Monday, October 4, 2021 7:41 PM

SVM

Classification

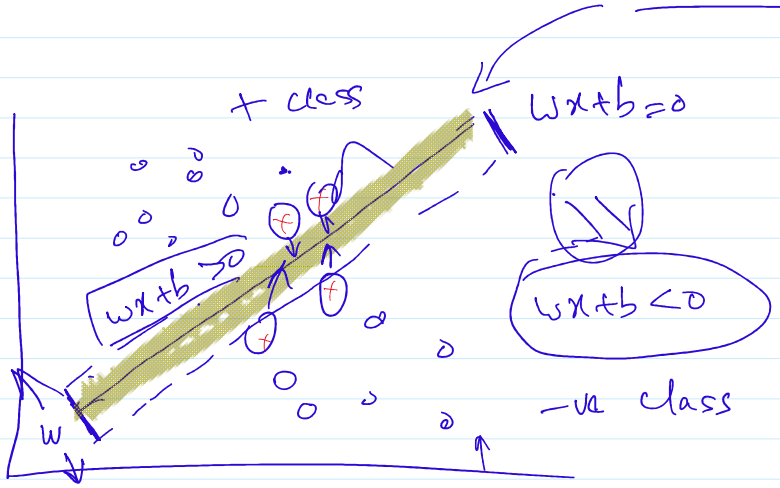
Linear
- LinearReg (Reg)
- Logistic (Class)
- Support Vector Machine

Tree
- Decision Tree
- Random Forest
- Ada boost

Data

Assumption

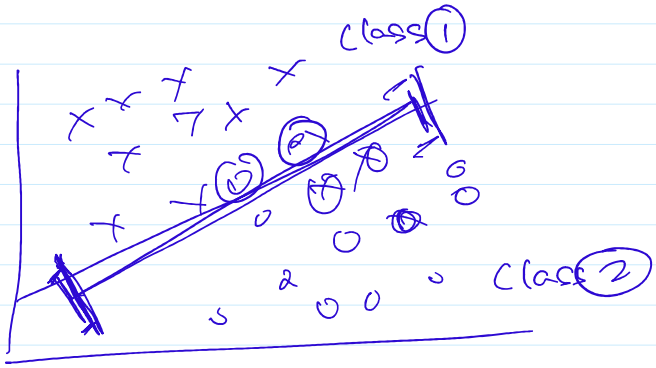
1) Data is linearly separable



Margin of a linear classifier is the width of hard around any classifier without touching examples

When is your model probable to make more mistakes?

The maximum margin linear classifier with max margin width



LSVM

Hard margin SVM

But, why max margin is better?

It maximizes the margin of hyperplane, reduces the generalization error (the most).

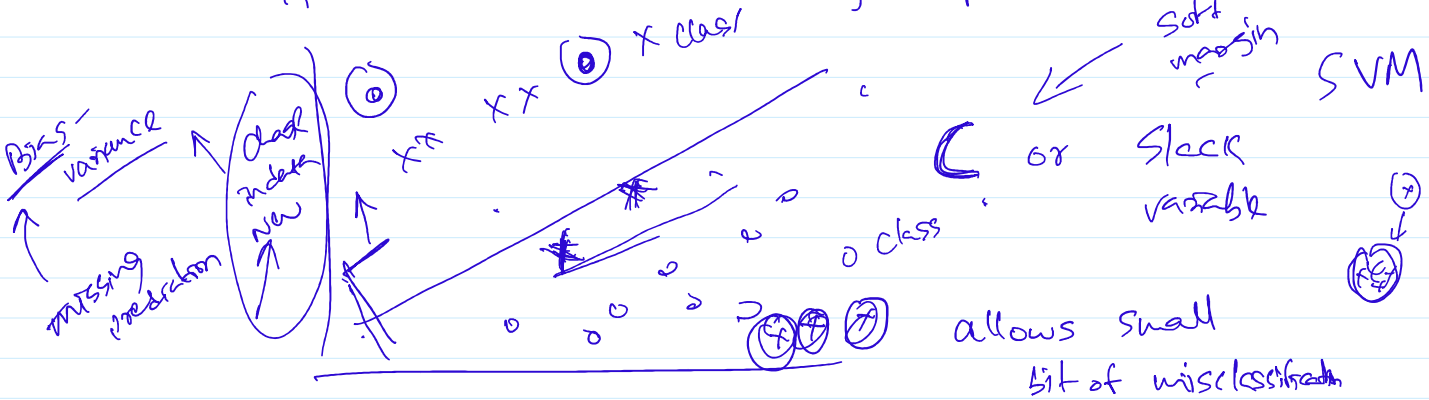
→ i.e. New data doesn't affect the model

Vapnik

What happens if data is not linearly separated?

