

# Machine Learning

**Let's have some fun**

# Overview

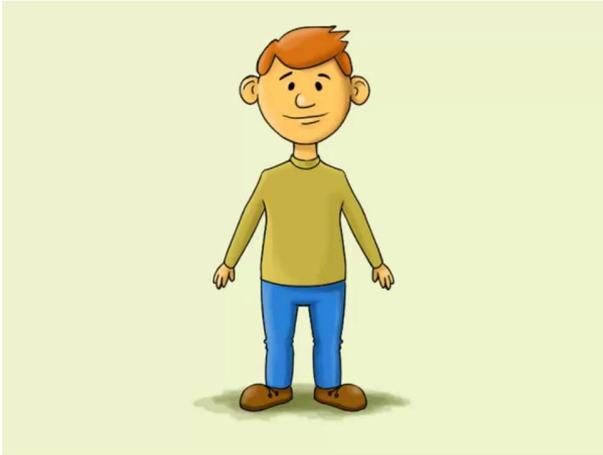
- Price of a house
- Spam !!!
- Recommendation
- Apply to a university
- Ninjas
- Apples and Oranges
- Pizza Parlour

# Overview

- **Price of a house**
  - **Spam !!!**
  - **Recommendation**
  - **Apply to a university**
- 
- Ninjas
  - Apples and Oranges
  - Pizza Parlour

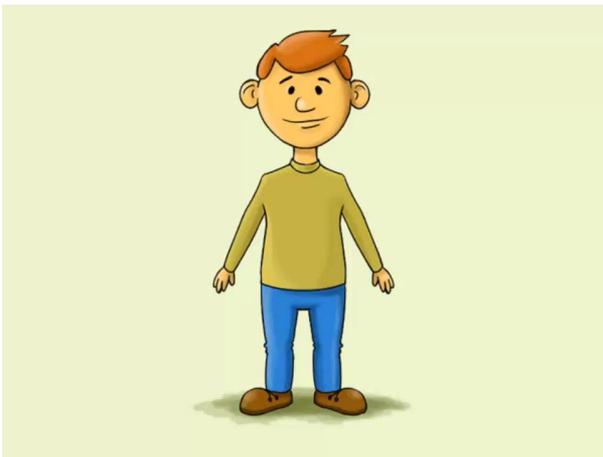
# What is Machine Learning?

Learn from experience



# What is Machine Learning?

Learn from experience

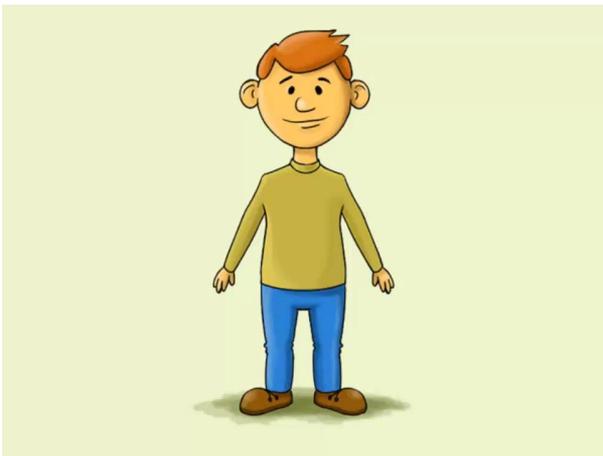


Follow Instructions

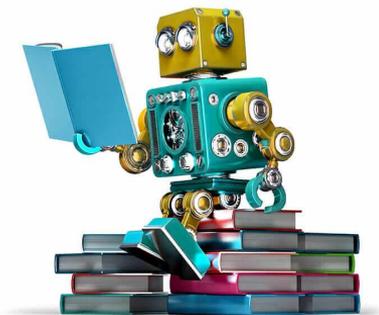


# What is Machine Learning?

Learn from experience



Learn from Data



Follow Instructions



# Learning

- ❑ Memorization (Declarative)
  - ❑ Accumulation of facts
  - ❑ Limitations
    - ❑ Time to memorize
    - ❑ Space to store the information
- ❑ Generalization (Imperative)
  - ❑ Deduce information from the old facts
  - ❑ Limitations
    - ❑ Accuracy of the deduction
    - ❑ Assumes Past  $\rightarrow$  Future

# Price of a house



Price : \$70,000

# Price of a house



Price : \$70,000



Price: \$160,000

# Price of a house



Price : \$70,000

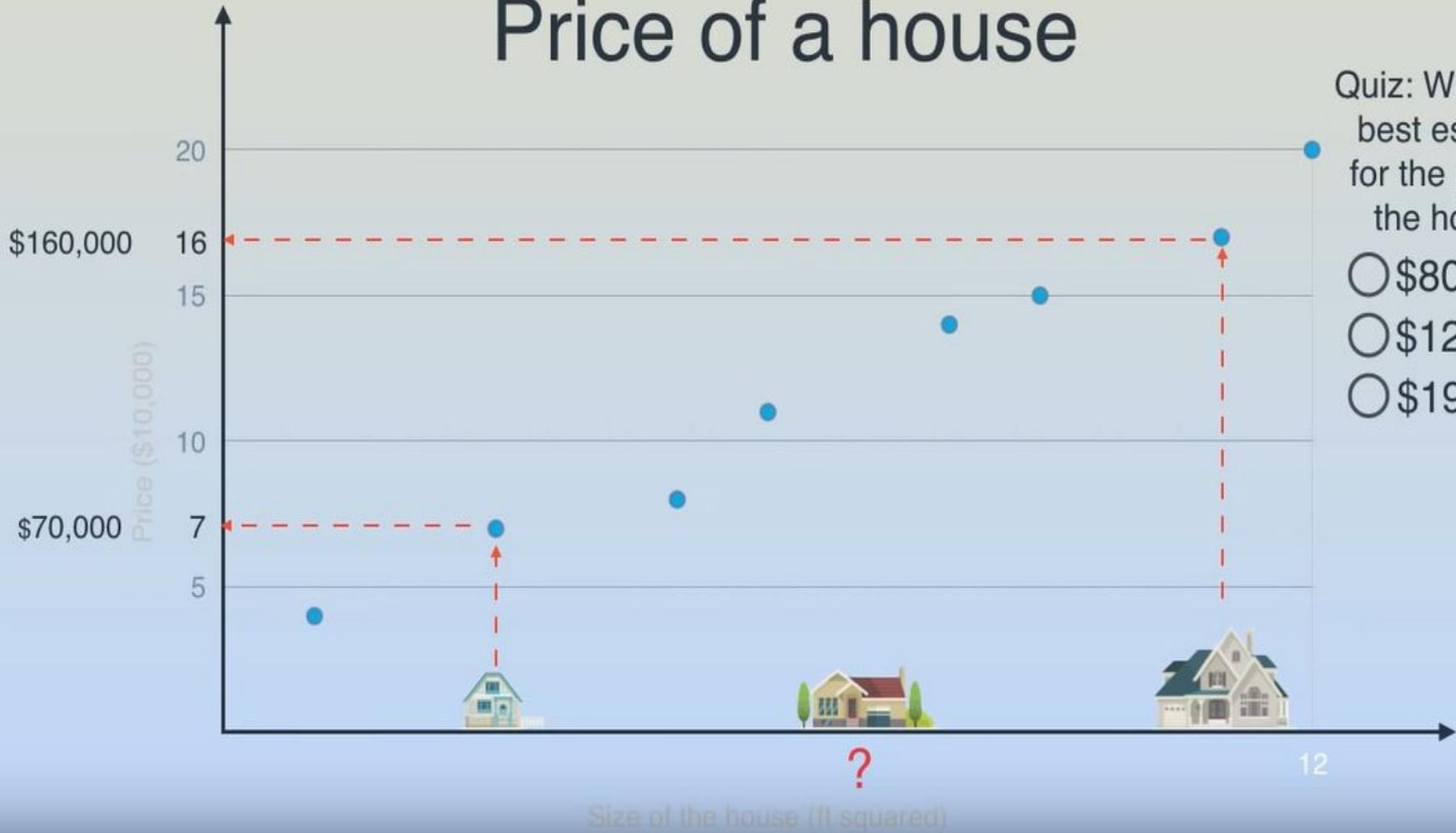


???



Price: \$160,000

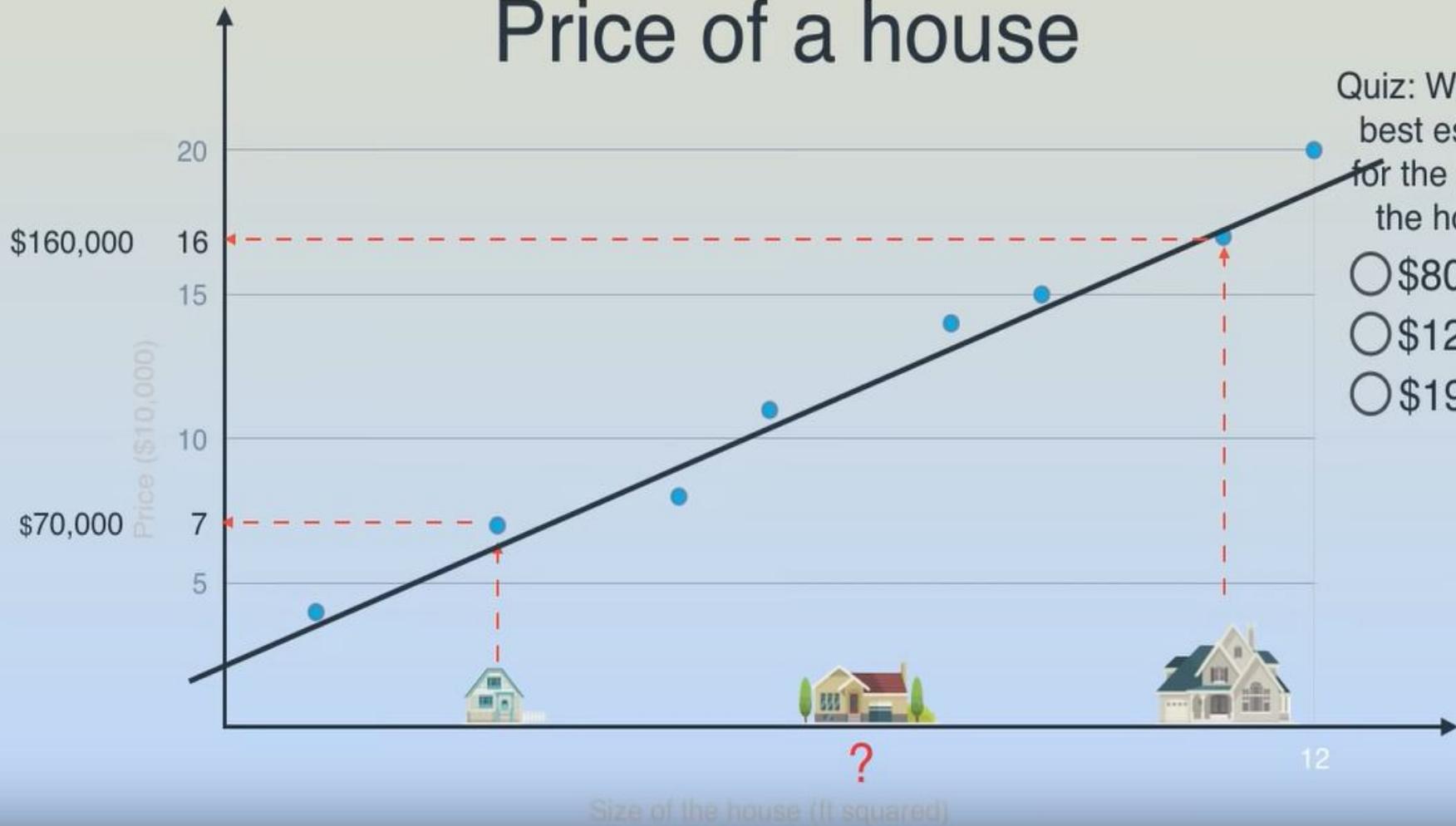
# Price of a house



Quiz: What's the best estimate for the price of the house?

- \$80,000
- \$120,000
- \$190,000

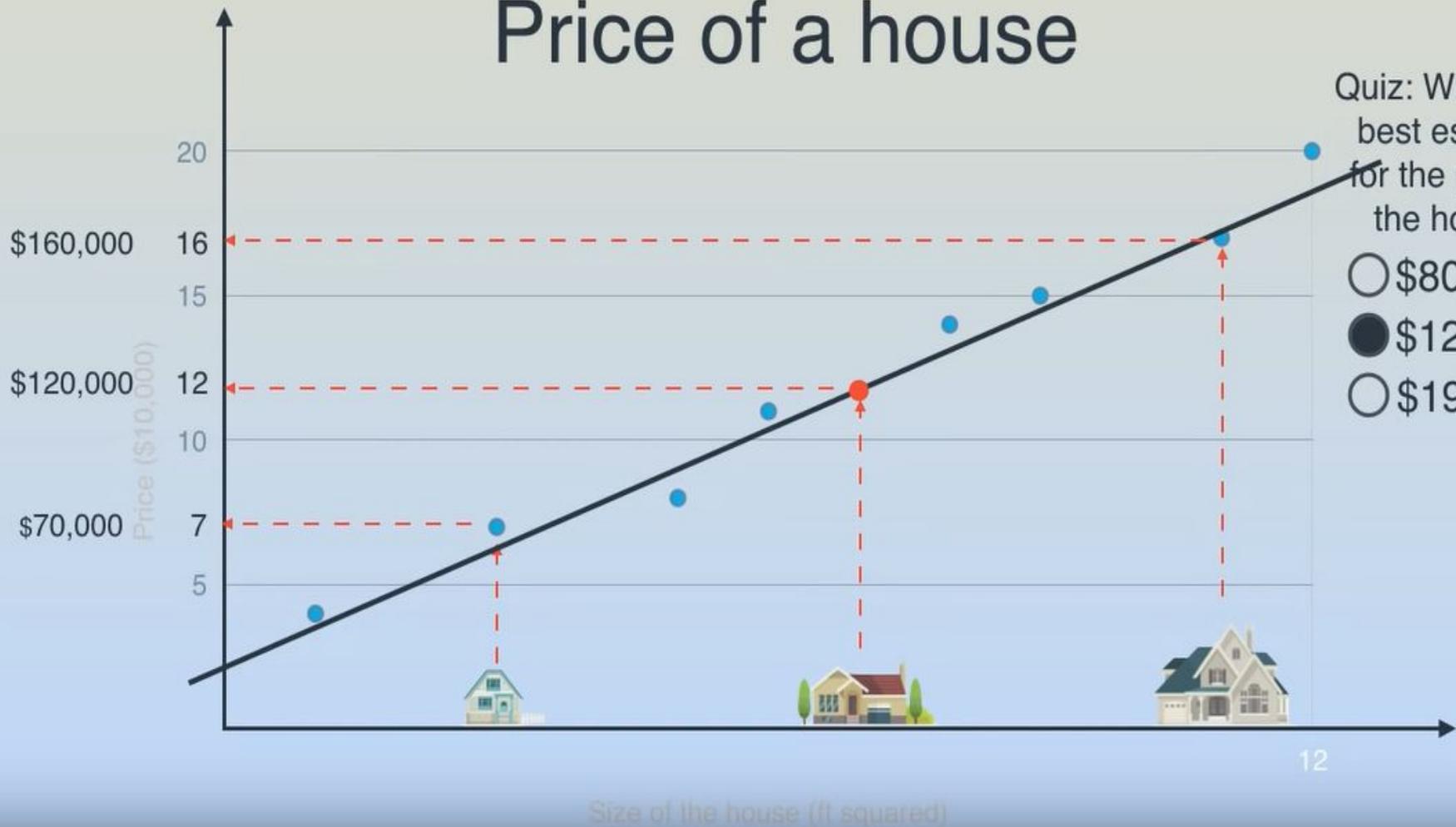
# Price of a house



Quiz: What's the best estimate for the price of the house?

- \$80,000
- \$120,000
- \$190,000

# Price of a house



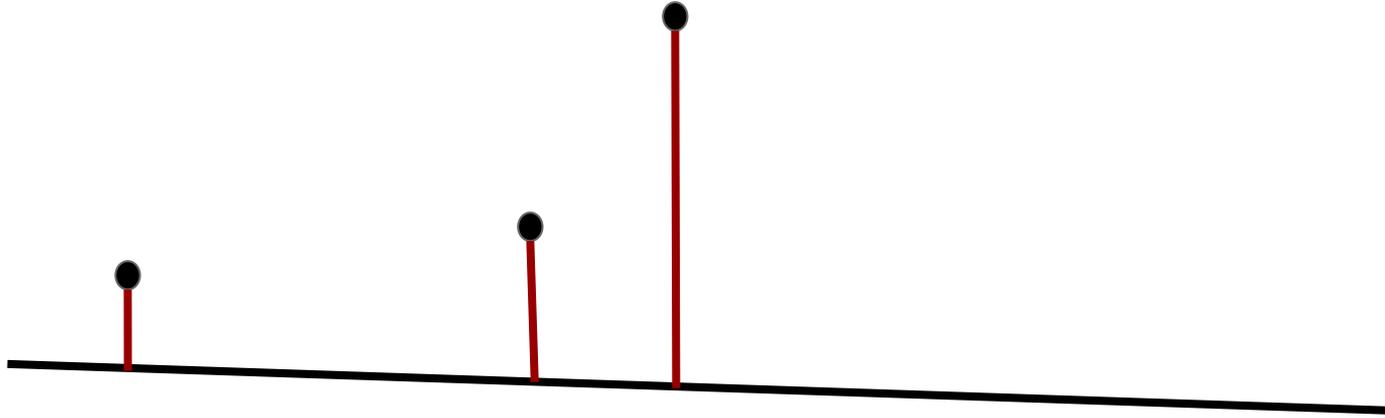
# Linear Regression

- How can we learn to predict the **prices of houses of other sizes** in the city, as a **function of their living area**?
- Example of supervised learning problem.
- When the target variable we are trying to predict is continuous, **regression** problem.

