Proposal

Title: Predicting the news articles that the user might be interested in reading based on user browsing.

Problem: How to engage users with user favorite news articles. Showing the articles of interest to a particular user will increase the user viewership/popularity of the website.

The goal is to implement an algorithm that will predict topics based on user reading patterns.

Who might care?: There are a lot of websites who would like to entice users to read articles. Those websites would like to increase traffic to a particular website with an increased number of viewership.

Data: The data is provided in a proprietary format in weblogs.

Modeling Approach: Topic modelling with cosine similarity is used to predict and show topics of similar interest to a particular user.

Possible Limitations: There could be limitations on the number of articles that can be displayed,for example if we limit to the number of articles displayed as 5, and if a user is browsing stories on sports and entertainment, we can display only 5 articles with both topics.

Deliverables:
1. Code for :
   - Data Acquisition
   - Data Cleaning
   - Data Exploration Analysis
   - Machine learning model development
2. Report/Presentation on the capstone project.