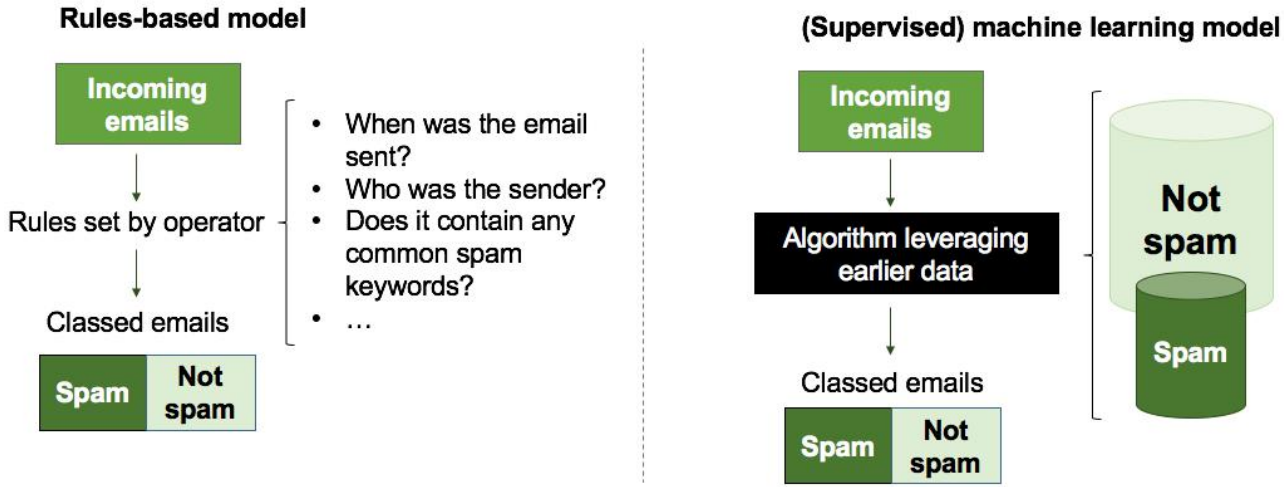
OBJECT DETECTION WITH MACHINE LEARNING

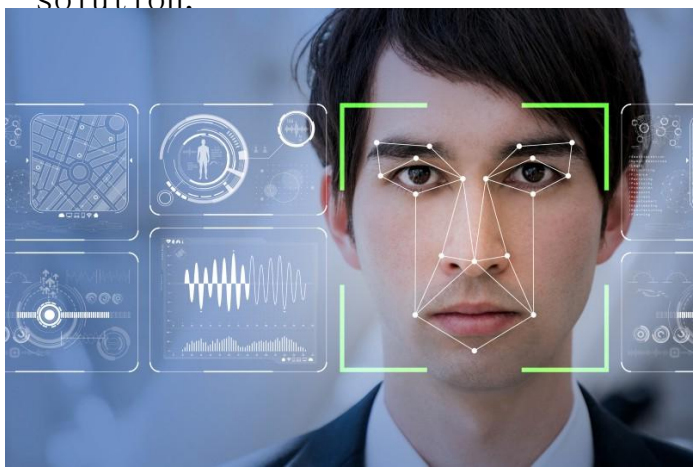
SPS
Swiss Post Solutions

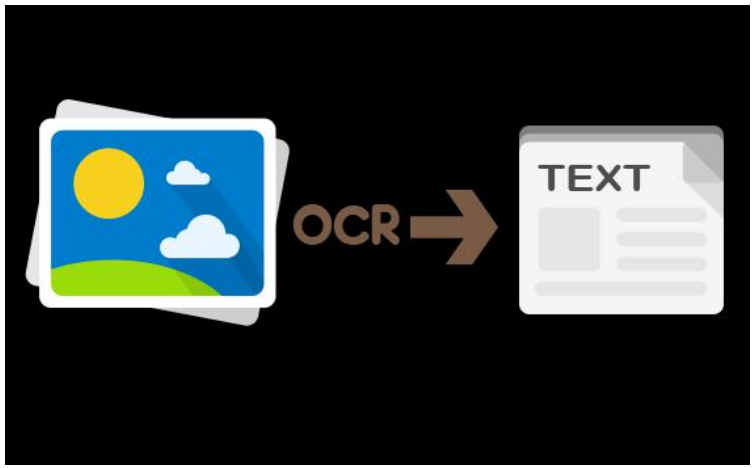
WHY WE NEED MACHINE LEARNING ?