# Linear Regression



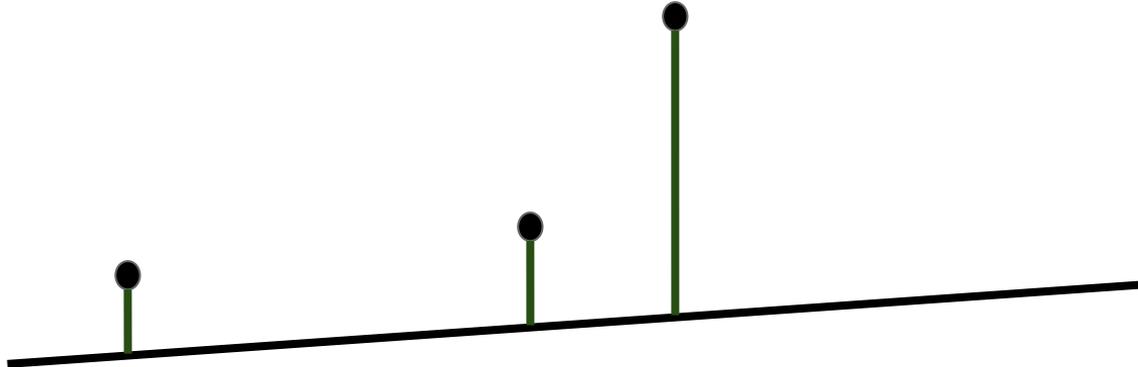
# Linear Regression



Error :



# Linear Regression

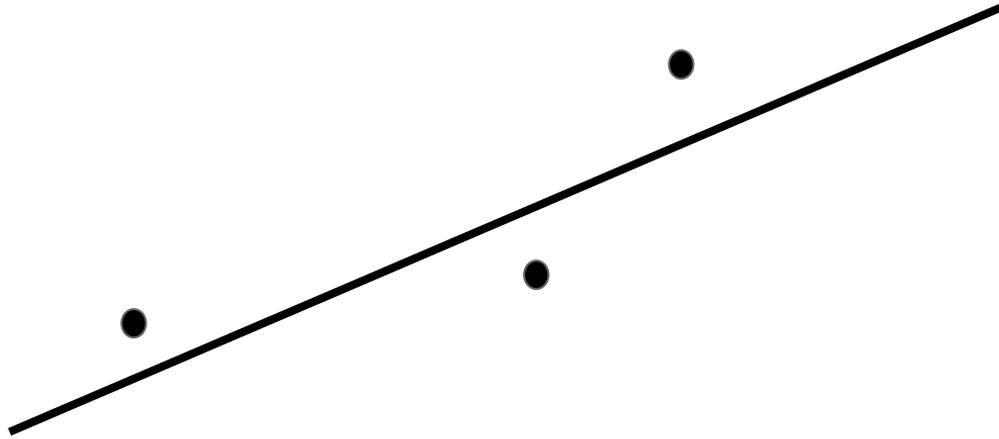


Error :



# Linear Regression

Finally:



- Gradient Descent
- Method of least square : Add the sum of squares of error at individual point.

# Gradient descent



Mount Erorest



# Gradient descent



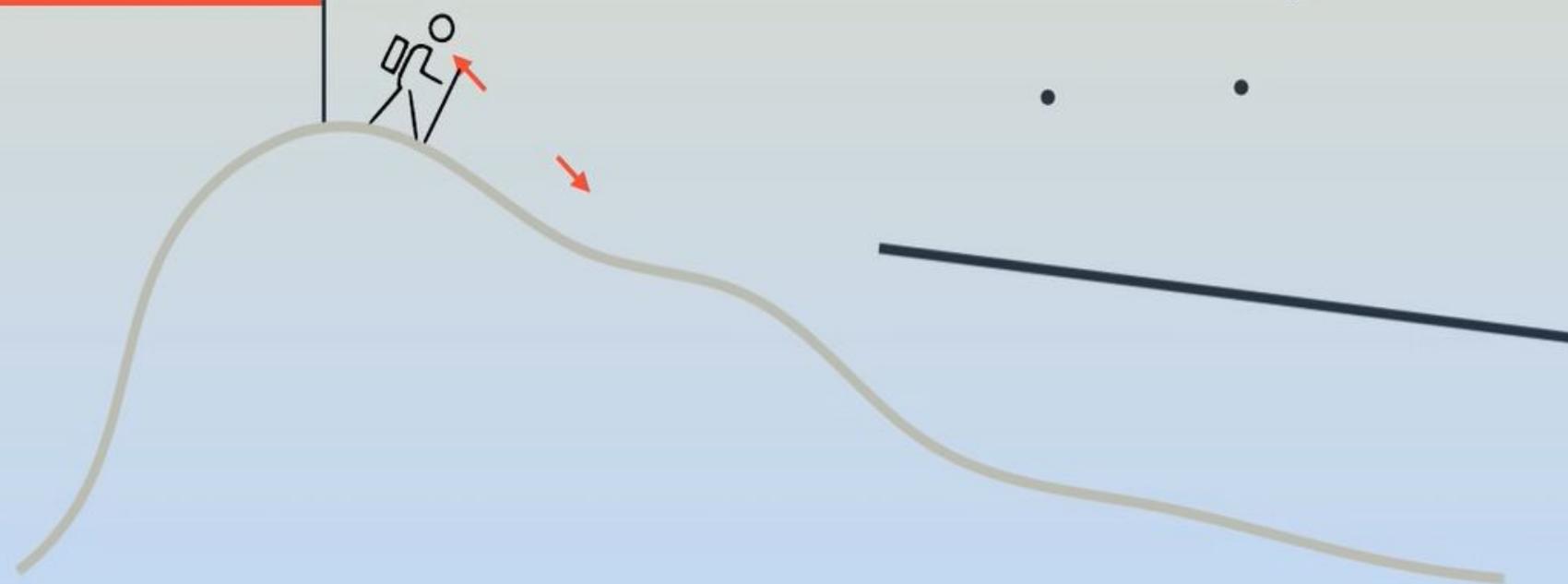
Mount Erorest



# Gradient descent



Mount Erorest



# Gradient descent



Mount Erorest



# Gradient descent

Mount Erorest



# Linear and Polynomial Regression



# Detecting Spam e-mails

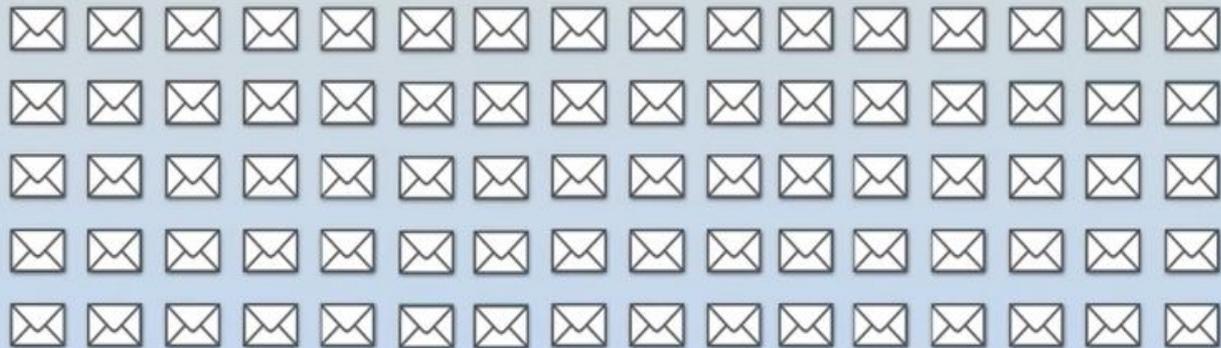


# Detecting Spam e-mails

Spam



Non-spam

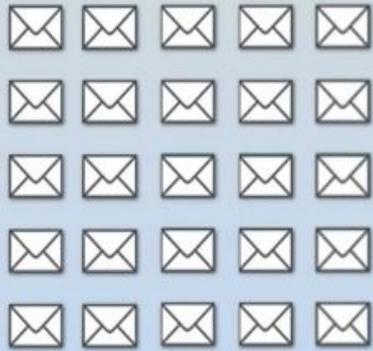


# Detecting Spam e-mails

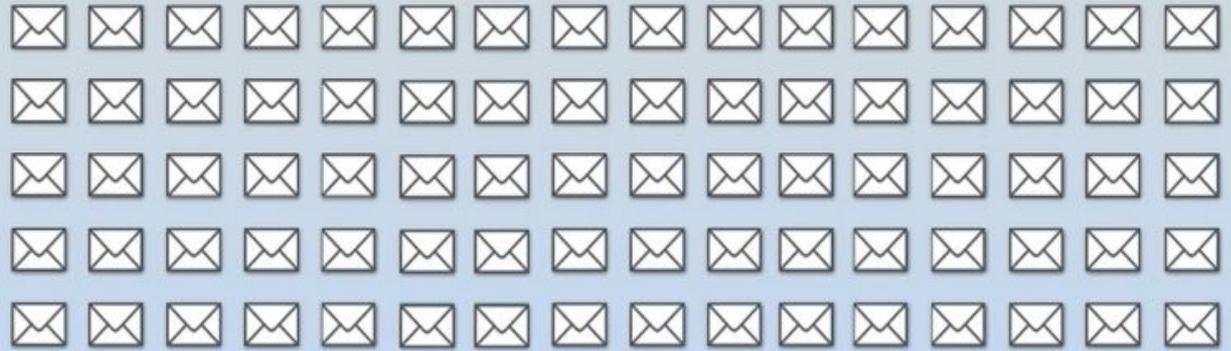


“Cheap”

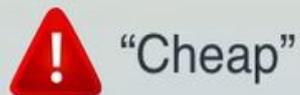
Spam



Non-spam



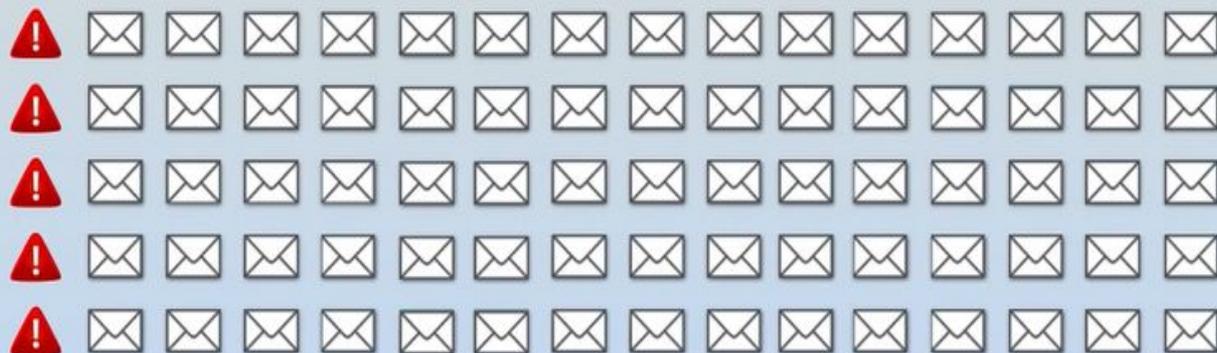
# Detecting Spam e-mails



Spam



Non-spam



# Detecting Spam e-mails



“Cheap”

Spam

Non-spam



# Detecting Spam e-mails



“Cheap”

Spam

Non-spam

**Quiz:** If an e-mail contains the word “cheap”, what is the probability of it being spam?



# Detecting Spam e-mails



“Cheap”

Spam

Non-spam



20



5

**Quiz:** If an e-mail contains the word “cheap”, what is the probability of it being spam?

- 40%
- 60%
- 80%

# Detecting Spam e-mails



“Cheap”

Spam

Non-spam



80%



20%

**Quiz:** If an e-mail contains the word “cheap”, what is the probability of it being spam?

- 40%
- 60%
- 80%

**Conclusion:**

If the e-mail contains the word “cheap”,  
The probability of it being spam is 80%

# Detecting Spam e-mails

 "Cheap" → 80%

 Spelling mistake → 70%

 Missing title → 95%

 etc...

**Quiz:** If an e-mail contains the word "cheap", what is the probability of it being spam?

40%

60%

80%

**Conclusion:**

If the e-mail contains the word "cheap",  
The probability of it being spam is 80%

# Naive Bayes Algorithm

 "Cheap" → 80%

 Spelling mistake → 70%

 Missing title → 95%

 etc...

**Quiz:** If an e-mail contains the word "cheap", what is the probability of it being spam?

40%  
 60%  
 80%

## **Conclusion:**

If the e-mail contains the word "cheap",  
The probability of it being spam is 80%

# Recommending Apps

Gender	Age	App
F	15	
F	25	
M	32	
F	40	
M	12	
M	14	

# Recommending Apps



Gender	Age	App
F	15	
F	25	
M	32	
F	40	
M	12	
M	14	

# Recommending Apps

Gender	Age	App
F	15	
F	25	
M	32	
F	40	
M	12	
M	14	

Quiz: Between Gender and Age, which one seems more decisive for predicting what app will the users download?

- Gender
- Age

# Recommending Apps

Gender	Age	App
F	15	
F	25	
M	32	
F	40	
M	12	
M	14	

Quiz: Between Gender and Age, which one seems more decisive for predicting what app will the users download?

- Gender
- Age

# Recommending Apps

Gender	Age	App
F	15	
F	25	
M	32	
F	40	
M	12	
M	14	

Quiz: Between Gender and Age, which one seems more decisive for predicting what app will the users download?

