



DeepL

订阅DeepL Pro以翻译大型文件。

欲了解更多信息，请访问www.DeepL.com/pro。

机器学习HW4

ML TAs

ntu-ml-2021spring-ta@googlegroups.com

概要

- 任务描述
- 数据集
- 数据分割
- Kaggle
- 准则

任务介绍

- 自我注意
 - 在GOOGLE的工作中提出，Attention是你需要的全部。它结合了RNN（考虑整个序列）和CNN（平行处理）的优势。
- 主要目标：学习如何使用变压器。

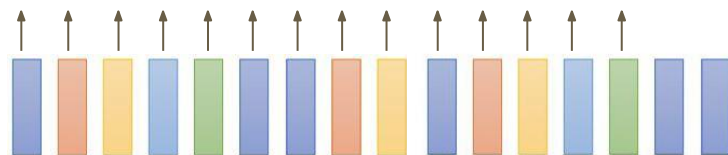
HW2:音素分类

任务。多级分类法

语音中的框架性音素预测。



m m ah ah sh sh ih ih ih n n n n ...



什么是音素？

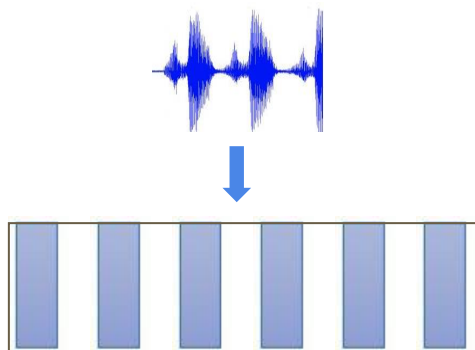
