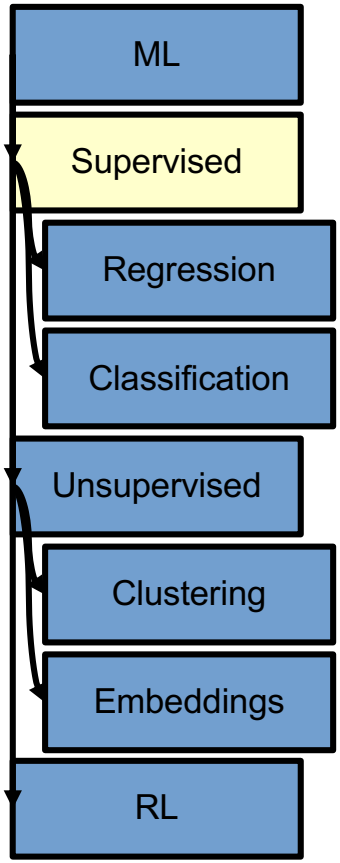
CS181: Introduction to Machine Learning

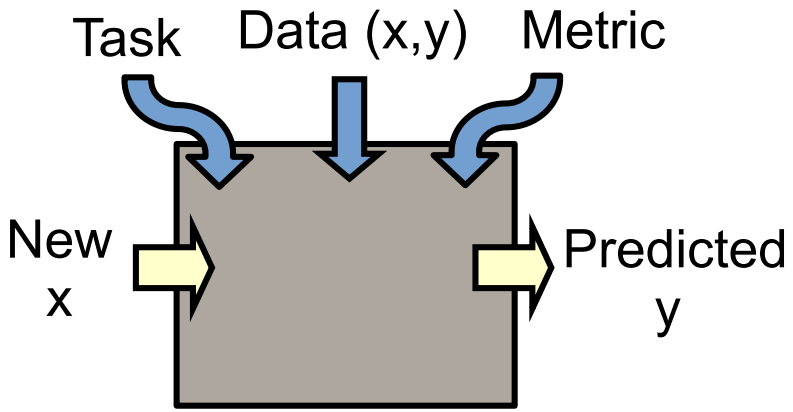
Lecture 5 (Probabilistic classification)

Spring 2021

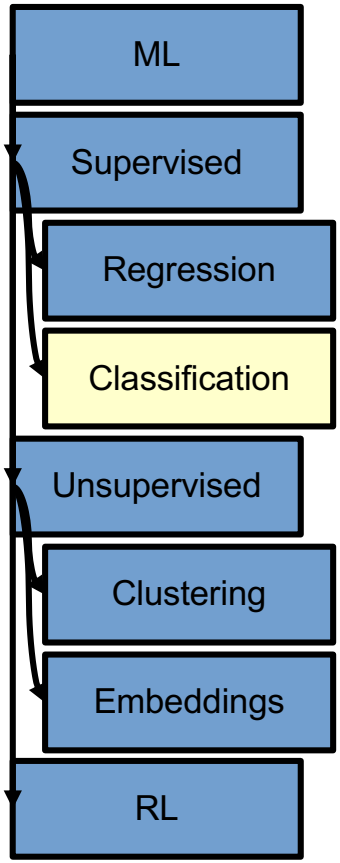
Finale Doshi-Velez and David C. Parkes
Harvard Computer Science

