

12. 自然语言处理II

自然语言推断

WU Xiaokun 吴晓堃

xkun.wu [at] gmail

2022/02/28

自然语言推断与数据集

果壳中的宇宙

Q: What is the universe?

A: Not Nut, but very close.

– Running like Clockwork

- 为什么“very close”? – *Universe in a Nutshell*

自然语言推断

研究假设 (hypothesis) 是否可以从前提 (premise) 推断出来

- 决定一对文本序列间的逻辑关系

自然语言推断

研究假设 (hypothesis) 是否可以从前提 (premise) 推断出来

- 决定一对文本序列间的逻辑关系

蕴涵: “A是好学生” \implies “A学业优秀”

自然语言推断

研究假设 (hypothesis) 是否可以从前提 (premise) 推断出来

- 决定一对文本序列间的逻辑关系

蕴涵: “A是好学生” \implies “A学业优秀”

矛盾: “A是好学生”, “A经常不及格”

自然语言推断

研究假设 (hypothesis) 是否可以从前提 (premise) 推断出来

- 决定一对文本序列间的逻辑关系

蕴涵: “A是好学生” \implies “A学业优秀”

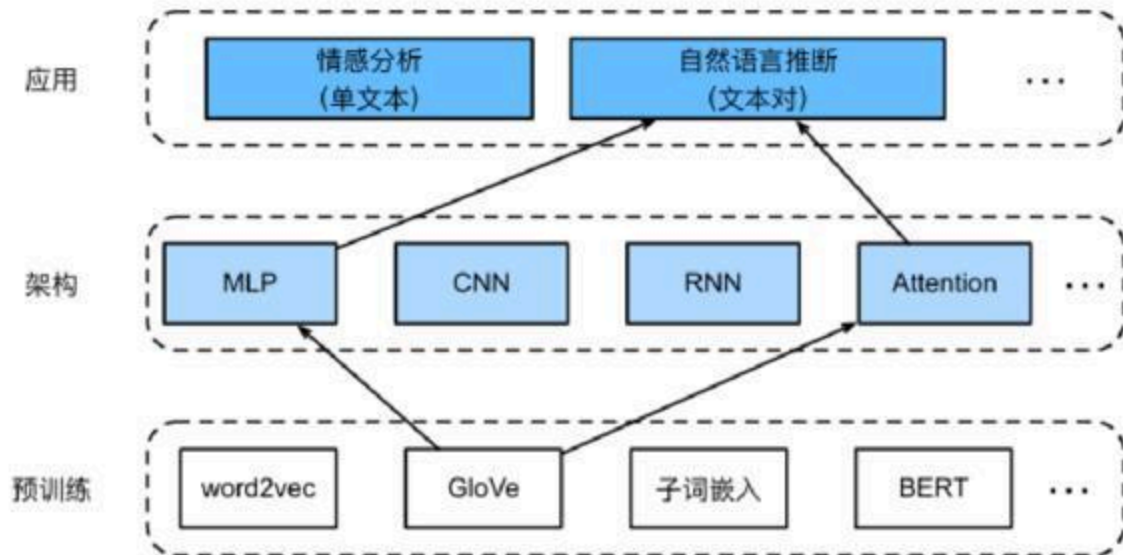
矛盾: “A是好学生”, “A经常不及格”

中性: “A是好学生”, “A爱唱歌”

