



基于多模态数据的情感分析技术研究

姓 名： 邱浙宇

学 号： 1023040814

专 业： 计算机科学与技术

目录

01 技术背景

02 模型框架

03 模型训练

04 功能演示与测试

05 总结与展望



技术背景

第一部分



情感分析介绍

情感分析技术指的是为给定的信息判定情感，如二分类则判定为积极或者消极。通过情感分析，可以分析各类数据隐藏的信息，从而达到更深层次地理解与分析。

传统的情感分析技术通常只使用文本数据进行分析，这可能会忽略其他可用的重要信息来源。





多模态情感分析



提出一种基于多模态数据的情感分析模型



模型架构

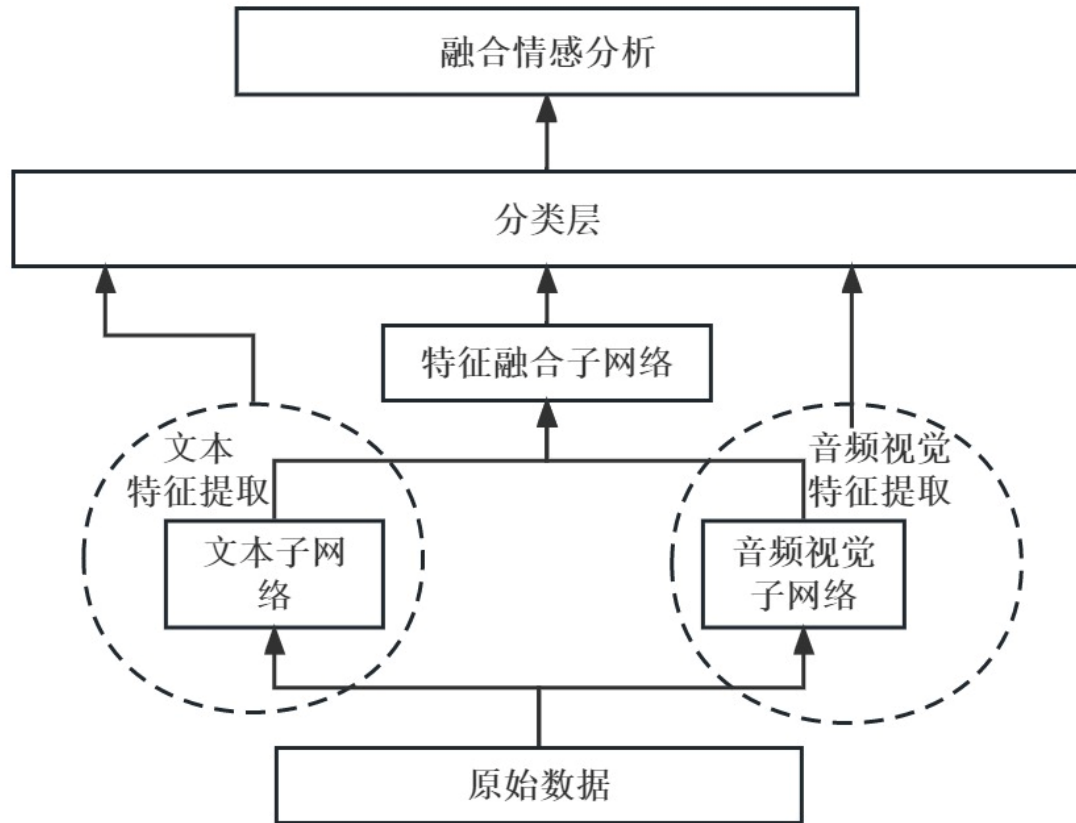
第二部分



模型整体架构

● 组成部分

- 文本子网络
- 音频视觉子网络
- 特征融合子网络
- 情感分类层





原始数据

- MOSI

MOSI数据集是一个英文多模态语料库，它包含了2199段意见视频片段。



