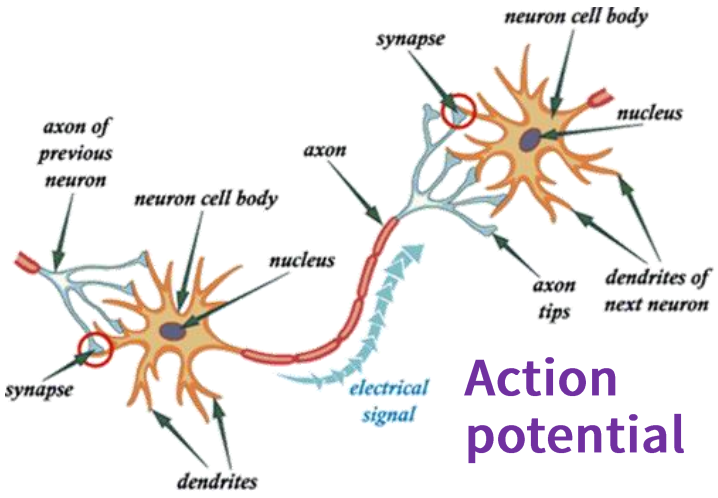
Idea of Artificial Intelligence (AI)

Human's brain and neuron

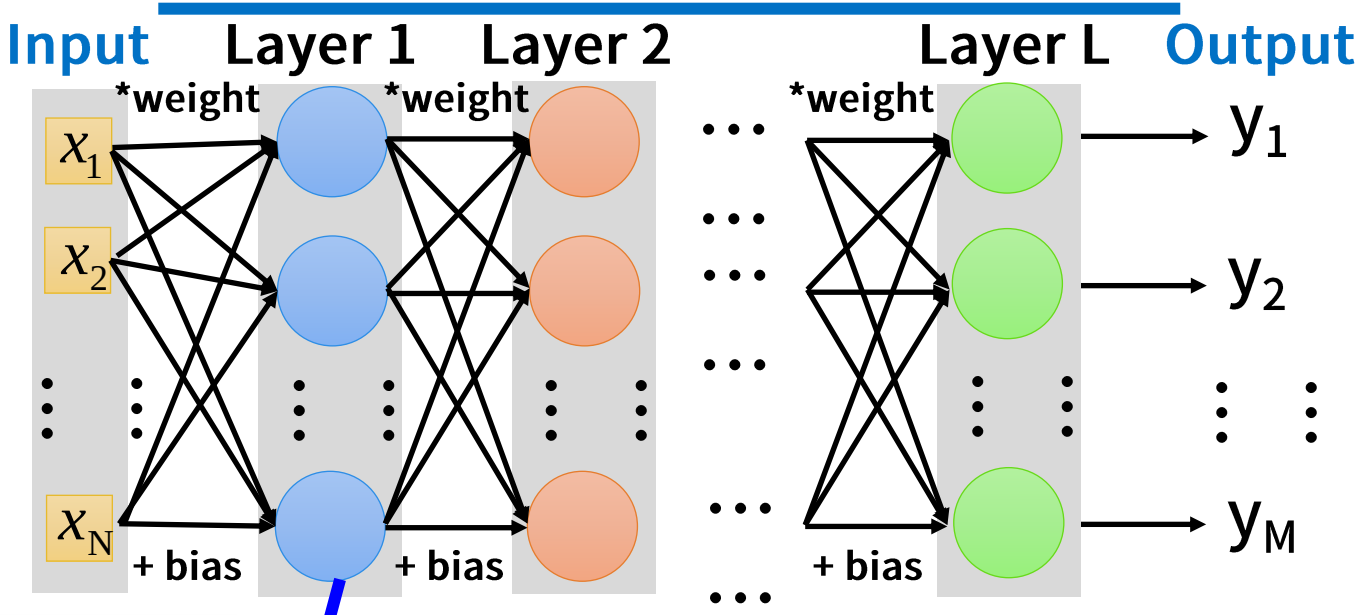
Input

Output

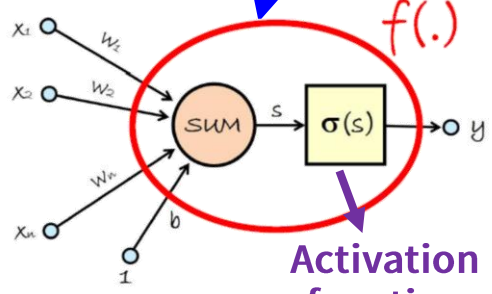


Action potential

Neuron Network and neuron



neuron



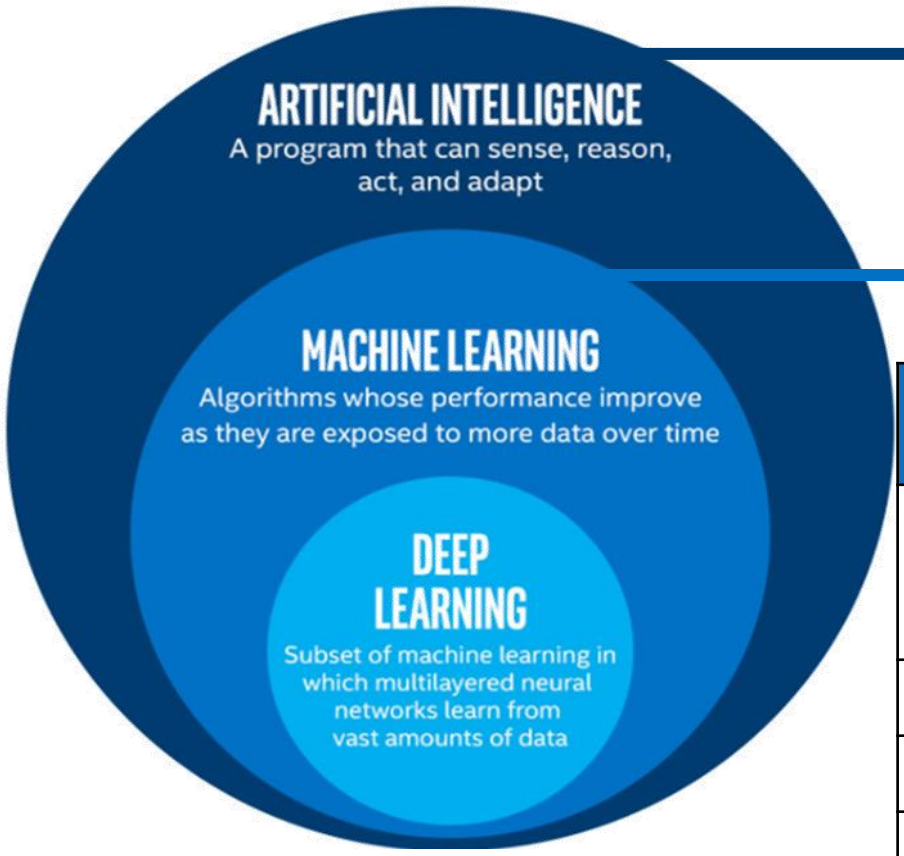
Activation function

$$y = \sigma(s) = f(\vec{x}\cdot\vec{w}+b)$$
$$= f(\vec{x},\vec{w},b) = f(\vec{x})$$