- Gender
- Age

# Recommending Apps

Gender	Age	App
F	15	
F	25	
M	32	
F	40	
M	12	
M	14	



# Recommending Apps

Gender	Age	App
F	25	
M	32	
F	40	



# Recommending Apps

Gender	Age	App
F	25	
M	32	
F	40	



# Decision Tree

Gender	Age	App
F	15	
F	25	
M	32	
F	40	
M	12	
M	14	



# Acceptance at a University



Test



Grades

# Acceptance at a University



Test



Grades

Student 1

Test: 9/10



Grades: 8/10

Student 2

Test: 3/10



Grades: 4/10

Student 3

Test: 7/10



Grades: 6/10

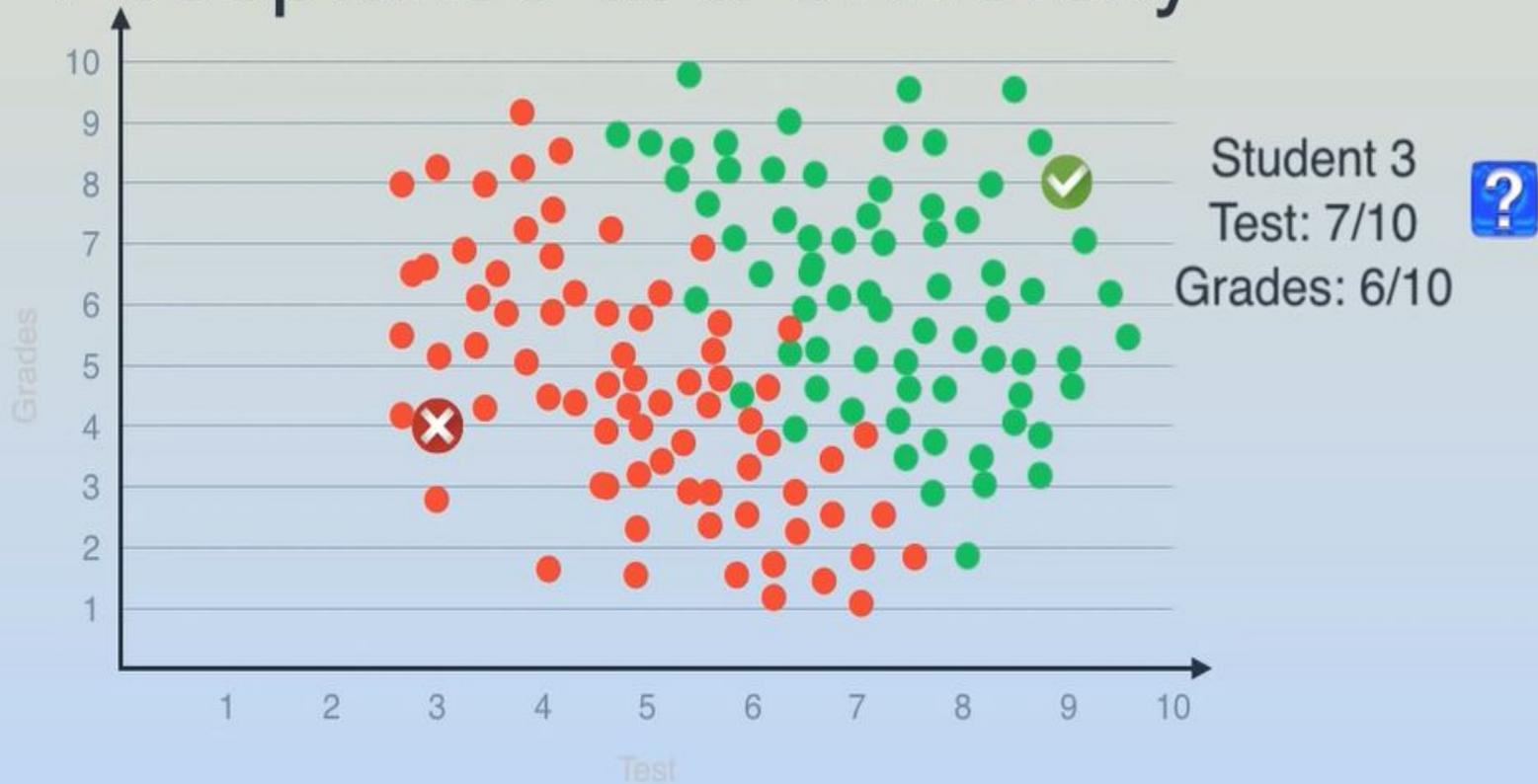
# Acceptance at a University



# Acceptance at a University



# Acceptance at a University



# Acceptance at a University



# Acceptance at a University



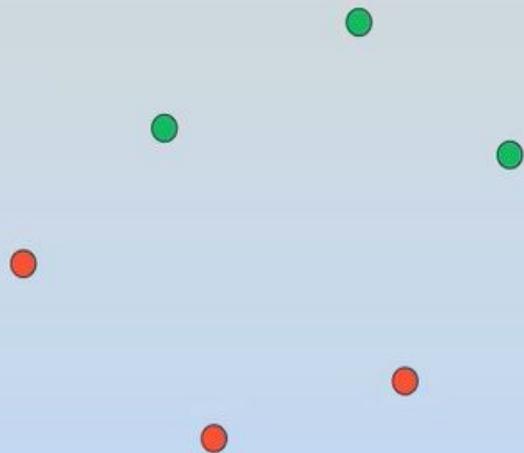
# Acceptance at a University



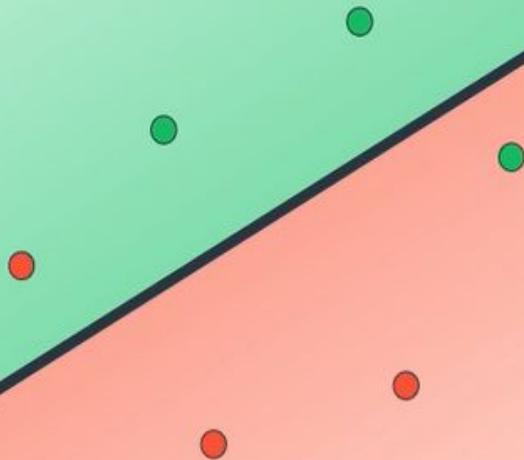
# Logistic Regression



# Logistic Regression



# Logistic Regression



# Logistic Regression

Errors: 2



# Logistic Regression

Errors: 2



Gradient descent

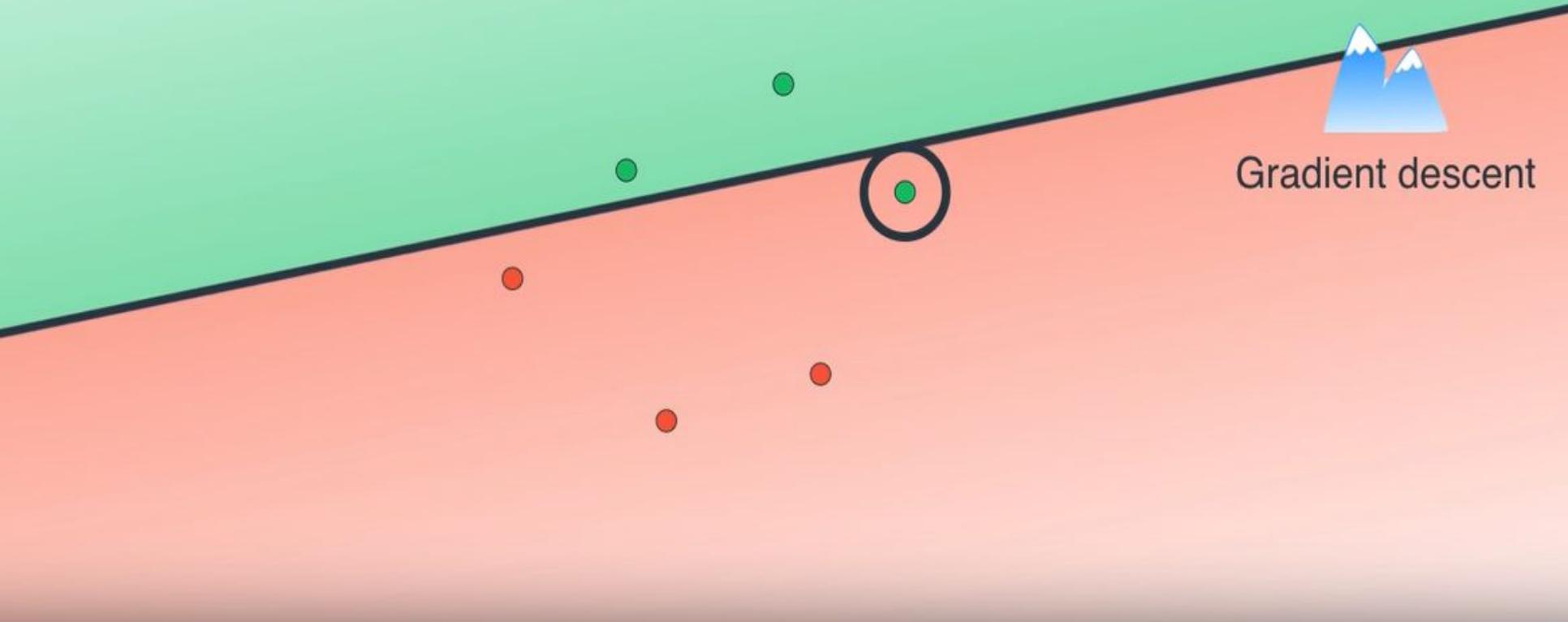