- Resolve problems for which existing solutions require a lot of hand-tuning or long list of rules

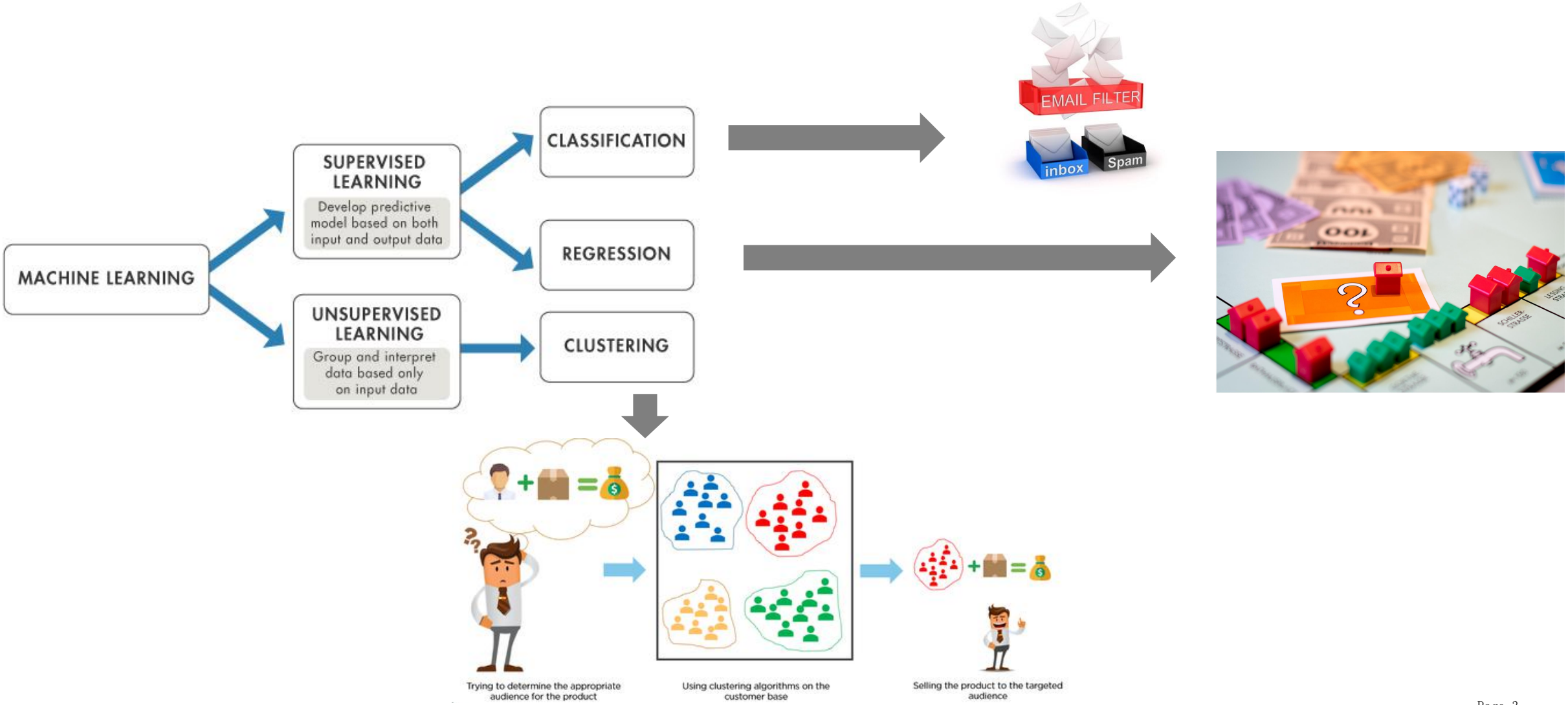


- Complex problems for which there is no good solution at all using a traditional approach. Machine Learning can help to find solution.

