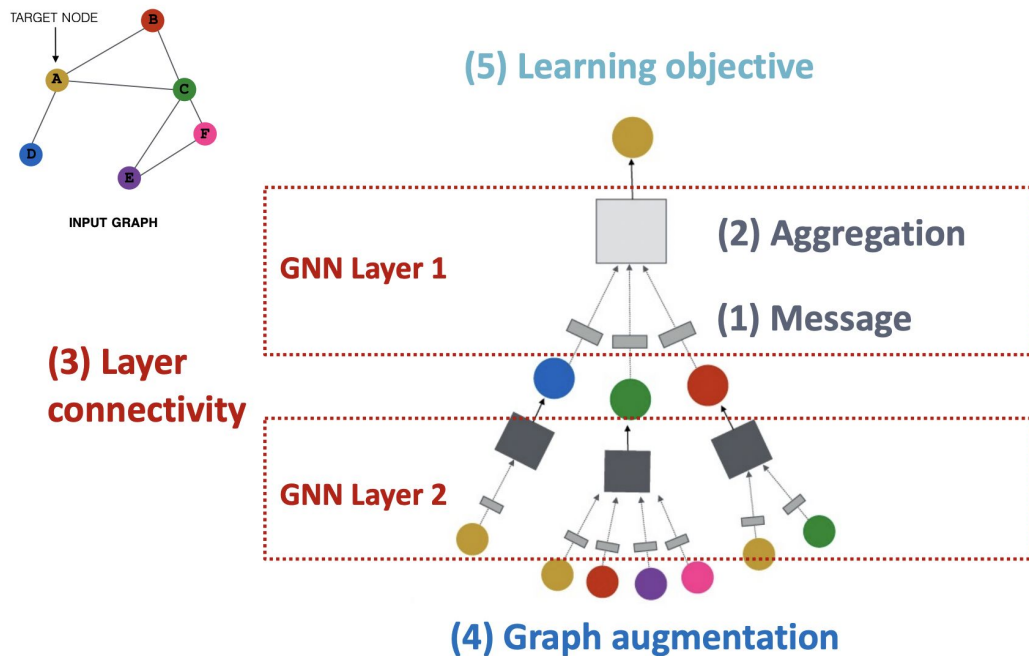


Глубинное обучение в анализе графовых данных

7. Аугментации и обучение

в предыдущих сериях...