Vape...

What happens if data is not linearly separated?



if C is large, the # of misclass ↑, generalization increases ↑ but has high bias ↑

if C is small, the classifier allows small misclassification, model may n't (over) generalize & has high variance ↑

if C=0 ?

C ranges 0.1 — 1000 ?

sine(x) cosine(x)

Gamma parameter of (RBF) Radial function → K(d) → N(d)

Basis

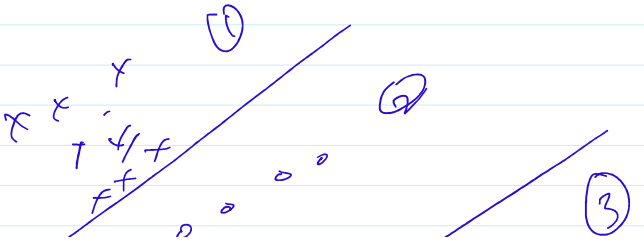
low gamma indicates large similarity radius → High bias → Low variance

(more pts grouped together)

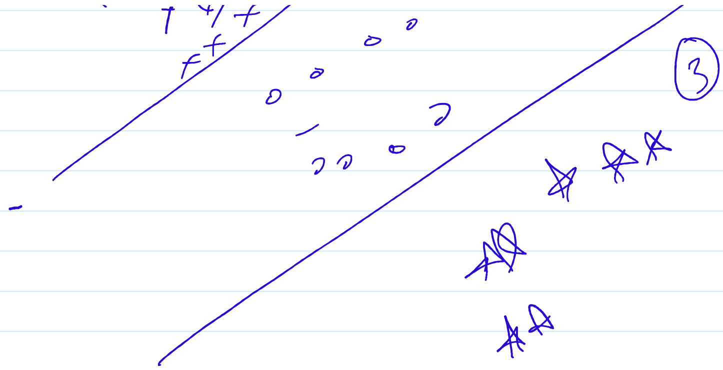
high gamma => only close points are grouped

