Generalised Aggregation for Graph Neural Networks

Ryan Kortvelesy, Steven Morad, Amanda Prorok

NEURAL INFORMATION PROCESSING SYSTEMS

Motivation

- In GNNs $y_i = \phi(x_i, \bigoplus \psi(x_i, x_j))$ we generally use a discrete set of options {mean, sum, max} as aggregators $\bigoplus$.
- We wish to parametrise the space of aggregation functions, creating a *learnable* aggregator "GenAgg".

Parameterisations

Aggregation Function	α	β	f
mean: $\frac{1}{n}\sum x_i$	0	0	$f(x) = x$
sum: $\sum x_i$	1	0	$f(x) = x$
product: $\prod \lvert x_i \rvert$	1	0	$f(x) = \log(\lvert x \rvert)$
min (magnitude): $\min \lvert x_i \rvert$	0	0	$f(x) = \lim_{p\to\infty} \lvert x \rvert^{-p}$
max (magnitude): $\max \lvert x_i \rvert$	0	0	$f(x) = \lim_{p\to\infty} \lvert x \rvert^{p}$
min: $\min x_i$	0	0	$f(x) = \lim_{p\to\infty} e^{-px}$
max: $\max x_i$	0	0	$f(x) = \lim_{p\to\infty} e^{px}$
harmonic mean: $\frac{n}{\sum \frac{1}{x_i}}$	0	0	$f(x) = \frac{1}{x}$
geometric mean: $\sqrt[n]{\prod \lvert x_i \rvert}$	0	0	$f(x) = \log(\lvert x \rvert)$
root mean square: $\sqrt{\frac{1}{n}\sum x_i^2}$	0	0	$f(x) = x^2$
euclidean norm: $\sqrt{\sum x_i^2}$	1	0	$f(x) = x^2$
standard deviation: $\sqrt{\frac{1}{n}\sum (x_i - \mu)^2}$	0	1	$f(x) = x^2$
log-sum-exp: $\log\left(\sum e^{x_i}\right)$	1	0	$f(x) = e^x$

A collection of special cases of GenAgg, and the corresponding parametrisations.

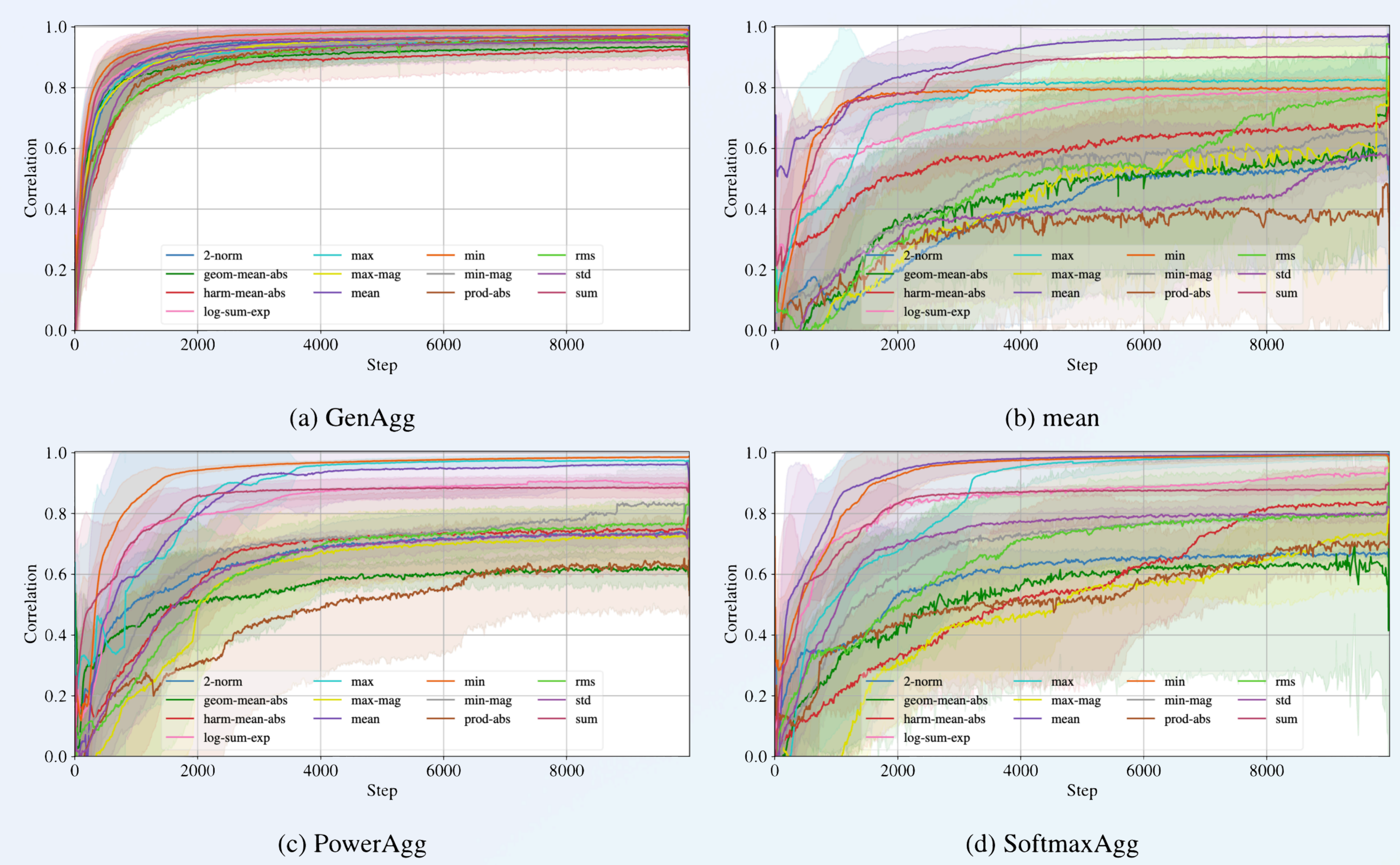
Method

$$\bigoplus x_i = f^{-1}\left(n^{\alpha-1} \sum_{i\in[1..n]} f(x_i - \beta\mu)\right)$$