Части GNN

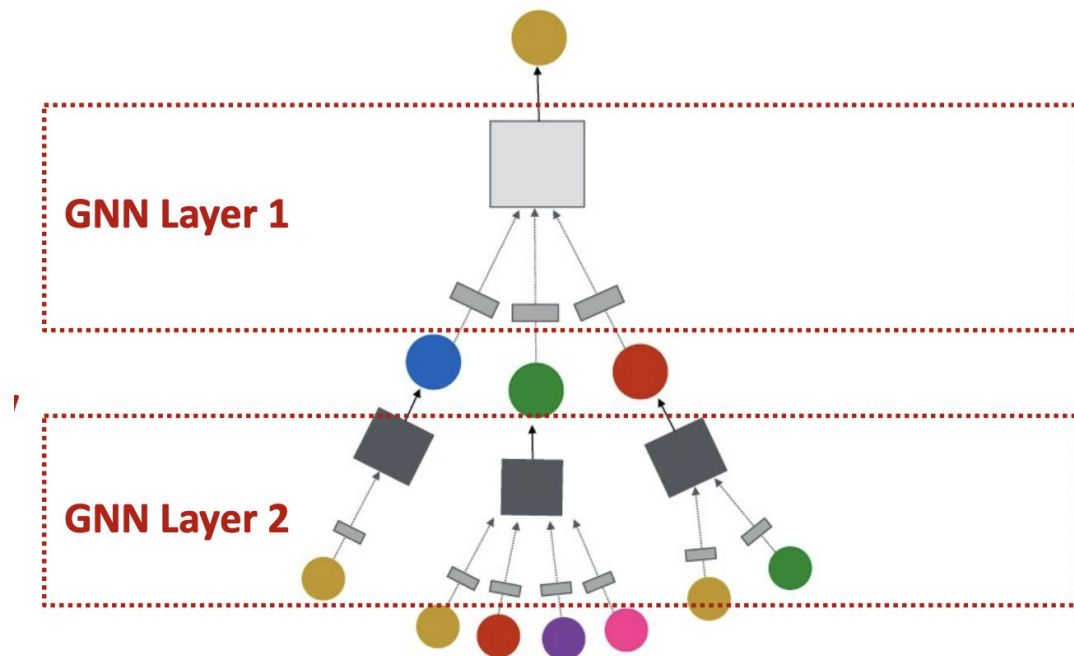


Различные архитектуры

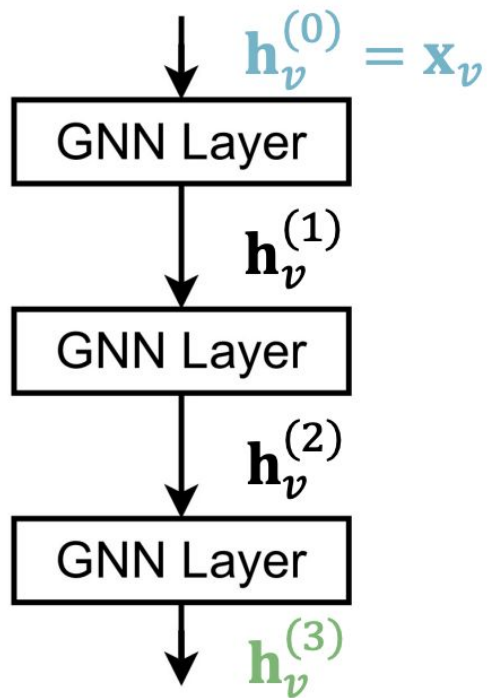
- GCN
- GraphSage
- GAT

Связь слоев

Связь слоев

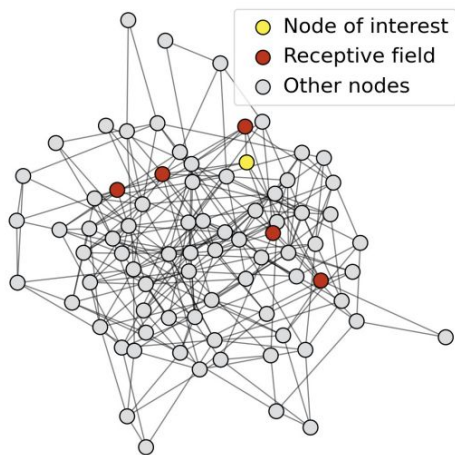


Последовательно

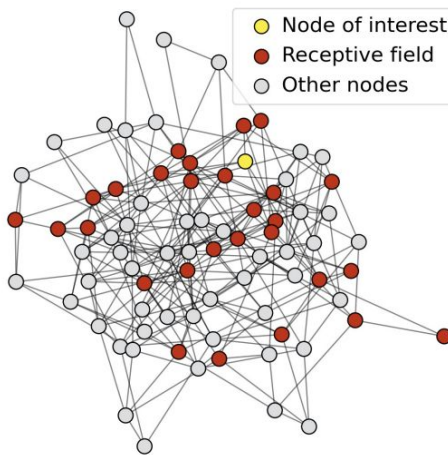


Рецептивное поле

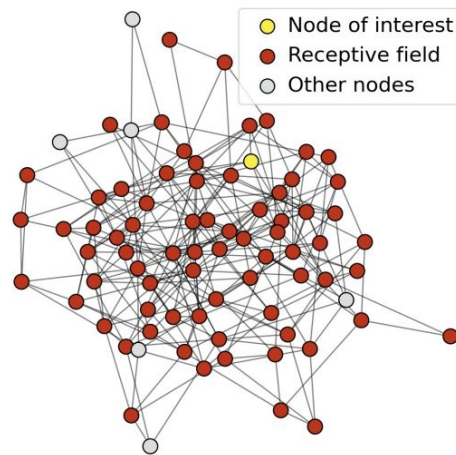
**Receptive field for
1-layer GNN**



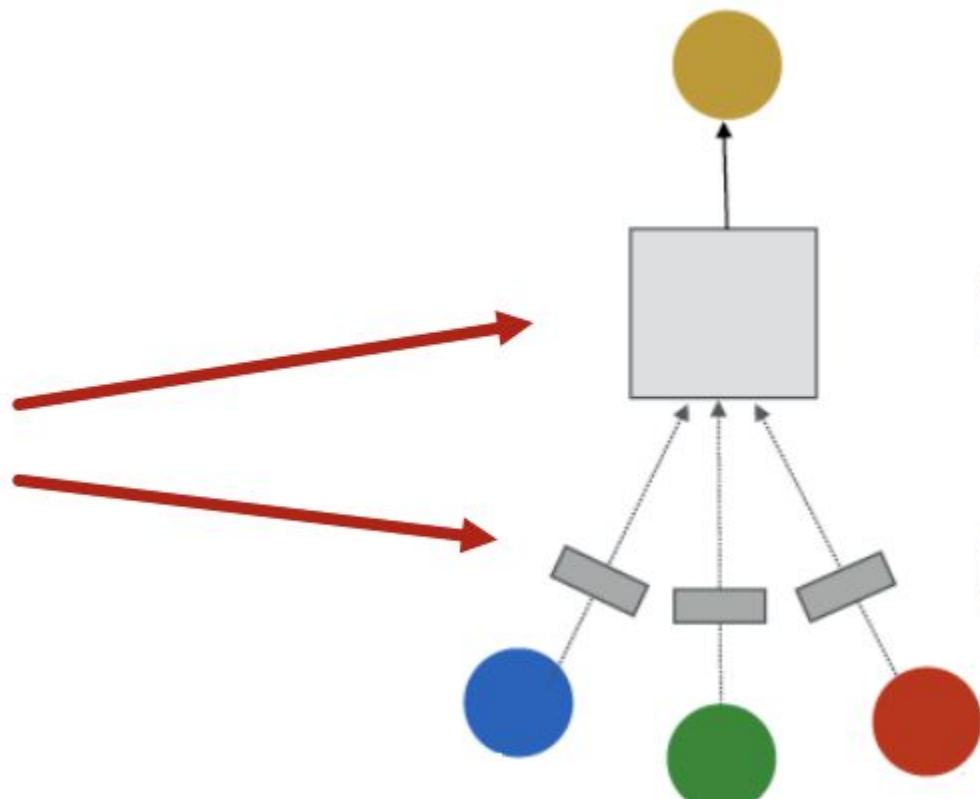
**Receptive field for
2-layer GNN**



**Receptive field for
3-layer GNN**



