22/10/2019 Meeting 3

Aim: To further discuss the literature survey and project plan

1. Do you have a preference in terms of American or English spelling?

- e.g. analysers vs analyzers
- English is better, but pick one and stick to it (be consistent)

2. Could we talk about some potential research domains of machine and deep learning?

 Look specifically for machine learning for NLP applications, not all algorithms are suitable for NLP - in summary, anything is useful if it's machine learning for NLP
 Convolutions are rarely used in NLP, mainly for images
 RNN - Recurrent neural networks
 LTSM - Good for "time-series data"
e.g. words are an example of **time series data** (their order affects their meaning)

3. Is F1 score a good enough method of evaluation - is it the state-of-the-art or should I consider other approaches

 F1 score is a good method of evaluation
 Sarcasm is a minority class
 F1 score takes overfitting into account

4. Minimum size of a dataset

 In the thousands, low thousands if they are accurate

5. Discuss the SARC dataset

- Allows for balanced and unbalanced data because there are so many data points
- Can the supercomputer cope with this volume of data? **No**
- SARC is fairly of noisy
- Getting own dataset by capturing tweets with #sarcasm

 In the paper, a small section will be Included on selecting a dataset
 For pre-training, it could be useful to use a big dataset
 Then, use a better quality smaller dataset for fine-tuning
 We ay work with multiple datasets
Try to pair dataset modelling and results

Corpus	Dataset	Sarcastic	Total
IAC	Joshi et al. '15	751	1502
	Oraby et al. '16	4.7K	9.4K
Twitter	Joshi et al. '16	4.2K	5.2K
	Bamman & Smith '15	9.7K	19.5K
	Reyes et al. '13	10K	40K
	Riloff et al. '13	35K	175K
	Ptáček et al. '13	130K	780K
Reddit	Wallace et al. '15	753	14124
	SARC	**1.34M**	**533M**