

玄学: Prompt-base Learning

# 自然语言处理的四个范式

范式的概念：

<https://en.wikipedia.org/wiki/Paradigm>

# Paradigm


🌐 61 languages ▾

[Article](#) [Talk](#)

[Read](#) [Edit](#) [View history](#)

From Wikipedia, the free encyclopedia

*For other uses, see [Paradigm \(disambiguation\)](#).*



This article's [lead section](#) may be too short to adequately [summarize](#) the key [points](#). Please consider expanding the lead to [provide an accessible overview](#) of all important aspects of the article. *(July 2020)*

In [science](#) and [philosophy](#), a **paradigm** (/ˈpærədaɪm/) is a distinct set of concepts or thought patterns, including theories, [research methods](#), postulates, and standards for what constitute legitimate contributions to a field. The word *paradigm* is Greek in origin, meaning "pattern", and is used to illustrate similar occurrences.

范式可以理解为“模式”。

模式就是一种规律，是存在于人们感知到的世界、人造设计或抽象思想中的规律。

# 什么是自然语言？

指一种自然地随文化演化的语言。

中文、英语、法语、西班牙语等都是自然语言。

方言也是自然语言。

世界语是人工语言。

# 处理文本或者语音序列

文本朗读  
语音识别  
实体识别  
文本分类

聊天机器人  
机器翻译  
文本纠错

.....

# 自然语言处理的四个范式

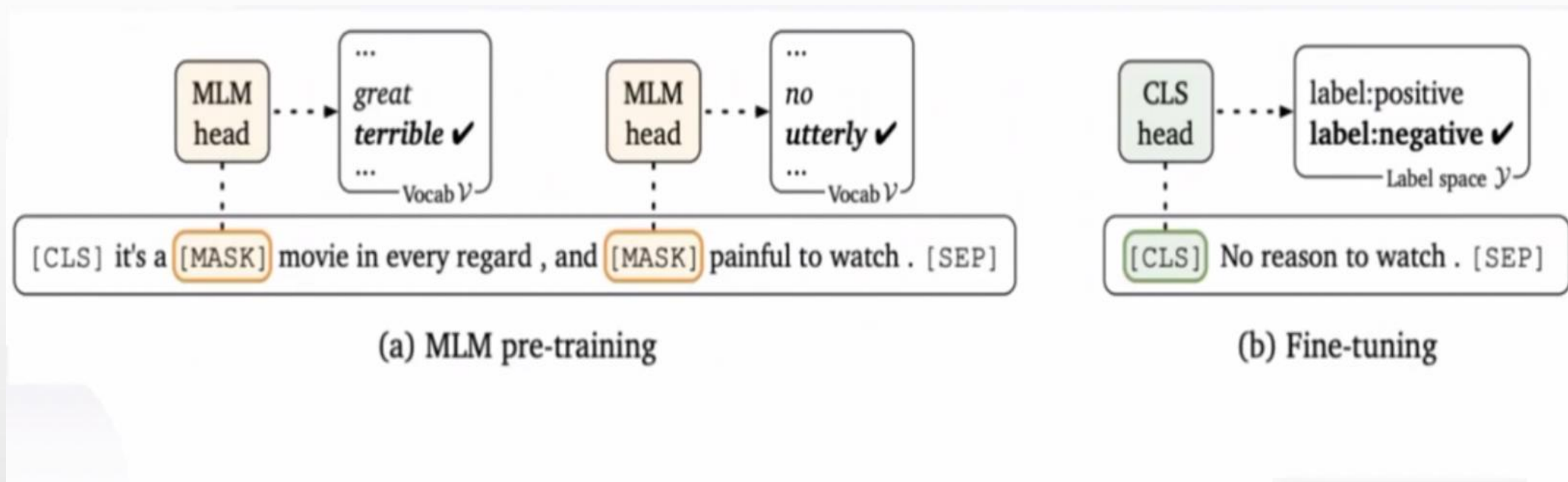
第一范式：基于传统机器学习模型的范式，如 **tf-idf** 特征+朴素贝叶斯等机器算法，需要做大量的特征工程工作；

第二范式：基于深度学习模型的范式，如 **word2vec** 特征 + **LSTM** 等深度学习算法，相比于第一范式，模型准确度有所提高，特征工程的工作也有所减少；

第三范式：基于预训练模型 + **finetuning**的范式，如 **BERT + finetuning** 的NLP任务，相比于第二范式，模型准确度显著提高，但是模型也随之变得更大

第四范式：基于预训练模型 + **Prompt** + 预测的范式，如 **BERT + Prompt** 的范式相比于第三范式，模型训练所需的训练数据显著减少。

# PTM+ Finetune 方式的缺陷



# PTM+ Finetune 方式的缺陷

[illegible]



# 完形填空

1. My name is Li Hua. I' m thirteen. I' m middle school student. I am\_1\_Class Five, Grade One. My English teacher \_2\_ Mr. Lin. He is \_3\_ old teacher.

