

Text entailment classification

Valle, Damian
s182091@student.dtu.dk

Barkat, Sacha
s182045q@student.dtu.dk

November 1, 2018

1 Introduction

NLI problem goal is to determine whether a natural language hypothesis h can justifiably be inferred from a natural language premise p .

One of the issue of this problem is the difficulty to label the data (even for human).

Considering for example the pair *A boat sank in the Pacific Ocean* and *A boat sank in the Atlantic Ocean*. We might labeled this pair as a contradiction if we assumes that they refer to the same single event, but on the other hand we could also labeled it as neutral if that assumption is not made.

Selecting one of this approach is a difficult choice Indeed if we opt not to assume that events are coreferent, then we will only ever find contradictions between sentences that make broad universal assertions, but if we opt to assume coreference, new counterintuitive predictions will emerge (add other examples here lol)

Using SNLI (The Stanford Natural Language Inference), we hope to adress these kinds of issues. In this method we are making sample based on the scenario to not have the same doubts as those we listed in the previous examples.

2 Related Work

3 Methods

4 Experimental setup

References

- [1] Samuel R. Bowman, Gabor Angeli, Christopher Potts and Christopher D. Manning. *A large annotated corpus for learning natural language inference*. Stanford Natural Language Processing Group, 2015.
- [2] Samuel R. Bowman, Jon Gauthier, Abhinav Rastogi, Raghav Gupta, Christopher D. Manning, Christopher Potts. *A Fast Unified Model for Parsing and Sentence Understanding*. Stanford Linguistics
- [3] Shuohang Wang and Jing Jiang. *Learning Natural Language Inference with LSTM*. Information Systems, Singapore Management University 2016
- [4] Tomas Mikolov, Kai Chen, Greg Corrado and Jeffrey Dean. *Efficient estimation of Word Representations in Vector Space*. Google Inc., Mountain View, CA, 2013
- [5] Jeffrey Pennington, Richard Socher and Christopher D. Manning *GloVe: Global Vectors for Word Representation*.
- [6] Chris McCormick. *Word2Vec: The Skip-Gram Model*. <http://mccormickml.com/2016/04/19/word2vec-tutorial-the-skip-gram-model>
- [7] Sepp Hochreiter and Jurgen Schmidhuber. *Long Short Term Memory*. Technische Universitat Munchen.