语言中的一个语音单位，可以起到区分一个词和另一个词的作用。

- 蝙蝠/拍子，坏的/床
- 机器学习 → M AH SH IH N L ER N IH NG

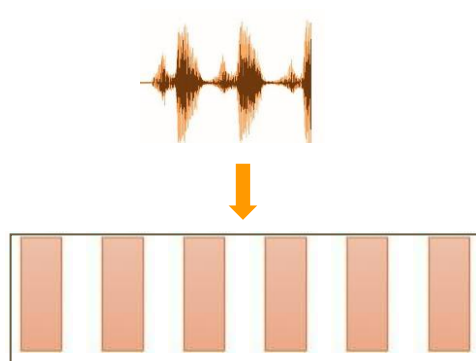
HW4: 扬声器分类

任务。多级分类法

从给定的演讲稿中预测演讲者等级。



发言人1

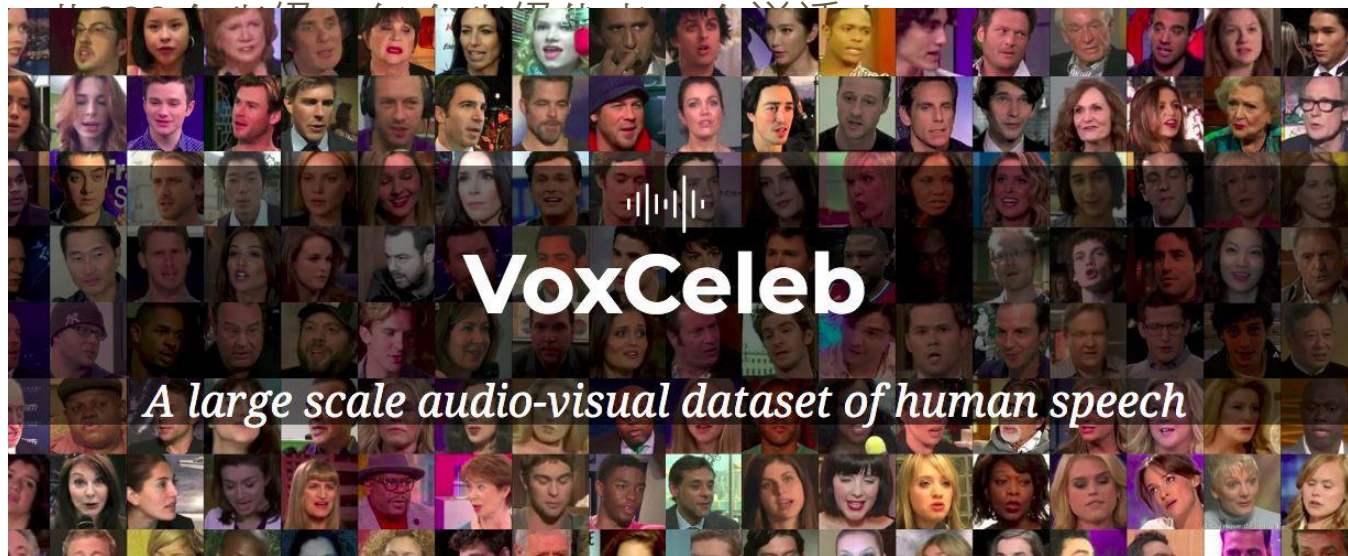


发言人2



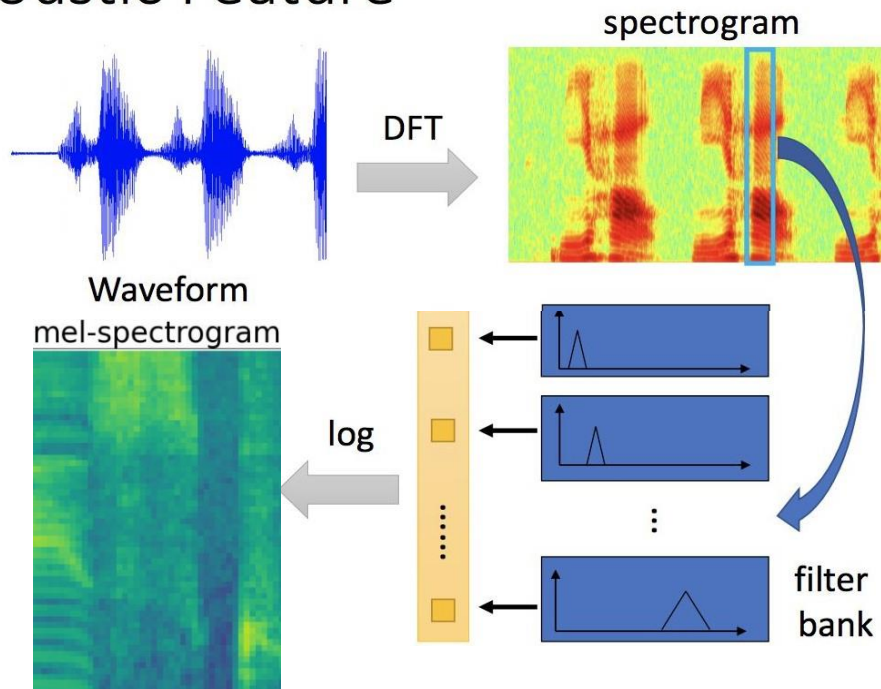
数据集

- 训练：69438个经过处理的带有标签的音频特征。
- 测试。6000个经过处理的音频特征，没有标签。
- 标签



数据预处理

Acoustic Feature



参考文献。

李鸿毅教授 [\[2020Spring](#)