自然语言推断：使用注意力

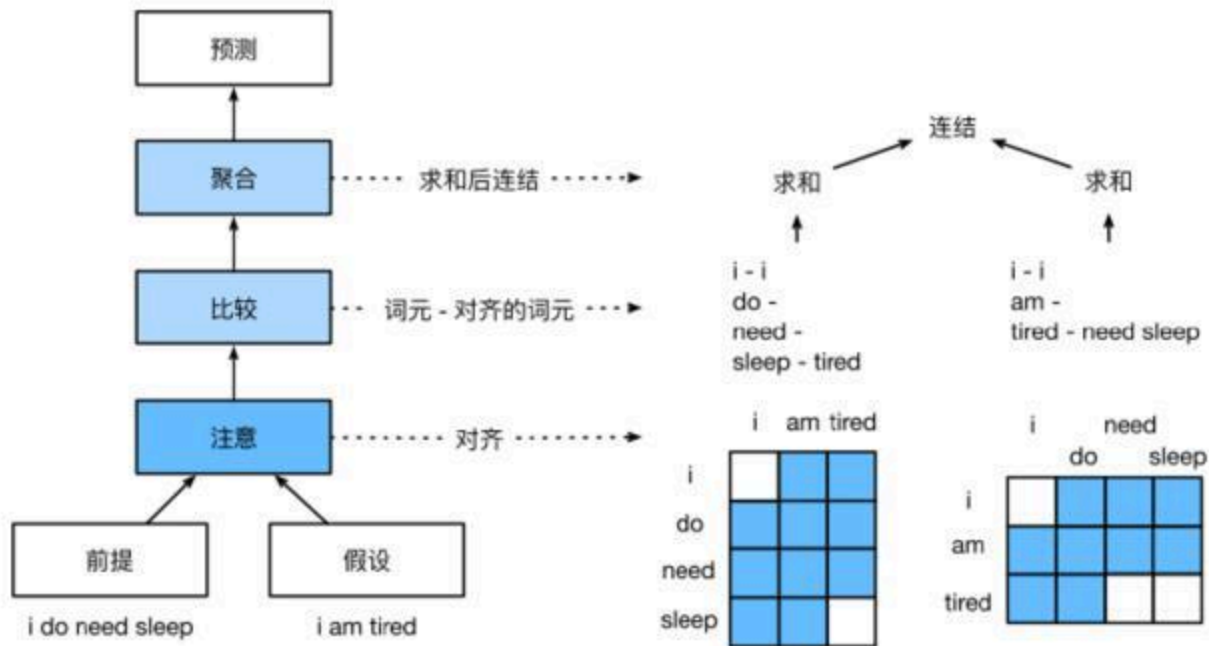
可分解注意力模型

[Parikh 2016] “可分解注意力模型”：没有循环/卷积层



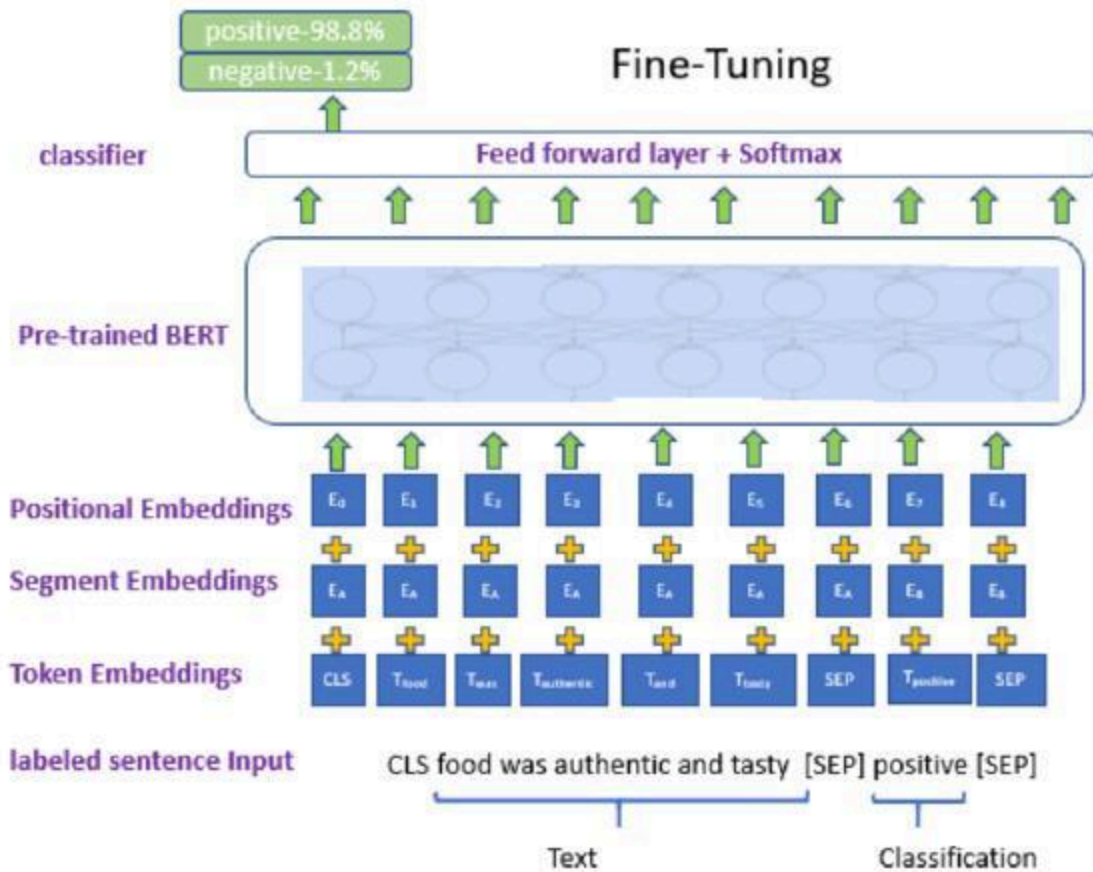
模型

将两个序列按词元对齐：比较、聚合信息用于预测



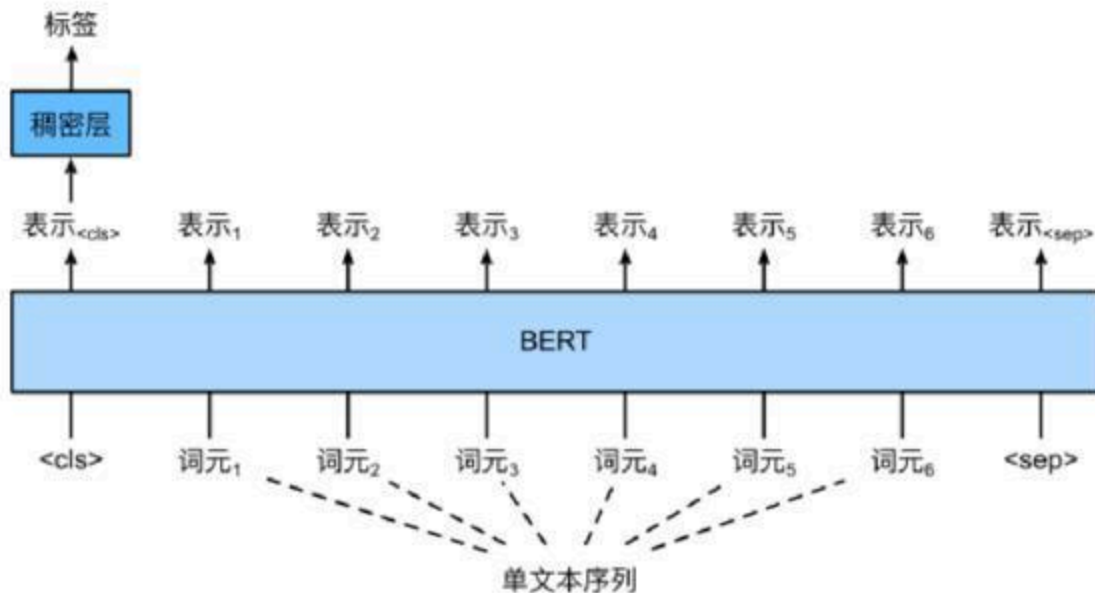
针对序列/词元级应用微调BERT

BERT: 微调架构图