x : input
y : output
w, b : adjustable variables

Algorithm : adjust w and b systematically

AI & ML



1950s

AI

Goal

Describing automated systems (machine) that can perform tasks considered to acquire "intelligence"

1980s

ML

Means

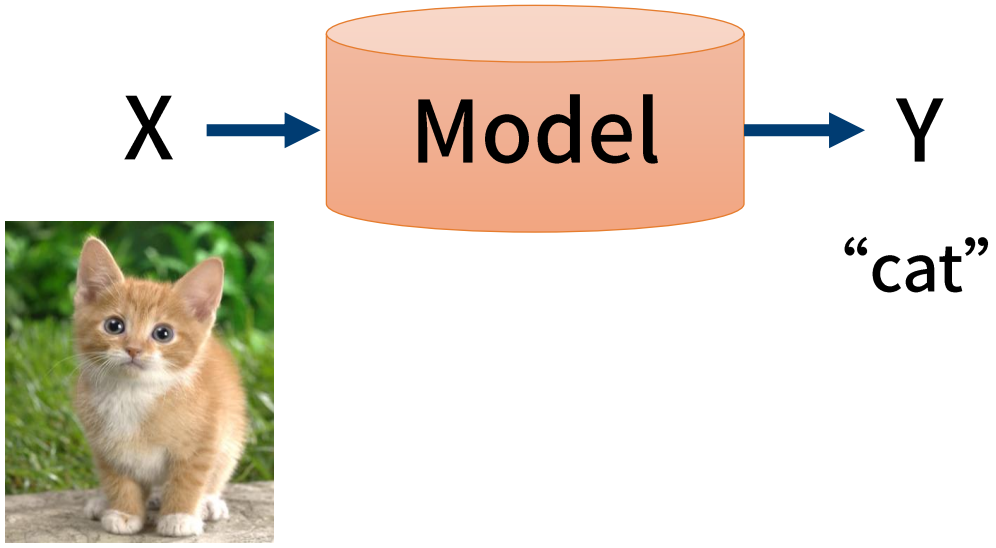
The process of developing machine with the ability to learn from and make predictions using data.