# Logistic Regression

Errors: 1



Gradient descent

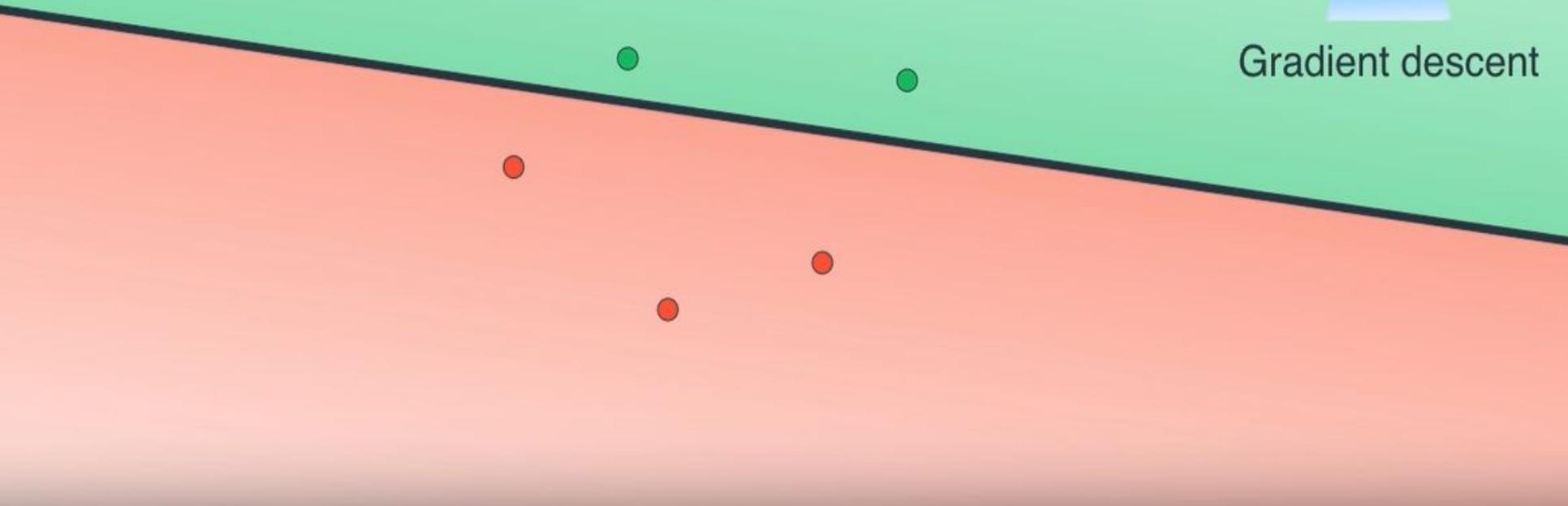


# Logistic Regression

Errors: 0



Gradient descent

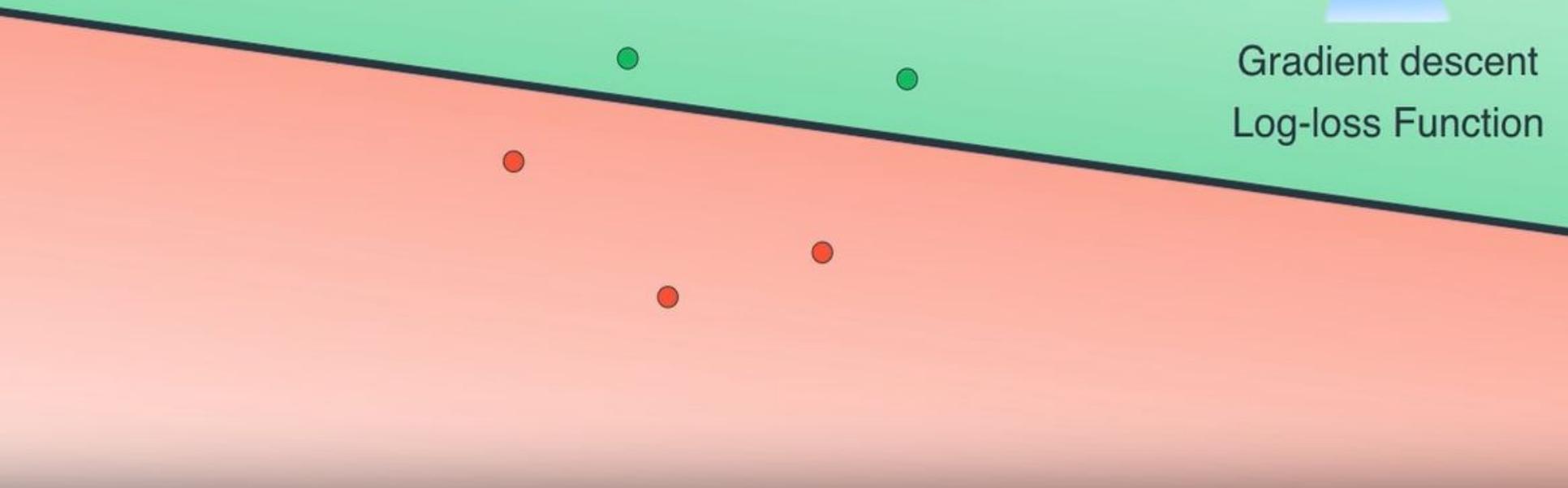


# Logistic Regression

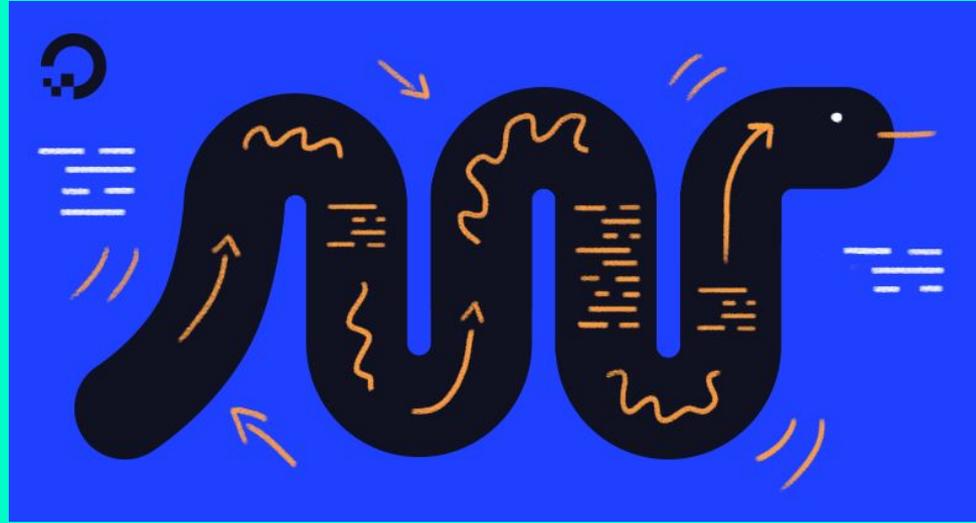
~~Errors:~~ 0



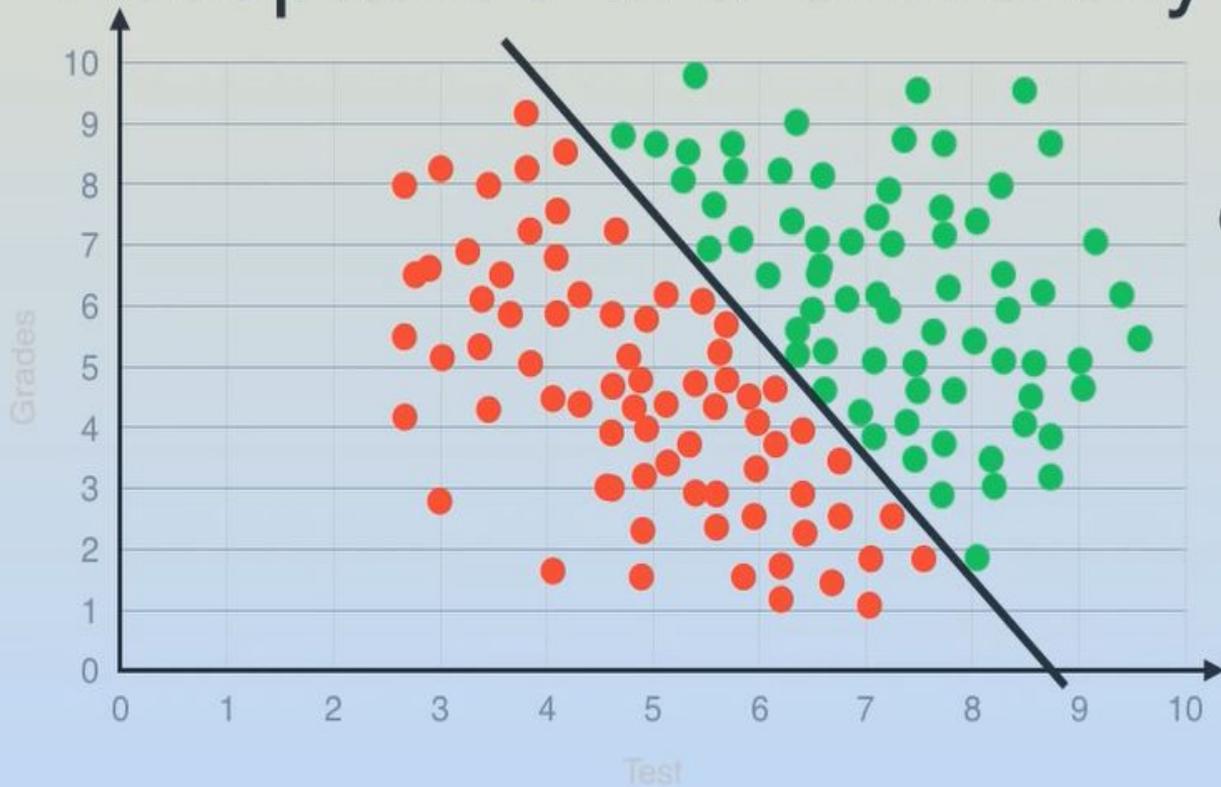
Gradient descent  
Log-loss Function



**ML without  
coding ???**



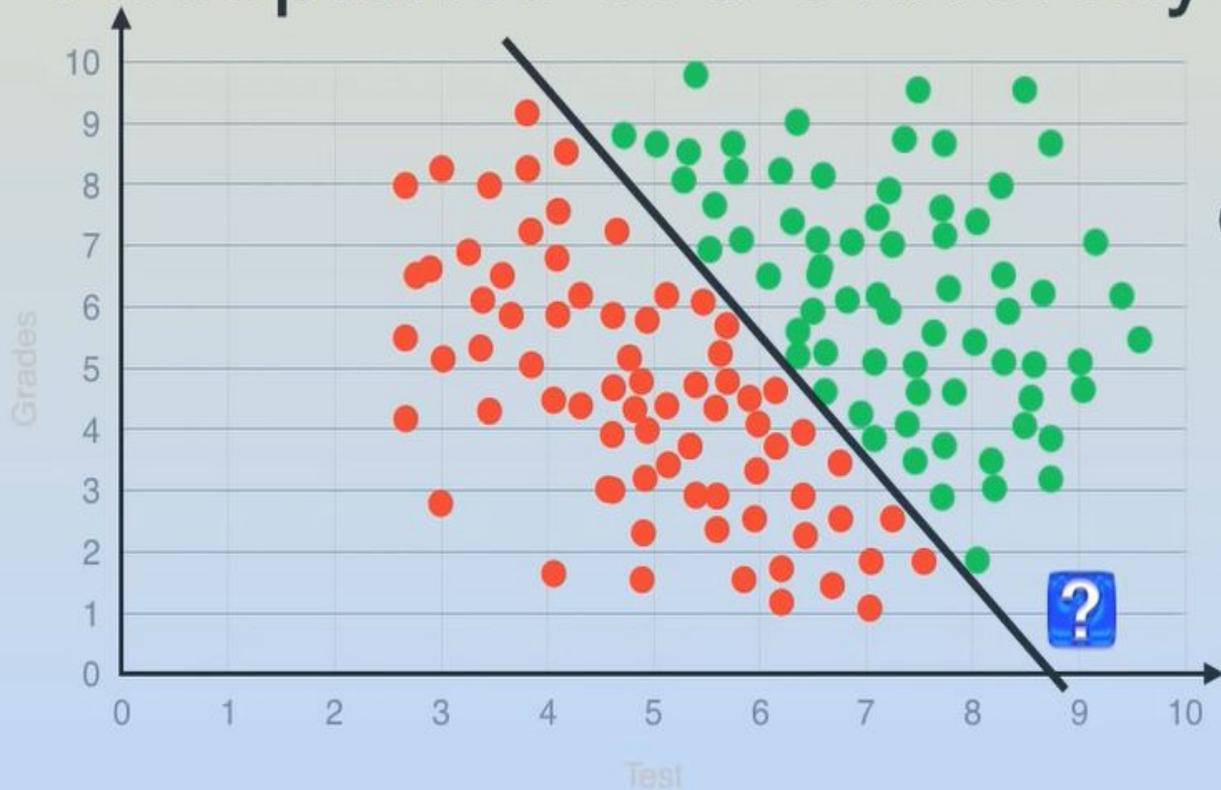
# Acceptance at a University



Student 4  
Test: 9/10  
Grades: 1/10

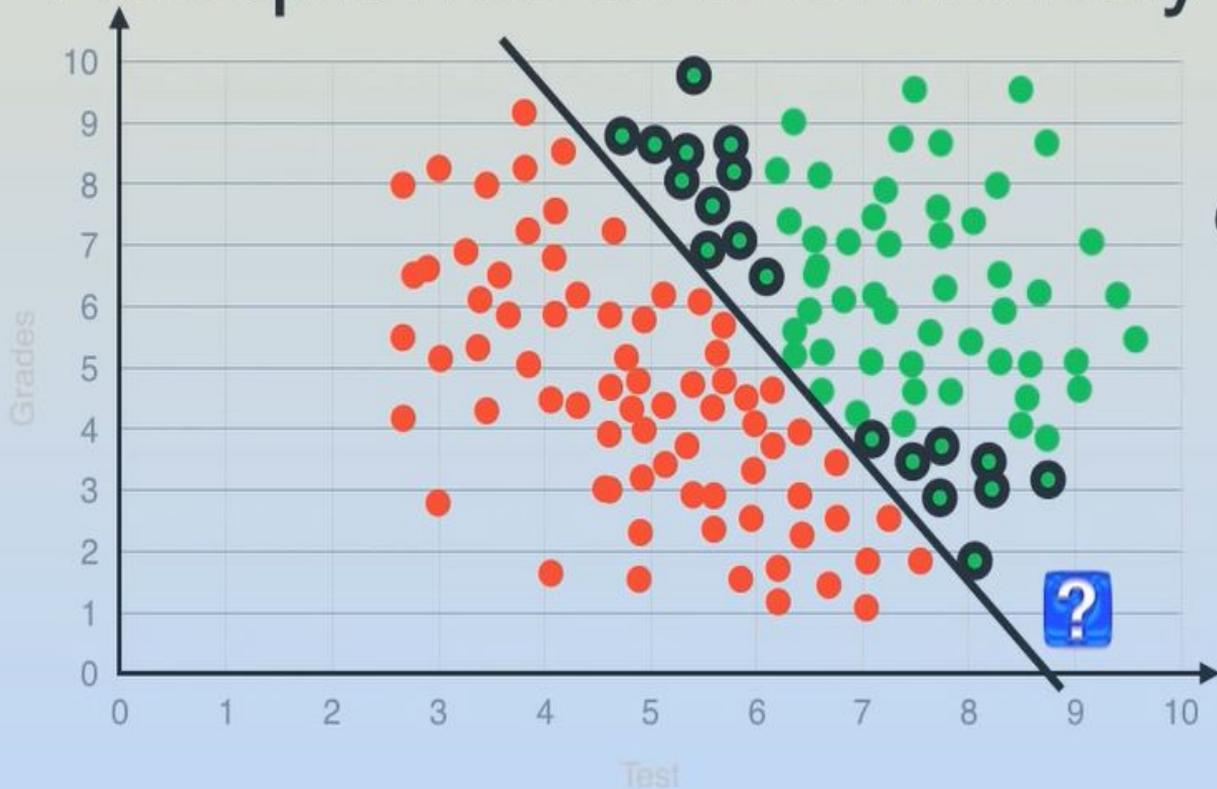


# Acceptance at a University



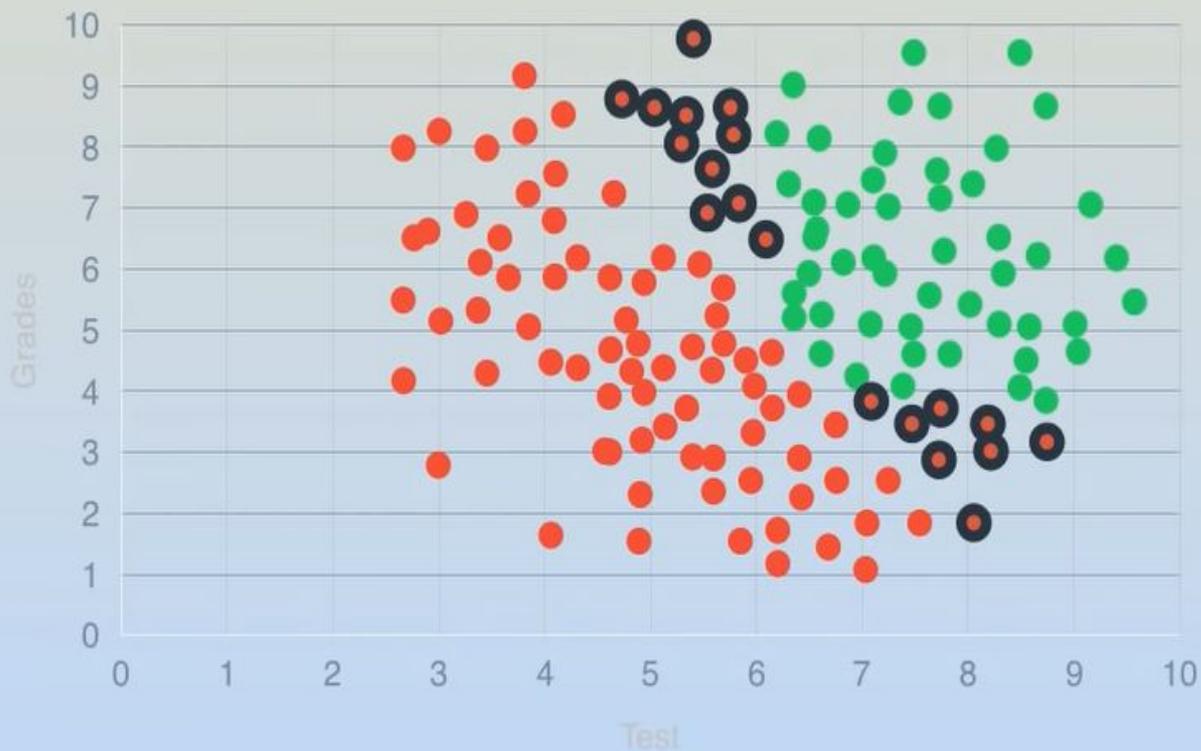
Student 4  
Test: 9/10  
Grades: 1/10

# Acceptance at a University

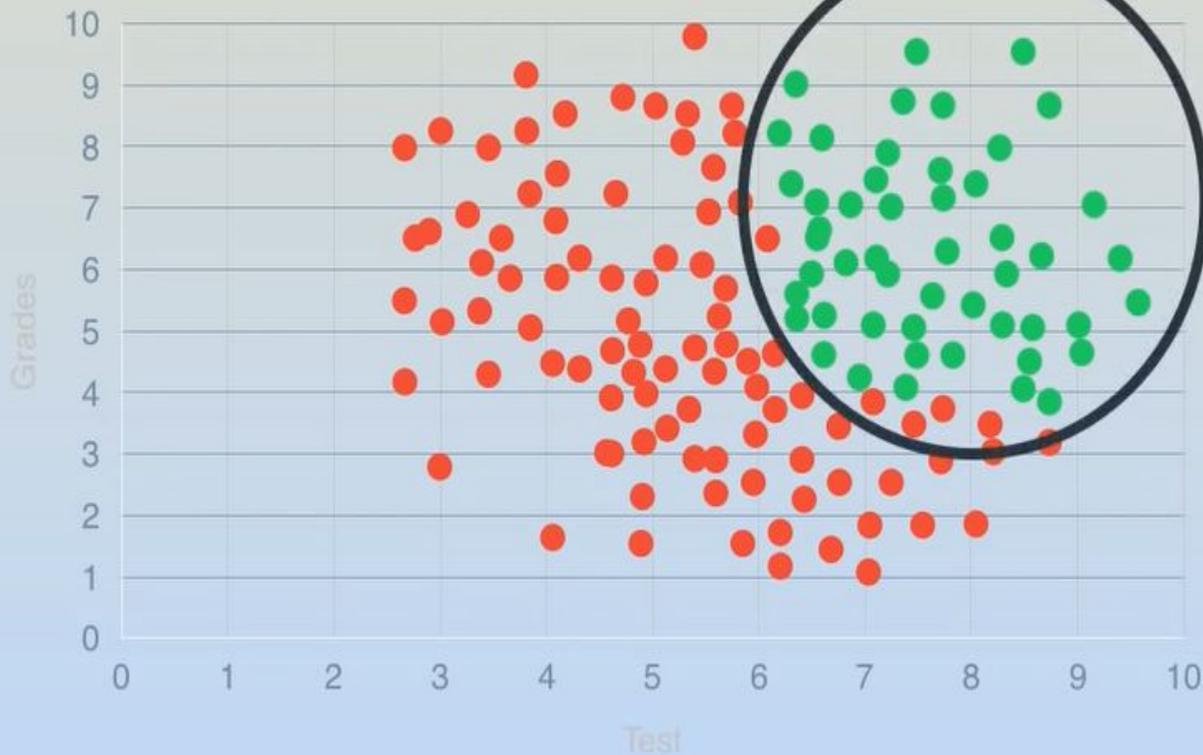


Student 4  
Test: 9/10  
Grades: 1/10

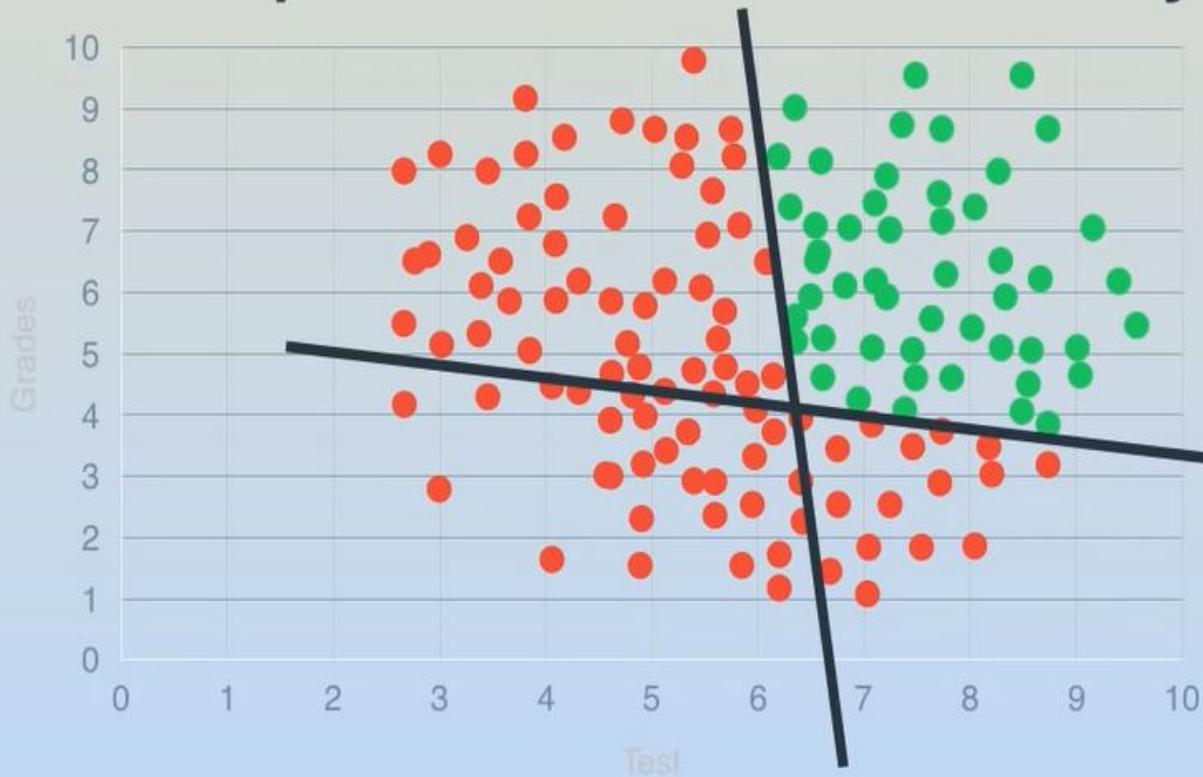
# Acceptance at a University



# Acceptance at a University

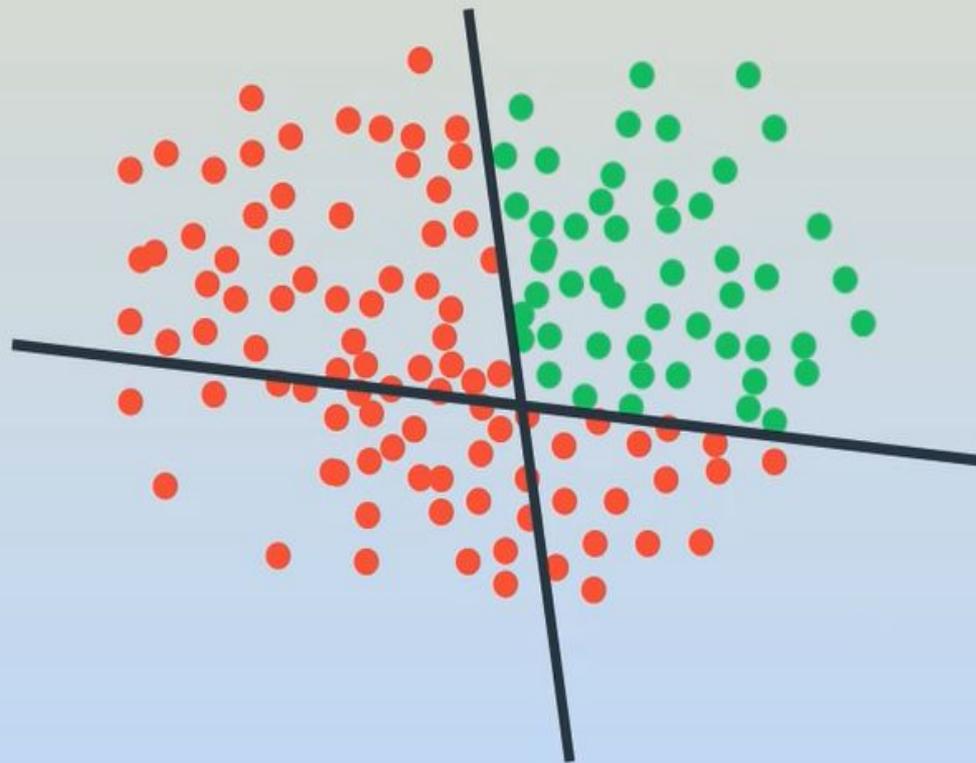


# Acceptance at a University

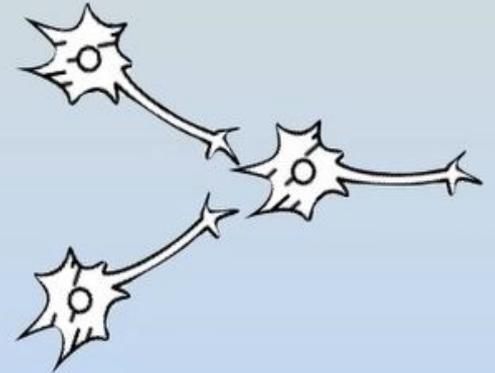
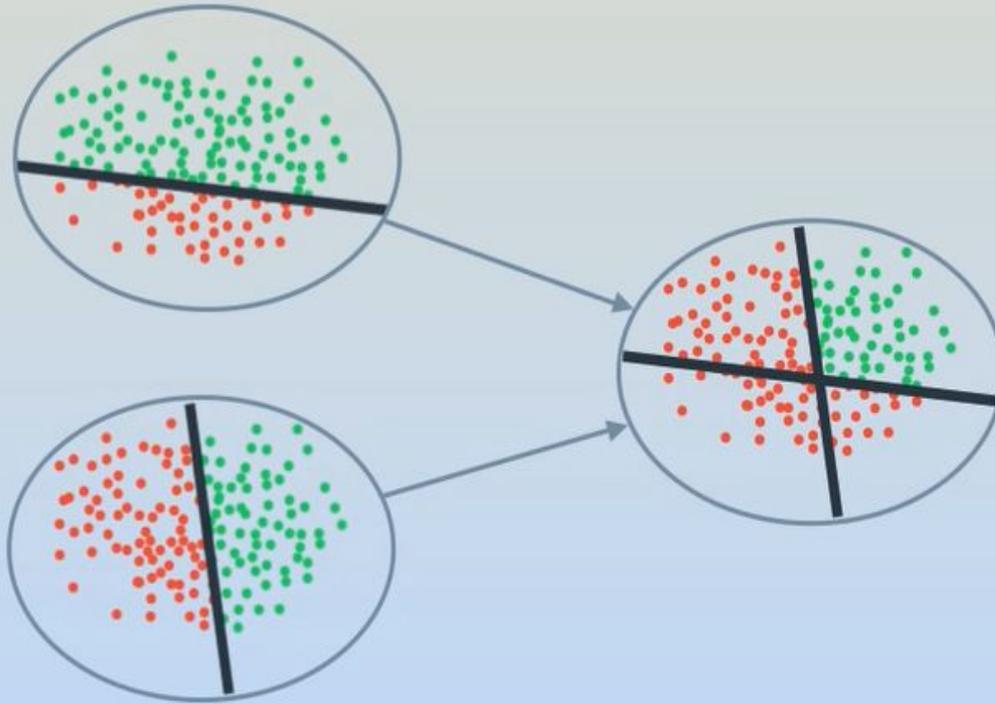


