Machine Learning

Did anyone trampoline over break?

Why not?

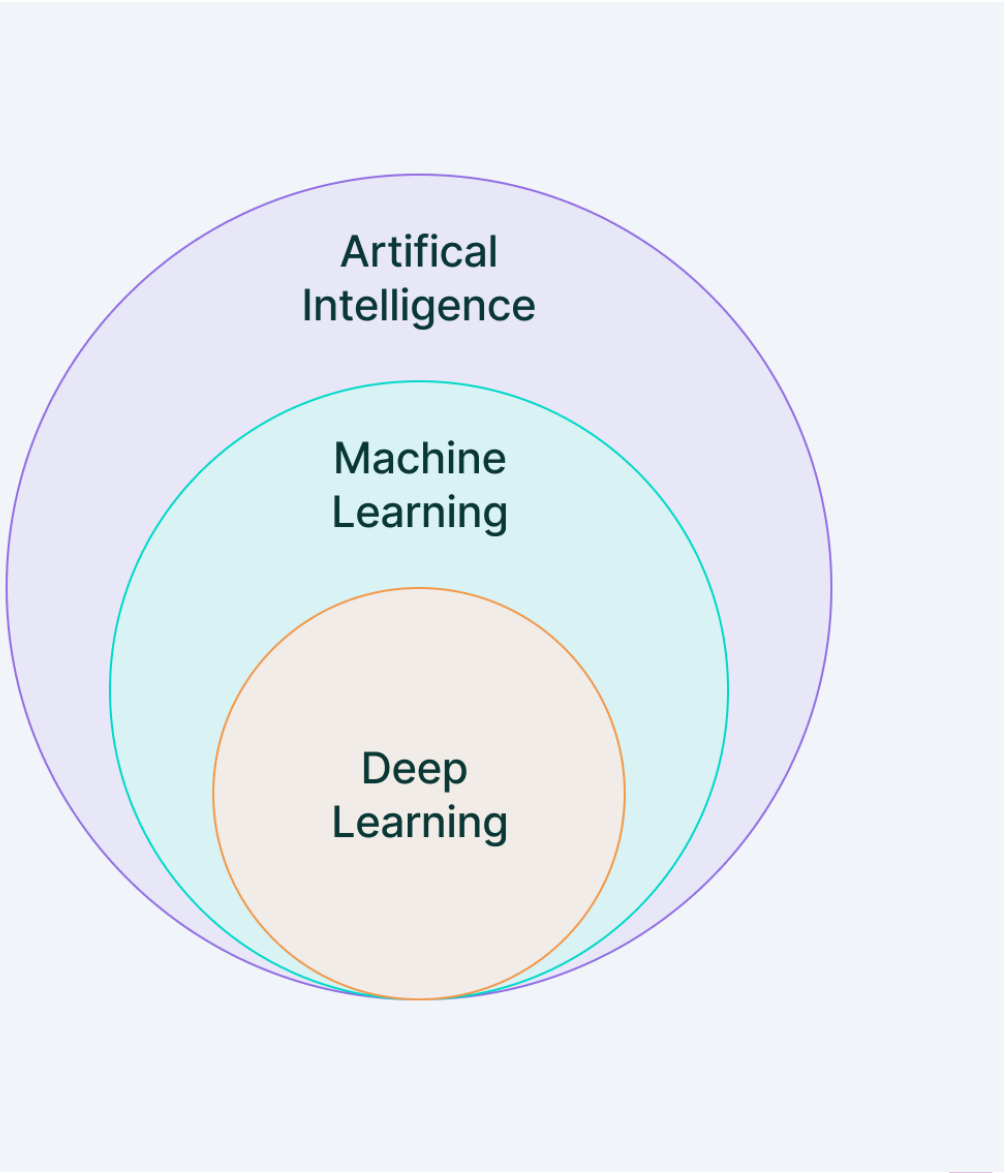
Because Spring Break

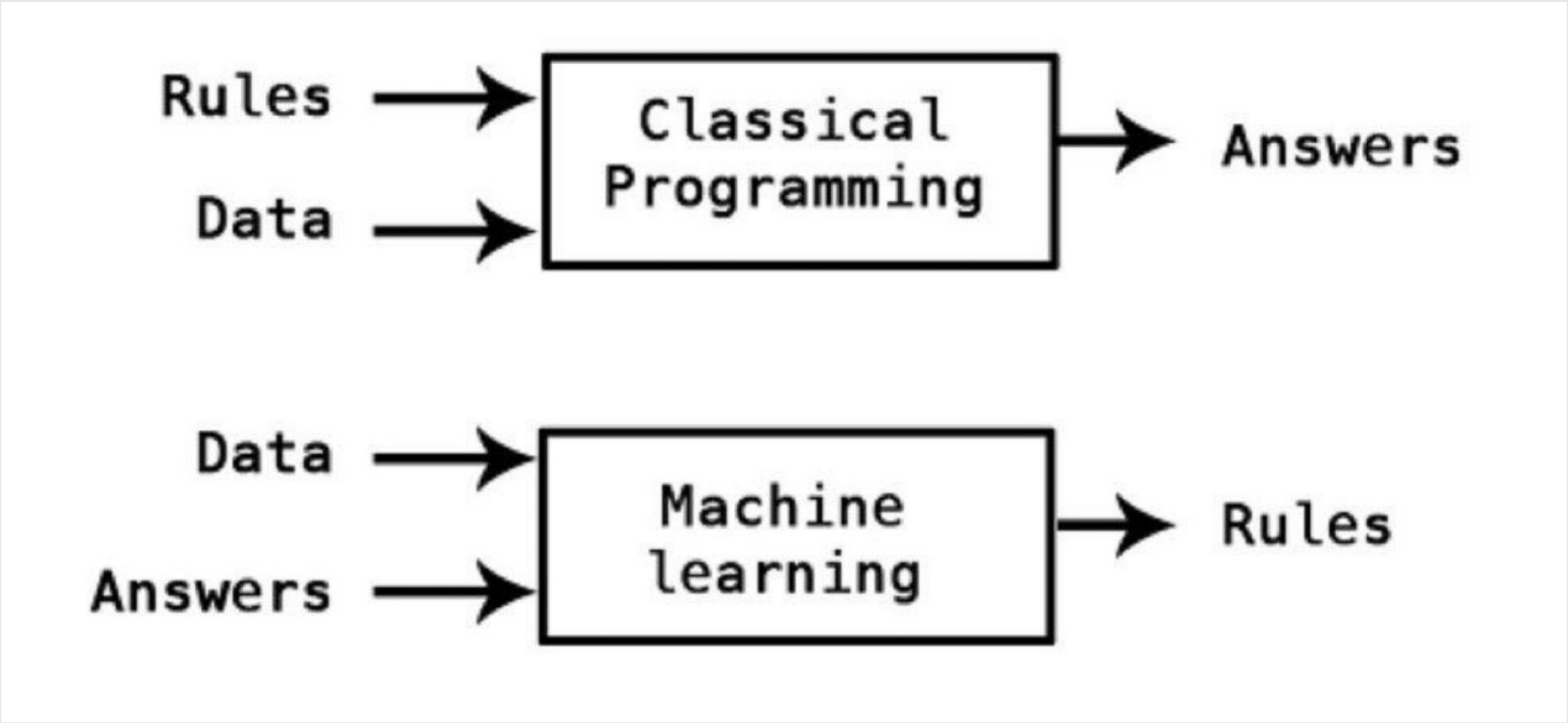
Now time for ML

What is this
really cool font?

What is Machine Learning?

Basically: structured prediction using data


Artifical
Intelligence
Machine
Learning
Deep
Learning


Rules
Data
Classical Programming
Answers
Data
Answers
Machine learning
Rules

Applications

- Agriculture
- Anatomy
- Adaptive website
- Affective computing
- Astronomy
- Banking
- Bioinformatics
- Brain–machine interfaces
- Cheminformatics
- Citizen science
- Computer networks
- Computer vision
- Credit-card fraud detection
- Data quality
- DNA sequence classification
- Economics
- Financial market analysis[75]
- General game playing
- Handwriting recognition
- Information retrieval
- Insurance
- Internet fraud detection
- Knowledge graph embedding
- Linguistics
- Machine learning control
- Machine perception
- Machine translation
- Marketing
- Medical diagnosis
- Natural language processing
- Natural language understanding
- Online advertising
- Optimization
- Recommender systems
- Robot locomotion
- Search engines
- Sentiment analysis
- Sequence mining
- Software engineering
- Speech recognition
- Structural health monitoring
- Syntactic pattern recognition
- Telecommunication
- Theorem proving
- Time-series forecasting
- User behavior analytics
- Behaviorism

What are some types of Machine Learning?

Let's start with an example…

Sq. Ft	House Price
500	50,000
1000	100,000
1500	150,000
2000	200,000
2500	250,000
3000	300,000

What do we need to create an ML model?

How do we predict 243,000 sq. ft?

