Human Created: Exploring the Effects of Strict Government Policy for Training Generative Machine Learning Systems.

Justin Cousins
Faculty of Computing Science
Northern University
District 2, New-Canada
jc.2014@northu.ca

Abstract

In 2035, humanity reached a tipping point where more content had been generated by artificial intelligence than humans throughout history. Soon after, next generation Generative Machine Learning (GML) systems began to be trained completely on machine generated content, producing state of the art results. This caused immense concern among both researchers and politicians as it became glaringly apparent that humans were training GML systems that were ultimately rendering human creativity obsolete. In an effort to combat this, governments across the globe agreed to make it a requirement that all new GML systems be trained using data that could be identified as "clearly human created", effectively banning the use of models being trained on data that they themselves had previously generated.

This paper explores the ethics of the so called "human created" law by asking two closely related questions:

1. Does enforcing a "human created" restriction on training data for generative models actually foster human creativity in any meaningful way?

2. Why shouldn't generative models be allowed to train on their own output? Is the act of taking inspiration from previously created work not what humans have been doing for thousands of years?

We provide answers to these questions through an analysis of current training data usage, a review of existing literature [1, 2], and discussions with leading experts in the GML space. This paper aims to provide a comprehensive background for decision makers to lean upon when tasked with defining the future of ethics for generative systems in creative fields.

Keywords

Generative Machine Learning, Data, Human Creativity, Ethics.

References

[1]: R. Nylander and W. O'Reilly, *Human Creativity in an increasingly computational world*, ICCC'34

[2]: S. Tavares and J. Keefe, *Content Matters: exploring the need for human generated content*, *ICC*C'29