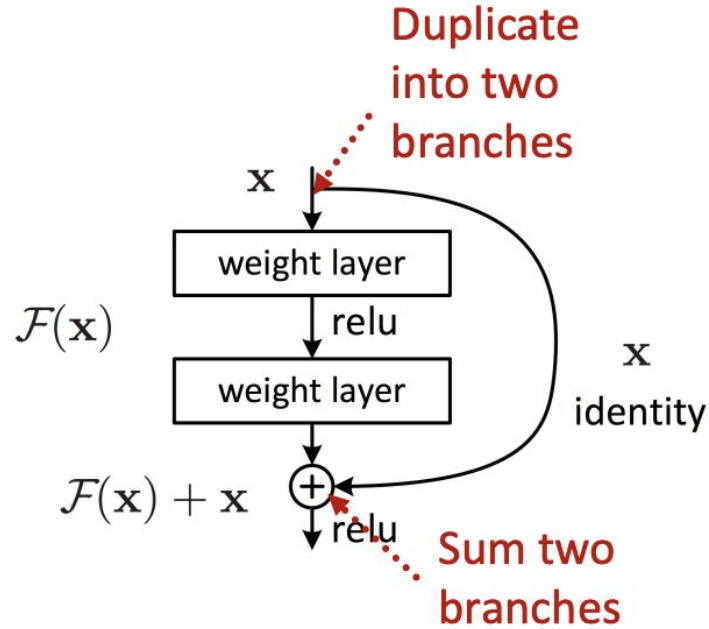
Решение 1



(2) Aggregation

(1) Transformation

Skip-connections



Пример

- A standard GCN layer

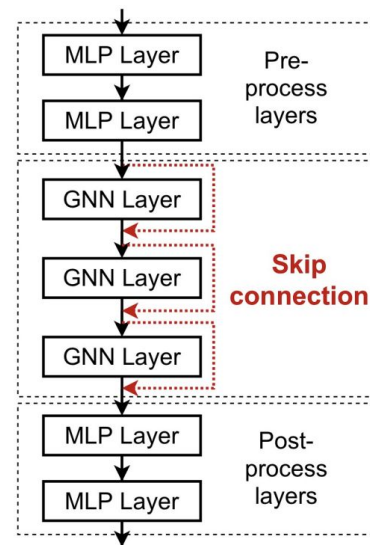
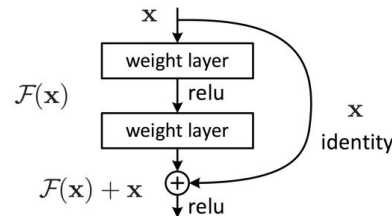
$$\mathbf{h}_v^{(l)} = \sigma \left(\sum_{u \in N(v)} \mathbf{W}^{(l)} \frac{\mathbf{h}_u^{(l-1)}}{|N(v)|} \right)$$

This is our $F(\mathbf{x})$

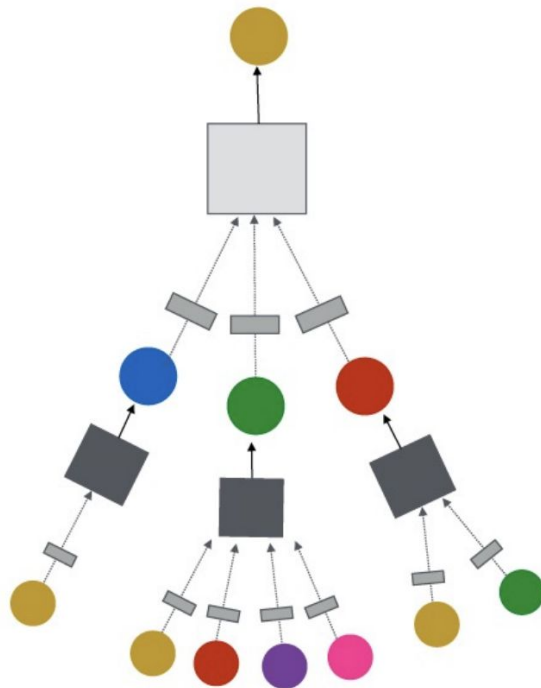
- A GCN layer with skip connection

$$\mathbf{h}_v^{(l)} = \sigma \left(\sum_{u \in N(v)} \mathbf{W}^{(l)} \frac{\mathbf{h}_u^{(l-1)}}{|N(v)|} + \mathbf{h}_v^{(l-1)} \right)$$

$F(\mathbf{x})$ + \mathbf{x}



Аугментации



(4) Graph augmentation

Проблемы

- нет фичей
- sparse
- dense
- large

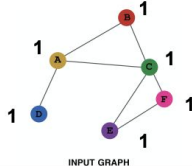
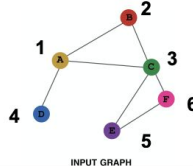
Решения

- нет фичей - аугментации
- sparse - добавить связей
- dense - использовать сэмплирование
- large - сэмплировать подграфы