[DLHLP\]](#) 语音识别

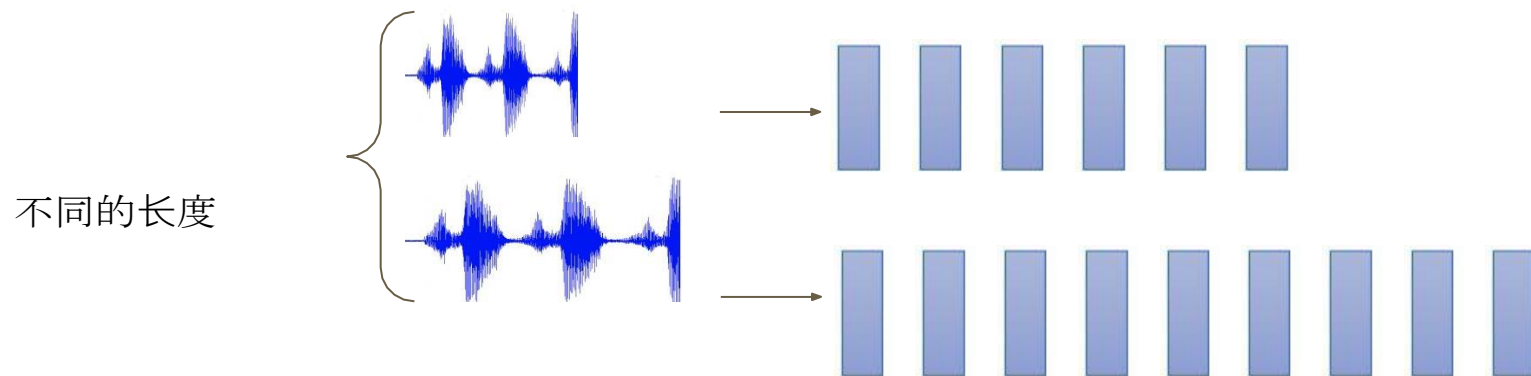
数据格式

- 数据目录
 - metadata.json
 - testdata.json
 - mapping.json
 - uttr-{random string}.pt
- 元数据中的信息
 - "n_mels"。Mel-spectrogram的尺寸。
 - "扬声器"。一本字典。
 - 关键：发言者的身份。
 - 值。"feature_path"和"mel_len"

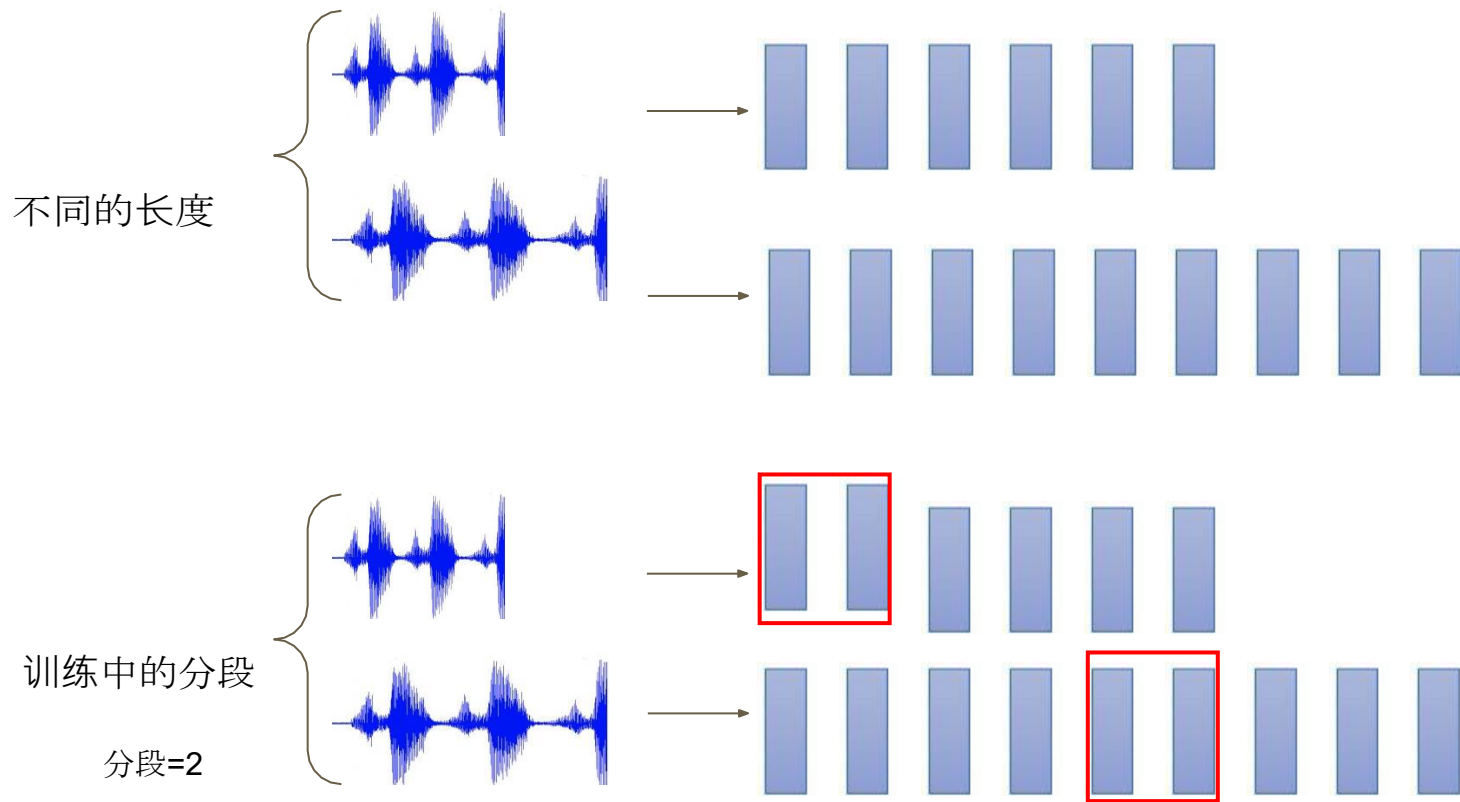
```
metadata.json
testdata.json
uttr-fff235bfc70d45b6b434c754a8136cd4.pt
uttr-fff284c8dfb94ed99010fb09208d7bcf.pt
uttr-fff286c666464b7ea2ca28811acf8f34.pt
uttr-fff3b487f8cd4905bca421b2d585bcf5.pt
uttr-fff461c64f7e4194b509b5246d2a1851.pt
```

```
{
  "n_mels": 40,
  "speakers": {
    "id10473": [
      {
        "feature_path": "uttr-5c88b2f1803449789c36f1",
        "mel_len": 652
      },
      {
        "feature_path": "uttr-022a67baccc54bfda3567a",
        "mel_len": 564
      },
      {
        "feature_path": "uttr-6a5c6e7231d642568633db",
        "mel_len": 952
      }
    ]
  }
}
```


