EECS 391: Introduction to AI

Soumya Ray

Website: http://vorlon.case.edu/~sray/eecs391_sp12/index.html

Email: sray@case.edu

Office: Olin 516

Office hours: Tue 2:30-4:00 or by appointment

Announcements

- HW5 is due
- Quiz 5 Thursday
  - Bring calculators

Today

- Machine Learning (Ch 18.1-2)
- Naïve Bayes (Ch 20.1-2.2, up to “Maximum Likelihood Parameter Learning”)
  - Additional notes will be on website

What is “machine learning”?

- “Machine”=autonomous system
  - Artifact
  - No (or very limited) human intervention
  - Computers, robots, software agents, etc.

What is "Learning"?

"Learning denotes changes in the system that enable the system to do the same task more effectively the next time."

- **Herbert Simon**

- And also, doing *related tasks*, though never encountered before

A Specification for a Learning System

- Given: Learning task, performance measure *P*, and examples *E*

- "Learning" or "Training" phase
  - Reason about the examples *E*
  - Formulate a concept that optimizes *P*
  - Could also use any prior knowledge

- "Evaluation" or "Testing" phase
  - Use learned concept on future, novel examples

Examples

- Learn to recognize lions ("categorization")
  - *E*: animals/images, annotated "lion" or "not-lion"
  - *P*: fraction of animals correctly recognized as lion/not-lion
- Learn to drive
  - *E*: sequence of road/traffic situations and correct steering/pedal operation
  - *P*: distance traveled before crash
- Learn to play chess
  - *E*: play games against opponent, record win/loss
  - *P*: fraction of games won

Note this is different from playing a *specific* chess game with minimax, for example

Three important ideas

- Inductive Generalization
- Target Concept
- Hypothesis Space

Inductive Generalization

- In all learning problems, need to reason from specific examples to a general case
  - Sort of like mathematical induction
  - Memorization $\neq$ Learning
- Example: “Learn to recognize lions”
  - You see three lions and several non-lions
  - In future, if you only ever identify those three lions as lions, you didn’t learn the concept “lion”
    - Could be dangerous if lions find you tasty

Target Concept

- The unknown underlying concept that solves the learning task
    - E.g., “has-fur” and “long-teeth” and “looks-scary” $\Rightarrow$ “lion”

- Typically, *P* will be a measure of difference between the learned and target concepts, with respect to *E*

Hypothesis Space

- Defines the space of concepts the learning system will consider when searching for the target
  - E.g., all possible conjunctions of animal properties when looking for concept identifying “lion”

- Ideally, target concept is a member of this space

Kinds of Learning Tasks

Learning with a "Teacher"

- Examples *E* are independently generated
- Teacher annotates *E* with target concept's output
- Learning system must find a concept that matches annotations
- "Supervised" Learning (e.g. categorization)

Learning without a Teacher

- No annotations-- “find interesting patterns in *E*”
- System defines what is interesting
- “Unsupervised” Learning or “Clustering”

Reinforcement Learning

- A sequential learning problem
  - e.g. learn to navigate a maze
- Teacher/environment provides intermittent feedback
  - e.g. you get a “reward” once you reach the exit to the maze

Example Representation

- For each learning setting, the examples for different learning problems can be represented in different ways
  - Example: “Learn to recognize lions”---how to represent the animals you see?

- Representation choice affects reasoning and the choice of hypothesis space

Attribute-Value Representation

- Examples are represented by "attribute-value" pairs ("feature vector")
- Number of attributes are fixed
- Can be written as an m-by-n matrix

	Attribute_1	Attribute_2	Attribute_3
Example_1	Value_{11}	Value_{12}	Value_{13}
Example_2	Value_{21}	Value_{22}	Value_{23}
Example_3	Value_{31}	Value_{32}	Value_{33}