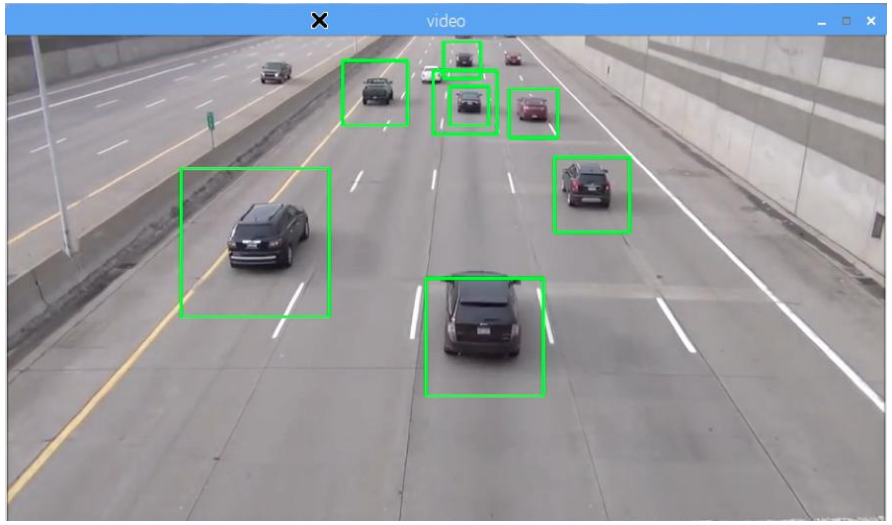




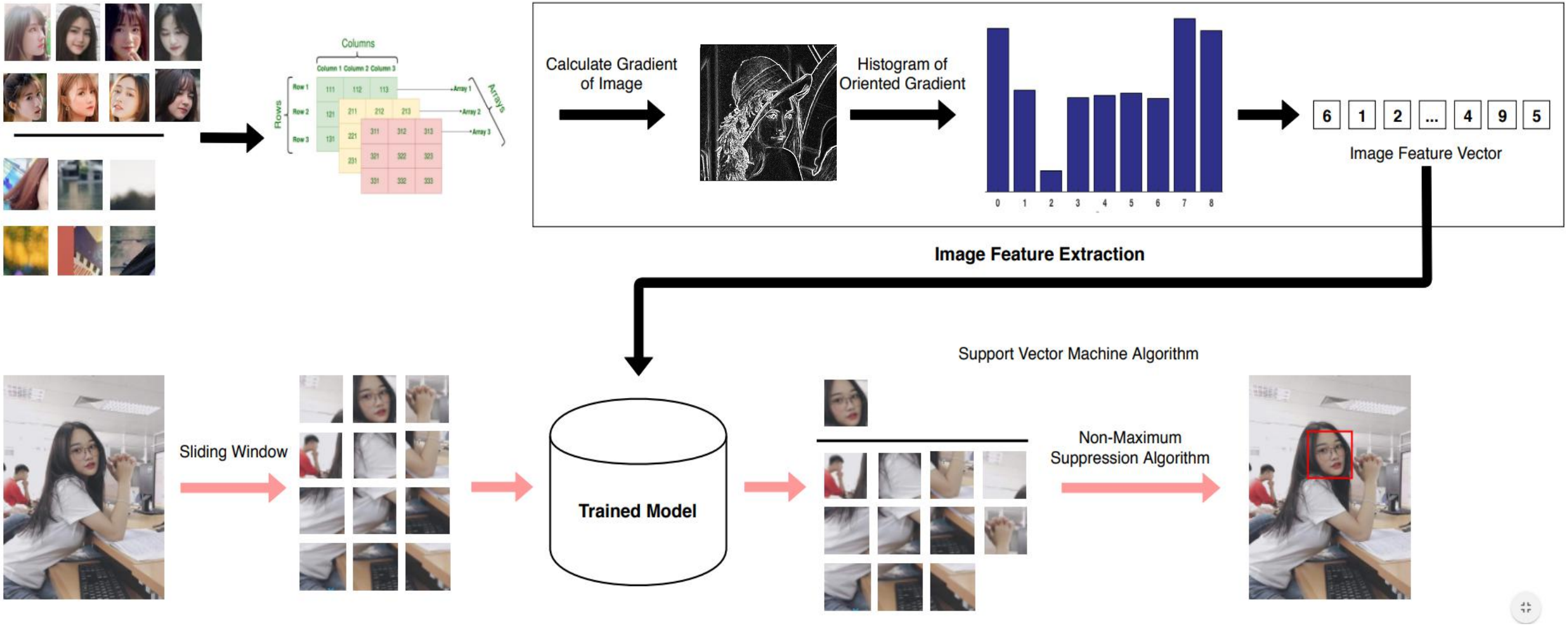
CATEGORIES OF MACHINE LEARNING



WHAT IS OBJECT DETECTION ?



HOW OBJECT DETECTOR WORKS ?



FEATURE EXTRACTION

HISTOGRAM OF ORIENTED GRADIENT

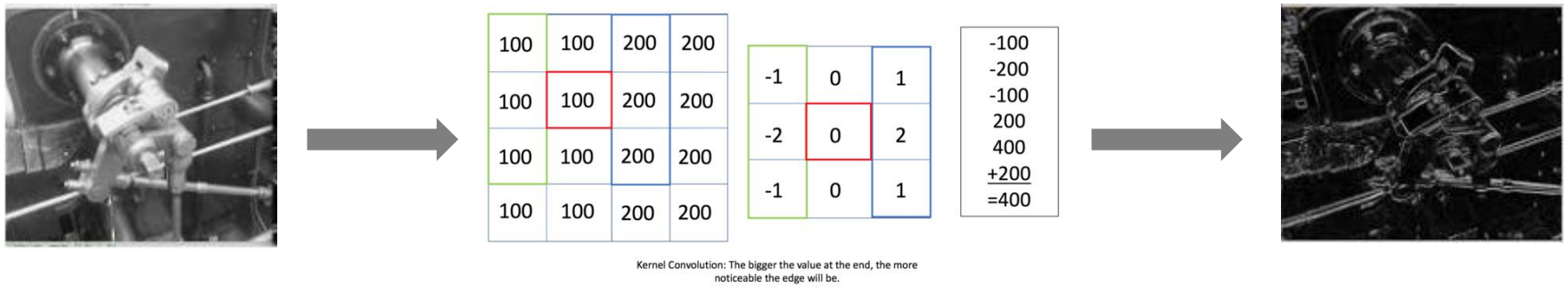
1. What is gradient ?

- Gradient of Image is a direction of change in the intensity or color in an image.

2. Edge Detection and Sobel

- Edge Detection is a image processing technique for finding the boundaries of objects within images. It works by detecting discontinuities in brightness.

- Sobel works by calculate gradient of each pixel of image. Therefore, we can see how abruptly or smoothly the image change at each pixel. Because of that, we can consider some pixels represent an edge.

