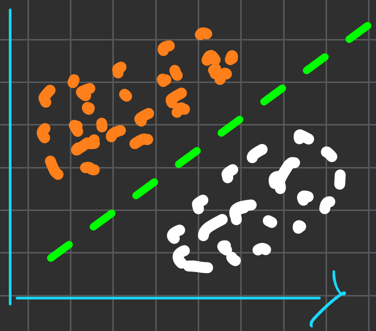
Support Vector Machines →

Linear/Logistic Regression → To come up with the best hyperplane that separates both class of points.

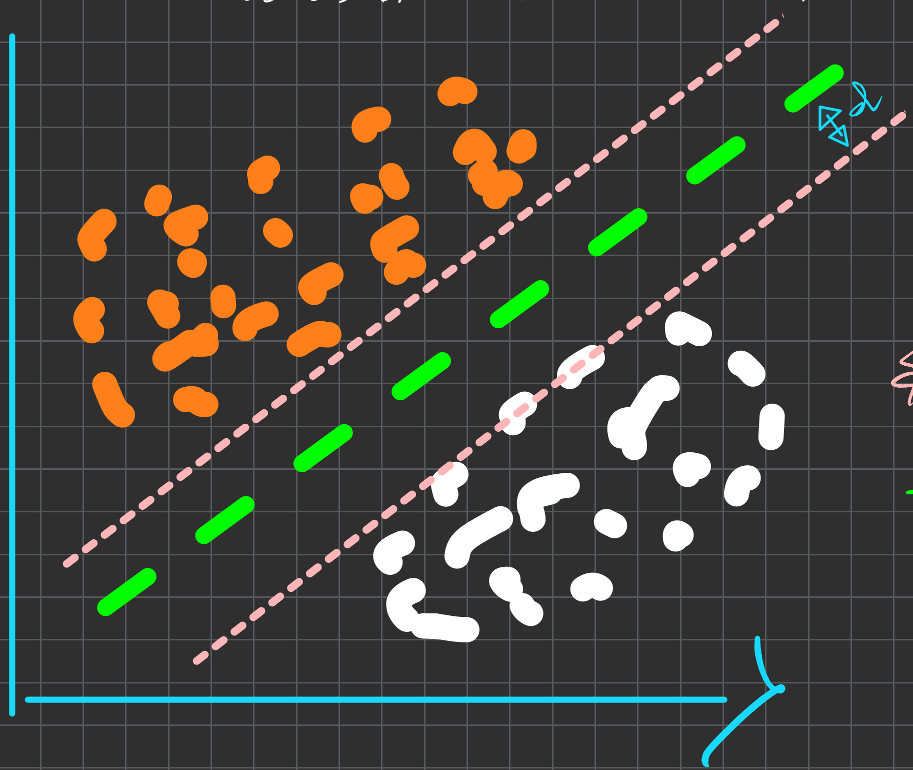


→ Linear Regression.

How to make it even better?

Ans → For more security & assurance

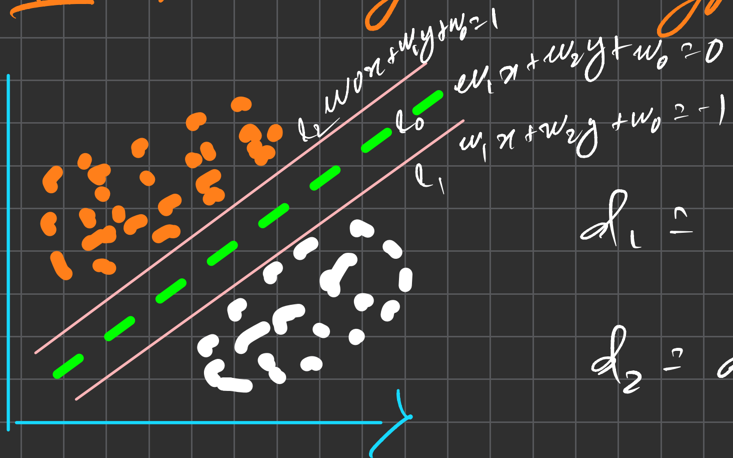
Mark two more lines on the boundary of each class that is nearest to the decision boundary.



