CS410 Final Project Proposal

Project Topic: Intelligent Learning Platform: Organize the scattered lectures into a coherent "multimedia textbook" and create an index

Group Name: Freelancer

Member: Jiecheng Zhao (jz109@illinois.edu)

1. **What are the names and NetIDs of all your team members? Who is the captain? The captain will have more administrative duties than team members.**

   One team member
   Name: Jiecheng Zhao (Captain)
   NetID: jz109

2. **What topic have you chosen? Why is it a problem? How does it relate to the theme and to the class?**

   The topic of this project is to organize the scattered lectures into a coherent "multimedia textbook" and create an index.

   A course typically covers many words, but only a few are key words and relates to the knowledge introduced in the course. The learner may like to quickly find the lecture or the location a specific topic is presented, or a specific key word is defined and explained.

   This topic relates to the text retrieval, language model, and topic analysis introduced in the text information system class.

3. **Briefly describe any datasets, algorithms or techniques you plan to use**

   The initial dataset will be used in this project is the lectures of this class (CS410 Text Information System). It is expected to extend to other classes on Coursera and other Massive Open Online Courses (MOOC) platforms.

   Technologies such as BM25, language model, and topic analysis (including EM algorithm)

4. **How will you demonstrate that your approach will work as expected? Which programming language do you plan to use?**

   The results of the initial work on CS410 will be compared with the key words listed in each week's overview part. The missing and adding key words will be manually checked. Furthermore, the effectiveness of the tool will be demonstrated on other MOOC.

   The programming language being used for this project will be Python. Toolboxes such as MeTA, coursera-dl, FLASK, will be used for this project.

5. **Please justify that the workload of your topic is at least 20*N hours, N being the total number of students in your team. You may list the main tasks to be completed, and the estimated time cost for each task.**

The tasks of this project and their corresponding workload are listed in the table below.

Task	**Workload (hours)**
Literature Review	2
Keywords Extraction Function Development	10
Keywords and Topic Indexing Function Development	10
Web Page Output Development	5
Test and Debug	8
Report and Demonstration	2
Total	**37**