	Supervised learning	Unsupervised learning	Semi-supervised learning
Definition	X → Model → Y	X → Model →	X → Model → (Y)
Labelled data	O	X	Little
Unlabelled data	X	O	Most
Methods	Classification, Regression	Clustering, 維度降低	Autoencoder

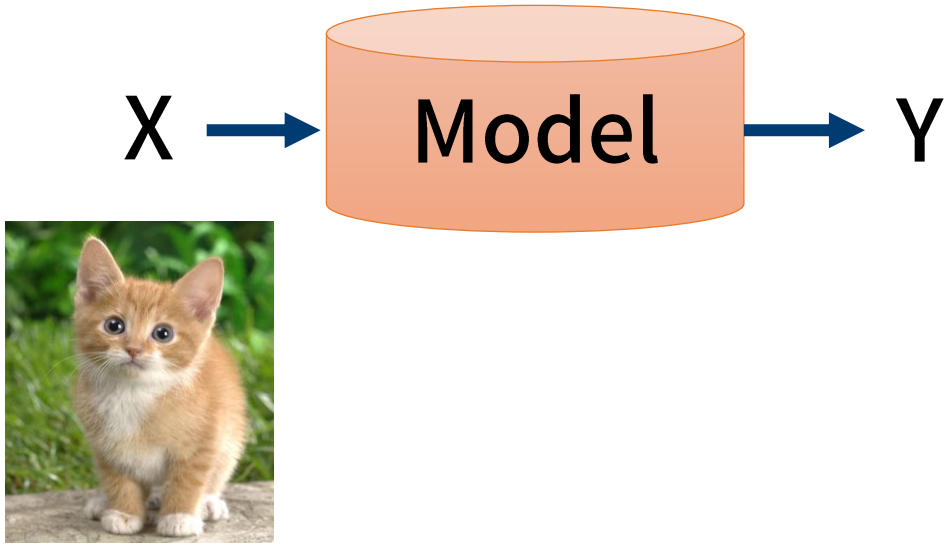
Richter A.N., Khoshgoftaar T.M. , A review of statistical and machine learning methods for modeling cancer risk using structured clinical data, Artif Intell Med. (2018) 90:1-14.
Yun Liu, et al. How to Read Articles That Use Machine Learning Users' Guides to the Medical Literature. JAMA. 2019;322(18):1806-1816.

