



Università
della
Svizzera
italiana

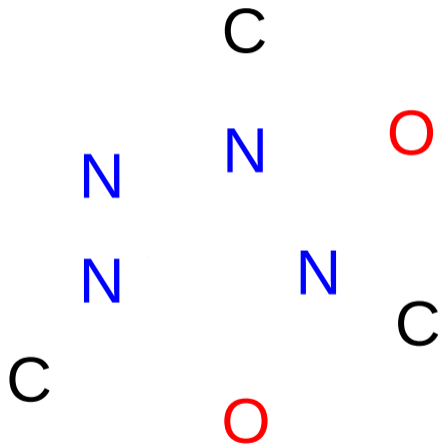
Graph Neural Networks e MinCut Pooling

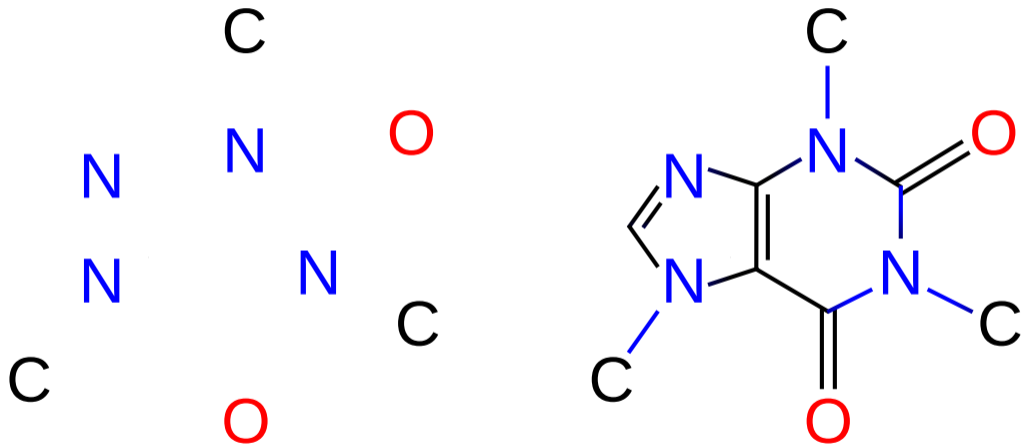
Daniele Grattarola (@riceasphait)

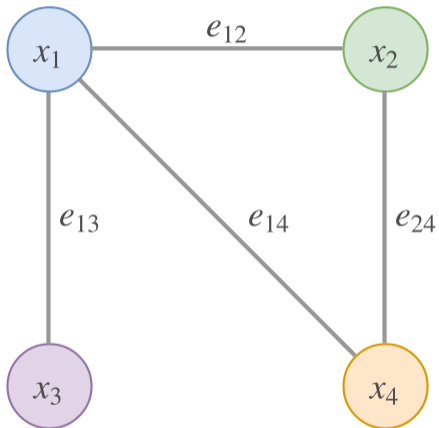
[danielegrattarola.github.io](https://github.com/danielegrattarola)

June 29, 2020

1. Graph Neural Networks.
2. Spectral Clustering with Graph Neural Networks for Graph Pooling (ICML 2020).
3. Esperimenti.





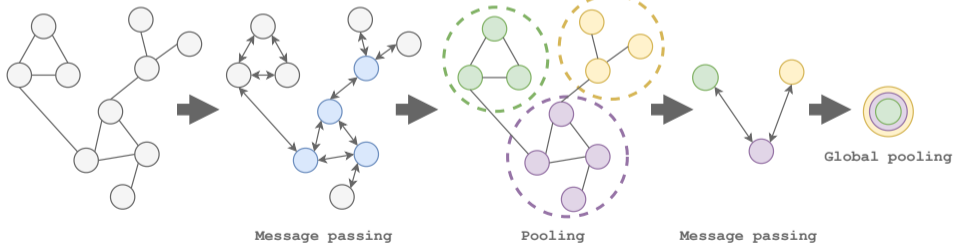
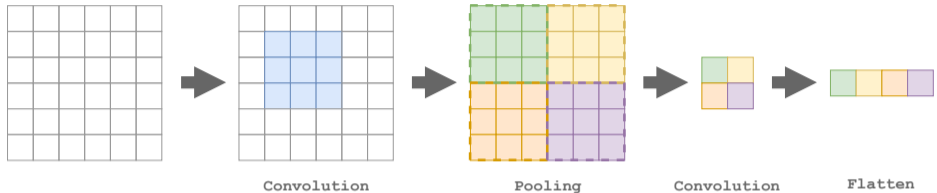


