CSC 510 Software Engineering: Project 1 [*]

Group-P [†]

Monica Metro North
Carolina State University
mgmetro@ncsu.edu

Zachery DeLong North
Carolina State University
2305 Horizon Hike Ct
Raleigh, NC
zpdelong@ncsu.edu

Zhangqi Zha North Carolina
State University
zzha@ncsu.edu

Bikram
Brookhaven Laboratories
bikram@ncsu.edu

ABSTRACT

Choosing a textbook for a class is something that many professors do several times a year, but the vast number of available textbooks for any given subject makes choosing one that both meets the needs of the class and meets the budget constraints of today's students is a challenge. In this paper, we propose two methods for automatically learning topics from books, a set of evaluations for those methods, and a system that will use those methods in a web app that attempts to suggest textbooks to a professor that balance hitting all topics the course requires while still being economical for students.

1. INTRODUCTION

One of the most basic tasks that a professor must do is identify what textbook to use for a given class. While some fields have textbooks that have become popular and are clearly the best in their subject matter, many fields (especially emerging ones) have no such exemplar. It would be a complex enough problem if there were not already a massive number of textbooks on sources such as Amazon, but self aware professors have attempted to use textbooks that are less expensive with hopes of allowing lower income students to attend more easily. This admirable intention serves to increase the time needed to research the given books, and there is no obvious heuristic to apply to searches to limit the field.

In the interests of making students lives less expensive, we propose exploring a system which can identify topics in

books automatically, and which, given a set of requirements from a user (usually a professor), can suggest options of varying expense and completeness for a given set of topics. To do this, we propose exploring word clusters generated by Doc2Vec, a family of common natural language processing (NLP) algorithms that attempt to cluster similar words, and topics generated by latent dirichlet allocation (LDP), another common NLP algorithm that directly attempts to identify topics in text, to automatically infer the content of a set of textbooks. We then intend to build a web app that will allow a user to specify a set of topics and which will use the mentioned algorithms to make suggestions while minimizing the cost of said textbooks.

2. LITERATURE REVIEW

This is a test of our lit review

3. PROPOSED SOLUTION

4. EVALUATION PLAN

This is a test of our evaluation plan sectione

5. CONCLUSION

This is a conclusion