Labelled data vs. Unlabelled data

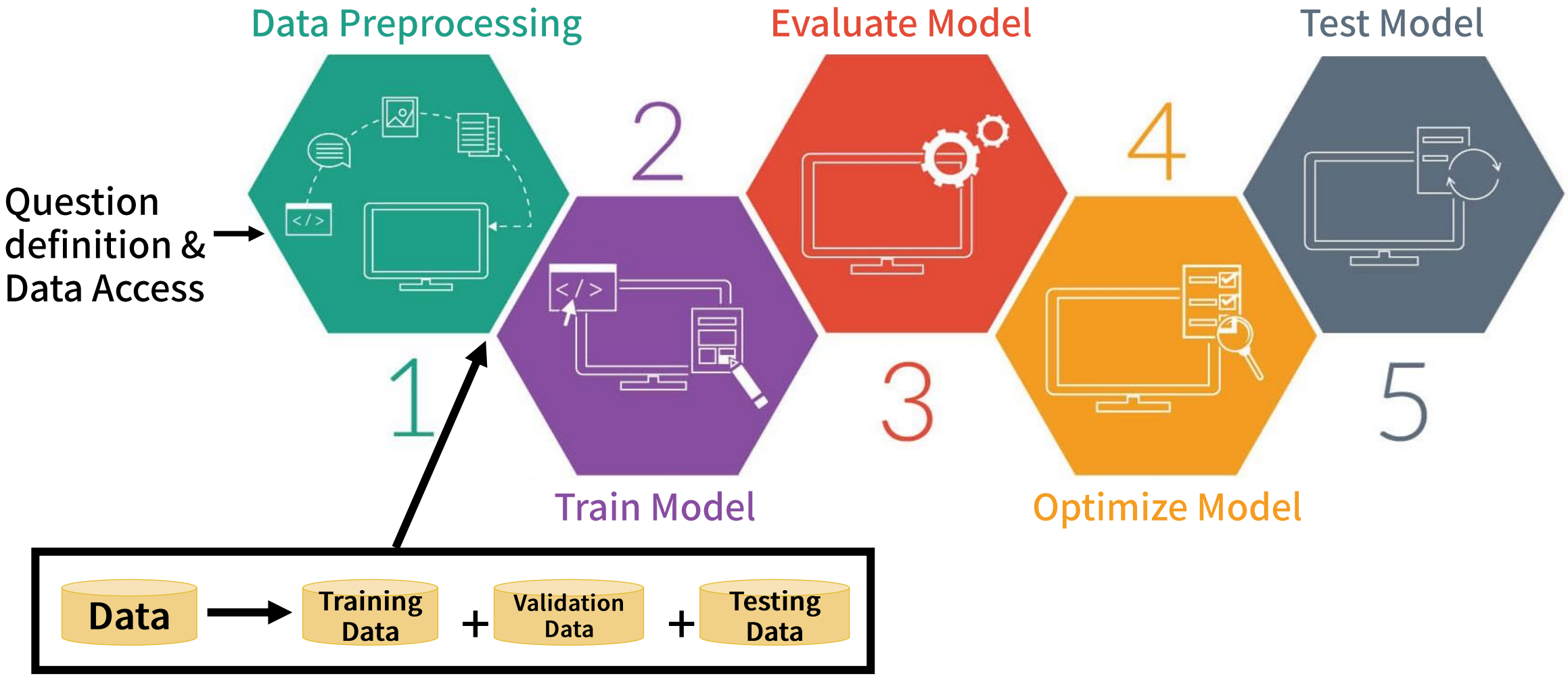
Labelled data



Unlabelled data



Machine learning workflow

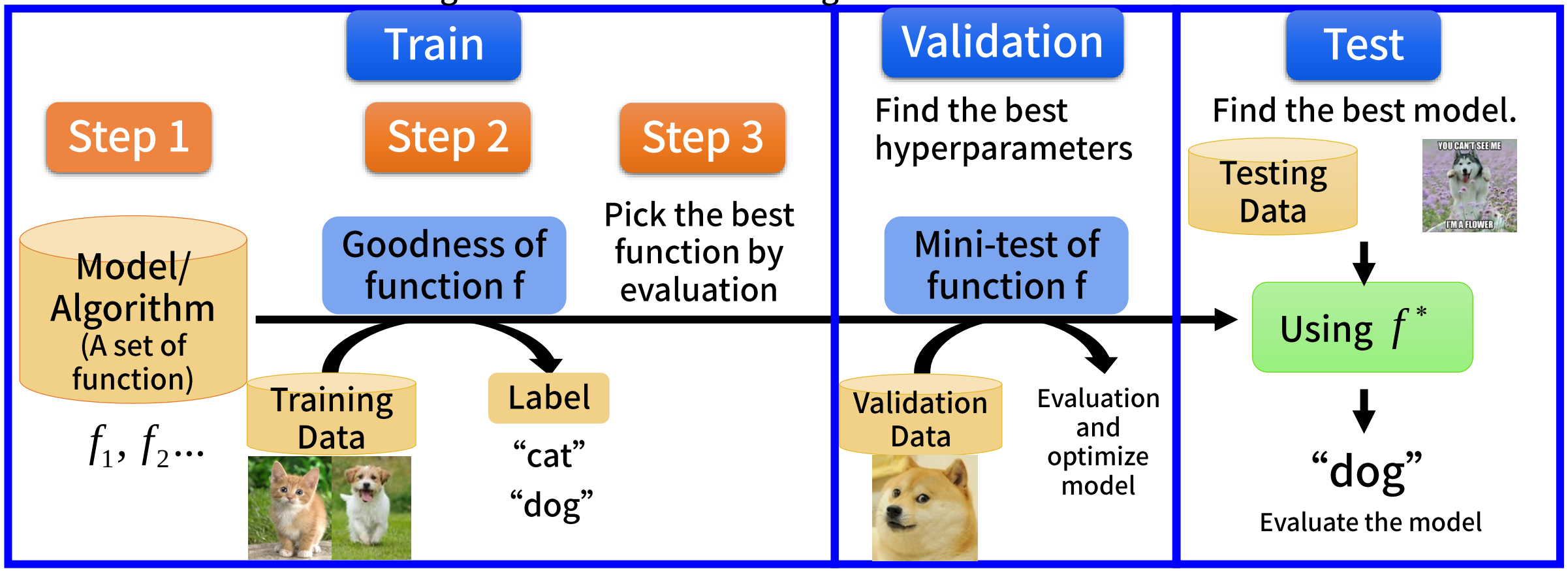


Training, validation, and testing data should be **independent** of **each other**.

Refer to https://www.mactores.com/services/aws-big-data-machine-learning-cognitive-services/

How to train the model in machine learning ?

- Machine learning is a process of look for the best function and features to solve the problem.
  - ➢ Problem: How to distinguish between cat and dog ?



☺ **Label** identifies what a collection of data (the model input) represents.

- ✓ Example: For a cancer model, it would be cancer present or absent.

Hung-Yi Li. Machine learning course in National Taiwan University. Retrieved on March 9th, 2020.

Evaluate the models

- Confusion matrix and the derivative metrics.

