

# 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

# 基于深度神经网络的情感分类

## 一、实验目的

1. 掌握基于循环神经网络和 TextCNN 的调用
2. 熟悉英文情感分类过程。

## 二、实验内容

基于循环神经网络和 TextCNN 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。

## 三、实验步骤

(1) 根据以下设置不同模型以及超参数, 打印模型摘要; (2) 用训练集对模型进行训练, 打印并观察模型的学习曲线, 讨论训练情况, 并就观察到的情况对训练做可能的相应调整; (3) 训练结束后输出测试集上的准确率, 并自己写一条评论进行测试, 评价所得结果。(4) 改变 batchsize、学习率、优化器等训练相关参数。(5) 对不同模型、不同超参数设置的结果进行对比和讨论。

## 四、模型设置

### 1. 循环神经网络

- 1) 单层 GRU: 单层 GRU (节点个数 128, 激活 tanh), 全连接 (1 个节点, 激活 sigmoid)。设置 batchsize=32。
- 2) 双层 GRU: 双层 GRU (每层节点个数 128, 激活 tanh), 全连接 (1 个节点, 激活 sigmoid)。
- 3) 把 GRU 改成 LSTM。

### 2. TextCNN

1) embedding 层输出维度 embed\_dim=200; 卷积层设置三种卷积核大小 kernel\_size=[2, 3, 4], 每种大小用 64 个卷积核, 即 filters=64; 倒数第二的全连接 dense 层输出 64 维, Relu 激活; 最后一个输出二分类结果, 1 个节点, sigmoid 激活。

- 2) 自己尝试改变以上模型设置。