Paper reports – Mohammad Osoolian

Paper Title: WINOGRANDE: An Adversarial Winograd Schema Challenge at Scale

Paper Link: [1907.10641] WinoGrande: An Adversarial Winograd Schema Challenge at Scale (arxiv.org)

Submit Date: 21 Nov 2019

What is paper about: This paper introduces a new dataset for WS challenge with 44 samples and an algorithms called AFLite for Adversarial Filtering of samples

Abstract:

WinoGrande is An Adversarial Winograd Schema Challenge at Scale" introduces WinoGrande, a large-scale dataset of 44,000 pronoun resolution problems inspired by the original Winograd Schema Challenge.

Background:

- WSC
- AF
- Annotation Artifacts

Challenge:

State-of-art models in WSC have reached the human accuracy on the provided datase. So there is a need for a new WS dataset that be bigger and more complex and de-biased.

There are other NLP datasets that have Annotation Artifacts and the previously presented algorithms like AF are not general and need to be redesigned for every dataset.

New Ideas:

The key steps in the construction of WinoGrande are large-scale crowdsourcing, followed by systematic bias reduction using a novel AFLITE algorithm that generalizes human-detectable word associations to machine-detectable embedding associations.

Results:

authors show that state-of-the-art models achieve considerably lower accuracy (59.4%-79.1%) on WinoGrande compared to humans (94%), confirming that the high performance on the original WSC was inflated by spurious biases in the dataset.

Also AFLite is introduced as an Adversarial Filtering algorithm that is faster and more general in contrast to AF.

My Idea for the challenge:

Collecting more samples for WC challenge and creating bigger database.

My Idea to improve this article:

Trying to use AFLite in other fields of NLP like NLI and Question Answering.