

# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
0- Recall and Plan

# Let's Python!



thanks  
for  
watching!

# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
1- What is Machine?

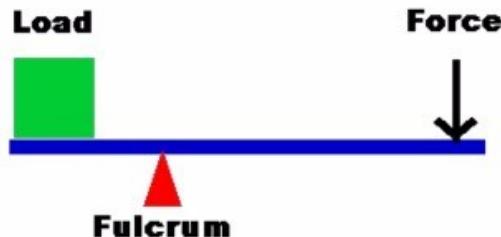
# ماشین چیست؟



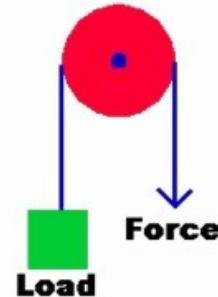
ماشین چیست؟

نیرو وارد کند و حرکت را  
کنترل کند...

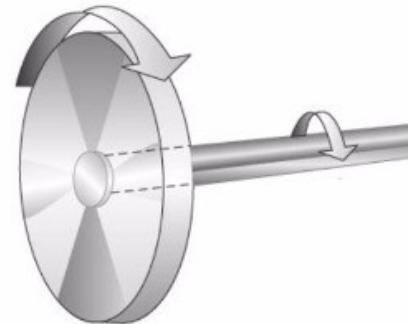
# ماشین چیست؟



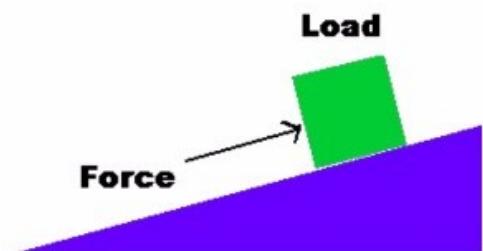
Lever



Pulley



Wheel & Axle



Inclined Plane

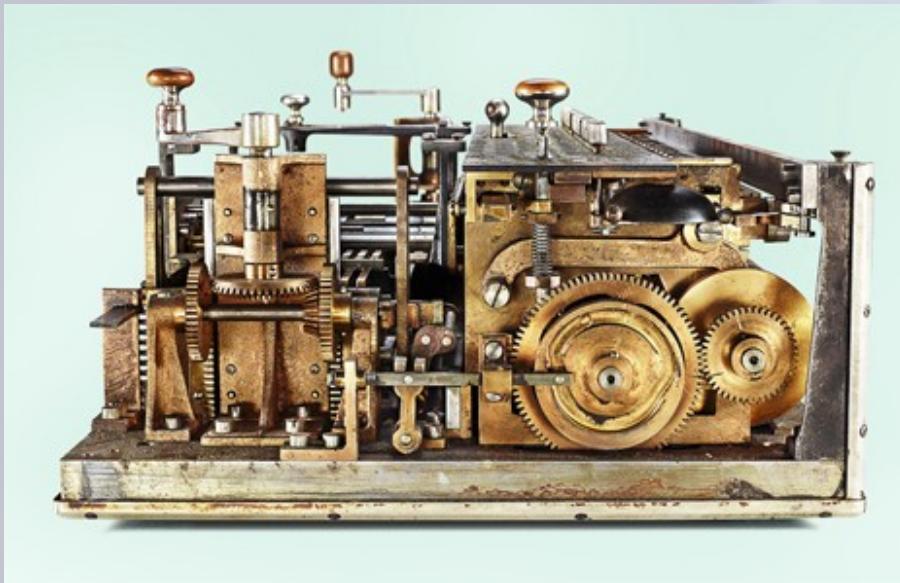


Wedge



Screw

# ماشین چیست؟



# ماشین چیست؟

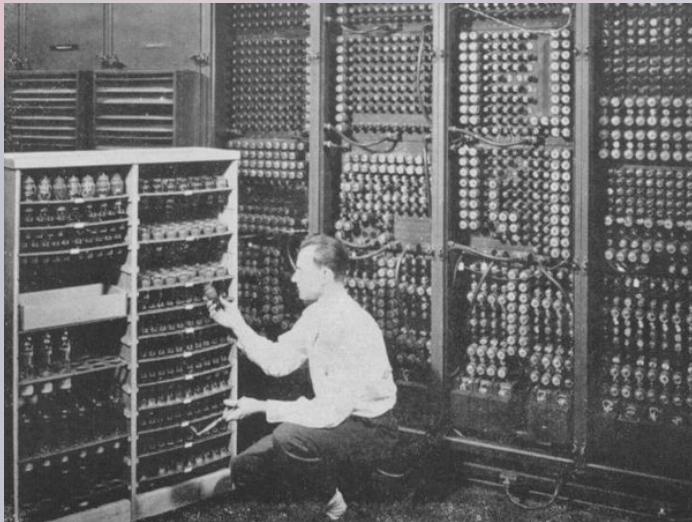


# ماشین چیست؟



- عمليات رياضي
- عمليات منطقى

# ماشین چیست؟



thanks  
for  
watching!

# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
2- What is Learning?

# یادگیری چیست؟



# یادگیری چیست؟

- اصلاح و تقویت
- رفتار
- دانش
- توانمندی
- انتخاب

# یادگیری چیست؟



# یادگیری چیست؟



# یادگیری چیست؟



# یادگیری چیست؟

- آنی نیست
- خودآگاهی نیاز ندارد

- کارکردگرایی
- شناختی
- تداعیگرایی
- نوروفیزیولوژیک
- تکاملی

نظریه های  
 مختلف

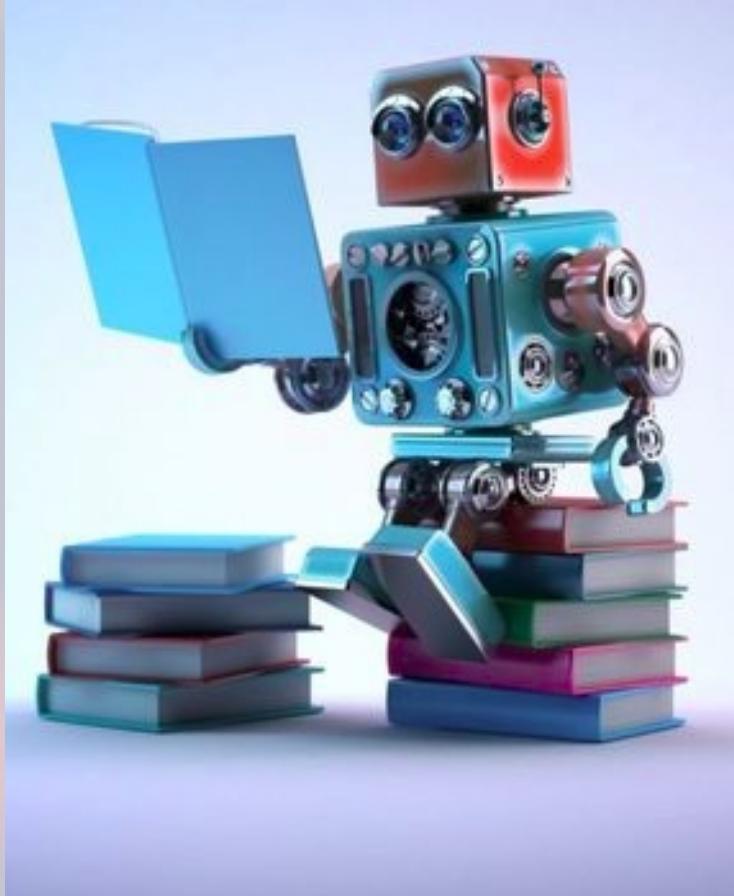
thanks  
for  
watching!

# MEDICAL DATA SCIENCE

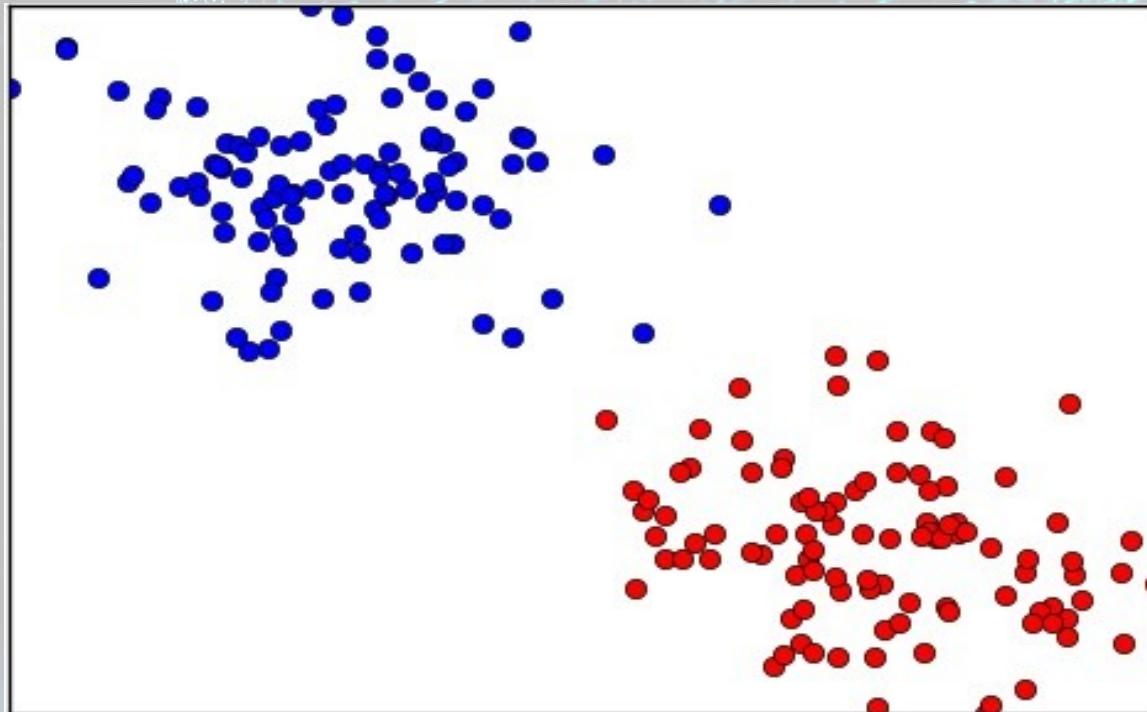
An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
3- What is ML?

# یادگیری ماشینی چیست؟

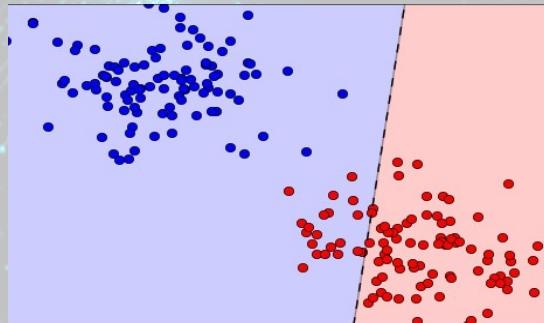
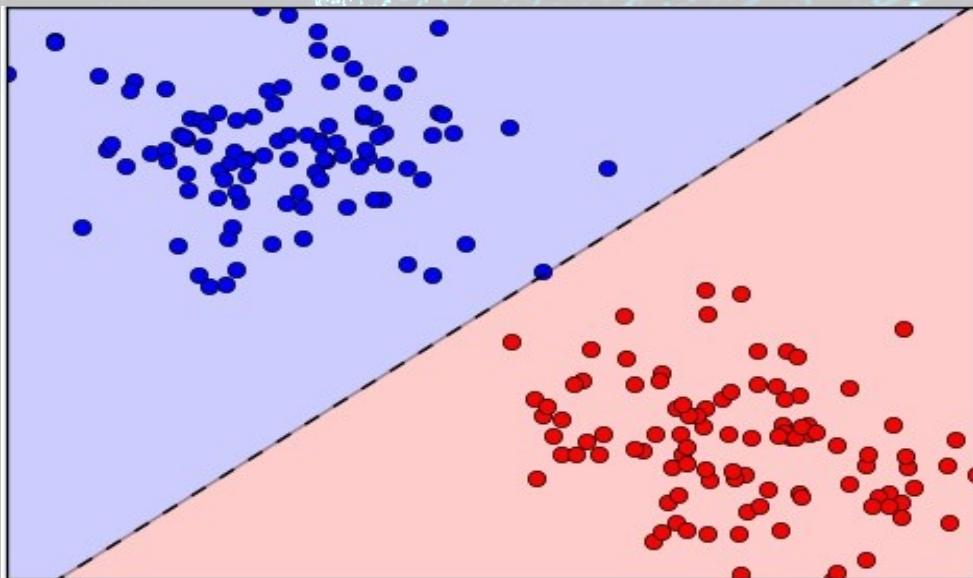


