

## Team 'Buddie' Project Proposal

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.
  - a. Aman Gupta, amang5 (Captain)
  - b. Venkata Sandeep Chillara, vsc5
  - c. Katie Shin, ks56

2. Which paper have you chosen?

Causal topic modeling

Hyun Duk Kim, Malu Castellanos, Meichun Hsu, ChengXiang Zhai, Thomas Rietz, and Daniel Diermeier. 2013. Mining causal topics in text data: Iterative topic modeling with time series feedback. In Proceedings of the 22nd ACM international conference on information & knowledge management (CIKM 2013). ACM, New York, NY, USA, 885-890. DOI=10.1145/2505515.2505612

3. Which programming language do you plan to use?

Python

4. Can you obtain the datasets used in the paper for evaluation?

- a. Yes there are two datasets mentioned in the paper, a New York Times article dataset and a stock time series dataset. The NYT dataset is available here: <http://www ldc.upenn.edu/Catalog/ CatalogEntry.jsp?catalogId=LDC2008T19>. A stock time series dataset can be requested and pulled from here: <https://finance.yahoo.com/quote/AAPL/history/>

5. If you answer “no” to Question 4, can you obtain a similar dataset (e.g. a more recent version of the same dataset, or another dataset that is similar in nature)?

N/A

6. If you answer “no” to Questions 4 & 5, how are you going to demonstrate that you have successfully reproduced the method introduced in the paper?

N/A