Lecture 6

Analytical Learning Discussion (3 of 4):
Learning with Prior Knowledge

Monday, January 31, 2000

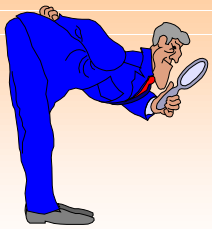
Aiming Wu
Department of Computing and Information Sciences, KSU
http://www.cis.ksu.edu/~awu8759

Readings:
Chown and Dietterich

Paper

- "A divide-and-conquer approach to learning from prior knowledge"
- Eric Chown and Thomas G. Dietterich
- Reviewed by Aiming Wu

What's the problem?



- Goal: calibrate the free parameters of MAPSS (Mapped Atmosphere-Plant-Soil System)
- Term: conceptual parameters--not directly measured, but summarize details.

MAPSS

- Purpose: predict the influence of global climate change on the distribution of plant ecosystems worldwide.
- Inputs: climate data from 1,211 USA weather stations and interpolating to 70,000sites.
- Outputs: amount of vegetation (LAI tuple: tree, grass, and shrub) and biome classification (Runoff?).

Calibration task

- Using manually chosen parametervalues, predict outputs for LAI and Runoff.
- Define a error function:
  J (s, θ) = sum (predicted - actual) ^2
- Find a value for θ so that sum of J(s, θ) over all sites s is minimal.
- Characteristics of the task:
  -- Imagine the computational burden.
  -- Non-linear nature: competition, threshold, exponential equations.

Approaches to attack the task

- Search approach:
  --gradient descent (hill-climbing)
  --simulated annealing
  --Powell's method
- Set-interaction approach
- Decomposition approach:
  decompose the overall problems into independent sub-problems.

How to solve the problem? A divide-and- conquer calibration method

- Pre-requirement: there are sites where only a subset of the parameters are relevant to the MAPSS' computations.
- Sub-problems:
  - --Identify "operating regions"
  - --Identify sites related to each region
  - --Calibrate parameters in each operating region

Identify operating regions

- Identify control paths through MAPSS program that involve the relevant parameters.
- Problems:
  - --MAPSS C program: hard to find a path.
  - --Iterative search for good LAI values make the # of paths infinite.
- Approaches:
  - --translate the MAPSS C program into a "declarative single-assignment, loop-free programming language" and analyze it using the partial evaluation techniques.
  - --Using the actual LAI values instead of searching for them.
  - --Start with M=1, and increasing.
  - --Stop when M unknown parameters have been found.

Identify training examples for a path (1)

- EM-style algorithm
  - --Initialize θ, compute the probability of each example belonging to each path.
  - --Hold the probabilities, modify θ to minimize the expected error
- Problems with the algorithm
  - --Global optimization, not good for large model.
  - --Local maxima.

Identify training examples for a path (2)

- Data Gathering Algorithm
  - --Filtering phase: initialize 40 random θs, get 40 training examples, compute J, select 20 models with the smallest J, test each example and determine its "pass" status until 40 examples have passed the filter.
  - --Calibration phase: one-to-one map the examples to models, simulated annealing search, update the 40 models.
- Characteristics of the approach
  - --Voted decision of the 40 models is more accurate than the decision of a single model.
  - --The filtering set is robust to bad models (20 models).
  - --The one-to-one match makes it robust to bad examples.

Calibrate the path

- Simulated annealing, start with parameter values from the best model found in the last filtering phase.
- Parameter temperature mechanism: calibrated parameters have low "temperatures", reduced exponentially as a function of the number of previous paths that calibrate the parameters.

Was the problem solved?

- According to the authors, yes.
- "The results agree closely with the target values except for the top soil level saturated drainage parameters, which are evidentally somewhat underconstrained by the model and the data."

Summary (1)

- Content critique
  --Key contribution: A decomposition approach of calibrating free parameters of large, complex, non-linear models
  --Strengths:
  •Identify control paths of a program.
  •Data Gathering Algorithm.
  --Weakness:
  •Total is larger than the sum of its parts. Is it true here?
  •What is the prior knowledge's role?

Summary (2)

- Presentation critique
  --Audience: machine learning expert, with interest in large, complex scientific models.
  --Positive points:
  •Detailed explanation of the MAPSS model and the divide-and-conquer approach.
  •Good comparison of the authors' approach with others'.
  --Negative points:
  •Tiger's head, snake's tail. Need elaboration on the final results so that they look more convincing.
  •Missed some important terms, such as Runoff, Vmin, etc.