UC Merced

Title

The Omniglot Jr. challenge; Can a model achieve child-level character generation and classification?

Permalink

https://escholarship.org/uc/item/0n09g2ck

Journal

ISSN

Authors

Kosoy, Eliza
Belyi, Masha
Snell, Charlie Victor
et al.

Publication Date

2021

Peer reviewed

The Omniglot Jr. challenge; Can a model achieve child-level character generation and classification?

Eliza Kosoy
UC Berkeley , Berkeley, California, United States

Masha Belyi
UC Berkeley, Berkeley, California, United States

Charlie Snell
UC Berkeley, Berkeley, California, United States

Brenden Lake
NYU, New York, New York, United States

Josh Tenenbaum
MIT, Cambridge, Massachusetts, United States

Alison Gopnik
University of California at Berkeley, Berkeley, California, United States

Abstract

Lake 2015 presented the Omniglot dataset to study how people generate and classify characters. They proposed a model for one-shot-learning of new concepts based on inferring compositionally structured generative models, and could transfer from familiar concepts to new ones. Their Bayesian-Programming-Language model was able to both classify and generate characters similar to adults. However, adults have years of experience, would a similar model apply to children without this prior knowledge? We introduce a new dataset called Omniglot_Jr., composed of Omniglot letters generated by children aged 3-6. Our results of training BPL with children's data in classification and generation find that BPL achieves higher classification accuracy and generates more adult looking letters. We propose the challenge of reproducing children's distinctive pattern of mistakes. Skills such as character recognition depend on child-like learning; this challenge should help us understand how that learning is possible, and how to simulate it in computational systems.