

附件 5

2023 年第五届
全球校园人工智能算法精英大赛
(算法专项赛)

赛
题
及
比
赛
规
则

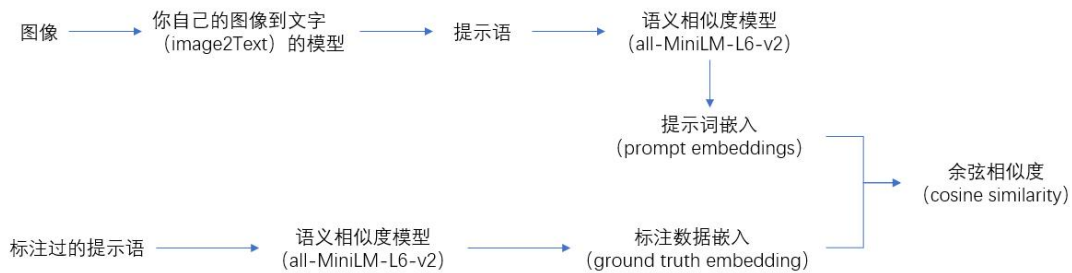
全球校园人工智能算法精英大赛组委会

2023 年 9 月

图像文字说明生成算法

一、赛题说明

1、图像文字说明生成算法比赛的目标为创建一个模型来预测给定生成图像的文本提示。参赛选手将在包含 **Stable Diffusion 2.0** 生成的各种（提示、图像）对的数据集上进行预测，通过了解潜在存在的提示、图像之间关系的可逆性。参赛选手通过构建一个模型来预测给定生成图像的文本提示。并把这个文本提示与标注过的文本提示进行对比。



2、提示工程（**Prompt Engineering**）是一种针对预训练语言模型（如 **ChatGPT**），通过设计、实验和优化输入提示来引导模型生成高质量，准确和有针对性的输出的技术。由此产生的提示学习是一种通过构建合适的输入提示来解决特定任务的方法。本赛题就是通过构建数据模型来研究如何提升提示工程的效果。

3、文本到图像模型的流行已经是基于提示工程的一个人工智能全新领域。用户体验的一部分是艺术，一部分是充满不确定性的数据科学，机器学习工程师正在迅速努力理解提示和它们生

成的图像之间的关系。在提示中添加“4k”是使其更具照片感的最佳方式吗？提示中的微小扰动是否会导致图像高度发散？提示关键字的顺序如何影响生成的场景？这项竞赛的任务是创建一个模型，该模型可以可靠地反转生成给定图像的扩散过程。

4、本赛题任务是预测用于生成目标图像的提示。这个挑战的提示是使用各种（未公开）方法生成的，从简单到复杂，都有多个对象和修饰符。使用 **Stable Diffusion 2.0 (768-v-ema.ckpt)** 根据提示生成图像，并以 768x768 像素的 50 步生成图像，然后将竞争数据集的图像缩小到 512x512。使用了此脚本，参见参考代码库。

5、为了以一种稳健的方式计算提示的相似性。

二、解题思路

可以通过多种图片生成文字模型形成自己的数据集，并训练了一个模型以有监督的方式或者无监督的方式生成句子嵌入。

为了提升相似性，可以采用模仿标注的数据集形成“填空问题回答”方式来完成相应题目。

参考文献：

1、A dataset and exploration of models for understanding video data through fill-in-the-blank question-answering

2、NICE 2023 Zero-shot Image Captioning Challenge

三、评价方式

使用预测的和实际的提示嵌入向量之间的平均余弦相似性

得分来评估提交。关于如何计算标注过提示的嵌入的精确细节，参阅组委会相关参考代码。

四、成绩提交

对于测试集中的每个图像，必须预测用于生成图像的提示，并将提示转换为 384 长度的嵌入向量。预测应该被展平为图像 (imgId) 和嵌入 (eId) 对 (imgId_eId) 的行。

```
imgId_eId,val
20057f34d_0,0.018848453
20057f34d_1,0.030189732
....
20057f34d_383,-0.007934112
227ef0887_0,0.017384542
etc.
```

五、赛事说明

(一) 省赛

- 1、省赛能够做运算力平台上活动比较好的成绩，依据系统自动评分获得奖励；（100分）
- 2、每个获奖队伍能够获得 1000 元的算能积分大礼包；
- 3、冠军队能够获得校招面试直通券；
- 4、获奖队伍活动实习终面直通券；
- 5、TPU 编程竞赛委员会加入邀请函；
- 6、企业参访计划。

(二) 总决赛

1、总决赛评分规则

(1) 省赛能够在边缘运算力平台上活动比较好的成绩，依据系统自动评分获得奖励（60分）；

(2) 能够基于这个提升工程应用，提出相应合理的应用场景（20分）；

(3) 完美现场（线上/线下）讲解和答辩（20分）。

2、提供丰厚奖品

奖项	奖金	奖品	名额
一等奖	5000元	每个队能够获得一个能跑大模型的边缘盒子（边缘算力盒子）	一名
二等奖	3000元	每个队能够获得一个能跑大模型的边缘盒子（边缘算力盒子）	二名
三等奖	1000元	每个队能够获得一个能跑大模型的边缘盒子（边缘算力盒子）	三名

(1) 每个获奖队伍能够获得1000元的算能积分大礼包；

(2) 冠军队能够获得校招面试直通券；

(3) 获奖队伍活动实习终面直通券；

(4) TPU编程竞赛委员会加入邀请函；

(5) 企业参访计划。

六、联系方式

参赛QQ群：791291365