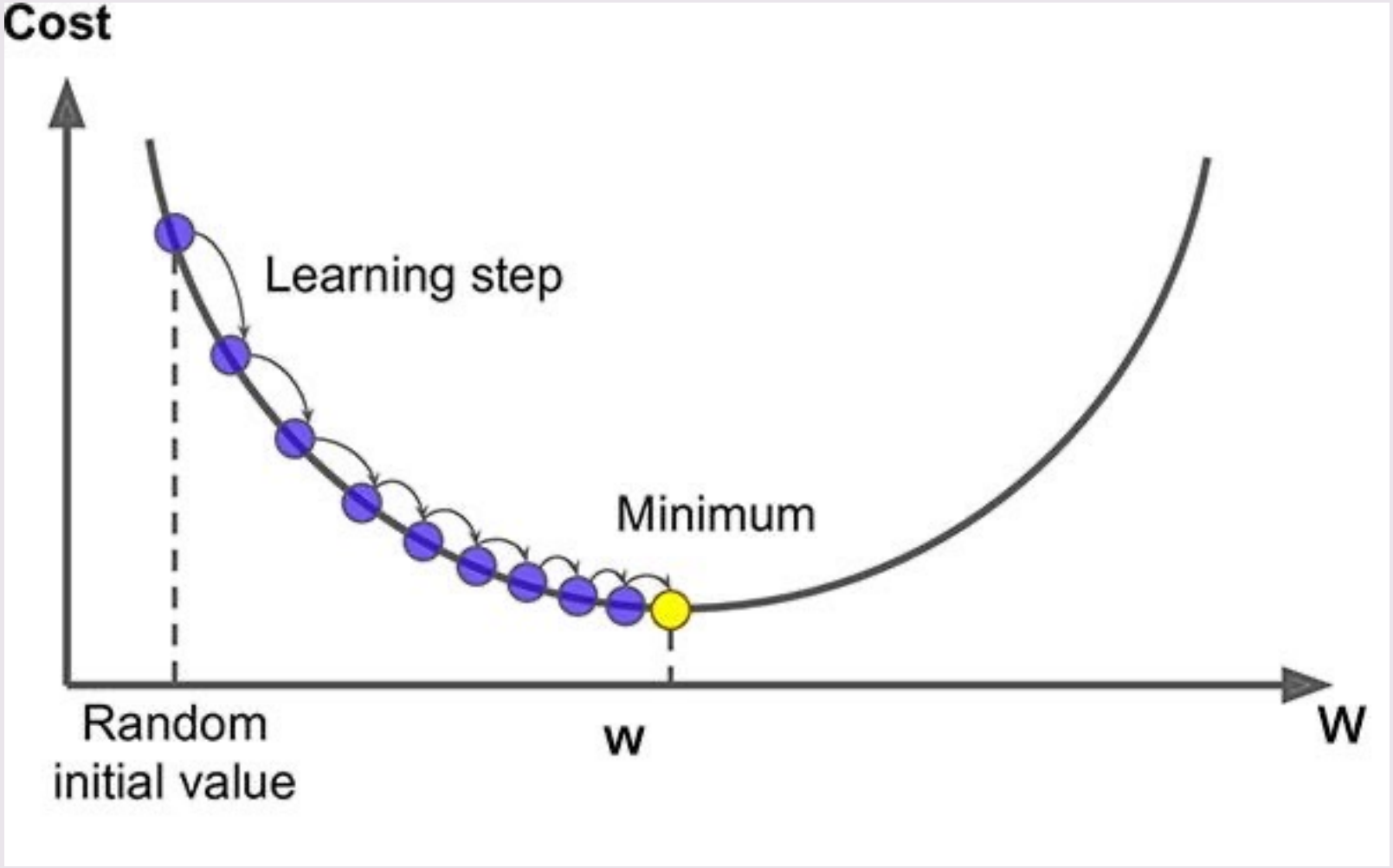
For simplicity let's pick two points

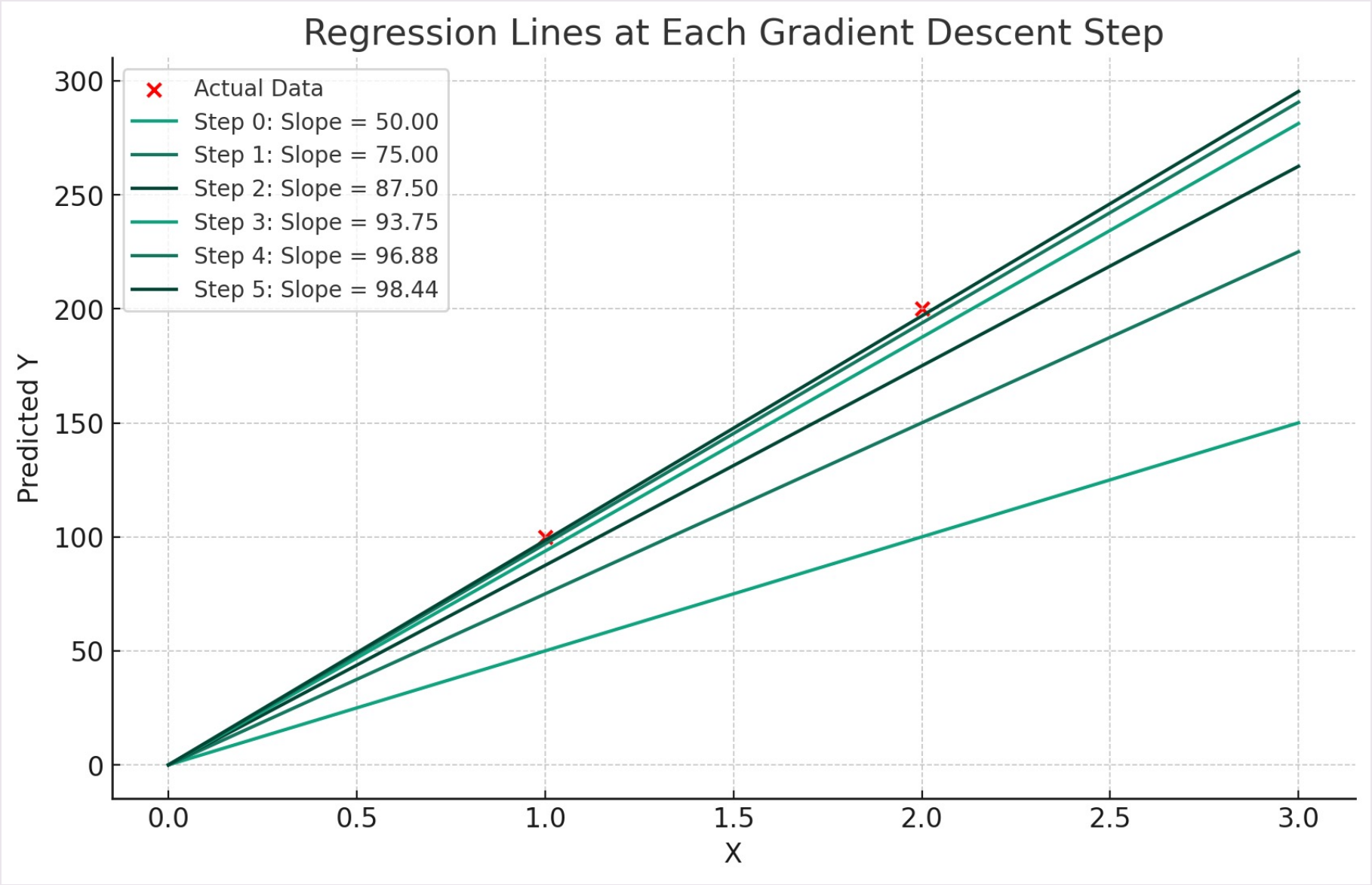
Sq. Ft	House Price
1000	100,000
2000	200,000

Let’s move to New York

Sq. Ft	Monthly Rent
1	100
2	200

Gradient Descent?