Low bias High Var

overfitting



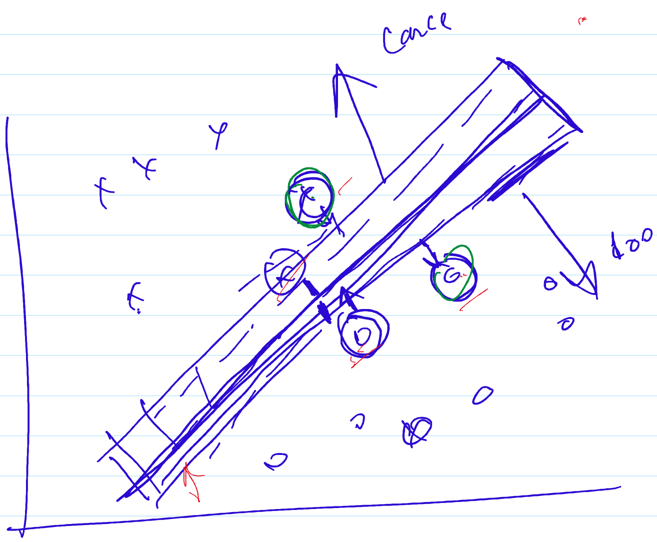
Low



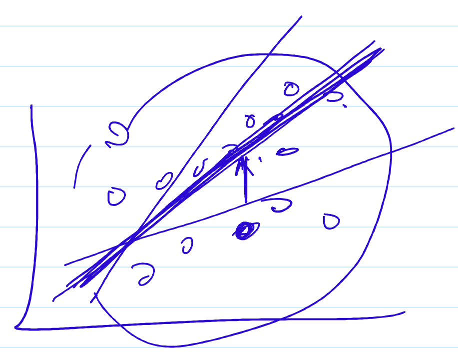
(3) Low gamma

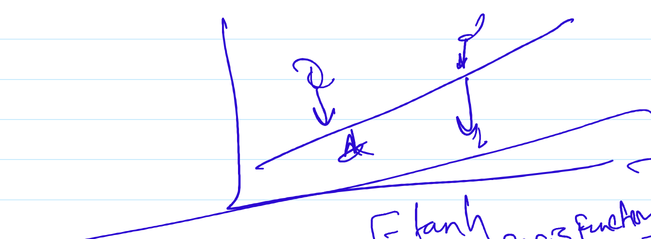
large gama

0.0001 < gamma < 10

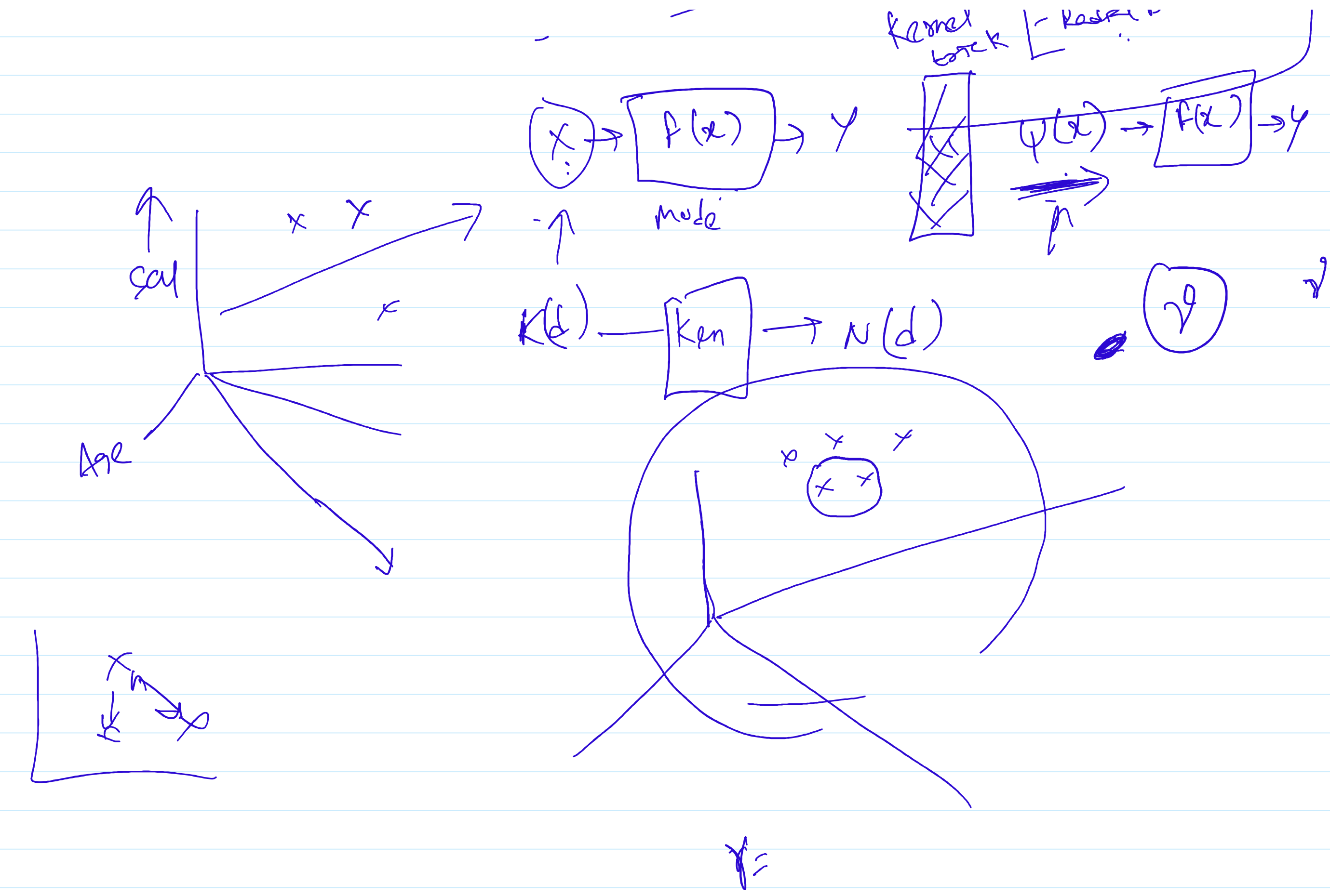


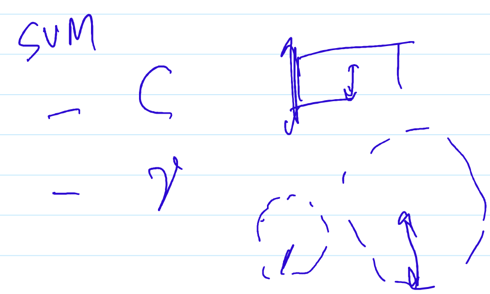
Support Vector





Kernel trick
- tanh
- Radial Basis function


Kernel trick
f(x)
Age


SVM