

Machine Learning

Linear Separability

Mostafa S. Ibrahim

Teaching, Training and Coaching for more than a decade!

Artificial Intelligence & Computer Vision Researcher

PhD from Simon Fraser University - Canada

Bachelor / MSc from Cairo University - Egypt

Ex-(Software Engineer / ICPC World Finalist)

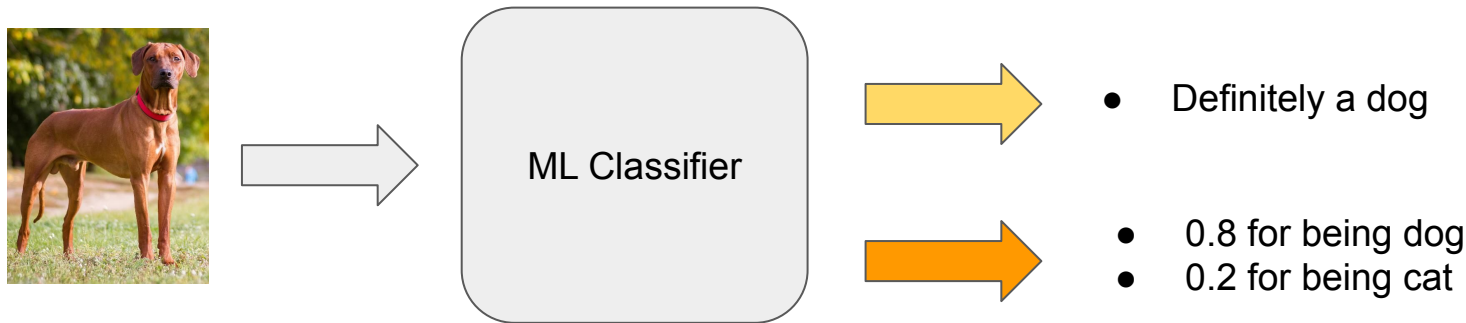


© 2023 All rights reserved.

Please do not reproduce or redistribute this work without permission from the author

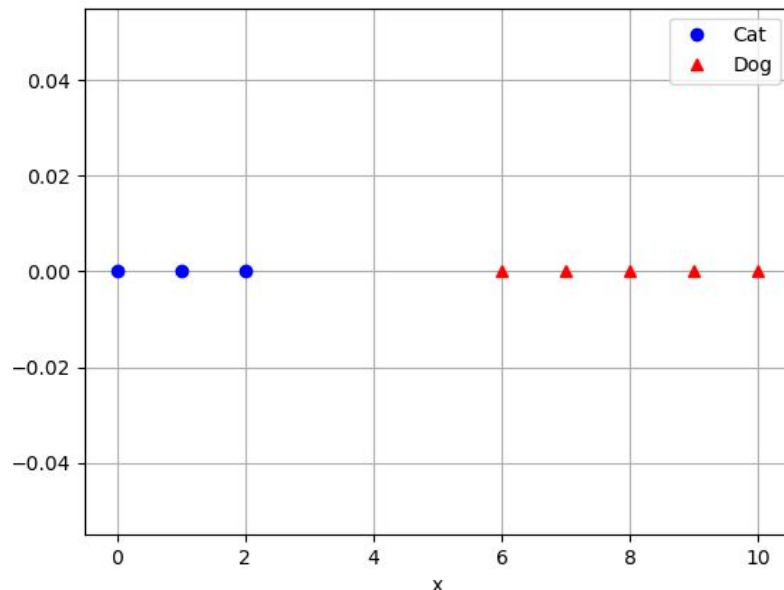
Classifiers

- There are many algorithm for classification:
 - **Logistic regression**, neural network, decision tree, random forest, SVM and KNN
 - Notice “Logistic regression is for classification” NOT regression!
- There are 2 popular output **styles** from algorithms
 - **Discrete** output (**hard**): Just output zero or one (e.g KNN)
 - **Probability** output (**soft**): continuous value in range [0 - 1], (e.g. Logistic regression)
 - The same in clustering: hard (k-means) or soft (Gaussian Mixtures)