Regression Lines at Each Gradient Descent Step
Actual Data
Step 0: Slope = 50.00
Step 1: Slope = 75.00
Step 2: Slope = 87.50
Step 3: Slope = 93.75
Step 4: Slope = 96.88
Step 5: Slope = 98.44
Predicted Y
X
0
50
100
150
200
250
300
0.0
0.5
1.0
1.5
2.0
2.5
3.0

Let’s code this up!

Let's code this up!

We need to code the loss function, evaluation, optimizer, …

Why don't we `import` it

Why don't we `import` it

Enter: **Sci-Kit Learn**

Machine Learning

what society thinks I do

what my friends think I do

what my parents think I do

$$L_p = \frac{1}{2}\|\mathbf{w}\|^2 - \sum_{i=1}^{l} \alpha_i y_i (\mathbf{x}_i \cdot \mathbf{w} + b) + \sum_{i=1}^{l} \alpha_i$$

$$\alpha_i \geq 0, \forall i$$

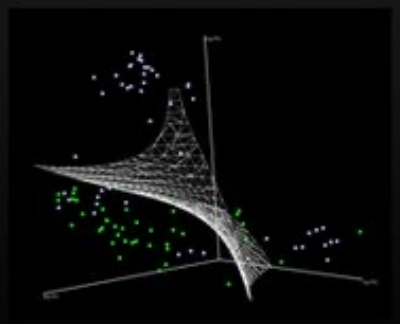
$$\mathbf{w} = \sum_{i=1}^{l} \alpha_i y_i \mathbf{x}_i, \sum_{i=1}^{l} \alpha_i y_i = 0$$

$$\nabla \hat{g}(\theta_t) = \frac{1}{n} \sum_{i=1}^{n} \nabla \ell(x_i, y_i; \theta_t) + \nabla r(\theta_t).$$

$$\theta_{t+1} = \theta_t - \eta_t \nabla \ell(x_{v(t)}, y_{v(t)}; \theta_t) - \eta_t \cdot \nabla r(\theta_t)$$

$$\mathbb{E}_{v(t)}[\ell(x_{v(t)}, y_{v(t)}; \theta_t)] = \frac{1}{n} \sum_{i} \ell(x_i, y_i; \theta_t).$$

what other programmers think I do



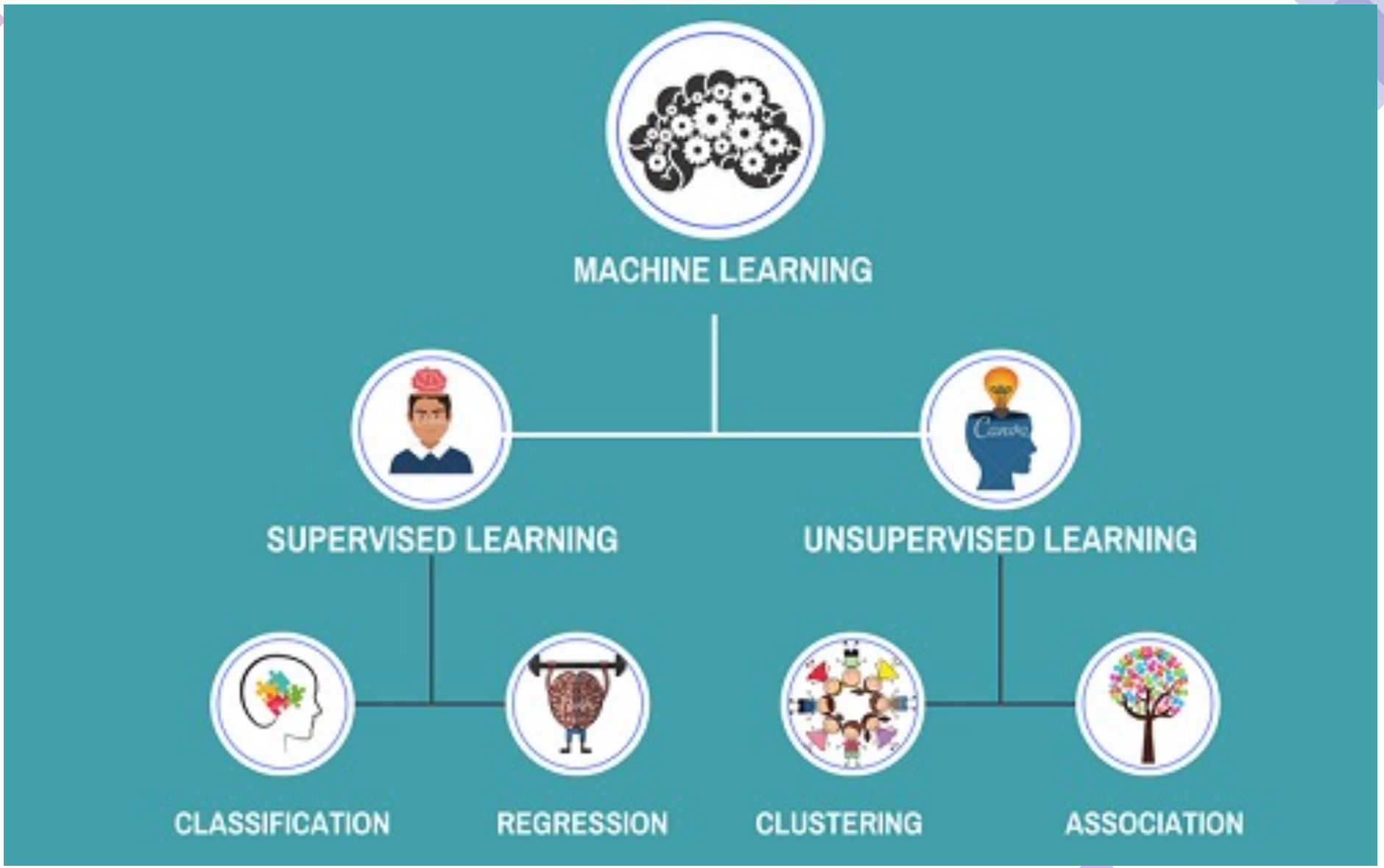
what I think I do

```
>>> from scipy import SVM
```