训练期间的数据分割



训练期间的数据分割



示例代码

Colab 链接: [链接](#)

- 基线。
 - 简单。运行示例代码并知道如何使用变压器。
 - 中。知道如何调整变压器的参数。
 - 困难：构建[保形器](#)，它是变压器的一个品种。

分级

- 评估指标 = @1 准确度。 +1 pt (示例代码)
- 简单的基线 (公开) +1 pt (示例代码)
- 简单的基线 (私人) +1点
- 中等基线(公共) +1点
- 中等基线(私营) +1点
- 硬基线 (公开) +1点
- 硬基线 (私人) +4分
- 上传代码到NTU COOL

共计：**10**分

提交格式

- "Id, Category "在第一行被','分割。
- 后面是6000行由','分割的 "filename, speaker name"。

```
Id|Category
uttr-7eadda33f5fe4c9fa884c30ca0c05381.pt|id11111
uttr-7e0673bd280e4d5e8f352c8b9b5872b3.pt|id22222
uttr-9681040a85a8490cb7486f968c26131a.pt|id33333
uttr-dc680bc998a84069835e4422e3b46324.pt|id44444
uttr-3184e679b6ab43d7a4b5016ac35b38cb.pt|id55555
```

截止日期

- Kaggle: **2021/04/16 23:59 (UTC+8)**
- ntu cool: **2021/04/18 23:59 (UTC+8)**

分级 - 奖金

- 如果你得到**10**分，我们就把你的代码公开给全班同学。
- 在这种情况下，如果你也提交一份**PDF**报告，简要地描述你的方法（<100字的英文），你会得到**0.5**个百分点的奖励（你的报告也将提供给所有学生）。
- 报告模板

