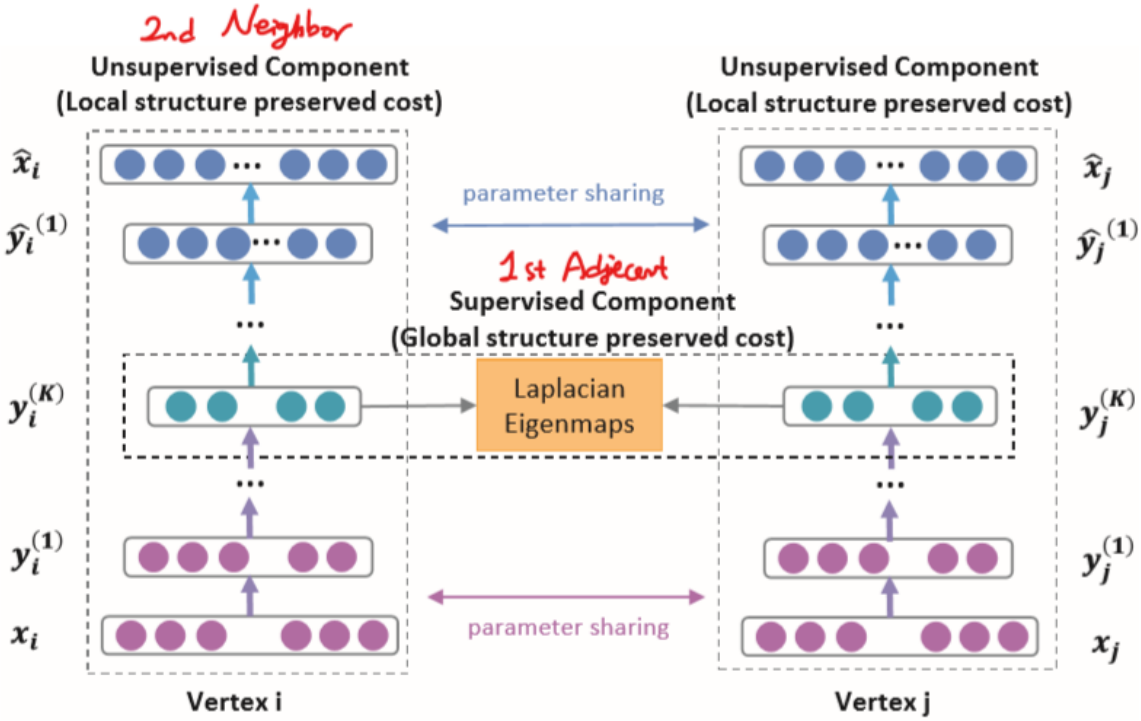
1 Structural Deep Network Embedding

1.1 Summary

This article [WCZ16] didn't pay much attention to data pre-processing, it focused on what features a model must to learn.



This article restricted loss functions to achieve its goals. It hoped that learnt features can preserve local and global information. Assume we have adjacent matrix of a graph A, where

$$A = \begin{bmatrix} - & a_1 & - \\ \vdots & \vdots & \vdots \\ - & a_n & - \end{bmatrix}$$

then, a_i is all the connection between node i and all other nodes, so, it's global structure of node i. On the other hand, since it adapts autoencoder model, a latent feature vector, or one can think it compress a_i to a lower dimensional vector in the middle layer $y_i^{(K)}$. As mentioned above, a_i must be reconstructed from it, and in addition, if node i and node j has connection, i.e., $a_{i,j} > 0$, then, the distance between $y_i^{(K)}$ and $y_j^{(K)}$ should be close, and this is the local information.

In summary, using an autoencoder model can encode or compress $\boldsymbol{a}_i$ to a lower dimensional $y_i^{(k)}$, and $y_i^{(k)}$ includes both local and global structure information of node

i, global information means $DEC(ENC(\boldsymbol{a}_i)) = \boldsymbol{a}_i$, and local information means a small $\|y_i^{(k)} - y_j^{(k)}\|_F^2$. So, for node i and node j we have loss function:

$$\|\widehat{\boldsymbol{a}}_i - \boldsymbol{a}_i\|_2^2 + \|\widehat{\boldsymbol{a}}_j - \boldsymbol{a}_j\|_2^2 + a_{i,j}\|y_i^{(K)} - y_j^{(K)}\|_2^2$$

References

[WCZ16] Daixin Wang, Peng Cui, and Wenwu Zhu. Structural deep network embedding. In *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1225–1234, 2016.