

Model-based Machine Learning

Chris Bishop

Microsoft Research Cambridge

Royal Society, March 2012









Traditional machine learning

- Logistic regression
- Neural networks
- K-means, mixture of Gaussians
- PCA, kernel PCA, ICA, FA
- Support vector machines
- Deep belief networks
- Decision trees and random forests
- ... many others ...

Model-based machine learning

Goal: a single modelling framework which supports a wide range of models

Traditional:

"how do I map my problem onto a standard algorithm"?

Model-based:

"what is the model that represents my problem"?

Realisation of model-based ML

- Bayesian framework
- Probabilistic graphical models
- Efficient deterministic inference

Movie recommender demo

Probabilistic graphical models



P(M, G, A) = P(M) P(G|M) P(A|M)

Graph structure captures domain knowledge

Efficient inference

$$\sum_{x} \sum_{y} xy = x_1y_1 + x_2y_1 + x_1y_2 + x_2y_2$$
$$= (x_1 + x_2)(y_1 + y_2)$$

Local message-passing

Maths (M)



What if distributions are intractable?







True distribution

Monte Carlo Variati

Variational Message Passing Loopy belief propagation

Expectation propagation

Algorithms → Models

$$\overline{\mathbf{x}} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{x}_n$$

$$\mathbf{S} = \frac{1}{N} \sum_{n=1}^{N} (\mathbf{x}_n - \overline{\mathbf{x}}) (\mathbf{x}_n - \overline{\mathbf{x}})^{\mathrm{T}}$$

$$\mathbf{S}\mathbf{u}_i = \lambda_i \mathbf{u}_i$$



M. E. Tipping and C. M. Bishop (1997)

C. M. Bishop (1999)









Childhood Asthma



Allergic Sensitisation Model



Comparison with traditional ML

Separation of model and training algorithm Auto-generated inference algorithm

Easy extension to more complex situations

Modify model, use the same inference algorithms Flexible as requirements change

Compact code

Easy to write and maintain

Transparent functionality

Many traditional methods are special cases One simple framework for newcomers to the field



Computational size vs. statistical size





Noisy ranking



Conventional approach to ranking: "Elo" single strength value for each player cannot handle teams, or more than 2 players

Bayesian Ranking: *TrueSkill*[™]



R. Herbrich, T. Minka, and T. Graepel; NIPS (2006)

Multi-player multi-team model



*TrueSkill*TM

Sept. 2005; 10s of millions of users; millions of matches per day

Convergence

Number of Games

- Specify your machine learning problem as a probabilistic model in a .NET program (typically 10-20 lines of code).
- 2. Use Infer.NET to compile the model into optimized runtime code.
- 3. Run the code to make inferences on your data automatically.

research.microsoft.com/infernet

research.microsoft.com/~cmbishop