- SIMS

SIMS数据集是一个中文多模态情感分析数据集，它包含了2,281个精炼的视频片段






知识图谱嵌入BERT

- WN11

WN11是从WordNet抽取的子集，共有38696个实体，11种关系。

- 嵌入方法

将WN11知识图谱当作原始数据，对BERT模型进行重新训练。



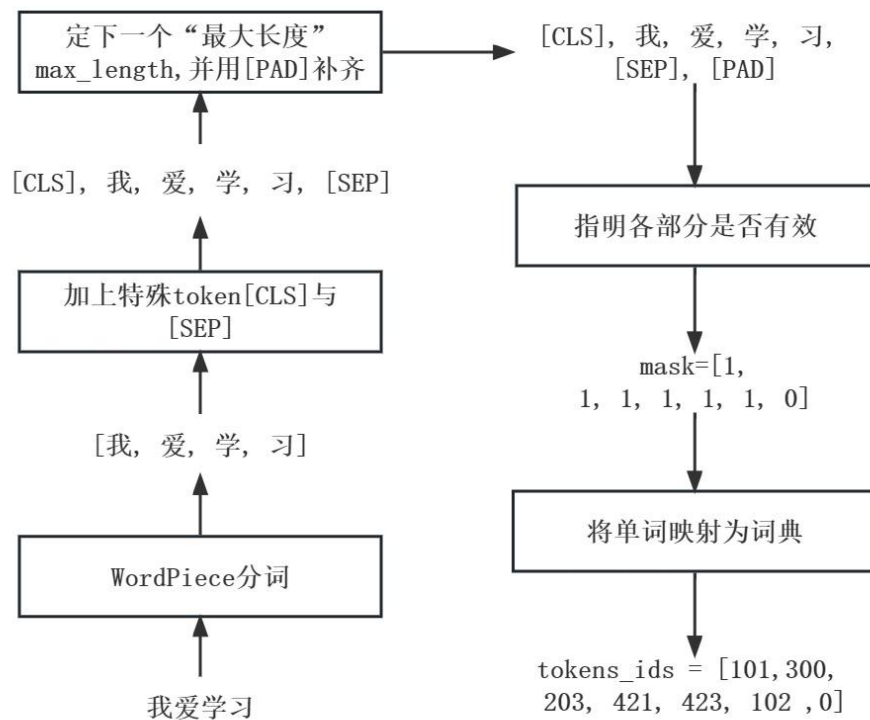
- config.json
- eval_results.txt
- pytorch_model.bin
- test_results.txt
- vocab.txt



使用知识增强的BERT模型进行数据处理

分词步骤

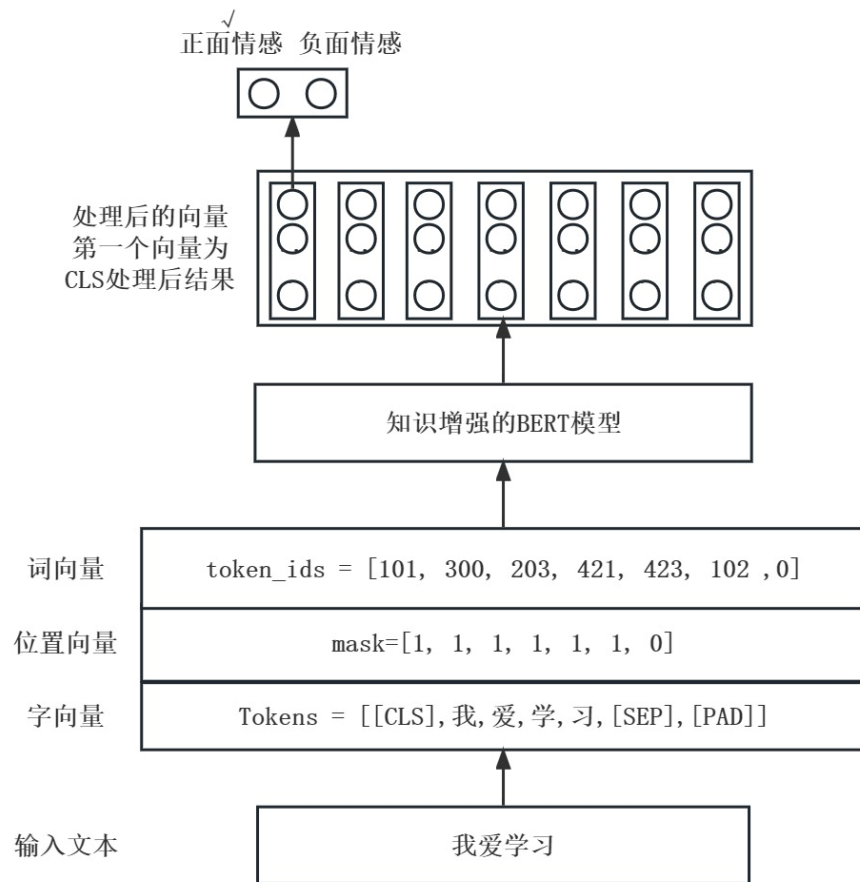
- 1、分词, Tokens = [我, 爱, 学, 习]
- 2、加上[CLS]与[SEP]
- 3、补全到最大长度, [PAD]填充
- 4、标注[PAD]位置
- 5、单词映射为数字





