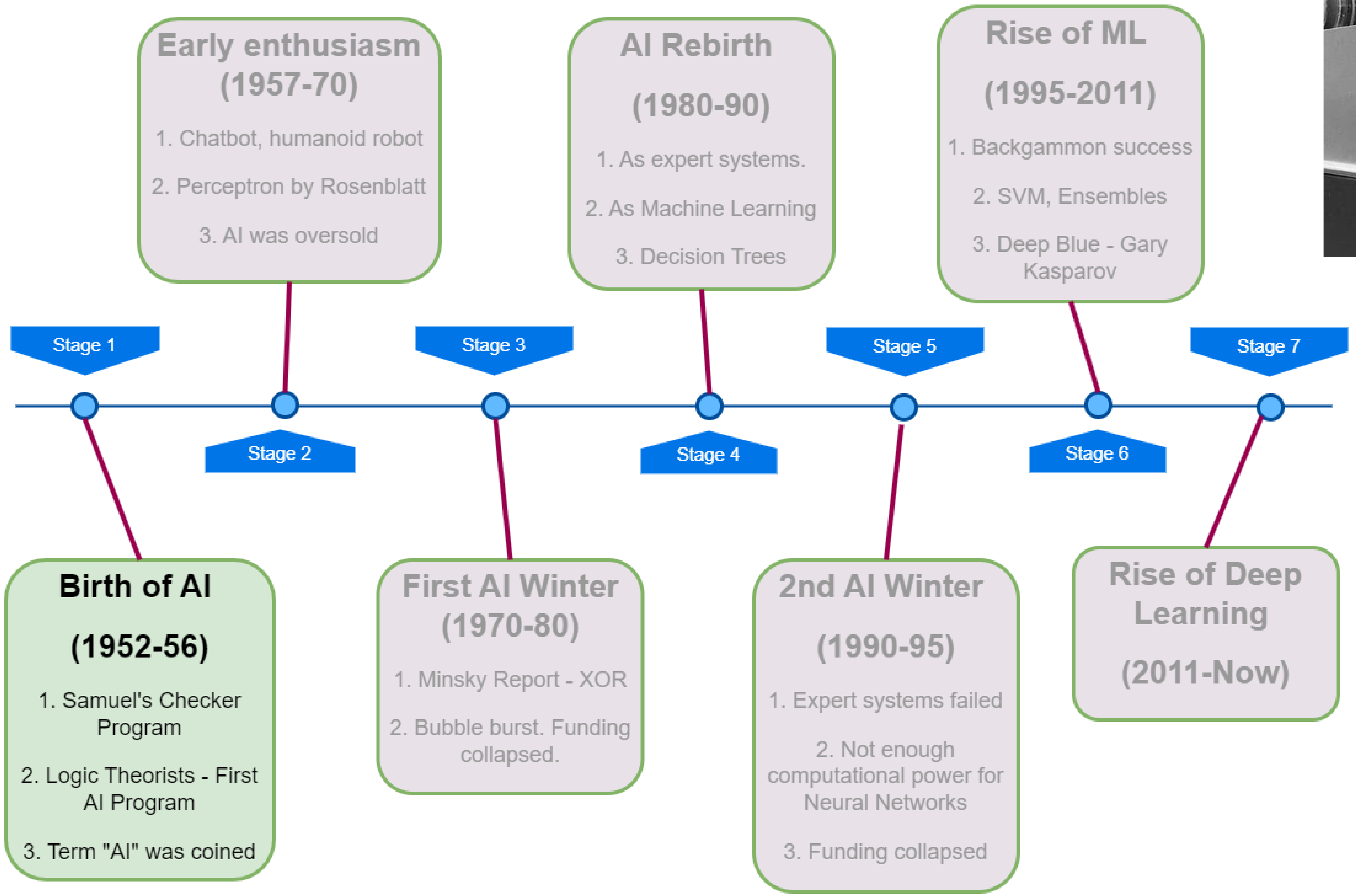


3. AI history from ML perspective

Goal of classic AI

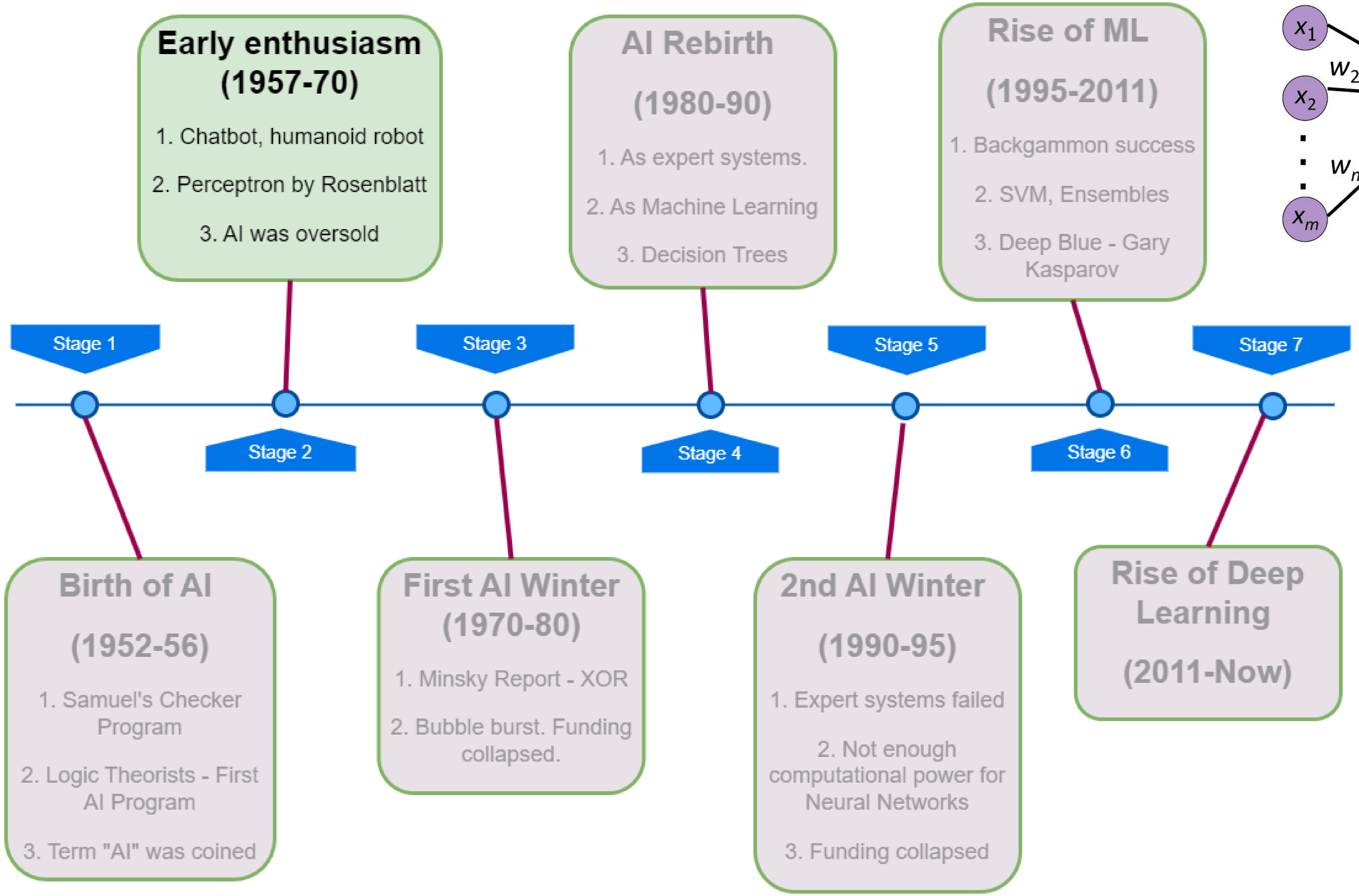
- To make computers
  - Think and act like humans
  - Make correct inferences
  - Be human like in decision making