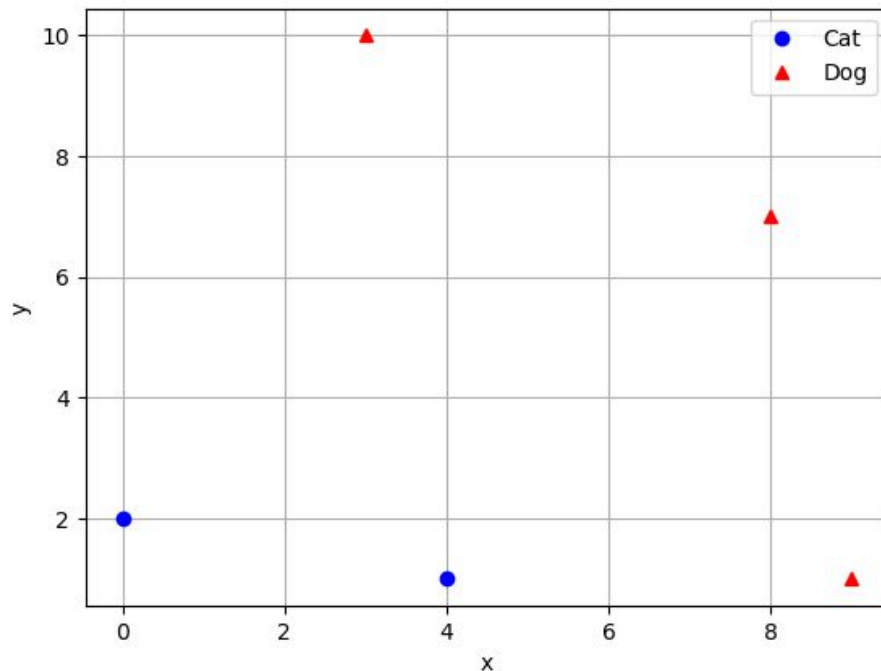
Line for Classification (1D input)

- Can you figure out how can we utilize a line that classifies data?
 - Observe, we draw here X only, not (X, Y) for 1D data
 - A point splits points before and after
 - We can find with a simple loop
- Let's move to input that is 2D



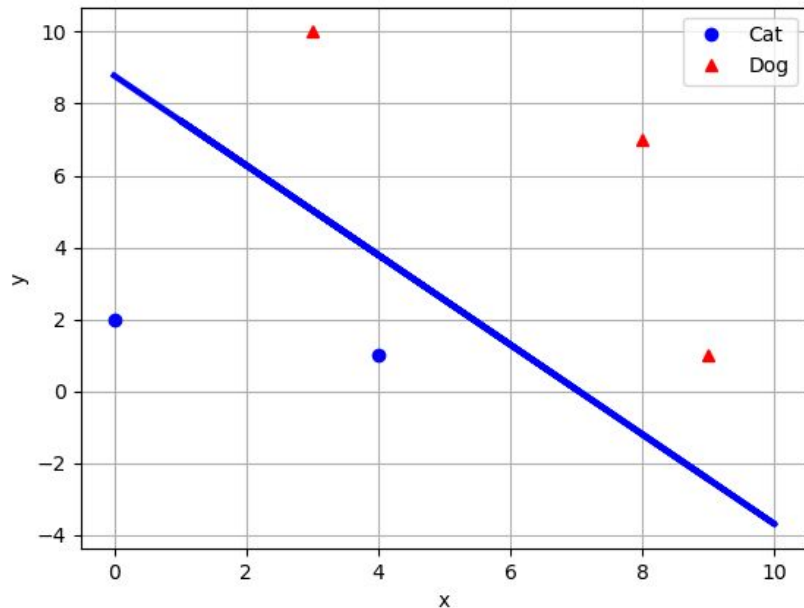
Line for Classification (2D input)

- Can you figure out how can we utilize a line that classifies data?