Nodi (X), archi (E)

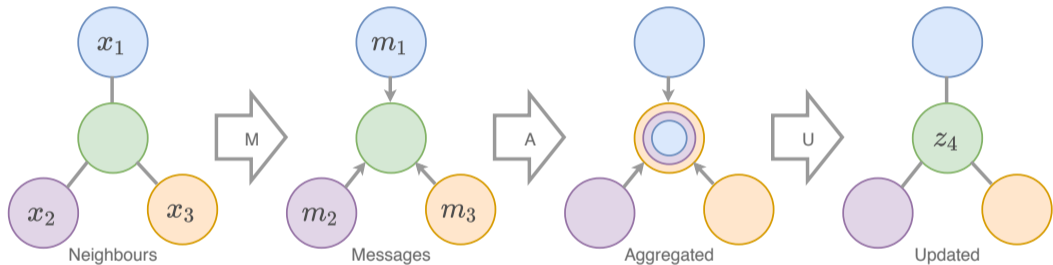
Attributi di nodo ($x_i \ 8i \ 2 \ X$)

Attributi di arco ($e_{ij} \ 8(i,j) \ 2 \ E$)

GNN = CNN Generalizzate



Message Passing [1]



[1] J. Gilmer *et al.*, "Neural message passing for quantum chemistry," 2017.

Strategie per Fare Message Passing

GraphConv

Kipf & Welling

ChebConv

Defferrard et al.

GraphSageConv

Hamilton et al.

ARMAConv

Bianchi et al.

ECConv

Simonovsky & Komodakis

GraphAttention

Velickovic et al.

GraphConvSkip

Bianchi et al.

APPNP

Klicpera et al.

GINConv

Xu et al.

DiffusionConv

Li et al.

GatedGraphConv

Li et al.

AGNNConv

Thekumparampil et al.

TAGConv

Du et al.

CrystalConv

Xie & Grossman

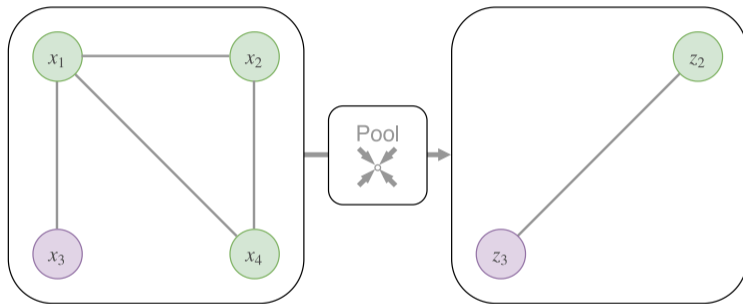
EdgeConv

Wang et al.

MessagePassing

Gilmer et al.

Pooling nelle Graph Neural Networks



Obiettivo: ridurre il numero di nodi.

Pooling nelle Graph Neural Networks

Model-free

Indipendenti dal task

Strategia predefinita

Teoria dei grafi

[2], [3]

Model-based

Dipendenti dal task

Imparano a ridurre i grafi

Basati su euristiche

[4], [5]

[2] I. S. Dhillon *et al.*, "Weighted graph cuts without eigenvectors a multilevel approach," 2007.

[3] F. M. Bianchi *et al.*, "Hierarchical Representation Learning in Graph Neural Networks with Node Decimation Pooling," 2019.

[4] R. Ying *et al.*, "Hierarchical Graph Representation Learning with Differentiable Pooling," 2018.

[5] H. Gao *et al.*, "Graph U-Nets," 2019.

Spectral Clustering with Graph Neural Networks for Graph Pooling

Filippo Maria Bianchi^{*1} Daniele Grattarola^{*2} Cesare Alippi^{2,3}

Abstract

Spectral clustering (SC) is a popular clustering technique to find strongly connected communities on a graph. SC can be used in Graph Neural Networks (GNNs) to implement pooling operations that aggregate nodes belonging to the same cluster. However, the eigendecomposition of the Laplacian is expensive and, since clustering results are graph-specific, pooling methods based on SC must perform a new optimization for each new sample. In this paper, we propose a graph clustering approach that addresses these limitations of SC. We formulate a continuous re-