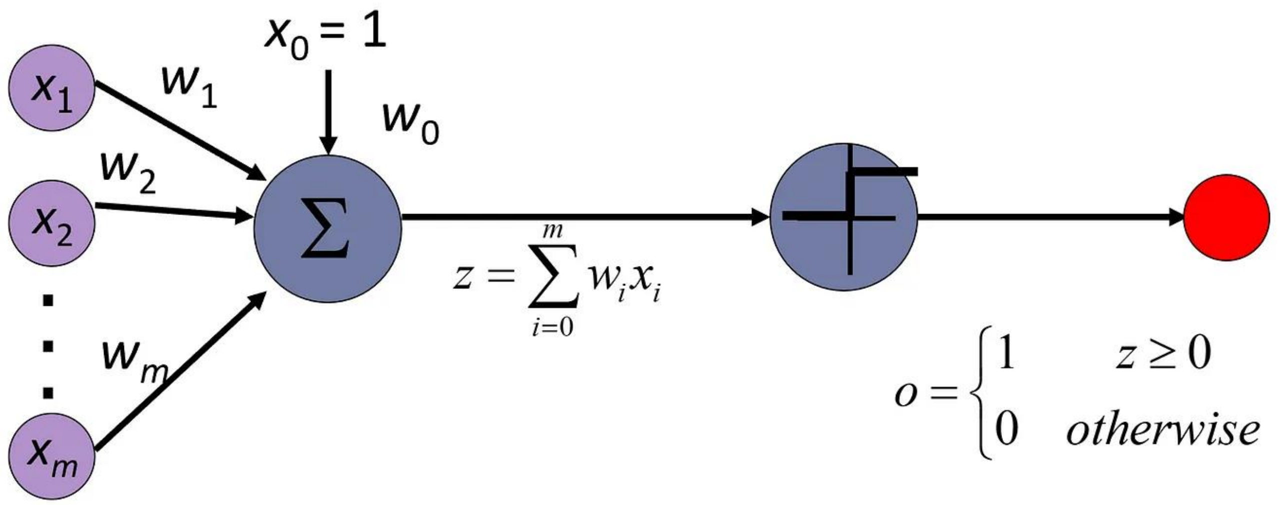
Stage 1 – Birth of AI



Minimax algorithm with brute force search in entire "move space"

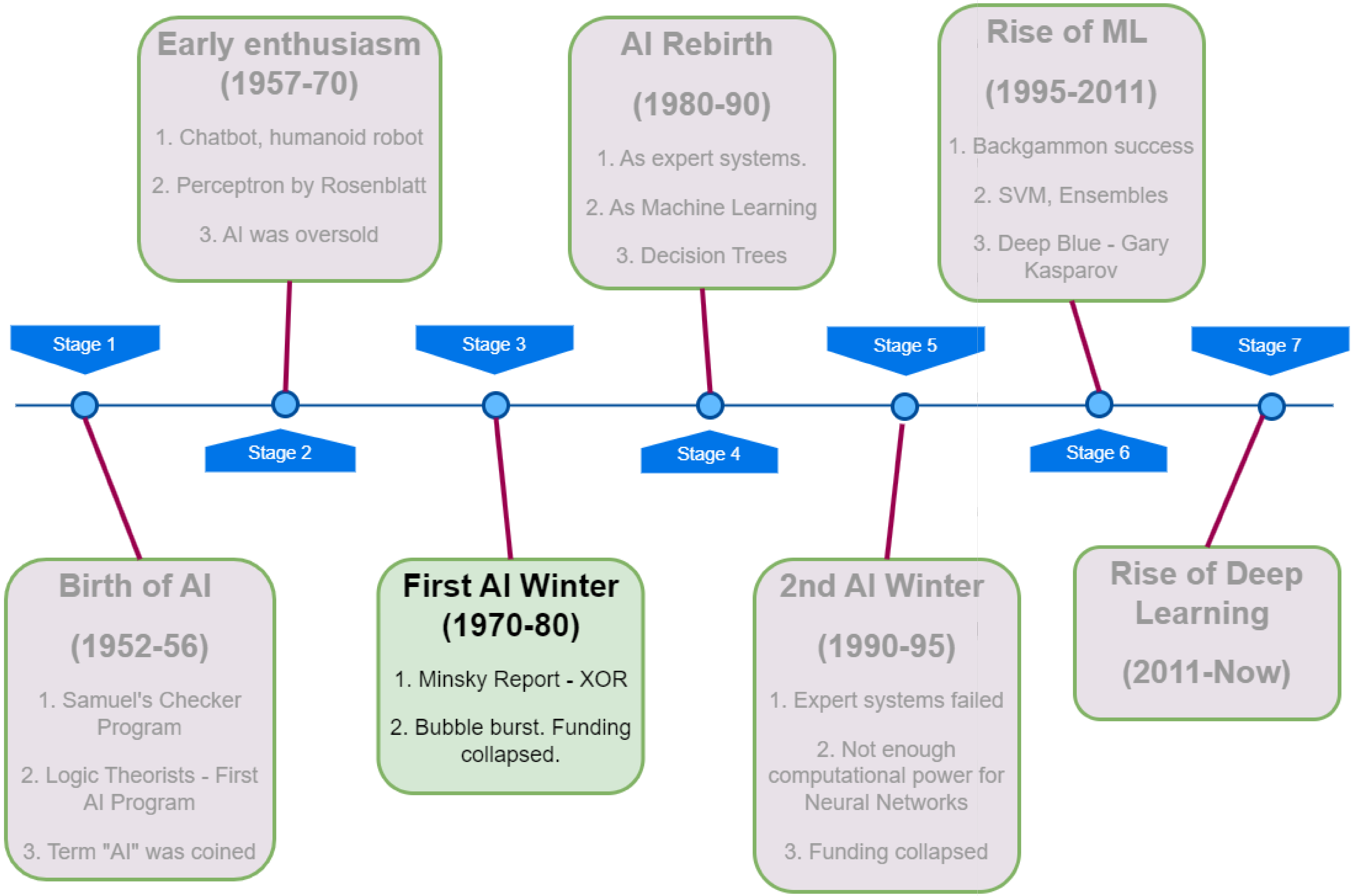
Stage 2 – Early enthusiasm





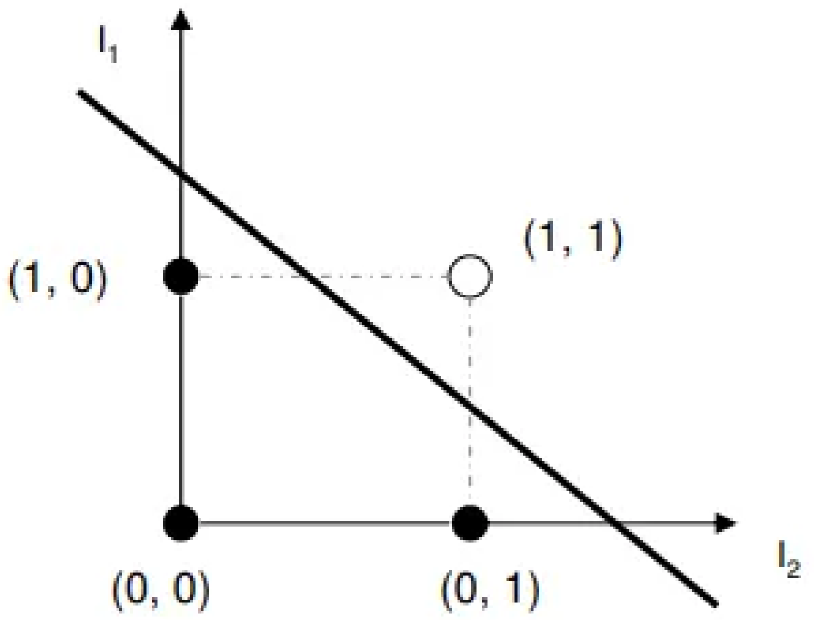
Perceptron is the cornerstone of neural networks

