Lecture 10 (Nov 16, 2020)

Assignment 10

Nov 23, 30

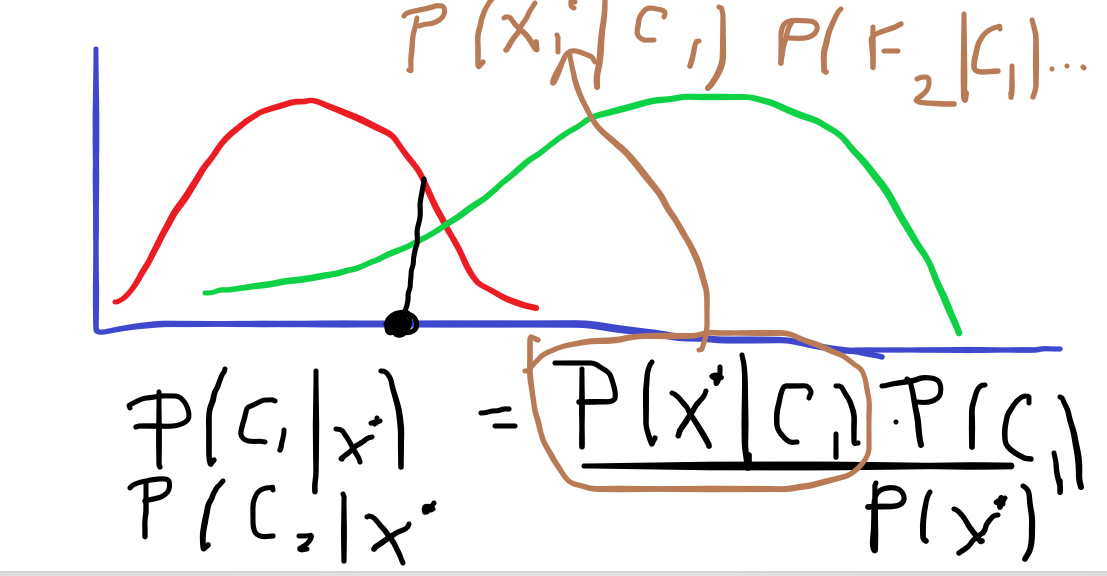
Project Presentations: Dec 7 (Monday)

Final exam: Dec 12-13 (Sat, Sun)

Code due by Dec 10, 2020

powerpoint presentation on Dec 7

* 1-2 page description how to run
* code
* - video recording about your project



For ensemble learning:

weak learners must classify with accuracy > 50%

suppose you have 3 learners

p1 probability of right answer from 1

p2 probability of right answer from 2

p3 prob of right answer from 3

prob all 3: p1\*p2\*p3

what is the probability that at least one classifier is correct:

p1 + P2 + p3

prob that all are incorrect:

(1-p1)\*(1-p2)\*(1-p3)

probability that at least one is correct is

1 –(1-p1)\*(1-p2)(1-p3)

suppose you are given n classifiers. Suppose each one is correct with probability p

Q: what us the probability that exactly k are correct?

A:

P(k successes) = C(n, k)p\*\*k \* (1-p)\*\*(n-k)

SUPPORT VECTOR MACHINES

Conceptually:

Original Points -------------🡪 Higher dimensional space (using kernels)

In the new space, construct the separating hyperplace

Each instance x\* in the original space is mapped into this higher dimensional space and classified

Homework Assignment

Part 1 : svm prediction of labels

SVM ONLY!!!!!

for Gaussian SVM you specify

kernel = “rbf”

(radial basis function)

Part #2 Compare classifiers that we learned when applied to your stock

|  |  |  |
| --- | --- | --- |
| Classifier | Accuracy | Portfolio Value for Year 2 |
| Buy and Hold |  |  |
| kNN with best k |  |  |
| logistic |  |  |
| Naïve bayesian |  |  |
| Linear models (best degree) |  |  |
| Decision tree |  |  |
| Random Forest |  |  |
| Linear discriminant |  |  |
| Quadratic discriminant |  |  |
| SVM (linear) |  |  |
| Ada Boost |  |  |
|  |  |  |

put the numbers in this table and round them

68.9134

68.9137

* round classifier accuracy to integers
* answer couple of questions:

what are the best 3 classifiers (by accuracy)

what are the worst 3 classifiers by accuracy