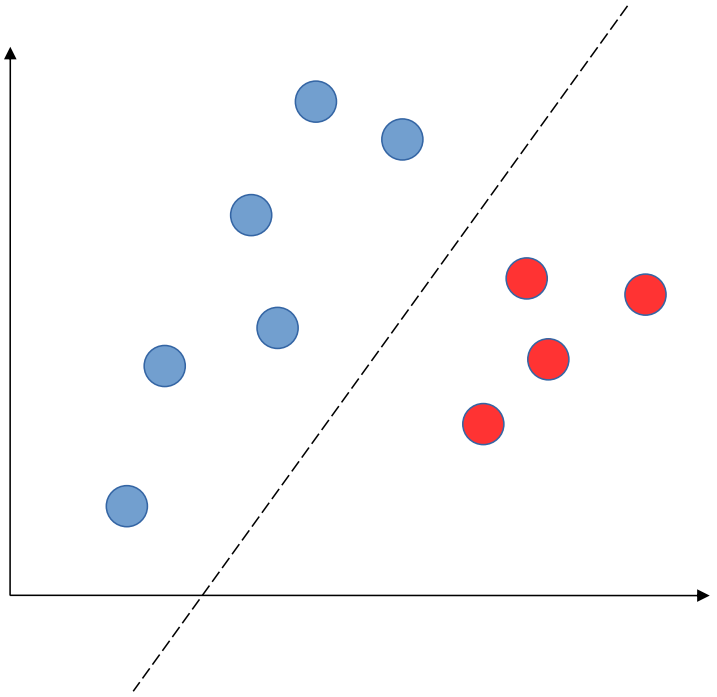
Machine Learning Taxonomy

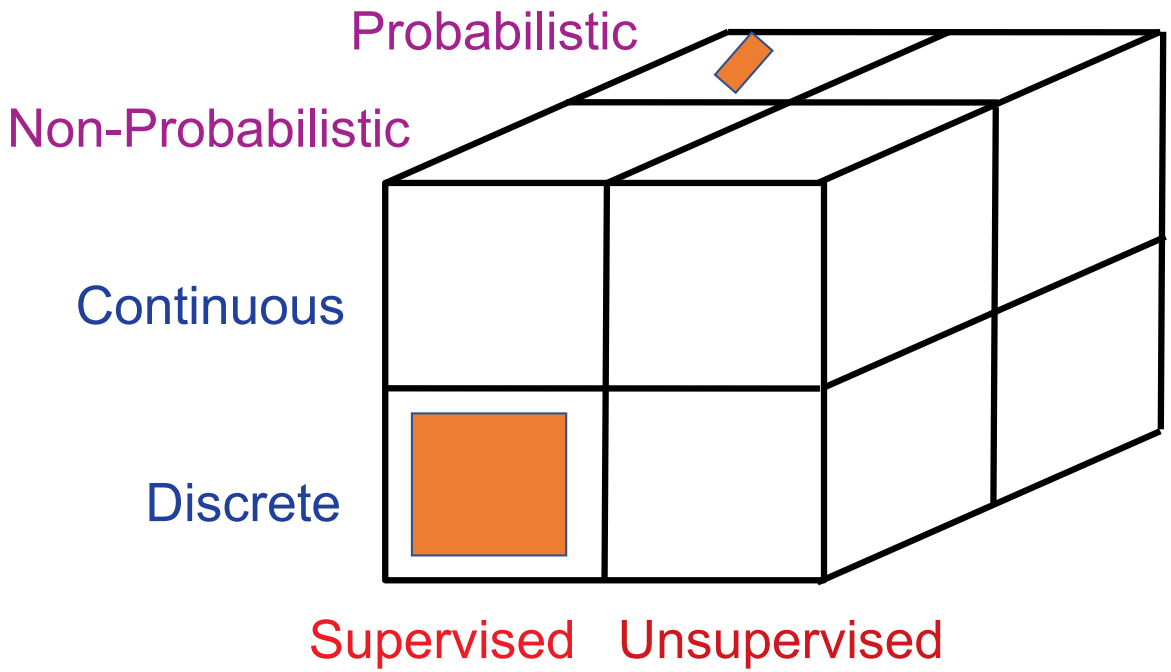




Terminology: Classification



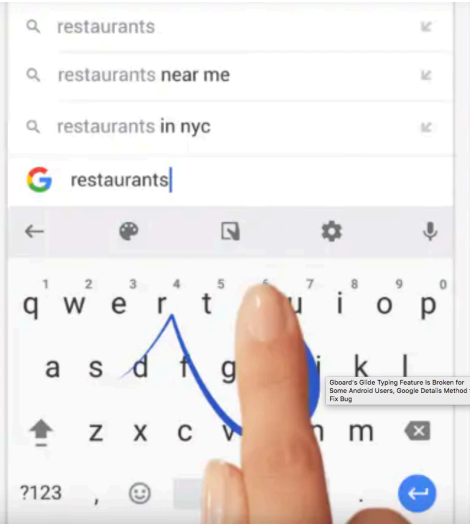




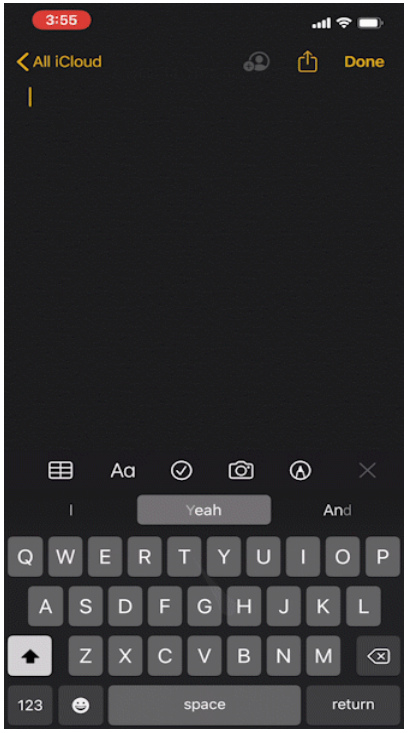
+ graphical models, reinforcement learning

Today: Probabilistic Classification

“Swipe typing”



Android Glide Typing



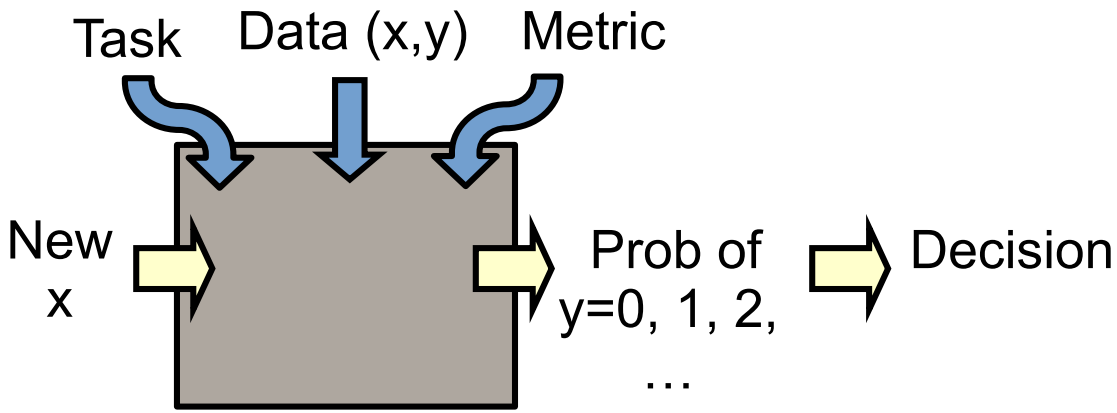
iOS QuickPath

Idea: what is the probability of word 1 vs word 2 vs word 3…? When high enough, then suggest word.

https://www.macworld.com/article/3432168/how-to-use-the-quickpath-swipe-keyboard.html
