**Measuring Medical Sentiment on Social Media**

**Master Research project**

**Student: Chen Wang**

On patient forums, patients discuss valuable information but also their own experiences and opinions on their disease. Sentiment analysis of these social media posts could reveal whether clinical outcomes or experiences are positive or negative in order to judge the impact of the medical condition on the patient’s overall wellbeing.

Sentiment analysis in the medical domain and on social media in particular offers distinct challenges. Most research into medical sentiment analysis has focused on analyzing clinical records, which are often written by nurses or doctors (Denecke & Deng, 2015). However, there are distinct differences between clinical records and patient forums that impact sentiment analysis for e.g. word usage is more subjective (Denecke & Deng, 2015).

Recently, Carrillo-de-Albornoz, Rodríguez Vidal, & Plaza (2018) has developed a dataset (eDiseases) for assessing medical sentiment in social media. They employ three classes: positive, negative and neutral. Importantly, they make a distinction between the sentiment of facts (so-called *polar facts),* opinions and experiences. Facts can have different polarities in the medical domain, as they can have positive or negative implications for patients. Their research in fact combines two traditional sentiment tasks: subjectivity analysis and polarity classification. Subjective analysis refers to classifying text into objective or subjective. Polarity classification refers to classifying text into having positive or negative connotations. Carillo-de-Albornoz *et al.*  are also the first to look extensively at which features (i.e. lexical, semantic, syntactic, word embeddings etc.) contribute to classifying sentiment in user-generated texts from patient forums. However, they only use simple machine learning algorithms.

Within this field of medical sentiment analysis, transfer learning based methods have not been applied as of yet, although they are the state-of-the-art in general sentiment classification. Thus, in this research project, you will investigate the application of transfer learning techniques to medical sentiment analysis, both for determining polarity and subjectivity of the social media posts of the eDiseases dataset.

Transfer learning methods rely on a language model pre-trained on a large corpus of text which is then fine-tuned on a domain-specific or task-specific body of text. Transfer learning is likely to be beneficial for this task, because the data sets are small but specialized. XLNet is the current state-of-the-art transfer learning approach for sentiment analysis. It is an attention-based model. This aspect could be useful for helping us to understand which features contribute to the sentiment classification.

**Preliminary research question:**

To what extent can transfer learning aid sentiment classification in medical social media?

**Possible sub-questions:**

* How does transfer learning compare to other methods for medical sentiment classification?
* Under which conditions does transfer learning optimally classify sentiment in medical social media and why?
* How is sentiment classification affected by considering subjectivity and polarity analysis as two separate steps or one single step? Or by their order in which they are performed?
* How is sentiment classification affected by considering the three diseases separately or together for fine-tuning the pre-trained model ?
* To what extent are the features found with transfer learning comparable to those found by Carillo-de-Albornoz *et al.* ?

**Dataset:** eDiseases data set (Carrillo-de-Albornoz et al., 2018)

**Details on the eDiseases dataset**

<https://zenodo.org/record/1479354#.XSyBgegzZPY>

The eDiseases dataset contains patient data from the MedHelp health site (http://www.medhelp.org/), where different communities share information and opinions about diseases. Each community consists of a number of conversations; a conversation being a sequence of comments posted by patients.

To build the dataset, we automatically extracted 10 conversations from each of the following three communities: allergies, crohn and breast cancer. We selected a set of diseases that, according to medical expert, show high heterogeneity concerning both the degree of medical understanding of the diseases and the profile of the users. The conversations were selected randomly, but we automatically filtered out conversations with less than 10 posts. In total, we extracted 146 posts for allergies, 191 posts for crohn, and 142 posts for breast cancer; which include 983 sentences for allergies, 1780 sentences for crohn, and 1029 sentences for breast cancer, covering a 6 years time interval. Three frequent users of health forums annotated each sentence in the dataset as:

Factuality: OPINION, FACT, EXPERIENCE.  
Polarity: POSITIVE, NEUTRAL, NEGATIVE.

In case of doubt, the annotators labeled the sentence as NOT\_LABELED. As a result, we collected 967 labeled sentences for allergies, 1,709 labeled sentences 294 for crohn, and 959 labeled sentences for breast cancer.

**Interesting blog post as an introduction to transfer learning**

<http://jalammar.github.io/illustrated-bert/>

**and to XLNet**

<https://mlexplained.com/2019/06/30/paper-dissected-xlnet-generalized-autoregressive-pretraining-for-language-understanding-explained/>

**Articles:**

Denecke, K., & Deng, Y. (2015). Sentiment analysis in medical settings: New opportunities and challenges. *Artificial Intelligence in Medicine*, *64*(1), 17–27. <https://doi.org/10.1016/J.ARTMED.2015.03.006>

Carrillo-de-Albornoz, J., Rodríguez Vidal, J., & Plaza, L. (2018). Feature engineering for sentiment analysis in e-health forums. *PloS One*, *13*(11), e0207996. <https://doi.org/10.1371/journal.pone.0207996>

Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R., & Le, Q. V. (n.d.). *XLNet: Generalized Autoregressive Pretraining for Language Understanding*. Retrieved from https://github.com/zihangdai/xlnet

Howard, J., & Ruder, S. (n.d.). *Universal Language Model Fine-tuning for Text Classification*. Retrieved from http://nlp.fast.ai/ulmfit.