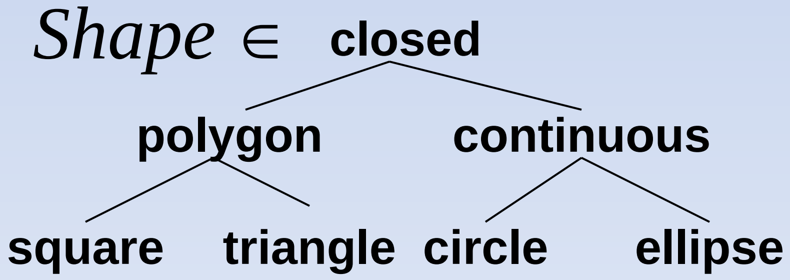
Example

	Has-fur?	Long-Teeth?	Scary?
Animal$_1$	Yes	No	No
Animal$_2$	No	Yes	Yes
Animal$_3$	Yes	Yes	Yes

Types of Attributes

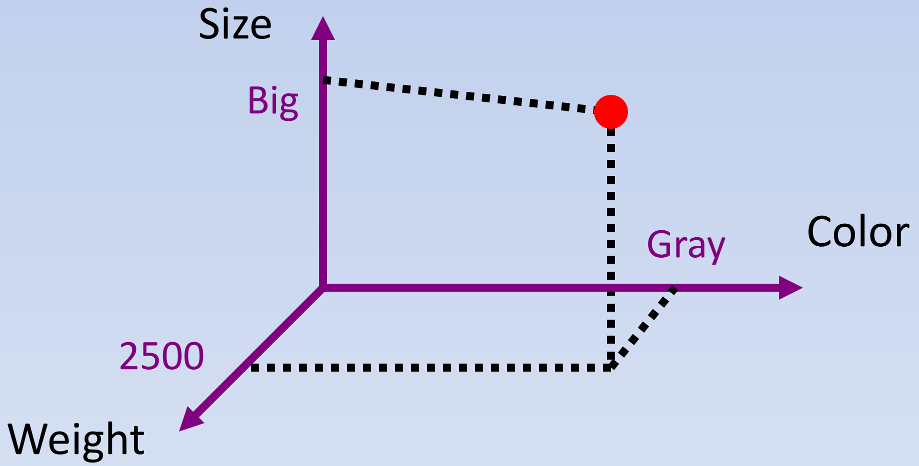
- Discrete, Nominal
- Continuous
- Discrete, Ordered
- Hierarchical

- *Color* $\in$ *(red, blue, green)*
- *Height*
- *Size* $\in$ *(small, medium, large)*
- *Shape* $\in$



Feature/Attribute Space

- We can think of examples embedded in an n dimensional vector space
  - Useful as a conceptual tool



The Binary Classification Problem

- Simplest supervised learning problem
- Target concept assigns one of two labels ("*positive*"/"*negative*", "*yes*"/"*no*") to all examples---the class label
- Can extend to "multiclass" classification, "regression"

Example

	Has-fur?	Long-Teeth?	Scary?	*Lion?*
Animal_1	Yes	No	No	No
Animal_2	No	Yes	Yes	No
Animal_3	Yes	Yes	Yes	Yes

Attribute-value representation ($\mathbf{X}$)