https://en.wikipedia.org/wiki/Swype
https://gadgets.ndtv.com/apps/news/gboard-for-android-glide-typing-bug-fix-google-1917639

Additional examples

- Email spam



Earn a Degree based on your Life Experience

Obtain a Bachelor's, Master's, MBA, or PhD based on your present knowledge and life experience.

No required tests, classes, or books. Confidentiality assured.

Join our fully recognized Degree Program.

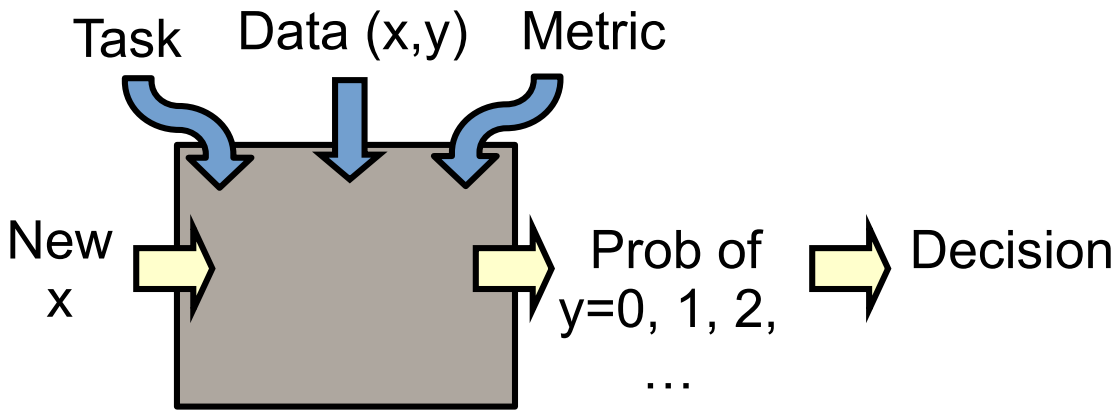
Are you a truly qualified professional in your field but lack the appropriate, recognized documentation to achieve your goals?

