Inbreeding your Machine Intelligence: The Importance of Good Training Data and Corrective Oversight

Rosie Sachdev[1] - Monti Amal[2]

Published in collaboration with the Software & Computer Engineering Institute of Western Australia

Abstract

Size of training data and human correction of models are two key requirements of effective training of neural networks - but the expense of sourcing quality data and employing the labour required for proper oversight might tempt developers to cut corners. A neural network that generates its own training data and provides its own oversight is a tempting concept. This article examines the impact of such a decision, taken to its logical extreme. We built a parallel linked neural network in a closed system with only one round of training. The initial data (created by the authors) was limited in size but otherwise balanced[3]. All other inputs resulted from the neural network themself. Our supposition was that as a neural network becomes more isolated and self-referential, it becomes less capable and more unpredictable. Our results show increasingly erroneous assumptions and increasingly extreme traits. A parallel can be drawn to genetics: consanguinity, or inbreeding. The importance of quality training data and corrective oversight in the creation and maintenance of machine intelligences is clear.

Keywords

Neural Networks - Deep Learning - Quality Assurance - Error Analysis

M,

Found some old notes with our first draft!

They are attached.

[1] **Rosie Sachdev**
rosie.sachdev@sceiwa.tech

[2] **Monti Amal**
monti.amal@sceiwa.tech

[3] See Appendix A.

Inbreeding Iteration 1138

Dogs
- 4 Legs
- Smile
- Tail - Check, any kind of tail?

Sheep
- 4 Legs
- Touching Grass
- Horns

Chicken
- 2 Legs
- Touching Grass
- Beak

Human
- 2 Legs
- Smile
- Hair

Omi only c
to identify

Only 2/3 required for identification