



## 个人信息:

---

姓名: 李泽宇

出生日期: 1998 年 9 月 18 日

性别: 男

政治面貌: 共青团员(入党积极分子)

邮箱: [lizeyujack@sjtu.edu.cn](mailto:lizeyujack@sjtu.edu.cn)

电话: 19821612550

专业: 计算机科学与技术

Github: <https://github.com/lizeyujack>

## 教育经历

---

哈尔滨工程大学 计算机科学与技术 本科 哈尔滨 中国 | 2017 年 9 月-2019 年 7 月

悉尼科技大学 数据工程 本科 悉尼 澳大利亚 | 2019 年 7 月-2020 年 11 月

上海交通大学 计算机科学与技术 硕士 上海 中国 | 2021 年 9 月-现在

## 概述

---

本人本科为哈尔滨工程大学计算机科学与技术、悉尼科技大学数据工程双学位。

在哈尔滨工程大学学习期间绩点为 3.58/4.0, 专业排名 1/51, 并连续 3 学期获得校级一等奖学金。在悉尼科技大学期间为全英文授课, 最终的 GPA 为 6.0/7.0。可流利使用英语进行交流, 并已通过英语六级。

2021 年, 保研至上海交通大学计算机科学与技术专业, 师从付宇卓教授; 主要研究方向为深度学习、计算机视觉以及自然语言处理等。

## 专业成绩

---

微积分 A (93) 离散数学一 (93) 离散数学二 (95) 数据结构 (83) 操作系统 (87)

Sensing, Actuation and Control (92)

Introduction to Data Engineering (90)

## 项目经历

---

### (1). 智能商场广告滚动播放程序 组员：3人 悉尼 澳大利亚|2020年2月-4月

- 图像采集：利用 beautiful soup 库爬虫在网络的图片进行爬取，并给予手动标签。
  - 数据扩充：利用对图片的上旋转来扩充数据的数量。
  - 图像识别：调用 tensorflow 预训练模型结合 mobilenet 网络训练人脸识别模型。
  - 项目结果：通过训练，模型的 A1 模型识别的准确率为 85%，A2 准确率为 92%。本项目作为新生引导项目向新同学展示模型分类的效果。
- 

### (2). 爱立信自然语言处理 队员：3人 悉尼 澳大利亚|2020年4月-6月

- 图像采集：爱立信技术团队负责从基站获取故障数据。
  - 数据预处理：通过 word2vector，将文字数据输入为向量的形式进行预处理。
  - 模型：分别尝试 K-means, k-nn 模型对数据进行聚类处理。通过 transformer 对文字信息进行编码以及训练。
  - 项目结果：聚类算法最高可以分类正确 68%的数据；transformer 模型极大的提高了模型的精度，模型分类准确率提升到了 95%。
- 

### (3) 澳大利亚疫情统计（可视化） 队员：4人 悉尼 澳大利亚|2020年4月-6月

- 项目描述：对于澳大利亚的新冠肺炎的疫情数据通过绘制线性图和饼状图对“患病人数，死亡人数和治愈人数”进行可视化。我们还将数据分不同的州进行整理，分别以坐标的形式显示在地图中。
  - 项目职能：从约翰霍普金斯大学的官方的 Github 中提取数据，使用 Python 中的 Bash 库，将数据可视化，显示在网站上。最后，使用 Nginx 文件部署到谷歌虚拟机的服务器中。
- 

### (4) 对抗生成网络白天转黑夜（可视化） 队员：4人 中国 上海|2021年10月-现在

基于 CoMoGAN 模型的循环生成不同时间段的模型，对 waymo 数据集首先进行二次筛选审核。通过改变 A 域以及 B 域的方向，同时改变回传 loss 值的旋转矩阵方程的旋转方向使模型加速收敛。通过训练 40 个 epoches，模型可以生成较为真实傍晚、黄昏、白天的视频样本。

---

(5) 自然语言处理(ijcai-workshop) 队员：4人 中国 上海|2022年1月-3月

- 作为第二作者，提出了一个名为 DCC-Net 模型。
- 方法：它通过对比学习来聚合对话，模型将话语引入相同的会话并在不同的会话中推送话语，因此相同会话中的话语彼此会更亲近。然后进行聚类采用方法生成预测聚类标签。
- 结果：为了缓解缺乏大量注释数据的问题，我们引入了一种策略选择伪样本进行无监督训练，无需任何人工注释。对电影对话进行的综合实验数据集和 IRC 数据集表明我们的模型以有监督和无监督的方式达到最先进的水平。

---

(6) 基于 Transformer 的 deit 模型可解释性研究 中国 上海|2022年4月-现在

- 该项目符合目前主流研究方向，主要是对基于 Transformer 的 DeiT 模型，运用目前的解释算法 calm 算法对其分类过程通过可视化的方式进行解释。
- 模型：使用 DEIT\_small\_16\_patch\_768 作为模型进行训练，使用 CUB 数据集。因为 deit 的模型是使用 ImageNet 数据集进行训练的，并且提供了对于 ImageNet 的预训练模型，通过加载 imagenet 预训练到 CUB 数据集进行训练。
- 成果：我将 deit 中 vit 进行逐层反向求逆获取了 calm 中 latent z 的信息并且将得到的位置信息 z 应用到 calm 中对图像进行热力图可视化解释。效果对比 CAM 具有更准确的边缘预测展示能力。

## 荣誉

---

2017年10月：六系新生辩论赛亚军；

2018年1月：国际大学生雪雕大赛志愿者；

2018年5月：获得春季学期校级一等奖学金；

2018年8月：学术型雅思考试 6.0；

2018年10月：校级创新训练项目（引导型）：良好；

2018年11月：获得秋季学期校级一等奖学金；

2019年5月：获得春季学期校级一等奖学金、校优秀共青团员；

2019年6月：普通话水平测试二级甲等证书；

2022年3月：大学英语六级证书。

## 专业特长

---

可熟练使用 python 语言进行数据分析、数据可视化，数据爬取以及机器学习等项目；

可熟练使用 c++；

熟悉 matlab 语言，了解 matlab 的基本操作；

会使用 SQL 语言对数据库进行基本操作；

会使用 RIMS 环境使用 C 语言进行嵌入式编程；

## 个人评价

---

本人性格开朗，乐于助人；热爱足球、网球、乒乓球等运动，追求上进，善于抓住机会；愿意为集体奉献，有责任感；会用实践的方式将所学的知识应用于现实生活中；专业基础扎实，具有良好的英语交流能力与阅读理解能力。希望可以加入到可以改变世界的技术团队，并且结交到志同道合的同学和朋友。