

Arda Bedoyan

Project Proposal

CS 410: Text Information Systems

Genre Analysis of Movie Scripts

1. Team Members: Arda Bedoyan (bedoyan2@illinois.edu) – Team Arda
 - a. I chose to work alone given the constraints of working with a team when juggling a full-time job, family, and two classes. This will also allow me to truly dive into the material and understand every aspect of the project.
2. The track I chose is a free topic and will be genre analysis of movie scripts. By using the topic analysis techniques discussed in class, I aim to create a tool that will take as input a movie script and output the top genres that the movie falls under. I find this topic interesting and important, because often times movies are labelled with incorrect genres on various websites across the internet. Additionally, defining a genre is subjective, so by automating this process, then we can at least determine that all movies analyzed with this tool will be comparable. As opposed to multiple movies on another website that may all be tagged with genres determined manually by different people. I am a big fan of movies and find this topic interesting personally and want to work on a project that I can be excited about. This project is different than other projects that will be making a movie recommender system, because those systems may be utilizing genre tags from other websites, whereas I will be analyzing the genres independently of what other websites may tag the movie as. I plan to create a web application that will act as a database for

screenplays. Then when a movie title is entered, the program will analyze the script and assign topics to it, which will then be used to determine and output the genre of the movie. I plan to use tools like Django or Electron.js, python, and appropriate APIs that can help to create the application and the database needed. I will evaluate the success of my project empirically. I will test movie scripts that I am very familiar with and see if the resulting genres given by the program match my expectations.

3. I plan to use the python programming language to write the topic analysis part of my code. When it comes developing and running the web application portion, I may use a combination of Node.js, Electron.js, and JavaScript, or I may invest time in learning how to use Django and use that tool instead.
4. Given that I am working alone and that I do not have a strong background in topic analysis or web application development, this project will certainly take me more than 20 hours of work.
 - a. 3 hours – Topic, algorithm, and web application development research
 - b. 10 hours – Python code to generate topic/genre analysis
 - c. 5 hours – Script database creation
 - d. 10 hours – Front-end application development
 - e. 2 hours - Evaluation