



现在的任务是输入 $w_{t-n+1}, \dots, w_{t-1}$ 这前 $n-1$ 个单词，然后预测出下一个单词 w_t

数学符号说明：

- $C(i)$: 单词 w 对应的词向量，其中 i 为词 w 在整个词汇表中的索引
- C : 词向量，大小为 $|V| \times m$ 的矩阵
- $|V|$: 词汇表的大小，即预料库中去重后的单词个数
- m : 词向量的维度，一般大于 50
- H : 隐藏层的 weight
- d : 隐藏层的 bias
- U : 输出层的 weight
- b : 输出层的 bias
- W : 输入层到输出层的 weight
- h : 隐藏层神经元个数

计算流程：

1. 首先将输入的 $n - 1$ 个单词索引转为词向量，然后将这 $n - 1$ 个向量进行 concat，形成一个 $(n - 1) \times w$ 的矩阵，用 X 表示
2. 将 X 送入隐藏层进行计算， $\text{hidden}_{out} = \tanh(d + X * H)$
3. 输出层共有 $|V|$ 个节点，每个节点 y_i 表示预测下一个单词 i 的概率， y 的计算公式为 $y = b + X * W + \text{hidden}_{out} * U$