Cardinality-Agnostic Universal Approximation for Neural Networks on Point Clouds

Christian Bueno*
University of California, Santa Barbara

Alan G. Hylton
NASA Glenn Research Center
Feed-Forward Neural Networks consume fixed-size ordered data.

E.g. vectors
Recurrent Neural Networks consume arbitrary-size ordered data. *i.e. sequences*
This Talk: Neural Networks that consume arbitrary-size un-ordered data. 

*l.e. sets*
Point Clouds

$\text{Fin}(\Omega)$

Point Clouds
$F : \text{Fin}(\Omega) \rightarrow \mathbb{R}^n$

Permutation-Invariant

Cardinality-Agnostic
PointNet and DeepSets
PointNet and DeepSets

\[ F_{PN}(A) = \psi \left( \max_{a \in A} \varphi(a) \right) \]  
(Qi et al. 2017)

\[ F_{DS}(A) = \psi \left( \sum_{a \in A} \varphi(a) \right) \]  
(Zaheer et al. 2017)
PointNet and DeepSets

\[ F_{PN}(A) = \psi \left( \max_{a \in A} \varphi(a) \right) \]  
(Qi et al. 2017)

\[ F_{DS}(A) = \psi \left( \sum_{a \in A} \varphi(a) \right) \]  
(Zaheer et al. 2017)

\[ F_{DS}(A) = \psi \left( \frac{1}{|A|} \sum_{a \in A} \varphi(a) \right) \]
Refactor

\[ F_{PN} = \psi \circ \max_f \]

\[(\max_f)_i = \max_{f_i} \]

\[ \max_{f_i} : \text{Fin}(\Omega) \rightarrow \mathbb{R} \]

\[ F_{DS} = \psi \circ \text{ave}_f \]

\[(\text{ave}_f)_i = \text{ave}_{f_i} \]

\[ \text{ave}_{f_i} : \text{Fin}(\Omega) \rightarrow \mathbb{R} \]
Continuity?

\[ F_{PN} = \psi \circ \max_f \]
\[ F_{DS} = \psi \circ \text{ave}_f \]

\[ (\max_f)_i = \max_{f_i} \]
\[ (\text{ave}_f)_i = \text{ave}_{f_i} \]

\[ \max_{f_i} : \text{Fin}(\Omega) \rightarrow \mathbb{R} \]
\[ \text{ave}_{f_i} : \text{Fin}(\Omega) \rightarrow \mathbb{R} \]

What topologies yields continuity on \( \text{Fin}(\Omega) \)?
Continuity?

\[ F_{PN} = \psi \circ \max_f \quad F_{DS} = \psi \circ \text{ave}_f \]

\[ (\max_f)_i = \max_{f_i} \quad (\text{ave}_f)_i = \text{ave}_{f_i} \]

\[ \max_{f_i} : \text{Fin}(\Omega) \rightarrow \mathbb{R} \quad \text{ave}_{f_i} : \text{Fin}(\Omega) \rightarrow \mathbb{R} \]

What topologies yields continuity on \( \text{Fin}(\Omega) \)?
Continuity?

\[ F_{PN} = \psi \circ \max_f \]
\[ (\max_f)_i = \max_{f_i} \]
\[ \max_{f_i} : \text{Fin}(\Omega) \rightarrow \mathbb{R} \]

\[ F_{DS} = \psi \circ \text{ave}_f \]
\[ (\text{ave}_f)_i = \text{ave}_{f_i} \]
\[ \text{ave}_{f_i} : \text{Fin}(\Omega) \rightarrow \mathbb{R} \]

What topologies yields continuity on \( \text{Fin}(\Omega) \)?

\( (\mathcal{K}(\Omega), d_H) \)
Space of nonempty compact subsets with Hausdorff metric \( d_H \)

\( (\mathcal{P}(\Omega), d_W) \)
Space of Borel probability measures with Wasserstein metric \( d_W \)
Upgrade: Unique Continuous Extension

\[ F_{PN} = \psi \circ \text{Max}_f \]

\[ \text{Max}_{f_i}(A) = \max_{a \in A} f_i(a) \]

\[ \text{Max}_{f_i} : (\mathcal{K}(\Omega), d_H) \to \mathbb{R} \]

\[ F_{DS} = \psi \circ \text{Ave}_f \]

\[ \text{Ave}_{f_i}(\mu) = \mathbb{E}_{x \sim \mu}[f_i(x)] \]

\[ \text{Ave}_{f_i} : (\mathcal{P}(\Omega), d_W) \to \mathbb{R} \]

\[ \text{Space of nonempty compact subsets with Hausdorff metric } d_H \]

\[ \text{Space of Borel probability measures with Wasserstein metric } d_W \]
Upgrade: Unique Continuous Extension

\[ F_{PN} = \psi \circ \text{Max}_f \]

\[ \text{Max}_{f_i}(A) = \max_{a \in A} f_i(a) \]

\[ \text{Max}_{f_i} : (\mathcal{K}(\Omega), d_H) \rightarrow \mathbb{R} \]

\[ F_{DS} = \psi \circ \text{Ave}_f \]

\[ \text{Ave}_{f_i}(\mu) = \mathbb{E}_{x \sim \mu} [f_i(x)] \]

\[ \text{Ave}_{f_i} : (\mathcal{P}(\Omega), d_W) \rightarrow \mathbb{R} \]