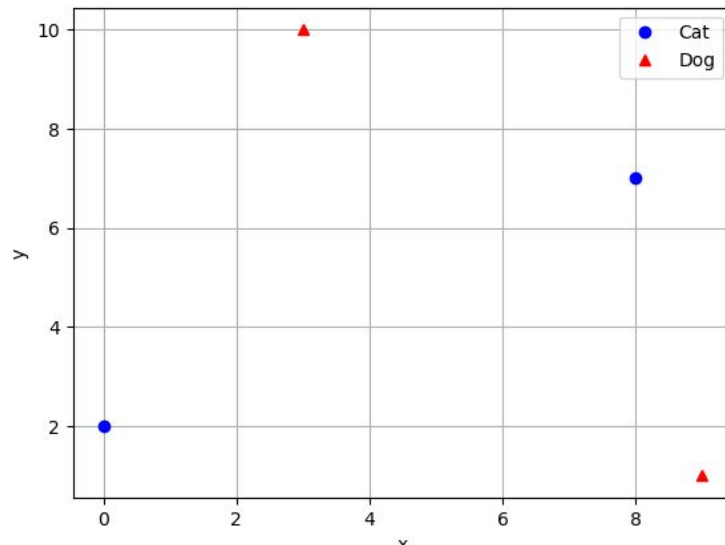
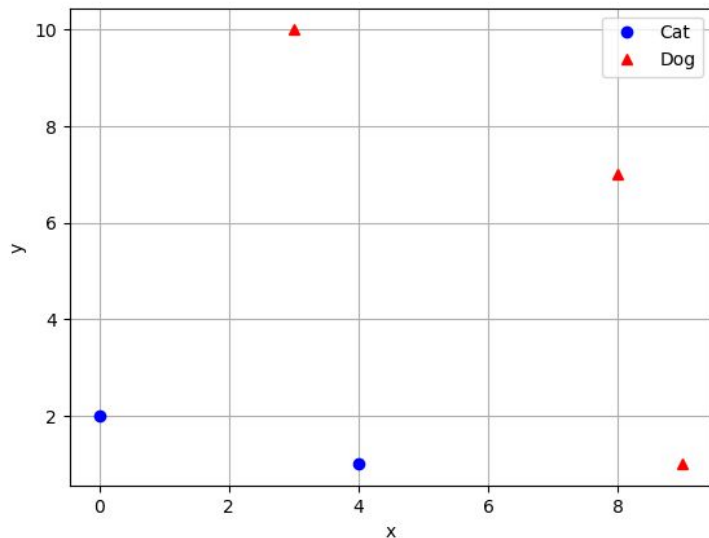
Line for Classification

- We simply use the line to **split** the points into 2 regions
 - Not fitting. Just split
- Points **above** the line are e.g. class #1
- Points **under** the line are e.g. class #0
- Observe, there are infinite number of lines to do so
 - **SVM** has an interesting way for finding this line
 - Large margin classifier



Linearly Separable Dataset

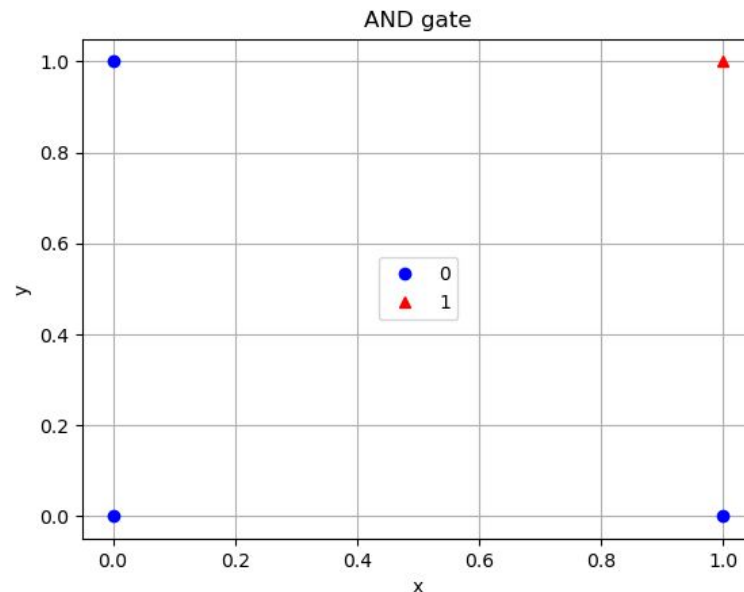
- **Linear Separability** implies the existence of a **hyperplane** separating the two classes (left example)
- There is no such line for the right image. We find a line that minimizes some cost function



Question!

- Assume we want to learn a classifier for **AND** gate
 - Binary input $(x, y) \Rightarrow$ single binary output
- Is this linearly separable data?