- $f: \mathbb{R} \to \mathbb{R}$ function that defines the relative impact of each input.
- α : Scaling factor for the cardinality of the input set.
- β : Variance factor, allowing the representation of centralised moments.

Results

Experiment 1: Tested the ability of GNNs with different aggregation functions to represent simple aggregators



(a) GenAgg (b) mean (c) PowerAgg (d) SoftmaxAgg

Experiment 2: Demonstrated that using GenAgg in a GNN increases performance on GNN Benchmarks

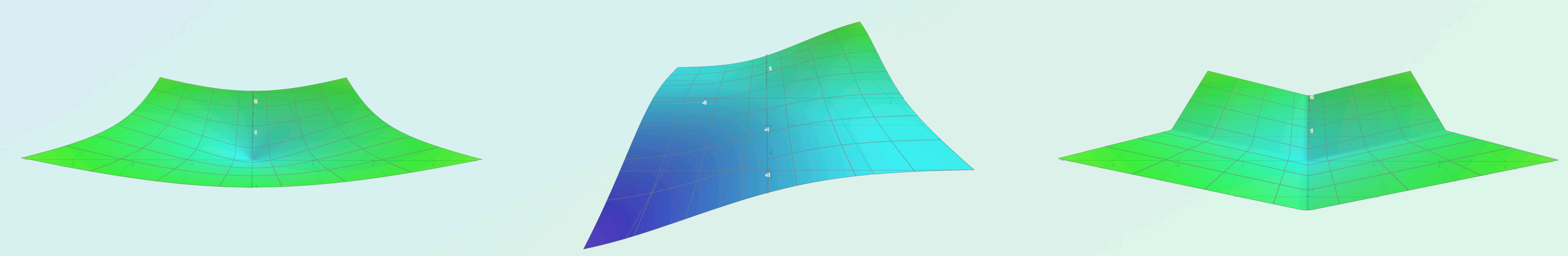
Dataset	GenAgg	P-Agg	S-Agg	PNA	mean	sum	max
CIFAR10	**0.540** ±0.09	0.451 ±0.14	0.467 ±0.13	0.473 ±0.14	0.502 ±0.13	0.435 ±0.15	0.471 ±0.13
MNIST	**0.928** ±0.07	0.877 ±0.07	0.864 ±0.06	0.717 ±0.19	0.847 ±0.08	0.844 ±0.08	0.831 ±0.08
CLUSTER	**0.627** ±0.02	0.610 ±0.01	0.611 ±0.01	0.168 ±0.01	0.602 ±0.01	0.170 ±0.02	0.501 ±0.01
PATTERN	**0.925** ±0.00	0.883 ±0.01	0.896 ±0.03	0.861 ±0.01	0.871 ±0.01	0.860 ±0.01	0.860 ±0.01

Distributive Property

GenAgg defines the *Generalised Distributive Property, which is satisfied for aggregator* $\bigoplus$ *with binary operator* ψ *if:* $\psi(c, \bigoplus x_i) = \bigoplus \psi(c, x_i)$

Any parametrisation of GenAgg is satisfied by the operator: $\psi(a,b) = f^{-1}(f(a)\cdot f(b))$

Aggregation Function	Distributive Operations $\psi(a,b)$
mean: $\frac{1}{n}\sum x_i$	$a+b, \quad a\cdot b$
sum: $\sum x_i$	$a \cdot b$
product: $\prod \lvert x_i \rvert$	$\lvert a \rvert^{\log \lvert b \rvert}$
min (magnitude): $\min \lvert x_i \rvert$	$\min(\lvert a \rvert, \lvert b \rvert)$
max (magnitude): $\max \lvert x_i \rvert$	$\max(\lvert a \rvert, \lvert b \rvert)$
min: $\min x_i$	$\min(a,b)$
max: $\max x_i$	$\max(a,b)$
harmonic mean: $\frac{n}{\sum \frac{1}{x_i}}$	$\frac{a\cdot b}{a+b}, \quad a\cdot b$
geometric mean: $\sqrt[n]{\prod \lvert x_i \rvert}$	$\lvert a \rvert \cdot \lvert b \rvert, \quad \lvert a \rvert^{\log \lvert b \rvert}$
root mean square: $\sqrt{\frac{1}{n}\sum x_i^2}$	$\sqrt{a^2+b^2}, \quad \lvert a \rvert \cdot \lvert b \rvert$
euclidean norm: $\sqrt{\sum x_i^2}$	$\lvert a \rvert \cdot \lvert b \rvert$
standard deviation: $\sqrt{\frac{1}{n}\sum (x_i-\mu)^2}$	$\lvert a \rvert \cdot \lvert b \rvert$
log-sum-exp: $\log\left(\sum e^{x_i}\right)$	$a+b$



UNIVERSITY OF CAMBRIDGE

PROROKLAB