what I really do

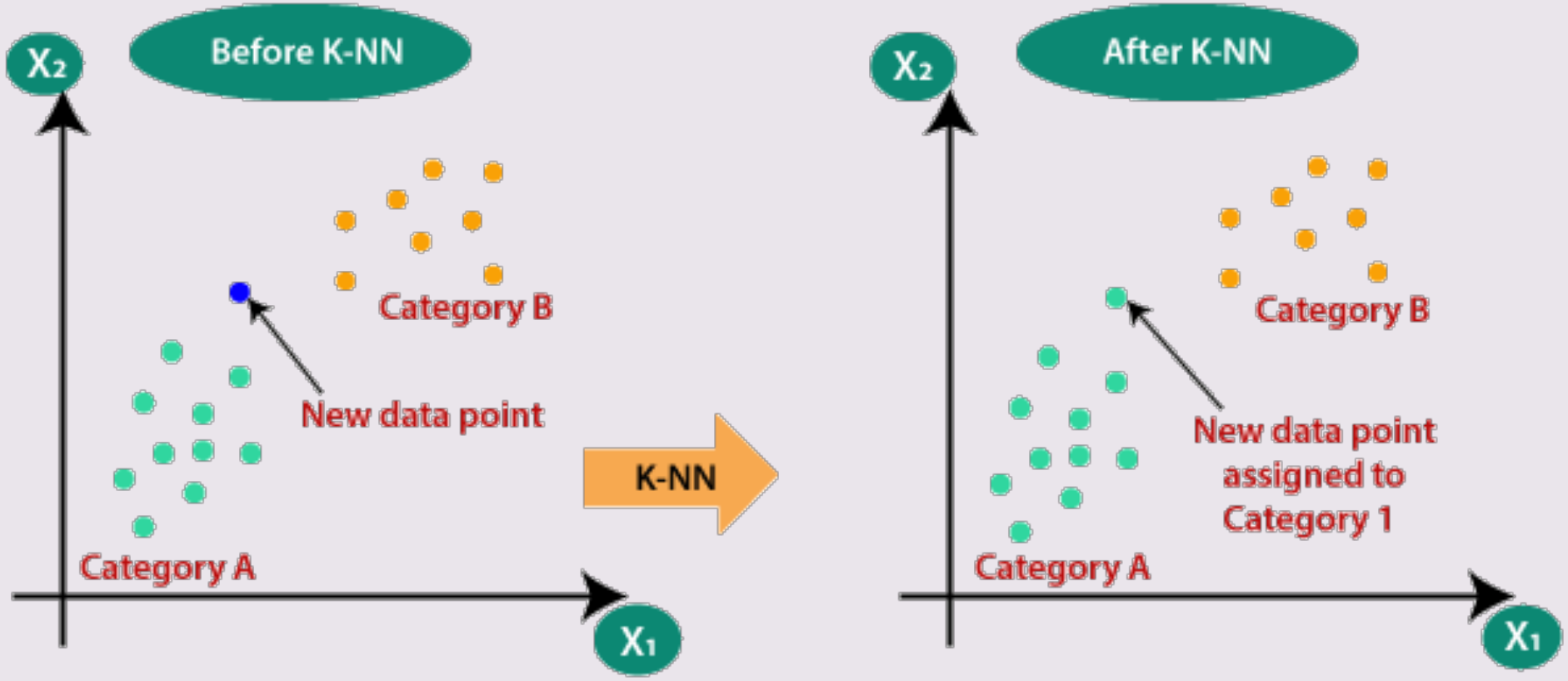
Great! That's Regression.