使用BERT进行处理

- 字向量: $[[CLS], \text{我}, \text{爱}, \text{学}, \text{习}, [SEP], [PAD]]$
- 位置向量: $[1, 1, 1, 1, 1, 1, 0]$
- 词向量: $[101, 300, 203, 421, 423, 102, 0]$





音频处理

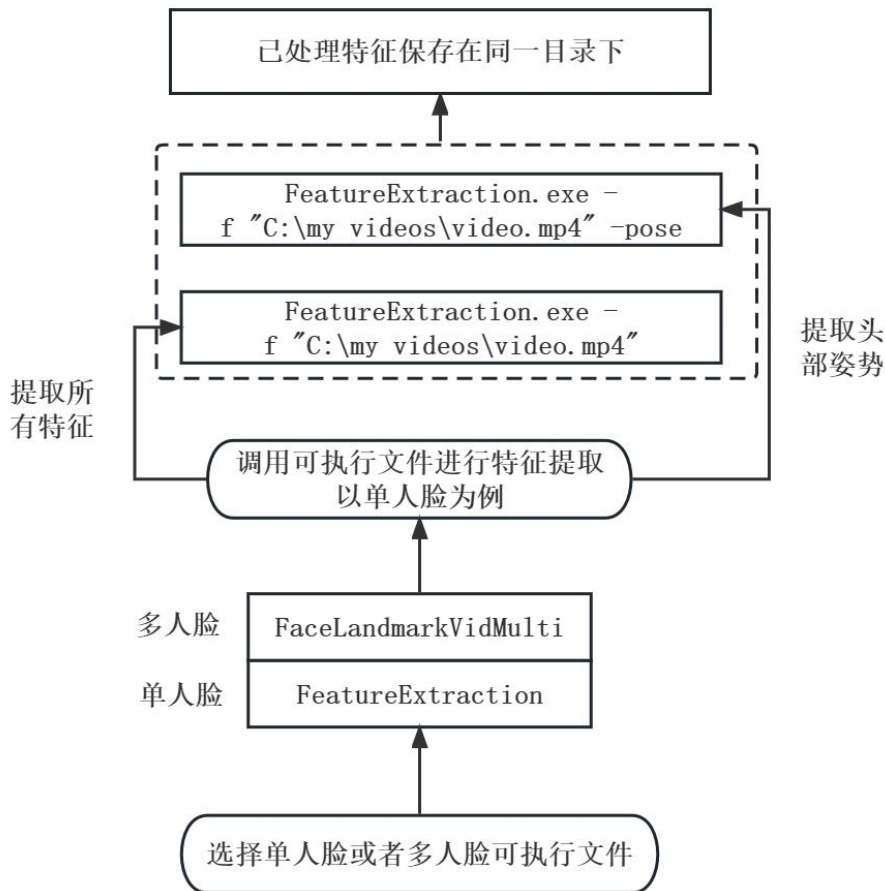
- OpenSmile (ffmpeg)
 - 1、设置路径
 - 2、设置配置文件，选择ComParE_2026(标准特征集数量最多)
 - 3、配置相关命令进行原始音频数据的处理