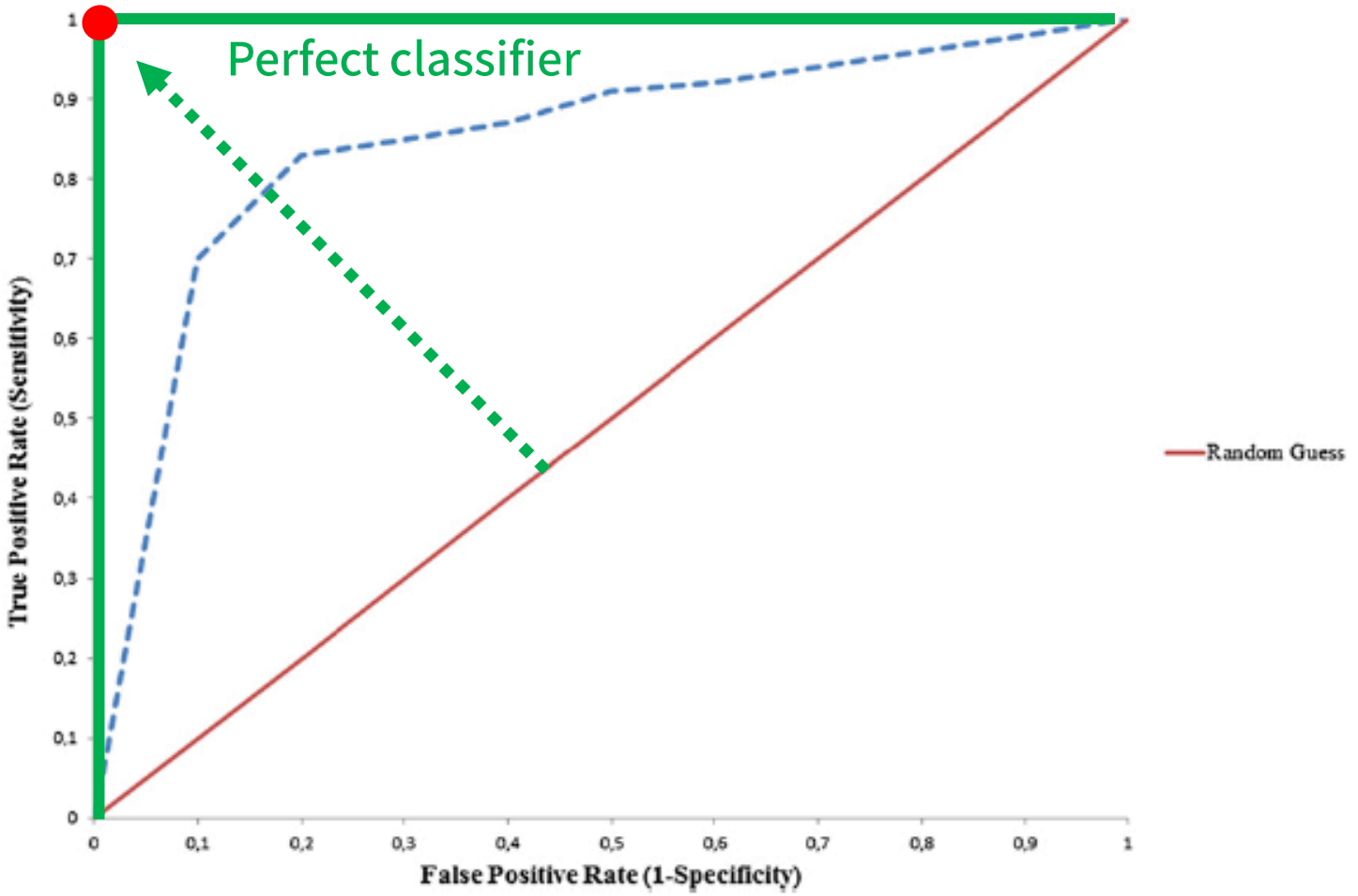
		Actual condition	
		Positive	Negative
Predicted condition	Positive	True Positive (TP)	False Positive (FP)
	Negative	False Negative (FN)	True Negative (TN)

