More Than A Program

- Usually, we think a program is something written by an experienced person.

- Often, the program isn’t complete without “experience” of its own.

- Example: If I had to write a compression program, I could do ok. But, using the character-count statistics from the Gettysburg Address makes it possible to produce the most compact code.
Machine learning is the idea of writing programs that use data (experience) to create better programs than people can write directly.

Supervised learning is a specific problem in which the experience is in the form of labeled examples and the learning system needs to learn to label novel examples.
In its simplest form, a learner’s job is to produce a classifier.

A classifier takes objects as input and assigns each one to a class.

Most simply, objects are represented as vectors of features and classes are 0/1.
Software Used To Predict Who Might Kill

Posted by kdawson on Mon Dec 04, 2006 02:11 AM
from the three-psychics-in-a-pool dept.

eldavojohn writes

"Richard Berk, a University of Pennsylvania criminologist, has worked with authorities to develop a software tool that predicts who will commit homicide. I could not find any papers published on this topic by Berk, nor any site stating what specific Bayesian / decision tree algorithm / neural net is being implemented."

From the article:

"The tool works by plugging 30 to 40 variables into a computerized checklist, which in turn produces a score associated with future lethality. You can imagine the indicators that might incline someone toward violence: youth; having committed a serious crime at an early age; being a man rather than a woman, and so on. Each, by itself, probably isn't going to make a person pull the trigger. But put them all together and you've got a perfect storm of forces for violence,' Berk said. Asked which, if any, indicators stood out as reliable predictors of homicide, Berk pointed to one in particular: youthful exposure to violence."
Example Classifier

- **Input**: A high school student
- **Output**: Will the student drop out of college?
- **Vector of Features**: Score on SATs, grades in Math/English/Science, age, parent’s income, years at current address, height
- Such a classifier might be useful as a tool for admissions or financial aid.
Learning: The Problem

- **Input**: A *training set* consisting of *labeled instances*, each of which is a feature vector and a desired class (1 = yes, 0 = no).

- **Output**: A classifier, which we hope will accurately assign new feature vectors to classes.

- A *learning algorithm* is a program that addresses this problem.
Informal to Formal

• Note that problem is not really defined at this point! Whether it does well on new examples isn’t directly measurable.

• One way to make it formal: find classifier with minimum number of mistakes on the training set.

• Usually not enough to create good classifiers: overfitting.
Learning Algorithms

- Decision trees
- Boosting
- Nearest neighbors
- Support Vector Machines
- Neural networks
- Naive Bayesian classifiers
- etc.
Some Applications

- Autonomous driving (camera $\Rightarrow$ steering wheel)
- Loan applications (age, income $\Rightarrow$ default?)
- Handwriting recognition (scanner $\Rightarrow$ digits)
- Speech recognition (audio $\Rightarrow$ phonemes)
Decision Trees

- Decision trees are another kind of classifier.
- Branch depending on test on attribute value.
- Can build such a tree incrementally and automatically using examples.
• Machine learning is almost always grounded in measurable data.
• Computer programs that can collect and use their own data can be quite powerful.
• Essentially, that’s what learning robots are.
Forever...

- What sorts of program would *purposely* have an infinite loop?
- Think about a software-controlled thermostat. It might have a program that looks something like:
Loop Forever

- operating systems
- user interfaces
- video games
- process controllers
- robots
Robot Basics

- From a software standpoint, modern robots are just computers.
- Typically, they have less memory and processing power than a standard computer.
- Sensors and effectors under software control.
Standard Robots

- Industrial manufacturing robots.
- Research /hobby robots.
- Demonstration robots.
- Home robots.
- Planetary rovers.
- Movie robots.
Manufacturing

- Often arms, little else.
- Part sorting.
- Painting.
- Repeatable actions.
Research / Hobby

- Pioneer
- Handy Board / Lego
- Segbot
- Stanley
Space Exploration

- Sojourner
- Deep Space Agent
Home Robots

- Roomba.
- Mowers.
- Moppers.
- Big in Japan.
- Nursebots.
- Emergency rescue bots, Aibo.
Demonstration Robots

- Honda: Asimo.
- Toyota: lip robot.
- Sony: QRio.
Sensors and Effectors

- Sensors:
  - bump
  - infrared
  - vision
  - light
  - sonar
  - sound

- Effectors:
  - motors
  - lights
  - sounds
  - graphical display
  - laser