我们如何让今天的计算机在文本数据上执行聚类，分类等等，因为我们知道它们在处理和处理字符串或文本方面通常效率低下，无法取得丰硕成果。

当然，一台电脑可以匹配两个字符串，并告诉你它们是否相同。但是当你搜索梅西时，我们如何让电脑告诉你足球或罗纳尔多？你怎么让电脑明白，“苹果是美味的水果”中的“苹果”是可以吃的水果，而不是公司？

上述问题的答案在于为单词创建一个表示形式，以捕捉它们的含义，语义关系以及它们所使用的不同类型的上下文。

所有这些都是通过使用Word嵌入或文本的数字表示来实现的，以便计算机可以处理它们。

Word嵌入是将文本转换为数字，并且可能有相同文本的不同数字表示。

许多机器学习算法和几乎所有深度学习架构都不能以原始形式处理字符串 或纯文本 。他们需要数字作为输入来执行任何工作，无论是分类，回归等。而以文本格式存在的大量数据，就必须从中提取知识并构建应用程序.

Co-occurrence : 对于一个给定的语料库，一对词的共同出现。对于w1和w2是它们在上下文窗口中一起出现的次数。

Context Window：上下文窗口由数字和方向指定。

问题：

为什么CBOW 用隐藏层和输出层之间的权重作为单词的单词向量表示？Skip gram可以用 输入层和隐藏层之间的权重作为训练后的单词向量表示？