Trade-off

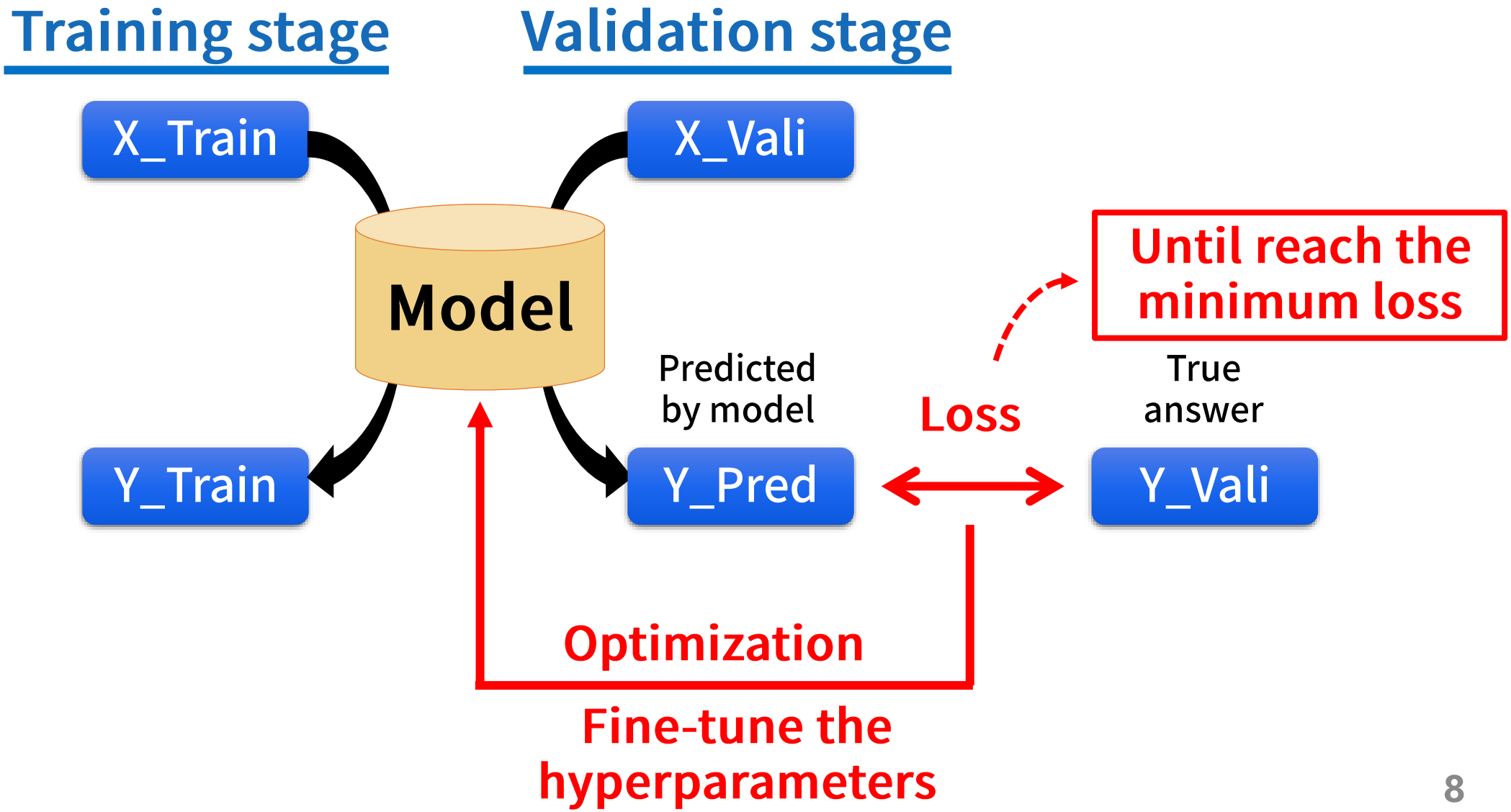
Metrics	Definition	Formula	
Accuracy	Predict the correct proportion (Meet paradox when unbalanced data)	$\frac{TP + TN}{TP + FN + TN + FP}$	
Precision	$\frac{\text{True positive}}{\text{Predicted positive}}$	$\frac{TP}{TP + \mathbf{FP}}$	
Recall	$\frac{\text{True positive}}{\text{Actual positive}}$	$\frac{TP}{TP + \mathbf{FN}}$	
F1 score	Harmonic mean of precision and recall (model performance)	$\frac{2 \times precision \times recall}{precision + recall}$	$\frac{2}{\frac{1}{precision} + \frac{1}{recall}}$

Evaluate the models

- Evaluate the model by Receiver operating characteristic (ROC) curve and its area under curve (AUC ; Concordance/C index)



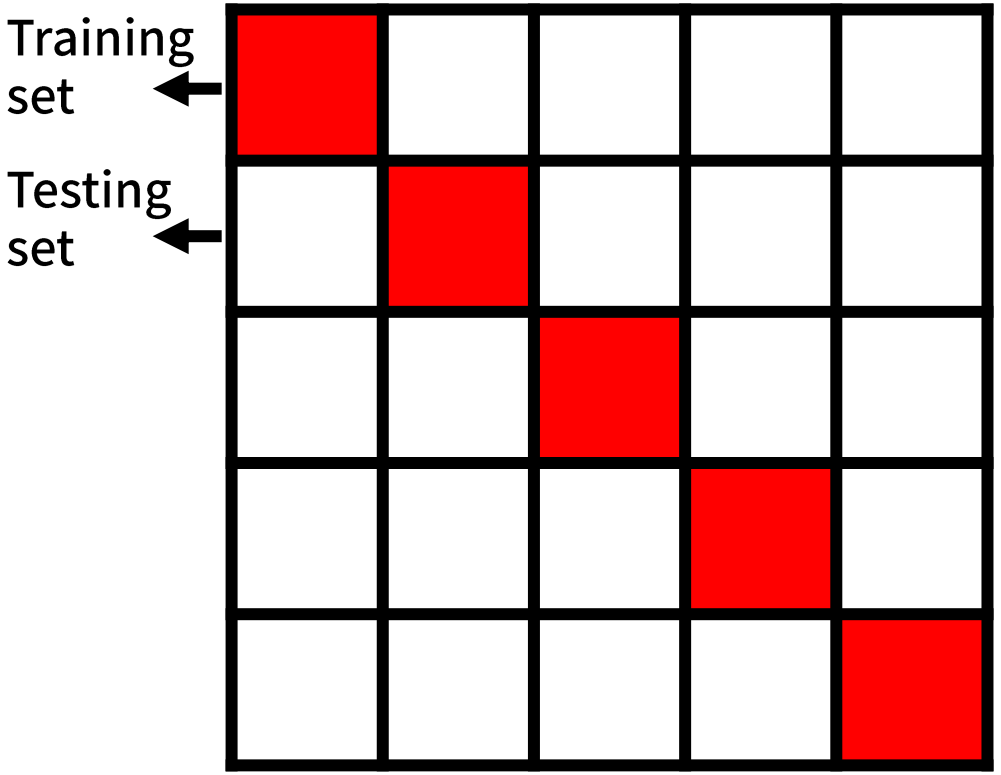
Optimize the models



Cross validation

Five-fold

Divided patients into 5 groups stratified by cisplatin/carboplatin and AKI/non-AKI

