

Starting Small, After All? Curriculum Learning with Child-Directed Speech

Mattia Opper

University of Edinburgh, Edinburgh, Midlothian, United Kingdom

Sydelle de Souza

University of Edinburgh, Edinburgh, Midlothian, United Kingdom

Abstract

The idea of curriculum learning, whereby a model is first exposed to simpler examples before an increase in complexity, has long fascinated the AI community. Unfortunately, the experimental successes of curriculum learning have been mixed, particularly applied to natural language, where a vast body of literature appears to evidence its failures. However, recent work has shown that language models trained on transcribed-child-directed-speech (CDS) learn more grammar compared to those trained on Wikipedia. To a lesser extent, the same trend has been observed through training on transcribed speech and simple text data. Motivated by these findings, we revisit the idea of curriculum learning starting from CDS, before moving to simple data, and finally finishing with complex long form text. Unfortunately, through experimentation with an array of models and training step sizes, only in the smallest models trained for the least steps does curriculum learning show any advantage over random sampling.