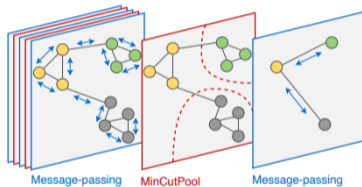
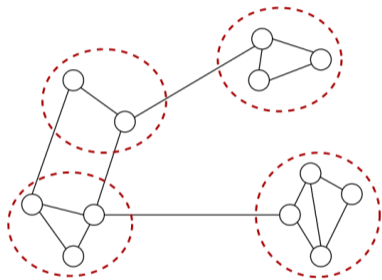
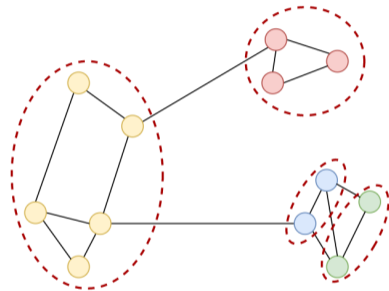


Figure 1. A deep GNN architecture where message-passing is followed by the MinCutPool layer.

Spectral Clustering



Spectral clustering standard.



Possiamo migliorarlo?

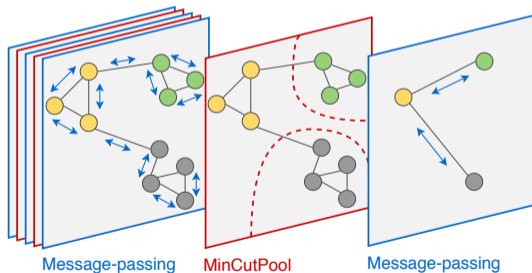
MinCut Pooling

Problemi di SC: non differenziabile, costoso, usa solo gli archi.

Idea: generare i cluster con una rete neurale.

Cerchiamo cluster simili a SC: usiamo il taglio minimo (MinCut) come loss.

La NN può trovare l'equilibrio tra la MinCut loss e la task loss.

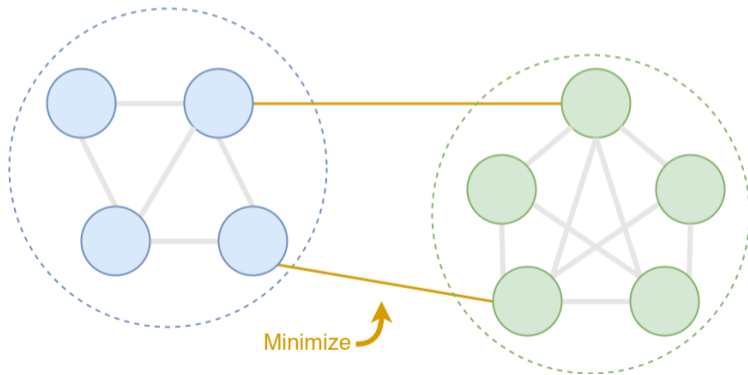


Spectral clustering / MinCut

MinCut: trovare K gruppi di nodi per cui

il volume tra i cluster è minimizzato;

il volume all'interno dei cluster è massimizzato;



Possiamo scrivere il problema MinCut come

$$\text{maximize } \frac{1}{K} \sum_{k=1}^K \frac{C_k^T A C_k}{C_k^T D C_k}, \text{ s.t. } \underbrace{C 1_K = 1_N}_{\substack{1 \text{ nodo} \quad ! \quad 1 \text{ cluster}}}$$

$C \in \{0, 1\}^{N \times K}$ è una matrice di clustering discreta.

Formulazione rilassata

$$\arg \max_{\mathbf{Q} \in \mathbb{R}^{N \times K}} \frac{1}{K} \sum_{k=1}^K \mathbf{Q}_k^T \mathbf{A} \mathbf{Q}_k,$$

$$\text{s.t. } \mathbf{Q} = \mathbf{C}(\mathbf{C}^T \mathbf{D} \mathbf{C})^{-\frac{1}{2}}, \underbrace{\mathbf{Q}^T \mathbf{Q} = \mathbf{I}_K}_{\text{Ortogonale}}, \underbrace{\mathbf{C} \mathbf{1}_K = \mathbf{1}_N}_{\text{Nodi divisi equamente}}$$

$\mathbf{C} \in \mathbb{R}^{N \times K}$ è una matrice di clustering continua.

\mathbf{D} è la matrice di grado.

Soluzione ottimale:

$$Q = U_K O$$

U_K è la base degli autovettori associati ai primi K autovalori
 O è una trasformazione ortogonale.

