Giovanni Petri and António Leitão

Topological Expressive Power Of Neural Networks

ISI Foundation & ISI Global Science Foundation

NOVA IMS Information Management School

We propose a topological description of neural network expressive power. We adopt the topology of the space of decision boundaries realized by a neural architecture as a measure of its intrinsic expressive power. By sampling a large number of neural architectures with different sizes and design, we show how such measure of expressive power depends on the properties of the architectures, like depth, width and other related quantities.

What is expressive power?

Expressive power is the number of different problems a Neural Network can solve. In our work we restrict ourselves to binary classification problems.

How is it measured?

We evaluate how **complex** and how **different**, topologically, are the decision boundaries that a given Neural Network architecture can express.

We measure how **spread out** is the space of persistence diagrams of the decision boundaries that an architecture can express.

1. Sample the space of Neural Networks

Consider a Neural Network architecture F of n parameters. It is trivial to see that there is a direct mapping from F to $\mathbb{R}^n$. We can therefore explore F by sampling $\mathbb{R}^n$.

In this poster we show results for architectures of 1-10 hidden layers with 5, 7, 9 and 10 neurons. For each we uniformly sample a set of 2000 parameters in $[-1, 1]^n$

For each parameter vector there exists a Neural Network and each Neural Network describes a decision boundary. Our work aims at mesuring expressive power by evaluating how "topologically diverse" are the decision boundaries, given a uniform distribution of parameter vectors.

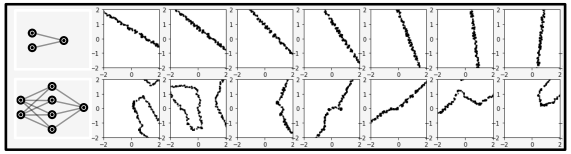


Figure 1 : The decision boundaries of two paths in the parameter space of two different architectures. A very simple one (top row) and one with one hidden layer (bottom row). Note how the simple architecture can only produce decision boundaries with trivial homology, no matter the parameter values. In our methodology the bottom row architecture is more expressive since the space of its decision boundaries is more "topologically diverse".

2. Evaluate the persistent homology of each decision boundary

2.1 Approximating decision boundary: The decision boundary for a given problem is not unique. We consider here one that maximizes the distance between each class. This is reasonable since any other disparity metric (for example support) is based on distance. Under this assumption, the decision boundary we aim to approximate is the union of the edges of adjacent Voronoi cells corresponding to points of different classes and is unique.

2.2 Topological properties of decision boundaries: For a given architecture we sampled 2000 parameter vectors. Each parameter vector defines a decision boundary captured in the step above. We compute its persistent homology. We do it using the python package Ripser (Tralie et all. 2018).

Since the input space is always 2 dimensions we compute only H_0 and H_1 homology.

For each architecture we now have associated 2000 persistence diagrams. We consider their pairwise Wasserstein distance and construct a metric space for each architecture.

Sampling Decision Boundary

We use a novel method to sample points from the decision boundary of two classes A and B. Our method is faster and more scalable than the one proposed by Ramamurthy et all 2018.

Our method uniformly samples points from the edges of Voronoi cells belonging to points of different classes.

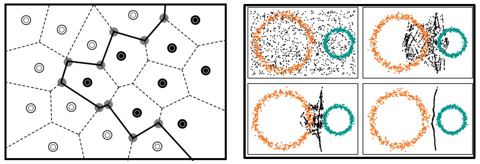


Figure 2: Left is an illustration of the decision boundary, the edges of Voronoi cells belonging to points of different classes (black and white). Right is an example of our method of sampling 1000 points (black) from this boundary given two classes (orange and green).

Bibliography

Simon Willerton. Spread: a measure of the size of metric spcaes 2015.
Leland McInnes, John Healy, and James Melville. Umap: Uniform Manifold Approximation and Projection for dimensionality reduction 2018.
Karthikeyan Natesan Ramamurthy, Kush R Varshney, and Krishnan Mody. Topological data analysis of decision boundaries with application to model selection. 2018
Ben Poole, Subhaneil Lahiri, Maithra Raghu, Jascha Sohl-Dickstein, and Surya Ganguli. Exponential expressivity in deep neural networks through transient chaos. 2016
Guido Montúfar, Razvan Pascanu, Kyunghyun Cho, and Yoshua Bengio. On the number of linear regions of deep neural networks, 2014.
Christopher Tralie, Nathaniel Saul, and Rann Bar-On. Ripser.py: A lean persistent homology library for python. 2018

What is spread?

Spread introduced by Willerton 2015 is a notion of diversity for metric spaces. For a metric space (X, d) spread is defined as

$$E_0(tX) = \sum_{x \in X} \left(\sum_{y \in X} e^{-td(x,y)} \right)^{-1}$$

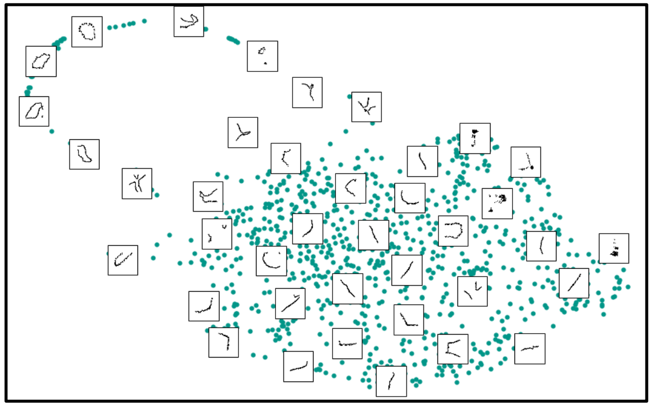
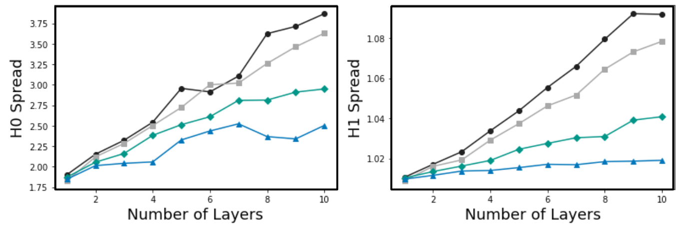


Figure 3: Embedding of the metric space (P, W_p) where P is the space of persistence diagrams of the decision boundaries obtained from different parameter vectors. Along with the Wasserstein distance between the persistence diagrams. The embedding into 2 dimensions was done using UMAP (McInnes et all 2018).

3. Measure the spread of the space

We then measure the spread (Willerton 2015) of each metric space associated with each architecture.

Our results show (among others) that: 1. Spread is weakly correlated with the number of parameters in the architecture. 2. Spread grows with depth and width, although not exponentially as observed in geometric properties by (Poole et all 2016). 3. Spread has complex correlations with other architecture parameters such as the presence of a bottleneck.



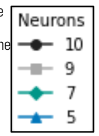


Figure 4: Spread values (y-axis) for different architectures. Each architecture always has the same number of neurons per layer. On the left we have the H_0 spread, that is the spread of the metric space constructed from the 0-dimensional persistence diagrams of the decision boundaries and the Wasserstein distance between them. On the right we have the H_1 spread which is constructed analogously.

https://openreview.net/forum?id=I44kJPuvqPD