# Hai Hu 02-19-2020

# Building a natural language inference dataset in Chinese

## What is NLI?

when you have to determine whether a hypothesis contradicts, entails from or is neutral towards a premise.

## Issues with SNLI

Turkers do not want contradiction to go both ways.

## Bias in hypotheses

If you train on SNLI on just the hypotheses, you get better than majority baseline.

**There's bias in the hypotheses** One thing is that sleeps contradicts almost any other action. Additional heuristics in the dataset probably introduced by the Turkers probably exist. By creating synthetic data that goies against the heuristics, the result is very very poor performance (19% accuracy for BERT was the best).

## XNLI:

- 15 languages

- translated from SNLI/MNLI
    - bad quality translation, lots of things that just don't translate well

# Our chinese NLI

- undergrads instead of turkers
- told to write 3 neutral, 3 contradiction, 3 entail as a way of getting them to introduce more variety.
- Students still apply heuristics.
- Issues that emerged:
    - phone call transcriptions are bad
    - use of questions in premises was confusing

# Todo

- how to get more variation in hypotheses?
- one annotator only writes Entailments not C/N

---