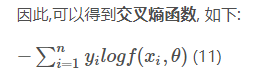
1. 集束搜索（beam search）

主要是用于测试时候进行句子翻译的方法

<https://blog.csdn.net/qq_16234613/article/details/83012046>

1. 交叉熵

主要是用于计算训练模型f(x,θ)中θ的最优值，利用最大似然函数还计算参数组θ的最优值



交叉熵函数可用于多分类multi-classification任务中的损失函数.

CNN 中, 输出层可以采用 softmax 来输出每一类的概率值, 再将 softmax 输出结果, 送入交叉熵损失函数.

Categorical CrossEntropy 多分类的对数损失函数

<https://blog.csdn.net/zziahgf/article/details/79928978>

1. 学习速率

机器学习中的学习速率是指每一步调整的步长，学习速率太小，收敛速度过慢，学习速率太大，会导致代价函数震荡，无法收敛