

作业一：基于HMM和BPE的分词

▶ 作业描述

- ▶ 基于附件提供的Jupyter Notebook文件，以代码填空的形式，实现HMM中文分词和BPE英文分词。填空完毕后，需展示代码中相应测试部分的输出结果。

▶ 成绩占比

- ▶ 9% + 9%

▶ 评分标准

- ▶ 代码与输出结果，同时参考代码注释和实验总结性文字

▶ 注意事项

- ▶ 鼓励学习并使用辅助编程工具（如Copilot和ChatGPT），但是需要在总结中注明自己使用的工具
- ▶ 鼓励与其他同学合作，但是一定要尊重彼此知识产权，不能抄袭
- ▶ 实验结果严禁造假

作业一：基于HMM和BPE的分词

▶ 截止时间

- ▶ 2023年11月19日23:59

▶ 宽限时间说明

- ▶ 本学期所有作业共享7天的宽限时间，即**所有作业的迟交时间之和不能超过7天**。如：第一次小作业迟交了3天，那么第二次小作业便至多迟交4天；第二次小作业再次迟交4天后，后面的作业便不再享有宽限时间。所有宽限时间用完后再迟交者会惩罚一定分数，具体的惩罚方案最终会由老师与助教讨论后确定并公布。同时，即使宽限时间未用完，最终的提交时间也不能晚于本学期最后的提交期限。

▶ QA文档

- ▶ <https://docs.qq.com/sheet/DRENMVWdFc3ZZaWRU?tab=BB08J2>