视频处理

- OpenFace

OpenFace 是一个能够进行面部标志检测、头部姿势估计、面部动作单元识别和眼睛注视估计的工具包





音视频特征处理

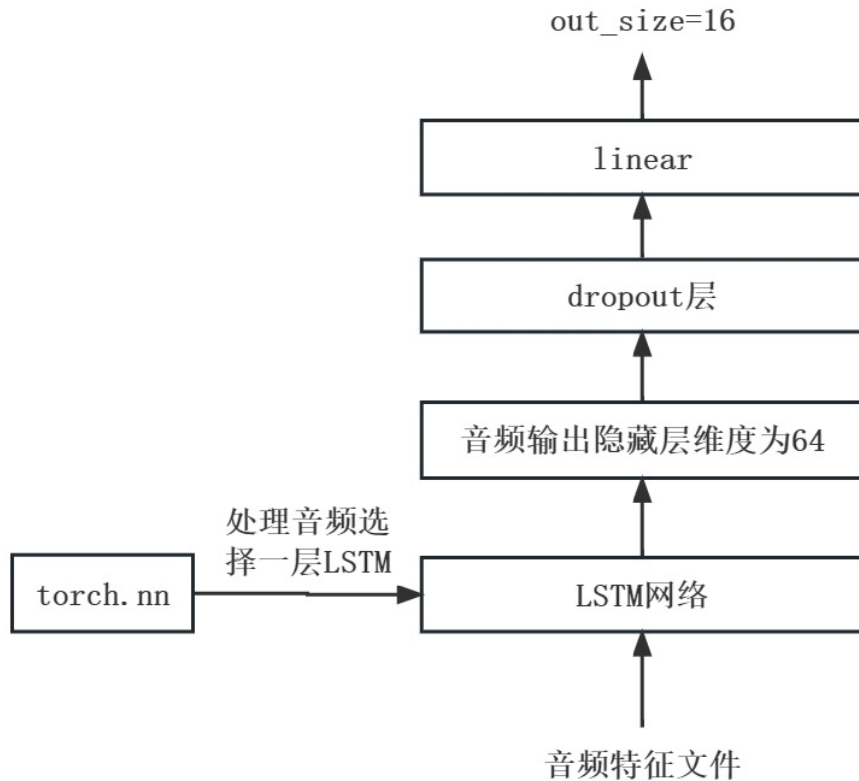
- LSTM

$$i_t = \alpha(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$f_t = \alpha(W_f \cdot [h_{t-1}, x_t] + b_f)$$

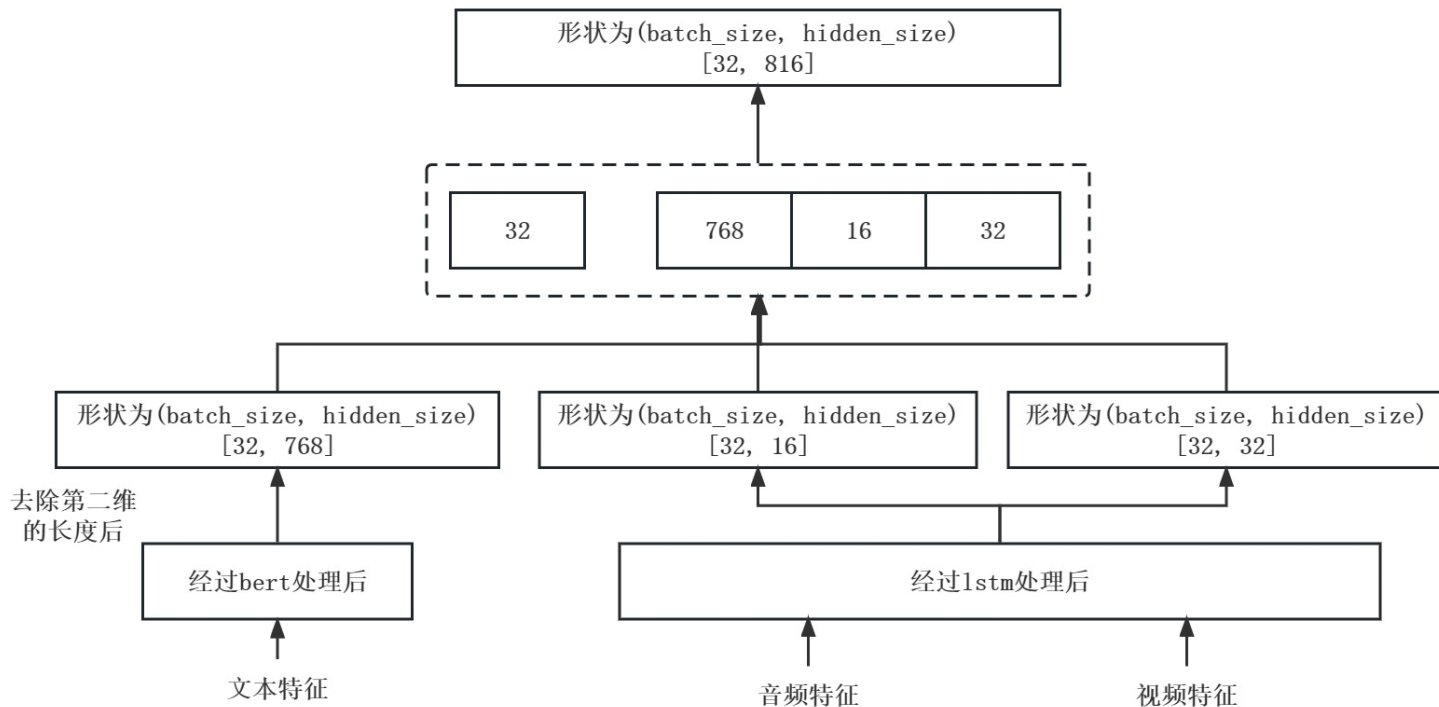
$$\tilde{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$$