# یادگیری ماشینی چیست؟



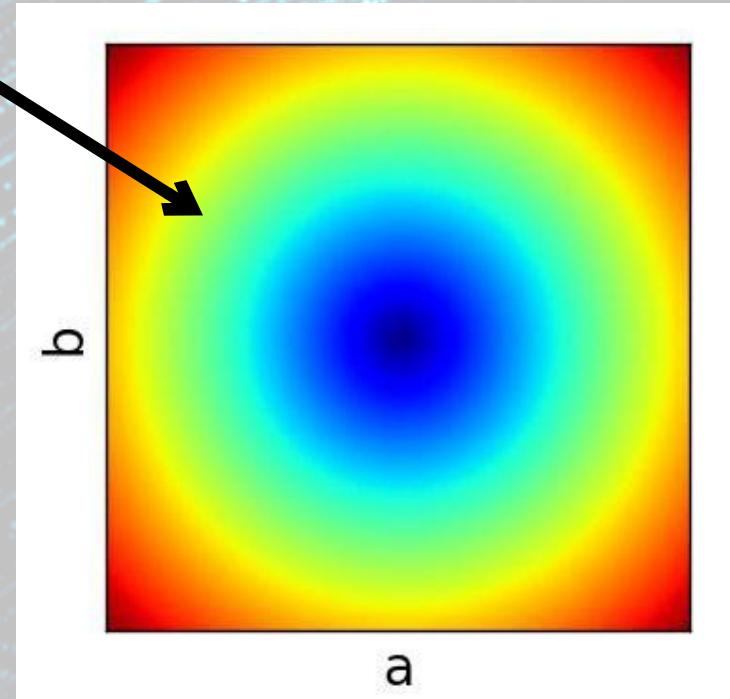
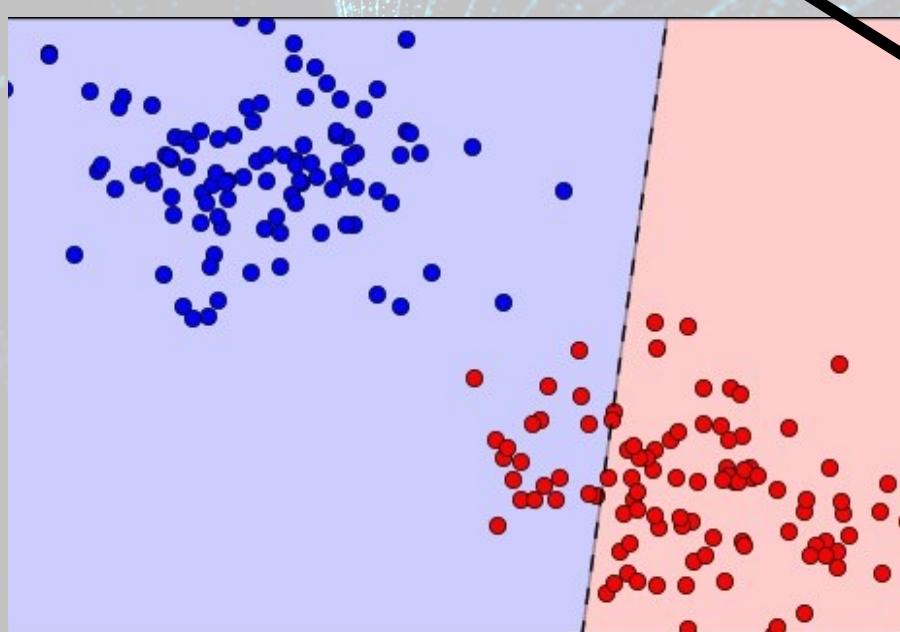
# یادگیری ماشینی چیست؟

$$y = ax + b$$



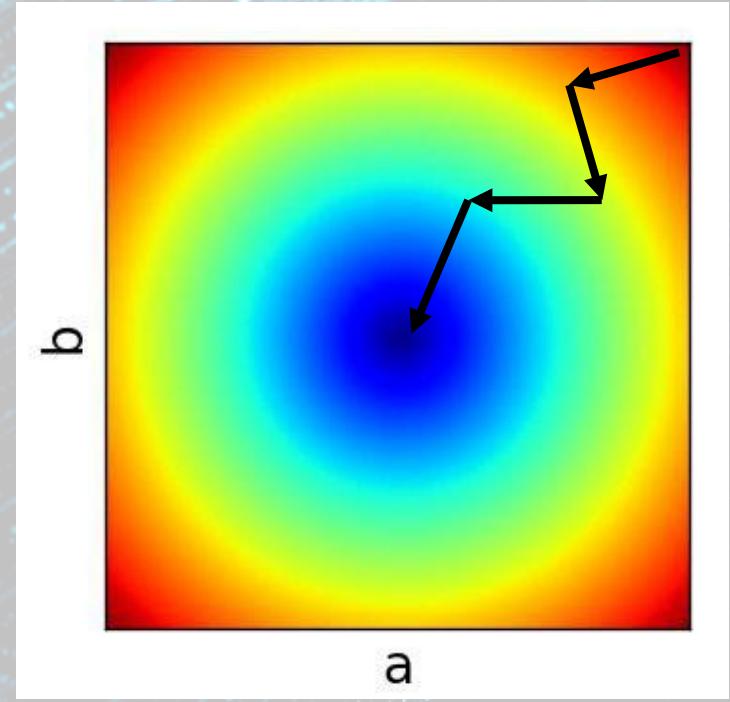
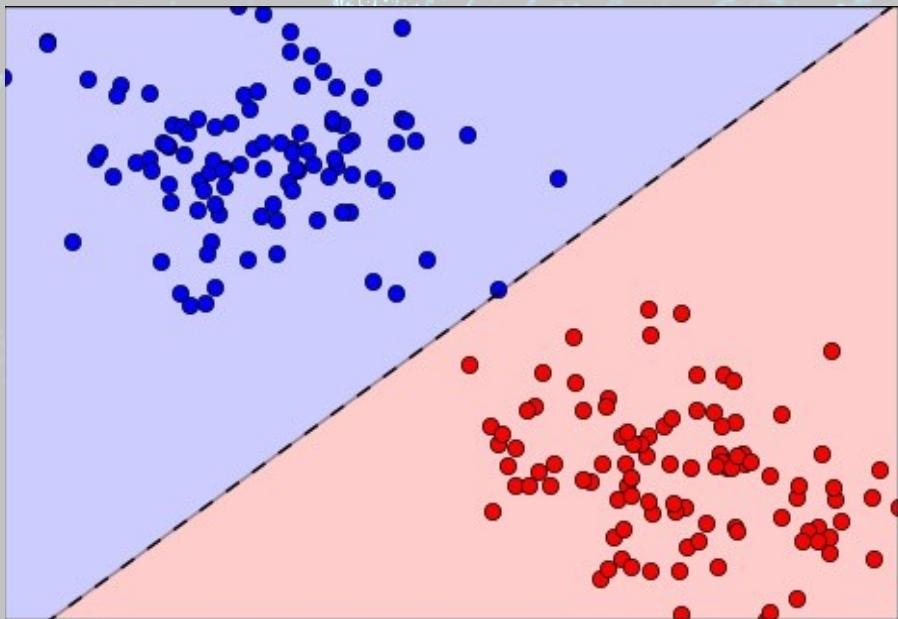
# یادگیری ماشینی چیست؟

$$y = ax + b$$

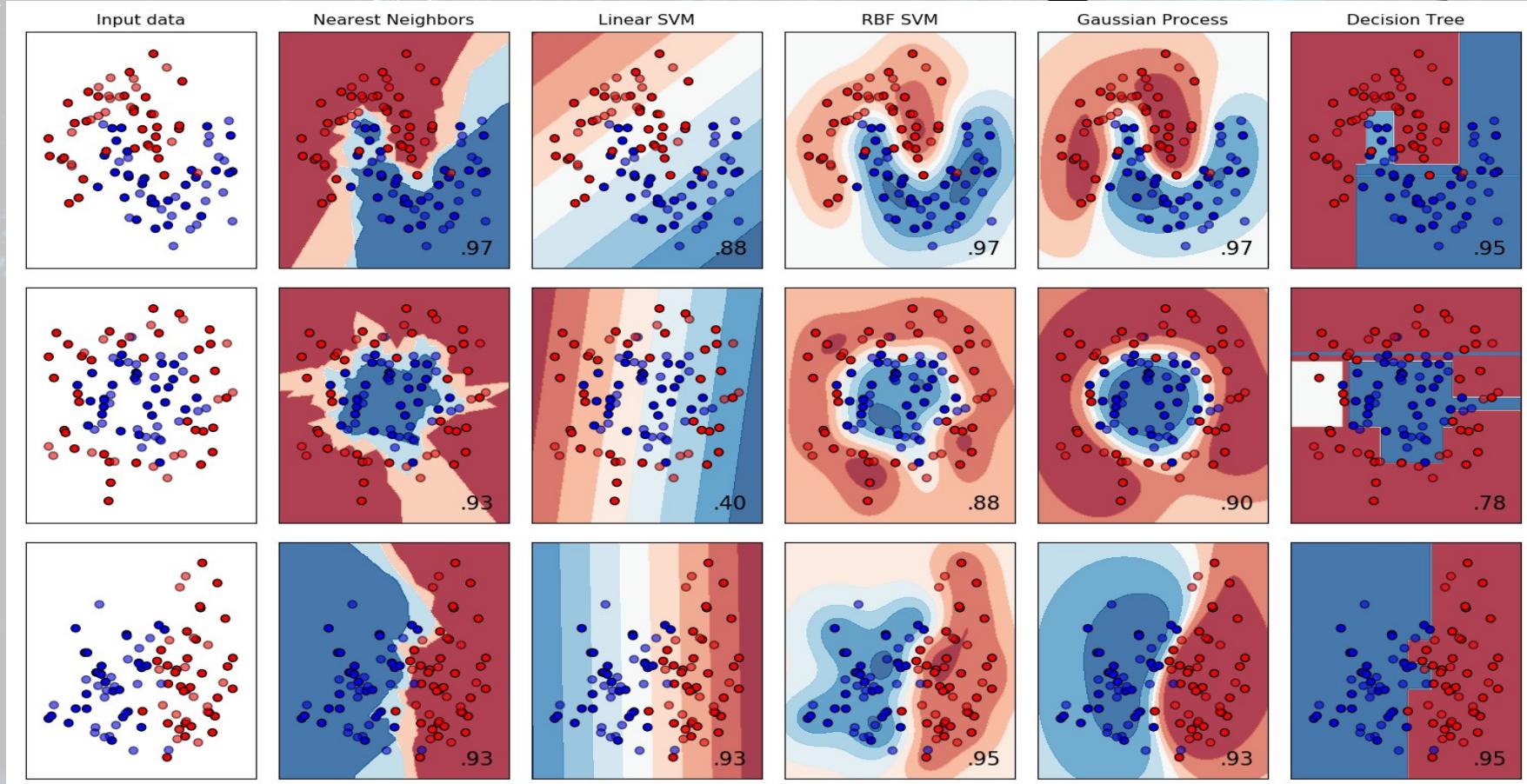


# یادگیری ماشینی چیست؟

$$y = ax + b$$



# یادگیری ماشینی چیست؟



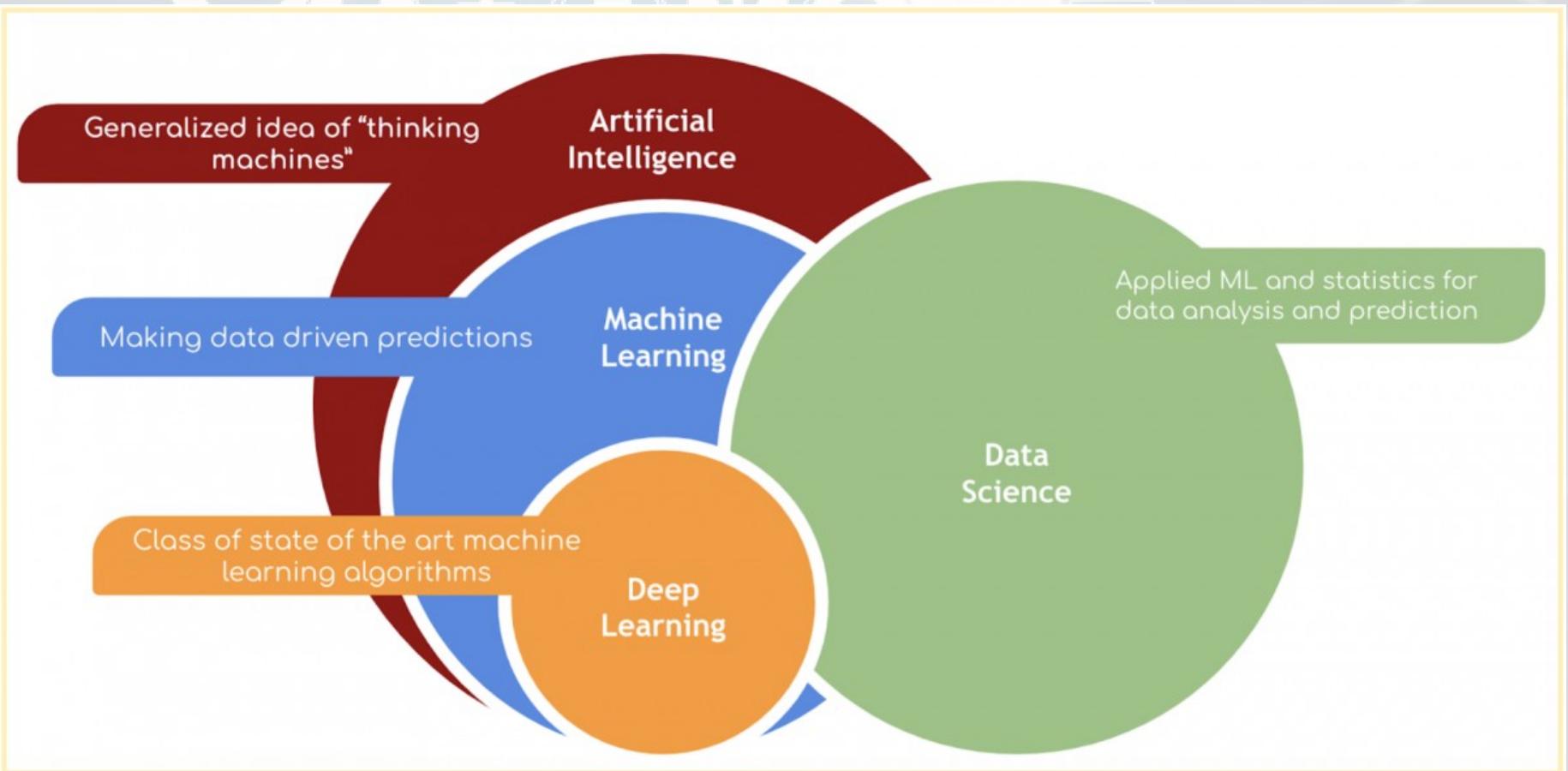
thanks  
for  
watching!

# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
4- What is AI then?

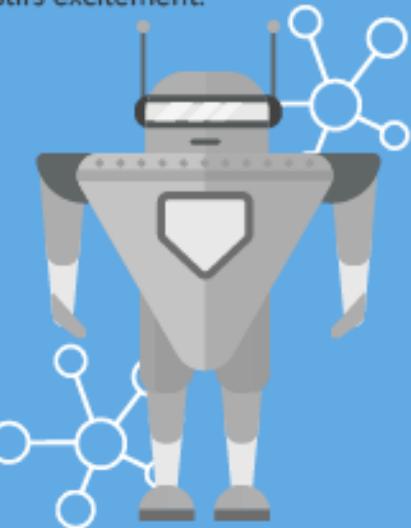
# DS vs ML vs DL



# DS vs ML vs DL

## ARTIFICIAL INTELLIGENCE

Early artificial intelligence stirs excitement.



1950's

1960's

1970's

1980's

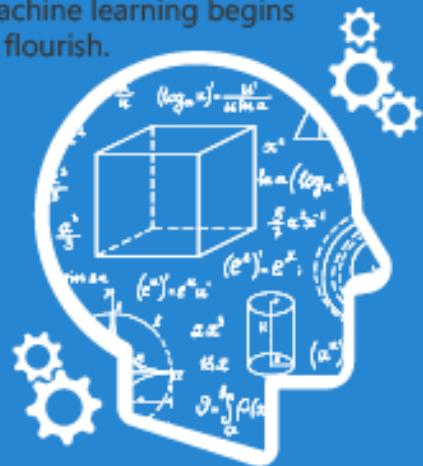
1990's

2000's

2010's

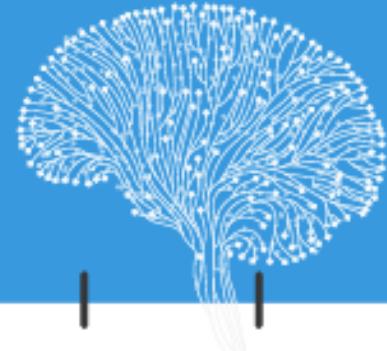
## MACHINE LEARNING

Machine learning begins to flourish.