- now each point can't go beyond its intended boundary.

SVM is
- A really good classifier.
- Robust to outliers.

Q → How to get this hyperplane with two support vectors?



Distance b/w two lines →

$$d_1 = d(l_0, l_1) = \frac{w_0 - (w_0+1)}{|w|}$$

$$d_2 = d(l_0, l_2) = \frac{w_0 - (w_0-1)}{|w|}$$

$$\text{margin} = d_1 + d_2 = 2/|w|$$

$$\max(\text{margin} = 2/|w|) = \min(|w|/2)$$

$$|w| = \sqrt{w_1^2 + w_2^2 + \dots w_n^2}$$

Constraints → 1) The classes are binary & are linearly separable.

$\min(|w|/2)$ → hard margin SVM.

- no misclassifications allowed.
- error = 0.

$|w|$ → L_2 error (basically)

Soft Margin & Hinge Loss →

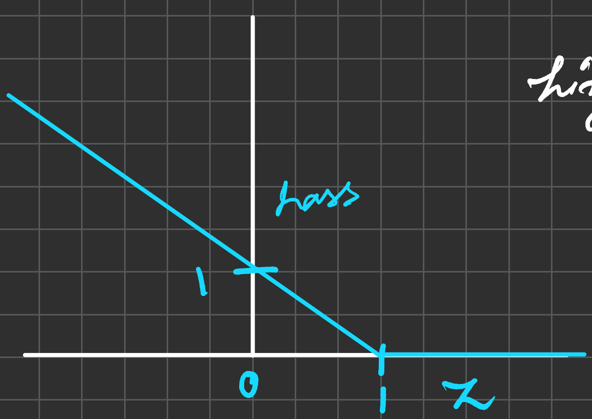
↳ Allows some misclassification.
↳ error → misclassified point.

Signed distance → $y(w^T x + w_0) = z$

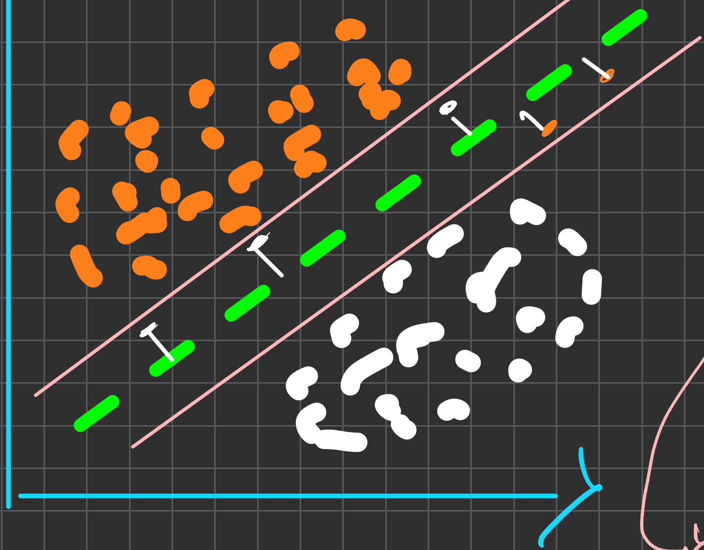
y	$w^T x + w_0$	$y(w^T x + w_0)$
-1	-1	1
-1	1	-1
1	-1	-1
1	1	1

Hinge Loss

1 → similar classifications.
-1 → misclassifications.



$$\text{hinge}(z) = \max(0, 1 - z)$$



Svm with hinge loss

=

Soft Svm.

$$\min\left(\|w\|/2 + c \sum \max(0, 1 - z_i)\right)$$

$c \propto$ overfit

↳ hyper-Parameter.

1) Introduces some error.
   - avoids overfitting.

2) minimizes this error.
   - avoids underfitting.

error $\propto$ underfit

new eqn of line becomes → $w_1 x_i + w_2 y_i + w_0 \geq 1 - e_i$

minimize e_i.

Find the combination with maximum margin.