Or are you venturing into a new field and need a boost to get your foot in the door so you can prove your capabilities?

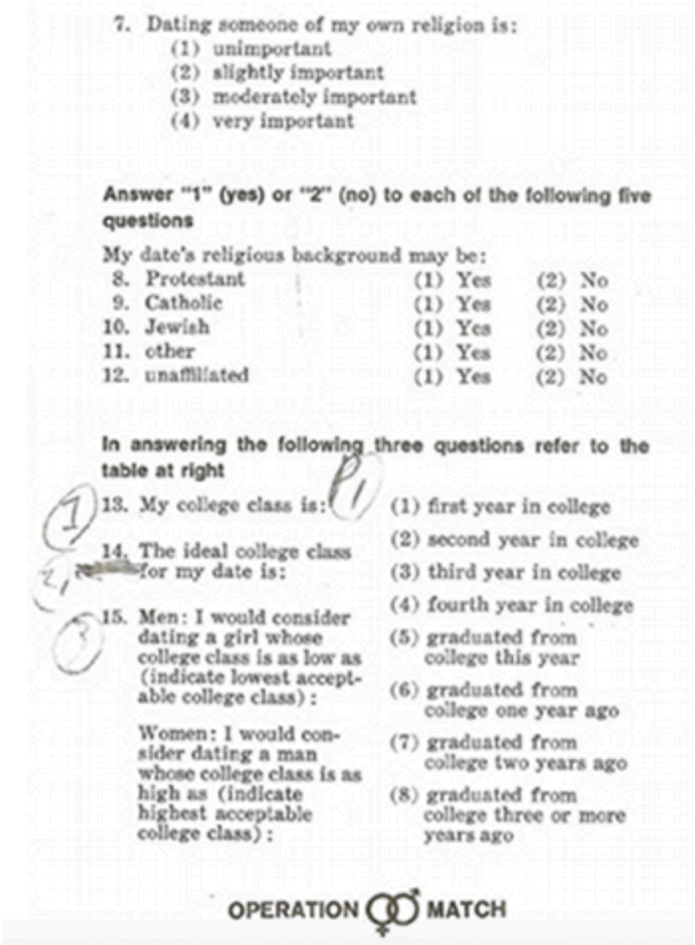
Call us for information that can change your life and help you to achieve your goals!!!

CALL NOW TO RECEIVE YOUR DIPLOMA WITHIN 30 DAYS

Additional examples



- Email spam
- Snow emergency, parking bans (prob "> 5inches snow")
- pCTR * CPC = bid (internet advertising)
- Quarantine (prob COVID-19)
- Dating platforms (prob will like each other)



7. Dating someone of my own religion is:
   (1) unimportant
   (2) slightly important
   (3) moderately important
   (4) very important

Answer "1" (yes) or "2" (no) to each of the following five questions

My date's religious background may be:

8.	Protestant	(1) Yes	(2) No
9.	Catholic	(1) Yes	(2) No
10.	Jewish	(1) Yes	(2) No
11.	other	(1) Yes	(2) No
12.	unaffiliated	(1) Yes	(2) No

In answering the following three questions refer to the table at right

13. My college class is:
14. The ideal college class for my date is:
15. Men: I would consider dating a girl whose college class is as low as (indicate lowest acceptable college class):

    Women: I would consider dating a man whose college class is as high as (indicate highest acceptable college class):

(1) first year in college
(2) second year in college
(3) third year in college
(4) fourth year in college
(5) graduated from college this year
(6) graduated from college one year ago
(7) graduated from college two years ago
(8) graduated from college three or more years ago

OPERATION MATCH

Though computer-dating was still a new concept in 1965-back then, the answers to personality questionnaires were converted into punch cards which were then fed into computers the size of small cars-two rival outfits had already popped up at Harvard: Operation Match and Contact Incorporated. Very little distinguished the two companies. Operation Match sold its questionnaires for $3. Contact charged $4. The Operation Match questionnaire was somewhat playful. Contact posed more serious questions.

