

# 学期汇报

汇报人: 刘家伟    导师: 林欣



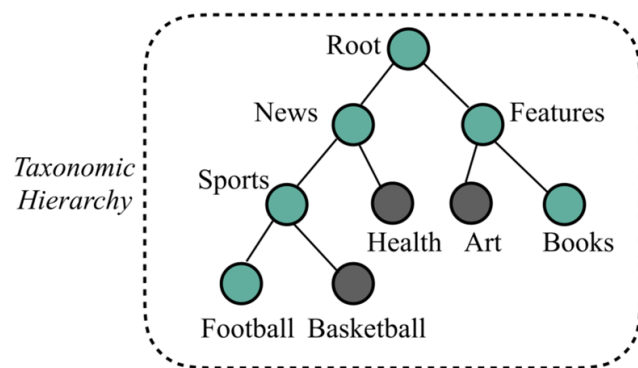


# 基础学习

- 阅读邱锡鹏《神经网络与深度学习》学习深度学习基础知识。
- 根据《Pytorch官方教程》掌握Pytorch语法，并实现文本分类，命名实体识别，机器翻译，图像生成等基础任务。
- 通过网络资源学习当前比较热门的模型架构，如Transformers, 生成对抗网络，图神经网络等。

# 科研探索

## Hierarchical Text Classification



Sentence: Durant couldn't play game because of an Achilles injury.

Labels: News、Sports、Health、Basketball

- 阅读文本分类相关论文  
包括小样本学习，半监督学习，引入图结构的分类模型等。
- 复现相关论文的代码  
通过复现论文加深对论文的理解，并掌握一些论文中没有详细说明的方法。

# 比赛实战

## SemEval 2021 Task 11: NLPContributionGraph

### 1 比赛内容

从NLP领域的学术论文中抽取关键信息构成知识图谱

### 2 方法策略

将任务拆分为**句子分类**、**命名实体识别**、**三元组拼接**三个部分

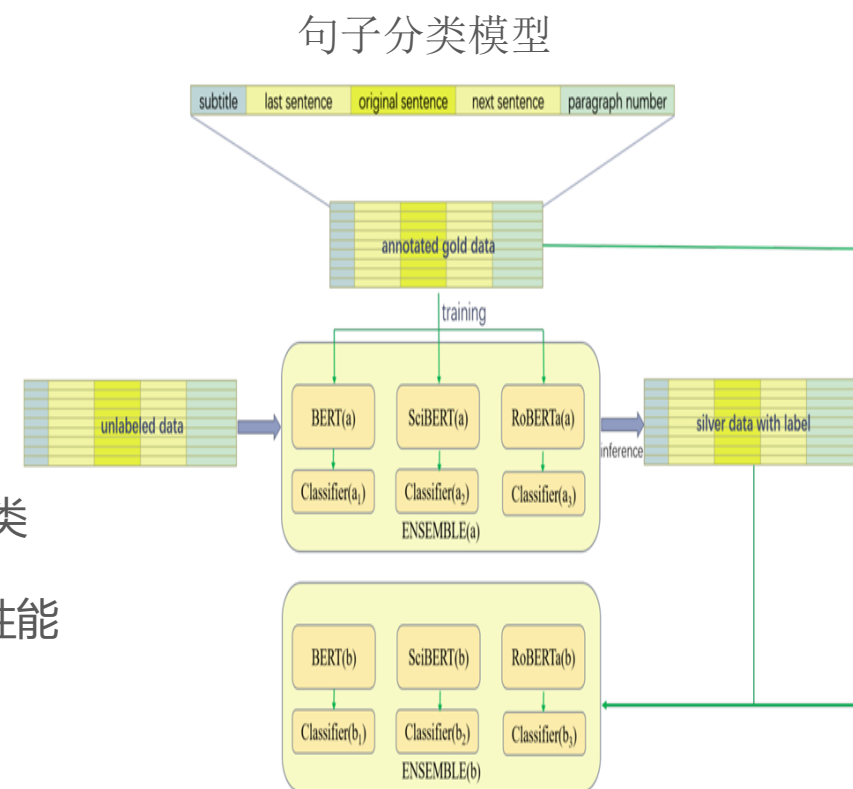
### 3 我的任务

任务：从论文中抽取具有学术贡献的句子，并将这些句子根据指定类别进行分类

方法：通过融入相关信息，数据增强，self-training, 集成学习等策略提高模型性能

### 4 比赛结果

在评测的三个子任务中分别取得 4th、2th、2th。



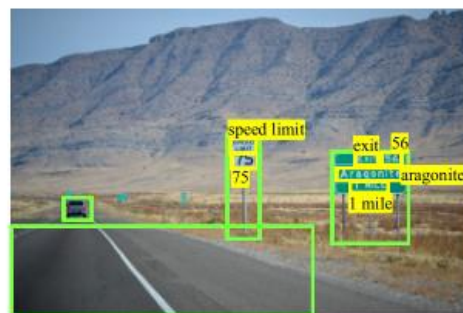
# 未来规划

- 继续跟进Hierarchical Text Classification方向，思考在小样本或注意力机制上的创新。
- 探索多模态方向，Visual Question Answering (VQA)任务,即给定一张图片和一个与该图片相关的自然语言问题，计算机能产生一个正确的回答。



Q: What is the color of the bus ?

A: yellow



question: what is the speed limit of this road ?

answer: 75 mph

detected objects: car road sign ...

OCR tokens: speed limit 75 exit ...

Thanks !