What about *classification*

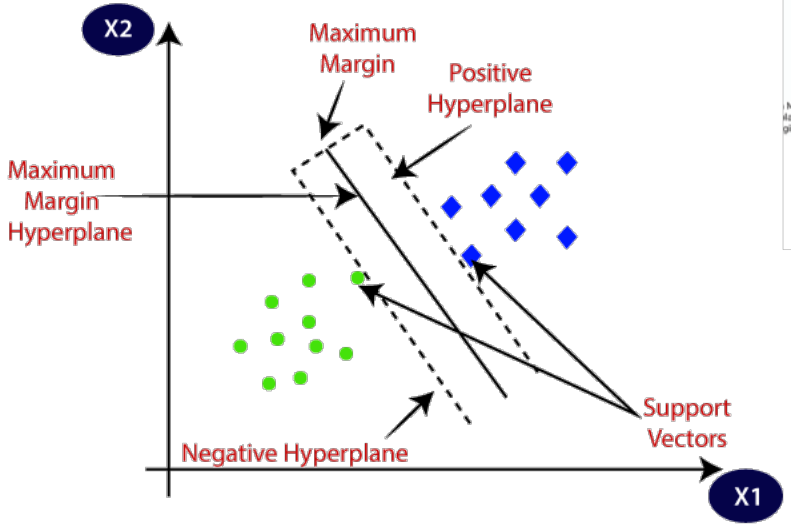

MACHINE LEARNING
SUPERVISED LEARNING
UNSUPERVISED LEARNING
CLASSIFICATION
REGRESSION
CLUSTERING
ASSOCIATION

Basic Classification Example?