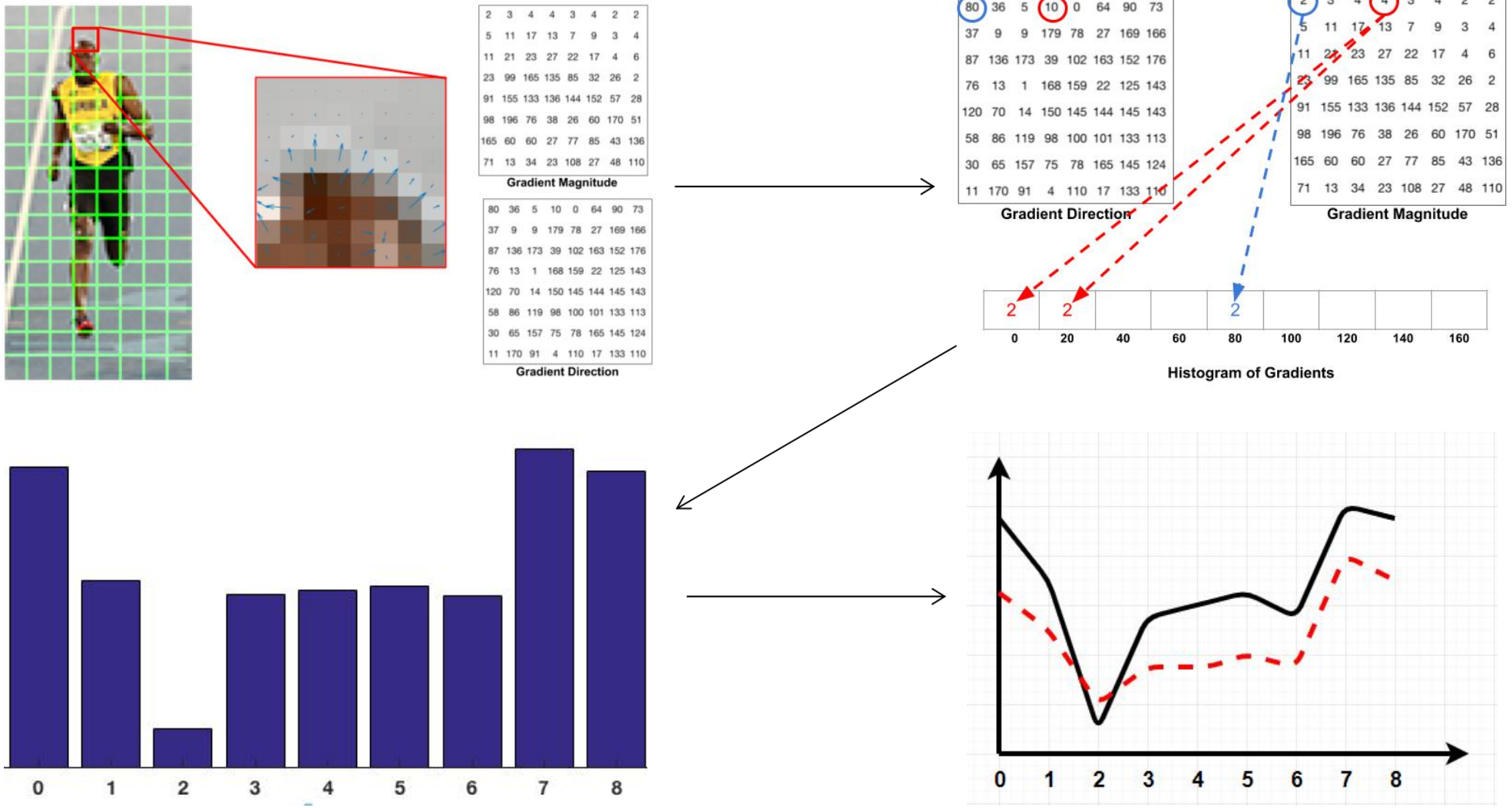


Kernel Convolution: The bigger the value at the end, the more noticeable the edge will be.