## DEEP LEARNING

Deep learning breakthroughs drive AI boom.



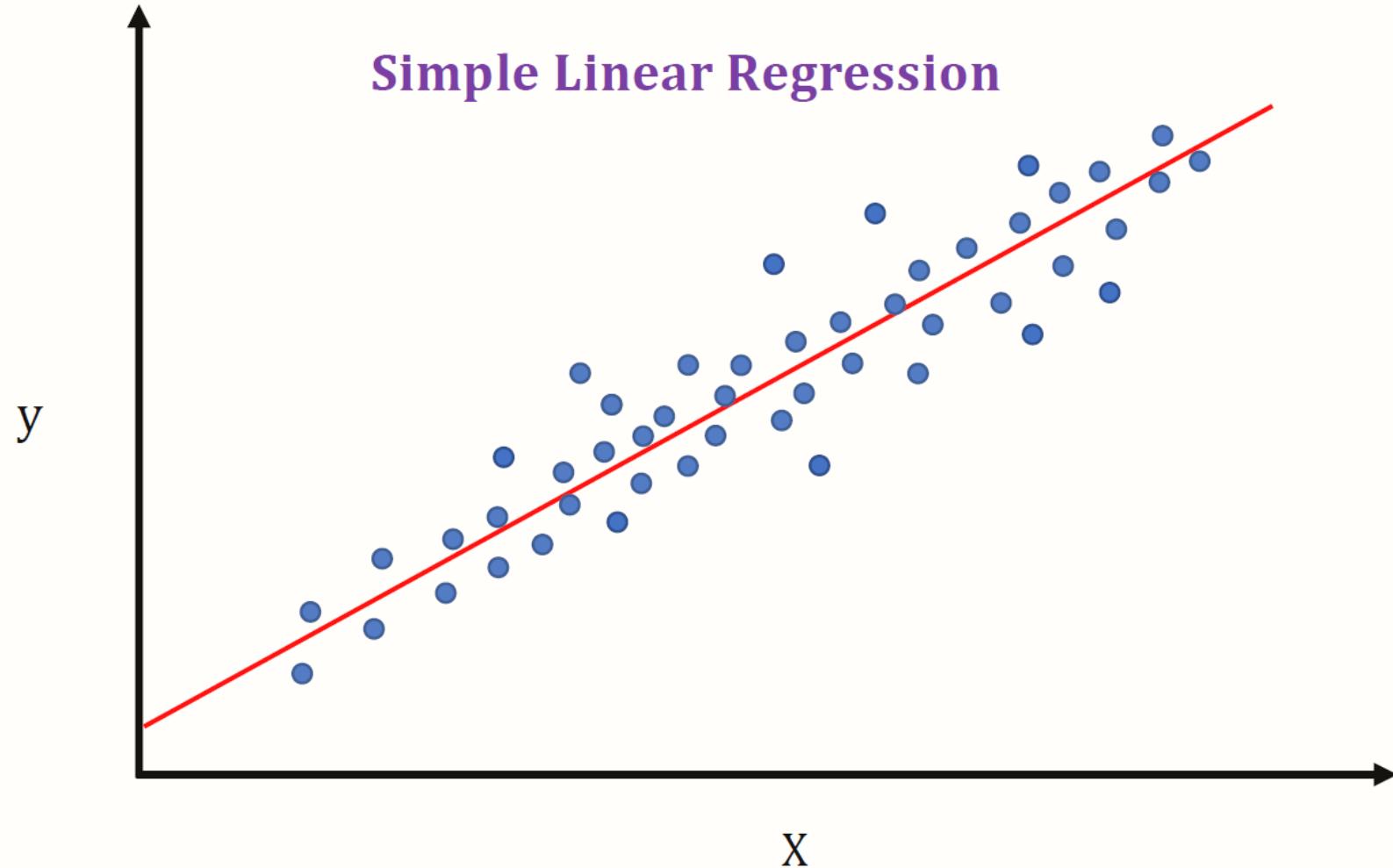
thanks  
for  
watching!

# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
5- Linear Regression

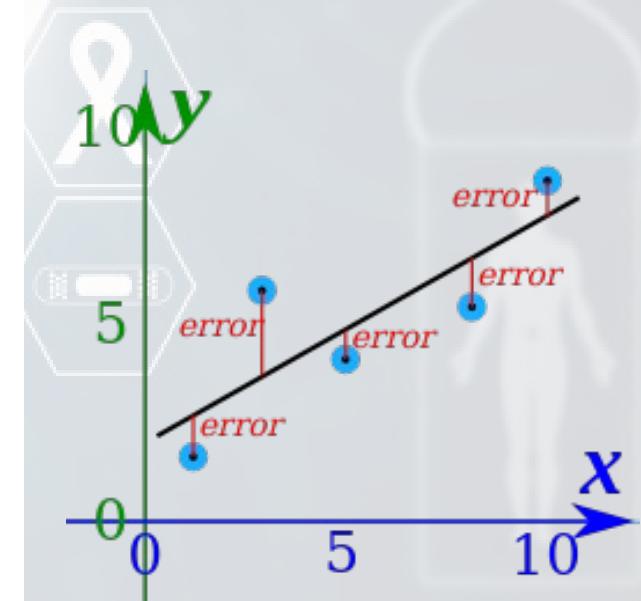
## Simple Linear Regression



## Independed variable

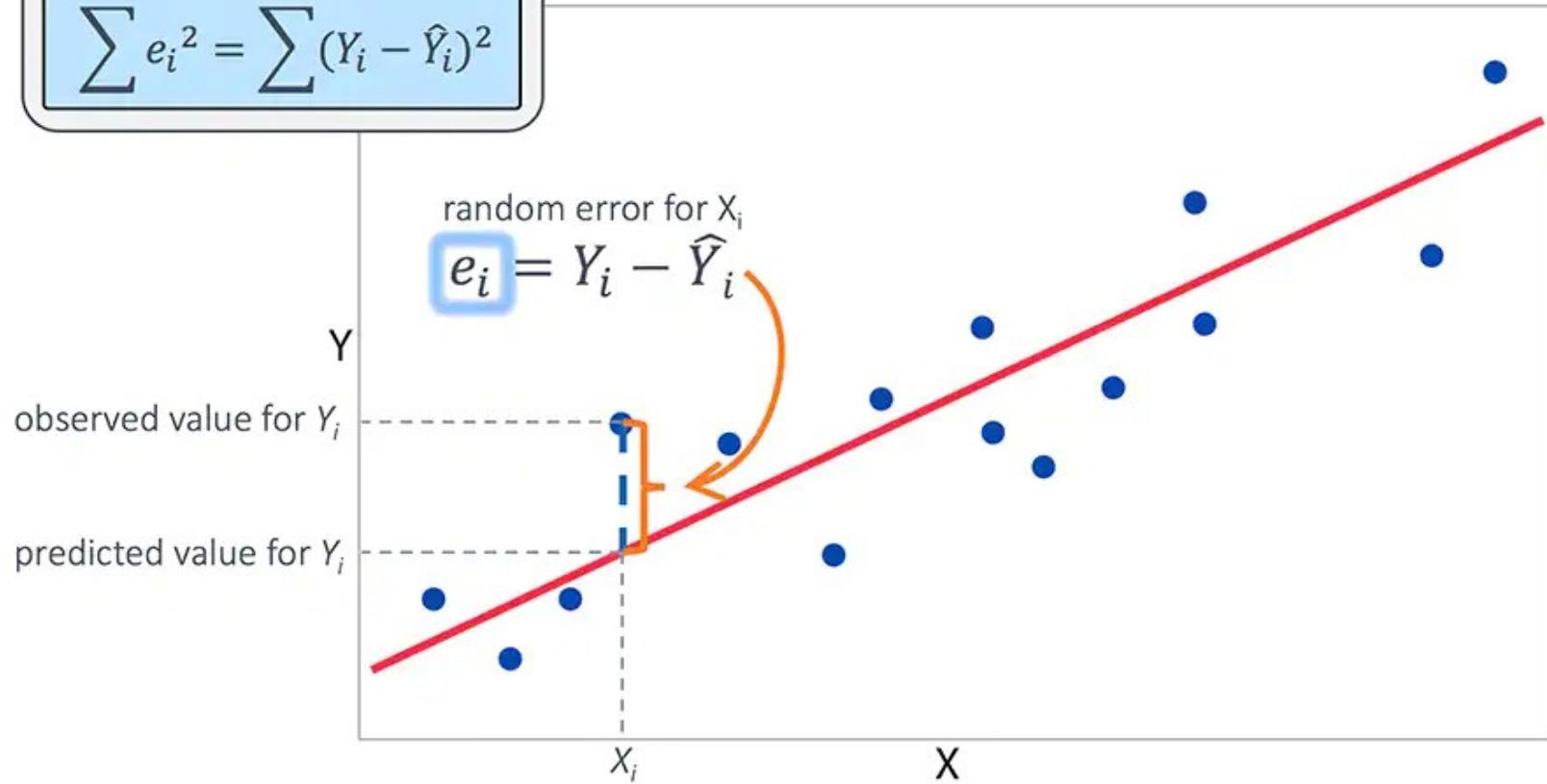
$$y = xw + b$$

↑  
Depended variable  
↓  
Intercept (bias)  
↑  
Slope



### Method of Least Squares

$$\sum e_i^2 = \sum (Y_i - \hat{Y}_i)^2$$



thanks  
for  
watching!

# MEDICAL DATA SCIENCE

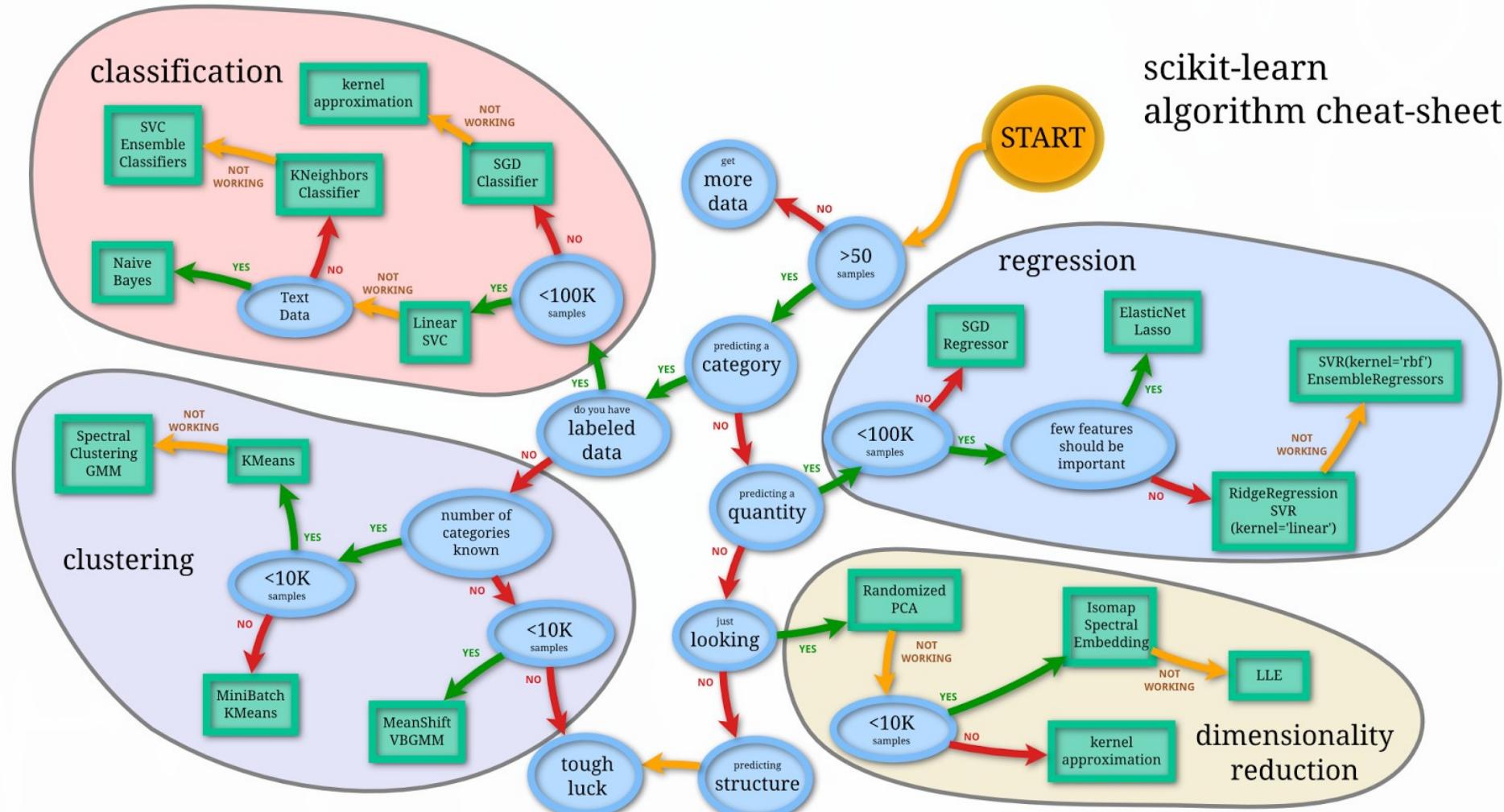
An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
6- SciKit-Learn



FOR  
DATA  
SCIENCE

# scikit-learn algorithm cheat-sheet



# Let's Python!



thanks  
for  
watching!

# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
6.5- Data Preparation

# Let's Python!



thanks  
for  
watching!

# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
7- Training a model

# Let's Python!



thanks  
for  
watching!

# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
8- Predict, Evaluate

# Let's Python!



thanks  
for  
watching!

# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
9- Regression Metrics

$$R^2 = 1 - \frac{\sum_{i=1}^n (\gamma_i - \hat{\gamma}_i)^2}{\sum_{i=1}^n (\gamma_i - \bar{\gamma})^2}$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right|$$

# Let's Python!



thanks  
for  
watching!

