

CS 410 - Final Project Progress Report

Team PDT

Name	NetID
Darius Nguyen (Captain)	huy2
Pritom Saha Akash	pakash2
Trisha Das	trishad2

1. Steps completed:

- Dataset collection: Both text and time series datasets are collected.
- Text data preprocessing: We have preprocessed text data for using it into the PLSA algorithm.
- Topic modeling: We have applied the PLSA algorithm to find topics from text data.
- Causal topic filtering: We have applied the granger test for finding significant casual topics based on time-series data.
- Word-level causality modeling: We have implemented the part for filtering causally related words (positive and negative) for each causally significant topic.

2. Steps outstanding:

- Topic prior generation
- Apply topic prior feedback to topic modeling

3. Challenges:

- We are facing some problems in understanding which distribution(negative or positive or both) for each causally significant topic to apply as prior to topic modeling.