FEATURE EXTRACTION

HISTOGRAM OF ORIENTED GRADIENT

3. Histogram of Gradient



FEATURE EXTRACTION

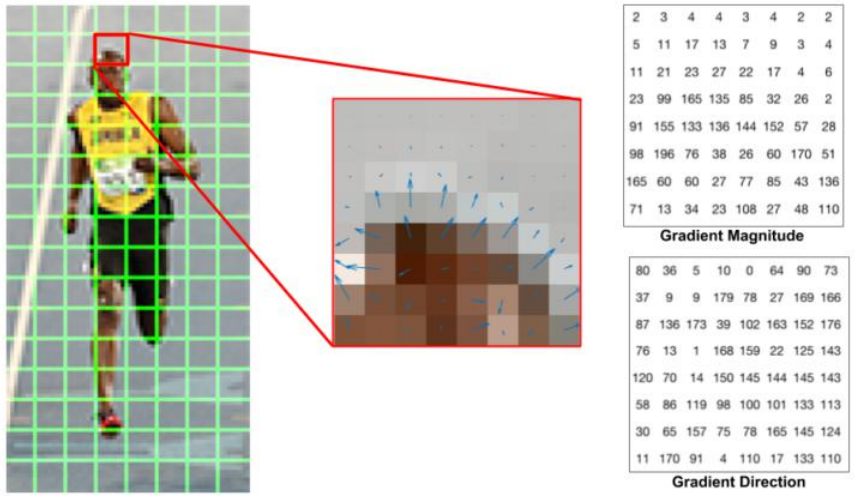
HISTOGRAM OF ORIENTED GRADIENT

4. Block Normalization

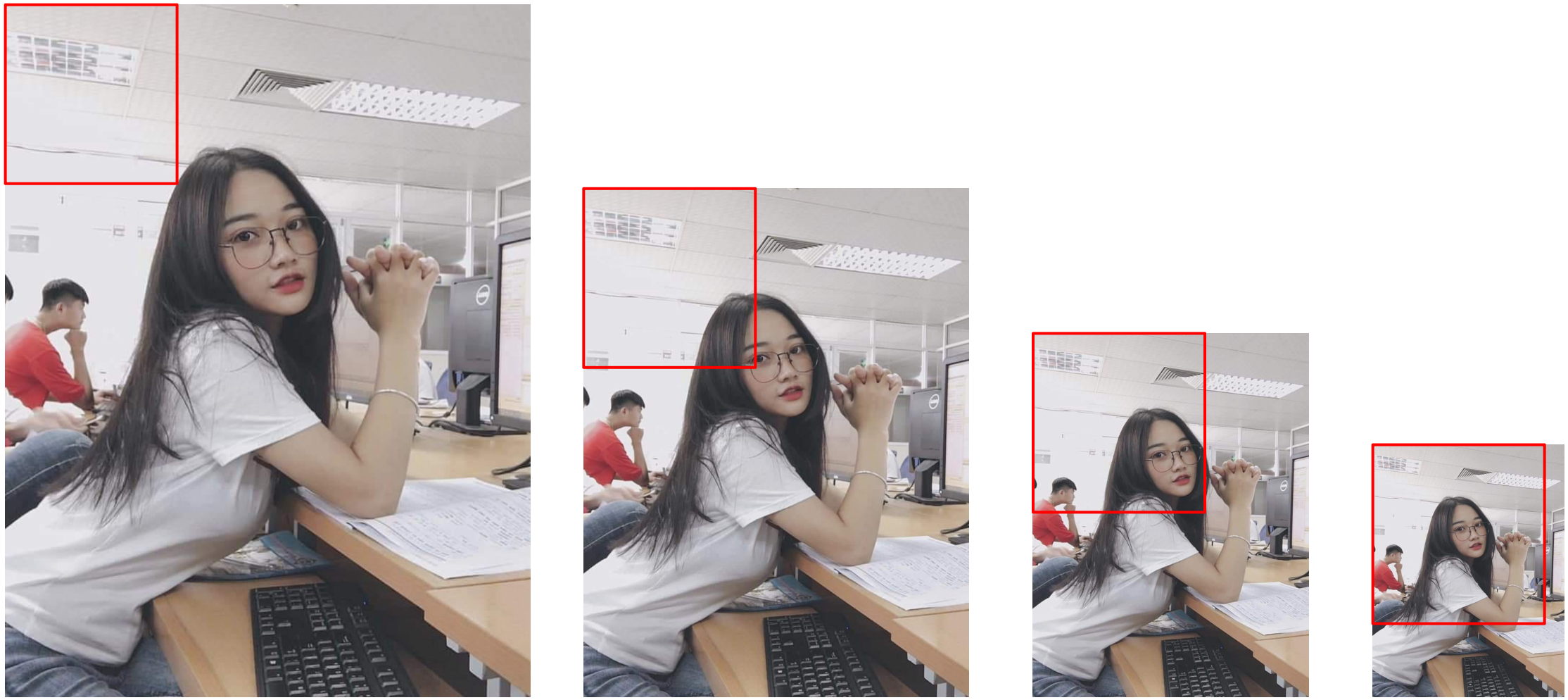
- Gradients of an image are sensitive to overall lighting.
- [128, 64, 32] is RGB color vector => length of vector = $\sqrt{128^2 + 64^2 + 32^2}$ = 146.64
- Dividing each element of this vector by 146.64 gives us a normalized vector [0.87, 0.43, 0.22].
- [256, 128, 64] = 2 * [128, 64, 32] after normalize, we still get [0.87, 0.43, 0.22]
- With this idea, we can concatenate 4 histograms = 4 * 9 bins. With this, we has a 36 dimensions vector