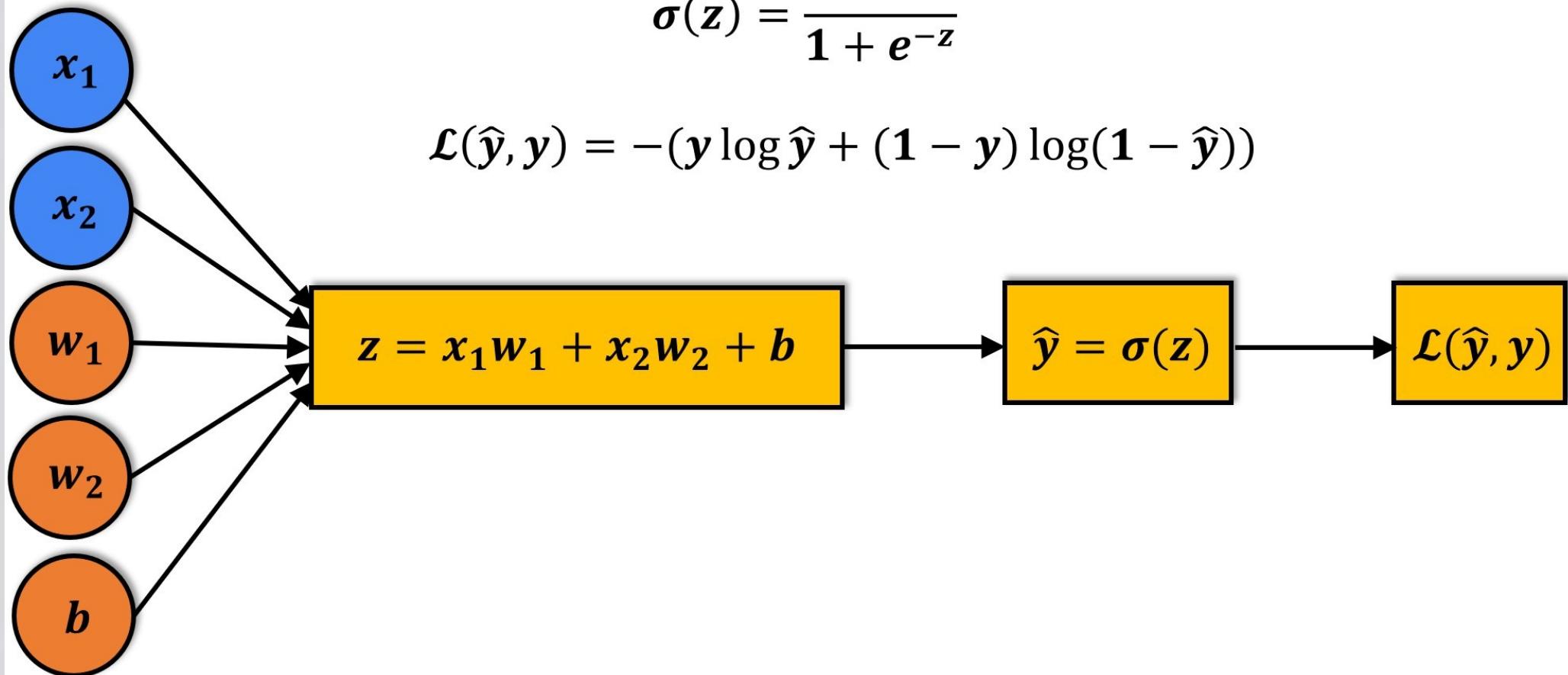
# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
10- Classification:  
Logistic regression

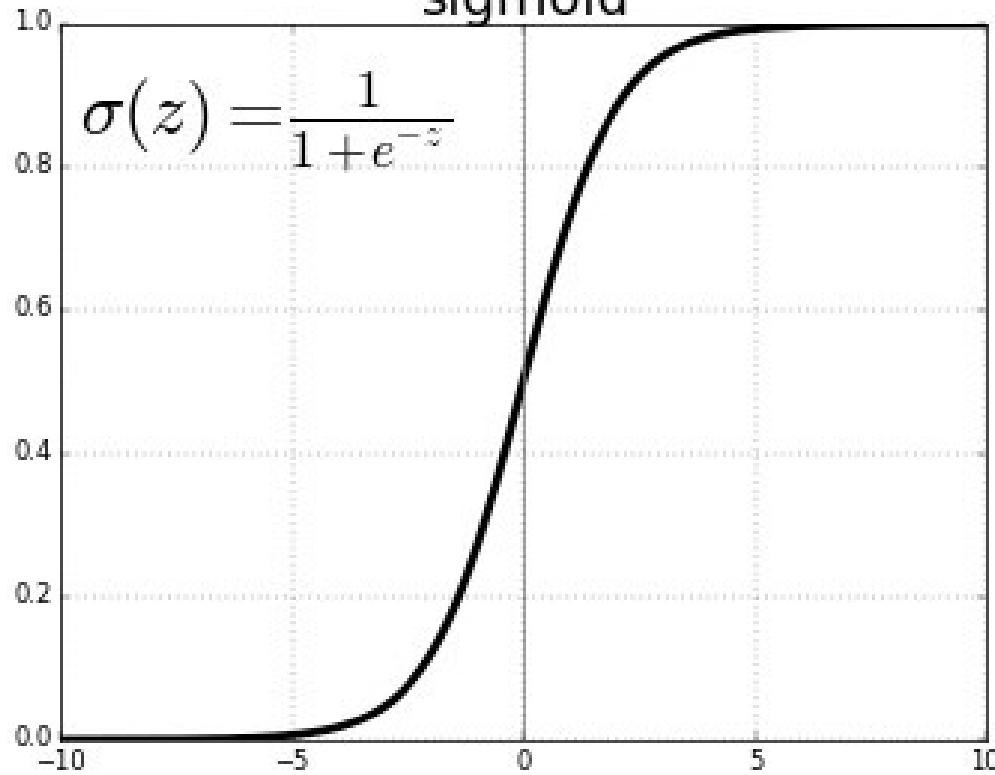
$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

$$\mathcal{L}(\hat{y}, y) = -(y \log \hat{y} + (1 - y) \log(1 - \hat{y}))$$



## sigmoid

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$



# Let's Python!



thanks  
for  
watching!

# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
11- Classification Metrics

relevant elements

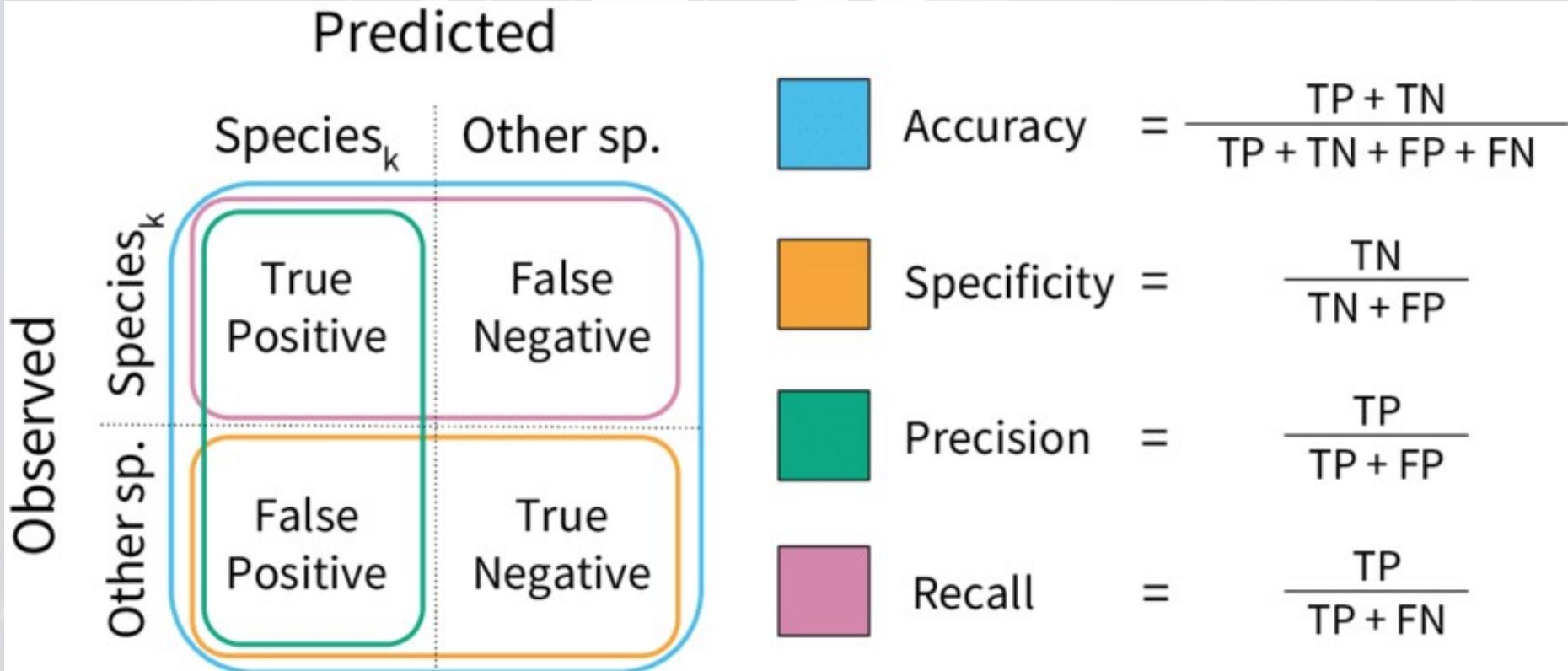
false negatives

true negatives

true positives

false positives

selected elements



	Sun	Cloud	Rain	Snow
Sun	TP	FN	FP	FP
Cloud	FP	TP	FP	FP
Rain	FP	FN	TP	FP
Snow	FP	FN	FP	TP

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

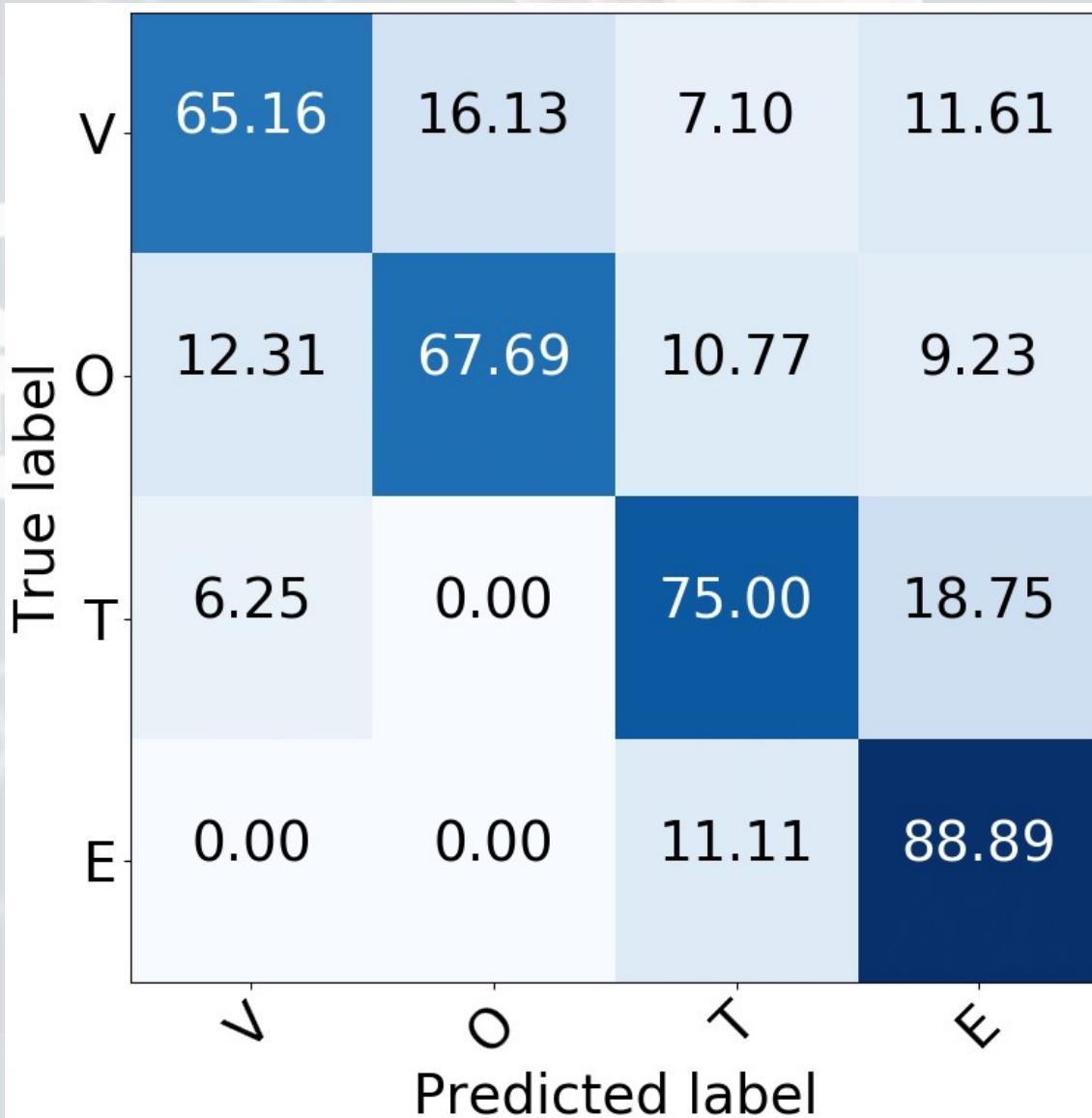
$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

Observed

	Sun	Cloud	Rain	Snow
Sun	TP	FP	FN	FP
Cloud	FP	TP	FP	FP
Rain	FN	TP	FP	FP
Snow	FP	FN	FP	FP

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$



# Let's Python!



thanks  
for  
watching!

# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
12- Predictive models

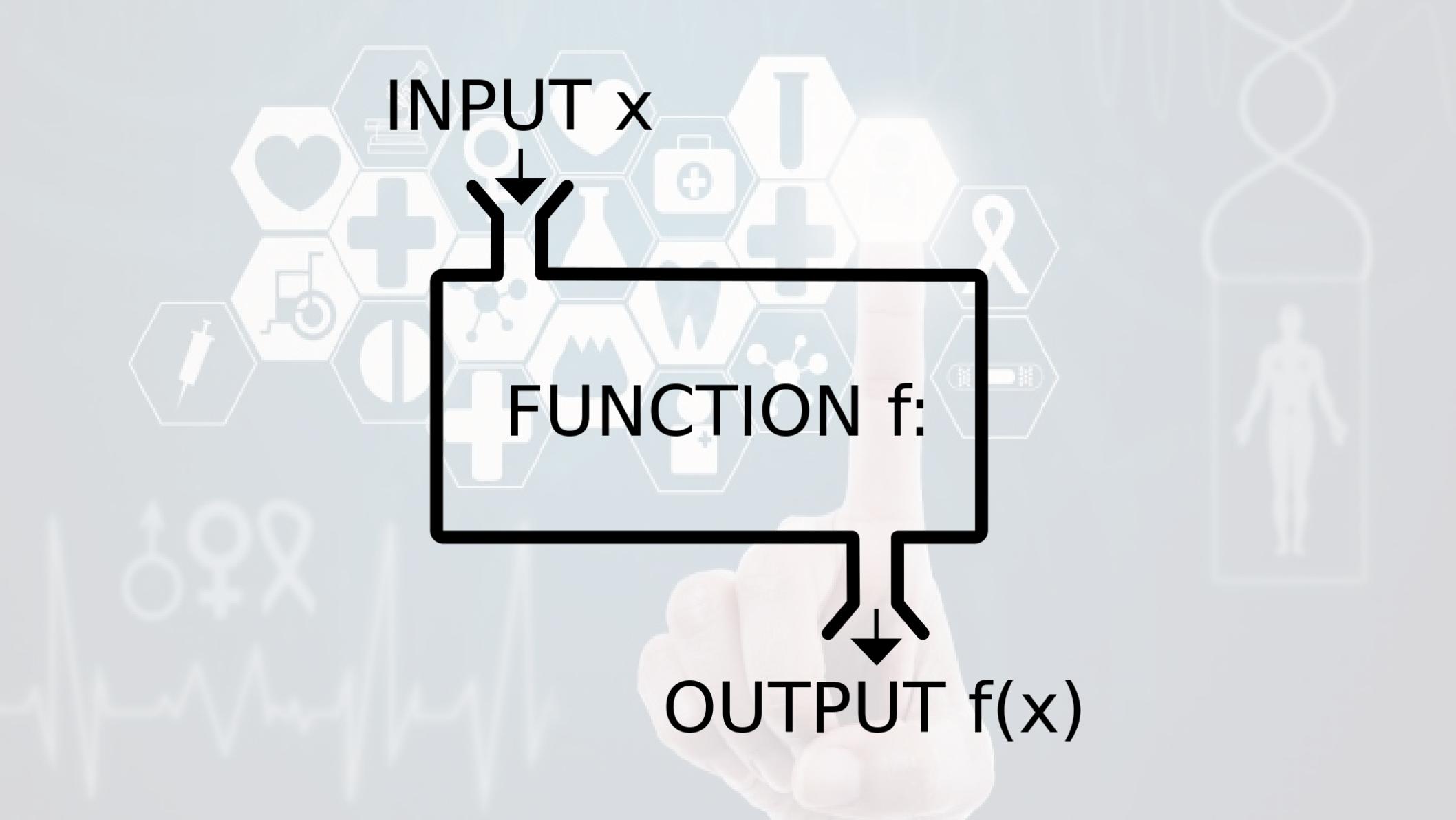
Input  
Value

$f(x)$  =  $x - 5$

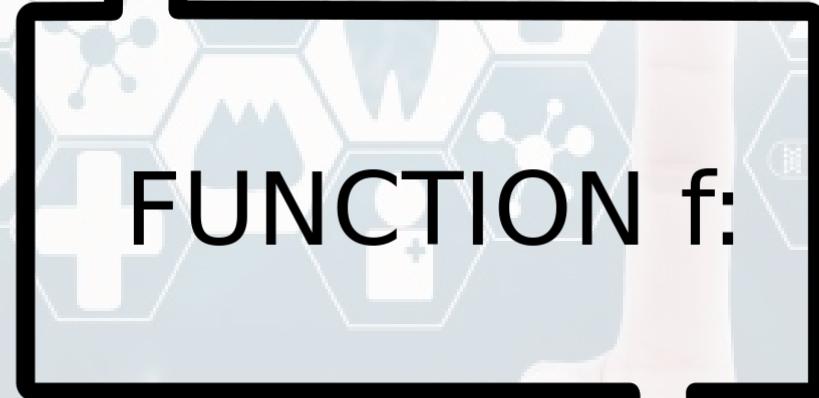
Function  
Name

Output  
Value

$$f(x) = x - 5$$



**INPUT  $x$**



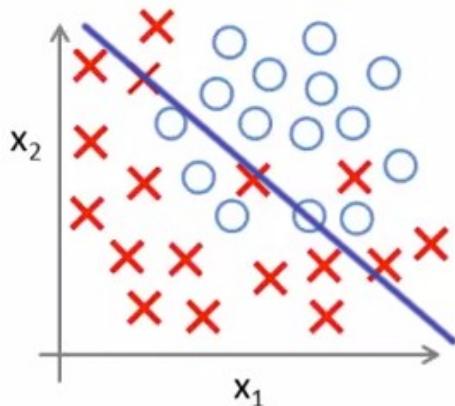
**OUTPUT  $f(x)$**

thanks  
for  
watching!

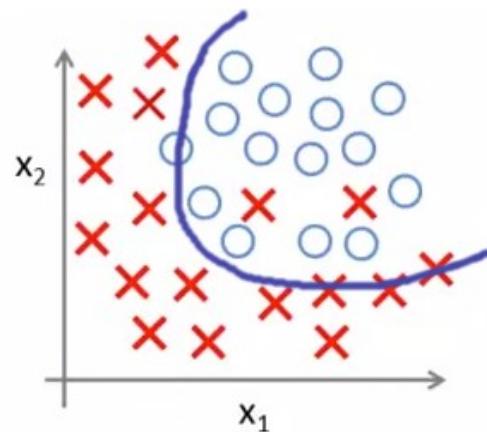
# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

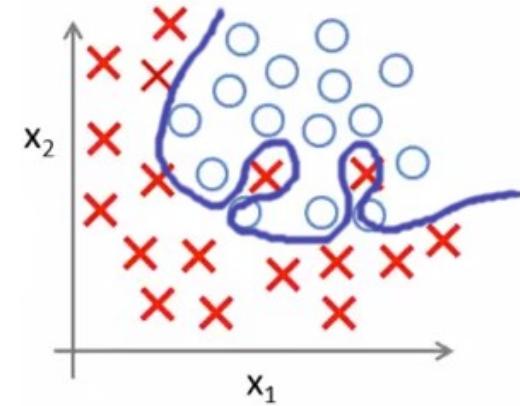
By: Alireza Vafaei Sadr  
13- Overfit



$$h_{\theta}(x) = g(\theta_0 + \theta_1 x_1 + \theta_2 x_2)$$

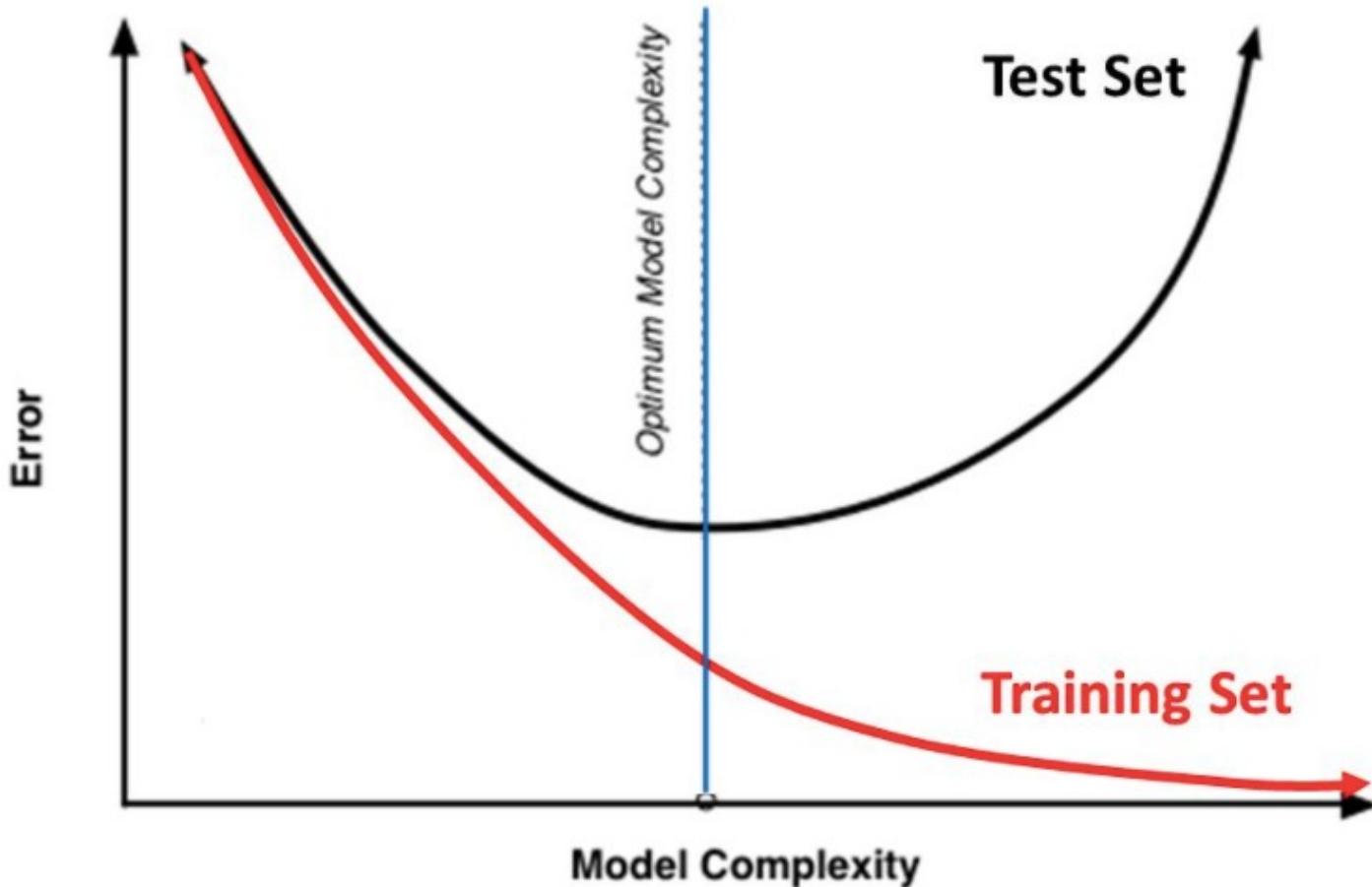


$$\begin{aligned}g(\theta_0 + \theta_1 x_1 + \theta_2 x_2 \\+ \theta_3 x_1^2 + \theta_4 x_2^2 \\+ \theta_5 x_1 x_2)\end{aligned}$$



$$\begin{aligned}g(\theta_0 + \theta_1 x_1 + \theta_2 x_1^2 \\+ \theta_3 x_1^2 x_2 + \theta_4 x_1^2 x_2^2 \\+ \theta_5 x_1^2 x_2^3 + \theta_6 x_1^3 x_2 + \dots)\end{aligned}$$

# Training Vs. Test Set Error



# Let's Python!

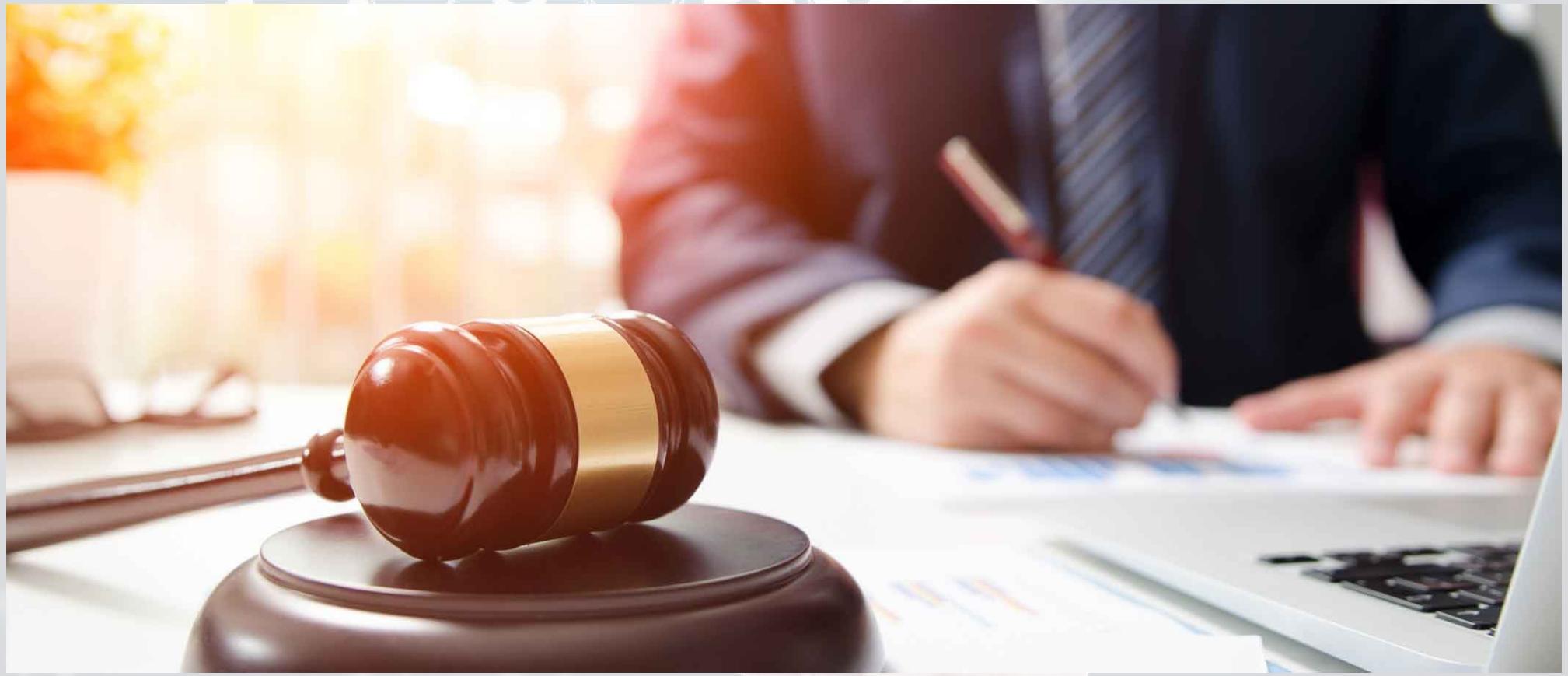


thanks  
for  
watching!

# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
14- Regularization





$$L1 \text{ Regularization} = (\text{loss function}) + \alpha \sum_{j=1}^p |b_j|$$

$$L2 \text{ Regularization} = (\text{loss function}) + \alpha \sum_{j=1}^p b_j^2$$

# Let's Python!

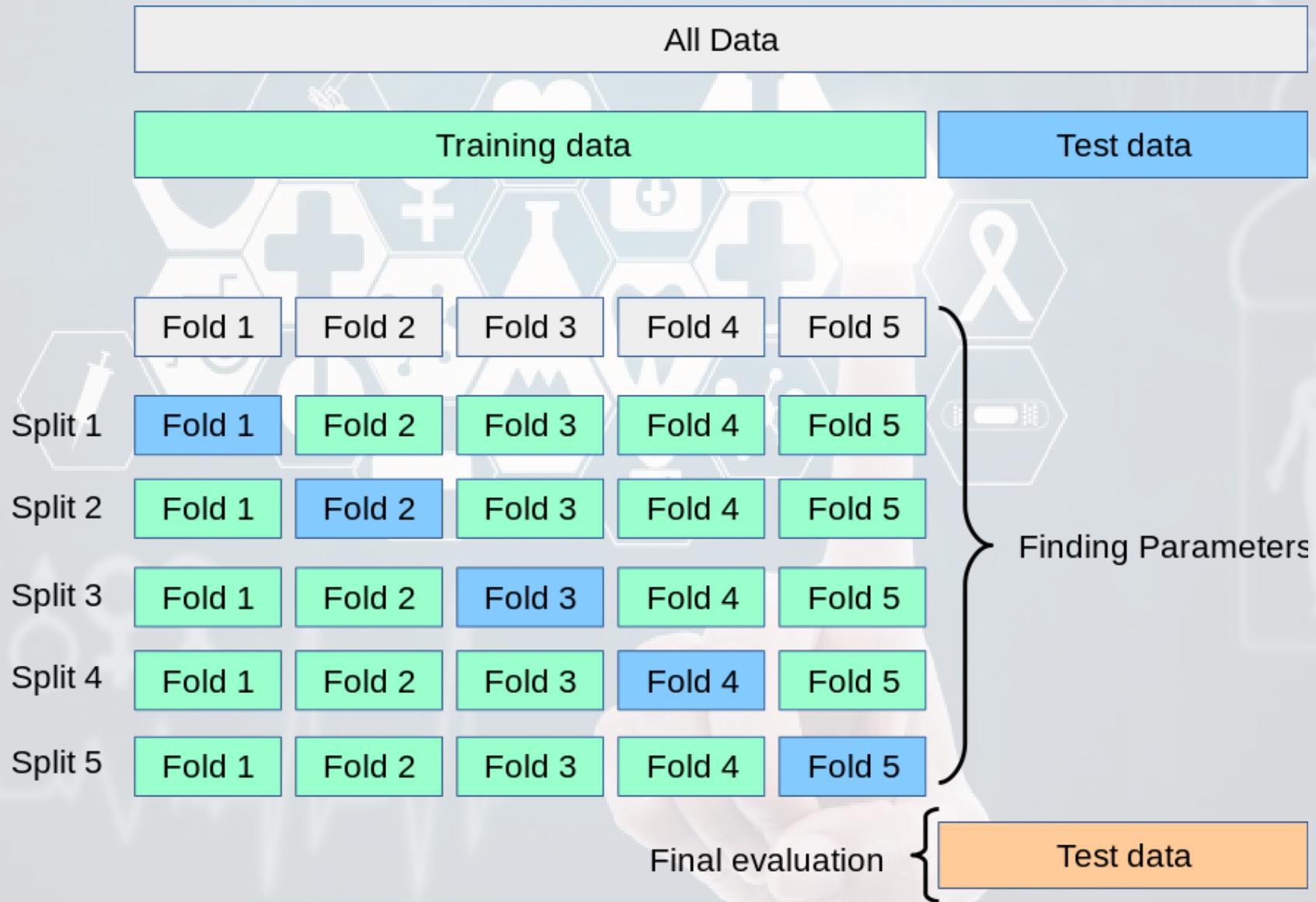


thanks  
for  
watching!

# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
15- Cross Validation



# Let's Python!

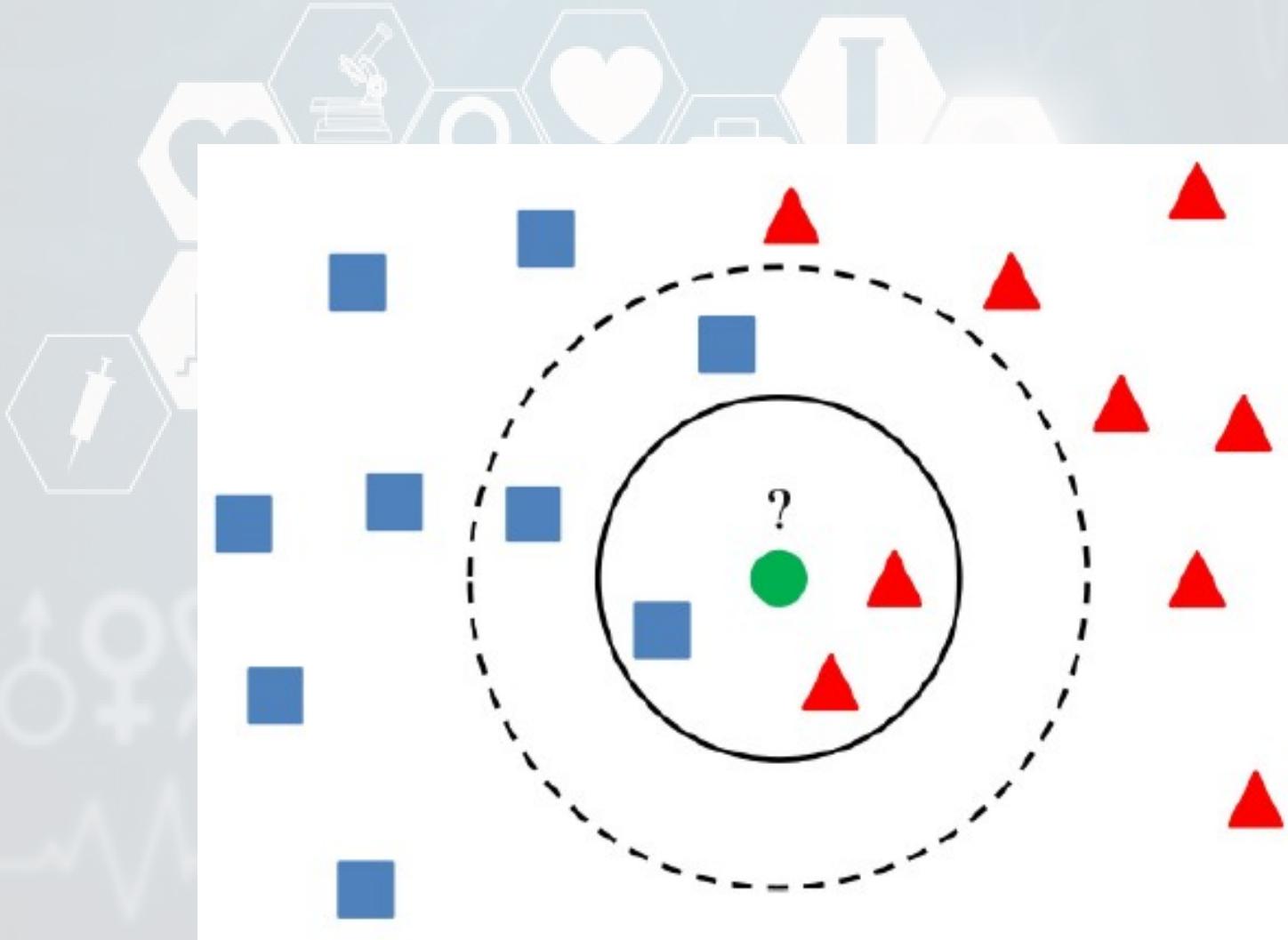


thanks  
for  
watching!

# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
16- K-Nearest Neighbours



# Let's Python!

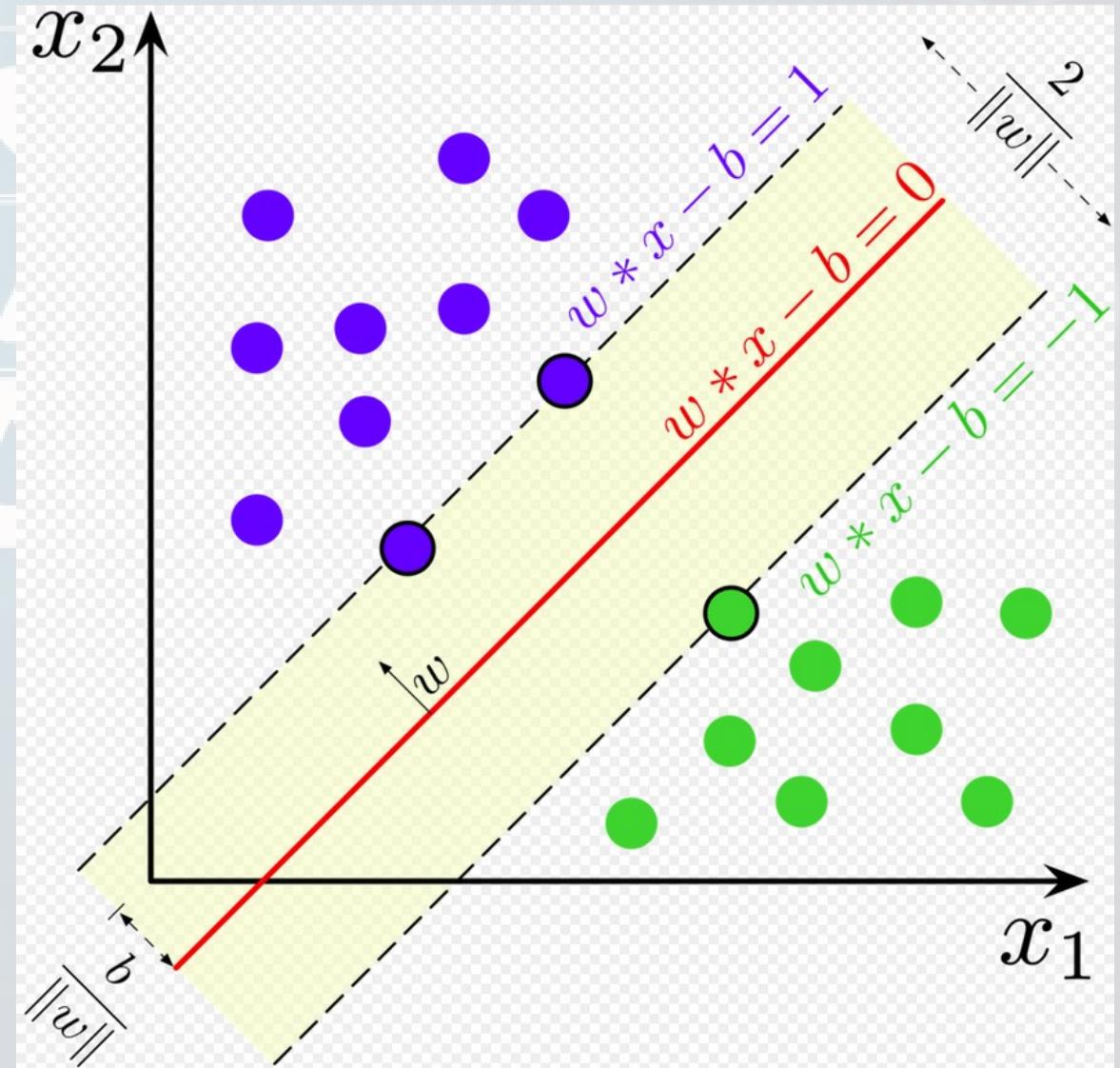
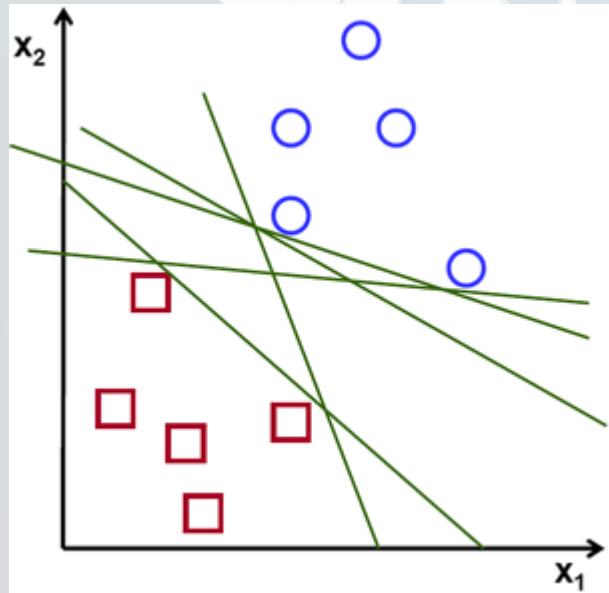


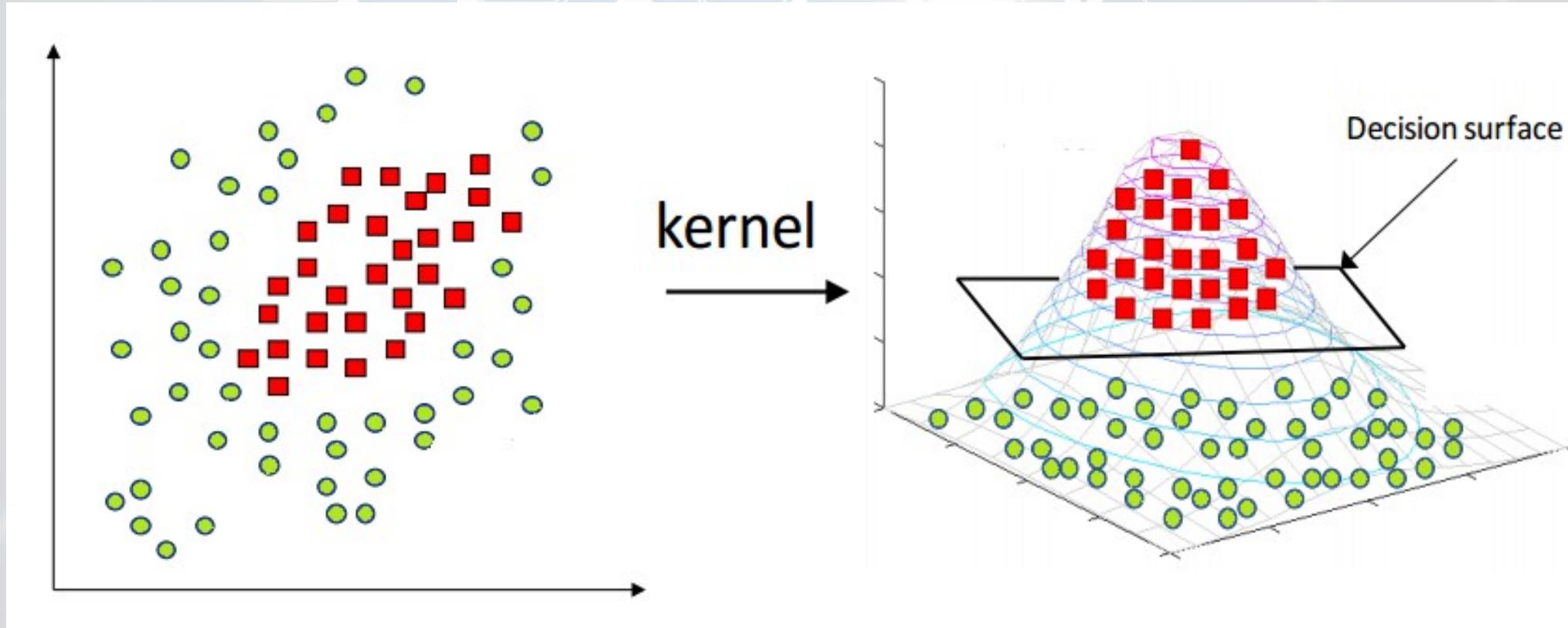
thanks  
for  
watching!

# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
17- Support Vector Machines





# Let's Python!

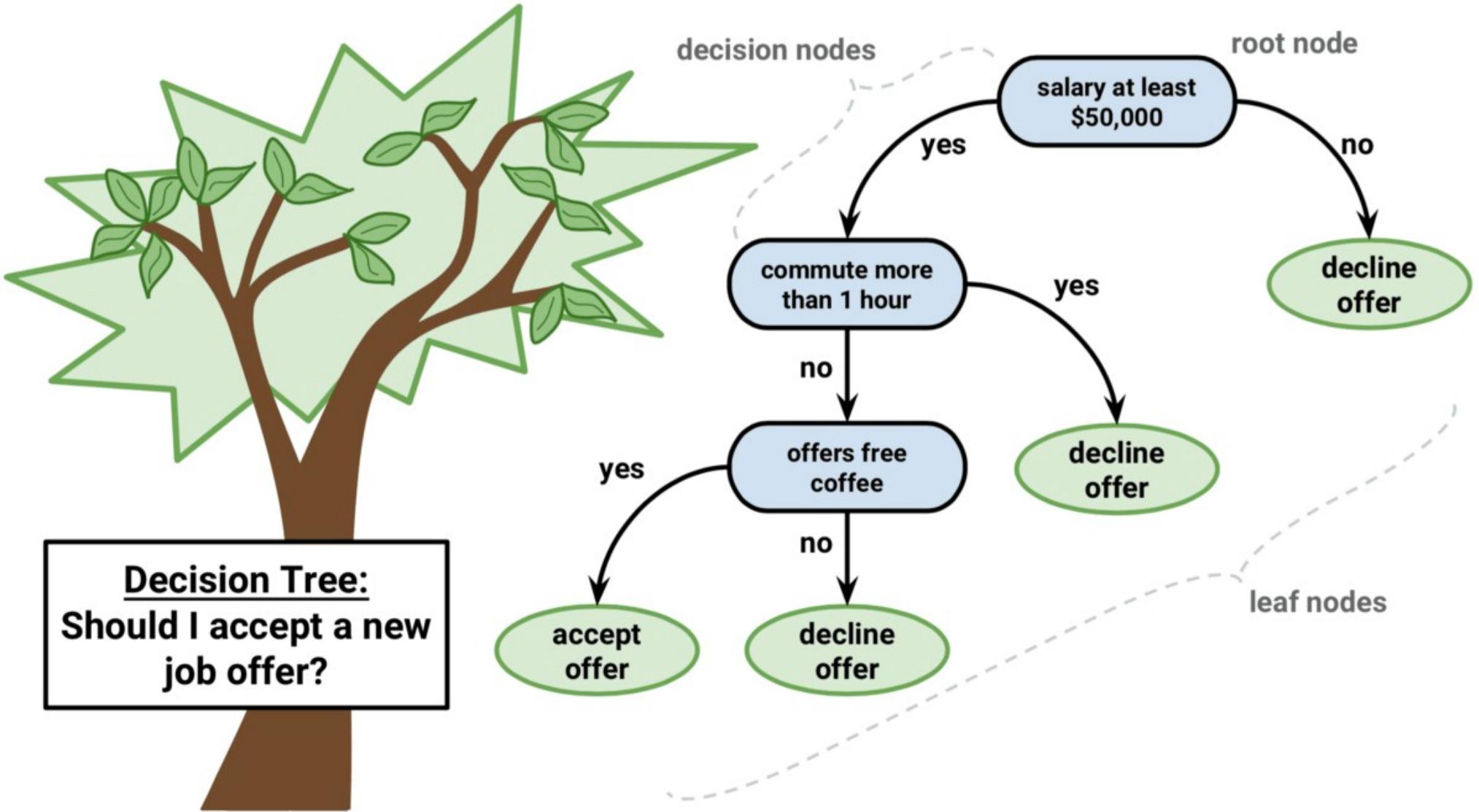


thanks  
for  
watching!

# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
18- Decision tree



# Let's Python!

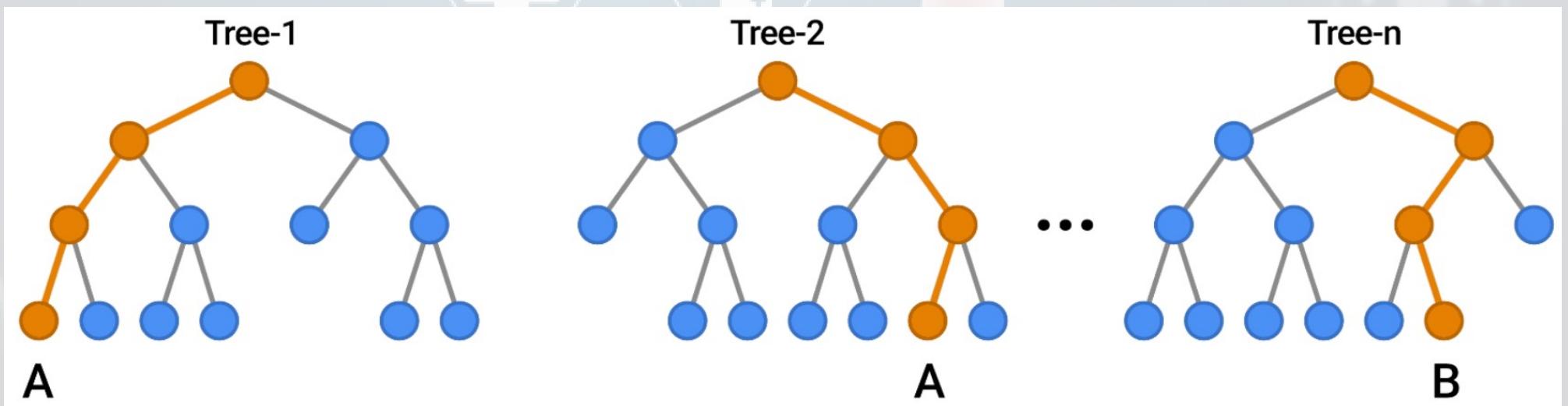
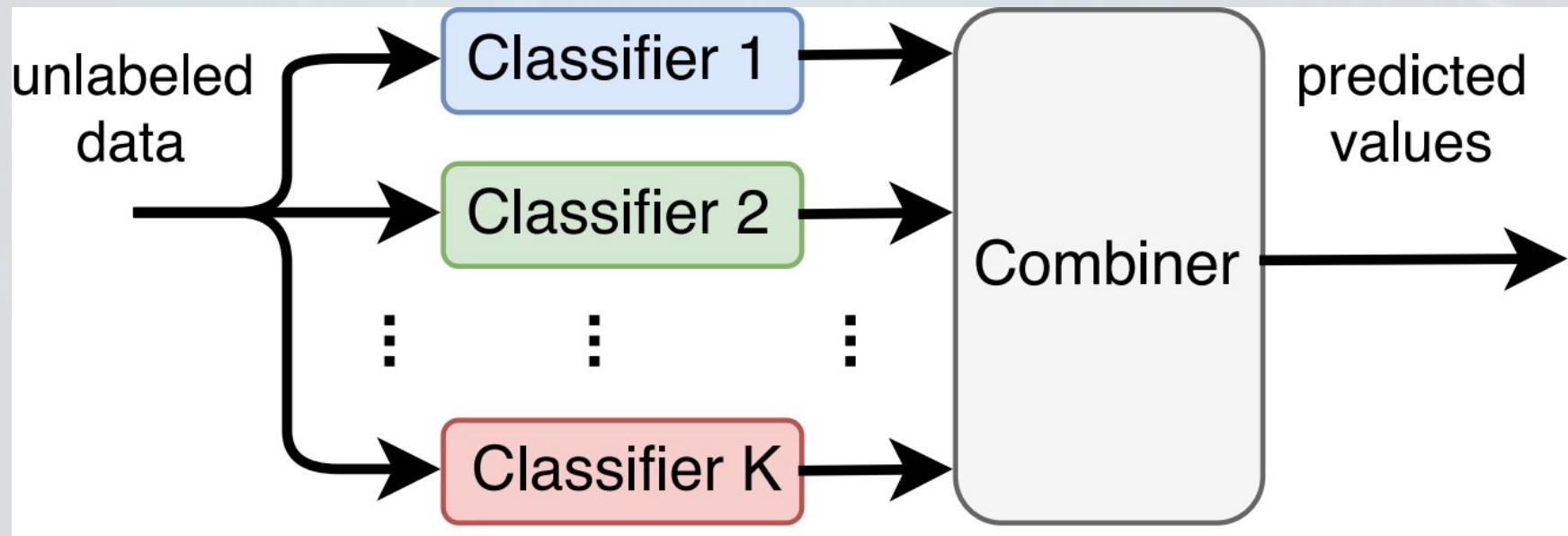


thanks  
for  
watching!

# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
19- Ensemble Learning  
Random Forests