Stage 3 – First AI winter



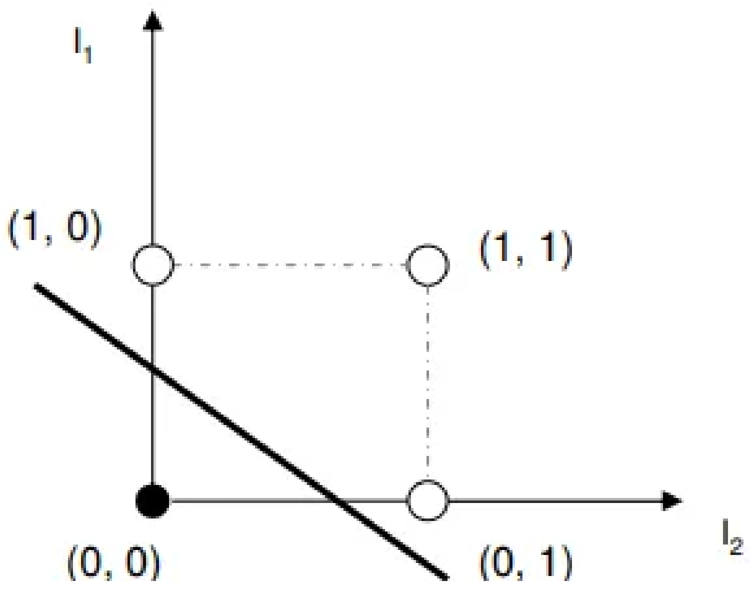
AND

I_1	I_2	out
0	0	0
0	1	0
1	0	0
1	1	1



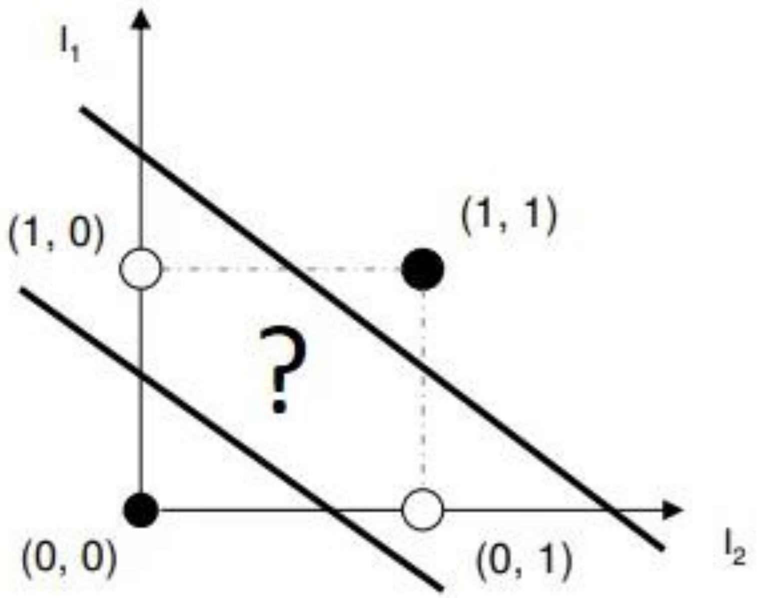
OR

I_1	I_2	out
0	0	0
0	1	1
1	0	1
1	1	1

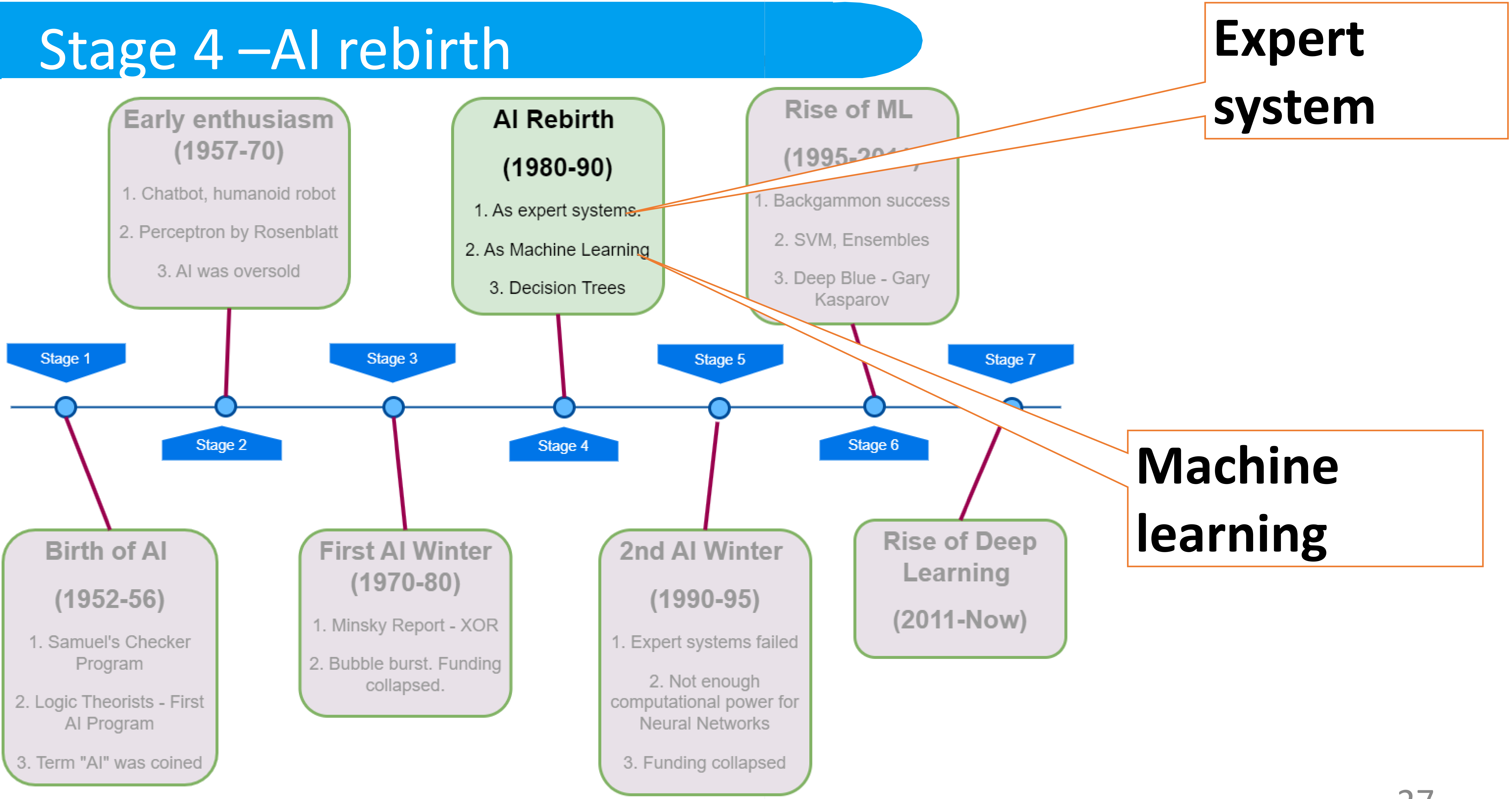


XOR

I_1	I_2	out
0	0	0
0	1	1
1	0	1
1	1	0

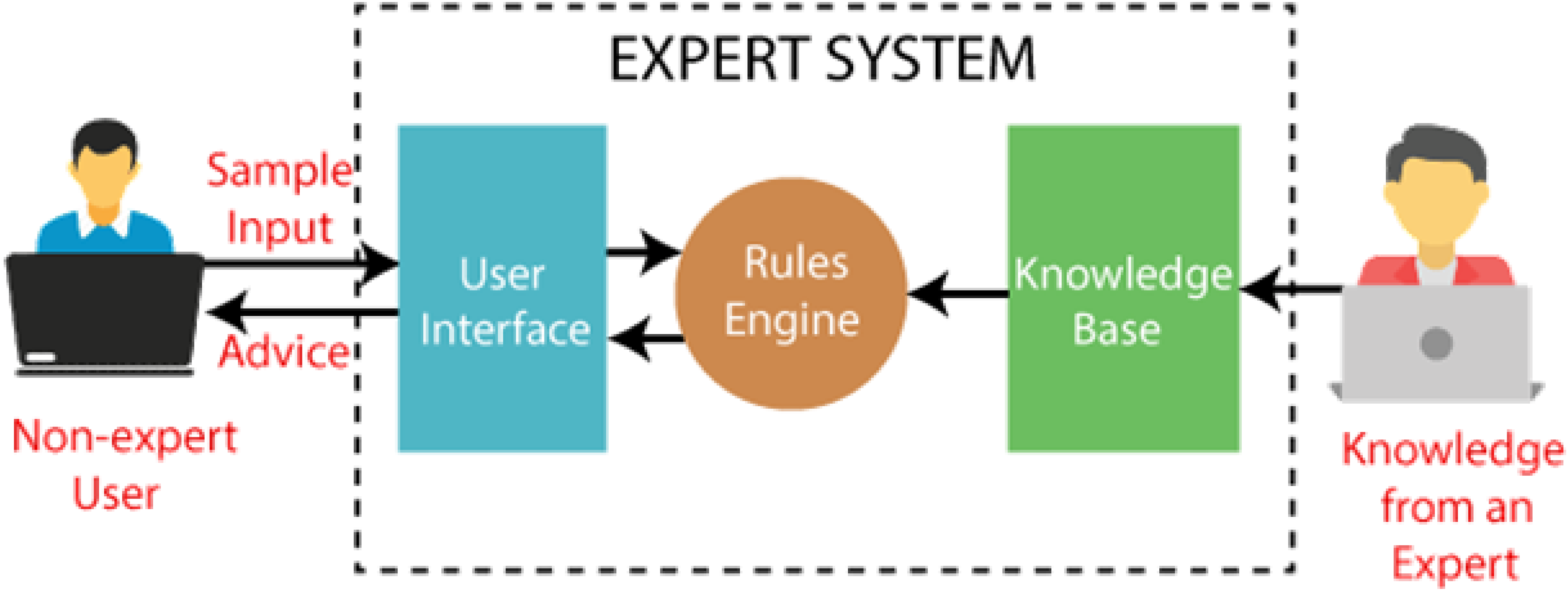


Stage 4 –AI rebirth



4. AI rebirth (contd.)

- Expert Systems
  - Emulate human decision making



4. AI rebirth – Expert Systems

- Used in healthcare. E.g. MYCIN
- Limitations
  - Needed experts to capture the domain knowledge
  - Cannot scale with knowledge explosion
  - Cannot apply for unknown areas

```
IF
   1. the infection is primary bacteria, and
   2. The site of the culture is one of the sterile sites and
   3. The suspected portal of entry is gastro intestinal
THEN
      Suggested evidence of 0.7 that the organism is a bactericide
```

