

实验5 综合实验（大作业）



考核方式

- 实验环节：40%
 - ✓ 实验1（Pytorch基本操作实验） 10%
 - ✓ 实验2（前馈神经网络实验） 10%
 - ✓ 实验3（卷积神经网络实验） 10%
 - ✓ 实验4（循环神经网络实验） 10%
- 综合实验+小组汇报：60%
 - ✓ 实验5（大作业，综合实验） 30%
 - ✓ 小组汇报（5人组随机汇报，时长不超过8分钟，组长提交ppt） 30%

标记清楚实验一、实验二、...，汇报ppt，分清章节提交，实验报告文件命名格式：学号_姓名.doc/.md/.pdf，坚果云收件箱提交

<https://send2me.cn/9fG83pfy/Ts6A67HVAyMXhw>

课堂报告总体要求

■ 本次课堂报告内容分组进行（可组内按顺序讲解或课堂上随机选一名组员负责讲解）

■ 从5个讲解材料中**任选**（不少于10分钟，不多于15分钟）

■ 本次实验共包含以下四道题目：

题目1：注意力机制

题目2：生成对抗网络

题目3：Transformer

题目4：图卷积网络

题目5：领域论文综述（任意领域）

总体要求

■ 本次实验要求专业课同学必做，平台课同学选做

■ 从4道题目中**任选1题**来做

■ 本次实验共包含以下四道题目：

题目1：出租车流量预测（时空数据学习领域）

题目2：疫情微博情绪分类（NLP领域）

题目3：单目标跟踪（计算机视觉领域）

题目4：顶会论文复现（任意领域）

题目一：出租车流量预测

■ 任务概述

车流量预测任务是一个回归任务，旨在根据区域历史的车流量情况来预测其未来某一段时间的车流量情况。使用的数据为纽约市出租车流量数据。**输入为纽约市各区域的历史车流量时间序列，输出为对应各区域的未来车流量的预测值。**

■ 数据集说明

纽约出租车流量数据集，时间跨度为从2015年1月1日到2015年3月1日。数据处理成为网格流量数据时的时间间隔设定为30分钟。后20天数据被划定为测试集，其余数据为训练集。

● 使用方式

```
volume_train=np.load(open(train_path, "rb"))["volume"]  
volume_test=np.load(open(test_path, "rb"))["volume"]
```

● 数据格式说明

以训练集为例，其shape=(1920*10*20*2) 代表有1920个时间段，10*20个区域，2个特征分别为区域的入流量与出流量

下载链接: <https://pan.baidu.com/s/1OB0p0fuMYFUR6zY8bs7Q3A> 密码: 1bbg

题目一：出租车流量预测

■ 评价指标

本任务以平均绝对误差(MAE)、均方根误差(RMSE) 作为评测指标

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

题目一：出租车流量预测

■ 任务要求

要求同学们以提高在测试集上的效果为目标，自己根据数据特点及需要进行数据预处理以及模型设计，本任务不对模型的设计和选择进行限制，但根据数据特点可结合CNN与RNN的特点进行模型设计。

本任务要求各位同学首先将**数据进行归一化处理并对数据进行滑动窗口划分**，具体划分要求为使用历史6个时间步预测未来一个时间步，模型的输出应该为纽约200个区域的**各区域未来半小时的出入流量**。在计算评价指标MAE和RMSE时应将数据**反归一化后再计算**。

实验报告中需要包含但不限于数据归一化的过程介绍、滑动窗口划分的过程介绍、模型图、模型中各部分的作用介绍或者使用理由、超参数设置、训练过程中各指标变化、实验结果分析等。

注：为保证公平，在训练过程中同样也不可引入测试数据，测试数据只用于最终评估。

题目二：疫情微博情绪分类

■ 任务背景

情感分析技术一直是自然语言处理领域研究的重点内容之一。2020年，新冠肺炎疫情成为了全国人民关注的焦点，众多用户针对此次疫情在新浪微博等社交媒体平台上发表自己的看法，蕴含了非常丰富的情感信息。