Gradient descent

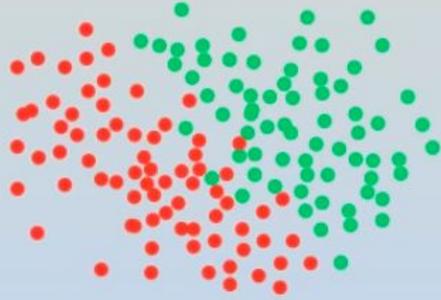
# Neural Network



# Neural Network



# Logistic Regression & Neural Networks

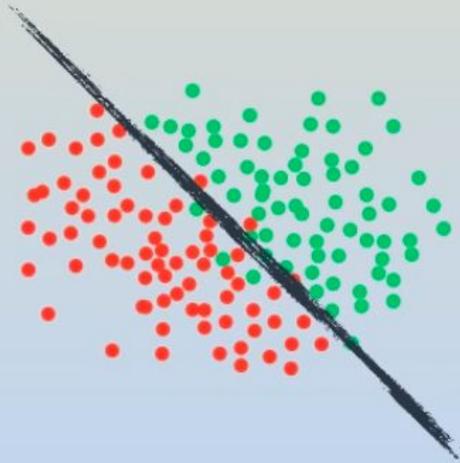


Logistic Regression

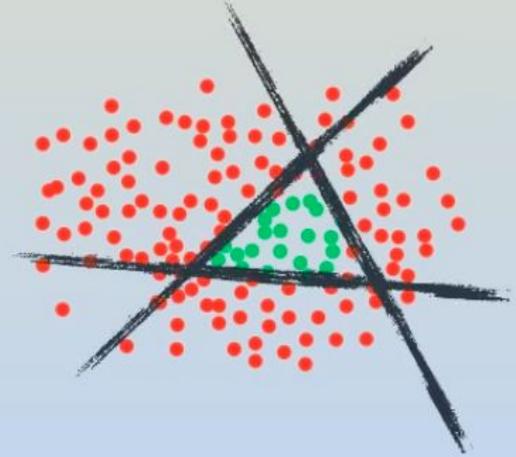
# Logistic Regression & Neural Networks



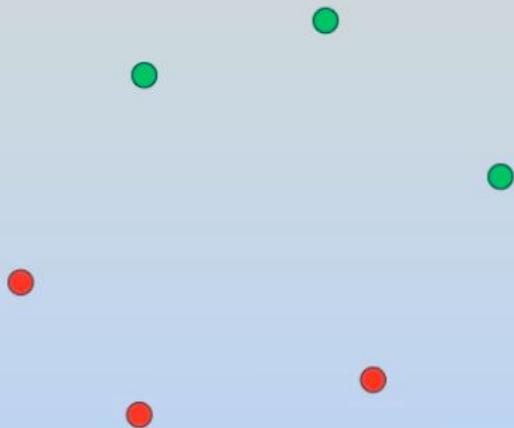
Logistic Regression



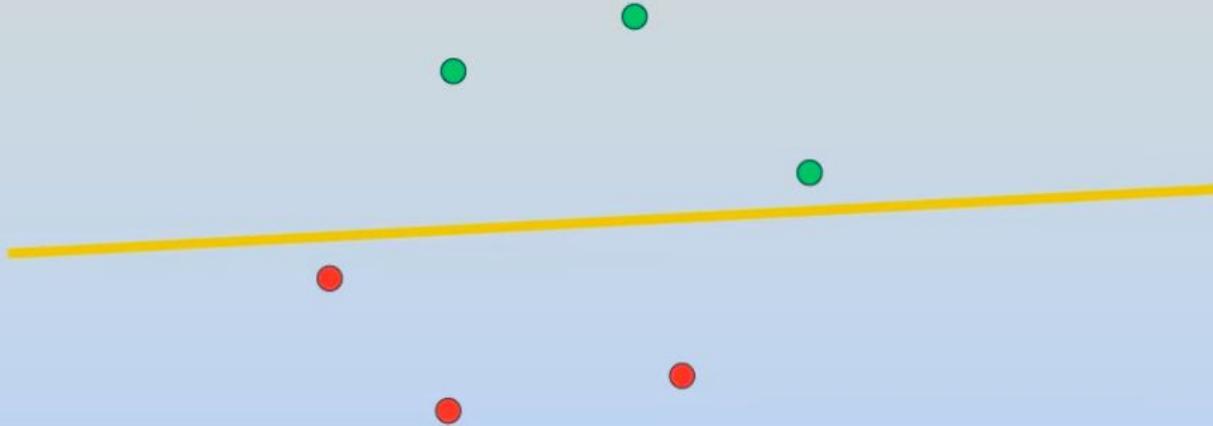
Neural Network



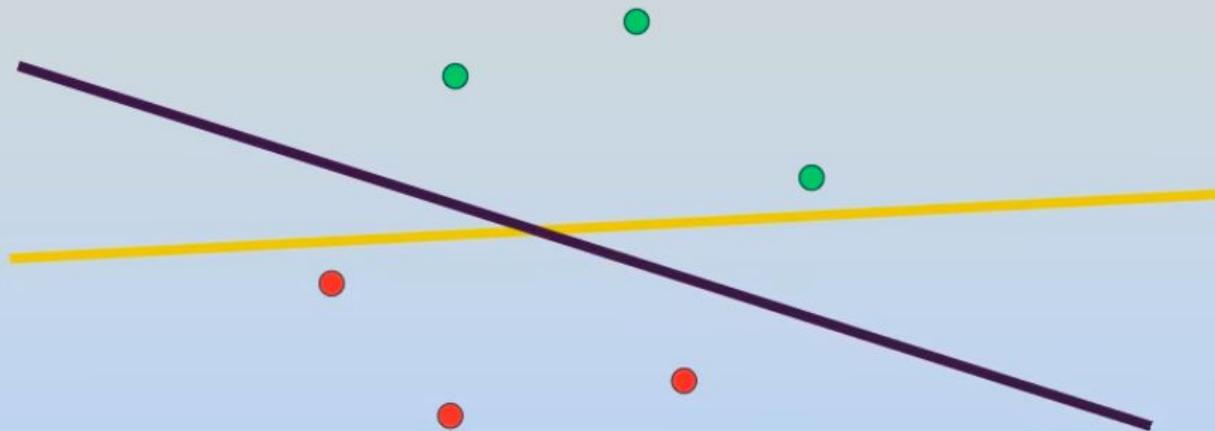
# Cutting data with style



# Cutting data with style

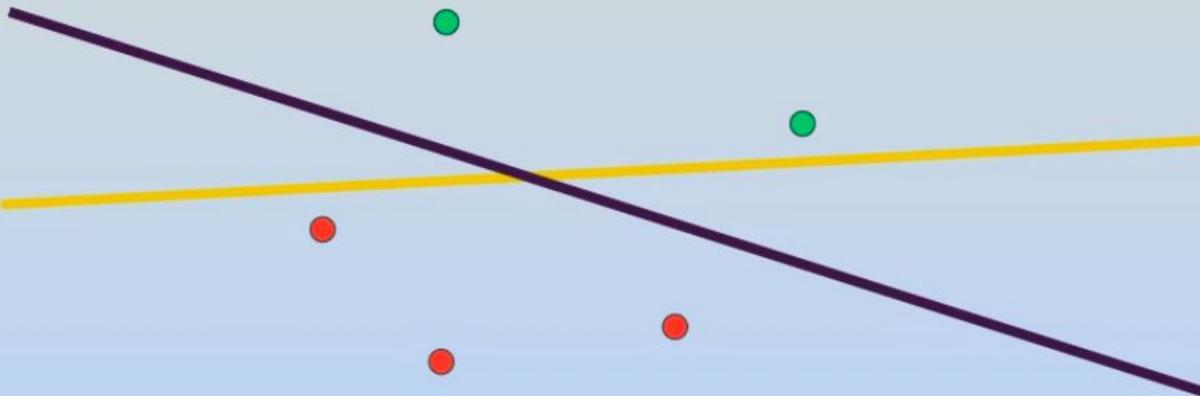


# Cutting data with style



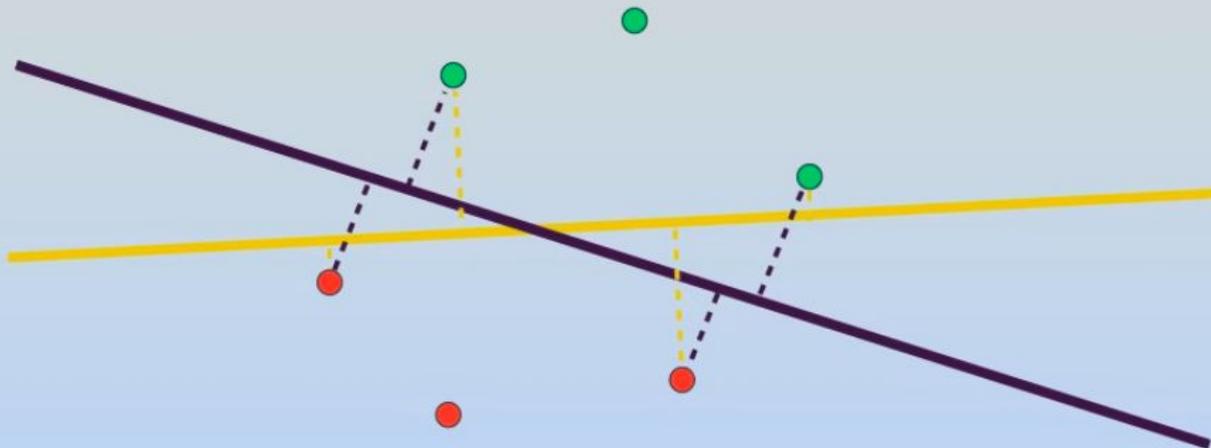
# Cutting data with style

Quiz:  
Which one is a  
better line?



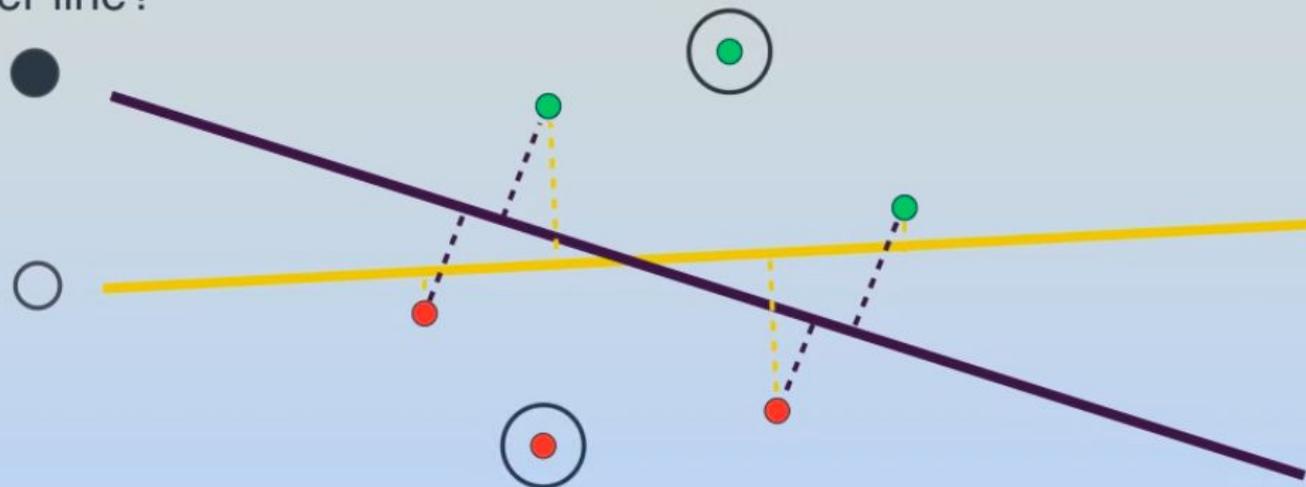
# Cutting data with style

Quiz:  
Which one is a  
better line?



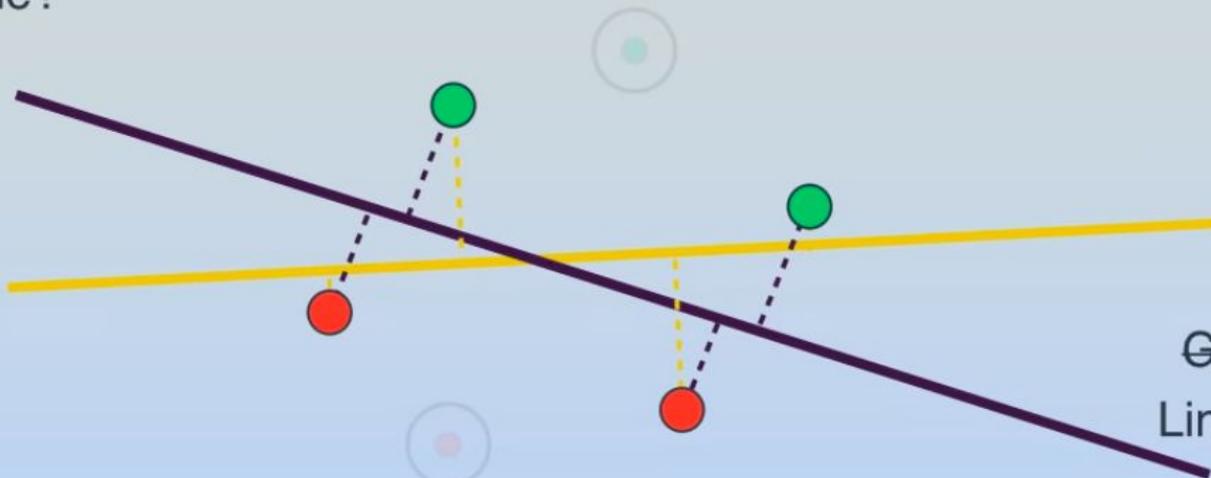
# Cutting data with style

Quiz:  
Which one is a  
better line?



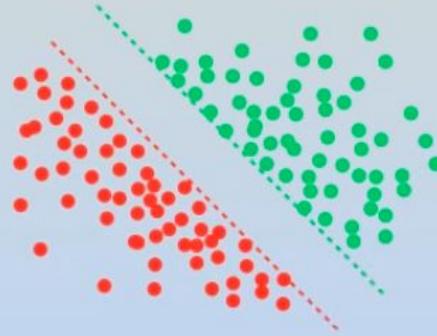
# Support Vector Machine

Quiz:  
Which one is a  
better line?

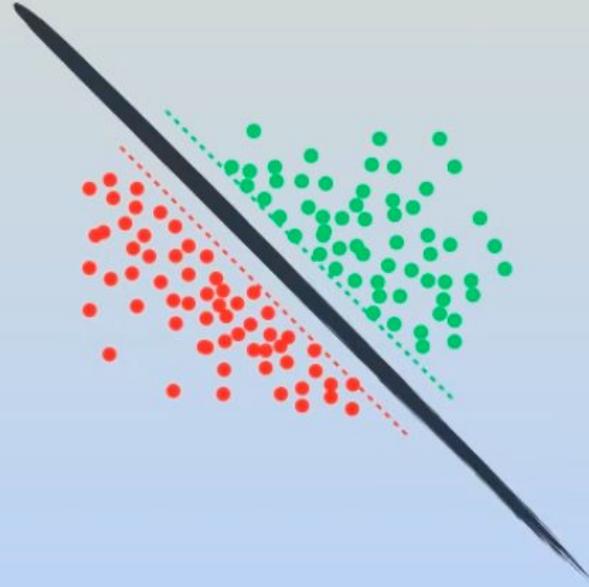


Gradient-descent  
Linear Optimization

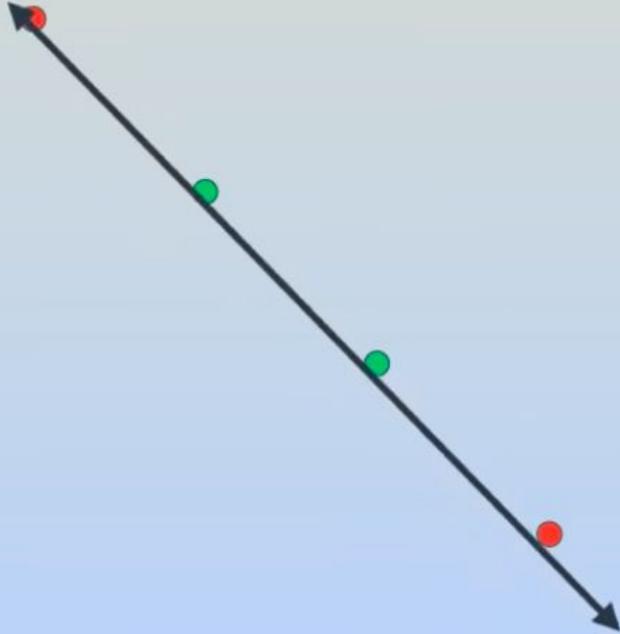
# Support Vector Machines



# Support Vector Machines



When a line is not enough...

