**Project Progress Report**

Team: Best Bots

|  |  |  |  |
| --- | --- | --- | --- |
| **S.No.** | **Task** | **Current Status** | **Remarks/Impediments** |
| 1 | Access to DataSet by all members | Completed |  |
| 2 | Understanding the dataset | In-Prog | How to deal with gaps in dataset. Raised @1273 on Piazza |
| 3 | Find parallel time series dataset | Completed |  |
| 4 | Finding the libraries required for implementation | Completed |  |
| 5 | Construct document\_topic\_prob and topic\_word\_prob matrix for initiate PLSA | In-Prog |  |
| 6 | Apply PLSA to Documents to generate topics | In-Prog | POC done, Raised @**1251** on Piazza |
| 7 | Use Causality measure to find topics with significance values | In-Prog |  |
| 8 | Apply Granger Test to determine causality relationship. | In-Prog | POC done |
| 9 | For each candidate topic apply Causality measure to find most significant causal words among top words in each Topic. | In-Prog |  |
| 10 | Record the impact values of these significance words using Pearson correlations | In-Prog |  |
| 11 | Separate positive impact terms and negative impact terms | In-Prog |  |
| 12 | If orientation of words in prior step is very weak, ignore minor group | In-Prog |  |
| 13 | Assign prior probabilities proportions according to significance levels | In-Prog |  |
| 14 | Apply PLSA to Documents using prior obtained | In-Prog |  |
| 15 | Repeat until satisfying stopping criteria (e.g. reach topic quality at some point, no more significant topic change). | In-Prog |  |
| 16 | Generating performance evaluation output graph using Granger tests | In-Prog | POC done |