

自然语言处理

2022年秋季

黄河燕、鉴萍

北京理工大学 计算机学院

hhy3, pjian@bit.edu.cn

实践作业一

15分

黄河燕、鉴萍

北京理工大学 计算机学院

hhy63, pjian@bit.edu.cn

提交内容、方式和时间

□ 提交内容

- 实验报告(Word或PDF文档)+源码
 - 报告内容包括：
 1. 采用模型或方法的原理
 2. 使用的数据、平台或工具
 3. 实验过程(训练和测试)、实现过程描述
 4. 实验结果
 5. 感想，如果有的话
-

提交内容、方式和时间

□ 提交方式和deadline

- 压缩包提交到乐学，命名：实践1_学号_姓名.xxx
- Deadline：10月3日

实践内容

□ 分别采用规则方法、概率统计方法、神经网络方法实现汉语句子的情感极性判断

□ 提示：

- 基于一部情感词典，实现一个基于词典或规则的情感分析系统；
- 基于标注了情感极性的语料，实现一个基于Naïve Bayes的情感分析系统；
- 基于标注了情感极性的语料，实现一个基于逻辑回归的情感分析系统，以汉语词向量为输入。

实践内容

□ 注：

- 直接调用汉语分词工具，比如jieba
- 自己实现Naïve Bayes
- 情感词典、情感标注语料老师提供
- 自己选择深度学习框架，建议PyTorch
- 老师可以提供训练好的汉语词向量
- 尽量自己实现逻辑回归，如果实在有困难，可以调用