TextMining

Jianchi Wei, Chongyu He & Chenyu Mao

CS 4624 Multimedia, Hypertext & Information Access
Edward Fox
Virginia Tech, Blacksburg VA
24061
May 10, 2022
Outline

- Project Review
- Work Timeline
- Project Design
- Front End
- Lessons Learned
- Future Work
- Acknowledgments
- References
Project Review

- Extract valuable information and keys from Electronic Theses and Dissertations (ETDs).
- Using LDA, ProdLDA, NeuralLDA, and CTM models for summarizing topics.
Work Timeline

**Text Mining Timeline**

**Models**
To get familiar with training model, and build basic HTML structure with FLASK

**Testing & Fixing**
After implementing all features, to test and fix bugs

**January (JAN)**
- Selecting Project Topic

**February (FEB)**
- Connecting & Data Processing
  To connect back-end and front-end by FLASK, and train four models by preprocessed data

**March (MAR)**
- Final Submission
  Finish all, including final report, code integration
Project Design

- Index page
- Topic page
- Document page
- Advanced research page
- Search keyword page
Front End Demo

The video for the demonstration is in the file TextMiningDemo.mp4.
Lessons Learned

- Front End with FLASK
- Combination of JavaScript, CSS, and HTML
- Testing data
Future Work

Deploy to server
Aman Ahuja
aahuja@vt.edu

Professor Edward A. Fox
fox@vt.edu
References

