

博学 审问 慎思 明辨 笃行

# 总结

SUN YAT-SEN UNIVERSITY



中山大學  
SUN YAT-SEN UNIVERSITY

陈和港

--2023/6/6

无标记文本

➤假设：弱监督信息 = 类的关键词

ID	Documents
$D_1$	I cheered for Lakers winning NBA.
$D_2$	I am sad that Heat lost.
$D_3$	Great news! Scientists discovered ...
$D_4$	The new film is not satisfactory.
	.....

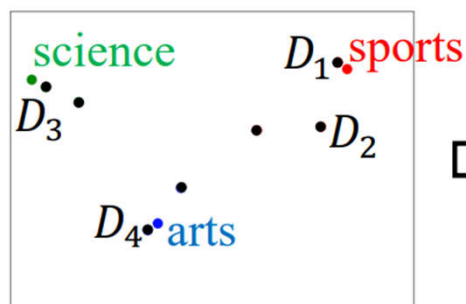
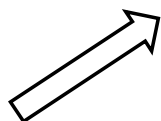
sports	NBA, basketball, ...
arts	film, music, ...
science	technology, computer, ...

类的关键词

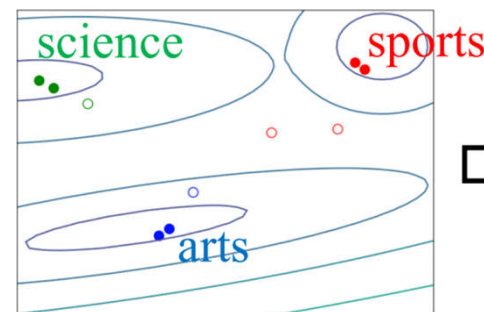
文本嵌入



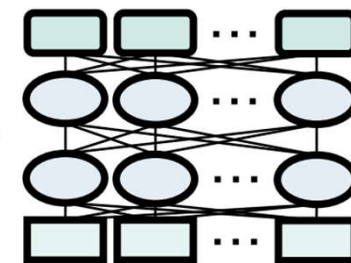
类嵌入



估计分类信心



自训练和预测



## LIME: Weakly-Supervised Text Classification Without Seeds

动机:

- 1) 先前基于关键词的伪标签生成方法依赖于样本是否存在代表关键词。此外受限于种子词数，或词嵌入质量等，模型鲁棒性不够高
- 2) 忽略了文本的上下文信息

# 研究背景

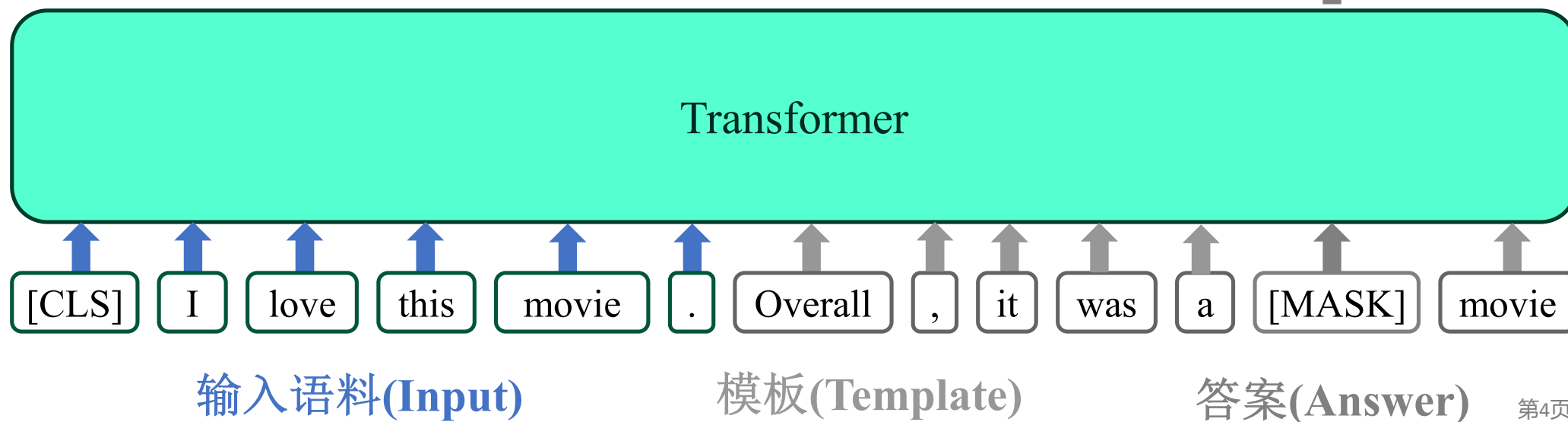
## Prompt Learning 提示学习

➤ 目标：拉近预训练与目标任务的距离

对应标签：

*positive*

完型填空：good / fantastic / interesting



➤形式化:



提示(Prompt):



Dataset	Type	# of Classes	Dataset size	Prompt
20News	News topic	5	17,871	<i>The text is about &lt;class label&gt;.</i>
AGNews	News topic	4	120,000	<i>The text is about &lt;class label&gt;.</i>
Yelp	Restarant review	2	38,000	<i>It was good. / It was bad.</i>
DBpedia	Wikipedia topic	14	560,000	<i>The text is about &lt;class label&gt;.</i>

Table 2: Statistics for benchmark datasets.

