

周报，2021年1月25日

屈原斌
首都师范大学
ybqu@cnu.edu.cn

1 上周计划

1. [***] 可视化结果观察
2. [***] 数据处理实验更新

2 上周计划执行情况

1. [✓]
2. [✓×] xLNet分类模型未完成

3 本周部分重点工作详述

3.1 可视化结果观察

- 结果见图1、2、3
- 分析结论：

？模型可能认为首句没有标点的句子最重要

```
=== 句子 attn
> 0.183015 记忆中的温暖，记忆中的温暖。
> 0.090344 我在自己的房间里说了好几次，我一起温暖就又幸福又温暖的但我又怕消失在我的眼前，我自己下了一个目眩，珍稀好每一
次，对妈妈好，在我有一次，我给妈妈洗脚时，我就很感动，很温暖。
> 0.008803 我就对妈妈说“妈妈在我小时候，你为我付出了许多——我长大，一定会还您。
> 0.027525 在我的记忆中永远存在一种东西——温暖。
> 0.010159 温暖可以打败挫折和困难。
> 0.011144 温暖在我有困难的时候它可以发挥力量。
> 0.015316 温暖可以战胜一切“
> 0.036825 温暖是有一种伟大的爱而形成的。
> 0.070738 那就是妈妈给我们的爱给了我们力量的温暖在心中是很有热量的，在我的记忆中温暖最伟大了，最宽宏了，最广阔亮了。
> 0.040418 在我记忆中的温暖是很热很热，很伟大而近思的，温暖给了我们力量和自强不息的爱，有妈妈温暖着我，我即使没有了热量，
但也是有热量的“妈妈温暖着我，爱抚着我，关爱着我，给了我的一切的，所有的爱来关爱着我，我只好像是一朵白云，妈妈就是太阳，每天的
照亮着我，我很温暖，我很幸福。
> 0.132267 人们都说，只要有了爱和温暖就等于有了——一切“就在心中非常的温暖，温暖就可以记在了心中，所以我们要珍稀好每一分
每一秒，叫温暖永远的存在在我的记忆中。
> 0.183603 记忆中的温暖 ← 题目
> 0.006942 在我中以长大的心，回忆起我就非常感动，回忆起了我小时候的顽皮。
> 0.002383 妈妈就哄着我，在我小时候，妈妈就没离开过我，还是每天的哄我、逗我玩——小时候的时光总是那么的快，妈妈也一天一天
的在变老，现在我长大了，我现要做的就是回报妈妈对我的一切。
> 0.001972 在我小时候到现在妈妈每天关心我的我现在十四岁了，妈妈照看了我十四年了，想一想，妈妈得多么累啊，我现在也不让妈妈
受苦受累，有妈妈“我永远是温暖的。
> 0.000171 时光过的真快呀！
> 0.020835 我心中有着回报妈妈的心”妈妈有着给我温暖的心。
> 0.002723 妈妈给我留下了无数的印象，妈妈对我的温暖是用什么也消失不了的。
> 0.044392 在我记忆中充满了温暖与幸福，我相信，妈妈给我的温暖永远的存在在我的心中。
> 0.110426 也在我心中记忆中永远的在这里。
```

Figure 1: 题目与中间句互换

```

> 0.321902 记忆中的温暖。
> 0.072950 在我的记忆中永远存在一种东西——温暖。
> 0.011596 温暖可以打败挫折和困难。
> 0.031564 温暖在我有困难的时候它可以发挥力量。
> 0.016795 温暖可以战胜一切“
> 0.022729 温暖是有一种伟大的爱而形成的。
> 0.029878 那就是妈妈给我们的爱给了我们力量的温暖在心中是很有热量的，在我的记忆中温暖最伟大了，最宽宏了，最辽阔了。
> 0.181804 在我记忆中的温暖是很热很热，很伟大而近思的，温暖给了我们力量和自强不息的爱，有妈妈温暖着我，我即使没有了热
但也是有热量的“妈妈温暖着我，爱抚着我，关爱着我，给了我一切的，所有的爱来关爱着我，我只好像是一朵白云，妈妈就是太阳，每
照亮着我，我很温暖，我很幸福。
> 0.014765 人们都说，只要有了爱和温暖就等于有了——一切“就在心中非常的温暖，温暖就可以记在了心中，所以我们要珍惜好每
每一秒，叫温暖永远的存活在我的记忆中。
> 0.210826 记忆中的温暖，记忆中的温暖。
> 0.005059 我在自己的房间里说了好几次，我一起温暖就又幸福又温暖的但我又怕消失在我的眼前，我自己下了一个目眩，珍稀好
次，对妈妈好，在我有一次，我给妈妈洗脚时，我就很感动，很温暖。
> 0.000113 我就对妈妈说“妈妈在我小时候，你为我付出了许多——我长大，一定会还您。
> 0.004427 在我中以长大的心，回忆起我就非常感动，回忆起了我小时候的顽皮。
> 0.000122 妈妈就哄着我，在我小时候，妈妈就没离开过我，还是每天的哄我、逗我玩。小时候的时光总是那么的快，妈妈也一天
的在变老，现在我长大了，我现要做的就是回报妈妈对我的一切。
> 0.007222 在我小时候到现在妈妈每天关心我的我现在十四岁了，妈妈照看了我十四年了，想一想，妈妈得多么累啊，我现在也不让
受苦受累，有妈妈“我永远温暖的。
> 0.000102 时光过的真快呀！
> 0.008716 我心中有着回报妈妈的心”妈妈有着给我温暖的心。
> 0.017376 妈妈给我留下了无数的印象，妈妈对我的温暖是用什么也消失不了的。
> 0.030368 在我记忆中充满了温暖与幸福，我相信，妈妈给我的温暖永远的存活在我的心中。
> 0.011686 也在我心中记忆中永远的在这里。