基于自然语言处理技术自动识别社交媒体文本中的情绪信息，可以帮助政府了解网民对各个事件的态度，及时发现人民的情绪波动，从而更有针对性地制定政策方针，具有重要的社会价值。尽管之前的社交媒体情感分析技术已经取得了不错的进展，但是如何将之前的研究成果快速高效地应用到疫情相关的数据当中，仍然是一个值得研究的问题。

在本任务当中，要求专门针对疫情相关微博进行情绪分析，来判断微博所表达的情感。

题目二：疫情微博情绪分类

■ 任务概述

微博情绪分类任务旨在识别微博中蕴含的情绪，输入是一条微博，输出是该微博所蕴含的情绪类别。在本次任务中，我们将微博按照其蕴含的情绪分为以下六个类别之一：**积极**、**愤怒**、**悲伤**、**恐惧**、**惊奇**和**无情绪**。

- 数据集的各类情绪微博举例如下表所示：

情绪	样例数据
积极(happy)	愿大家平安、健康[心]#致敬疫情前线医护人员# 愿大家都健康平安
愤怒(angry)	这次的疫情，哪家旅游app做的良心，我以后就专门用哪家了。被扣除了手续费真不爽！ ??
悲伤(sad)	救救武汉吧，受不了了泪奔，一群孩子穿上大人衣服学着救人 请官方不要瞒报谎报 耽误病情[泪][泪][泪][泪]
恐惧(fear)	对着这个症状，没病的都害怕[允悲][允悲]
惊奇(surprise)	我特别震惊就是真的很多人上了厕所是不会洗手的。。。。
无情绪(neural)	辟谣，盐水漱口没用。

题目二：疫情微博情绪分类

■ 数据集说明

● 数据集来源

本数据集（疫情微博数据集）内的微博内容是在疫情期间使用相关关键字筛选获得的疫情微博，其内容与新冠疫情相关。

链接: <https://pan.baidu.com/s/1OB0p0fuMYFUR6zY8bs7Q3A> 密码: 1bbg

● 数据集标签

每条微博被标注为以下六个类别之一：neural（无情绪）、happy（积极）、angry（愤怒）、sad（悲伤）、fear（恐惧）、surprise（惊奇）。

● 数据集规模

疫情微博训练数据集包括6,606条微博，测试数据集包含5,000条微博。

● 数据集形式

数据集为json格式，包含三个字段：数据编号（id），文本（content），情绪标签（label）。

示例：{"id": 11, "content": "武汉加油！中国加油！安徽加油！", "label": "happy"}

题目二：疫情微博情绪分类

■ 评价指标

本任务以宏精准率(macro_P)、宏召回率(macro_R)、宏F1值(macro_F1)作为评测指标

$$macro_P = \frac{1}{n} \sum_{i=1}^n P_i$$

$$macro_R = \frac{1}{n} \sum_{i=1}^n R_i$$

$$macro_F1 = \frac{2 \times macro_P \times macro_R}{macro_P + macro_R}$$

其中, $n = 6$, 为情绪类别数 $P_i = \frac{TP_i}{TP_i + FP_i}$ $R_i = \frac{TP_i}{TP_i + FN_i}$

题目二：疫情微博情绪分类

■ 任务要求

要求同学们以提高在测试集上的效果为目标，自己根据数据特点及需要进行数据预处理以及模型设计。

本任务不对模型的选择和设计进行限制，同时**要求训练所用数据不脱离给定的数据集范围，不可引入外部语料**（例如通过外部语料得到的预训练词向量等），可以以CNN、RNN等模型为基础进行设计

实验报告中需要包含但不限于数据的处理过程介绍（例如分词等）、词向量、模型图、模型中各部分的作用介绍或者使用理由、超参数设置以及训练过程中各指标变化和实验结果分析等。