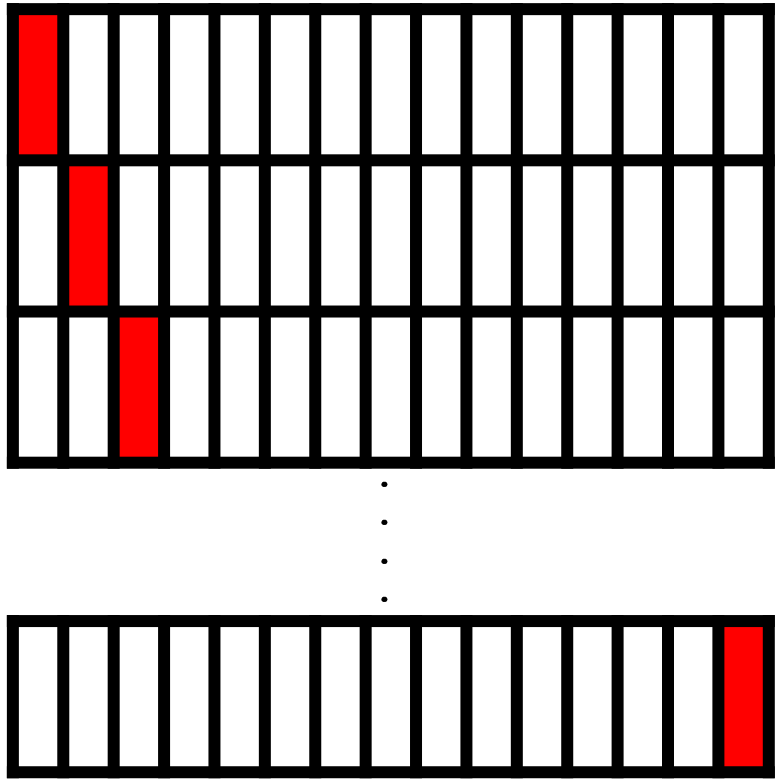


Metrics = A1 A2 A3 A4 A5

Average metrics = (A1 + A2 + A3 + A4 + A5)/5

Leave-one-out

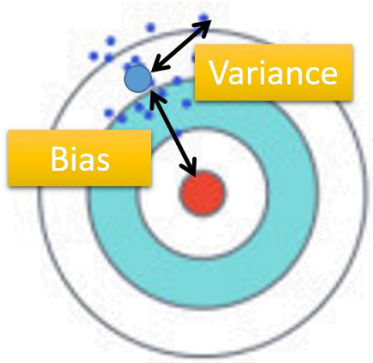
Every individual patient as a group

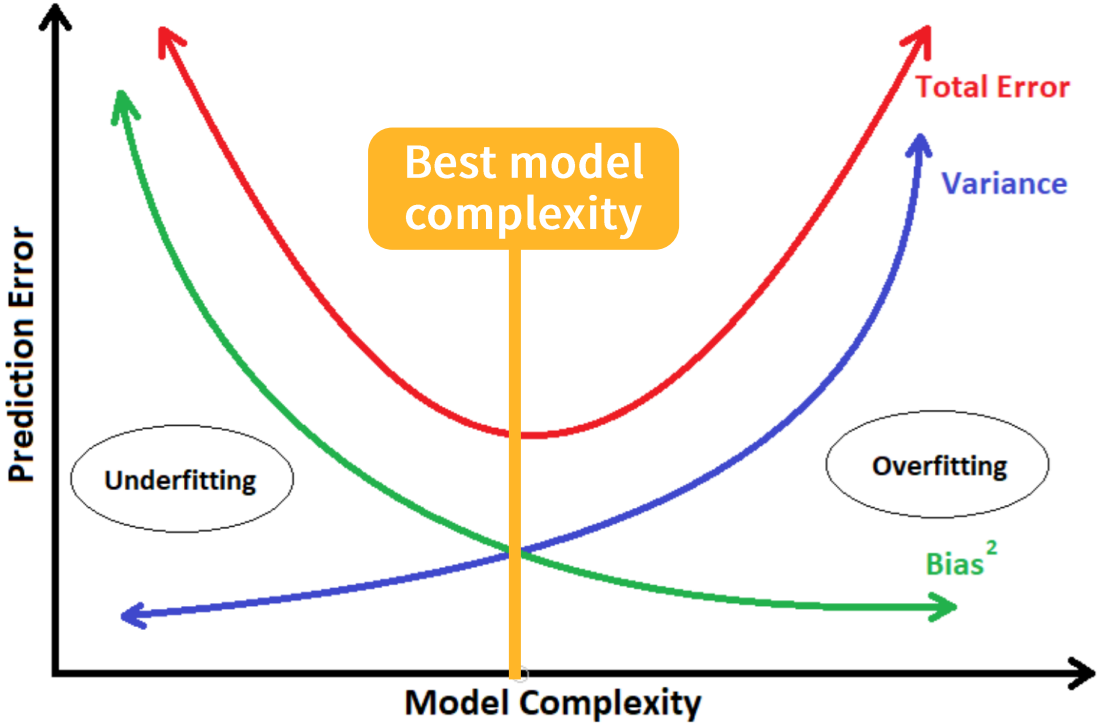


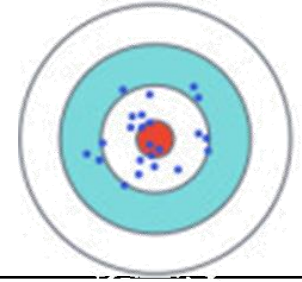
Average metrics = (A1 + A2 + ⋯ + AN)/N

Bias-variance trade-off

- Solution to overfitting: lower the prediction error
  - Bias-variance trade-off decomposition: analyze the expected the prediction error





	Underfitting	Overfitting
Bias	High	Low
Variance	Low	High
Solutions	Redesign model	• More data • Regularization • Adjust features based on algorithms
Pattern		

Bias-Variance Tradeoff. https://jason-chen-1992.weebly.com/home/-bias-variance-tradeoff . Retrieved on March 9th, 2020.

Hung-Yi Li. Machine learning course in National Taiwan University. Retrieved on March 9th, 2020.

Konstantina Kourou, et al. Machine learning applications in cancer prognosis and prediction. Computational and Structural Biotechnology Journal 13 (2015) 8–17

