



THE UNIVERSITY OF TENNESSEE

College of Engineering

The Min H. Kao Department of Electrical Engineering and Computer Science

- Announcement of a Doctoral Dissertation Defense -

Claxton 203 (Conference Room)
Friday, November 5, 2010 @ 2:00 P.M.

“A Visual Approach to Automated Text Mining and Knowledge Discovery”

Andrey Puretskiy, Ph.D. Candidate
Dr. Michael W. Berry, Major Professor

Abstract

The focus of this dissertation has been on improving the non-negative tensor factorization technique of text mining. The improvements have been made in both pre-processing and post-processing stages, with the goal of making the non-negative tensor factorization algorithm accessible to the casual user. The improved implementation allows the user to construct and modify the contents of the tensor, experiment with relative term weights and trust measures, and experiment with the total number of algorithm output features.

Non-negative tensor factorization output feature production is closely integrated with a visual post-processing tool, FutureLens, that allows the user to perform in depth analysis and has a great potential for discovery of interesting and novel patterns within a large collection of textual data.