“We discard pseudo-labels with confidence under 50%.”

Model	20News	AGNews	Yelp	DBpedia
Supervised	96.45 / 96.42	93.99 / 93.99	95.70 / 95.70	98.96 / 98.96
Entailment classifier	67.95 / 67.50	79.94 / 79.99	94.79 / 94.79	80.14 / 79.27
WeSTClass	71.28 / 69.90	82.30 / 82.10	81.60 / 81.6	81.42 / 81.19
ConWea	75.73 / 73.26	74.60 / 74.20	71.40 / 71.20	N/A
LOTClass	73.78 / 72.53	86.89 / 86.82	87.75 / 87.68	86.66 / 85.98
X-Class	78.62 / 77.76	85.74 / 85.66	90.00 / 90.00	91.32 / 91.17
LIME	<b>79.74 / 79.56</b>	<b>87.21 / 87.16</b>	<b>95.22 / 95.22</b>	<b>92.19 / 92.20</b>

Table 3: Experiment results on 4 classification benchmarks. All reported scores in the form *micro-F1 / macro-F1*. Baselines are quoted from (Wang et al., 2021).

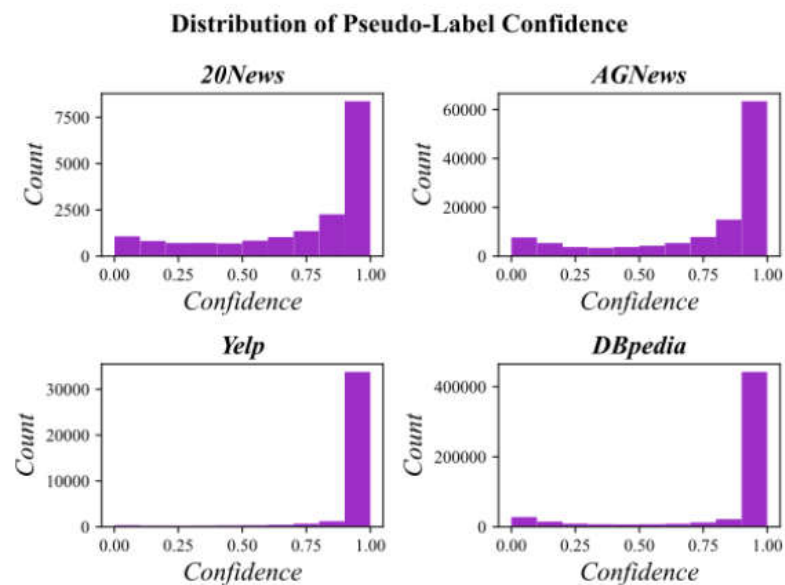


Figure 1: Histogram of pseudo-label confidence. More confident pseudo-labels result in more accurate classification self-training.

**Average F1 Score vs. Pseudo-Label Confidence Thre**

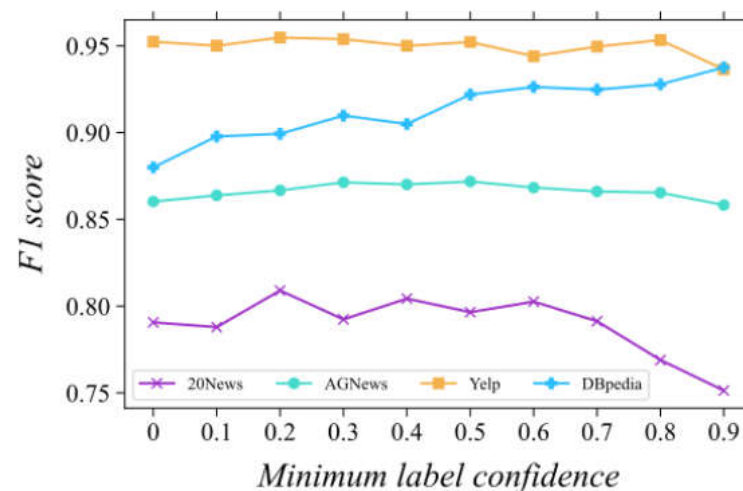


Figure 2: Effect of varying confidence thresholds on self-training F1 scores.



# 期末项目：弱监督文本分类

## ➤ 基本要求：设计并实现一个弱监督文本分类模型

- 报告模型在两个数据集上的性能（须使用同一套参数）
- 使用至少两个评价方法
- 推荐采用自训练算法
- 实现基于聚类策略的同学需要展示聚类图

## ➤ 进阶要求：未知类型识别

- dataset文件（包括dataset.zip和dataset\_unk.zip）中的标签仅用于测试，不能用于模型训练。dataset\_unk.zip用于【进阶要求：未知类型识别】，其中的数据集各自删去了1-2个类的名称及其关键词（keywords.txt和classes.txt）。

课堂展示的同学，准备8分钟左右的PPT演示

6月13号

杨若凡  
叶彦烈  
凌国明  
陈乐豫  
郑北辰  
张雄  
郑圳毅  
郭晓炀

6月20号

陶俊成  
陈俊儒  
温子健  
谢文龙  
黄兰舒  
叶抒锐  
郑梓霖  
梁爽