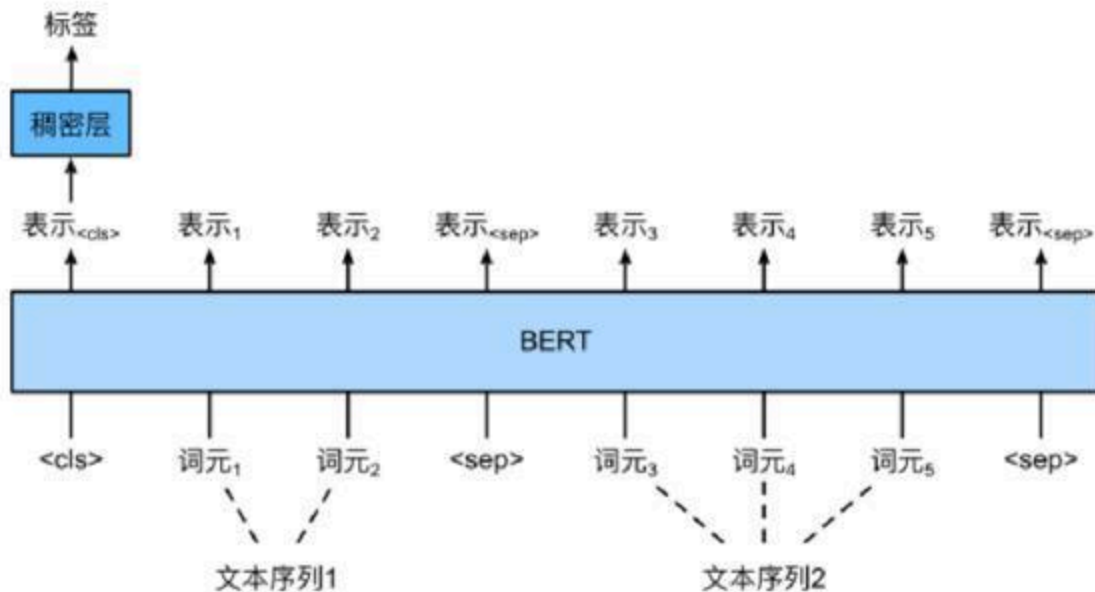
单文本分类

输出：单/多个类别概率



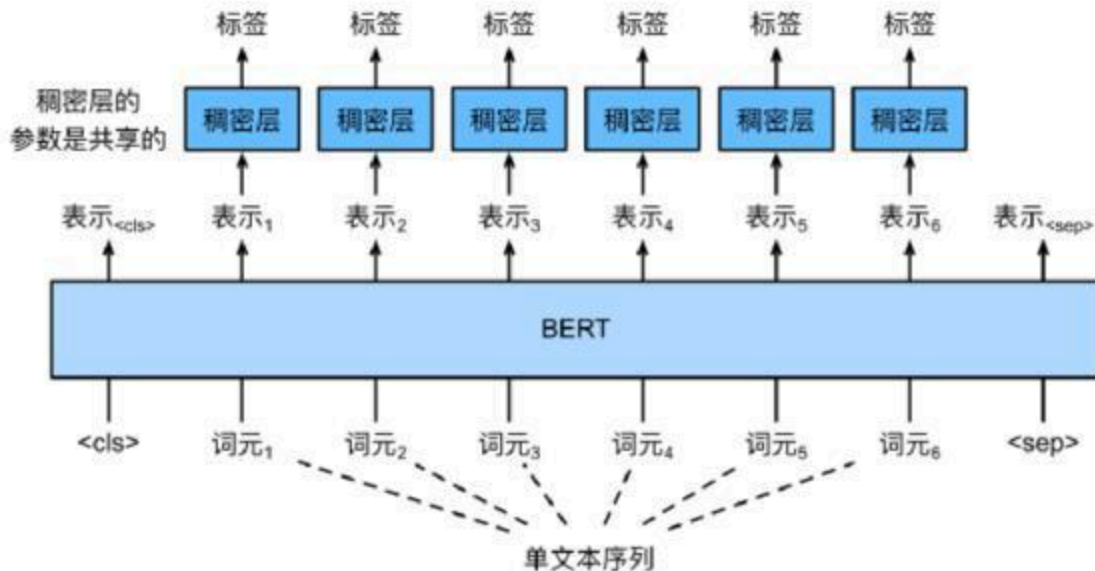
文本对分类或回归

输出：相似度



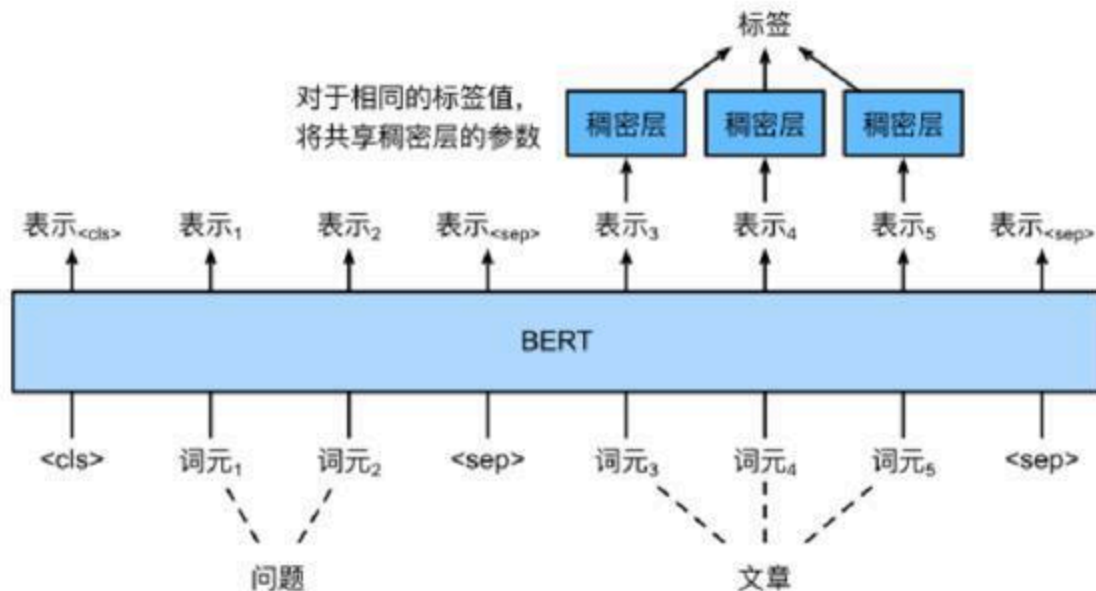
文本标注

输出：每个词元的标注概率



问答

输出：段落起止位置



“Some experts report that a mask’s efficacy is inconclusive. However, mask makers insist that their products, such as N95

respirator masks, can guard against the virus."

Q: "Who say that N95 respirator masks can guard against the virus?" -- A: mask makers

自然语言推断：微调BERT

微调BERT

微调BERT：只需一个额外的基于多层感知机的架构