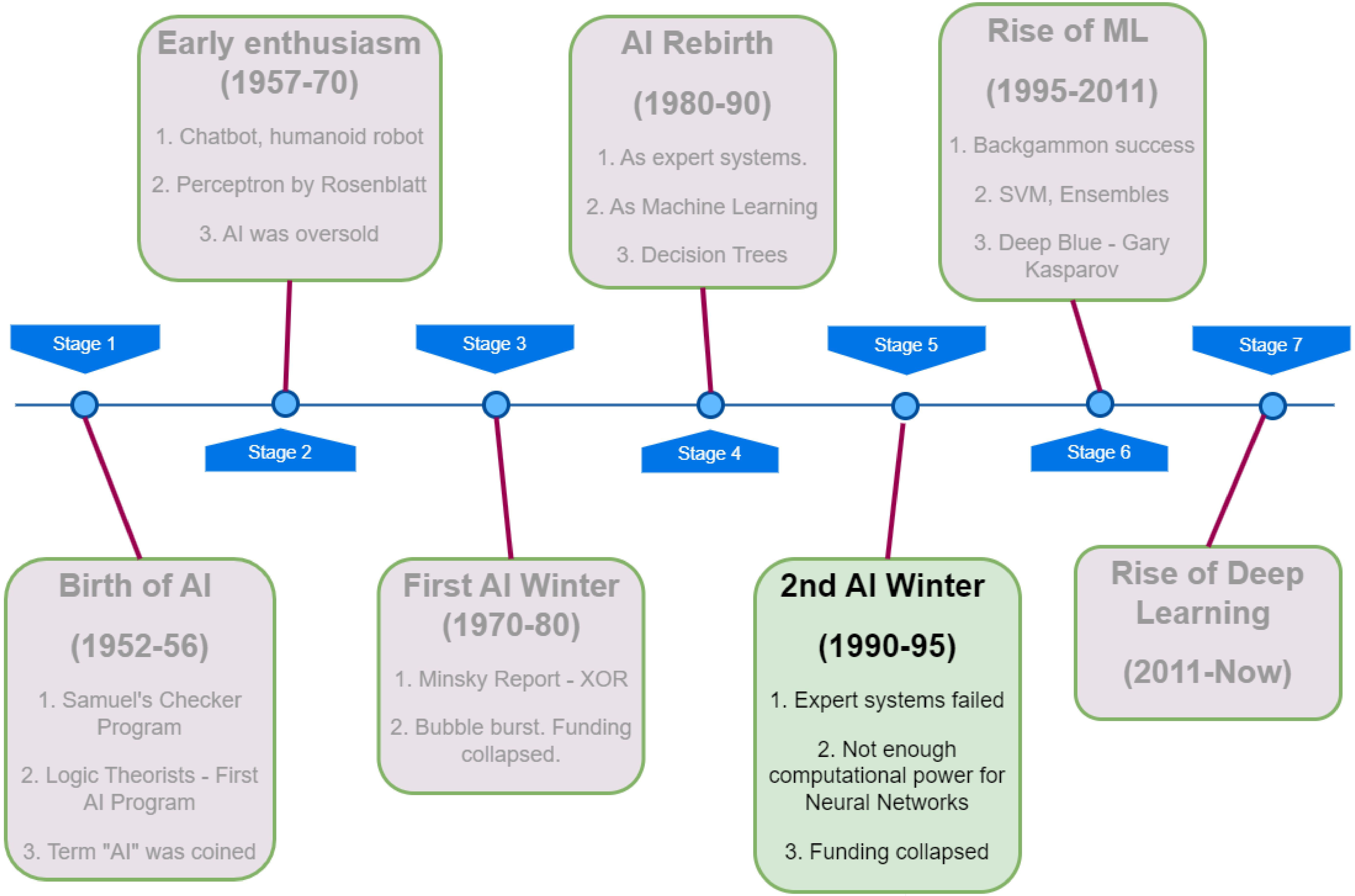
Stage 4. AI rebirth

- Machine Learning: new term coined for funding

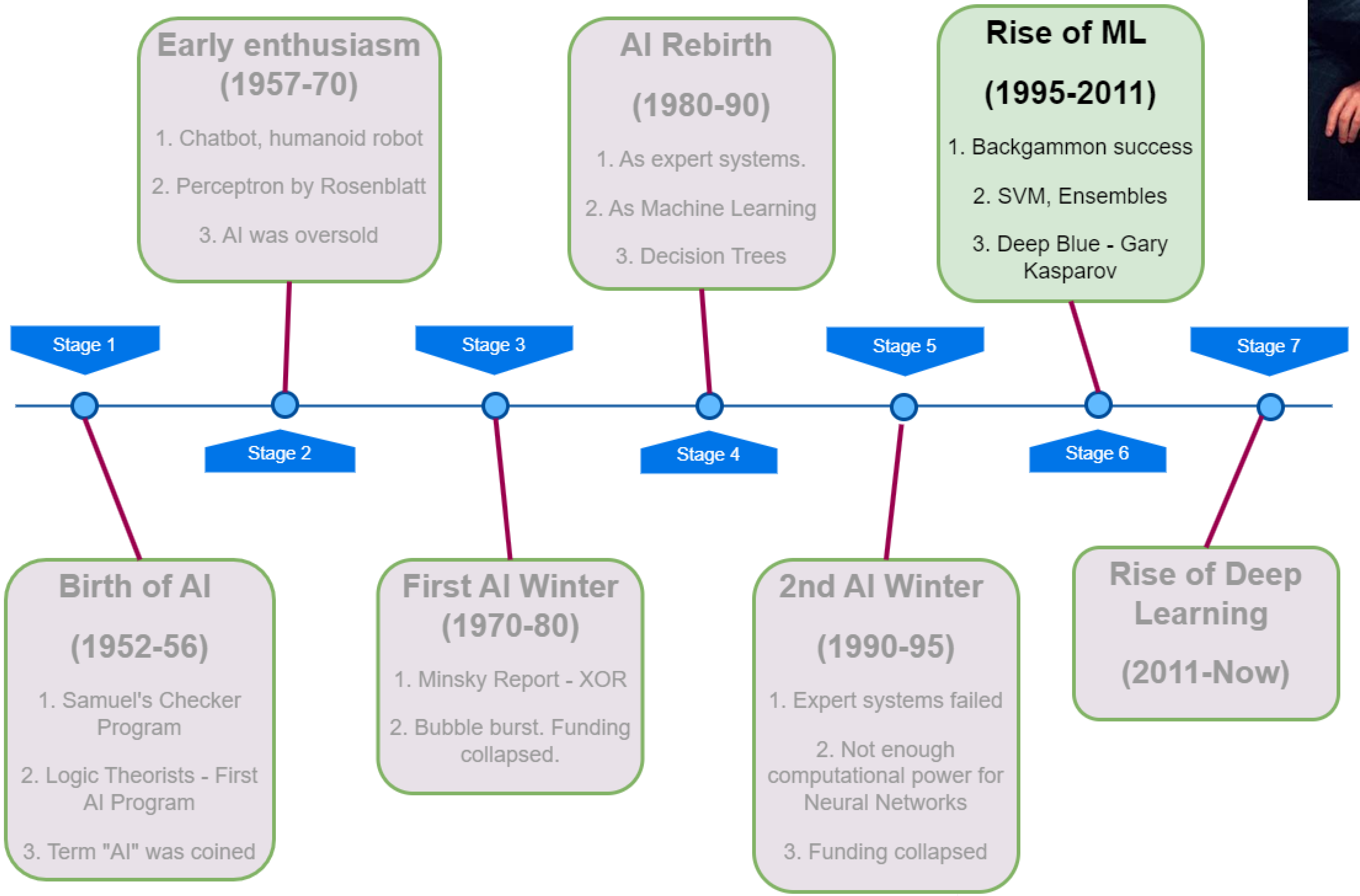
Key Differences

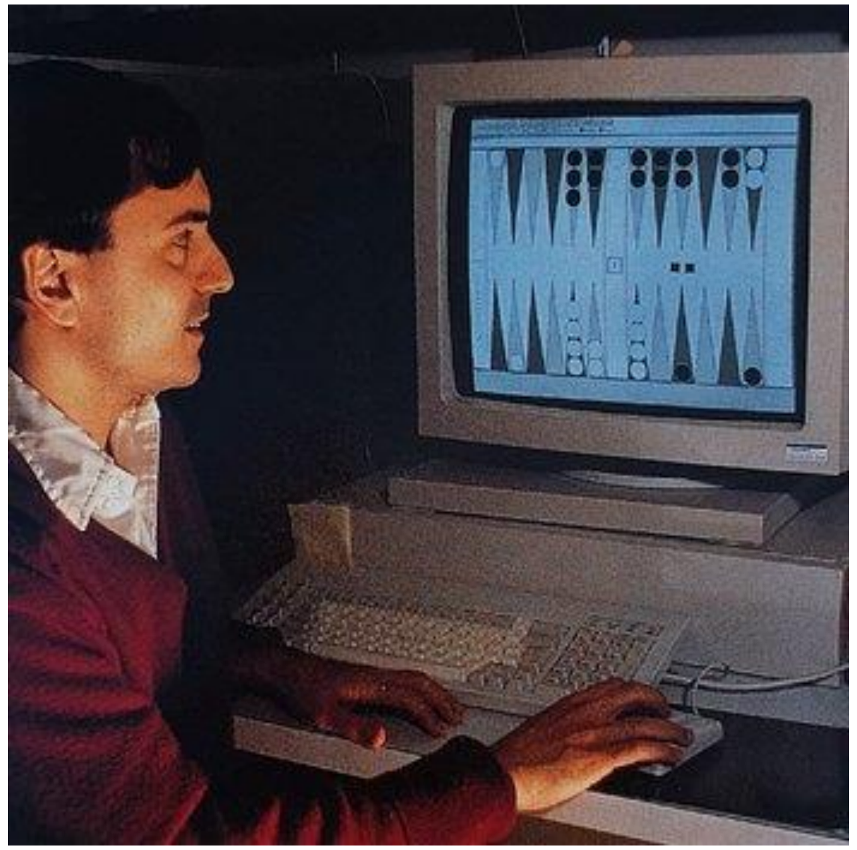
	Artificial Intelligence	Machine Learning
APPROACH	Top down	Bottom up
GOAL	Grand goal	Practical smaller goals
BASED ON	Logic	Statistics & Optimization

