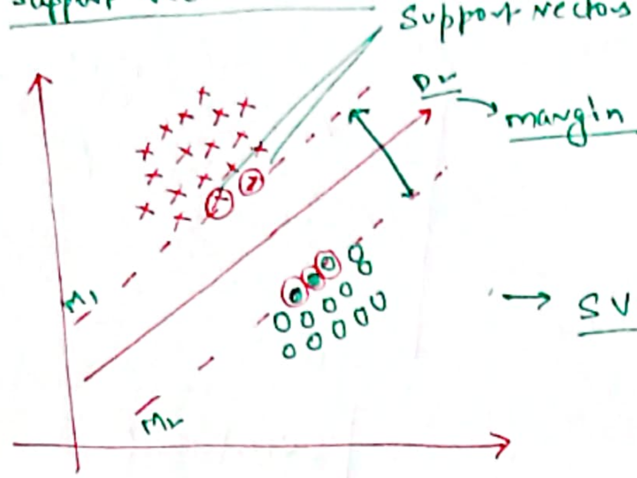
Designing Data Intensive Application — Martin Kleppmann

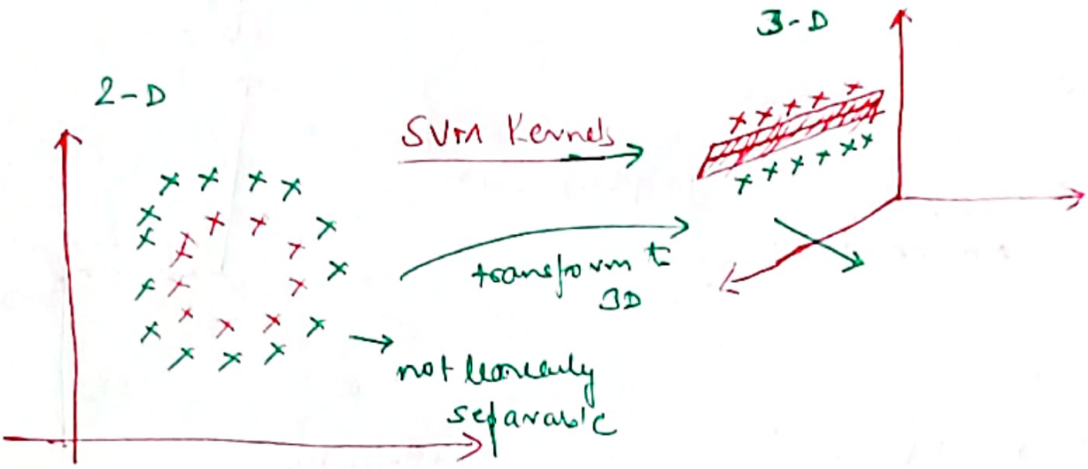
System design — Hussain Naseer

Support-Vector Machine

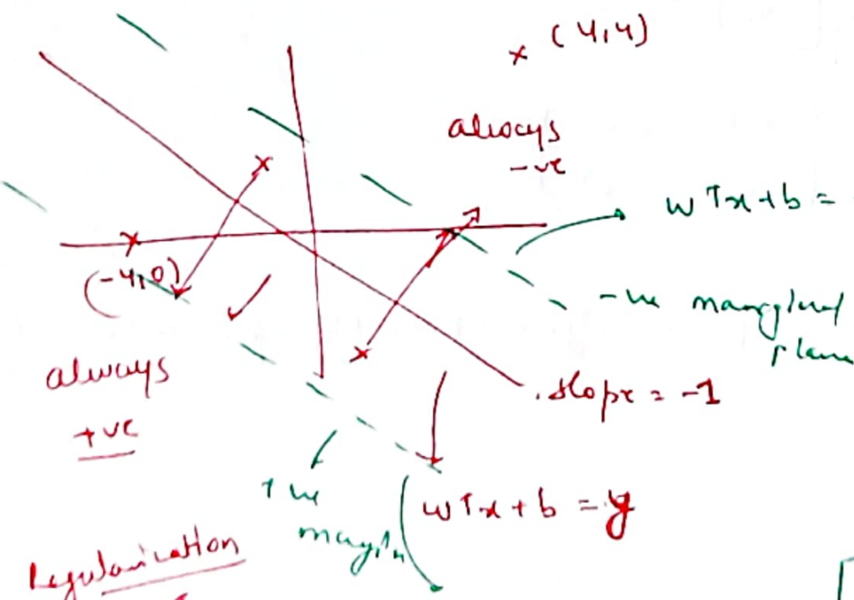


- we should maximise D_2 ↑↑.
- Support vectors are the one that touches the M_1 & M_2.

→ SVM classifier.



SVM maths Intuition



$w^Tx+b=-1$

$w^Tx+b=y$

$w^Tx+b=1$

$w^T(x_2-x_1)=2$

$\frac{w^T}{\|w\|}(x_2-x_1)=\frac{2}{\|w\|}$

↑ Maximise this

or. $\min \frac{\|w\|}{2}$

now can use GD SGD & other minimisation

$y_i = \begin{cases} +1 & w^Tx+b \geq 1 \\ -1 & w^Tx+b \leq -1 \end{cases}$

$\boxed{y_i * (w^Tx_i+b_i) \geq 1}$

if this not the case this is a mis classification.