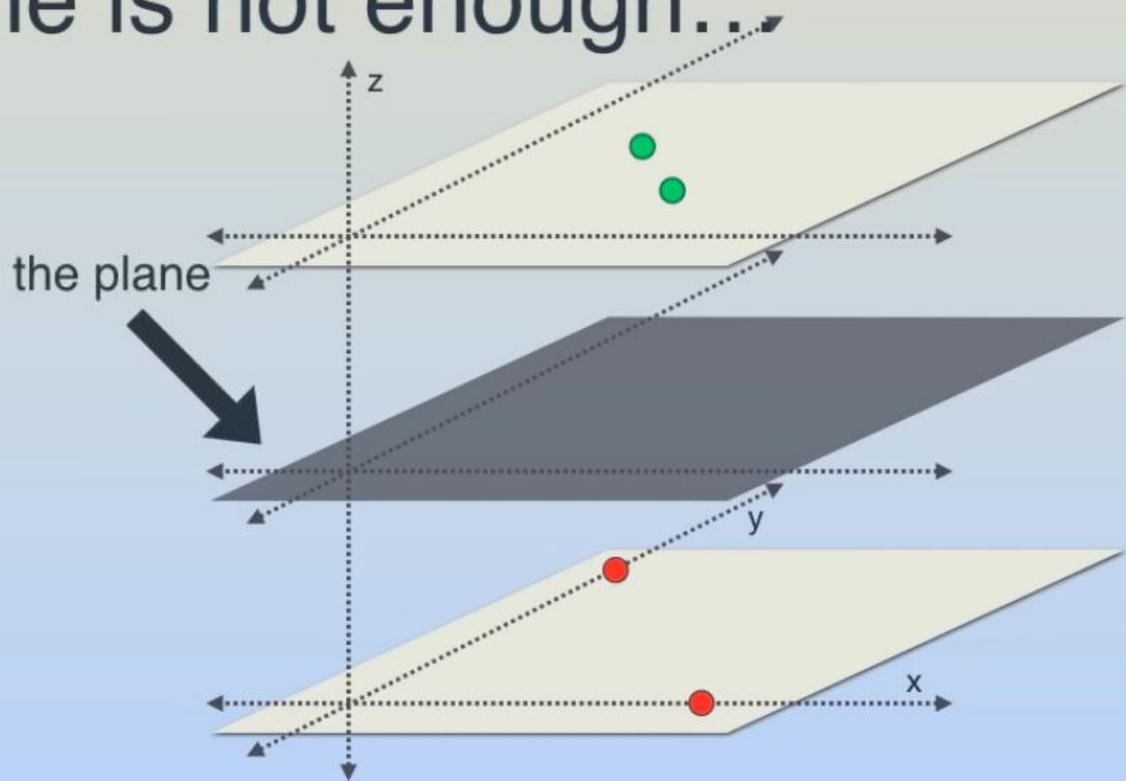
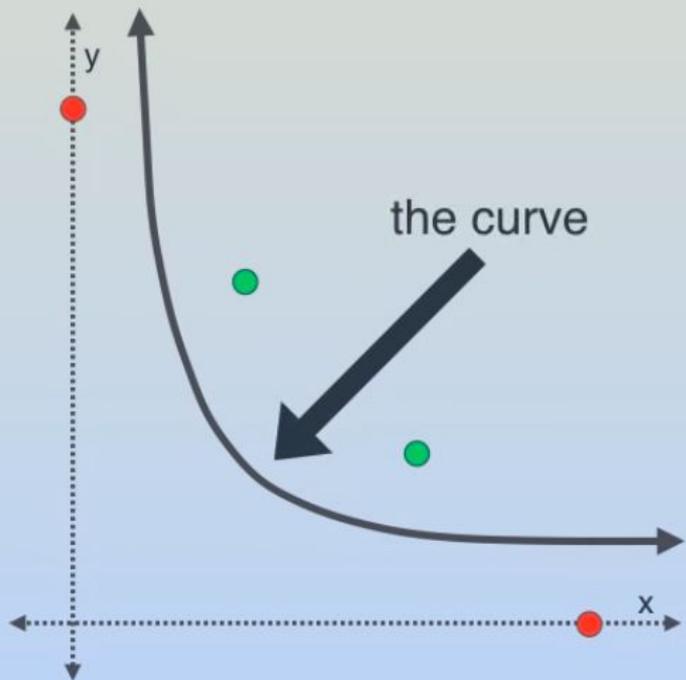




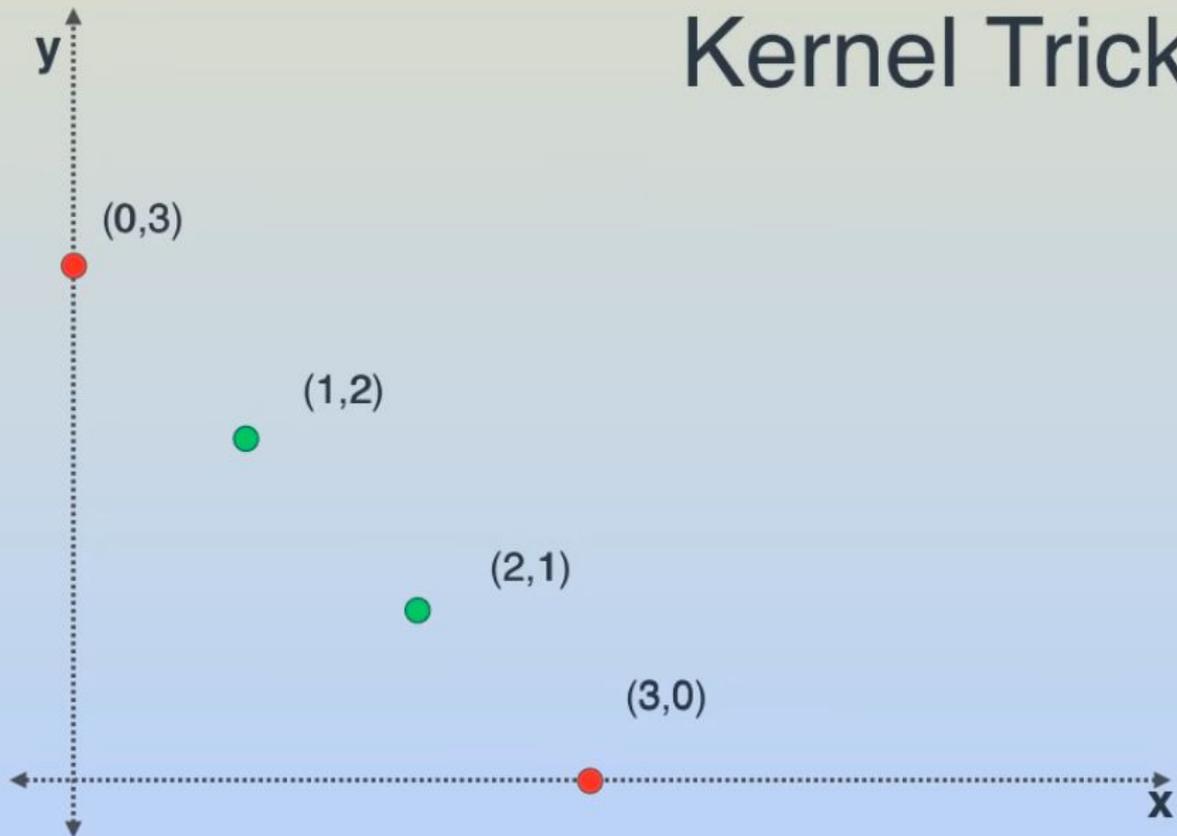
When a line is not enough...



# When a line is not enough...



# Kernel Trick



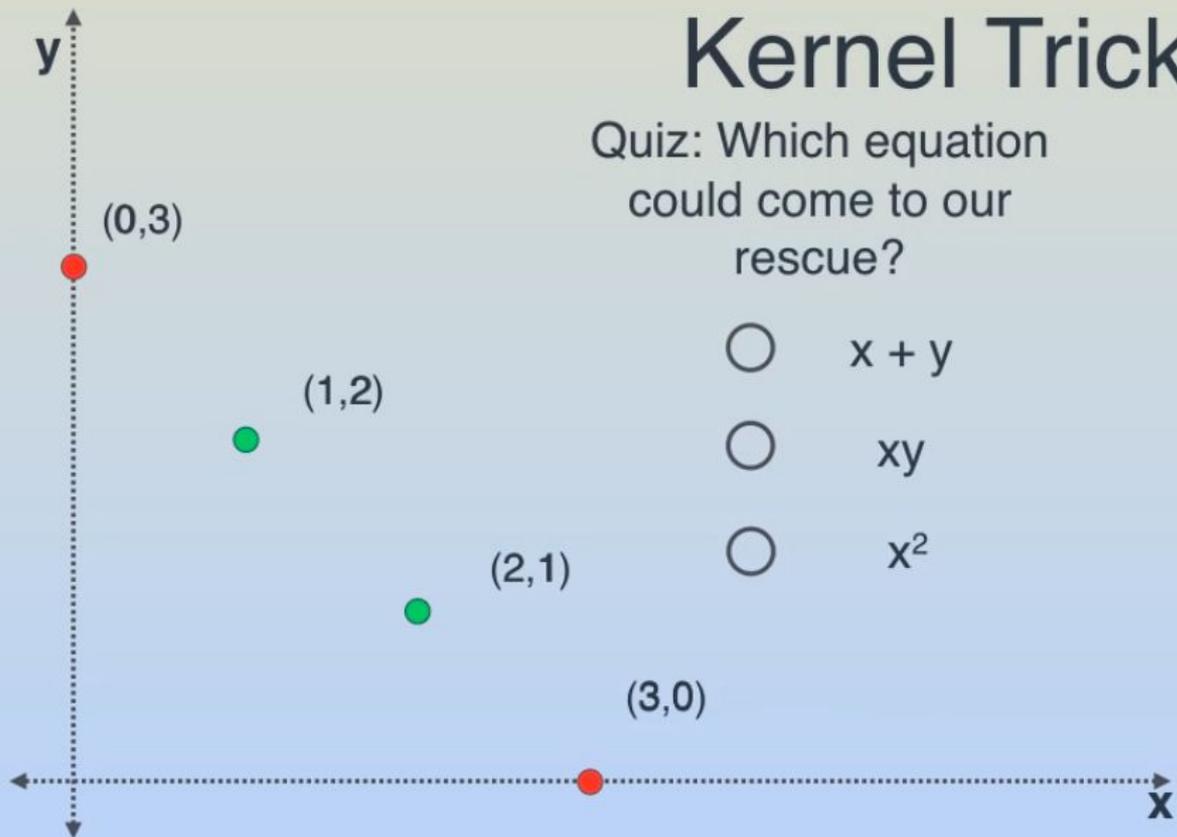
# Kernel Trick

Quiz: Which equation  
could come to our  
rescue?

$x + y$

$xy$

$x^2$



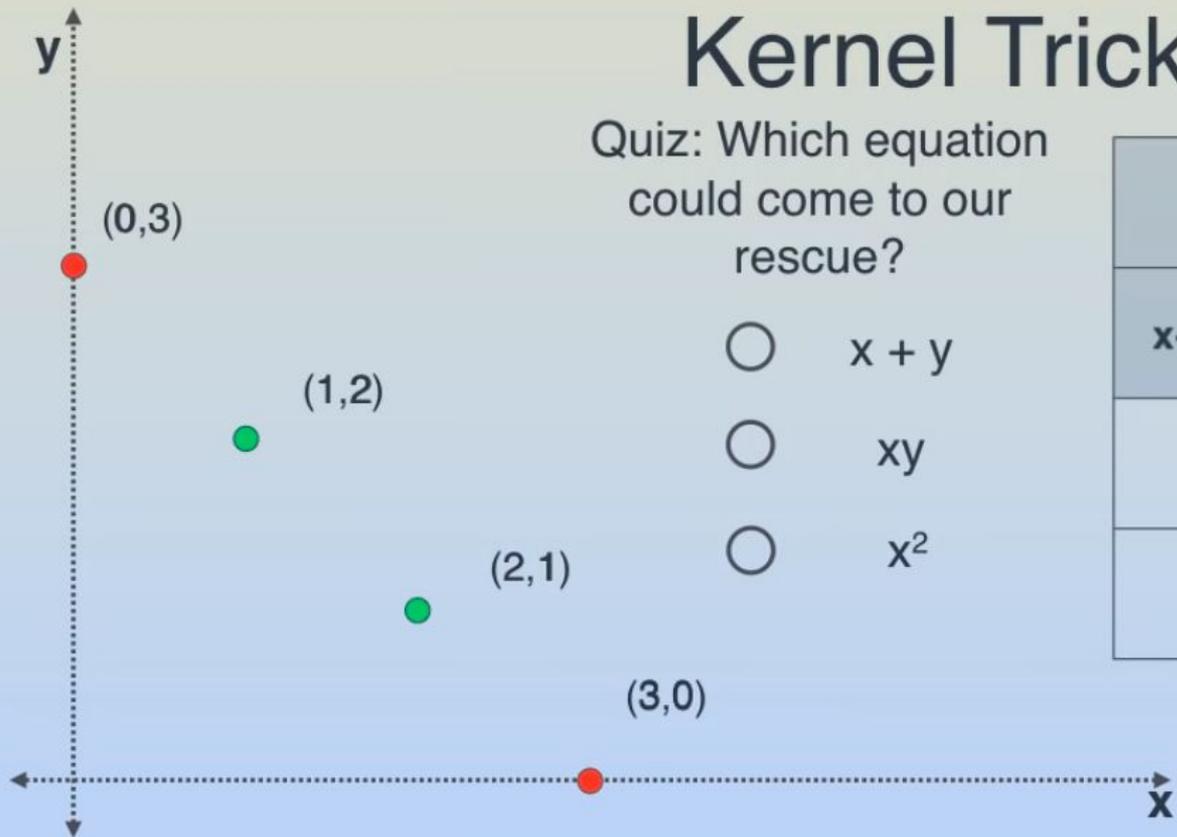
# Kernel Trick

Quiz: Which equation could come to our rescue?

$x + y$

$xy$

$x^2$



	(0,3)	(1,2)	(2,1)	(3,0)
x+y	3	3	3	3

# Kernel Trick

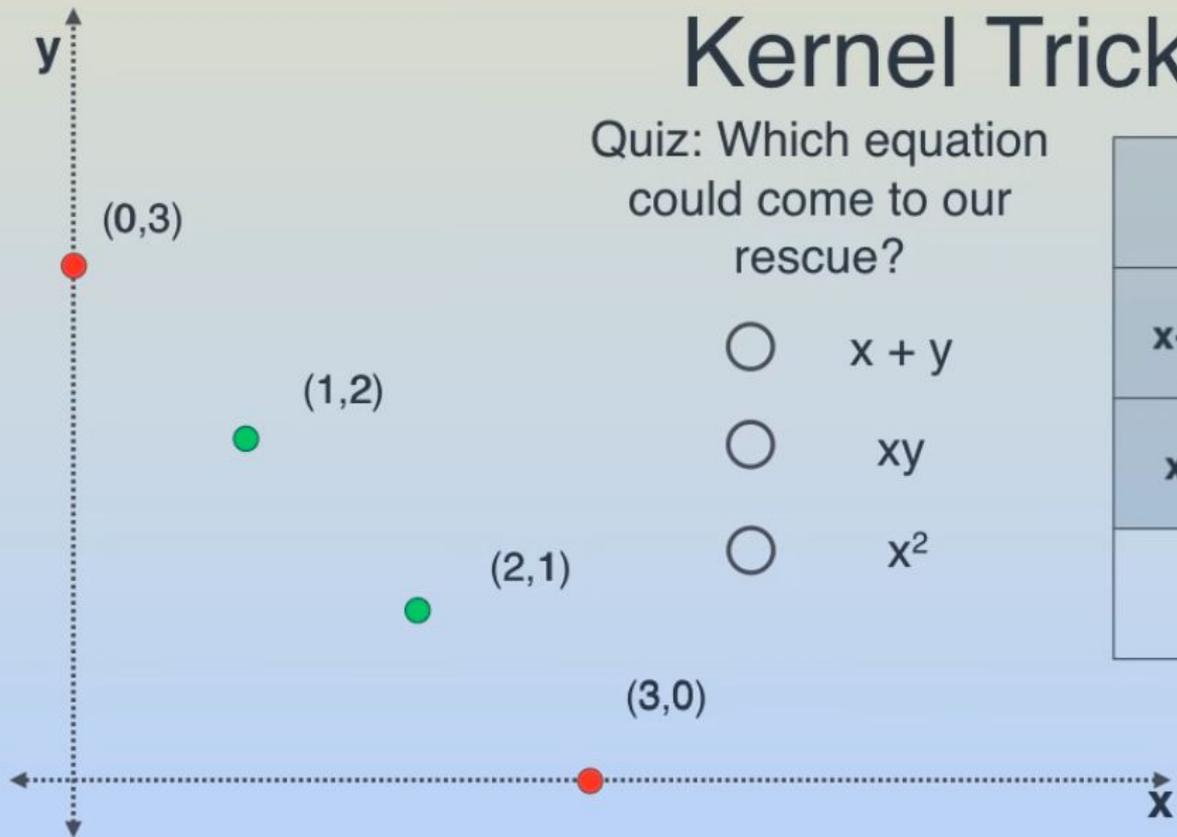
Quiz: Which equation could come to our rescue?