K Nearest Neighbors

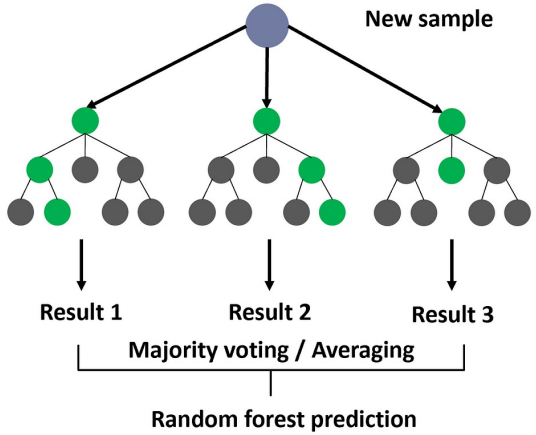


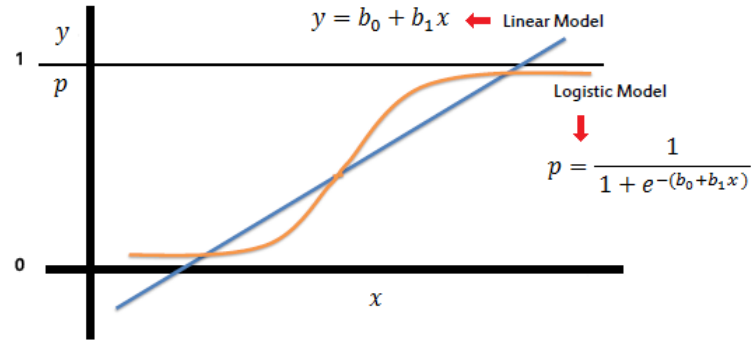
Other Options



Support Vector Machines

Random Forest