Regularisation

$\min \frac{\|w\|}{2} + C_i \sum_{i=1}^{n} \xi_i$ — value of the error.

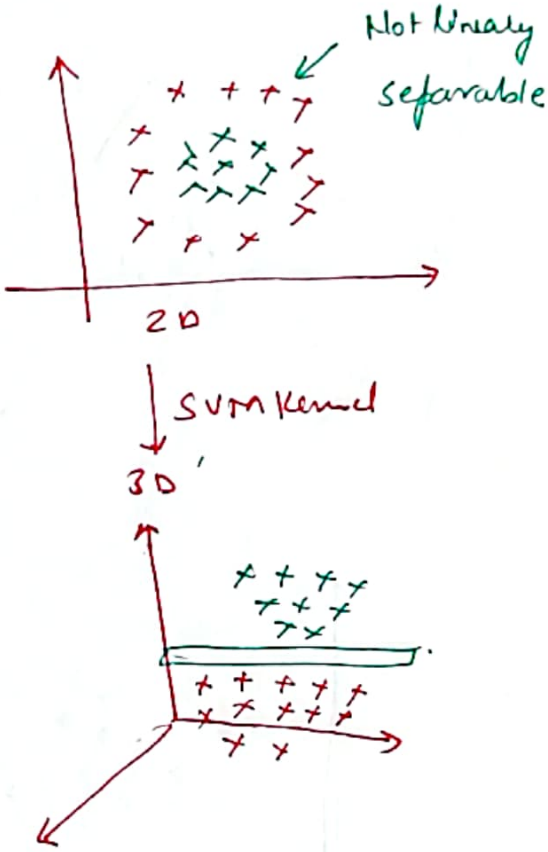
How many error my model can

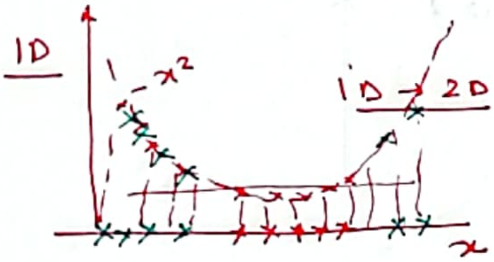
SVM Kernels

[Soft margin - C > 1
[hard margin

- Polynomial Kernels
- RBF Kernels
- Sigmoid Kernels

Some kind of transformation from lower dimension to higher dimensions.





$y = f(x) = x^2$

Polynomial Kernel:

$$f(x_1, x_2) = (x_1^T x_2 + 1)^d \rightarrow \text{dimension}$$

x_1	x_2	y_0	x_1^2	x_2^2	$x_1 x_2$

$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} [x_1 \; x_2]$

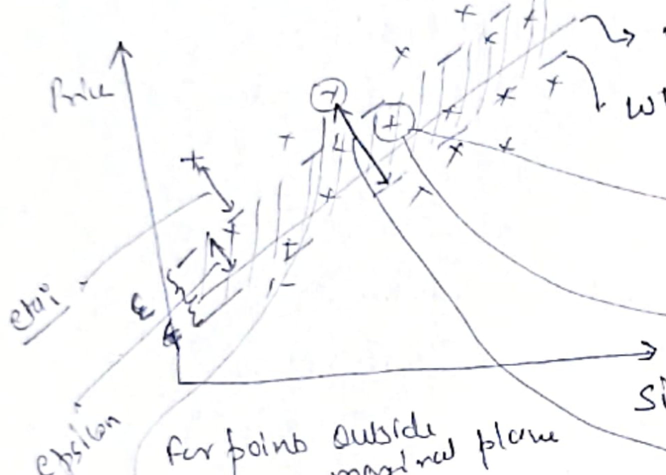
$\begin{bmatrix} x_1^2 & x_1 x_2 \\ x_1 x_2 & x_2^2 \end{bmatrix}$

what to use polynomial, rbf or sigmoid? → hyperparameter tuning.

KNN limitations :-

1. Cannot be used efficiently with huge data sets.
2. Outliers impact this algorithm hugely. sensitive to outliers
3. Sensitive to missing values.

SVM Regression



→ $w^Tx + b + \epsilon = 0$

→ $w^Tx + b = 0$

→ $w^Tx + b - \epsilon = 0$

1. fit a line like a linear regression to the data
2. find a marginal plane

→ for points inside marginal plane

cost f^n = $\underset{(w,b)}{\text{minimize}} \frac{||w||}{2}$

constraint

→ $|y_i - \underbrace{w^Tx_i}_{\hat{y}}| \le \epsilon$ (epsilon)

It is a good pred^n

for points outside marginal plane

$$\min_{(w,b)} \frac{||w||}{2} + C \sum_{i=1}^{n} |\xi_i|$$

↳ etta.

↑ how many points we can adjust outside the marginal plane

distance b/w the point & marginal plain

$\underbrace{C, \xi}$

↳ hyperparameters / hinge loss.

constraint

$$|y_i - w^Tx_i| \le \epsilon + |\xi_i \text{ (etai)}|$$