$x + y$

$xy$

$x^2$

	(0,3)	(1,2)	(2,1)	(3,0)
x+y	3	3	3	3
xy	0	2	2	0



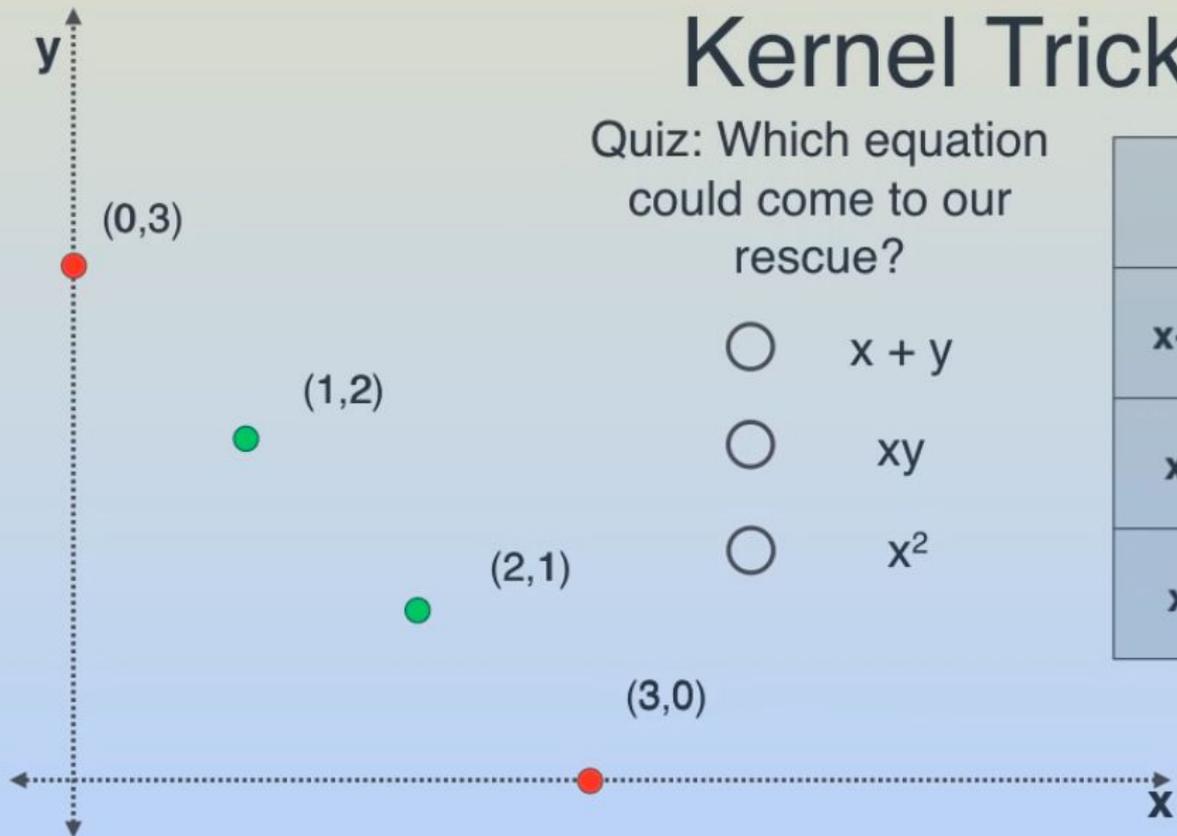
# Kernel Trick

Quiz: Which equation could come to our rescue?

$x + y$

$xy$

$x^2$



	(0,3)	(1,2)	(2,1)	(3,0)
$x+y$	3	3	3	3
$xy$	0	2	2	0
$x^2$	0	1	4	9

# Kernel Trick

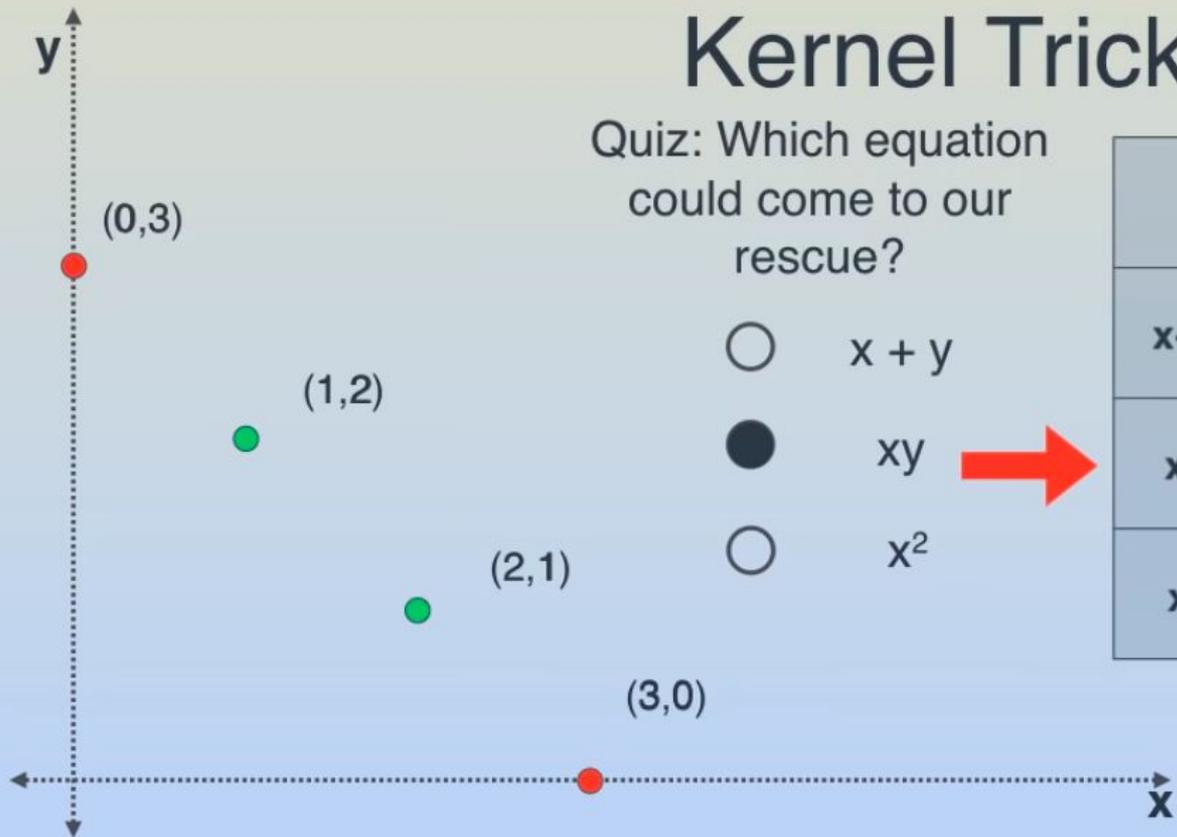
Quiz: Which equation could come to our rescue?

$x + y$

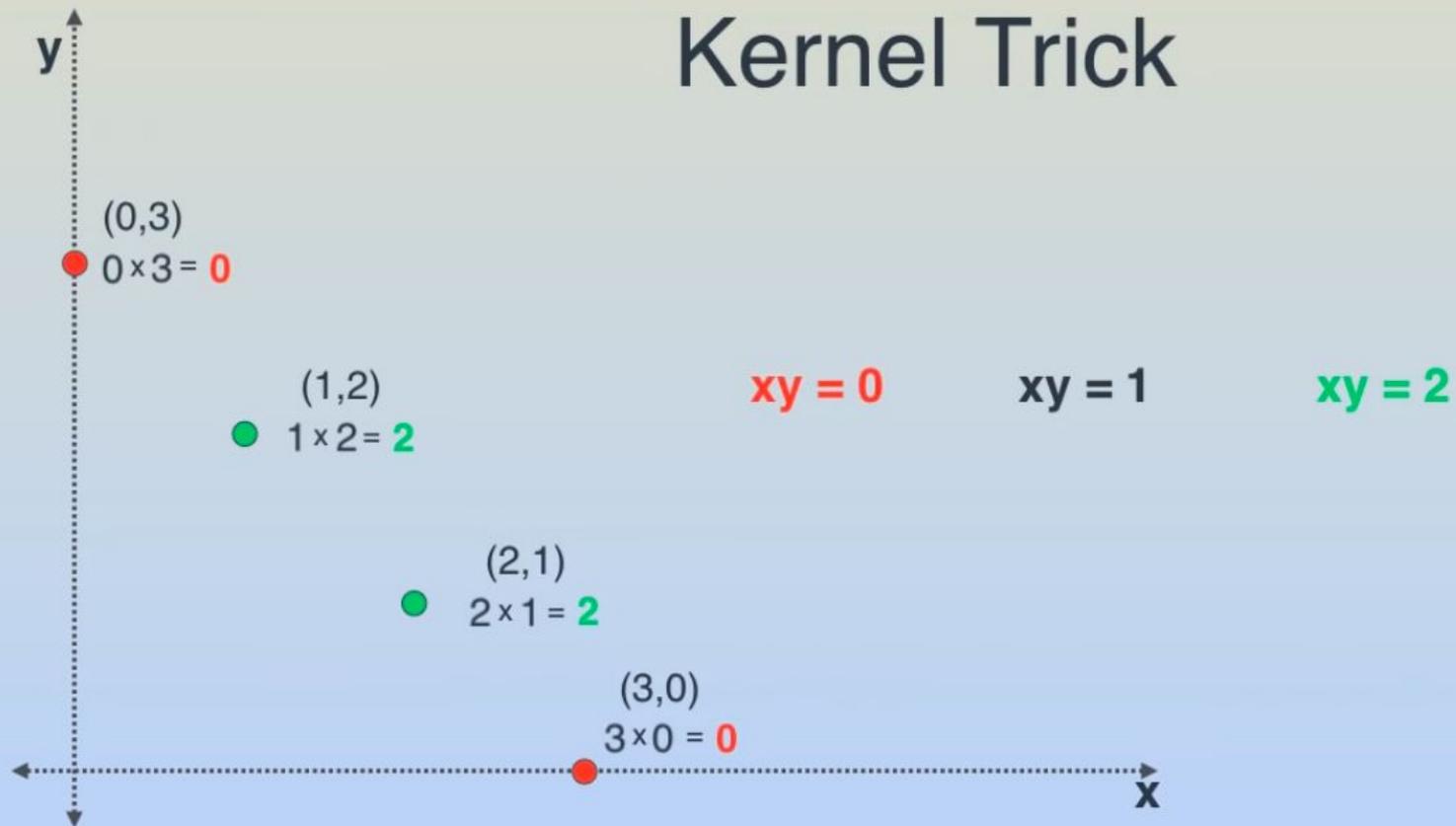
$xy$  

$x^2$

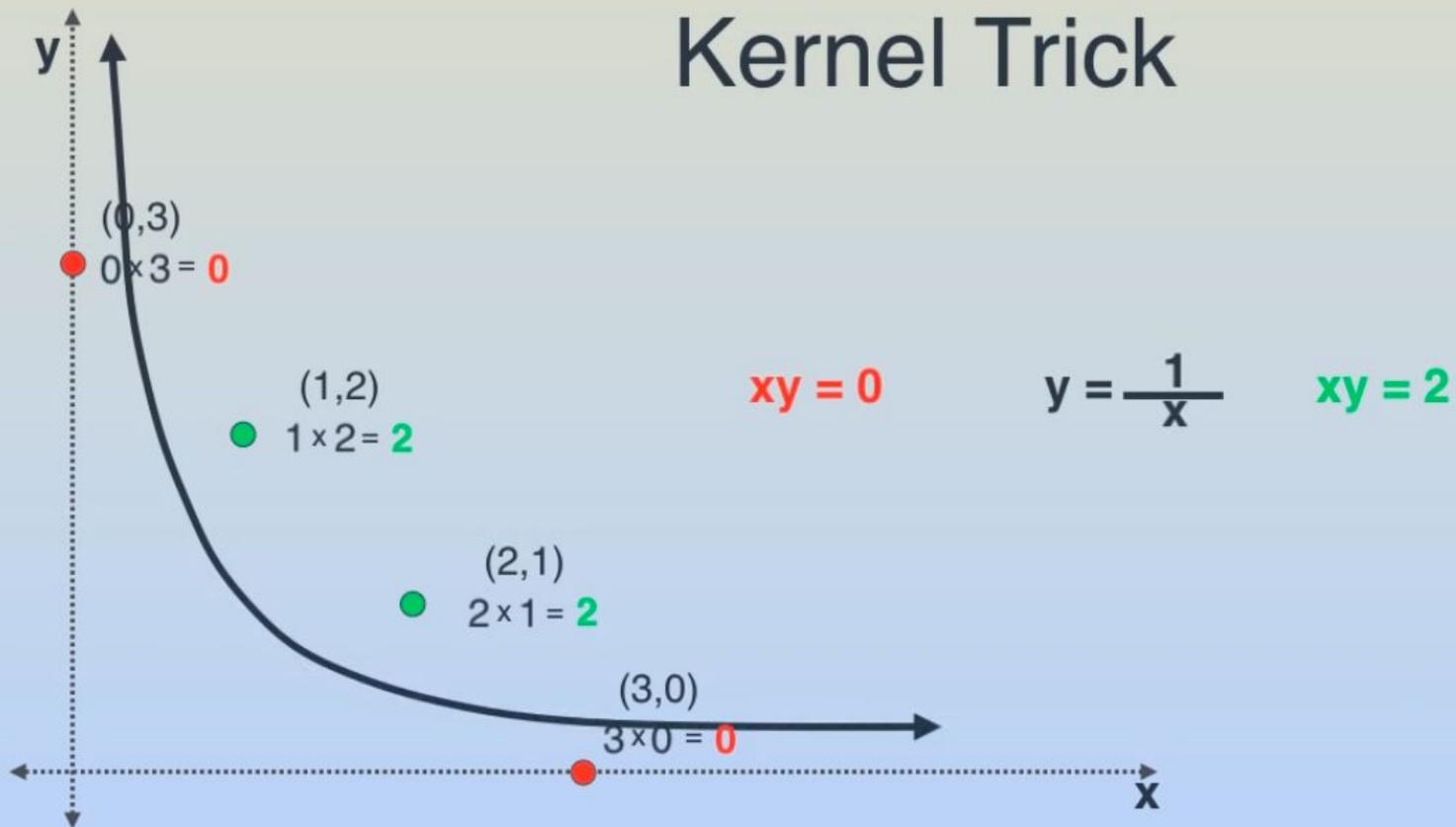
	(0,3)	(1,2)	(2,1)	(3,0)
$x+y$	3	3	3	3
$xy$	0	2	2	0
$x^2$	0	1	4	9



# Kernel Trick



# Kernel Trick



# Kernel Trick

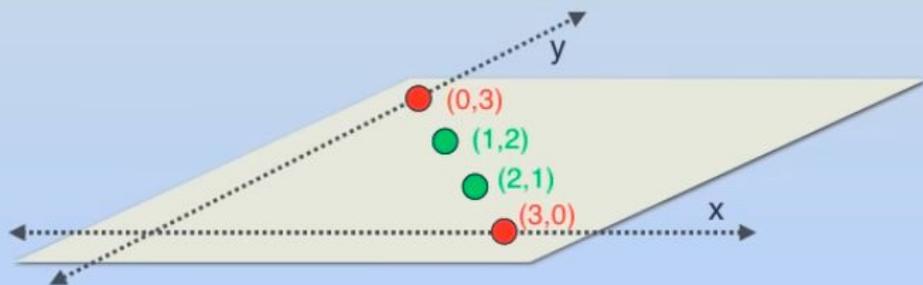
$(x,y)$

$(0,3)$

$(1,2)$

$(2,1)$

$(3,0)$



# Kernel Trick

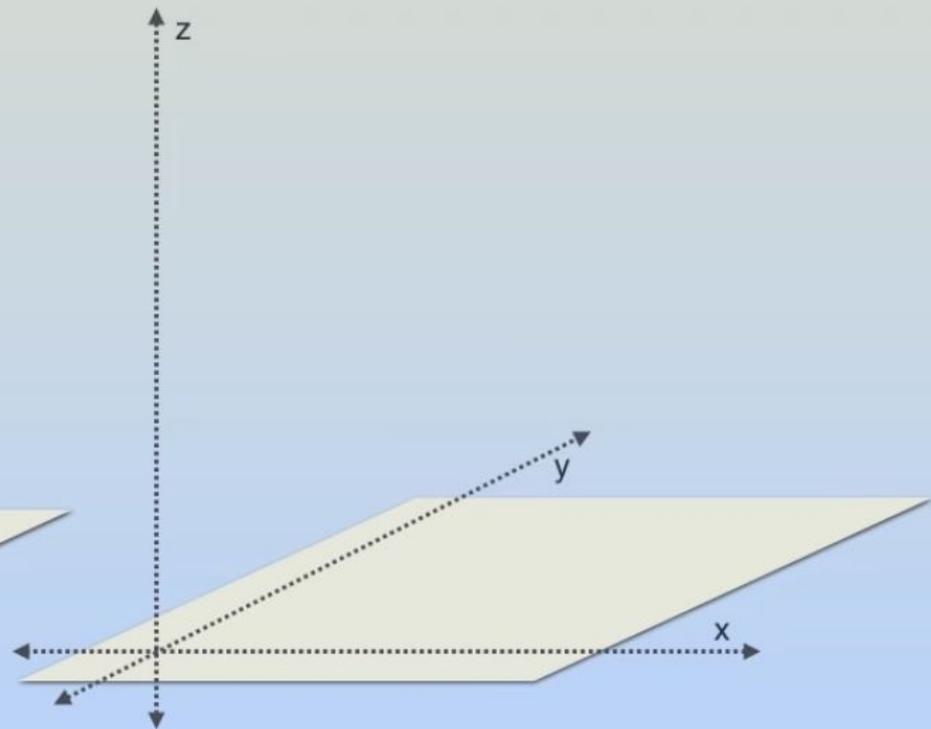
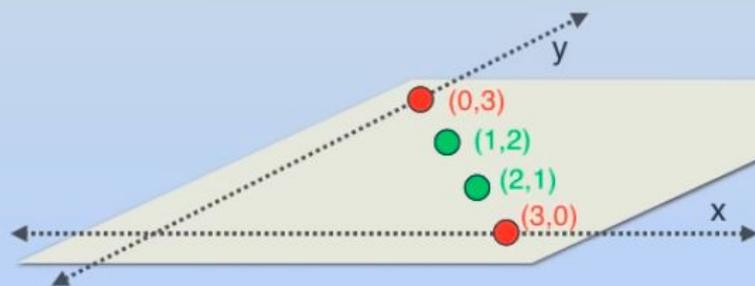
$(x,y) \longrightarrow (x,y,xy)$