5. Calculate the HOG feature vector

- To calculate the final feature vector for the entire image patch, the 36×1 vectors are concatenated into one giant vector.

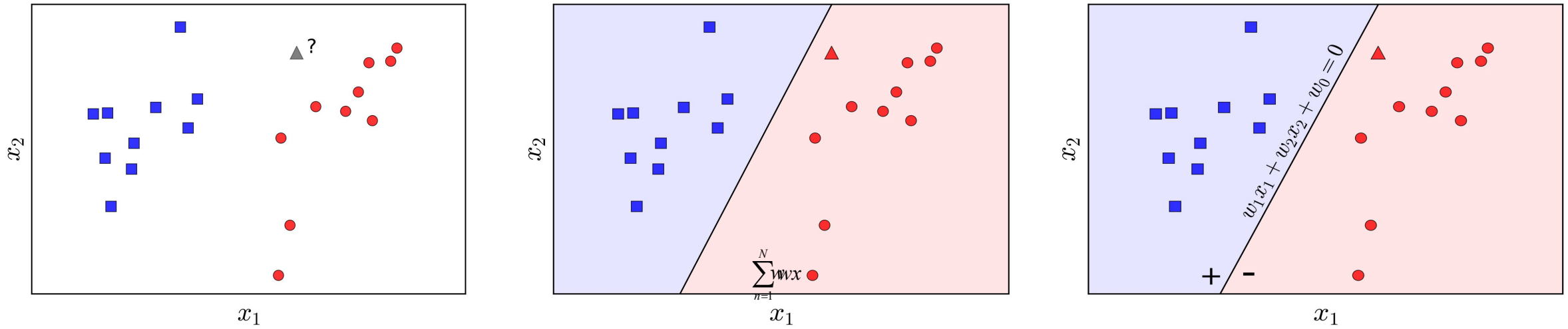


SLIDING WINDOW



CLASSIFICATION

PERCEPTRON ALGORITHISM



f(x) = w_1x_1 + w_2x_2 + ... + w_nx_n là phương trình của siêu mặt phẳng phân chia 2 class xanh và đỏ.

Với w được generate random lúc đầu, trong mỗi lần lặp, ta thế x vào phương trình trên để tính được f(x)

Khi đã tính được f(x), ta dùng cost function để xác định độ chính xác của w

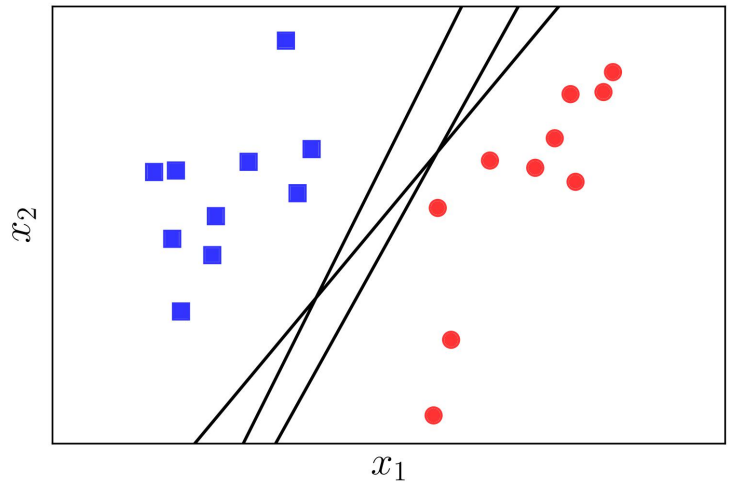
Cost function được dùng là $$RMSE = \frac{1}{2}\sqrt{(f(x) - y)^2}$$

