NEURAL NETS AND UNIVERSAL ALGEBRA PROJECT DESCRIPTION

CHARLOTTE ATEN

As discussed in chapter 20 of Understanding Machine Learning by Shai Shalev-Shwartz and Shai Ben-David[3] as well as in, for example, the pioneering work of Cybenko in the late 1980s[2], neural nets are able to approximate a wide range of functions arbitrarily well under certain assumptions on their architecture and activation functions.

Similar results had already appeared in the universal algebra[1] literature in the 1960s and 1970s. In particular we have the following theorem, discussed in chapter chapter 6 of [1]:

Theorem 1 (Murskiĭ, 1970s)**.** *Let P be the property that an algebra is primal. If ρ is a similarity type which contains at least two basic operations, at least one of which is not unary, then* $\mathrm{Pr}_\rho(P) = 1$.

The content of this theorem is that if we take a sufficiently large finite set A and a randomly-selected finite collection of functions $f_1, f_2, \ldots, f_k$ of the form $f_i \colon A^{n_i} \to A$ where $k \geq 2$ and at least one n_i is at least 2, we will find that, with probability 1, any function of the form $g \colon A^m \to A$ can be written as a composite of the randomly-chosen functions. In other words, a randomly chosen finite set of finitary operations almost certainly has the maximum possible capacity to express other functions.

If we reword a little bit so that our set A becomes data input into a neural net and those operations f_i become activation functions, then Murskiĭ's Theorem takes on the following form:

Theorem 2 (Murskiĭ, 1970s, interpreted for neural nets)**.** *If ρ is a similarity type which contains at least two basic operations, at least one of which is not unary, then a randomly-selected finite algebra $\mathbf{A}$ of signature ρ has (with probability 1) the property that given any target function $h \colon A^n \to A^m$ there exists a discrete neural net $(V_1, \ldots, V_r, E, \Phi)$ whose activation functions are all basic operations of $\mathbf{A}$ or projections which represents h.*

In keeping with the theme of this REU, we would like to bridge the gap between this existing theory and practice. See this video or the accompanying slides for more information on the formulation of neural nets used here. Specifically, this project would address questions such as:

(1) Are discrete neural nets capable of performing learning tasks to a level at all comparable with traditional feed forward neural nets using gradient descent and sigmoid activation functions?

Date: 2021 july 25.

This is a research project description for the the 2021 TRIPODS/STEMforAll REU program.

[1]Universal algebra is a discipline within abstract algebra and logic which is concerned with the study of general algebraic structures and classes thereof. Clifford Bergman has a good introductory text on the subject[1].

(2) How many layers are typically needed before a randomly-chosen set of activation functions can represent any function?
(3) How does the depth of a neural net with randomly-chosen activation functions affect its performance on learning tasks?
(4) Do specific finite algebras of interest, like the finite fields of prime order $\mathbb{F}_p$, perform better or worse than randomly chosen ones at common learning tasks?

The initial stages of the project would include going over the relevant theory to the extent which time allows and building tools for experimenting with this variant of the neural net concept.

References

[1] Clifford Bergman. *Universal Algebra: Fundamentals and Selected Topics*. Chapman and Hall/CRC, 2011. ISBN: 978-1-4398-5129-6 (cit. on p. 1).
[2] G. Cybenko. "Approximation by Superpositions of a Sigmoidal Function". In: *Math. Control Signals Systems* 2 (1989), pp. 303–314 (cit. on p. 1).
[3] Shai Shalev-Shwartz and Shai Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. 32 Avenue of the Americas, New York, NY 10013-2473, USA: Cambridge University Press, 2014. ISBN: 978-1-107-05713-5 (cit. on p. 1).

Mathematics Department, University of Rochester, Rochester 14627, USA
URL: aten.cool
E-mail address: charlotte.aten@rochester.edu