

自然语言处理第二次实验

实验环境

python3.5 + jieba + pytorch1.0.0 + numpy

实验目标

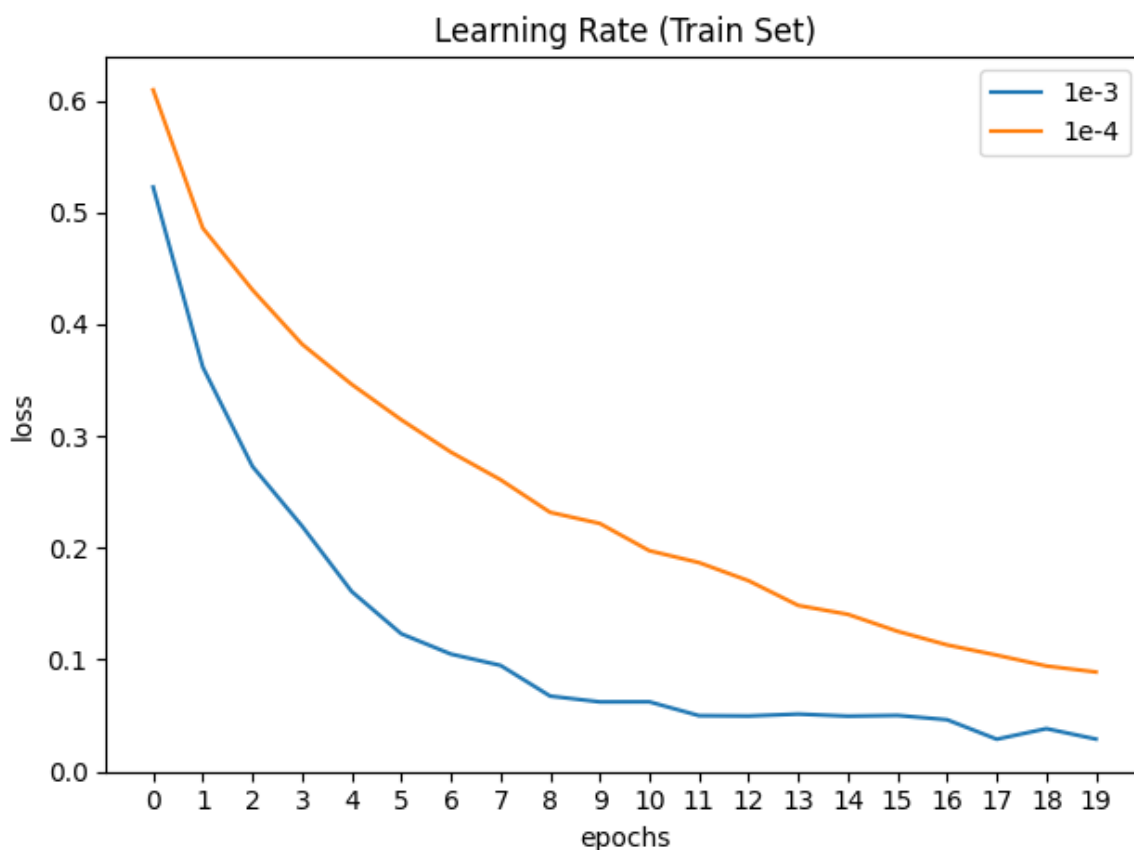
- 了解神经网络中的基础模型。
- 了解深度学习框架PyTorch的使用。
- 了解使用深度学习解决文本情感分类任务基本流程。

实验内容

1. 阅读代码，运行并得到结果。
2. 根据已有的 TextCNN 模型格式，写出 BiLSTM 模型。
3. 调整 batch size 参数，取值分别为 [8, 16, 32, 64]，画出 TextCNN 和 BiLSTM 训练集和验证集的 loss 折线图。
4. 调整 embedding size 参数，取值分别为 [64, 128, 256]，画出 TextCNN 和 BiLSTM 训练集和验证集的 loss 折线图。
5. 调整 hidden size 参数，取值分别为 [64, 128, 256]，画出 TextCNN 和 BiLSTM 训练集和验证集的 loss 折线图。
6. 调整 learning rate 参数，取值分别为 [1e-2, 5e-3, 1e-3, 5e-4, 1e-4]，画出 TextCNN 和 BiLSTM 训练集和验证集的 loss 折线图。
7. 调整 dropout 参数，取值分别为 [0.1, 0.2, 0.3, 0.4, 0.5]，画出 TextCNN 和 BiLSTM 训练集和验证集的 loss 折线图。
8. 选出最好的参数组合（不局限于以上参数组合），列出参数并将实验结果写入表格进行对比分析，例如：

模型	准确率
TextCNN	86.80%
BiLSTM	87.68%

loss图示例：



附加实验

1. 使用 word2Vec 算法对实验数据中的训练集进行预训练，并使用预训练的词向量进行 TextCNN 和 BiLSTM 的情感分类实验。
2. 使用 [开源的词向量](#) 进行实验，来提升模型效果。
3. 使用 [注意力机制](#) 来提升 BiLSTM 模型的效果。

提交时间

- 基础实验：11月3号截止
- 附加实验：第18周截止

各班班长或学习委员收集班内所有同学的实验报告和实验代码后发送到 1327793532@qq.com 邮箱中。文件命名方式：姓名-学号-第X次实验

从10月19号开始所有完成的附加实验的人，需要写一个新的实验报告或文档，将文件和代码打包发送到 1327793532@qq.com 邮箱中，文件命名方式为 姓名-学号-第X次实验附加实验。

实验要求

- 完成所有实验内容
- 良好的代码风格
- 完整的实验报告

参考资料

1. [pytorch1.0.0官方文档](#)
2. [《神经网络与深度学习》](#)
3. [Convolutional Neural Networks for Sentence Classification](#)