2. 天气\_4\_, 今天风和日丽。

( ) 1. A. in          B.at          C.do          D.not

( ) 2. A. are          B.am          C.is          D.省略

( ) 3. A. a          B.an          C.this          D.very

( ) 4. A. 好          B.差          C.不好不坏          D.强

# Prompt Learning的基本构成:

- 模板构造    Prompt
- 预训练模型 PTM
- 标签映射    Verbalizer

原始输入：今天风和日丽。

Prompt 输入：天气[MASK]，今天风和日丽。

输出：{ “好， 晴， 说” }



<https://huggingface.co/bert-base-chinese>

# prompt 设计

天气[MASK], 今天风和日丽。-好  
今天风和日丽。天气[MASK] -好

# verbalizer: vocabulary 和 label 的 多对一映射

{ “不错”，“好”，“晴”，“强” } --> { “好” }

{ “糟糕”，“不好”，“差”，“弱” } --> { “差” }

verbalizer就是表意，本质是将词映射到类别。人工构建的verbalizer不太好，现在有算法自动构建。

# 实验结果： 50条数据训练

- 天气[MASK], 日和风暖的一天
- 好
- 天气[MASK], 爱就像蓝天白云, 晴空万里
- 好
- 天气[MASK], 细雨斜风敲打着我心房
- 差
- 天气[MASK], 好热好热好热好热
- 差
- 天气[MASK], 今天日丽风清
- 好
- 天气[MASK], 天寒地冻的世界, 动物必须冬眠
- 差

- 天气[MASK], 外面碧空万里, 出去玩吧
- 好
- 天气[MASK], 乍暖乍寒, 真搞不懂这天气
- 差
- 天气[MASK], 今天也是风和日丽的一天捏
- 好
- 天气[MASK], 出去一会儿感觉寒风侵肌
- 差
- 天气[MASK], 实在是太热了, 出去一会儿汗如雨下
- 差
- 天气[MASK], 刚打开水龙头就滴水成冰, 太冷了
- 差
- 天气[MASK], 皓月千里
- 好

# 为什么说它是玄学？

科学证明：

- 模板不同，效果不同
- 标签不同，效果不同
- 标签位置不同，效果也不同

科学也无法证明：

- 它为啥会这样？

- 固定表意，轻微地调整模板的位置，得到的准确率不同。

Pattern	verbalization	Accuracy
$x_1?$ [MASK], $x_2$	Yes/Maybe/No	<b>77.2</b>
$x_1$ . [MASK], $x_2$	Yes/Maybe/No	76.2
$x_1?$ [MASK] $x_2$	Yes/Maybe/No	74.9
$x_1$ $x_2$ [MASK]	Yes/Maybe/No	65.8
$x_2?$ [MASK], $x_1$	Yes/Maybe/No	62.9
$x_1?$ [MASK], $x_2$	Maybe/No/Yes	60.6
Fine-tuning	-	48.4



- 固定 prompt，轻微地调整表意的位置，得到的准确率不同。

Pattern	verbalization	Accuracy
x It was [MASK].	great/terrible	<b>92.7</b>
x It was [MASK].	good/bad	92.5
x It was [MASK].	cat/dog	91.5
x It was [MASK].	dog/cat	86.2
x It was [MASK].	terrible/great	83.2
Fine-tuning	-	81.4



Prompt Engineer  
is not a real job?

# 探讨:

1. 标签是不同长度的字，如何构建模板和标签映射。
  - 假设分类有“金融”，“大健康”。需要用multi-mask吗？  
用[MASK]代替“金融”还是[MASK][MASK]代替“金融”？  
参考：<https://discuss.huggingface.co/t/multiple-mask-tokens/174/2>
2. 有人用prompt-based learning做过其他任务吗？可以交流一下。
3. 下游统一任务如何做？

参考:

Pattern-Exploiting Training (PET) <https://github.com/timoschick/pet>