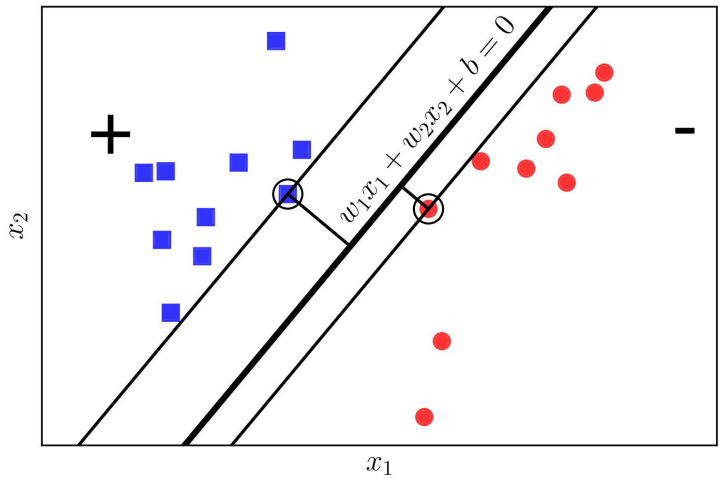
Ta thế f(x) đã tính được và y (label) vào công thức RMSE, ta sẽ biết được w cần phải tăng hay giảm để phù hợp với bộ dữ liệu

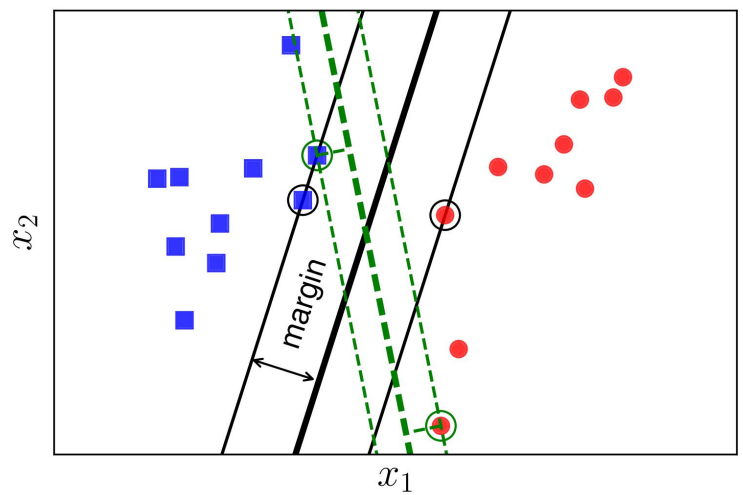
Khi nghiệm của hàm RMSE là thấp hơn mức chúng ta cần, thì xem như model đã được train thành công.

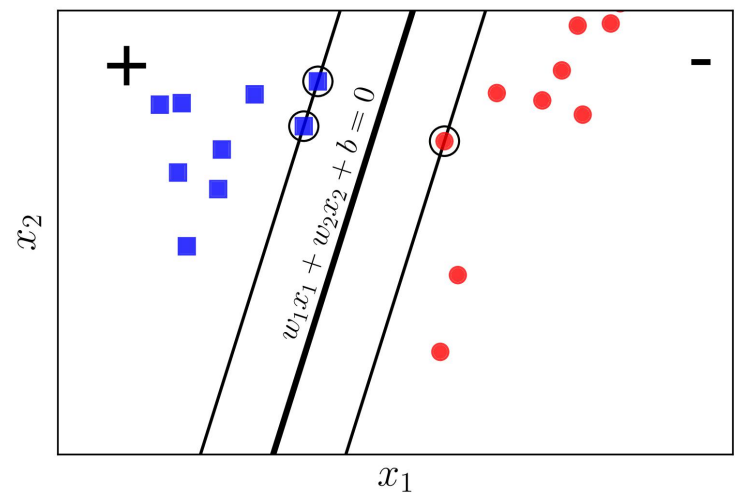
CLASSIFICATION

SUPPORT VECTOR MACHINE









Q & A

SPS
Swiss Post Solutions