Overfitting

- Common problem: Overfitting (Over-learning)

Definition

The model is trained to fit the training data too well and could not generalize to testing data.

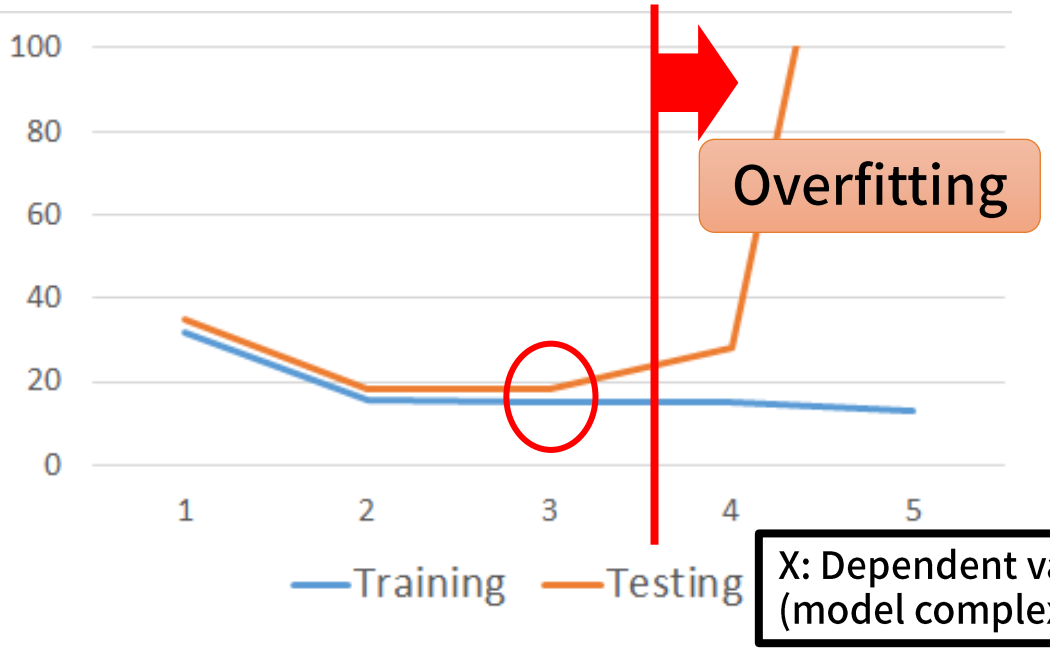
Characteristics

- ✓ Testing error rates significantly increase.
- ✓ Often related to model complexity.

Reason

To find the best model complexity from a high-variance low-bias complex model.

Y: Prediction average error rates



	Training	Testing
1	31.9	35.0
2	15.4	18.4
3	15.3	18.1
4	14.9	28.2
5	12.8	232.1

X: Dependent variable to the nth power (model complexity)

Hung-Yi Li. Machine learning course in National Taiwan University. Retrieved on March 9th, 2020.

Konstantina Kourou, et al. Machine learning applications in cancer prognosis and prediction. Computational and Structural Biotechnology Journal 13 (2015) 8–17

Hyperparameter selection: Grid search

- To find the best prediction performance with best hyperparameters combination of ANN, LR, RF and SVM.