Abstract

Kyushu Institute of Technology Graduate School of Information Engineering, Department of Interdisciplinary Information Informatics, Honda Lab

Student Number	19677504-1	Name	Alexandre Louvet
Title	Neural network expressiveness		

I   Introduction

With the current use of neural networks to solve many problems ranging from pattern recognition to medical diagnosis, the necessity to get an understanding on the theory behind these algorithms grew. Neural network expressiveness is the field that aims to provide tools to better understand the efficiency of the neural networks. It aims to use topological tools to categorize the different classes that can be approximated by a certain neural network architecture. In addition to that we used topological data analysis, a group of computing tools obtained from classical topology, to try to give a quantitative understanding view on networks expressiveness using topological data analysis and especially Betti numbers as the measure to do so.

II   Mathematical Definitions

For the purpose of our research we need to introduce various mathematical concepts starting with homology. Let X be a topological space. The purpose of homology is to distinguish topological features of X. Visually this concept is assimilated to the concept of "holes". Suppose X is of dimension $n \in \mathbb{N}$. The betti numbers of X will be an element of $B \in \mathbb{N}^n$ with $(B_i)_{1 \leq i \leq n} = \{$The number of i-dimensional homology in $X\}$.

Simple examples of homology are the balls and spheres. A topological ball in X of radius r and center $c \in X$ is: $\mathcal{B}(c,r) = \{x \in X \mid ||x,c||) \leq r\}$ with $|| \cdot || : X^2 \mapsto \mathbb{R}_+$ a distance on X. The topological sphere in X of radius r and center $c \in X$ is: $\mathcal{B}(c,r) = \{x \in X \mid ||x,c||) = r\}$. A result of homology study is that a topological ball of dimension m is a 0-homology for $m \geq 1$ and a topological sphere of dimension $m \geq 2$ is an $(m-1)$-homology.

The computational analogue of homology called persistent homology aims to compute the homology of discrete datasets. Persistent homology relies on the computation of topological feature using the computation of structures called Δ-complexes in the dataset. These are the smallest polygons containing $n+1$ points not lying in an hyperplane of dimension less than n. An example of it is the pyramid when $n = 3$.

The last concept we will need is the

mathematical knot. A knot is an embedding of a circle in a 3-dimensional euclidean space $\mathbb{R}^3$. One of the critical question of knot theory is to universally discriminate knots. Though this question has no answer today, various discriminants have been created by mathematicians, the simplest one being the knot determinant. It is an integer that can be computed from the components and crossings of a knot and is equal for two representations of the same knot.

III   Experiments

Relying on the works of Guss [1] and Raghu [2], we propose two experiments. The first one aims to understand from a knot theory perspective the evolution of knot trajectories inside neural networks of various depth. To do so we input points array representing circles, trefoils and figure-eight knots inside neural network and look at their evolution at each layer of the networks.

The second one aims to give a quantitative evaluation of the capacity of neural networks of different depth to properly characterize datasets depending on their topological features (homology). To do so we generate datasets with points classified in two classes and train neural networks on them. The first one random points forming the "background" of the dataset and the second one representing various homologies. We then compare the topological feature of the original data and the datasets obtained by

making neural networks evaluate an unlabelled set.

IV   Results

The results of the first experiment is that all trajectories will, with enough layers, end up as entangled unknots. The number of experiments ran is greatly insufficient to come to any general conclusion and this could lay ground to larger experiments to find the specificity of various neural networks architecture.

The second experiment saw various computational problems and would need more data to be relevant. A general tendency is as expected that deeper and wider networks have greater chances to properly characterize a higher homology dataset.

V   Short Bibliography

References

[1] W. H. Guss and R. Salakhutdinov, "On Characterizing the Capacity of Neural Networks using Algebraic Topology," *arXiv:1802.04443 [cs, math, stat]*, Feb. 2018. arXiv: 1802.04443.

[2] M. Raghu, B. Poole, J. Kleinberg, S. Ganguli, and J. Sohl-Dickstein, "On the Expressive Power of Deep Neural Networks," *arXiv:1606.05336 [cs, stat]*, June 2017. arXiv: 1606.05336.