参考链接：

http://blog.csdn.net/eastmount/article/details/50637476

**神经网络概率语言模型**

神经网络概率语言模型通过学习训练语料获取词向量和概率密度函数，词向量是多维实数向量，向量中包含了自然语言中的语义和语法关系，词向量之间余弦距离的大小代表了词语之间关系的远近，词向量的加减运算则是计算机在"遣词造句"。

词向量具有良好的语义特性，是表示词语特征的常用方式。词向量的每一维的值代表一个具有一定的语义和语法上解释的特征。故可以将词向量的每一维称为一个词语特征。词向量用Distributed Representation表示，一种低维实数向量。

NLP中最直观、最常用的词表示方法是One-hot Representation。每个词用一个很长的向量表示，向量的维度表示词表大小，绝大多数是0，只有一个维度是1，代表当前词。但这种One-hot Representation采用稀疏矩阵的方式表示词，在解决某些任务时会造成维数灾难，而使用低维的词向量就很好的解决了该问题。低维的词向量让相似或相关的词在距离上更加接近。

词向量是一个稠密、低维的实数限量，它的每一维表示词语的一个潜在特征，该特征捕获了有用的句法和语义特征。其特点是将词语的不同句法和语义特征分布到它的每一个维度上去表示。