代码提交

- **NTU COOL (4分)**

- 将你的代码和报告压缩到

<学生ID>_hw4.zip

* 例如：

b06901020_hw4.zip

- 我们只能看到你的最后一次提交。
- 请不要提交你的模型或数据集。

- 如果你的代码不合理，你的学期成绩 $\times 0.9$ 。

代码提交

- 你的.zip文件应该只包括
 - 代码：.py或.ipynb
 - 报告：.pdf（只针对获得10分的人）
- 例子。



链接

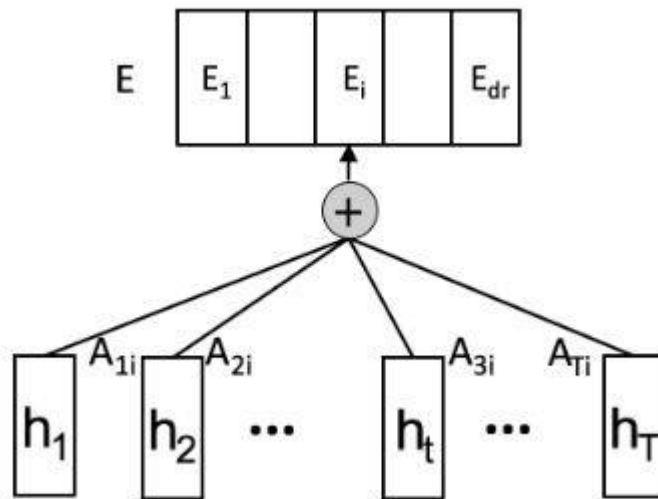
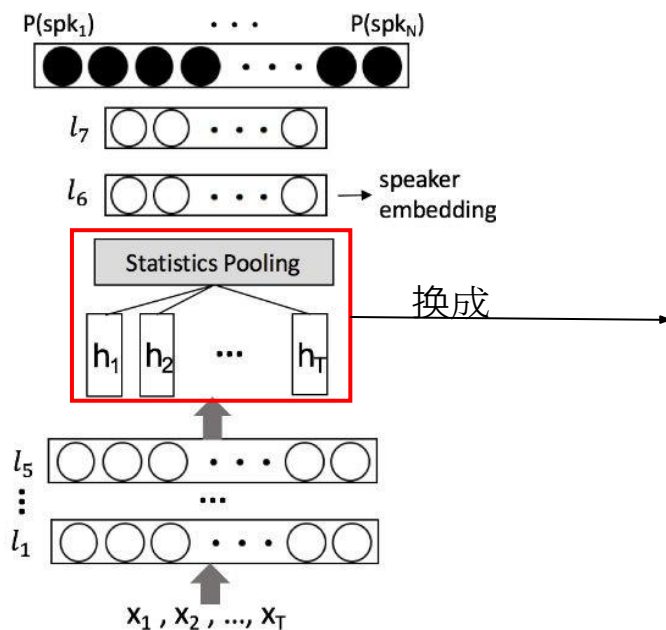
Kaggle: [link](#)

Colab: [link](#)

Data: [link](#)

提示

自觉的演讲者嵌入：[链接](#)

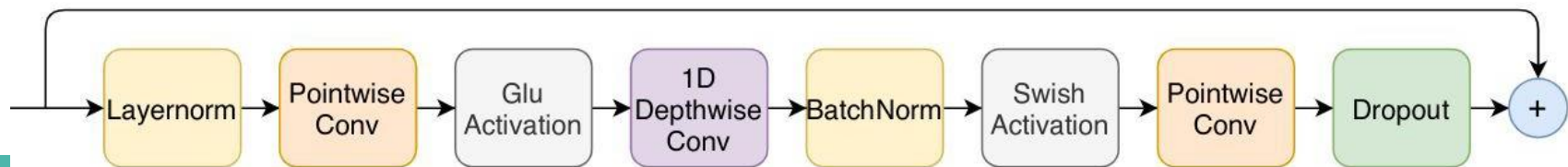
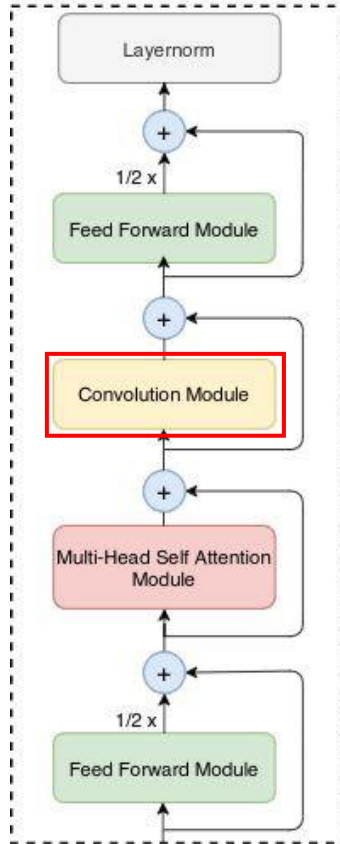


