tf的构建流程：

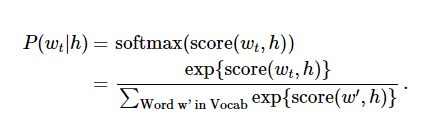
构造模型🡪loss函数🡪初始化变量🡪启动图(Session)

# 字词的向量表示

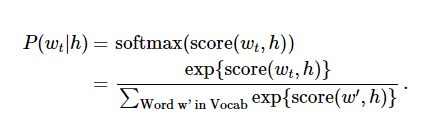
Word2vec是一种可以进行高效率词嵌套学习的预测模型。其两种变体分别为：连续词袋模型（CBOW）及Skip-Gram模型。从算法角度看，这两种方法非常相似，其区别为CBOW根据源词上下文词汇（'the cat sits on the'）来预测目标词汇（例如，‘mat’），而Skip-Gram模型做法相反，它通过目标词汇来预测源词汇。Skip-Gram模型采取CBOW的逆过程的动机在于：CBOW算法对于很多分布式信息进行了平滑处理（例如将一整段上下文信息视为一个单一观察量）。很多情况下，对于小型的数据集，这一处理是有帮助的。相形之下，Skip-Gram模型将每个“上下文-目标词汇”的组合视为一个新观察量，这种做法在大型数据集中会更为有效。本教程余下部分将着重讲解Skip-Gram模型。

## 处理噪声对比训练

神经概率化语言模型通常使用[极大似然法](https://en.wikipedia.org/wiki/Maximum_likelihood) (ML) 进行训练，其中通过 [softmax function](https://en.wikipedia.org/wiki/Softmax_function" \t "_blank) 来最大化当提供前一个单词 **h** (代表 "history")，后一个单词的概率 http://www.tensorfly.cn/tfdoc/images/vr1.png (代表 "target")，



当 **score(w\_t,h)** 计算了文字 **w\_t** 和 上下文 **h** 的相容性（通常使用向量积）。我们使用对数似然函数来训练训练集的最大值，比如通过：



## Variables: 创建，初始化，保存，和恢复

TensorFlow Variables 是内存中的容纳 tensor 的缓存。这一小节介绍了用它们在模型训练时(during training)创建、保存和更新模型参数(model parameters) 的方法。

当训练模型时，用[变量](http://www.tensorfly.cn/tfdoc/api_docs/python/state_ops.html)来存储和更新参数。变量包含张量 (Tensor)存放于内存的缓存区。建模时它们需要被明确地初始化，模型训练后它们必须被存储到磁盘。这些变量的值可在之后模型训练和分析是被加载。