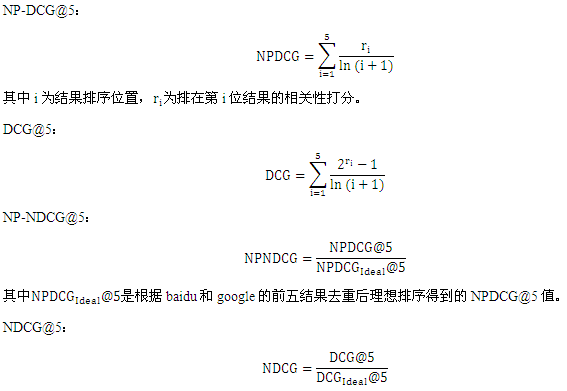
**背景**

DCG（及其衍生指标）被业界广泛用来评估整体（前n整体）搜索结果满意度。由于人工评估效率较低，评估样本数量有限，且对评估人员的专业性有严格要求，机器DCG指标（我们称之为auto-DCG）在大规模数据评估及挖掘中，便有了较大的作用空间。

**Auto-DCG**

Auto-DCG的关键在于如何对每条结果自动评分，有了评分之后，直接基于公式计算即可得到整体（前n）DCG值。此前评估过，click quality与人工打分的一致性约为85%，不一致case的主要原因是，未考虑点击率因素。例如搜“肯德基”，展现麦当劳结果，点击率较低，但是一旦点击，click quality并不差（按照人工评估标准，这种竞品分值较低）。因此，综合考虑ctr和click quality，即可得到与人工评估高度一致的评分结果，进而得到高度一致的DCG值。实际中，选取 score = ctr \* clickq （后续优化思路包括引入指数权重，对ctr进行position bias normalization等）。

DCG及其衍生指标的计算公式如下（来自webpm周敏）：



Auto-DCG的相关应用如下。

**强商业query挖掘**

从用户体验角度出发，判断给定query下是否适合展示广告（尤其是多媒体广告），最重要的评估标准是，展现广告与否时，搜索结果整体（前n）满意度如何变化。如果展示广告使得整体满意度大幅下降，则该query下不适合展现广告。

上述判断逻辑可以通过auto-DCG实现：

* 基于历史数据，选定该query下score最高的广告（只选取中高频广告）
* 计算展现该条广告与否时，整体（前n）结果DCG的差
* Ctr，clickq使用统计值，自然结果展现较为稳定，可以基于“query-位次”统计
* 差值大于一定阈值时，该query下适合展示广告（对整体用户体验正向影响，或负向影响较小）

**基于广告数据优化自然结果**

由于广告的targeting过程包含了广告主的智慧（广告主会针对性地购买相关拍卖词），因此，针对一些变体、缩写、错拼等情况，广告有时能够更准确地进行识别。基于auto-DCG挖掘（展现广告时，DCG值明显高于无广告的情况）的一些case如下：





**搜索满意度预估**

Auto-DCG反映了搜索结果整体满足程度，可以作为 [搜索满意度](http://wiki.baidu.com/pages/viewpage.action?pageId=86030043)预估模型的重要特征。

Auto-DCG与搜索满意度指标的定位区别在于：Auto-DCG首先针对每条结果计算score，再合成整体score，而搜索满意度本身为整体指标。因此，Auto-DCG可以用于拼凑不同次展现的结果，适合用于强商业query挖掘的场景（判断展现最好的广告时，整体满足度如何变化），但只能针对中高频结果进行计算。而搜索满意度指标更适合用于整体评估，无论query及其展现结果是否低频。