

自然语言处理实验 2——哈工大语言技术平台（LTP）使用

【相关理论知识提要】

LTP，全称哈工大语言技术平台（访问地址：<http://ltp.ai/index.html>）是国内最优秀的中文自然语言处理开源综合平台。LTP 提供了一系列中文自然语言处理工具，用户可以使用这些工具对于中文文本进行分词、词性标注、命名实体标注、完全句法分析、依存句法分析等工作。

【实验目的】

1. 复习自然语言处理关键底层技术的基本概念、知识与模型；
2. 熟悉哈工大语言技术平台（LTP）提供的技术资源及文档说明；
3. 掌握哈工大语言技术平台（LTP）本地安装及常用功能调用方法；

【实验内容】

1. 在线文档学习

任务一、认真阅读在线文档的“快速上手”部分

- 熟悉 LTP 安装方法
- 熟悉 LTP 核心模块的调用方法

任务二、认真阅读在线文档的“附录”部分

- 熟悉 LTP 所使用的词性标注集、命名实体标注集、依存句法关系集、语义角色标注集
- 了解 LTP 所使用的语义依存关系集

2. 使用 LTP 的在线演示平台

任务一、使用平台上默认的样例数据进行各项自然语言处理分析

- 认真阅读分析结果，并查看不同视图下的分析结果

任务二、自设一整段文本，进行各项自然语言处理分析

- 认真阅读分析结果，并查看不同视图下的分析结果

3. 本地安装后调用各项功能

任务一、按照在线文档的“快速上手”部分进行复现

- 完成 LTP 安装
- 完成 LTP 各核心模块的调用

任务二、编写一个完整程序，实现文件读入、使用 LTP 进行各项自然语言处理，并将结果写入到文件中（将本任务的完整代码与结果文件打包，发到我的作业邮箱 txu_homework@163.com 中，邮件名称和压缩包名称用“2023 自然语言处理实验 2 + 本人姓名”方式命名）

- 自己补充少量代码，形成一个完整程序，实现对实验数据集的完整自然语言处理，并将结果自动另存成一份文件