Proactive Information Retrieval: relevance feedback from eye movements?

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?

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.

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-specific feature vector.

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 behavior patterns seem to work.


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.