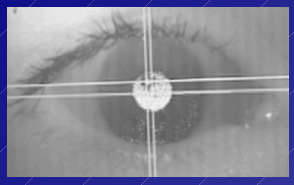
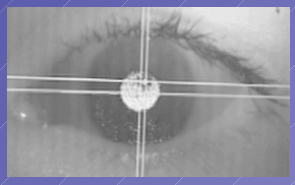
Proactive Information Retrieval: relevance feedback from eye movements ?

Jarkko Salojärvi, Kai Puolamäki and Samuel Kaski
Neural Networks Research Centre, Helsinki University of Technology,
P.O. Box 9800, FIN-02015 HUT, Finland.
Email: {forename.surname}@hut.fi





Motivation

Information retrieval benefits from user feedback.
- Explicit feedback is coarse-grained and laborious.
- How about implicit feedback, eye movements during reading for example ?
  - Rich source of information.
  - Very noisy.

This is a feasibility study using standard methods.

Research problem

Since relevance of a document is subjective, we first designed a controlled setup where relevance is known:
- Find an answer to a question from a list of titles.
- Each title is known to be either:
  - I (irrelevant for the question)
  - R (relevant for the question)
  - O (correct answer)

The goal of this work is to try to predict the known relevance of a title from eye movements.

Our research questions are:

1. Can relevance be predicted from eye movements at all?
2. Do the models benefit from the time series nature of the data?
3. Do discriminative models help?
4. Does modeling of the global scanning behavior help in predicting relevance?
5. Is it possible to discover reading strategies of the user with different HMM structures?

Experimental setup. Left: The eye movements are measured with a head-mounted eye tracker. The tracker consists of a helmet with two cameras; one monitors the eye and the other one the visual field of the subject. Right: The eye movement pattern during reading plotted on the assignment. Pattern consists of rapid eye movements, *saccades* (lines in the picture), followed by *fixations* where the eye is fairly stable (circles). (Matlab reconstruction)

Feature extraction

- Fixations were assigned to the nearest word.
- For each word, a set of 21 standard features were computed.
- Features most responsible for discrimination were sought by a Bayesian MLP using an ARD prior.
- The variables are:
  - I One fixation or many (Binary)
  - II log of total fixation duration (Gaussian)
  - III Reading behavior (Multinomial): skip next word, go back to already read words, read next word, jump to an unread line, or last fixation in an assignment.
- For LDA and SVM (below), the data was averaged to a title-spesific feature vector.

Models

Dumb Classifier

63.2 %

•Assign all to largest class.

1 Linear Discriminant Analysis

69.2 %

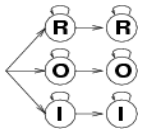
•Simplest classifier.

Support Vector Machine

75.0 %

Relevance can be inferred to some extent. Averaged features cannot explain user behavior.

2 Separate Hidden Markov Models

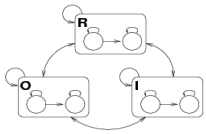


71.3 %

•Optimize HMMs for each class.
•MAP prediction.

Compared to 1: Time series modeling improves results somewhat. Data is not hopelessly noisy.

3 Discriminative Chain of Hidden Markov Models

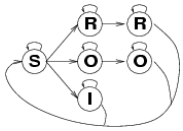


76.4 %

•1st level models transitions between titles.
•2nd level models transitions between words.
•Discriminative training using Extended Baum-Welch with Viterbi approximation in the 2nd level.

Compared to 2: Discriminative learning helps.

4 Global Hidden Markov Model



75.8 %

•Add a common state 'S' (scanning).
•First optimize each branch separately.
•Then optimize the whole model, keeping emission distributions of the 'R', 'O', 'I' branches fixed.

Modeling the whole trajectory improves results. The HMM structure models user behavior? (indication of 5).

Summary

Relevance can be inferred to some extent already with SVMs of average features. We still move to using HMMs to ultimately be able to model/discover patterns of user behavior. Discriminative learning and modeling of whole behaviour patterns seem to work.

More information: http://www.cis.hut.fi/projects/mi/prima.html

Acknowledgments:
The work was carried out together with Center for Knowledge and Innovation Research of Helsinki School of Economics and supported by Academy of Finland, grant 200836.
Thanks to: Jaana Simola, Ilpo Kojo, Janne Sinkkonen and Nelli Salminen.