Stage 5 –2nd AI winter



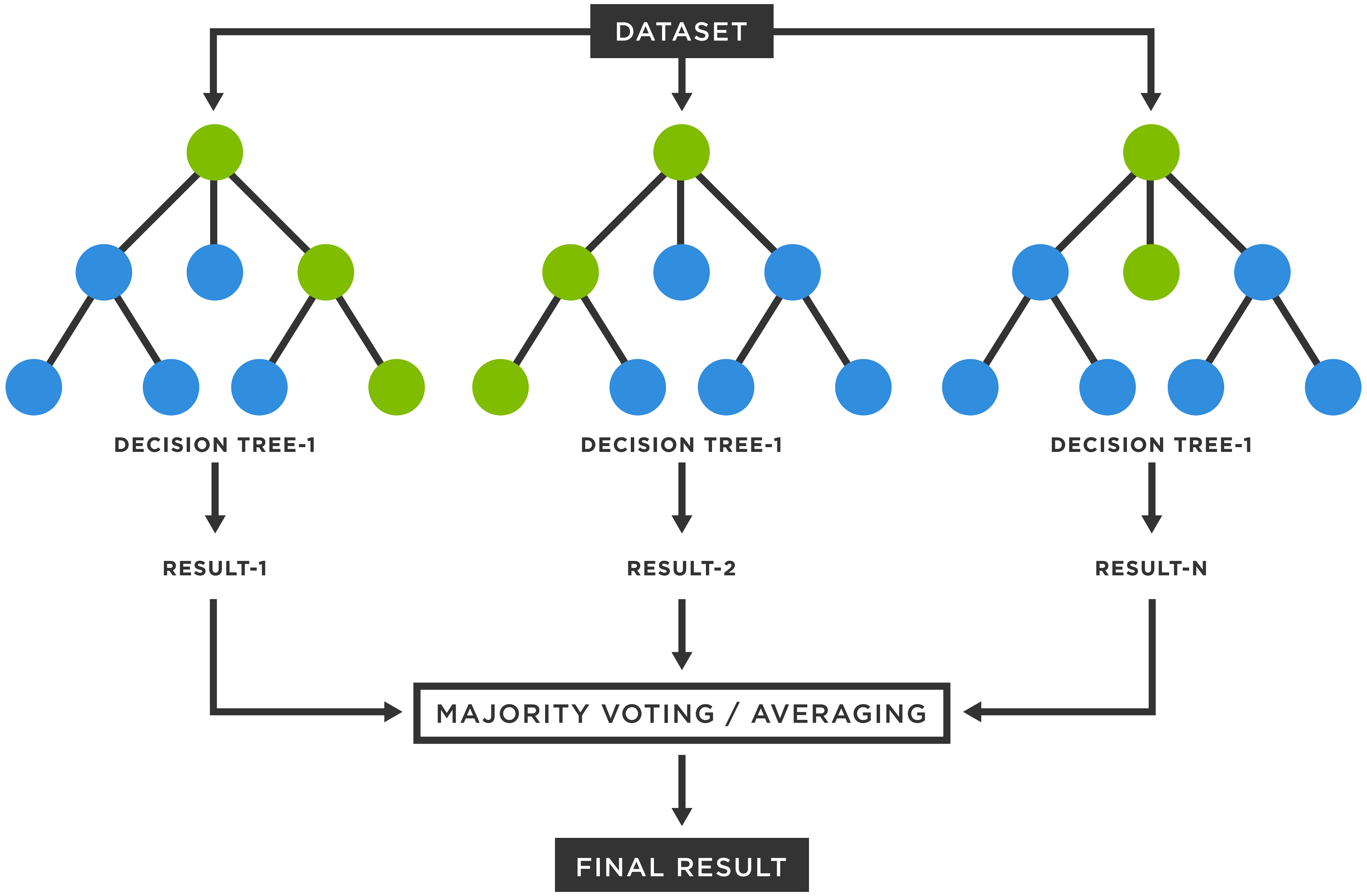
Stage 6 –Rise of ML

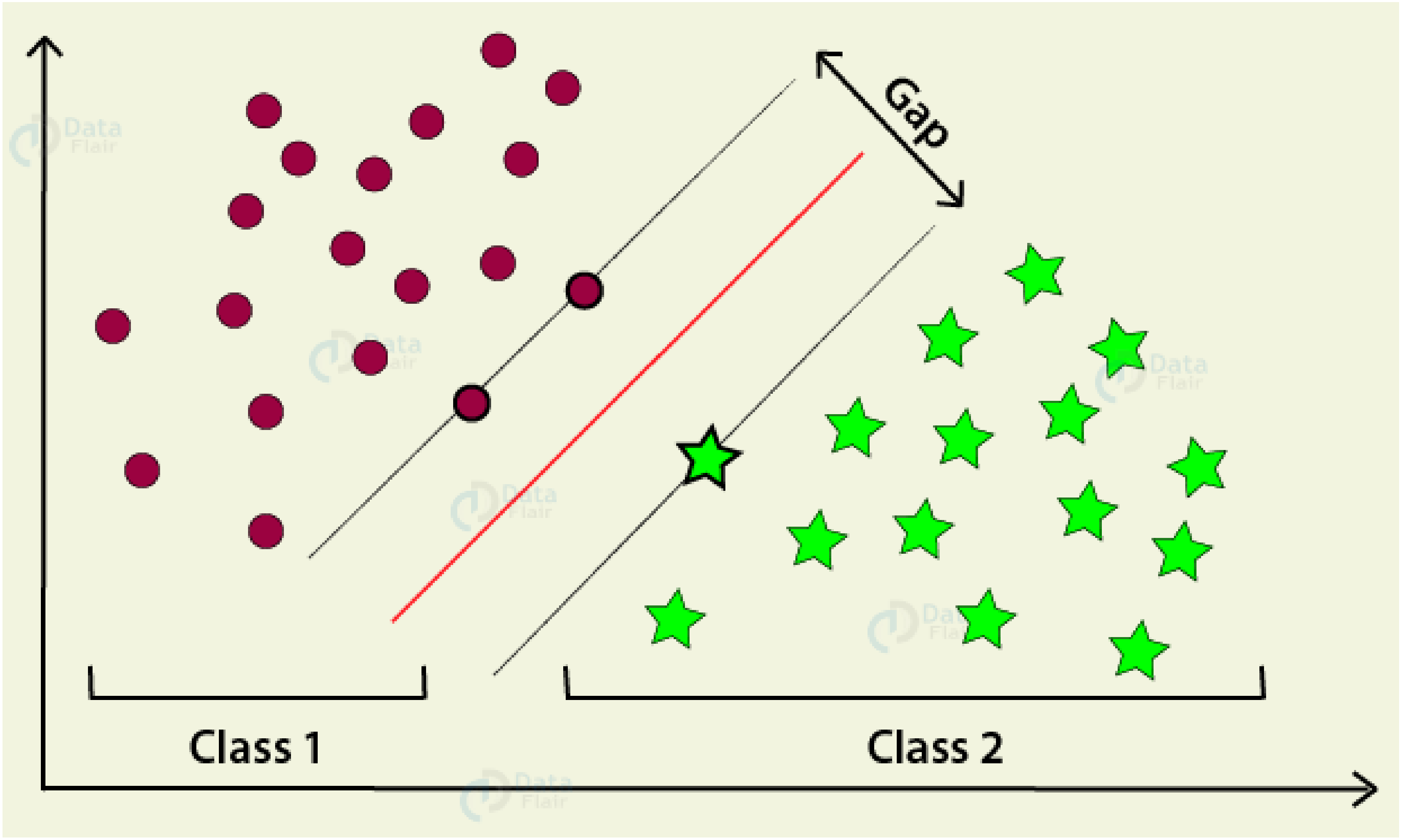




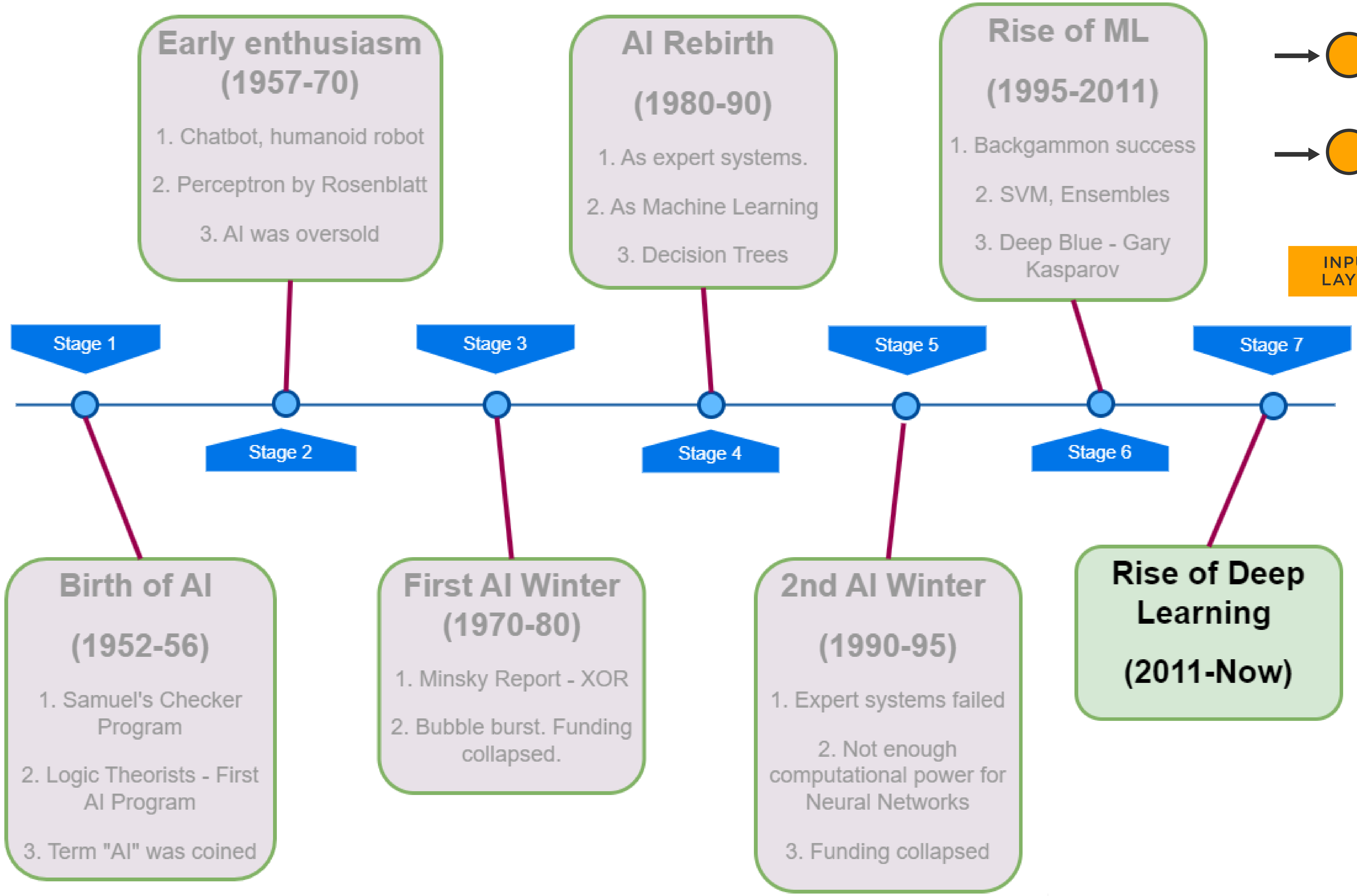
Compare with Alpha Zero - 2017

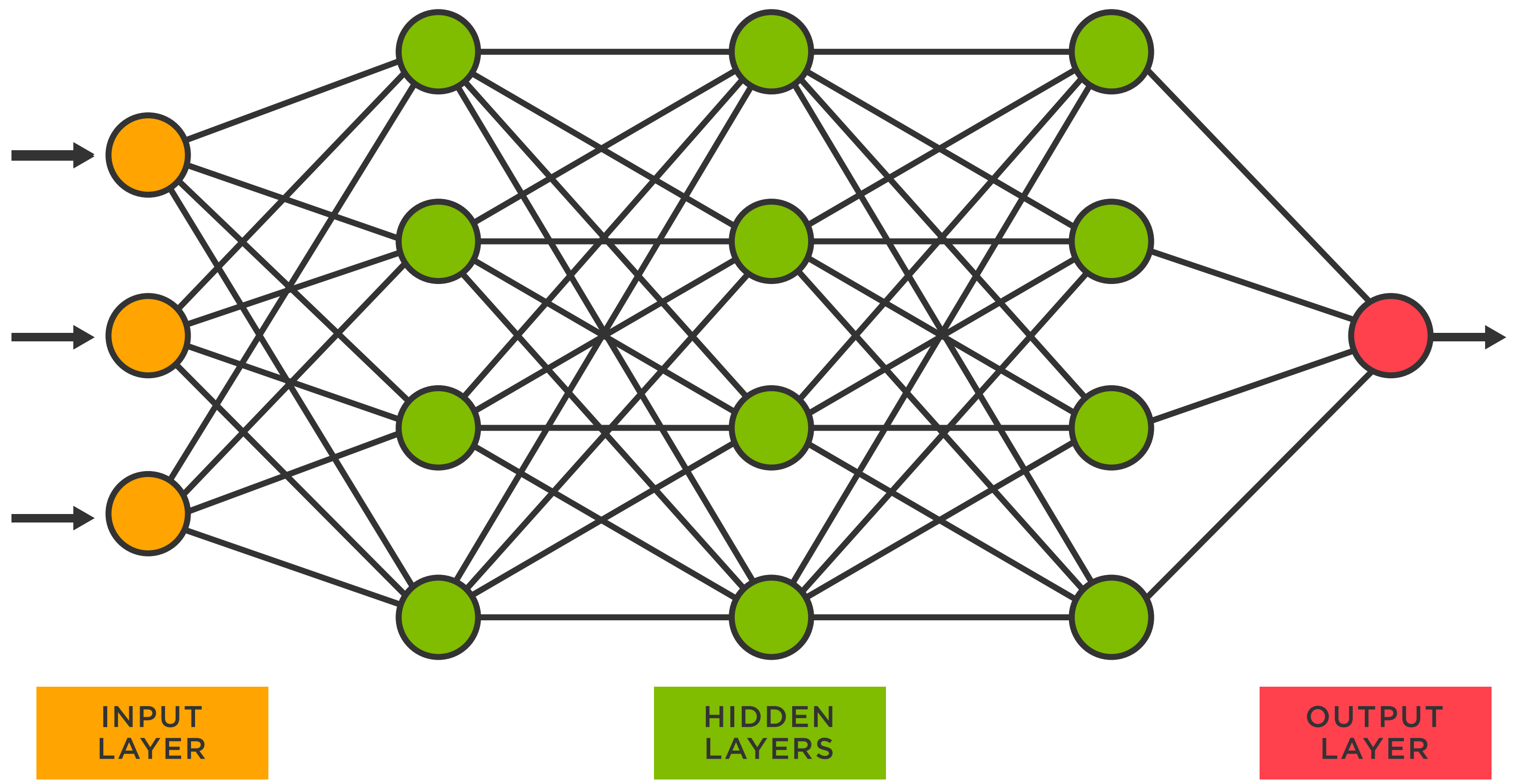
Stage 6 – Rise of ML

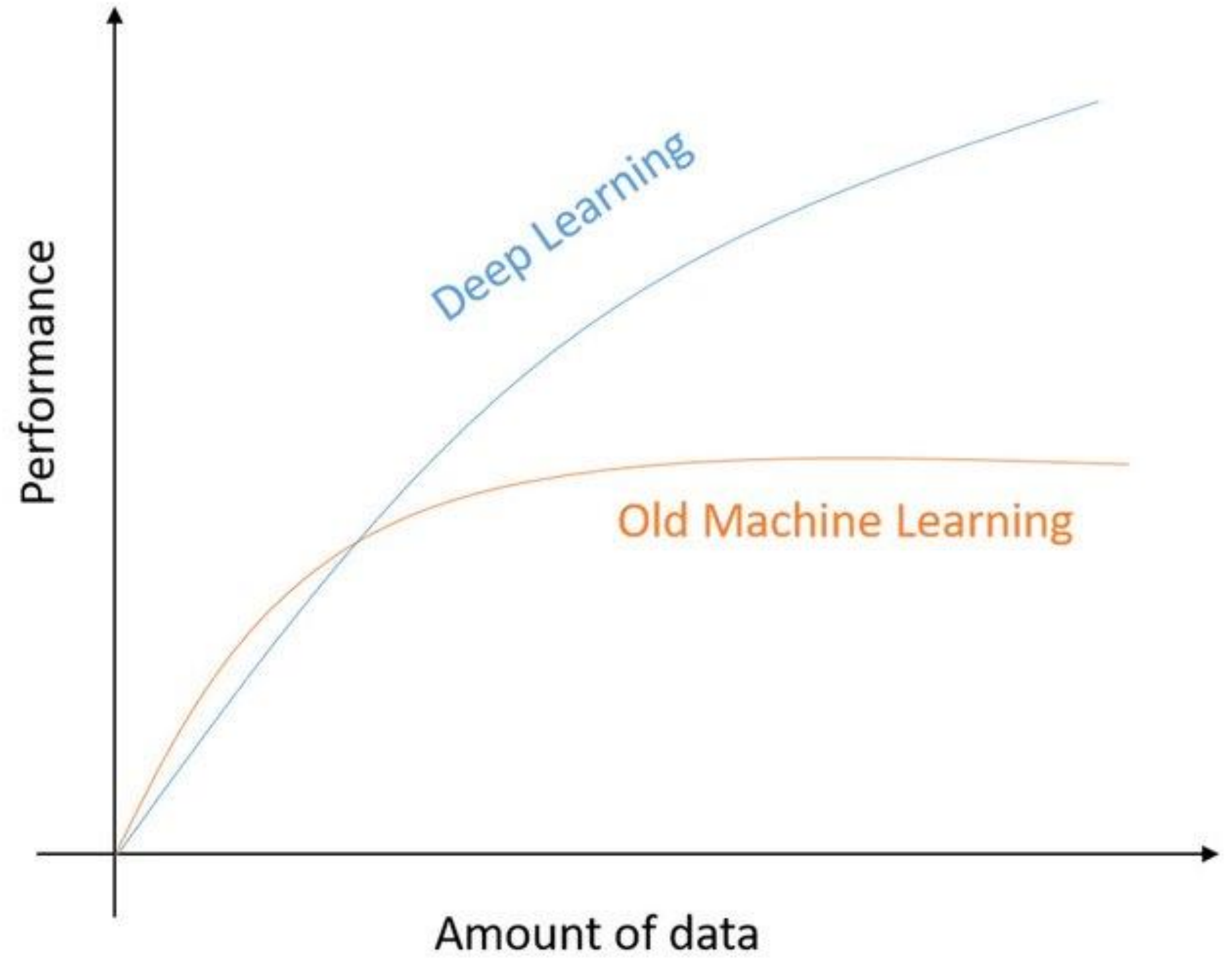




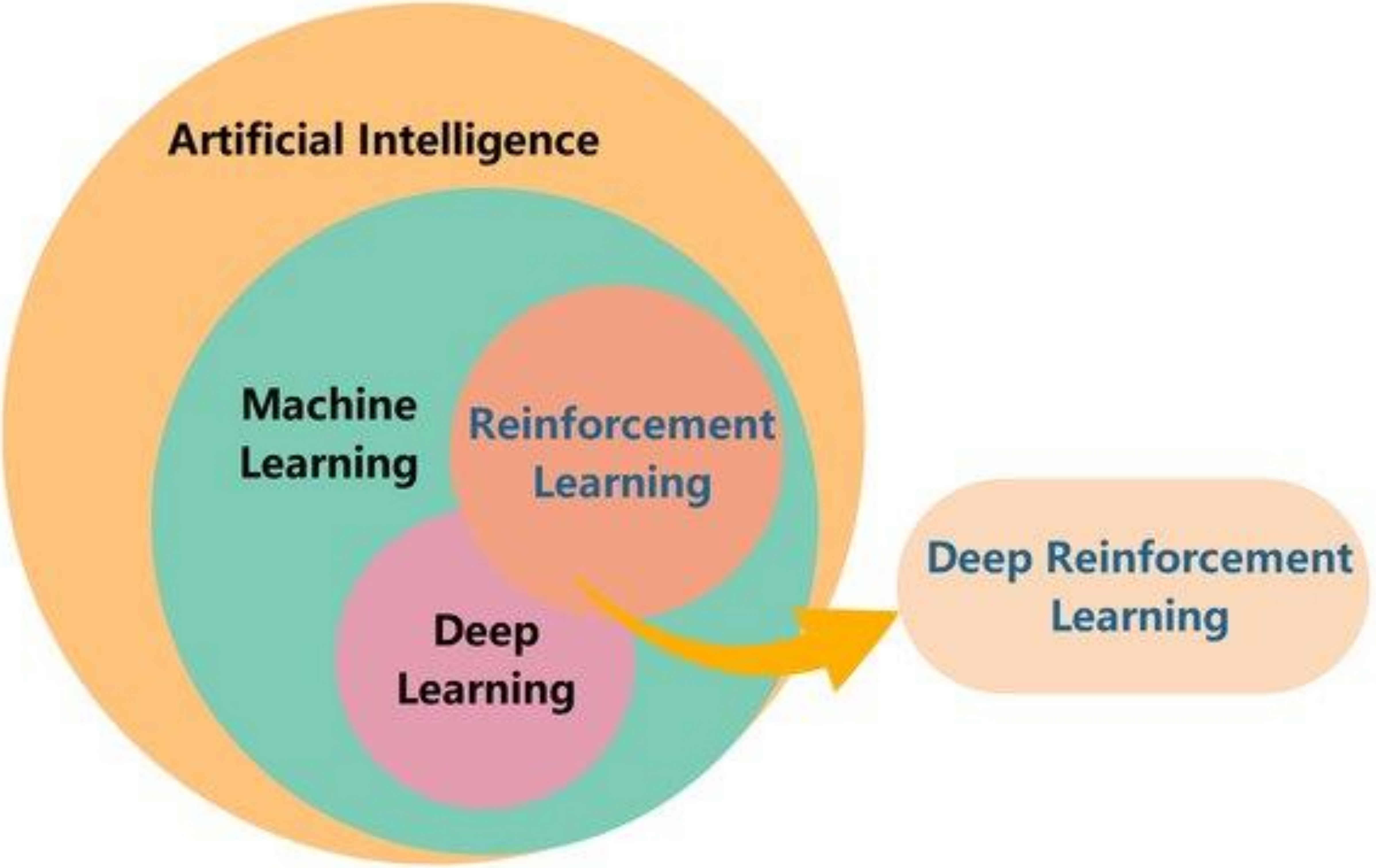
Stage 7 –Rise of Deep Learning







ML, DL v/s AI


Artificial Intelligence
Machine Learning
Reinforcement Learning
Deep Learning
Deep Reinforcement Learning

Takeaways

- Compare & Contrast
  - Traditional programming with machine learning
  - How ML scales by learning the pattern
- Brief history of AI from ML perspective