Space of nonempty compact subsets with Hausdorff metric \(d_H\)

Intuition

\[ p : \mathbb{Q} \rightarrow \mathbb{Q} \]

polynomial

Space of Borel probability measures with Wasserstein metric \(d_W\)
Upgrade: Unique Continuous Extension

\[ F_{PN} = \psi \circ \text{Max}_f \]

\[ \text{Max}_{f_i}(A) = \max_{a \in A} f_i(a) \]

\[ \text{Max}_{f_i} : (\mathcal{K}(\Omega), d_H) \rightarrow \mathbb{R} \]

\[ F_{DS} = \psi \circ \text{Ave}_f \]

\[ \text{Ave}_{f_i}(\mu) = \mathbb{E}_{x \sim \mu} [f_i(x)] \]

\[ \text{Ave}_{f_i} : (\mathcal{P}(\Omega), d_W) \rightarrow \mathbb{R} \]

\[ p : \mathbb{R} \rightarrow \mathbb{R} \]

polynomial

Intuition

Space of nonempty compact subsets with Hausdorff metric \( d_H \)

Space of Borel probability measures with Wasserstein metric \( d_W \)
Upgrade: Unique Continuous Extension

\[ F_{PN} = \psi \circ \text{Max}_f \]
\[ \text{Max}_{f_i}(A) = \max_{a \in A} f_i(a) \]
\[ \text{Max}_{f_i} : (\mathcal{K}(\Omega), d_H) \to \mathbb{R} \]

\[ F_{DS} = \psi \circ \text{Ave}_f \]
\[ \text{Ave}_{f_i}(\mu) = \mathbb{E}_{x \sim \mu}[f_i(x)] \]
\[ \text{Ave}_{f_i} : (\mathcal{P}(\Omega), d_W) \to \mathbb{R} \]

**Intuition**

\[ p : \mathbb{C} \to \mathbb{C} \]

polynomial
Upgrade: Unique Continuous Extension

\[ F_{PN} = \psi \circ \text{Max}_f \]

\[ \text{Max}_f : \left( \mathcal{K}(\Omega), d_H \right) \to \mathbb{R} \]

- Space of nonempty compact subsets with Hausdorff metric \( d_H \)

\[ F_{DS} = \psi \circ \text{Ave}_f \]

\[ \text{Ave}_f : \left( \mathcal{P}(\Omega), d_W \right) \to \mathbb{R} \]

- Space of Borel probability measures with Wasserstein metric \( d_W \)

\( (\Omega, d) \) compact \( \implies \) \( (\mathcal{K}(\Omega), d_H) \) and \( (\mathcal{P}(\Omega), d_W) \) compact
Stability of Extension

Theorem. Suppose $\Omega \subseteq \mathbb{R}^N$ is compact. Then every PointNet and normalized-DeepSet network with Lipschitz continuous activation functions is Lipschitz continuous on $(\mathcal{K}(\Omega), d_H)$ and $(\mathcal{P}(\Omega), d_W)$ respectively.

$$\|F_{PN}(A) - F_{PN}(B)\| \leq K_{F_{PN}} d_H(A, B)$$

$$\|F_{DS}(\mu) - F_{DS}(\nu)\| \leq K_{F_{DS}} d_W(\mu, \nu)$$
Classical UAT $\rightarrow$ Topological UAT

**Theorem.** Let $X$ be compact Hausdorff and $\sigma \in C(\mathbb{R})$ not a polynomial. If $S \subseteq C(X)$ separates points and has a nonzero constant, then $\text{span}(\sigma \circ \text{span} S)$ is dense in $C(X)$. If $S$ is a linear subspace, then $\text{span}(\sigma \circ S)$ is dense in $C(X)$. 
Topological UAT $\rightarrow$ UAT for Extension

**Theorem.** Let $X$ be compact Hausdorff and $\sigma \in C(\mathbb{R})$ not a polynomial. If $S \subseteq C(X)$ separates points and has a nonzero constant, then $\text{span}(\sigma \circ \text{span } S)$ is dense in $C(X)$. If $S$ is a linear subspace, then $\text{span}(\sigma \circ S)$ is dense in $C(X)$.

Letting $S_{PN} = \{\text{Max}_f | f \in \mathcal{N}^\tau\}$ and $S_{DS} = \{\text{Ave}_f | f \in \mathcal{N}^\tau\}$ works!

This yields a UAT for generalized PointNet and DeepSets on $\mathcal{K}(\Omega)$ and $\mathcal{P}(\Omega)$. 
**Theorem.** Let $X$ be compact Hausdorff and $\sigma \in C(\mathbb{R})$ not a polynomial. If $S \subseteq C(X)$ separates points and has a nonzero constant, then $\text{span}(\sigma \circ \text{span } S)$ is dense in $C(X)$. If $S$ is a linear subspace, then $\text{span}(\sigma \circ S)$ is dense in $C(X)$.

