

浙江工业大学



文本分析与挖掘

上机实验

计算机科学与技术学院

实验七、基于循环神经网络的情感分类

一、实验目的

1. 掌握基于进行分类的模型搭建和训练
2. 掌握不同 RNN 类型的调用
3. 熟悉英文情感分类过程。

二、实验内容

基于循环神经网络对英文影评 IMDB 数据集进行分类。可以直接从 keras 导入预处理好的数据。

1. 单层 GRU

- a) 模型 1: 单层 GRU (节点个数 128, 激活 tanh), 全连接 (1 个节点, 激活 sigmoid)。
- b) 设置 batchsize=32, 训练模型 1, 得到测试集上的准确。

2. 单层 LSTM

模型 2: 把模型 1 中的 GRU 改成 LSTM, 重复 1 (b) 实验, 对比结果。

3. 双层 GRU

模型 3: 双层 GRU (每层节点个数 128, 激活 tanh), 全连接 (1 个节点, 激活 sigmoid)。重复 1 (b) 实验, 对比结果。

4. 分别设置 batchsize=16 和 64, 重复 1 (b) 实验, 对比结果。

5. 模型 4: embedding 层 (128 个节点), 单层 GRU (节点个数 128, 激活 tanh), 全连接 (1 个节点, 激活 sigmoid)。重复 1 (b) 实验, 对比结果, 讨论降维的作用。

6. 尝试其他调参: 学习率、优化器等, 观察并讨论结果。

7. 尝试用预训练词嵌入作为模型 1 的输入, 对比模型 4 的结果。