**词向量共生概率**

实际上，思科（Cisco）进行过评估，在2016年， 互联网上超过85%的信息都是以像素的形式呈现在我们眼前，实际上我们已经进入一个视觉的时代，图片和视频的信息时代。由于互联网作为信息的载体和大量视觉传感器的实际使用如智能手机、行车记录仪、监控摄像头等，造成如今的互联网视觉信息大爆发。李飞飞教授称视觉信息为“互联网中的暗物质”，类比于银河系中的暗物质，据说银河系中85%的质量属于暗物质，但它们非常难以被检测和观察到。

YouTube服务器上，每60秒就会接收150小时以上的视频上传，Facebook上，每天新上传数亿张照片，如此庞大的数据量，人类靠眼睛完全无法完整浏览这些数据，给如此大量的数据进行标记、分类、索引是一项复杂且庞大的工作，依靠人类手工地去处理如此大量的数据是不可能完成的，唯一的希望就是依靠计算机视觉技术。能够对图片进行标记、分类，处理视频中的每一帧。

一些技术发展促进了前所未有的图像共享。首先是通过相机，手机和其他可穿戴传感设备以数字格式捕获图像的便利性。第二种是互联网，允许将数字图像内容传输到世界上任何地方的任何人。最后，最近，通过大规模采用社交网络平台，数字图像的共享达到了新的高度。所有突然的图像都带有标签，任何数字图像的标记，评论和评级都成为一种常见的习惯。尽管图像和标签倾盆大雨，搜索和查找特定图像的问题仍然很大程度上尚未解决。相反，它通过可靠和客观的图像标签的需求来扩大问题。

WordNet表面上类似于词库，因为它根据词的含义将词汇组合在一起。但是，有一些重要的区别。首先，WordNet不仅仅是单词形式 - 字母串 - 而是单词的特定意义。结果，在网络中彼此非常接近的单词在语义上被消除歧义。其次，WordNet标记了单词之间的语义关系，而同义词库中单词的分组并不遵循除意义相似性之外的任何显式模式。

1967年美国社会心理学家[米尔](https://baike.baidu.com/item/%E7%B1%B3%E5%B0%94" \t "_blank)[格伦](https://baike.baidu.com/item/%E6%A0%BC%E4%BC%A6" \t "_blank)（Stanley Milgram）提出了一个“六度分离”理论（Six Degrees of Separation）：该理论认为在人际交往的[脉络](https://baike.baidu.com/item/%E8%84%89%E7%BB%9C" \t "_blank)中，任意两个陌生人都可以通过“亲友的亲友”建立联系，这中间最多只要通过五个朋友就能达到目的。也就是说，最多通过五个人你就能够认识世界上任何一个陌生人，不管对方在哪个国家，属哪类人种，是哪种肤色。以此类推，我们可以认为：对于任意两个标签词之间，可以通过多个中间标签词找到他们之间的联系，中间间隔的中间标签词定义了两个标签词的语义距离。我们可以认为语义距离越小的两个标签词出现在同一张图片上的概率越大。

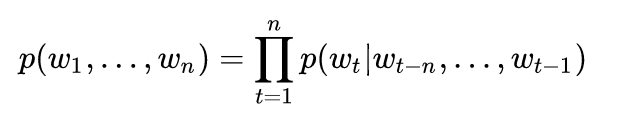
**分布式表示**

所有词都彼此孤立显然是不符合现实情况的。与独热编码不同，使用特定维度的低维实数向量来表示单词的方法称分布式表示。是一种稠密的表示方式。这种向量一般是这样的：

[0.712 -0.05 -0.343 0.156 ...]（数字随便编的）

其最大的好处在于将语义信息融合进了词向量中，使得基于衡量向量距离的方法（如欧式距离、余弦距离等）可以用来度量单词间的相似度。直观上来说，就是语义上相近或相似的词，在向量空间上的距离会更接近，而语义相反的词距离会更远。

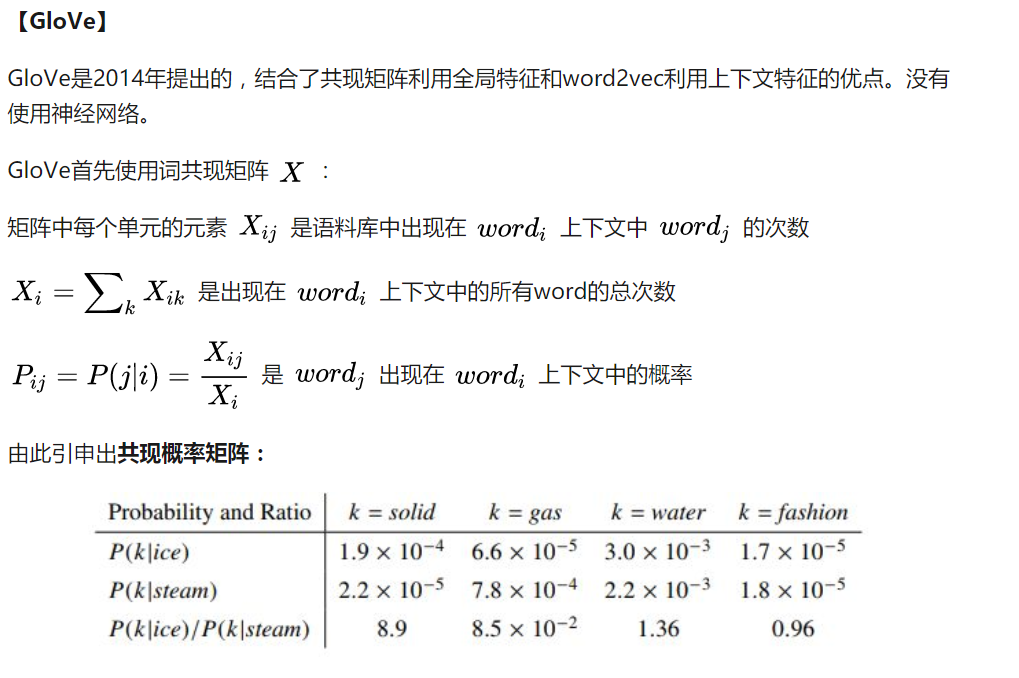
常见的统计语言模型为N-gram Model。句子得分概率的计算为：



上面这条公式右半部分的条件概率是大多数模型的目标。

用这种方法能一定程度上捕获到上下文信息。

词向量的分布式表示的核心思想由两部分组成：一、选择一种方式描述上下文；二、选择一种模型刻画目标词与其上下文之间的关系。



向量空间模型(VSM)的余弦定理公式(cos)

https://blog.csdn.net/whzhaochao/article/details/50637300

在ImageNet上1000类分类训练好的参数的基础上，根据我们的分类识别任务进行特定的微调。

ResNet不仅是优秀的图像分类器，它也是很有用的图像特征提取器。