中文句子质量评价：

（相似度分析）

1. 英文title数量较少，可以手动转为中文

无监督方法：

主题模型（LDA、PLSA、LSA、lda2vec .etc）