演讲者分类系统自我

注意集合

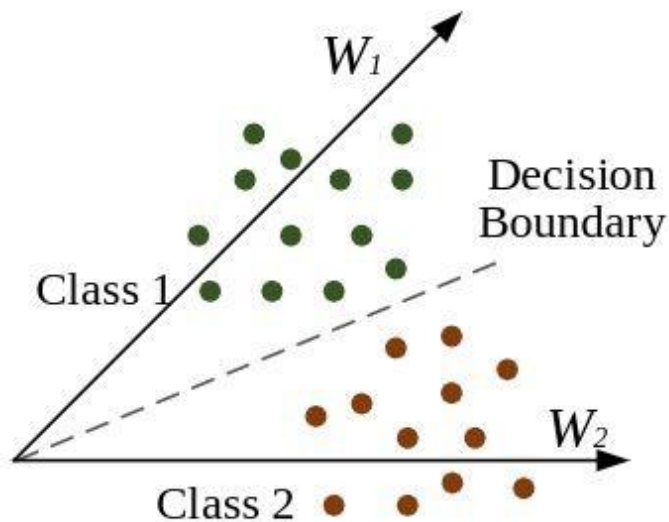
提示

顺应者：[链接](#)

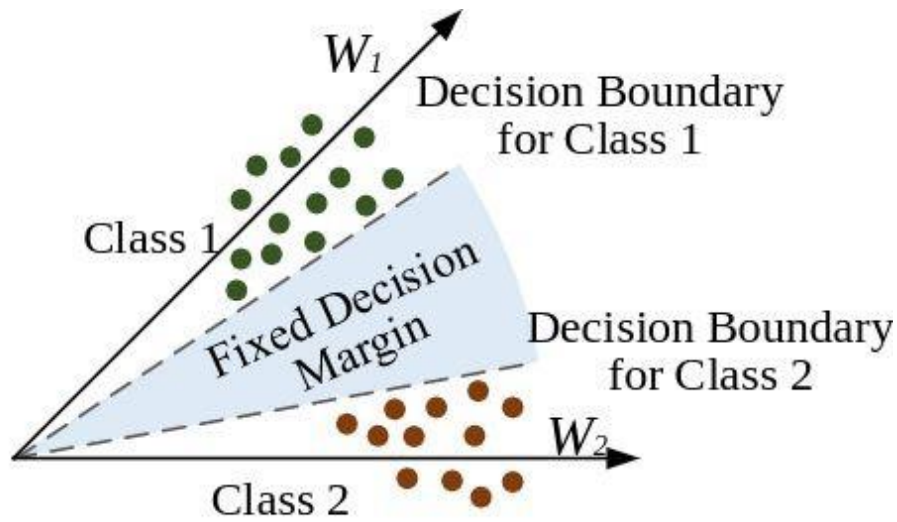


提示

加成边际软性极限: [链接](#)



Original Softmax



Additive Margin Softmax

规章制度

- 你不应该抄袭，如果你使用任何其他资源，你应该在参考文献中引用它。(*)
- 你不应该手动修改你的预测文件。
- 不要与任何生物分享代码或预测文件。
- 不要使用任何方法来提交你的结果，每天超过5次。
- 不要搜索或使用额外的数据或预训练的模型。
- 如果你违反了上述任何规则，你的**最终成绩×0.9**。
- 李教授和助教保留改变规则和成绩的权利。

(*) 科学技术部的《研究人员学术道德准则》。

如果有任何问题，你可以通过以下方式向我们提出。

- NTU COOL（推荐）。
 - <https://cool.ntu.edu.tw/courses/4793>
- 电子邮件
 - ntu-ml-2021spring-ta@googlegroups.com
 - 标题应以"[hw4]"开头。
- TA小时
 - 每周五上课时