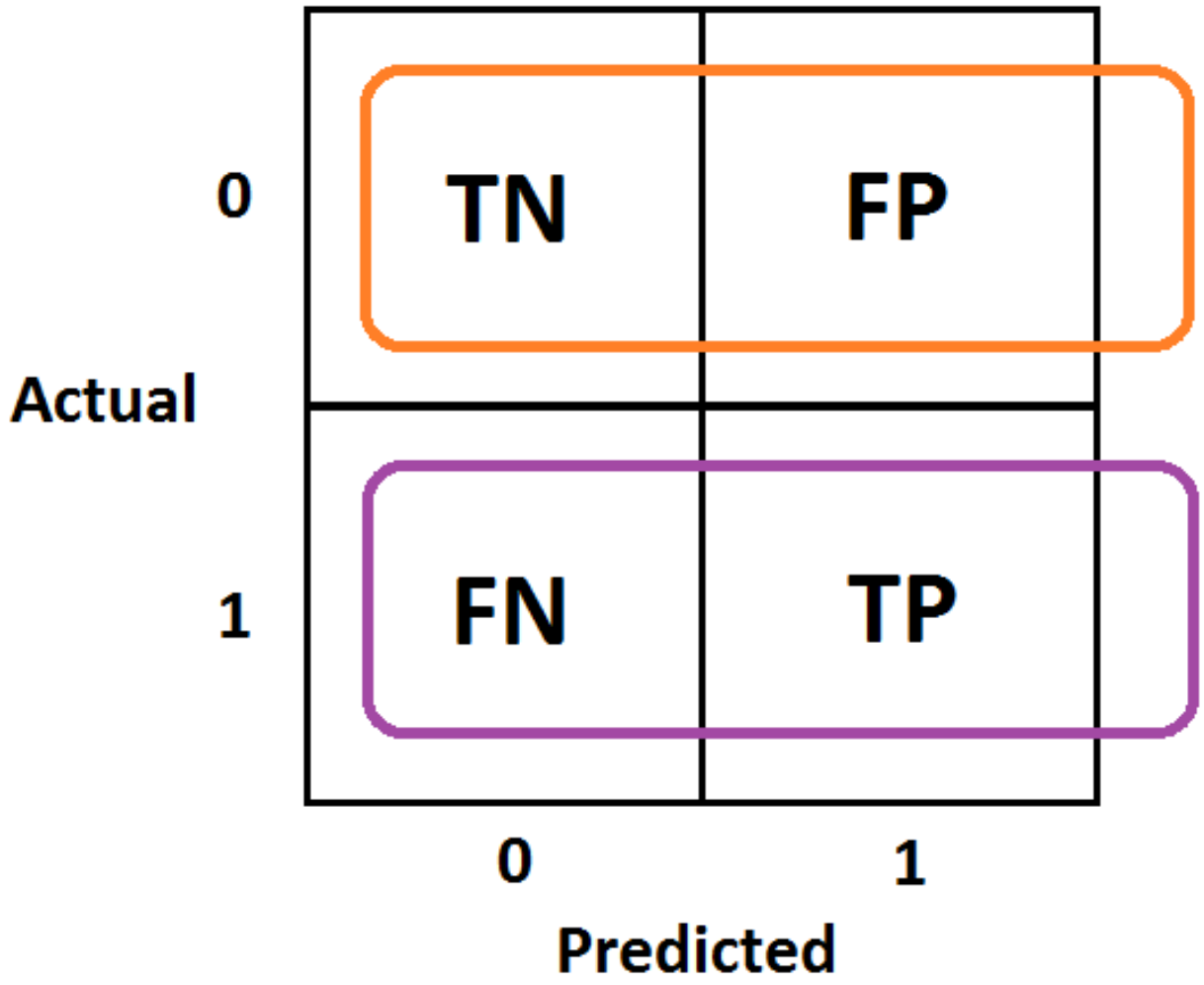


Logistic Regression

¡Vamos a implementar!

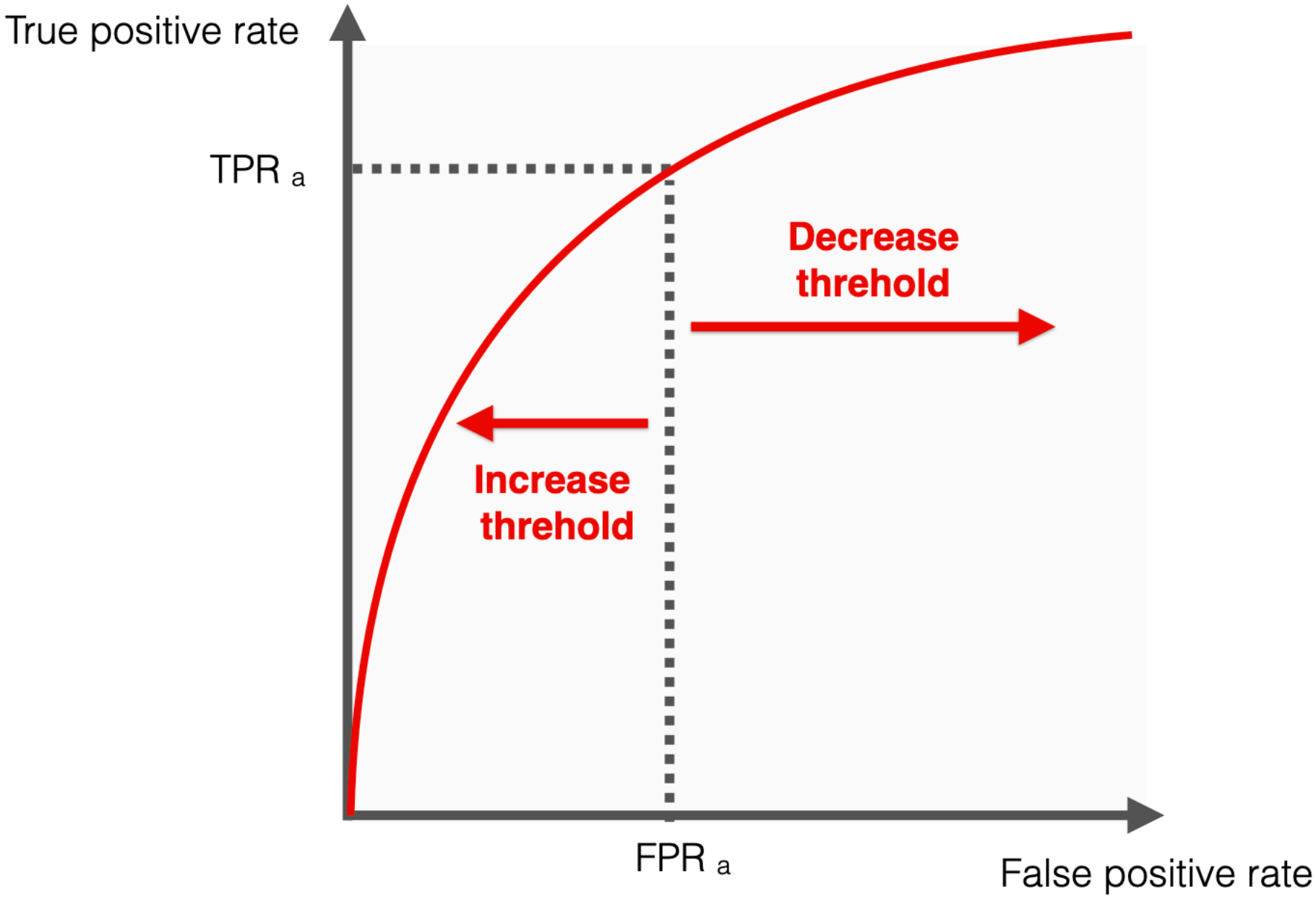
Let’s implement it!