Spectral clustering

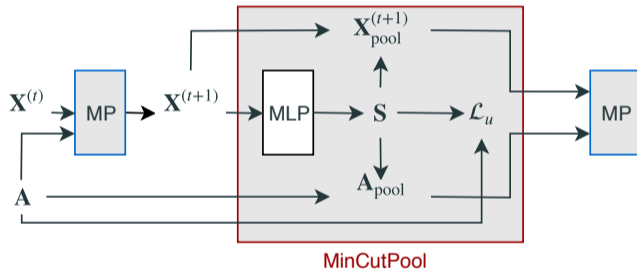
K -means sulle righe di U_K per ottenere C discreta.

MinCut pooling

Impara a generare i cluster:

$$S = \text{MLP}(X)$$

$$\text{Softmax} \Rightarrow S1_K = 1_N$$



MinCut loss:

$$L_c = \frac{\text{Tr}(S^T A S)}{\text{Tr}(S^T D S)}$$

Clustering ottimale ($L_c = 1$):

$s_i = [0.25, 0.25, 0.25, 0.25]$ non va bene

$s_i = [1.00, 0.00, 0.00, 0.00]$

Orthogonality loss (previene i minimi degeneri di L_c):

$$L_o = \left\| \frac{S^T S}{K^T S K_F} - \frac{I_K}{K} \right\|_F$$

$$L_u = L_c + L_o = \underbrace{\frac{\text{Tr}(S^T \tilde{A} S)}{\text{Tr}(S^T \tilde{D} S)}}_{L_c} + \underbrace{\left\| \frac{S^T S}{K S^T S k_F} \frac{I_K}{K} \right\|_F}_{L_o}$$

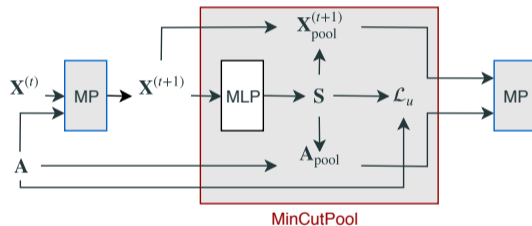
Loss non supervisionata di MinCutPool.

Pooling:

$$A^\theta = S^>AS$$

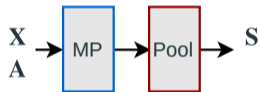
$$X^\theta = S^>X$$

Sommiamo la loss ausiliaria \mathcal{L}_u alla loss supervisionata.

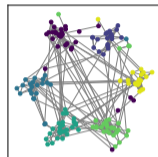


Esperimenti

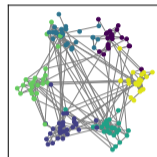
Clustering (grafi semplici)



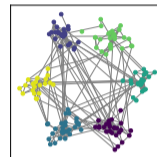
GNN per clustering and segmentazione.



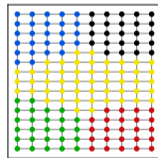
(a) SC



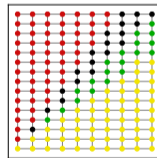
(b) Di Pool



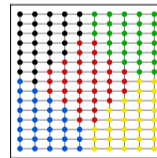
(c) MinCutPool



(d) SC



(e) Di Pool



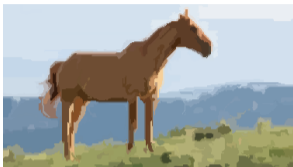
(f) MinCutPool

Clustering di grafi semplici.

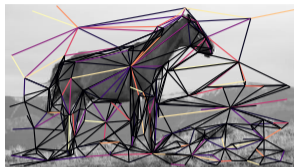
Segmentazione di Immagini



(a) Original



(b) Oversegmentation



(c) Region Adjacency Graph



(d) SC



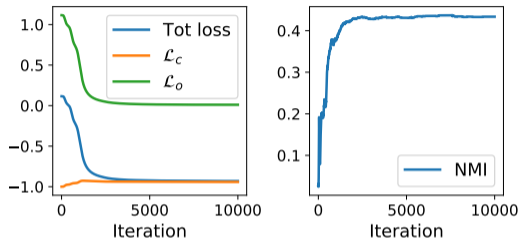
(e) Di Pool ($K = 4$)



(f) MinCutPool ($K = 4$)

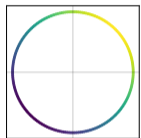
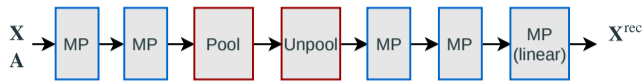
Segmentazione di un'immagine tramite il clustering del suo Region Adjacency Graph.