$$O_t = \alpha(W_o \cdot [h_{t-1}, x_t] + b_o)$$





特征融合





情感分析

- 分类层

三个线性变换层

两个Relu激活函数

- 两类情感

二分类：积极、消极

五分类：积极、弱积极、中性、弱消极、消极

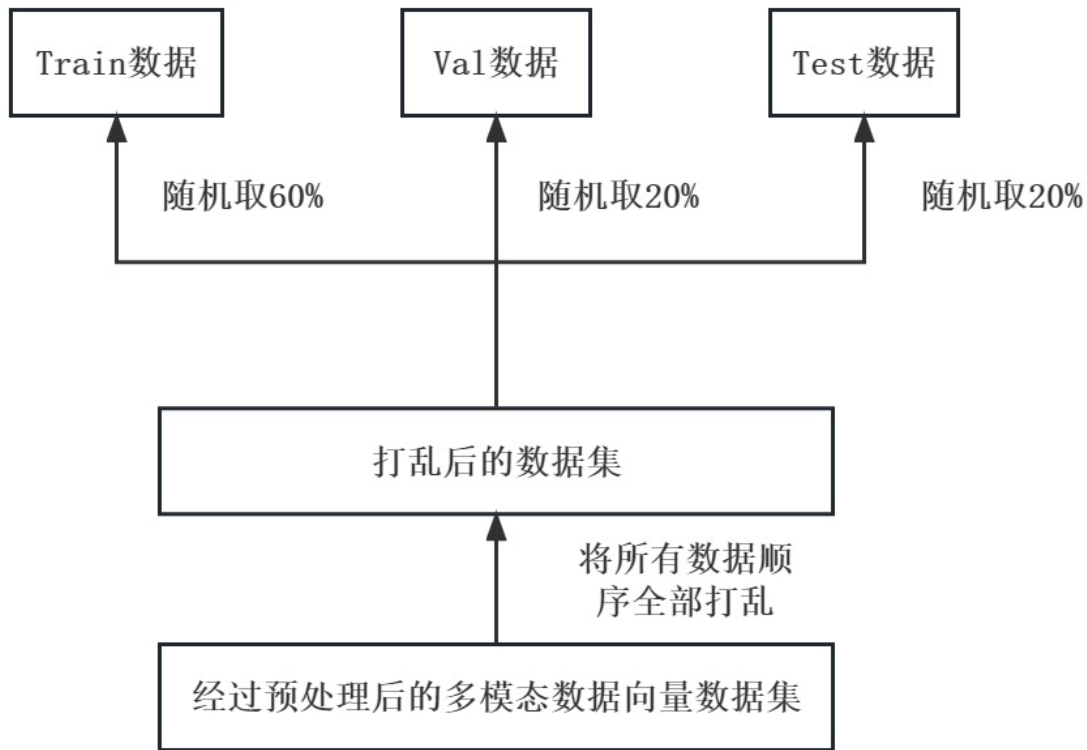


模型训练

第三部分

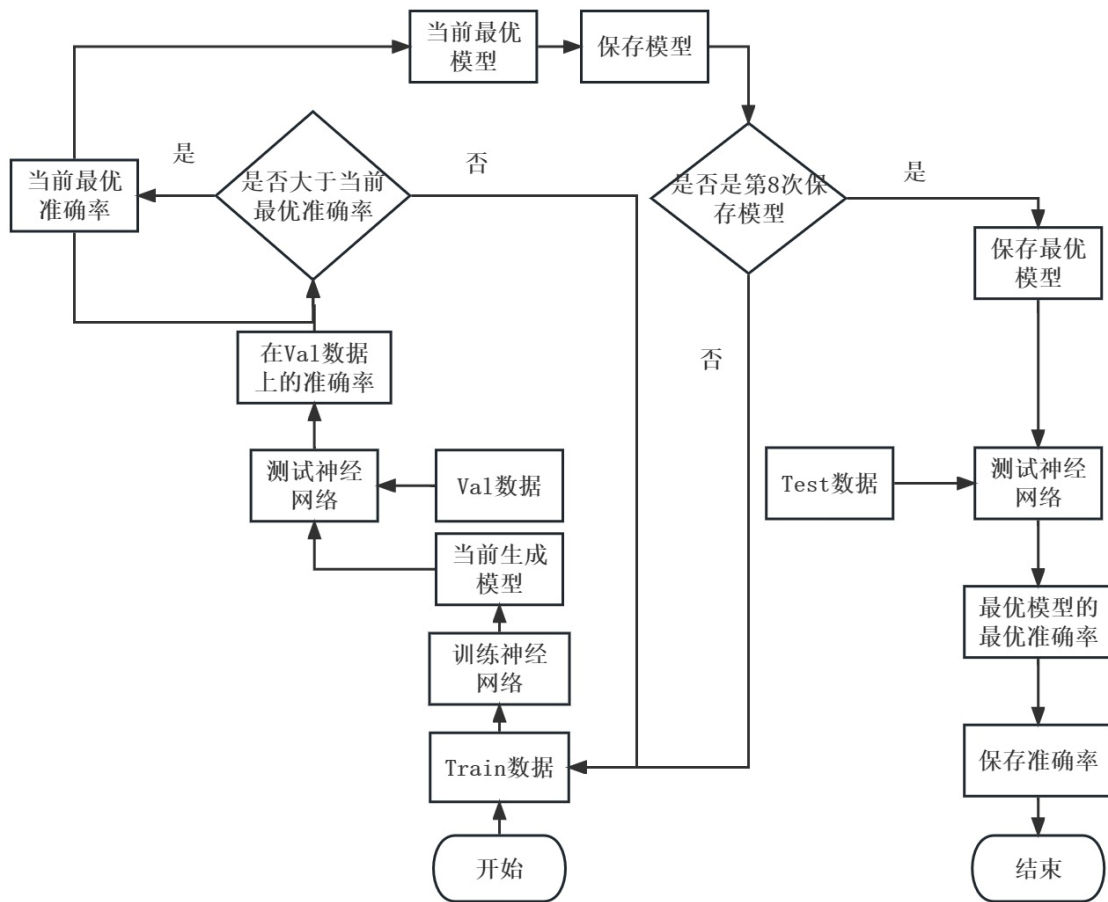


数据集分割





模型训练





功能演示与测试

第四部分



实验环境

主机环境	
操作系统	Windows 11
处理器	12th Gen Intel(R) Core(TM) i7-12700H 2.30 GHz
内存	16.0 GB (15.7 GB 可用)
显卡	NVIDIA GeForce RTX 3060 LapTop GPU 6G
虚拟环境	
Python	3.9.16
Pytorch	1.13.1
Numpy	1.23.5
以及一些其他用于python计算和训练所需的基础包	



实验结果

• 保存的模型

▼ saved_model
boo-sims.pth

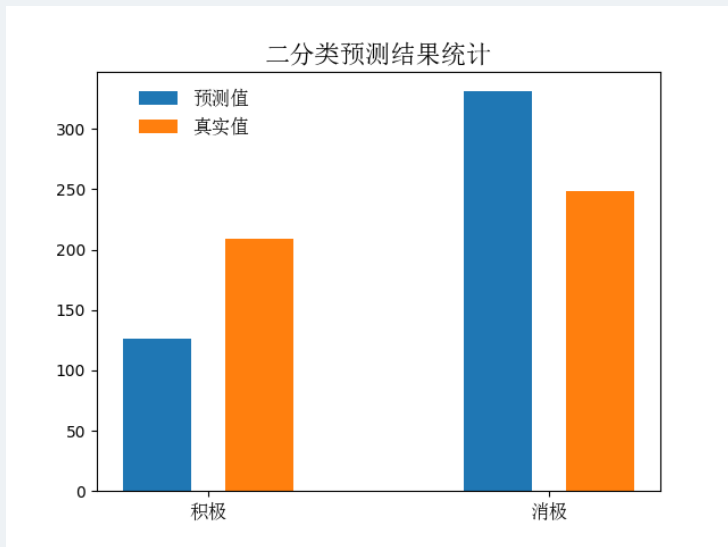
• 测试结果图

