AI as Function Approximation

© Chris Atkeson 2004
RI & HCII, CMU
www.cs.cmu.edu/~cga
cga@cmu.edu

Classification, Regression, Density Estimation

- Classification: identify class [yes/no, male/female, healthy/has-flu/has-sars, …], basically identify a discrete quantity
- Regression: estimate continuous quantity [tomorrow's temperature, stock market index, robot arm acceleration]
- Probability/Density Estimation: estimate the probability or density function of outcomes

Input: Features

- Discrete features: gender, nationality, blood type
- Continuous features: location in room, barometric pressure, value of dollar
- Multiple features can be organized in a vector:

x = [male, US, AB, 3.5, 2.4, 28.9]

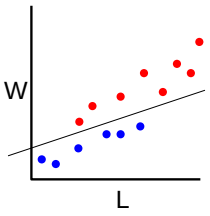
- One view is that AI is all about finding the right features. If you choose good features, a problem is easy and almost any solution method works. If you choose badly, nothing works.

AI = Function Approximation or Learning = Function Approximation

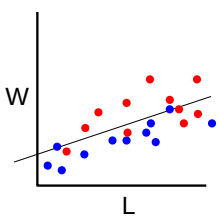
- Classification: find function that maps inputs into classes: class = f(x)
- Regression: find function that maps inputs into values: value = f(x)
- Probability/Density Estimation: find a function that maps inputs to probabilities/densities f(x)
- The function might be a formula with adjustable parameters, and learning involves adjusting the parameters so that the function correctly classifies/estimates on a training set.

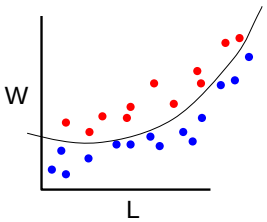
A (made up) example: Toad gender determination

- Features: Length, Weight
- Males are red dots, females blue in graph
- Separate clouds of points with a line aL + bW + c = 0
- Male: aL + bW + c > 0
- Female: aL +bW + c < 0
- Learning: find parameters a, b, and c to minimize error



What about overlap?





Sometimes a more complex function does better (is the problem shape or noise?)

I will teach you about tools (functions) to try

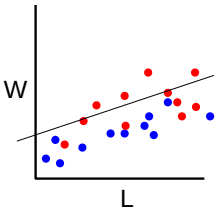
- Decision trees, neural nets, memory-based approaches, …
- http://www.cs.waikato.ac.nz/~ml is a nice source for information and software
- Your first assignment will be to classify email as spam/non-spam

Issue is how tools are used

- Just choosing a tool is usually not the best approach. How tools used is often more important than the quality of the tool.
- Example: Doctors don't just do a single classification. They can ask for more tests. Tests cost money. How should the entire process work to get the best outcome?
- Example: What is the best way to filter spam? Put email classified as spam in a folder? Rank order email in incoming list so more spam-like mail comes last? Let the users classify? Don't classify, just check for duplicate email across many users.

Often we don't want the truth, we want a good decision

- If male toads are a delicacy, but female toads are fatal to eat, we would never want to mistake a female toad for a male, but it is okay to make the opposite mistake.
- This is known as bias.



Two philosophies

- Probabilistic: Issue is noise, want to make the best classification/decision in the presence of uncertainty. Try to model probability distributions.
- Empirical/curve-fitting: Want to make the best classification/decision. Just fit the data, use what works best.
- We will first talk about empirical approaches. Later we will take a more probabilistic point of view.

Complex points

- Often continuous features are discretized (location in a room is measured on a grid, for example), or discrete features behave like continuous quantities (number of people in a room), issue is similarity, but we can induce a similarity on discrete values