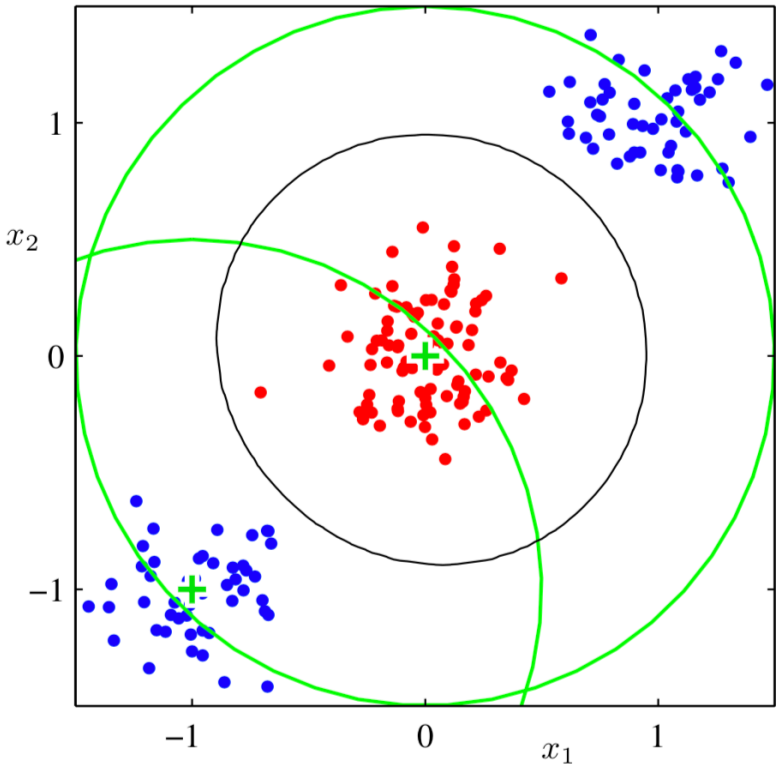
Still, both aimed to expand the campus dating pool from Wheaton to Wellesley, from Pembroke to Mount Holyoke.

