**Determining Inference Using the MultiNLI Dataset** Chris Huber

CSC 820, Natural Language Technology 02/26/2022

Prof. Anagha Kulkari

**Abstract**

This paper describes the process of classifying inference from sets of sentences. There are several models capable of this including RoBERTA[[1]](#footnote-1) and XLNet[[2]](#footnote-2) which can be fine-tuned to produce increasingly better accuracy.

**Background**

Natural language inference revolves around determining whether a hypothesis is an entailment, contradiction, or neutral. For example:

|  |  |  |
| --- | --- | --- |
| Premise | Label | Hypothesis |
| A man inspects the uniform of a figure in some East Asian country. | contradiction | The man is sleeping. |
| An older and younger man smiling. | neutral | Two men are smiling and laughing at the cats playing on the floor. |
| A soccer game with multiple males playing. | entailment | Some men are playing a sport. |

[[3]](#footnote-3)

The Mulit-Genre Natural Language Inference (MNLI) corpus has approximately 433,000 pairs of hypotheses and premises to test with. It spans a range of genres of spoken and written text and has cross-genre comparison sets. There is also a SciTail corpus which is slightly less curated.

**Implementation**

A preliminary test against the train dataset using RoBERTa computed a 0.3123 accuracy or slightly worse than guessing.

1. https://arxiv.org/pdf/1907.11692.pdf [↑](#footnote-ref-1)
2. https://arxiv.org/pdf/1906.08237.pdf [↑](#footnote-ref-2)
3. http://nlpprogress.com/english/natural\_language\_inference.html [↑](#footnote-ref-3)