MACHINE LEARNING BASICS

NO-FREE LUNCH THEOREM

↳ No one model works best for every problem

So always try different models and different training methods i.e train linear regression w normal equations or gradient descent

PARAMETRIC VS NON-PARAMETRIC MODELS

parametric - model has a fixed # of parameters
- computationally faster
- makes strong assumptions of the data

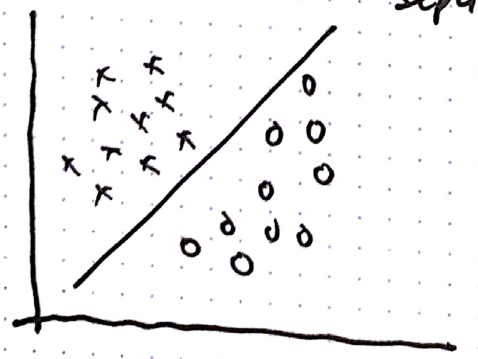
NON-PARAMETRIC - model does not make strong assumptions about the data
- sometimes results in better performance
- needs a lot of data
- slower to train

REGULARIZATION

- REGULARIZATION is making the weights as small as possible

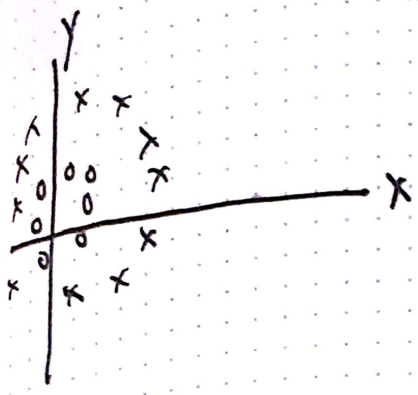
SUPPORT VECTOR MACHINES

SVMs separate classes of data, optimal hyperplane that separates the data



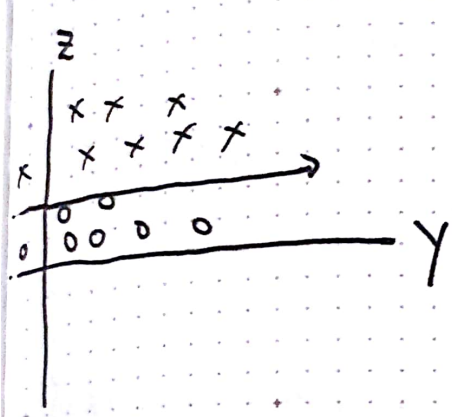
KERNELS

what line separates this data?



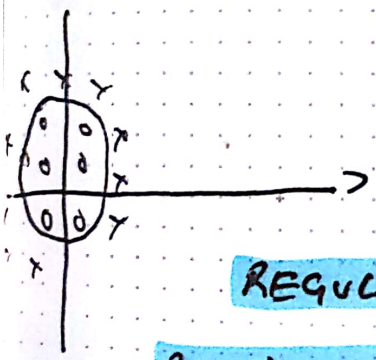
Apply transformations

1. Define a new axis 'z'
   $z = x^2 + y^2$



← plot data against new data and now a separating hyperplane can be found

← re-plot the original line onto the original axis to separate

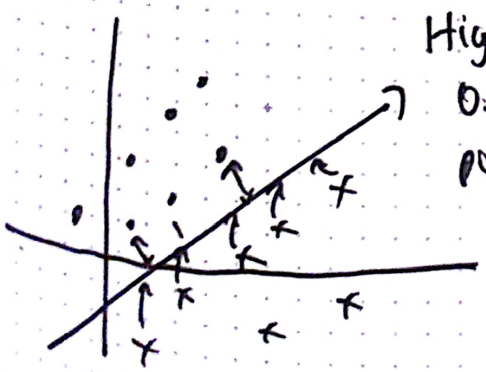


REGULARIZATION PARAMETERS

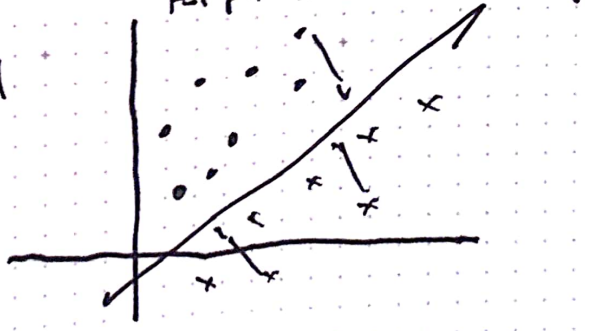
Regularization parameter ('C' in SKlearn) is how much you don't want to misclassify a training example

Gamma

This parameter is for how much of an impact a point will have on the decision hyperplane



High Gamma
Only nearby points are considered



Low Gamma
Far points are also considered