Нет фичей

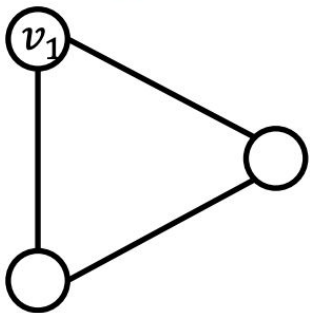
- 1) добавить всем константные фичи
- 2) добавить айдишники

Сравнение

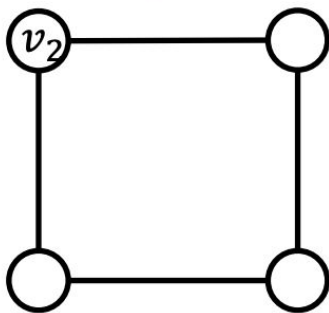
	Constant node feature 	One-hot node feature 
Expressive power	Medium. All the nodes are identical, but GNN can still learn from the graph structure	High. Each node has a unique ID, so node-specific information can be stored
Inductive learning (Generalize to unseen nodes)	High. Simple to generalize to new nodes: we assign constant feature to them, then apply our GNN	Low. Cannot generalize to new nodes: new nodes introduce new IDs, GNN doesn't know how to embed unseen IDs
Computational cost	Low. Only 1 dimensional feature	High. $O(V)$ dimensional feature, cannot apply to large graphs
Use cases	Any graph, inductive settings (generalize to new nodes)	Small graph, transductive settings (no new nodes)

Циклы

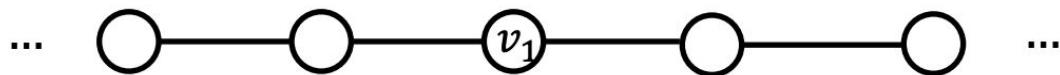
v_1 resides in a cycle with length 3



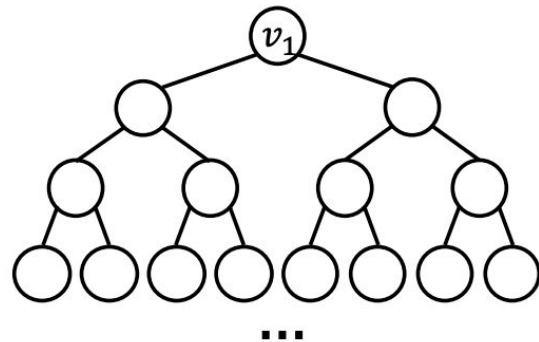
v_1 resides in a cycle with length 4



v_1 resides in a cycle with infinite length



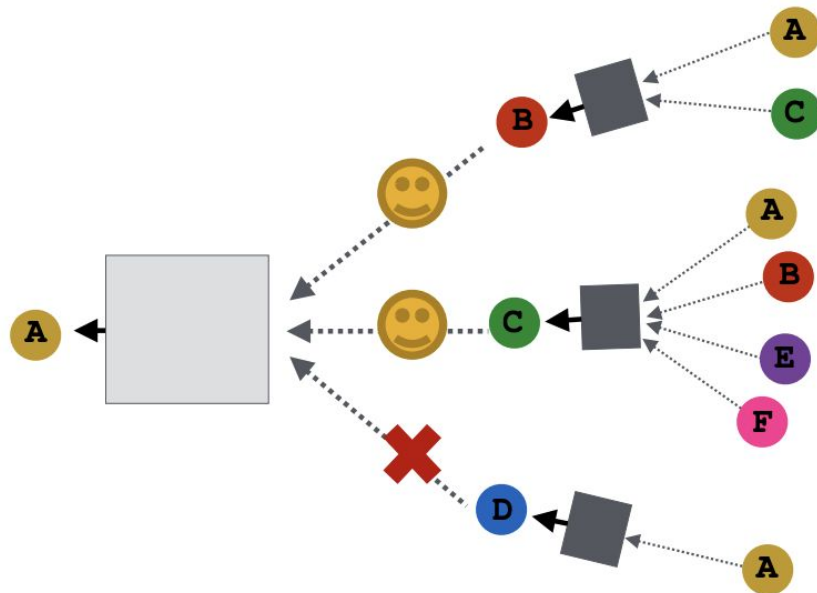
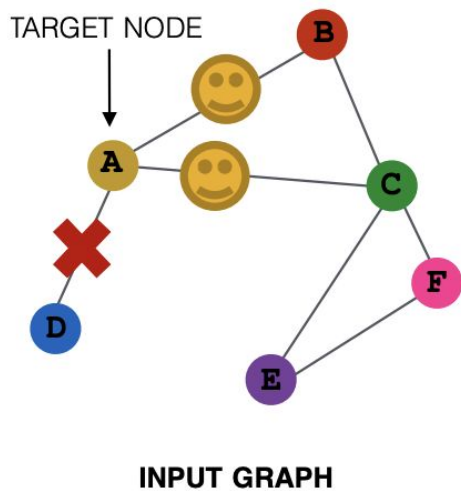
The computational graphs for node v_1 are always the same