https://www.seas.harvard.edu/news/2016/02/alumni-profile-jeff-tarr-ab-66

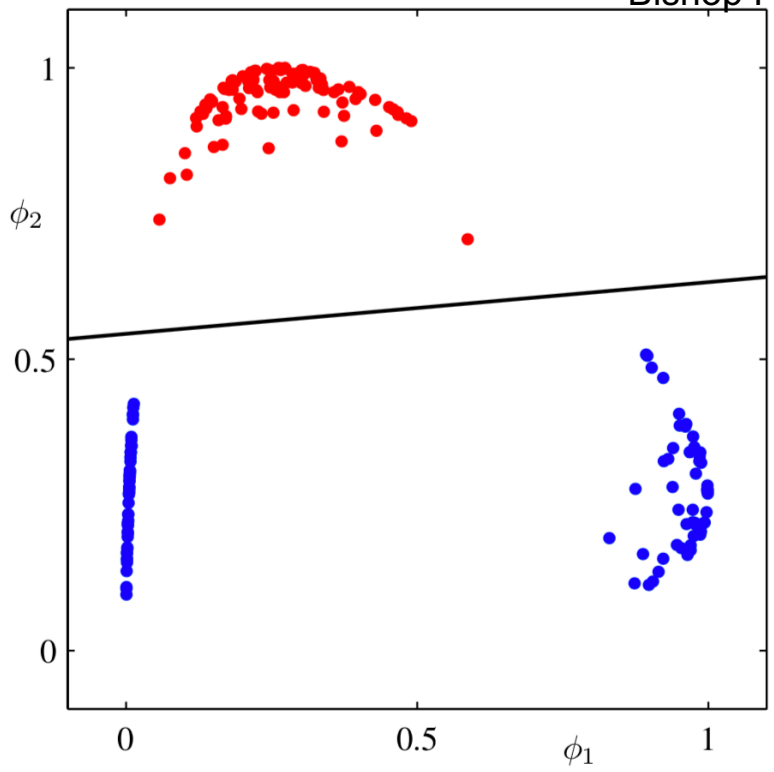
Discriminative Approach

Logistic regression with basis function

Bishop Fig 4.12



Original space, showing Gaussian basis functions and nonlinear boundary

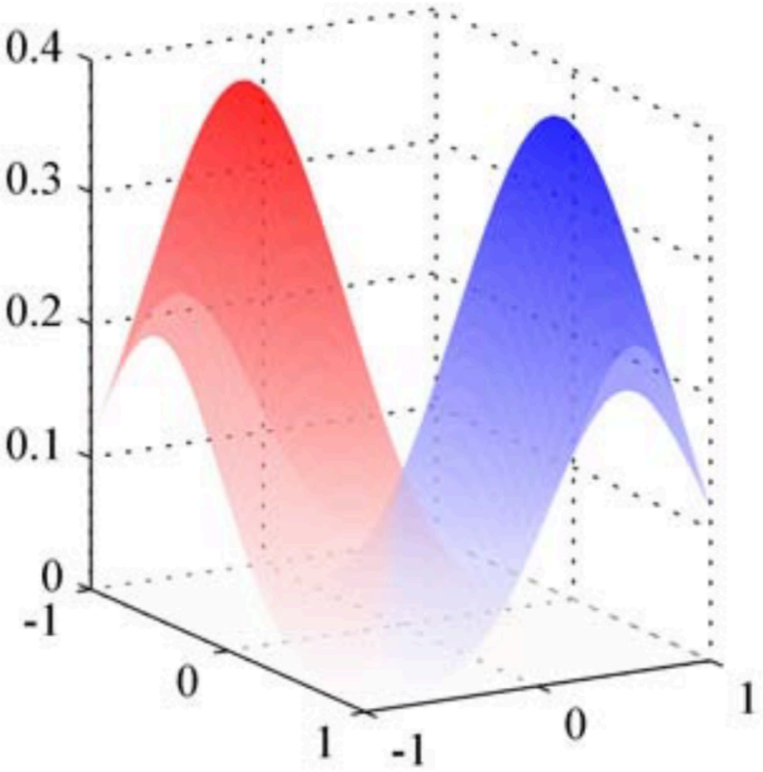


Linear decision boundary from logistic regression in basis space

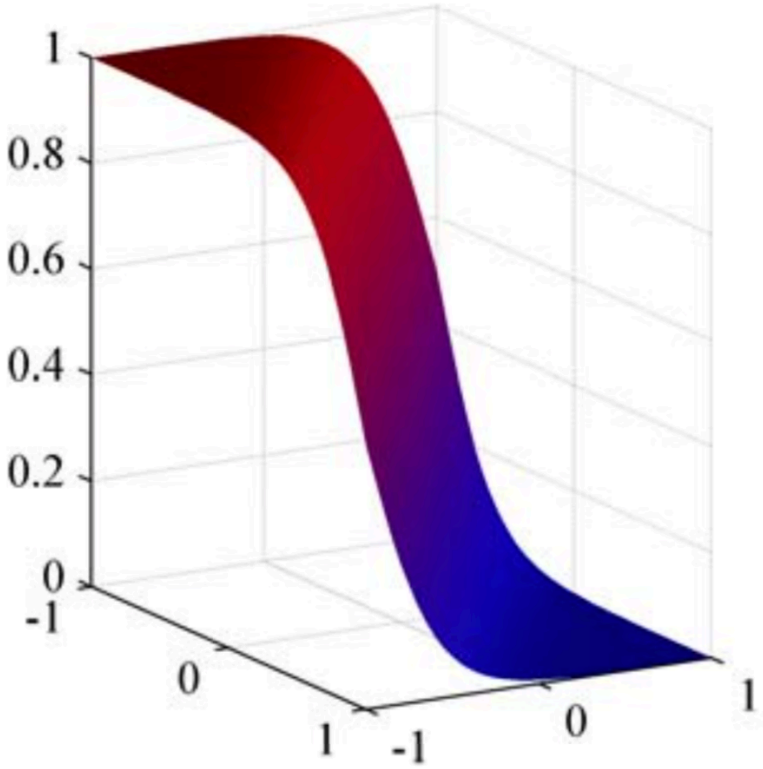
Generative Approach

Class-conditionals, binary classification

