Meeting Notes

**Thursday September 10th 11:30am – 1pm**

* Joined in group during lab time
* Discussed meeting times
* Organised GitHub repository
* Discussed datasets and agreed to individually research datasets and come with idea to next meeting
* Next meeting scheduled for 7pm Monday

**Monday September 14th 7pm-820pm**

Discussed two potential datasets and the machine learning practices we potentially will use.

Movie recommender system

Fake news detection

* Chose this idea

Datasets

* Labelled tweets from @politifact
  + Twitter data not usable in its form
* News articles
  + Fake news datasets from Kaggle, true news dataset from Signal Media

Potential Methods

* Reliability of source
* Sentiment analysis
* Deep learning
  + Neural network
* GANs <- a neural network to try trick our own neural network

Learnt how to pull and push with Github

To Do:

* Work on basics of proposal
* More research into datasets and machine learning methods
* Scheduled meeting for Thursday night to write methods and finalise proposal.

**Thursday September 17th 11:30am – 1pm**

* Met Luke in person!

Worked on project proposal

* Brought forward researched ideas since last meeting; datasets and method ideas
* Generated more analysis ideas
* Outlined and explored more method ideas as a group
* Went through datasets, deciding on which to use
* Defined visualisations for analysis and evaluation techniques
* Went through and refined and edited writing for datasets and task
* Organised method

To do:

* Individually write sentences for proposal supporting and referencing the reasons for method choices
* Formatting
* Meeting planned for 7pm tonight (17/09) to finalise proposal

**Thursday September 17th 7pm-7:35pm**

* Met and added final sentences that we had worked on individually
* Read through the proposal together to proofread
* Submitted proposal

To do:

* Meeting planned for next Thursday (24/09) during lab time

**Thursday September 24th 11am-1pm**

* Discussing the timeline for the project
  + Pre-processing sept 30/oct 1st
  + Models 12/13th Oct
  + Tweaking model, Evaluation, analysis 19/20th Oct
  + Video on 23rd Oct
  + Report on 30th Oct
* Assigning roles to each group member: pre-processing (Krystal and Laura), embeddings (Luke) and classifiers (ML- Laura, NN- Charlotte and Krystal).
* Going over the datasets to determine which are best suited
  + Confirming datasets and features.
  + Not including source in the model to focus on the text analysis, removes bias in the model based on the source. Potentially more interesting just looking at NLP.

To-Do:

* Pre-processing assigned tasks
* Embeddings
* Setting up google co-lab
* Next meeting Friday 25 12:45pm to discuss co-lab, and Monday 28 1pm to discuss completed pre-processing.

**Thursday September 24th 11am-1pm**

* Charlotte explained how to use google colab.

To-Do:

* Pre-processing assigned tasks by Laura and Krystal
* Next meeting Monday 12:39pm to discuss completed pre-processing.

**Monday September 28th 12:30pm-1pm**

Checked in to cover and explain work that had been done since previous meeting.

Laura- Edited all datasets to merge together in correct format, some duplicate and missing vals

Krystal- Data cleaning for outliers + noise, missing / incorrect vals + duplicates

To-Do:

* Embeddings by Luke
* Next meeting Friday 2nd Oct 8pm to discuss completed embeddings

**Friday October 2nd 8pm-8:15pm**

* Luke explain the embeddings he had done, doc2vec and pytorch; to be fully completed by Sunday

To-Do:

* Classifiers by Laura, Charlotte, and Krystal
* Next meeting Monday 12th to go over completed work