

Paper Implementation

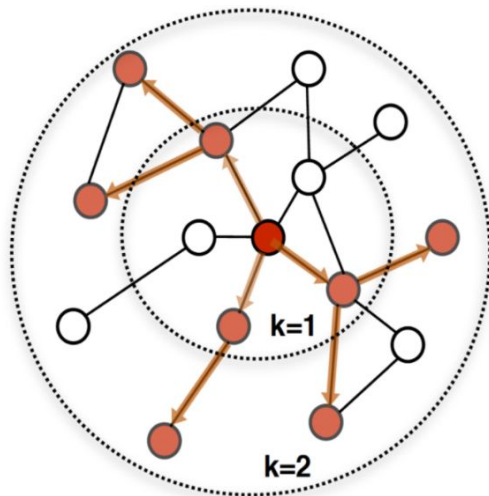
GraphSAGE: Inductive Representation Learning on Large
Graphs

—

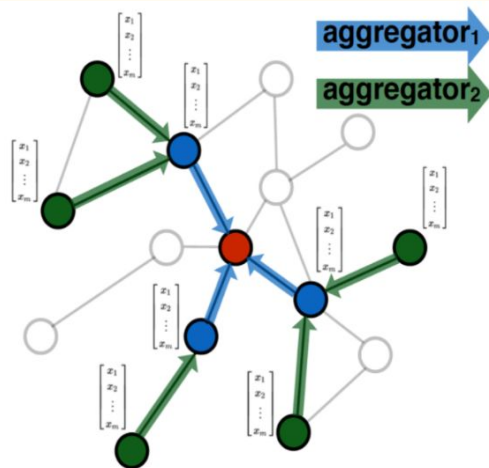
By Ahmed A. A. Elhag

General inductive
framework that leverages
node features to efficiently
generate node embeddings
for unseen datasets

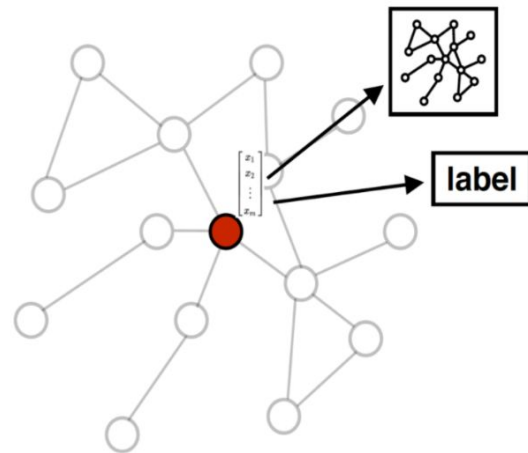
Main Idea



1. Sample neighborhood



2. Aggregate feature information from neighbors



3. Predict graph context and label using aggregated information

Aggregator Architectures

$$\begin{aligned}\mathbf{h}_{\mathcal{N}(v)}^k &\leftarrow \text{AGGREGATE}_k(\{\mathbf{h}_u^{k-1}, \forall u \in \mathcal{N}(v)\}); \\ \mathbf{h}_v^k &\leftarrow \sigma \left(\mathbf{W}^k \cdot \text{CONCAT}(\mathbf{h}_v^{k-1}, \mathbf{h}_{\mathcal{N}(v)}^k) \right) \\ &\cdot\end{aligned}$$

- Mean Aggregator
- LSTM Aggregator
- Pooling Aggregator

Implementation approach

GraphSAGE paper

- Define one GraphSAGE layer:
depending on the aggregator
type
 - Stack two GraphSAGE layers,
use nonlinearity between them.
 - Build the train function:
(optimizer, loss...)
 - Experiments
-

The Experiments

—

Results

Dataset	Type	Nodes	Edges	Classes	Features	Mean Aggregator	MaxPooling Aggregator
Cora	Citation network	2,708	5,429	7	1,433	80.0 \pm 1	81.0 \pm 1
Citeseer	Citation network	3,327	4,732	6	3,703	68 \pm 1	69.5 \pm 0.8

Improvement

ADVANCING GRAPHSAGE WITH A DATA -DRIVEN NODE SAMPLING

Jihun Oh et al., (ICLR 2019)

- A new data-driven sampling approach to reason about the real-valued importance of a neighborhood
 - Use a non-linear regressor, as a criterion for subsampling neighborhoods.
 - The regressor is learned using a value based reinforcement learning.
-

Conclusion

- GraphSAGE: inductive framework, leverages node feature information.
- Generate node embeddings for previously unseen data.
- Work on Large Scale graphs!

Thank you!

—