数据表格				
序号	音频编号	二分类结果	五分类结果	原始文本
1	video_0003\$_0006	消极	弱消极	这事结婚前咱俩不是说好了？
2	video_0010\$_0038	消极	消极	你妈说来磊儿来了影响方一凡学习。
3	video_0023\$_0007	消极	消极	你看我行吗？
4	video_0021\$_0005	积极	积极	那也让我见识下你的本事。
5	video_0045\$_0009	消极	中性	以我对你的判断我觉得你应该会先派黄橙橙
6	video_0027\$_0039	消极	弱消极	镇长，您把心放宽，多保重。
7	video_0019\$_0002	消极	弱消极	无所谓，乱世。
8	video_0024\$_0031	积极	中性	老马，平时在里边儿都爱喝啥啊
9	video_0016\$_0022	积极	弱积极	现在好了，汽水不冰了，可我的心却是冰冰的
10	video_0024\$_0054	消极	弱消极	倒是也行，你是法人，你看着办。
11	video_0051\$_0034	消极	弱消极	娘
12	video_0010\$_0036	积极	积极	你看他发挥多稳定啊，说明他心理素质特别好。
13	video_0016\$_0033	消极	消极	夏洛夏洛，我赢了，夏洛你知道吗，我为了打败你这套连招我练了多少