Letting $S_{PN} = \{\text{Max}_f \mid f \in \mathcal{N}^\tau\}$ and $S_{DS} = \{\text{Ave}_f \mid f \in \mathcal{N}^\tau\}$ works!

This yields a UAT for generalized PointNet and DeepSets on $\mathcal{K}(\Omega)$ and $\mathcal{P}(\Omega)$.

<table>
<thead>
<tr>
<th>Theorem.</th>
<th>Restricting to $\text{Fin}(\Omega)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{span } (\sigma \circ \text{span } S_{PN})$</td>
<td>$\text{UnifCont}(\text{Fin}(\Omega), d_H)$</td>
</tr>
<tr>
<td>$\text{span } (\sigma \circ S_{DS})$</td>
<td>$\text{UnifCont}(\text{Fin}(\Omega), d_W)$</td>
</tr>
</tbody>
</table>
The Overlap and Limitations

- \( d_H \)-continuous on \( \text{Fin}(\Omega) \)
- Uniformly \( d_H \)-continuous on \( \text{Fin}(\Omega) \)
- \( d_W \)-continuous on \( \text{Fin}(\Omega) \)
- Uniformly \( d_W \)-continuous on \( \text{Fin}(\Omega) \)

PointNet

DeepSets
The Overlap and Limitations

- $d_H$-continuous on $\text{Fin}(\Omega)$
- $d_W$-continuous on $\text{Fin}(\Omega)$
- Uniformly $d_H$-continuous on $\text{Fin}(\Omega)$
- Uniformly $d_W$-continuous on $\text{Fin}(\Omega)$
- Constant Functions...

PointNet

DeepSets
The Overlap and Limitations

Only Constant Functions

\( d_H \)-continuous on \( \text{Fin}(\Omega) \)

\( d_W \)-continuous on \( \text{Fin}(\Omega) \)

PointNet

DeepSets
The Overlap and Limitations

- PointNet: Uniformly $d_H$-continuous on $\text{Fin}(\Omega)$
- DeepSets: Uniformly $d_W$-continuous on $\text{Fin}(\Omega)$

Only Constant Functions
The Overlap and Limitations

Uniformly $d_H$-continuous on $\text{Fin}(\Omega)$

Uniformly $d_W$-continuous on $\text{Fin}(\Omega)$

Only Constant Functions

e.g. Diameter

PointNet

DeepSets
The Overlap and Limitations

- **Uniformly $d_H$-continuous on $\text{Fin}(\Omega)$**
  - **e.g. Diameter**

- **Uniformly $d_W$-continuous on $\text{Fin}(\Omega)$**
  - **e.g. Center-of-Mass**

- **Only Constant Functions**

**PointNet**

**DeepSets**
Is the Problem at Infinity?

\[ \leq k \text{ points} \]
Is the Problem at Infinity?

...Not Quite

≤ \( k \) points
Is the Problem at Infinity?

...Not Quite

PointNet still can’t learn center-of-mass even with fixed cloud size.
Center-of-Mass, PointNet, & Fixed Size Sets

Two $d_H$-continuous paths with same limit...  
...But different limiting centers.
Error Lower Bound for $\text{ave}_f$

**Theorem.** Let $\Omega \subseteq \mathbb{R}^n$ be the unit ball, $k \geq 3$, and $f : \Omega \rightarrow \mathbb{R}^n$ continuous. Then for any distinct $p, q \in \Omega$ and $0 < \tau < 1$ there exists a $k$-point set $A$ with $p, q \in A \subseteq \Omega$ so that

$$
\|F_{PN}(A) - \text{ave}_f(A)\| > (1 - \tau) \left(\frac{k - 2}{2k}\right) \|f(p) - f(q)\|
$$

for any PointNet-type $F_{PN}$, regardless of depth/width/training/etc. Thus,

$$
\|F_{PN} - \text{ave}_f\|_{L^\infty(\text{Fin}^k(\Omega))} \geq \left(\frac{k - 2}{2k}\right) \text{Diam}(f(\Omega))
$$

Moreover, we can construct such geometric “adversarial” examples
Test of Error Lower Bound (Center-of-Mass)
Test of Error Lower Bound (Center-of-Mass)
Test of Error Lower Bound (Center-of-Mass)
Test of Error Lower Bound (Center-of-Mass)
Summary

- PointNet & normalized-DeepSets uniquely continuously extend to $\mathcal{K}(\Omega)$ and $\mathcal{P}(\Omega)$ respectively.

- PointNet & normalized-DeepSets can uniformly approximate the uniformly continuous functions on $\text{Fin}(\Omega)$ w.r.t. $d_H$ and $d_W$ resp. They cannot uniformly approximate anything else.

- PointNet & normalized-DeepSets are Lipschitz if activations are Lipschitz

- Constants are only functions mutually approximable by PointNet and DeepSets on $\text{Fin}(\Omega)$.

- PointNet cannot uniformly approximate averages of continuous functions (even for fixed cloud size) and geometric adversarial examples are abundant and easily constructed.
References


Thank You!