

# 千言：文本生成一致性评测

## 1. 任务描述

本任务计划使用三个任务数据集测试参赛系统的生成能力，包括文案生成<sup>[1]</sup>、摘要生成<sup>[2]</sup>和问题生成<sup>[3]</sup>：

- 文案生成根据结构化的商品信息生成合适的广告文案；
- 摘要生成是为输入文档生成简洁且包含关键信息的简洁文本；
- 问题生成则是根据给定段落以及答案生成适合的问题。

显然，这三个任务对生成结果的事实一致性均有较高要求。

本次比赛的赛题设置有几个特点：

- 面向应用。三个任务数据都是从实际应用场景构建的任务，任务有很强的实用性。
- 任务形式多样。覆盖数据到文本生成，摘要生成，问题生成等多类任务。
- 评估方式更全面。同时评估生成文本的流畅性和事实准确性。

## 2. 评测指标

我们将同时评估参赛系统生成结果的事实一致性和文本流畅性，并取系统在所有任务上所有指标的平均值作为最终成绩。

事实一致性指标：由于任务 1 的输入有明确的事实数据，所以采用专门的评估指标，对任务 2、3 采用通用的评估指标，具体如下：

- a) 对任务 1，采用 PARENT 指标<sup>[4]</sup>，同时将生成的句子和参考文本、输入表格信息比较，在兼顾参考答案的同时，评价生成内容是否忠于输入表格的信息
- b) 对任务 2、3，利用提前构建的文本蕴含模型，通过衡量参考答案与文本生成结果的蕴含关系，作为衡量事实一致性的通用评估指标

文本流畅性指标：使用 BLEU-4，基于参考答案和预测结果，计算 n-gram 的匹配度。

## 基线系统

本次竞赛将提供基于飞桨框架 PaddlePaddle2.0 和 PaddleNLP 的开源基线系统，提供丰富的高层 API，全面支持动态图开发模式，从开发、训练到预测部署提供优质的整体体验。推荐您参照基线方案，基于 PaddleNLP 的丰富 API 库进行二次开发、模型调优和方案创新。

GitHub 基线系统：

[https://github.com/PaddlePaddle/PaddleNLP/tree/v2.0.8/examples/text\\_generation/uni-mo-text](https://github.com/PaddlePaddle/PaddleNLP/tree/v2.0.8/examples/text_generation/uni-mo-text)

AI Studio 基线系统：

<https://aistudio.baidu.com/aistudio/projectdetail/2303979>

AI Studio 提供 100 小时免费 V100 GPU 算力，算力码申请链接：

<https://aistudio.baidu.com/aistudio/questionnaire?activityid=1404>

## 参考文献

[1] Zhihong Shao, Minlie Huang, Jiangtao Wen, Wenfei Xu, Xiaoyan Zhu. 2019. Long and Diverse Text Generation with Planning-based Hierarchical Variational Model. In Proceedings of EMNLP 2019.

[2] Baotian Hu, Qingcai Chen, Fangze Zhu. 2015. LCSTS: A Large Scale Chinese Short Text Summarization Dataset. In Proceedings of EMNLP 2015.

[3] Wei He, Kai Liu, Jing Liu, Yajuan Lyu, Shiqi Zhao, Xinyan Xiao, Yuan Liu, Yizhong Wang, Hua Wu, Qiaoqiao She, Xuan Liu, Tian Wu, Haifeng Wang. 2019. DuReader: a Chinese Machine Reading Comprehension Dataset from Real-world Applications. In Proceedings of ACL 2018 MRQA Workshop.

[4] Bhuwan Dhingra, Manaal Faruqui, Ankur Parikh, Ming-Wei Chang, Dipanjan Das, William W. Cohen. 2019. Handling Divergent Reference Texts when Evaluating Table-to-Text Generation. In Proceedings of ACL 2019.