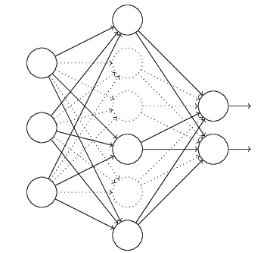
**dropout**

//参考 深度学习7.12

Dropout提供了一种廉价的Bagging集成近似,能够训练和评估指数级的神经网络。具体而言,Dropout训练的集成包括所有从基本的基础网络除去非输出单元形成子网络

超参数的采样概率为 1，隐藏层的采样概率通常为 0.5，输入的采样概率通常为 0.8。

****

**dropout为什么有效：**

1.bagging 2.参数共享 3.噪声

在Bagging的情况下，每一个模型在其相应训练集上训练到收敛。在Dropout的情况下，通常大部分模型都没有显式地被训练。而是在单个步骤中我们训练一小部分的子网络，参数共享会使得剩余的子网络也能有好的参数设定。除了这些，Dropout与Bagging算法一样。Dropout强大的大部分原因来自施加到隐藏单元的掩码噪声。权值的更新不再依赖于有固定关系的隐含节点的共同作用，阻止了某些特征仅仅在其它特定特征下才有效果的情况。迫使网络去学习更加鲁棒的特征。这可以看作是对输入内容的信息高度智能化、自适应破坏的一种形式，而不是对输入原始值的破坏。