提示：可使用python中的gensim库在给定语料上进行词向量训练

注：为保证公平，在训练过程中同样也不可引入测试数据，测试数据只用于最终评估

题目三：单目标跟踪

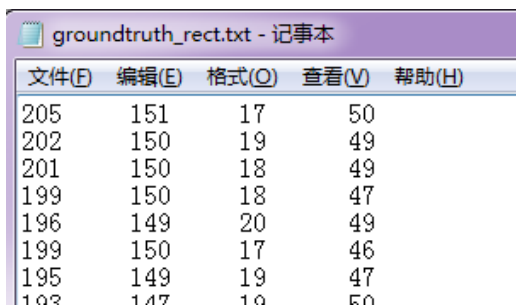
■ 任务概述

在单目标跟踪任务中，第一帧将目标物体用矩形框给定，需要跟踪算法在后续帧紧跟住这个目标物体。输入是视频的第一帧目标物体矩形框，输出是该视频后续其他帧中目标物体的矩形框。

■ 数据集说明

一共10个视频，每个视频的每一帧的标注（目标物体的矩形框位置）包含在内。能用于训练的数据只有每个视频的第一帧。其他帧的标注不能用于训练，只能用于测试目标跟踪算法。

● 数据格式说明



文件(F)	编辑(E)	格式(O)	查看(V)	帮助(H)
205	151	17	50	
202	150	19	49	
201	150	18	49	
199	150	18	47	
196	149	20	49	
199	150	17	46	
195	149	19	47	
102	147	10	50	

标注数据在每个视频文件夹的groundtruth_rect.txt文件中，文件的第一行是第一帧的标注，每一行有四个数据分别代表该矩形框的左上角点的x方向坐标、y方向坐标，矩形框的宽度w，矩形框的高度h。

下载链接: <https://cowtransfer.com/s/a5d4ba504dab47> 密码: 20200830

注: 有效期至9.6

题目三：单目标跟踪

■ 评价指标

本任务以Precision plot 、 Success plot 作为评测指标

Precision Plot: 计算跟踪算法估计的目标位置矩形框（bounding box）的中心点与人工标注（ground-truth）的目标的中心点的距离，这两者的距离小于给定阈值的视频帧的百分比。不同的阈值，得到的百分比不一样，因此可以获得一条曲线。

Success Plot: 首先定义重合率得分（overlap score, OS），跟踪算法得到的bounding box（记为a），与ground-truth给的box（记为b），重合率定义为： $OS = |a \cap b| / |a \cup b|$ ， $| \cdot |$ 表示区域的像素数目。当某一帧的OS大于设定的阈值时，则该帧被视为成功的（Success），总的成功的帧占有所有帧的百分比即为成功率（Success rate）。OS的取值范围为0~1，因此可以绘制出一条曲线。

提示：大家可以下载OTB2013benchmark工具，里面有实现好的画Precision plot和Success plot。

题目三：单目标跟踪

■ 任务要求

要求同学们结合CNN和RNN模型进行设计，以提高在测试集上的效果为目标，自己根据数据特点及需要进行数据预处理以及模型设计。

要求跟踪器的训练和更新不可使用测试数据集的未来帧（例如当前跟踪到第50帧，不可使用51帧及以后的图像数据），并且只能使用视频的第一帧标注数据。

实验报告中需要包含但不限于数据的处理过程介绍、模型图、模型中各部分的作用介绍或者使用理由、超参数设置以及训练过程中各指标变化、对比实验和实验结果分析等。

题目四：顶会论文复现

■ 任务要求

1、挑选一篇2018年以后发表在CCF推荐A类会议上的regular paper，对论文提出的模型进行复现

CCF推荐会议目录：https://www.ccf.org.cn/Academic_Evaluation/By_category/

2、针对论文中给出的baseline方法，至少选择其中一种基于深度学习的方法进行复现

3、将自己复现出的结果与论文公布的结果进行比较和分析

注意：

① 如选做本题，最好跟自己的导师商讨，选择对自己后续科研有帮助的论文来复现

② 如论文未公布源码，则注重模型的实现过程；如论文已公布源码，则注重参数调整和实验比较过程。需在实验报告中明确说明论文是否已公布源码。