- Compare & Contrast AI and ML
- Where DL wins against ML

Approach to Machine Learning

- Four pillars
  - Linear Algebra
  - Calculus & Optimization
  - Probability & Statistics
  - Programming

Sequence of Topics

- Analyzing Machine Learning from many perspectives
- Prototype Learning, Instance Learning – kNN
- Distance measures, scaling, applications, curse of dimensionality
- Linear Classifiers with Perceptron
- Metrics – Precision, Recall, F-1, F-Beta

Sequence of Topics

- Linear Regression
  - Simple & multiple, polynomial regression
- Analytical and numerical (Gradient Descent),
- Vector Calculus
- Bias-Variance, overfitting,
- Regularization (Lagrange Multipliers)
- Multicollinearity, Heteroskedasticity

- Support Vector Machines

Sequence of Topics

- Information Theory, Divergences
- Feature Selection, Feature Importance
- Probabilistic modelling – Generative/Discriminative
- Distributions – uni & multivariate
- MLE, MAP using Tensorflow Probability
- Naïve Bayes (various types)
- Bayesian
- Logistic Regression, Softmax Regression
- Handling Imbalanced data – SMOTE, SMOTE-Tomek

Sequence of Topics

- Unsupervised Learning
- Clustering:
  - K-Means, K-Means++, K-Medoid, GMM
  - Hierarchical
- Dimensionality Reduction - PCA
- Decision Trees
- Ensembles – Bagging, Boosting, Stacking
- AdaBoost, Gradient Boost, XGBoost


QUESTIONS

Literature Survey