Clustering (citation networks)

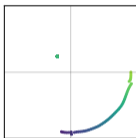


Dataset	K	Spectral clustering				Di Pool				MinCutPool			
		NMI		CS		NMI		CS		NMI		CS	
Cora	7	0.025	0.014	0.126	0.042	0.315	0.005	0.309	0.005	0.404	0.018	0.392	0.018
Citeseer	6	0.014	0.003	0.033	0.000	0.139	0.016	0.153	0.020	0.287	0.047	0.283	0.046
Pubmed	3	0.182	0.000	0.261	0.000	0.079	0.001	0.085	0.001	0.200	0.020	0.197	0.019

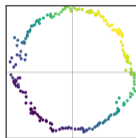
Autoencoder



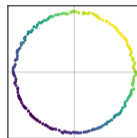
(g) Original



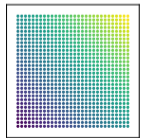
(h) Top-K



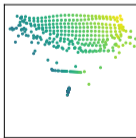
(i) Di Pool



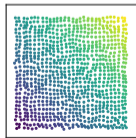
(j) MinCutPool



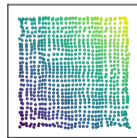
(k) Original



(l) Top-K



(m) Di Pool



(n) MinCutPool

Classificazione di Grafi

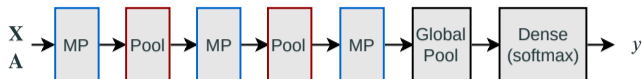


Table 1: Accuracy di classification. I risultati significativamente migliori ($p < 0.05$) sono in grassetto.

Dataset	WL	Dense	No-pool	Graclus	NDP	Di Pool	Top-K	SAGpool	MinCutPool									
Bench-easy	92.6	29.3	0.3	98.5	0.3	97.5	0.5	97.9	0.5	98.6	0.4	82.4	8.9	84.2	2.3	99.0	0.0	
Bench-hard	60.0	29.4	0.3	67.6	2.8	69.0	1.5	72.6	0.9	69.9	1.9	42.7	15.2	37.7	14.5	73.8	1.9	
Mutagenicity	81.7	1.1	68.4	0.3	78.0	1.3	74.4	1.8	77.8	2.3	77.6	2.7	71.9	3.7	72.4	2.4	79.9	2.1
Proteins	71.2	2.6	68.7	3.3	72.6	4.8	68.6	4.6	73.3	3.7	72.7	3.8	69.6	3.5	70.5	2.6	76.5	2.6
DD	78.6	2.7	70.6	5.2	76.8	1.5	70.5	4.8	72.0	3.1	79.3	2.4	69.4	7.8	71.5	4.5	80.8	2.3
COLLAB	74.8	1.3	79.3	1.6	82.1	1.8	77.1	2.1	79.1	1.5	81.8	1.4	79.3	1.8	79.2	2.0	83.4	1.7
Reddit-Binary	68.2	1.7	48.5	2.6	80.3	2.6	79.2	0.4	84.3	2.4	86.8	2.1	74.7	4.5	73.9	5.1	91.4	1.5

Nuovo metodo di pooling per le GNN.

Impara a fare il pooling, ma ha una base teorica.

Risolve problemi di spectral clustering.

Funziona molto bene in pratica.

Speaker

Daniele Grattarola (@riceasphalt)

Contatti

dani ele. grattarol a@usi . ch

fi l i ppombi anch i @gmai l . com

Codice

Disponibile su Spektral e Pytorch Geometric.

<https://github.com/FilippoMB/>

Spectral-Clustering-with-Graph-Neural-Networks-for-Graph-Pooling

- [1] J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl, "Neural message passing for quantum chemistry," *arXiv preprint arXiv:1704.01212*, 2017.
- [2] I. S. Dhillon, Y. Guan, and B. Kulis, "Weighted graph cuts without eigenvectors a multilevel approach," *IEEE transactions on pattern analysis and machine intelligence*, vol. 29, no. 11, pp. 1944–1957, 2007.
- [3] F. M. Bianchi, D. Grattarola, L. Livi, and C. Alippi, "Hierarchical representation learning in graph neural networks with node decimation pooling," *arXiv preprint arXiv:1910.11436*, 2019.
- [4] R. Ying, J. You, C. Morris, X. Ren, W. L. Hamilton, and J. Leskovec, "Hierarchical graph representation learning with differentiable pooling," *arXiv preprint arXiv:1806.08804*, 2018.
- [5] H. Gao and S. Ji, "Graph u-nets," *CoRR*, vol. abs/1905.05178, 2019. arXiv: 1905. 05178. [Online]. Available: <http://arxiv.org/abs/1905.05178>.