《人工智能》课程设计一

任务说明：

提供2000张标注了的车辆场景分类信息的高分辨率图片，请使用这些数据，在MindSpore等框架下建立并训练模型，并将此模型运用于测试数据集的图像分类标注。

提示：

该题目主要考核各位同学使用较小的模型和小规模的训练数据下的模型设计和训练的能力。需要同学充分利用迁移学习等方法，以解决训练数据少的问题，同时需要思考和解决深度学习中过拟合的问题。所提供的场景定义为比较单一的车辆场景，一定程度上平衡了难度系数。

标签信息：

0，巴士，bus

1，出租车，taxi

2，货车，truck

3，家用轿车，family sedan

4，面包车，minibus

5，吉普车，jeep

6，运动型多功能车，SUV

7，重型货车，heavy truck

8，赛车，racing car

9，消防车，fire engine

文件信息：

train.rar：训练集

val.rar：验证集

test.rar：测试集

《人工智能》课程设计二

**任务说明：**

（1）基于序列到序列（Seq2Seq）学习框架，在MindSpore等框架下设计并训练一个中英文机器翻译模型，完成中译英和英译中翻译任务。具体模型选择可以参考如LSTM，GRU，Transformer等，但不做限制；

（2）实验数据集为WMT18新闻评论数据集[News Commentary v13](https://aistudio.baidu.com/datasetdetail/181862)[[1]](#footnote-0)，整个语料库分训练集（约252,700条）、验证集和测试集（分别约2,000条）三部分，每部分包含中英文平行语料两个文件，全部语料已预分词处理；

（3）根据指定的评价指标和测试集数据评价模型性能。

**评价指标：**

（1）BLEU1/2/3：

BLEU（Bilingual Evaluation Understudy）计算同时出现在系统译文和参考译文中的n元词重叠程度，实验要求计算n=1/2/3。

（2）Perplexity：

Perplexity是衡量语言模型生成句子能力的评价指标，计算perplexity的公式如下：

其中：

，参考译文

，参考译文中的第i个单词

，参考译文句长，即译文中包含的单词数

**提交内容：**

（1）模型源码；

（2）实验报告：包括问题分析，算法设计，数据处理，实验分析及结果等。

1. https://aistudio.baidu.com/datasetdetail/181862 [↑](#footnote-ref-0)