

Project: Market Trend Analysis using News articles

Team: TeamAnand (individual)

Document Type: Progress Report

Github repository: <https://github.com/atodkar/CourseProject>

Progress made so far

1. Explored the REST APIs from [newsapi](#), [gnews](#) and [newsdata](#). Created a free account and generated API keys to access those public APIs (1 hour)
2. Read ReactJS and found it useful for the project. Accessed tutorials to understand the framework for the specific support needed in the project (1 hour)
3. Understood the PLSA, LDA and EM algorithm more in detail to be used for the assignment of topic extraction. Explored libraries which support these implementations and found gensim spacy and NLTK (3 hour)
4. Implemented a basic POC to verify the gensim (30 mins)

Remaining tasks

1. Need to build a web app using ReactJS and access REST APIs from server to show data as well as to show the relevant articles along with trends (12 hours)
2. topics extraction using the algorithms using gensim and other libraries on server side (3 hours)
3. Integration testing (4 hours)
4. Write documentation, prepare powerpoint slides, and record presentation (5 hours)

Challenges and Issues so far

1. Challenges to get the algorithm work correctly and being used over the REST endpoint. I see the algorithm is producing results but binding it to respond on the REST API is a bit of a gray area which needs to be tackled.
2. Spent a long time in understanding the core algorithm and its corresponding implementations available in the popular libraries. This is anyway a core part of the project and deserves more attention but I need to see that the completeness of the project is also dependent on the Web module.