Project Brief

Texts of Trade Agreements:

https://github.com/mappingtreaties/tota/tree/master/xml (Dataset)

http://mappinginvestmenttreaties.com/rta/

https://unctad.org/topic/trade-analysis/text-as-data

Outline and Objectives:

Examining the text of trade agreements may give insights into how the trade deal will play out. This project aims to use machine learning methods to determine the effects of a trade agreement on trade between the involved countries, and which features of a trade agreement are most important.

Approach:

The trade agreements will be broken down into different parts and important information like countries and start dates will be extracted. We will work with trade data to see how trade differs after each agreement and create a dataset with the desired information for each trade agreement. We must be careful to measure the effectiveness of the agreement whilst controlling for general economic factors, like recessions. We will then train neural network models on the data and experiment with different tabular models to achieve our objectives using existing libraries such as PyTorch and fastai.

Data Sources:

Texts of trade agreements:
https://github.com/mappingtreaties/tota/blob/master/xml/pta_1.xml

- Includes the countries involved, date signed, date effective, text body

Trade data:

- Will include amount of trade, perhaps an index of improvement in trade somehow?

General macroeconomic variables:

- Will include GDP and other general indicators

Timeline:

- Data preprocessing and exploration (1 week)

- Feature selection and dimensionality reduction (2 weeks)

- Building, training and selecting models (4 weeks)

- Write up and report findings (1 week)