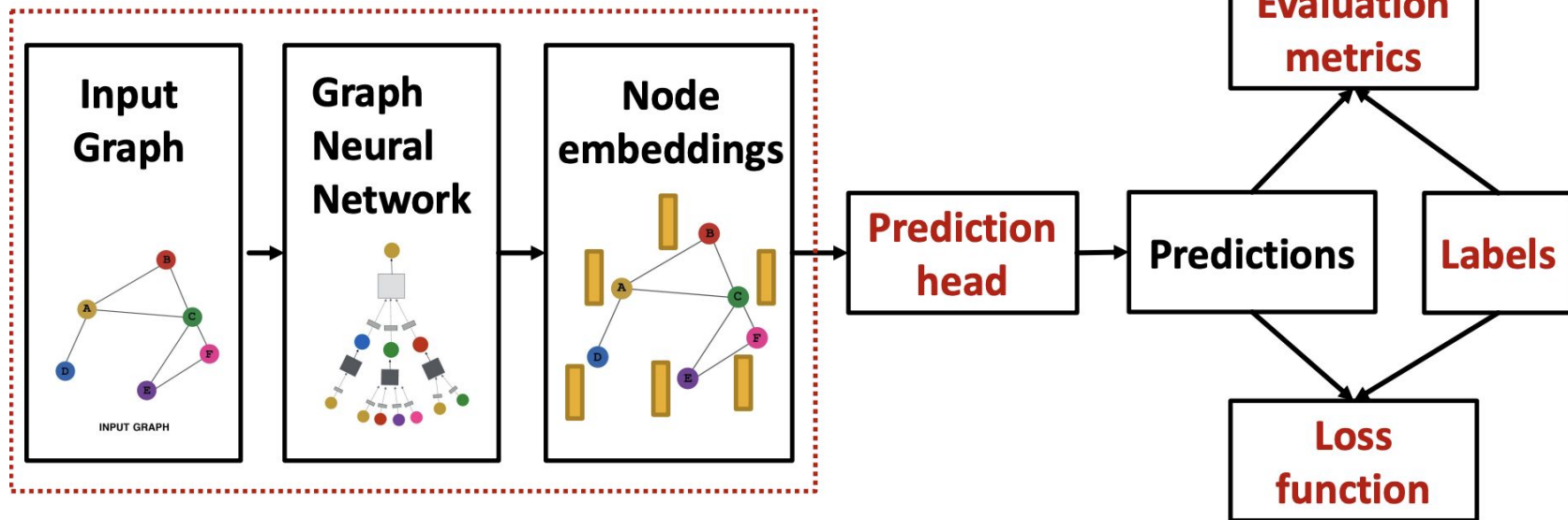
Добавление вершин/ребер

- сокращение дистанции

Сэмплирование соседей



Обучение



Output of a GNN: set of node embeddings

Иерархический global pooling

$$\{-1, -2, 0, 1, 2\}$$

$$\{-10, -20, 0, 10, 20\}$$

Splitting