Bishop Fig 4.10



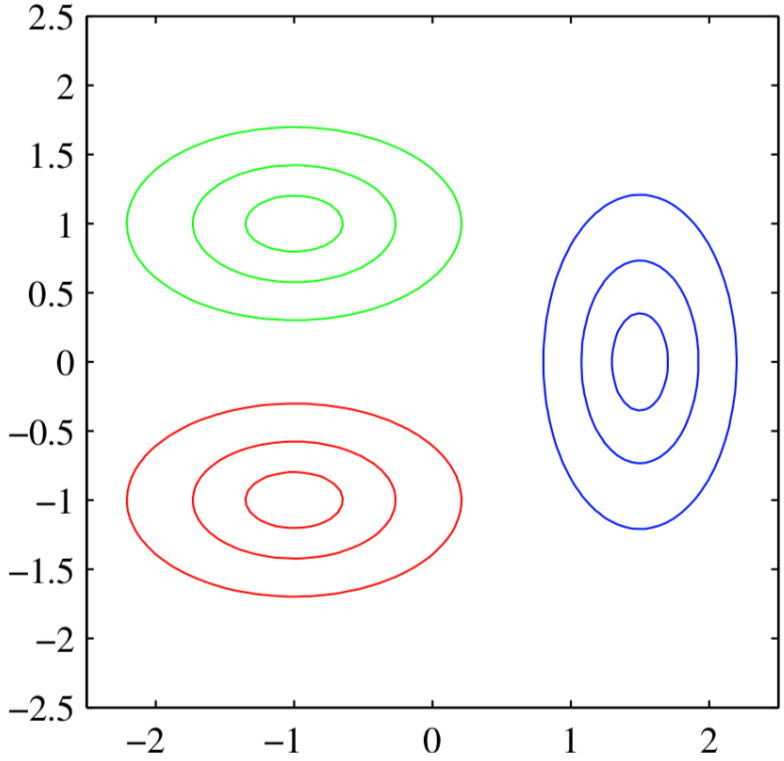
Two classes, multi-variate Gaussian conditionals



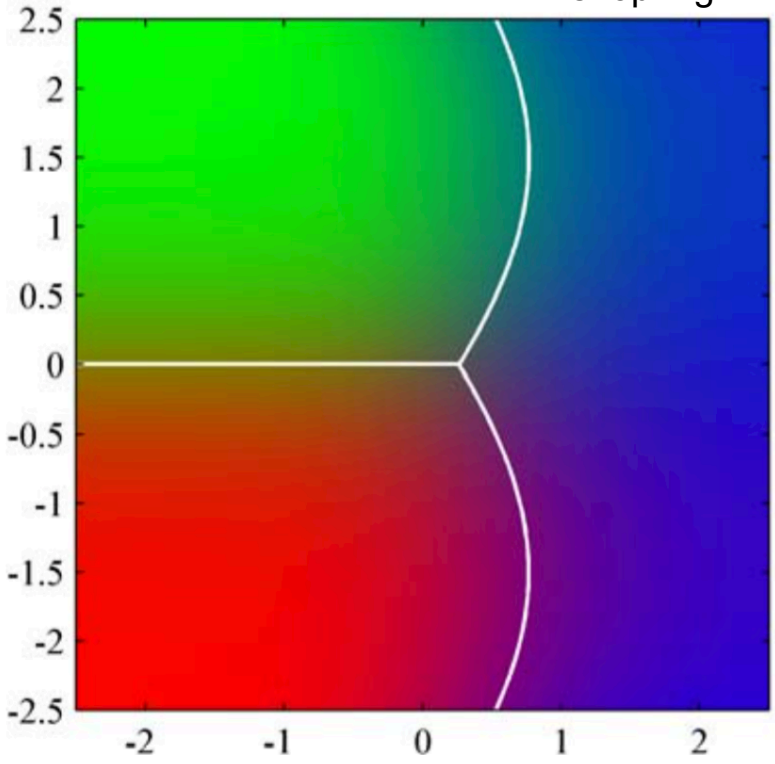
p(C_red|x) red ink, p(C_blue|x) blue ink

Class-conditional densities, 3 classes

Bishop Fig 4.11



3 class-conditionals, where red and green have same covariance matrix



p(C|x) for each class, also showing decision boundaries. Red-Green boundary is linear, whereas other classes is quadratic…