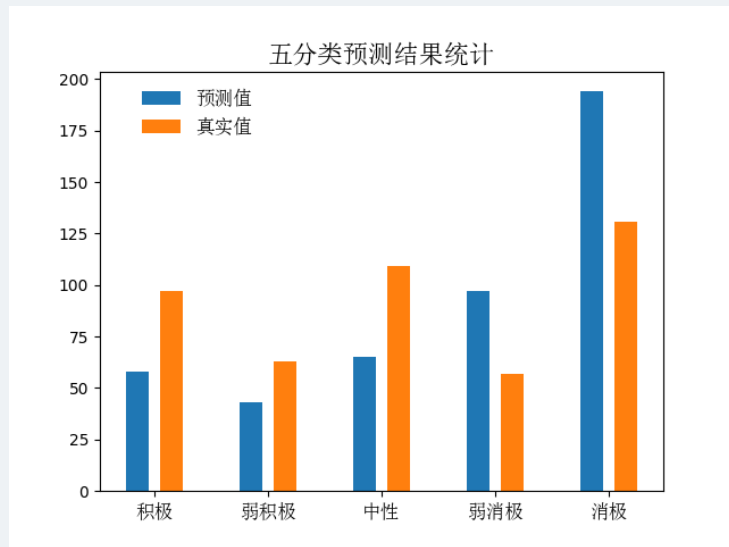


指标参数

- 二分类准确率(SIMS) • 五分类准确率(SIMS)