$(0,3)$

$(1,2)$

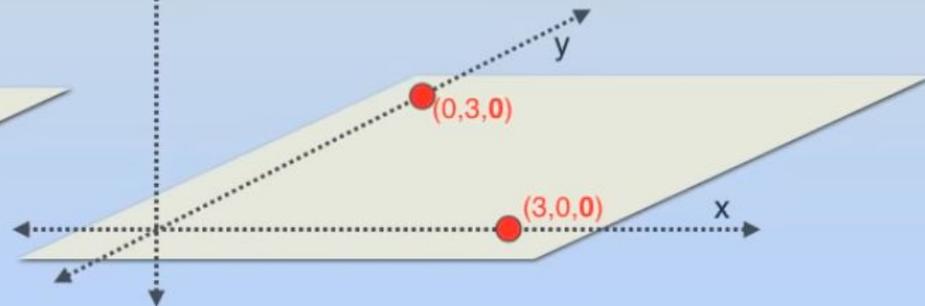
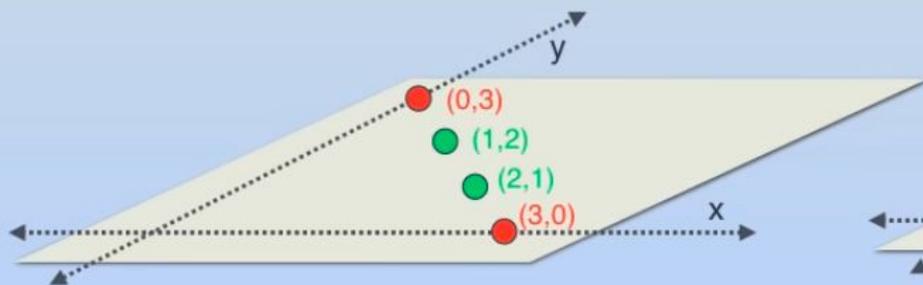
$(2,1)$

$(3,0)$



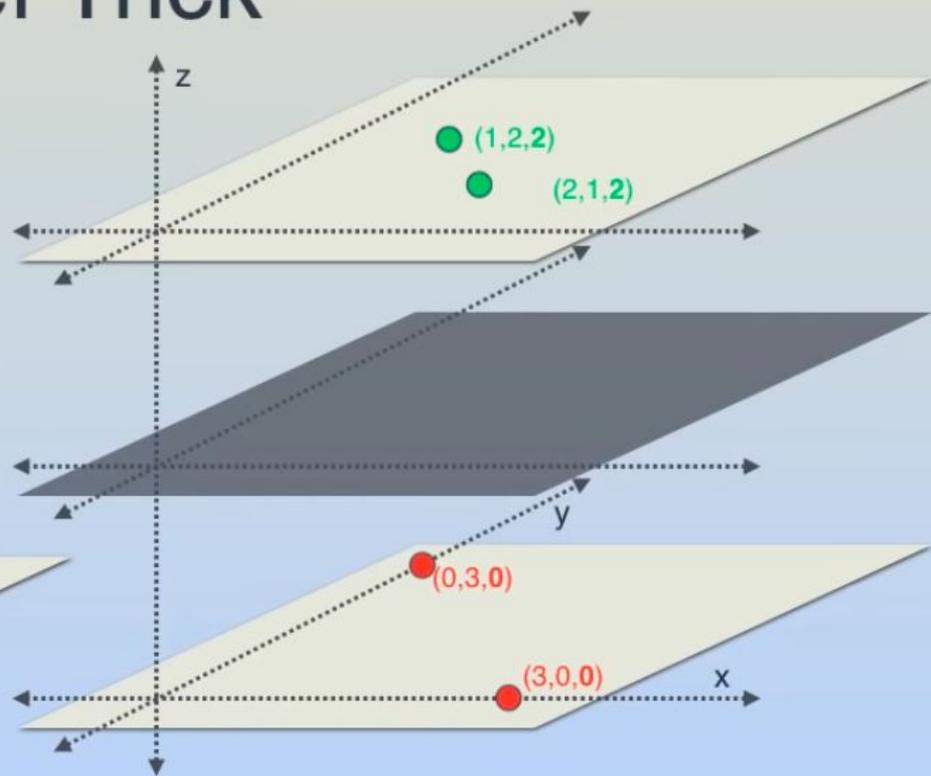
# Kernel Trick

$(x,y) \longrightarrow (x,y,xy)$   
 $(0,3) \longrightarrow (0,3,0)$   
 $(1,2) \longrightarrow (1,2,2)$   
 $(2,1) \longrightarrow (2,1,2)$   
 $(3,0) \longrightarrow (3,0,0)$



# Kernel Trick

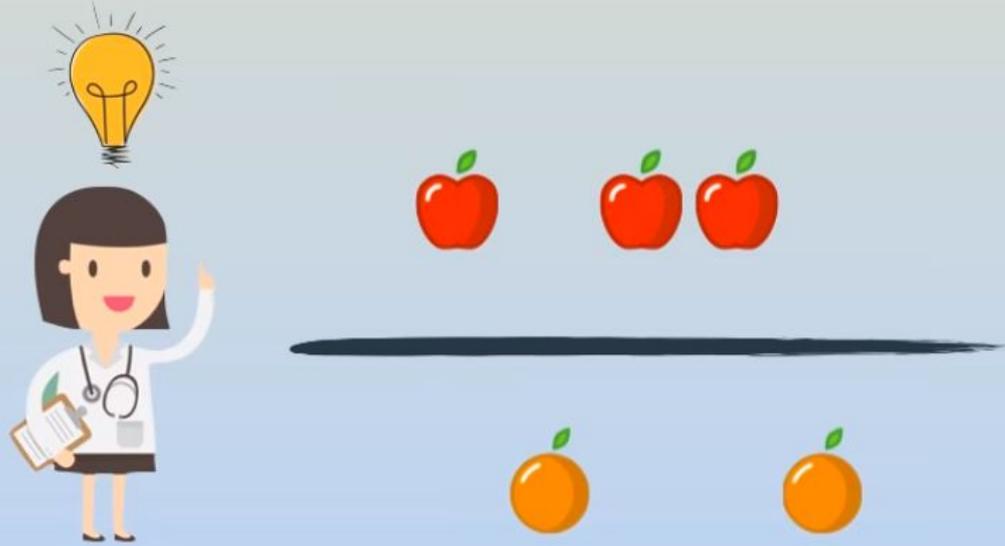
$(x,y) \longrightarrow (x,y,xy)$   
 $(0,3) \longrightarrow (0,3,0)$   
 $(1,2) \longrightarrow (1,2,2)$   
 $(2,1) \longrightarrow (2,1,2)$   
 $(3,0) \longrightarrow (3,0,0)$



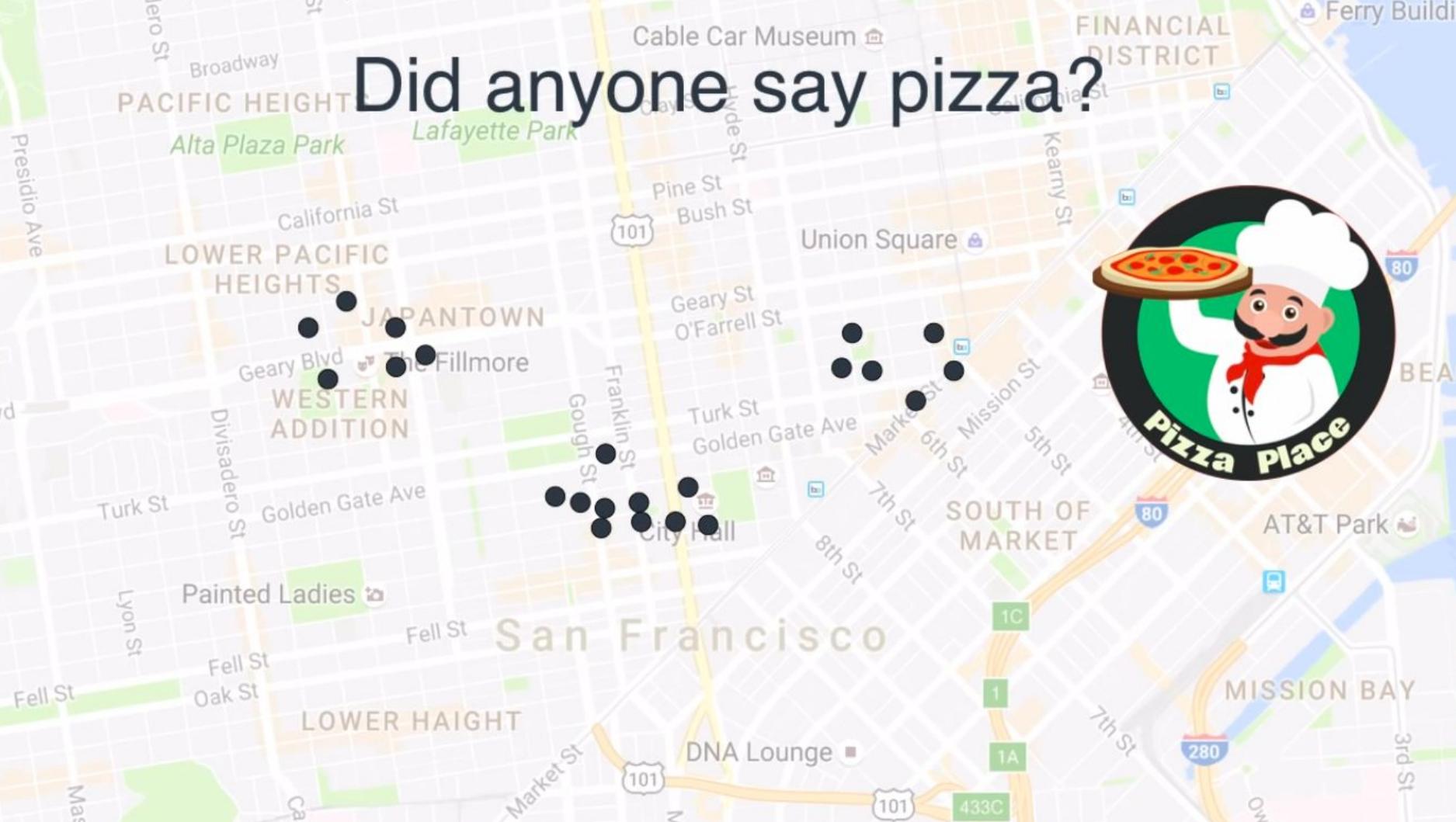
# Support Vector Machine Kernel Trick



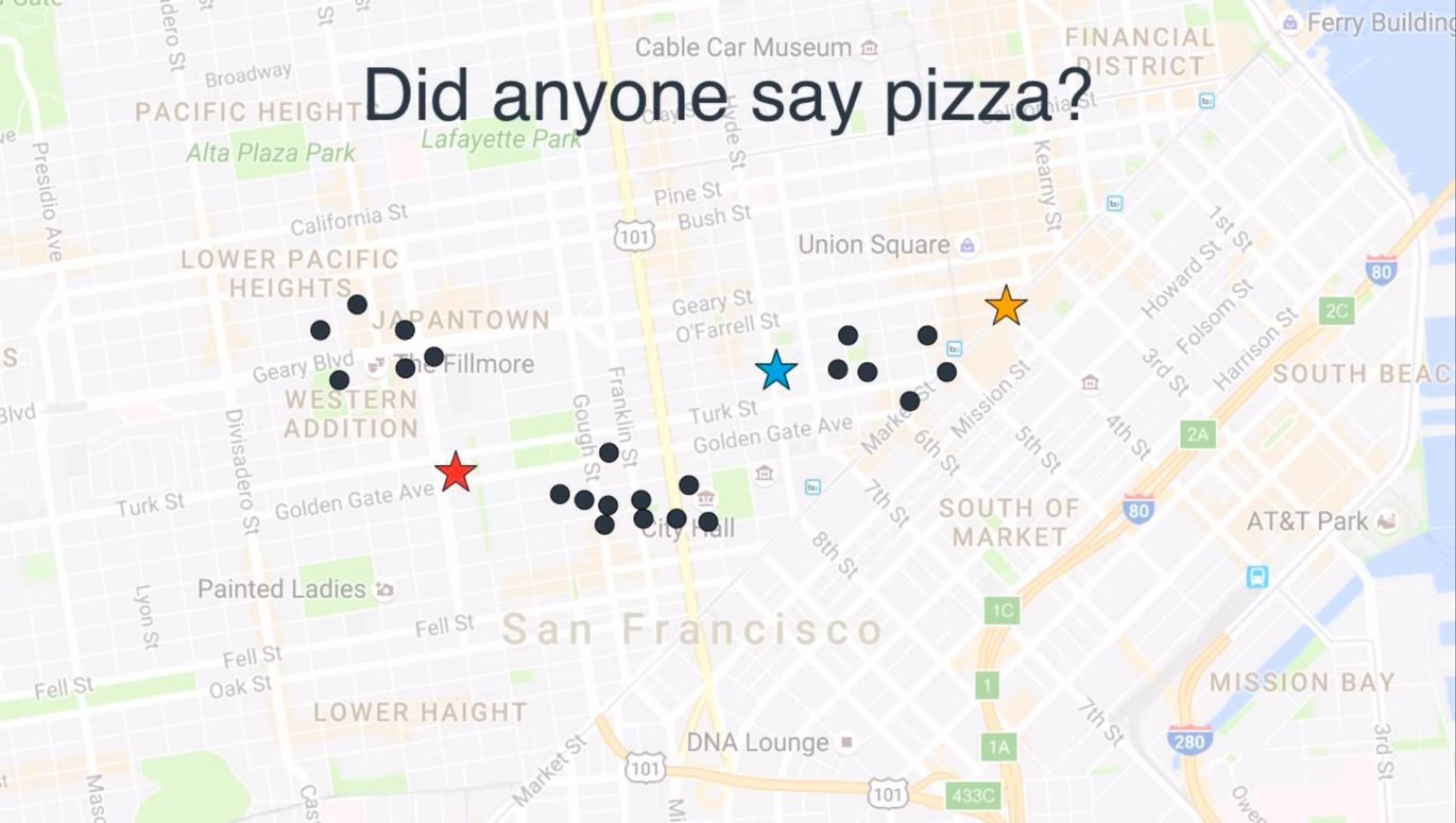
# Support Vector Machine Kernel Trick



# Did anyone say pizza?



# Did anyone say pizza?



# Did anyone say pizza?



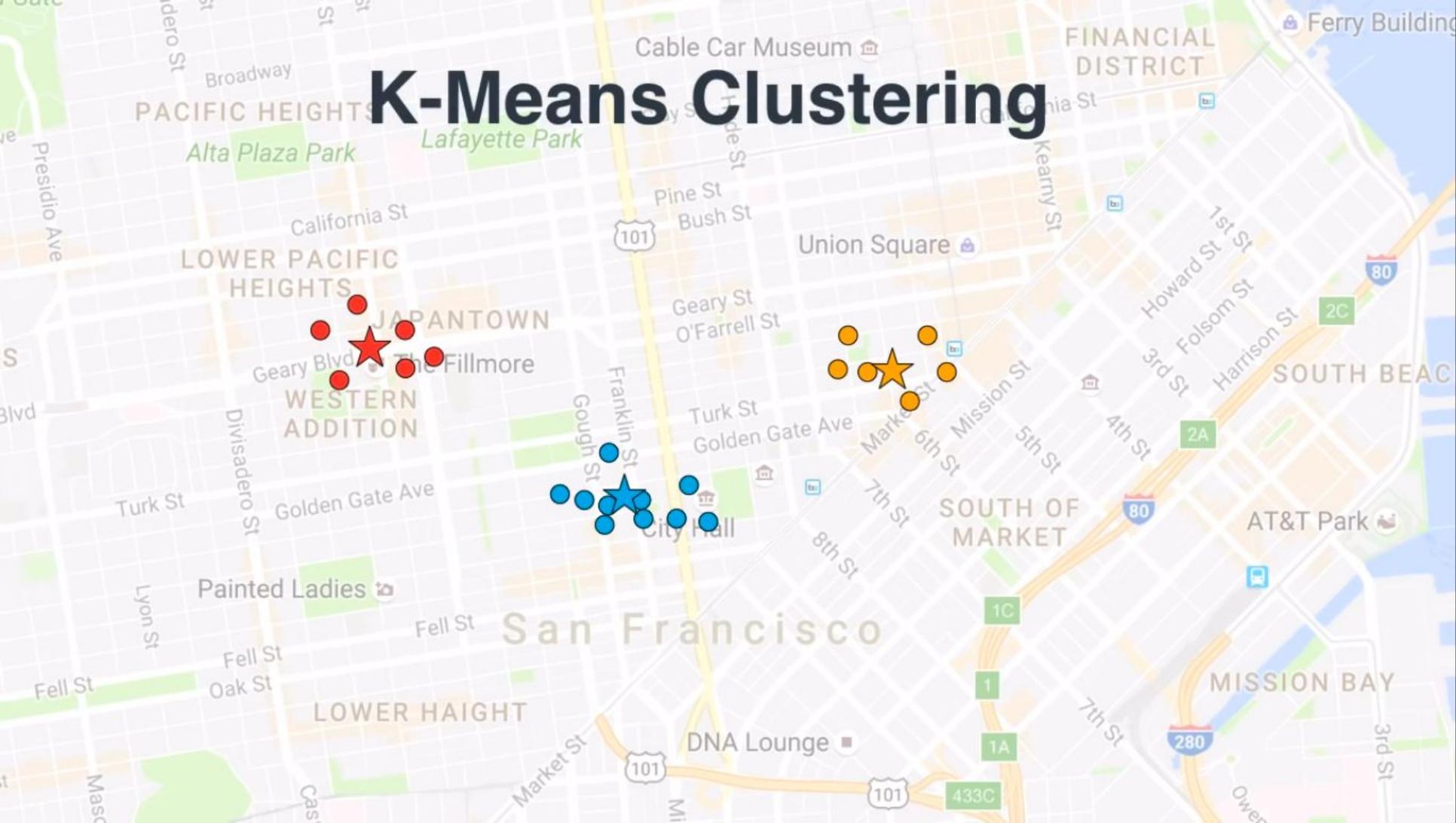
# Did anyone say pizza?



# Did anyone say pizza?



# K-Means Clustering



# Summary

- ❑ Linear Regression
- ❑ Naive Bayes
- ❑ Decision Tree
- ❑ Logistic Regression
- ❑ Neural Network
- ❑ SVN (Kernel Trick)
- ❑ Clustering (K-Means)