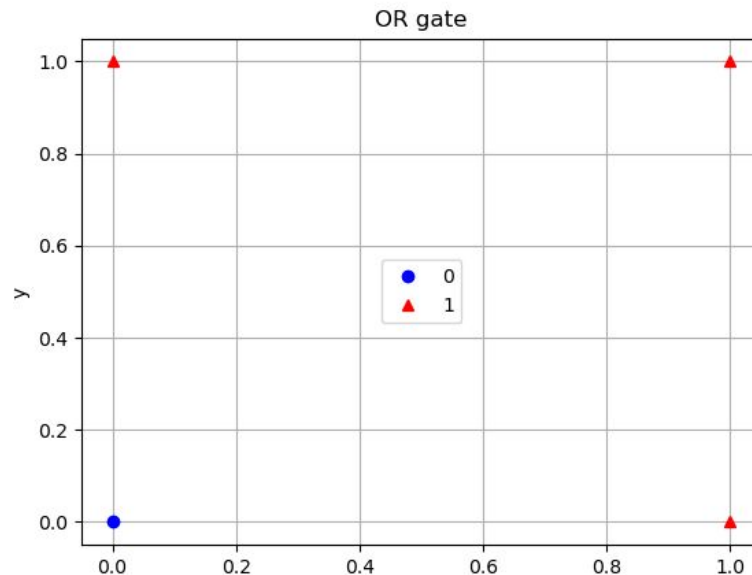
X	Y	$AND(X,Y)$	$OR(X,Y)$	$XOR(X,Y)$
0	0	0	0	0
0	1	0	1	1
1	0	0	1	1
1	1	1	1	0



Question!

- Assume we want to learn a classifier for **OR** gate
 - Binary input $(x, y) \Rightarrow$ single binary output
- Is this linearly separable data?

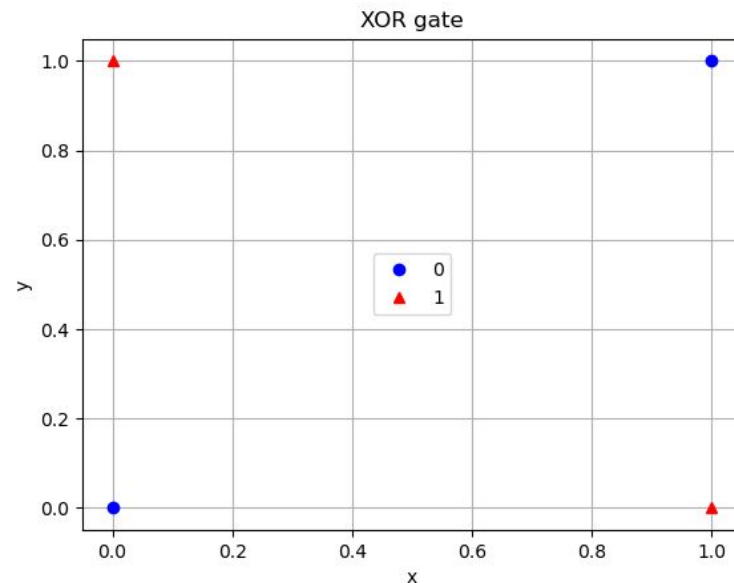
X	Y	$AND(X,Y)$	$OR(X,Y)$	$XOR(X,Y)$
0	0	0	0	0
0	1	0	1	1
1	0	0	1	1
1	1	1	1	0



Question!

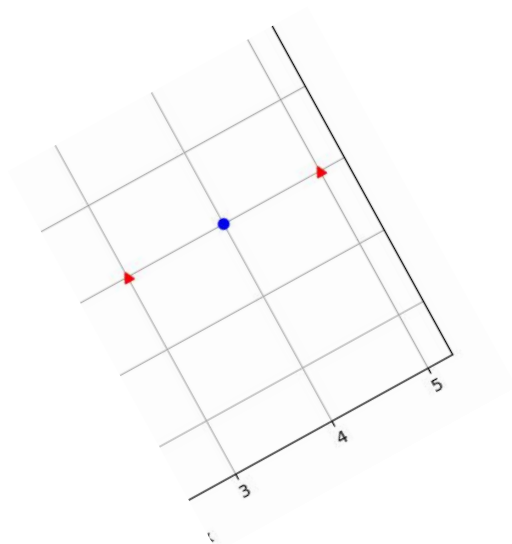
- Assume we want to learn a classifier for **XOR** gate
 - Binary input $(x, y) \Rightarrow$ single binary output
- Is this linearly separable data?

X	Y	$AND(X,Y)$	$OR(X,Y)$	$XOR(X,Y)$
0	0	0	0	0
0	1	0	1	1
1	0	0	1	1
1	1	1	1	0



Question!

- We see with even just 4 points from XOR, we can have non-separable data
- Is there a 3-points dataset that is non-separable?



“Acquire knowledge and impart it to the people.”

“Seek knowledge from the Cradle to the Grave.”