78.12%

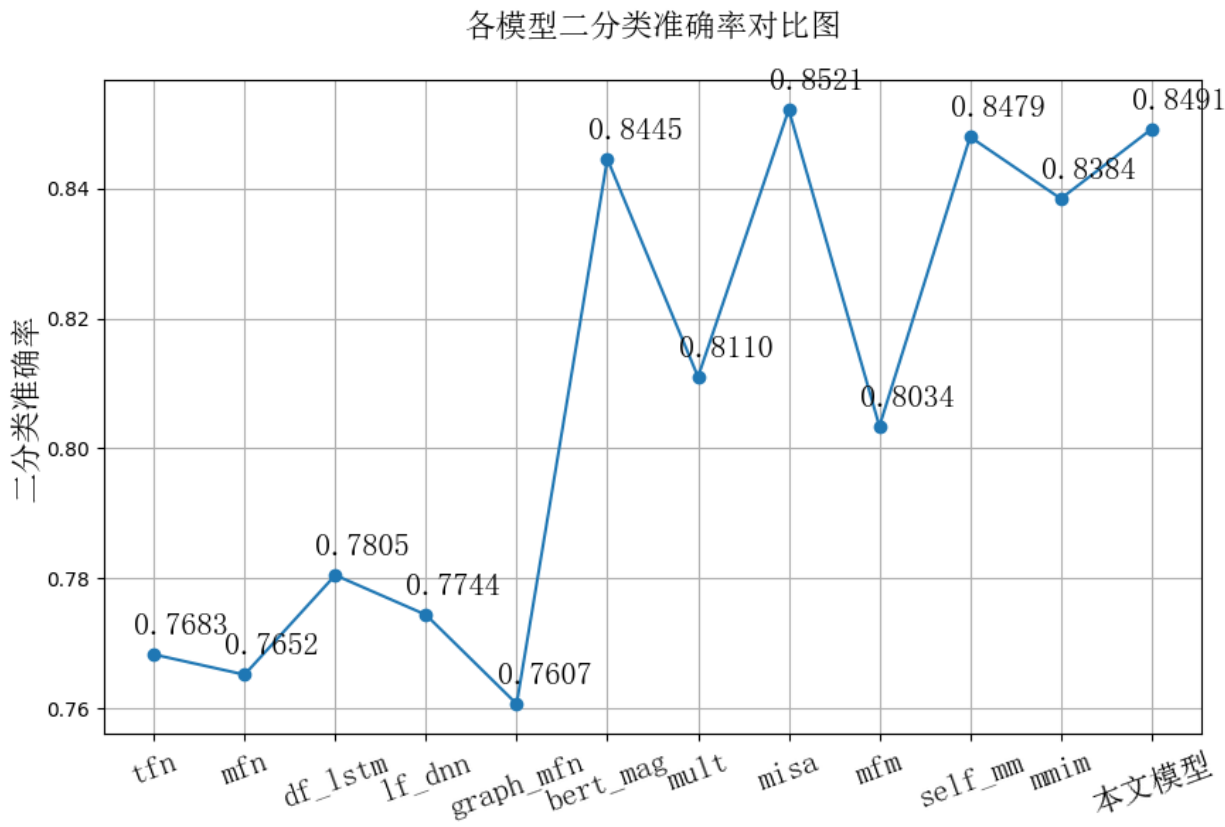


44.86%



模型对比

• 二分类准确率(MOSI)

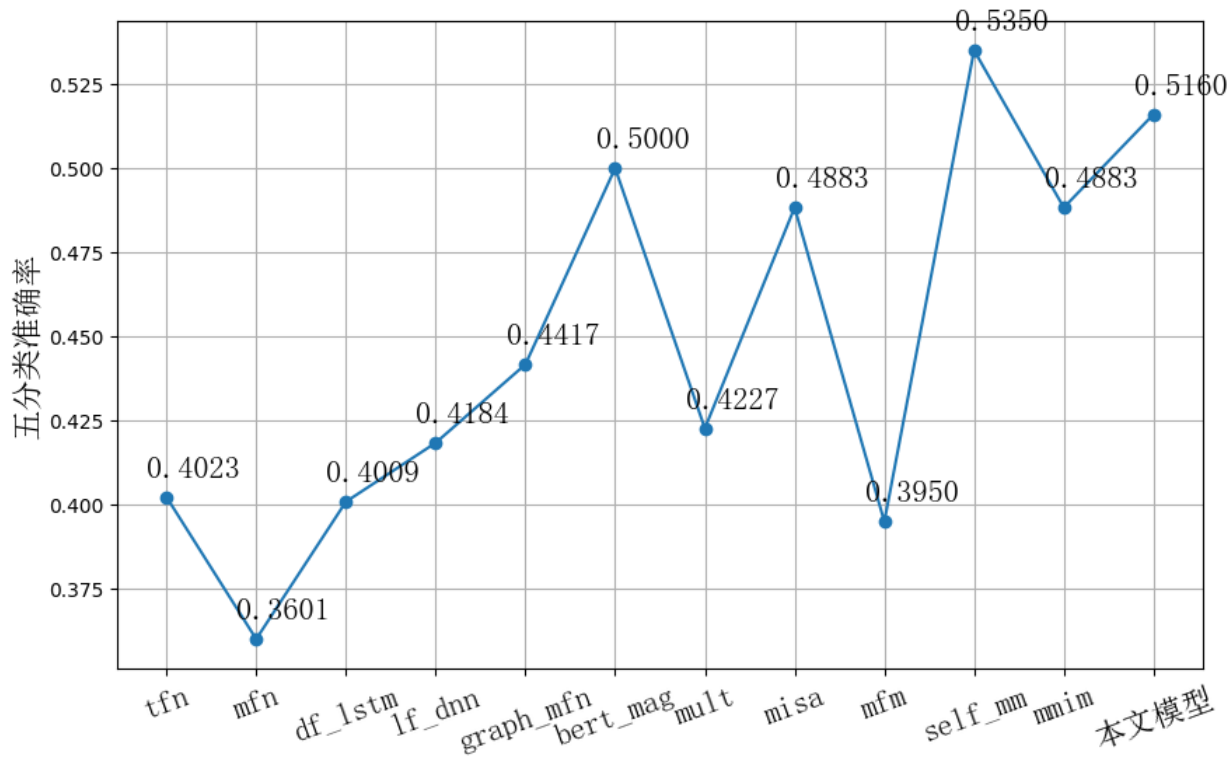




模型对比

• 五分类准确率(MOSI)

五分类准确率对比图





总结与展望

第五部分



总结

- 1、总结了相关研究，提出了一种基于多模态数据的情感分析方法。
- 2、文本：知识增强BERT；音视频：OpenSmile、OpenFace、LSTM。
- 3、对模型性能进行了测试，并与现存模型进行了比较。



展望

- 1、特定研究方向的知识图谱
- 2、特定研究方向的数据集
- 3、更高效的各部分处理方法