Perplexity—ppl

语言模型，给出一句话的前k个词，希望它可以预测第k+1个词是什么，即给出一个第k+1个词可能出现的概率的分布。

**Perplexity定义**

ppl是用在自然语言处理领域(NLP)中，衡量语言模型好坏的指标。它主要是根据每个词来估计一句话出现的概率，并用句子长度做normalize，公式为：

其中S代表句子sentence，N是句子长度，是第i个词的概率。第一个词就是，而是start，表示句子的起始，是个占位符。

这个式子可以这样理解，ppl越小，p()则越大，依据我们期望的sentence出现的概率也就越高。

还有一些人说，ppl可以认为是average branch factor(平均分支系数)，即预测下一个词时可以有多少种选择。比如ppl下降到90，可以直观的理解为，在模型生成下一句话时下一个词有90个合理选择，可选词数越少，可以认为模型越准确。

**Perplexity的影响因素**

* 训练数据集越大，PPL会下降的更低。
* 数据中的标点会対模型的PPL产生很大影响，一个句号能够让PPL波动几十，标点的预测总是不稳定。
* 预测语句中的“的，了”等词也对PPL有很大影响，可能“我借你的书”比“我借你书”的指标值小几十，但是从语义分析上有没有这些停用词并不能完全代表句子生成的好坏。