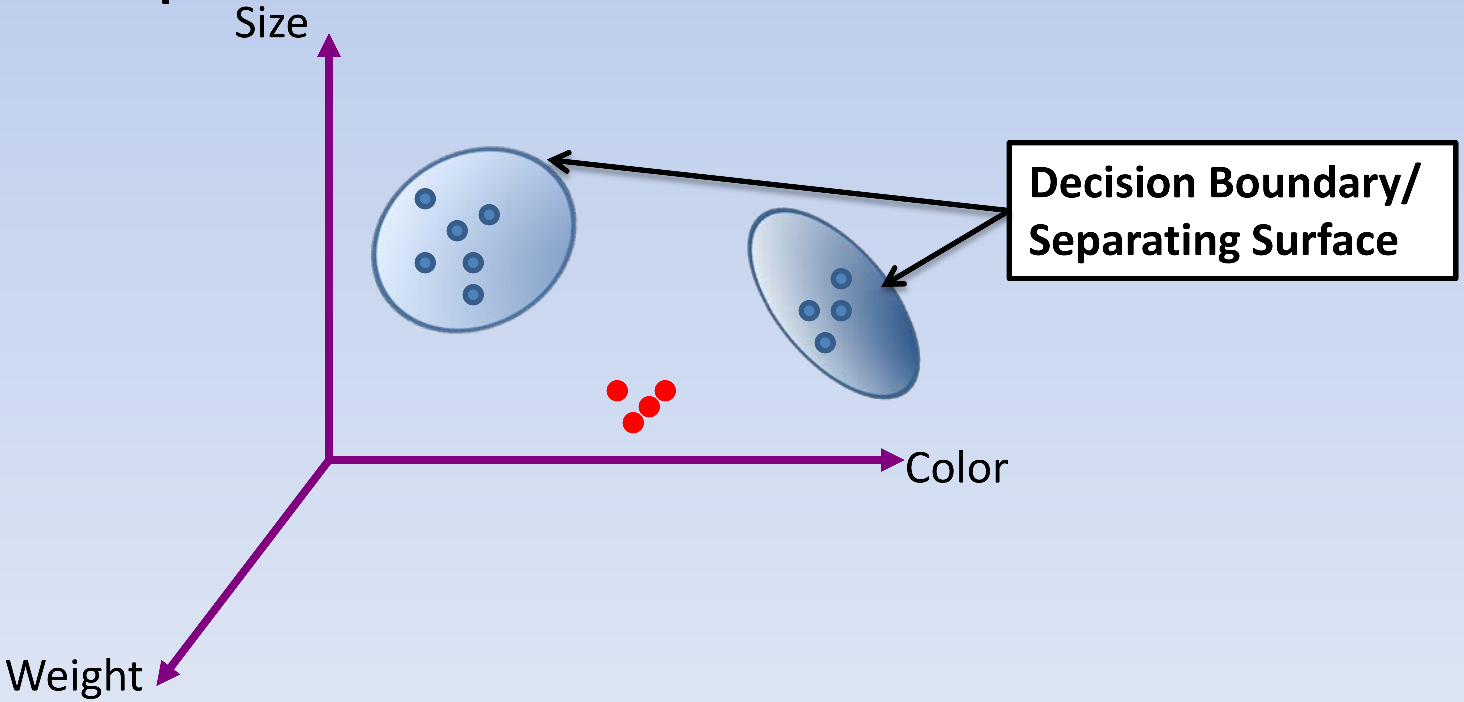
Binary Class Label (y) (assigned by target concept/ teacher)

Learning Problem

- Given such *training data*, we want to produce a concept (classifier) that will accurately classify new examples (*test data*)
  - Predict y given input $\mathbf{x}$
- The algorithms that do this have the property that given more training data, they produce better classifiers (more accurate predictions)
  - These are "learning algorithms"

Classifier Geometry

- Often useful to think of classifiers as N-dimensional volumes (possibly disjoint) in feature space



Probabilistic Classification

- Given examples as before, one way to solve the learning problem is:
  - Treat the features ($\mathbf{X}$) and the class label (Y) as random variables
  - Construct a probabilistic model of $p(\mathbf{X}=\mathbf{x}, Y=y)$
    - This involves estimation
  - Given a new example, for which $\mathbf{X}$ is known and has value $\mathbf{x}$, calculate $p(Y=y|\mathbf{X}=\mathbf{x})$
    - This involves inference
    - Return the y with highest probability

One problem

- Suppose examples are described by n Boolean (yes/no) attributes and a class label
  - How many atomic events are there? 2^{n+1}
- We can't store (or estimate) so many numbers, so we need to *simplify* the representation of $p(\mathbf{X}=\mathbf{x}, Y=y)$
- We've already seen how to do this
  - As in Bayesian networks

Naïve Bayes

- Simple probabilistic classifier for discrete data

$$p_{\mathbf{X},Y}(\mathbf{x}, y) = p(\mathbf{X} = \mathbf{x} \mid Y = y)\, p(Y = y)$$

