Tommi Jaakkola and Tomaso Poggio

Project Overview

- We seek to fundamentally improve the search of relevant information from large relatively unstructured databases
- Our goal is to develop an optimal interactive retrieval system
  - We substantially extend and integrate existing machine learning methods to solve the associated challenges
  - We design and develop proof of concept tools

Tommi Jaakkola and Tomaso Poggio

Progress Through June 2002

- Efficient active learning method for large domains
- Generally applicable continuation method for stable integration of heterogeneous data sources
- Information regularization approach to exploiting large numbers of unlabeled documents in filtering tasks
- Generalized low-rank matrix factorization approach to collaborative filtering

Research Plan for the Next Six Months

- Selected tasks:
  - Extension of the overall active learning approach to filtering problems involving multiple simultaneous topic categories (with Dr. Ueda)
  - Public release of software for stable integration of heterogeneous data sources
  - Extension, testing, and efficient implementation of information regularization for large filtering tasks
  - Concurrently solving multiple related retrieval tasks through a new matrix factorization approach to collaborative filtering