Furthermore, the generalization error gap of GCNs tends to be enlarged with more layers, which can interpret why GCNs with deeper layers have relatively poorer performance in test datasets.

此外，随着层数增加GCNs的泛化误差差距逐渐增大，这就可以解释了较深层的GCNs在测试数据集中的性能相对较差的原因。

Since Convolutional Neural Networks(CNNs) have relatively limited performance in processing such graph data, Graph Neural Networks(GNNs) are gaining fast development, which can operate directly on non-Euclidean data.

由于卷积神经网络（CNNs）在处理此类图数据方面的性能相对有限，因此图神经网络（GNNs）正在获得快速的发展，它可以直接处理非欧几里得数据。

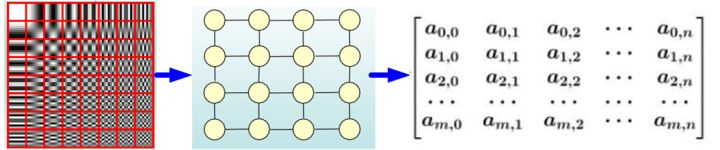
总的来说，数据类型可以分为两大类：

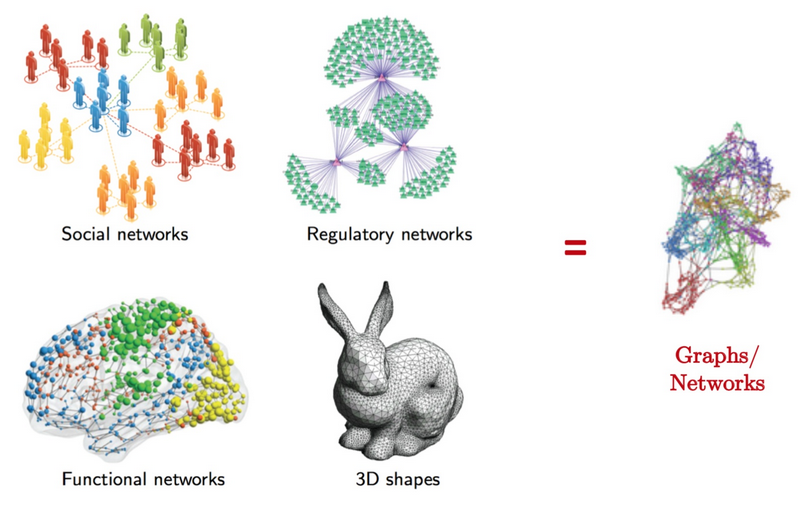
欧几里德结构数据(Euclidean Structure Data)

非欧几里德结构数据(Non-Euclidean Structure Data)

所谓的欧几里德数据指的是类似于grids（网格）, sequences（序列）… 这样的数据，例如图像就可以看作是2D的grid数据，语音信号就可以看作是1D的grid数据。但是现实的处理问题当中还存在大量的 Non-Euclidean Data，如社交多媒体网络（Social Network）数据，化学成分（Chemical Compound）结构数据，生物基因蛋白（Protein）数据以及知识图谱（Knowledge Graphs）数据等。

Euclidean Structure Data



Non-Euclidean Structure Data

α-Lipschitz continuous：α为系数

v -Lipschitz smooth：v为系数

dataset S：数据集S

T iterations：T次迭代

m samples：m个样本

Q is the upper bound of the loss function defined in assumption B, and M is the upper bound of the weights defined in assumption C.

Q是假设B中定义的损失函数的上界，M是假设C中定义的权值的上界。

the expectation Esgd is taken over the randomness inherent in SGD, Q is a constant depending on the loss function,

期望Esgd取代了SGD中固有的随机性，Q是一个依赖于损失函数的常数，

In the GCN model, SGD is selected to minimize the loss function by optimizing the parameters with initialization, which means that the loss will not increase tremendously during the training procedure.

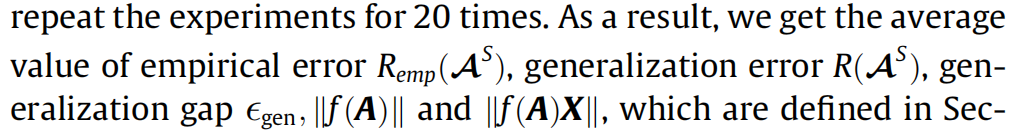
在GCN模型中，通过初始化优化参数，选择SGD来最小化损失函数，这意味着在训练过程中损失不会显著增加。

It means if a randomized algorithmis uniform stable, the generalization gap can be controlled. Therefore, if the uniform stability coefficient can be determined, the generalization gap can be computed.

这意味着，如果一个随机算法是均匀稳定的，则可以控制泛化间隙。因此，如果可以确定均匀稳定系数，就可以计算出泛化间隙。

For a GCN with L layers, the generalization gap is connected tightly with L. The gap may increase exponentially with the depth L, which may be the reason why GCN cannot be too deep.

对于具有L层的GCN，泛化间隙与L紧密相连，间隙可能随着深度L的增加呈指数增长，这可能是GCN不能太深的原因。



经验误差 泛化误差 泛化间隙

The norm of graph filters and its product with node features should be small to obtain a relatively small uniform stability and generalization guarantees, while the number of GCN layers should be selected approximately to obtain enough expressive as well as generalization ability.

图滤波器及其节点特征乘积的范数应较小，以获得相对较小的均匀稳定性和泛化保证，而应近似选择GCN层的数量，以获得足够的表达能力和泛化能力。