Any-Dimensional Invariant Universality

Shengtai Yao, Eitan Levin, Mateo Díaz

Stanford | The University of Chicago | ICML International Conference On Machine Learning

arXiv ICML

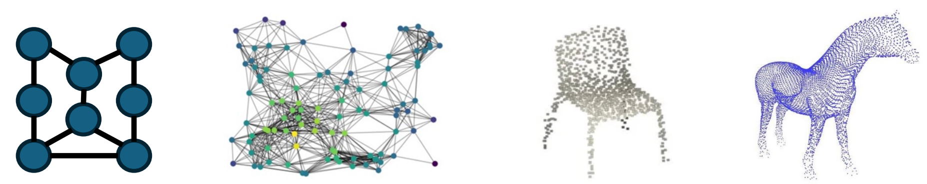
Motivation

- Classical Supervised Learning: $\mathbb{R}^m \to \mathbb{R}^n$ (Fixed Dimensions)

  Universality (Expressivity of models):
  Neural Networks can approximate any continuous function $\mathcal{C}(\mathbb{R}^m, \mathbb{R}^n)$ on a compact domain $K \subseteq \mathbb{R}^m$ to arbitrary precision[1]:

$$\sup_{x\in K} |f(x) - f_{\theta(\varepsilon)}(x)| \leq \varepsilon$$

- Any-Dimensional Objects:
  Sets, Graphs, Point clouds with varying sizes



Any-Dimensional Models:
DeepSets[2], GNNs[3], PointNet[4],……

Can any-dimensional models universally approximate functions across dimensions?

[1] Hornik, et al. (1989) "Multilayer feedforward networks are universal approximators." *Neural networks*
[2] Zaheer, et al. (2017) " Deep sets." *NeurIPS*
[3] Gori, et al. (2005) "A new model for learning in graph domains." *IJCNN*
[4] Qi, et al. (2017) "PointNet: Deep learning on point sets for 3d classification and segmentation." *CVPR*

Any-dimensional universality challenge

Finite parameters $\Longrightarrow$ Infinite functions $f_1, f_2, f_3, \dots$
(defined on growing sized spaces V_i)

Example: Graph parameters (total weight of all edges)

$$f(W) = \sum_{1\leq i\leq j\leq n} W_{ij},\ W \in \mathbb{R}^{n\times n}_{\text{sym}},\ \forall n \in \mathbb{N}$$

$$f_1: \mathbb{R}^{1\times 1}_{\text{sym}} \to \mathbb{R}$$
$$f_2: \mathbb{R}^{2\times 2}_{\text{sym}} \to \mathbb{R}$$
$$f_3: \mathbb{R}^{3\times 3}_{\text{sym}} \to \mathbb{R}$$
$$\vdots$$

The functions do not have a single input space!

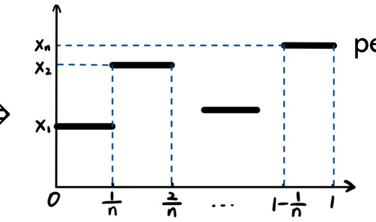
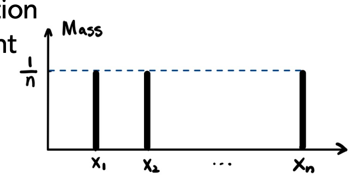
How to think functions across dimensions?

Construct ***Limit space V_∞*** (the single input space), containing each V_i as subspace (formal construction in the paper)

$$V_\infty \xrightarrow{\text{project to n-dimensional space}} V_n$$

Examples:

- Sets: $V_n = \mathbb{R}^n$ with normalized ℓ_p-norm $\|x\|_{\bar{p}} = \left(\frac{1}{n}\sum_{i=1}^n |x_i|^p\right)^{1/p}$

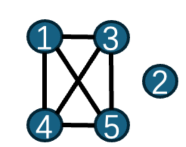
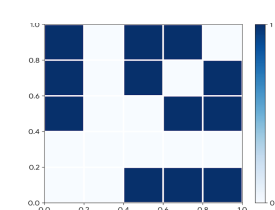
$(x_1, x_2, \dots, x_n) \Longleftrightarrow$  $\overset{\text{permutation invariant}}{\Longleftrightarrow}$ 

Any-dimensional vectors | Step functions on $[0,1]$ | Empirical measures

$$\overline{V_\infty} \cong L^p([0,1]) = \left\{ f: [0,1] \to \mathbb{R} \text{ measurable } : \int_0^1 |f(t)|^p dt < \infty \right\}$$

Permutation invariant $\Rightarrow$ $\mathcal{P}_p(\mathbb{R})$ (the probability space with W_p)

- Graphs: $V_n = \mathbb{R}^{n\times n}_{\text{sym}}$ with cut norm

 $\Longleftrightarrow$  $\overset{\text{permutation invariant}}{\Longleftrightarrow}$ Relabeling Orbit

Any-dimensional graphs | Step functions on $[0,1]^2$ | Graphon[5] space with cut distance

$$\overline{V_\infty} \cong \{W: [0,1]^2 \to [0,1] \text{ measurable symmetric}\}$$

An any-dimensional function (model) corresponds to a *continuous*[6] function defined on V_∞

Remark: the invariance makes compact sets sufficient rich

[5] Lovász, László. (2012) "Large networks and graph limits." *American Mathematical Soc*
[6] Levin, Ma, et al. (2025) "On Transferring Transferability: Towards a Theory for Size Generalization." *NeurIPS*

Recipe for any-dimensional universality

1. Characterize ***compact*** sets in the limit space
2. Construct ***continuous*** models on the limit space
3. Establish universality based on ***Stone-Weierstrass*** argument (subalgebra + separating points)

Universal models

- Set functions (permutation invariant)

  Normalized DeepSets[7] is universal

$$\overline{\text{DeepSets}_\infty}(\mu) = \sigma_\theta\left(\int \rho_\theta \, d\mu\right),\ \sigma_\theta \in \mathcal{C}(\mathbb{R}^r, \mathbb{R}),\ \rho_\theta \in \mathcal{C}(\mathbb{R}^k, \mathbb{R}^r)$$

- Graph functions (permutation invariant)

  Pointwise nonlinearity is not continuous w.r.t. cut distance

  Construct universal model via homomorphism densities

$$h_m(W) = \int_{[0,1]^m} \prod_{1\leq i<j\leq m} \left(a_{ij} W(x_i, x_j) + b_{ij}\right) d\mathbf{x}$$

- Point clouds functions (permutation and rotation invariant)

  By Gram map and Invariant Graphon Networks (IGN)[8]

$$X \longmapsto \frac{1}{2R^2}\langle X(x), X(y)\rangle + \frac{1}{2} \triangleq W_X(x,y) \longmapsto \text{IGN}(W_X)$$

[7] Bueno, et al. (2021) "On the representation power of set pooling networks." *NeurIPS*
[8] Herbst, et al. (2025) "Higher-Order Graphon Neural Networks: Approximation and Cut Distance." *ICLR*

Future work

- Establish any-dimensional ***equivariant*** universality
- Extend the framework to ***sparse graphs***