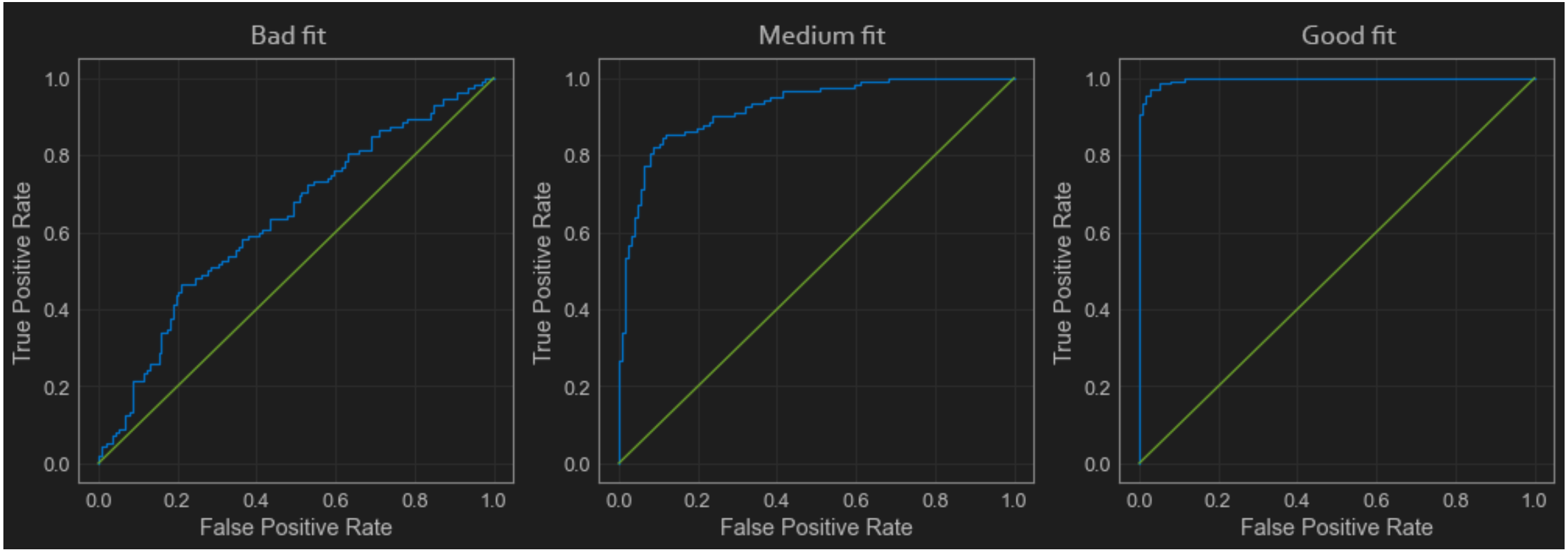
How do we measure classification models?

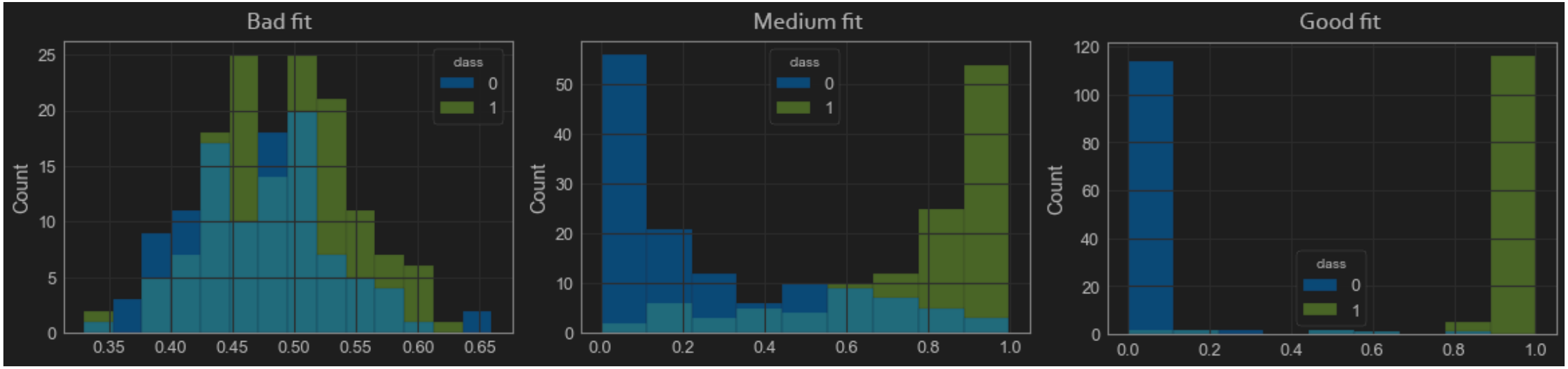


False Positive Rate (FPR) = $\frac{FP}{(FP + TN)}$

True Positive Rate(TPR) = $\frac{TP}{(TP + FN)}$


True positive rate
TPR a
Decrease threhold
Increase threhold
FPR a
False positive rate


Bad fit
Medium fit
Good fit
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0


Bad fit
Count
class
0
1
0.35
0.40
0.45
0.50
0.55
0.60
0.65
Medium fit
Count
class
0
1
0.0
0.2
0.4
0.6
0.8
1.0
Good fit
Count
class
0
1
0.0
0.2
0.4
0.6
0.8
1.0

To Jupyter Notebook!

Can we use KNN for Regression?

But there's so many neighbors, how do we decide how to maximize our *hyperparameters*?

What do birds call other birds
that live nearby?

Nest-door neighbors!

What if the data has different scales?

How do we validate our model?
What if we *get lucky* with our split?

How do we put this all together?

Let’s talk Unsupervised Learning

What is PCA

What is K Means?