# Let's Python!

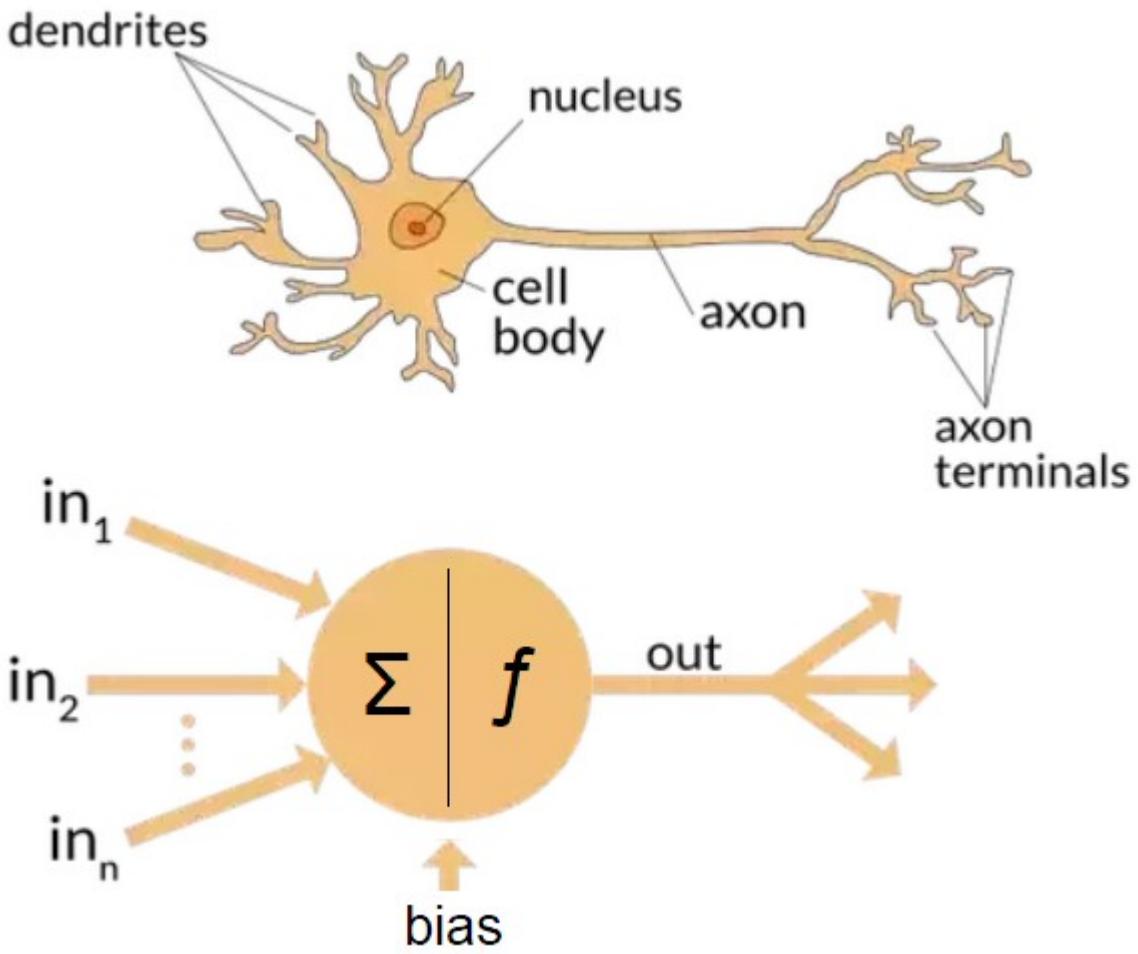


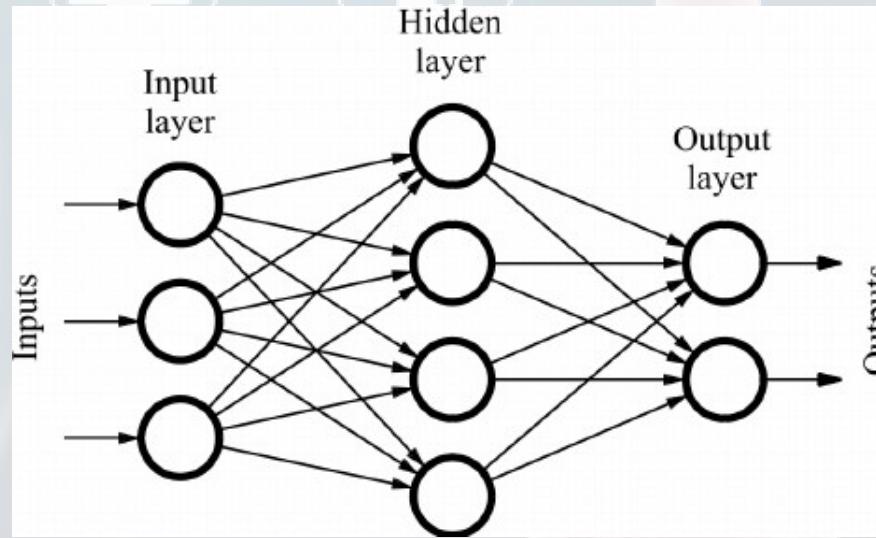
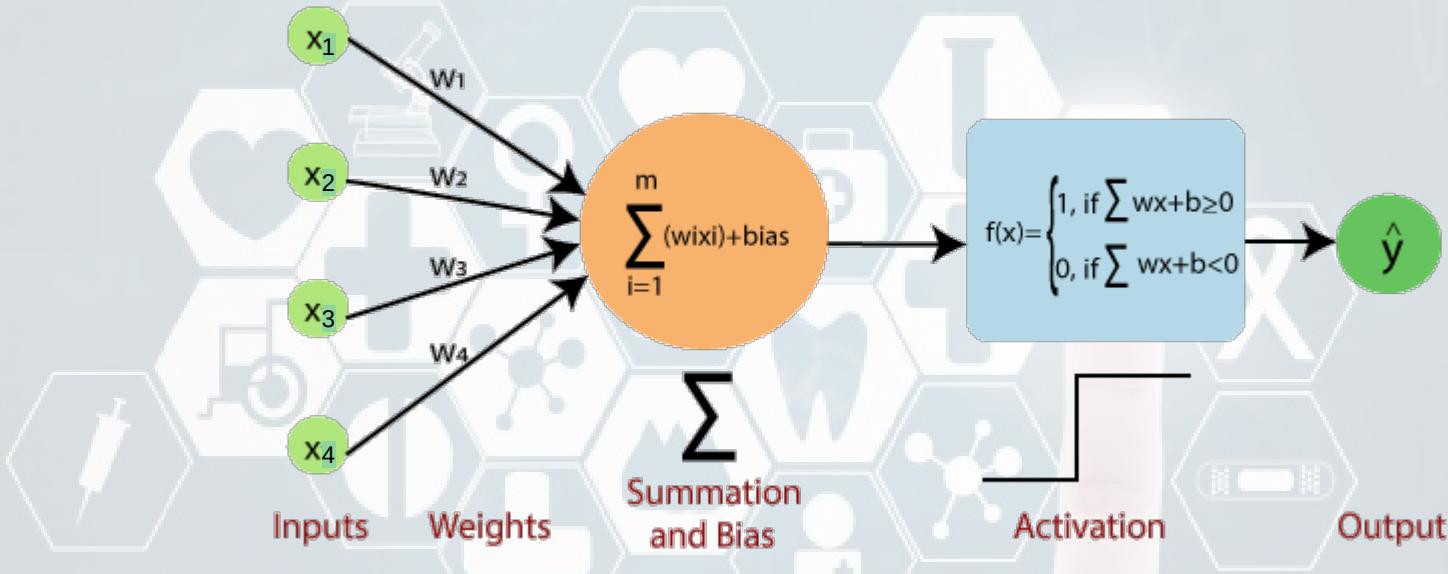
thanks  
for  
watching!

# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
19- Neural Networks I  
Dense layers





# Let's Python!



thanks  
for  
watching!

# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
20- Keras I  
Dense layers

# Let's Python!

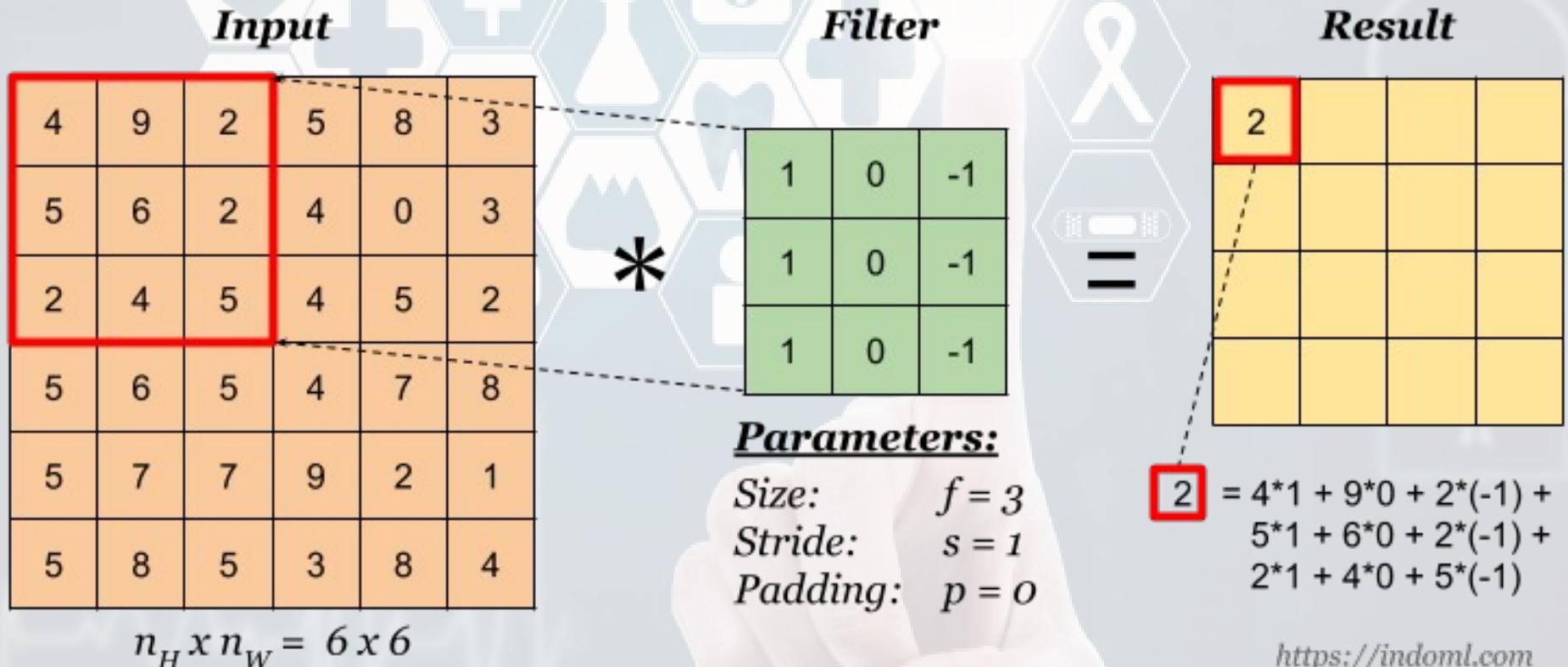


thanks  
for  
watching!

# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
21- Neural Networks II  
Convolutions



thanks  
for  
watching!

# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
22- MNIST!

0      1      2      3      4      5      6      7      8      9      0

0 1 2 3 4 5 6 7 8 9

0 1 2 3 4 5 6 7 8 9

0 1 2 3 4 5 6 7 8 9

0 1 2 3 4 5 6 7 8 9

0 1 2 3 4 5 6 7 8 9

0 1 2 3 4 5 6 7 8 9

0 1 2 3 4 5 6 7 8 9

0 1 2 3 4 5 6 7 8 9

0 1 2 3 4 5 6 7 8 9

0 1 2 3 4 5 6 7 8 9

# Let's Python!



thanks  
for  
watching!

# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
23- Keras II  
Convolutions

# Let's Python!



thanks  
for  
watching!

# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

By: Alireza Vafaei Sadr  
24- Augmentation



(a) Original



(b) Crop and resize



(c) Crop, resize (and flip)



(d) Color distort. (drop)



(e) Color distort. (jitter)



(f) Rotate  $\{90^\circ, 180^\circ, 270^\circ\}$



(g) Cutout



(h) Gaussian noise



(i) Gaussian blur



(j) Sobel filtering

# Let's Python!

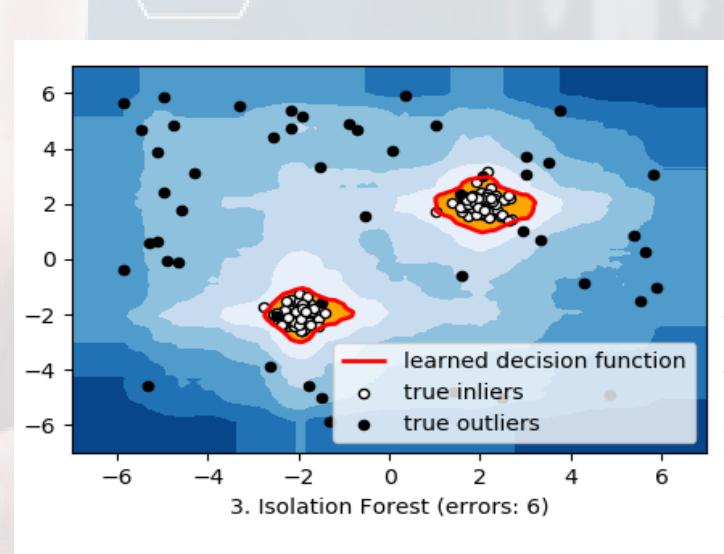
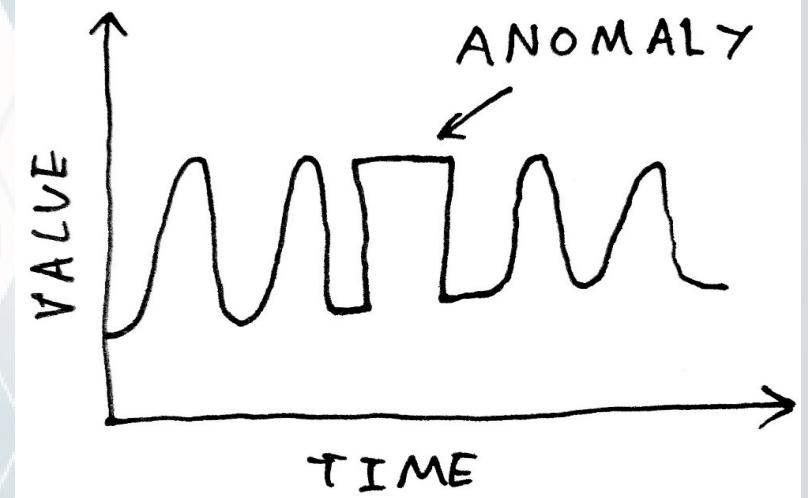


thanks  
for  
watching!

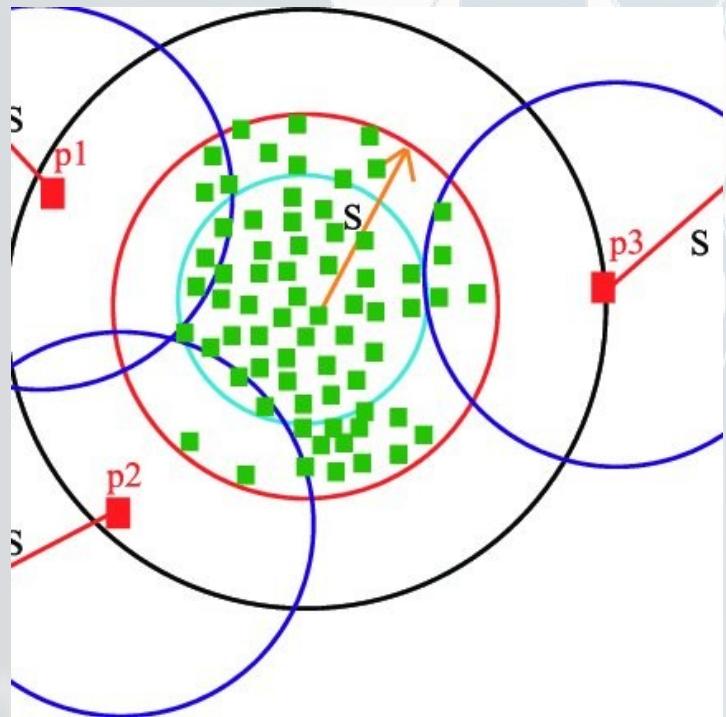
# MEDICAL DATA SCIENCE

An intermediate overview of:  
data science applications in healthcare

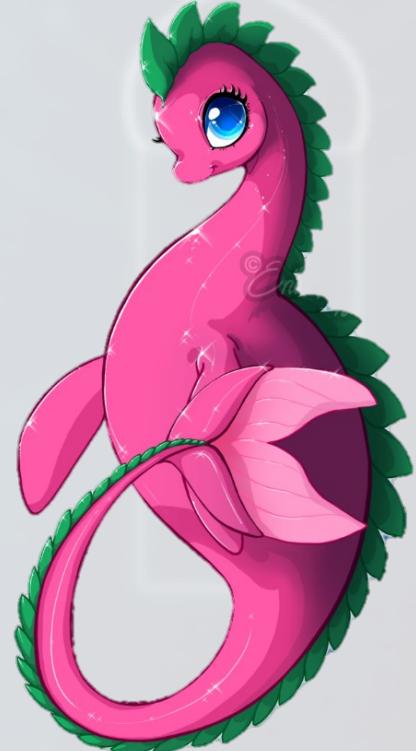
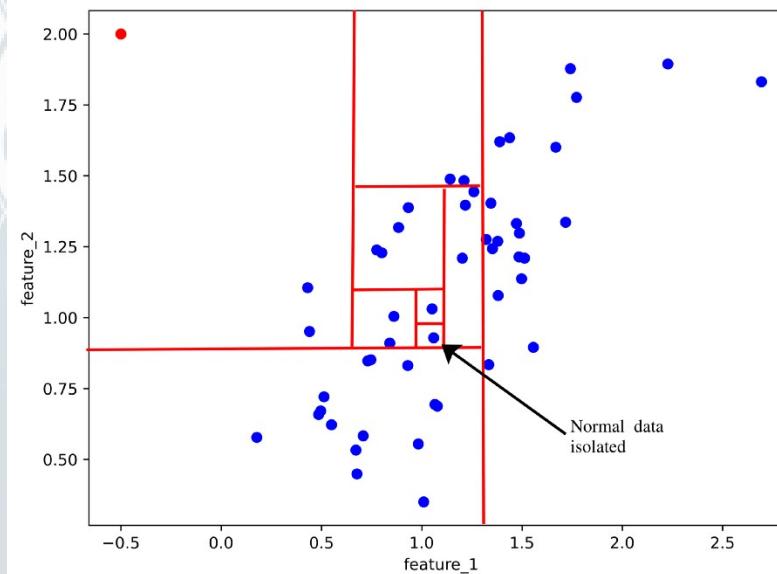
By: Alireza Vafaei Sadr  
25- Anomaly Detection



# local outlier factor



# Isolation Forest



# Let's Python!



thanks  
for  
watching!



**KEEP  
CALM  
AND  
ENJOY  
CODING**



FOR  
A  
CA  
C

