Learning with Kernels

Matthieu Bulté

2017-06-06

What are we talking about

Classification

We have data We have labels We want to learn how to label new data (decision function)

Decision function / Boundary (Plots of many different decision boundaries)

ML is about finding the right transformations for making it trivial to

Make the classification task trivial (linear)

Margin max

(or how to pick the best trivial decision function)

Demo 1 - moving linear decision boundary

support vectors

H only depends on (w, b)

Non linear projection: idea

Demo 2 - projection making linear trivial

Kernel trick

$k(x, y) = \langle \phi(x), \phi(y) \rangle$

Demo 3 - some kernels

the good

- E[P(error)] <= ...
- performance on small data sets
- domain knowledge

the bad

- parameters tuning
- training time
- domain knowledge (what do I know when talking about very complex problems?)

conclusion

?