```

Figure 2: 题目添加句号

```

=== 用于attn
> 0.527675 记忆中的温暖，记忆中的温暖
> 0.064634 我在自己的房间里说了好几次，我一起温暖就又幸福又温暖的但我又怕消失在我的眼前，我自己下了一个目眩，珍稀好
次，对妈妈好，在我有一次，我给妈妈洗脚时，我就很感动，很温暖。
> 0.004943 我就对妈妈说“妈妈在我小时候，你为我付出了许多——我长大，一定会还您。
> 0.019251 在我的记忆中永远存在一种东西——温暖。
> 0.005851 温暖可以打败挫折和困难。
> 0.006904 温暖在我有困难的时候它可以发挥力量。
> 0.009518 温暖可以战胜一切“
> 0.021522 温暖是有一种伟大的爱而形成的。
> 0.043495 那就是妈妈给我们的爱给了我们力量的温暖在心中是很有热量的，在我的记忆中温暖最伟大了，最宽宏了，最辽阔了
> 0.020203 在我记忆中的温暖是很热很热，很伟大而近思的，温暖给了我们力量和自强不息的爱，有妈妈温暖着我，我即使没有了
但也是有热量的“妈妈温暖着我，爱抚着我，关爱着我，给了我一切的，所有的爱来关爱着我，我只好像是一朵白云，妈妈就是太阳，
照亮着我，我很温暖，我很幸福。
> 0.089736 人们都说，只要有了爱和温暖就等于有了——一切“就在心中非常的温暖，温暖就可以记在了心中，所以我们要珍惜好
每一秒，叫温暖永远的存活在我的记忆中。
> 0.049543 记忆中的温暖。
> 0.004687 在我中以长大的心，回忆起我就非常感动，回忆起了我小时候的顽皮。
> 0.001731 妈妈就哄着我，在我小时候，妈妈就没离开过我，还是每天的哄我、逗我玩。小时候的时光总是那么的快，妈妈也一
的在变老，现在我长大了，我现要做的就是回报妈妈对我的一切。
> 0.001426 在我小时候到现在妈妈每天关心我的我现在十四岁了，妈妈照看了我十四年了，想一想，妈妈得多么累啊，我现在也不
受苦受累，有妈妈“我永远温暖的。
> 0.000127 时光过的真快呀！
> 0.014287 我心中有着回报妈妈的心”妈妈有着给我温暖的心。
> 0.002132 妈妈给我留下了无数的印象，妈妈对我的温暖是用什么也消失不了的。
> 0.031186 在我记忆中充满了温暖与幸福，我相信，妈妈给我的温暖永远的存活在我的心中。
> 0.081149 也在我心中记忆中永远的在这里。

```

Figure 3: 题目与中间句互换且去除句号

3.2 数据处理实验更新

- 分类实验结果见表1
- 离题实验更新：
 - 方案：取题目下全部作文表示平均作为题目表示，针对每个题目调节阈值
 - 实验结果见表2
 - 实验结论：
 - * 去除题目后指标下降

| | Accuracy | Precision | Recall | F1-score |
|---------------------|----------|-----------|--------|----------|
| HABiLstm_W | 0.7379 | 0.698 | 0.6344 | 0.653 |
| HABiLstm_W(加入测试集训练) | 0.7208 | 0.6831 | 0.6306 | 0.6467 |
| BERT | 0.6261 | 0.6524 | 0.517 | 0.5532 |
| xLNet | - | - | - | - |

Table 1: 分类模型指标

| | | 离题 | | | 不离题 | | |
|---------------------|-------|-----------|--------|----------|-----------|--------|----------|
| | | Precision | Recall | F1-score | Precision | Recall | F1-score |
| HABiLstm_W | 未处理数据 | 0.1044 | 0.2198 | 0.1416 | 0.9772 | 0.9467 | 0.9617 |
| | 处理后数据 | 0.1012 | 0.0603 | 0.0756 | 0.9759 | 0.9861 | 0.981 |
| HABiLstm_W(加入测试集训练) | 未处理数据 | 0.1888 | 0.2043 | 0.1963 | 0.9774 | 0.9752 | 0.9763 |
| | 处理后数据 | 0.1324 | 0.0681 | 0.0899 | 0.9762 | 0.9884 | 0.9823 |
| BERT | 未处理数据 | 0.1486 | 0.1991 | 0.1702 | 0.9771 | 0.9677 | 0.9724 |
| | 处理后数据 | 0.0865 | 0.0486 | 0.0623 | 0.9756 | 0.9867 | 0.9811 |

Table 2: 离题指标更新

4 下周计划

1. [***] 完成分类模型
2. [***] 补充对比实验
3. [***] 分析实验结果