$$= p(x_1, \ldots, x_n \mid Y = y)\, p(Y = y)$$

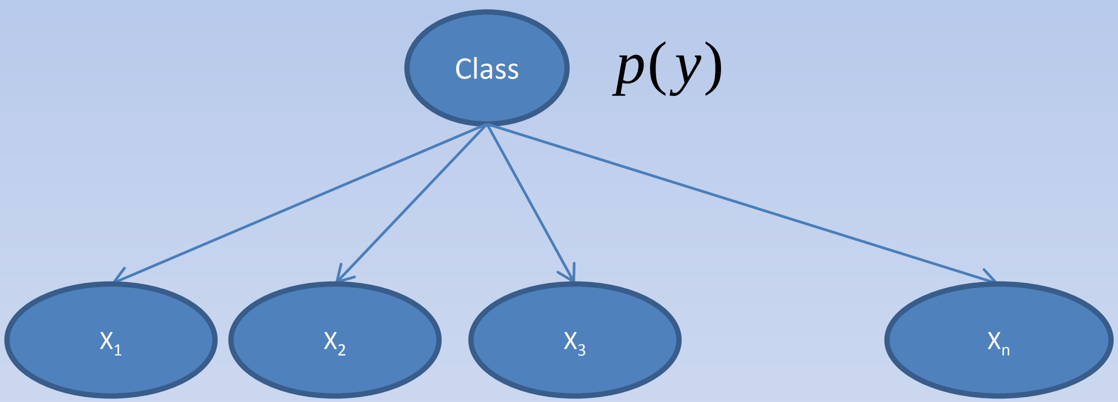
$$= \prod_i p(X_i = x_i \mid Y = y)\, p(Y = y)$$

Naïve Bayes assumption: Attributes are conditionally independent given the class

Naïve Bayes parameters: Instead of storing probabilities for each atomic event, we will only store these conditional probabilities and use this formula to calculate the probability for an atomic event (example).

Naïve Bayes as a Graphical Model

- The naïve Bayes model is a special case of a Bayesian network



$p(y)$

$p(X_i \mid y)$

$$p_{\mathbf{X},Y}(\mathbf{x}, y) = \prod_i p(X_i = x_i \mid Y = y)\, p(Y = y)$$

Example

	Has-fur?	Long-Teeth?	Scary?	*Lion?*
$Animal_1$	Yes	No	No	No
$Animal_2$	No	Yes	Yes	No
$Animal_3$	Yes	Yes	Yes	Yes

Naïve Bayes parameters:
p(Lion), p(Has-fur|Lion), p(Not-Has-fur|Lion), p(Long-Teeth|Lion), p(Not-Long-Teeth|Lion), p(Scary|Lion), p(Not-Scary|Lion)
p(Not-Lion), p(Has-fur|Not-Lion), p(Not-Has-fur|Not-Lion), p(Long-Teeth|Not-Lion), p(Not-Long-Teeth|Not-Lion), p(Scary|Not-Lion), p(Not-Scary|Not-Lion)

How many parameters?

- Two for $p(Y=y)$
- One each for $p(X_i=x_i|Y=y)$
  - Suppose X_i is Boolean
- $2(2n+1)$ total---much better than 2^{n+1}
  - Of these, need to estimate only $2n+1$

Classification with Naïve Bayes

- For a new example, calculate $p(\mathbf{x},Y=\text{“}positive\text{”})$ and $p(\mathbf{x},Y=\text{“}negative\text{”})$ and choose whichever is greater

$$p_{\mathbf{X},Y}(\mathbf{x}, pos) = \prod_i p(X_i = x_i \mid Y = pos)\, p(Y = pos)$$

- Earlier we said we wanted $p(y|\mathbf{x})$??

Example

	Has-fur?	Long-Teeth?	Scary?
Animal$_1$	Yes	No	No

p(Has-fur=Yes|Lion)=0.5, p(Has-fur=Yes|Not-Lion)=0.1
p(Long-Teeth=Yes|Lion)=0.9, p(Long-Teeth=Yes|Not-Lion)=0.5
p(Scary=Yes|Lion)=0.8, p(Scary=Yes|Not-Lion)=0.5
p(Lion)=0.1

Probability Estimation for Naïve Bayes

- Given a set of observations:

	Has-fur?	Long-Teeth?	Scary?	*Lion?*
Animal_1	Yes	No	No	No
Animal_2	No	Yes	Yes	No
Animal_3	Yes	Yes	Yes	Yes

- Estimate parameters $p(X_i=x_i|Y=y)$ and $p(Y=y)$
- We will use a method called "Maximum Likelihood Estimation"