邹啸天

邮箱: xtzou@ucdavis.edu Github: https://github.com/XiaotianZou 领英: Xiaotian Zou

教育背景

加州大学戴维斯分校 2019年9月 — 2021年3月(预计)
计算机科学硕士 (GPA: 3.74/4.0)
核心课程：*计算理论，计算机架构，图论，分布式数据系统，计算机图形学，数据可视化，线性代数等.*
课程项目包括:

- 利用**Unity**开发了一个VR地图可视化应用
- 重现了一篇数据可视化论文，本人工作主要利用**Django**构建后端系统获取Twitter数据

中山大学 2015年8月 — 2019年6月
软件工程学士 (GPA: 3.5/5.0)
核心课程： *数据结构与算法，操作系统，计算机网络，数据库系统，线性代数, Web 2.0程序设计，现代操作系统应用开发，多核程序设计等.*
课程项目包括(部分):

- 参与开发了一个二维码签到系统，主要采用**Vue.js**开发前端
- 安卓程序"Sports Expert"，主要用**Express.js**搭建后台，并且用**Java**完成前后端交互的API部分
- 运行在Windows 10系统的宿舍管理系统"宿舍宝"，主要采用**C#**开发
- 一个小游戏"Tank Honor"，采用**cocos2d**开发
- 一个简单的基于以太坊智能合约的投票系统，前端逻辑部分用**Vue.js**完成

工作经验

研究助理 2020年2月 — 2020年10月
加州大学戴维斯分校
在Gerald Quon博士的实验室中，我主要做生物相关的机器学习项目，涉及到包括单细胞数据以及脑信号的处理。所有代码都用**Tensorflow 2**和**R**完成。项目包括，

- 采用Transformer模型来捕获外部刺激与脑信号之前的关系，从而分析下游任务。
- 设计了一个预处理方法能够更好的捕获单细胞数据集中高表达基因。
- 设计了一个多任务自编码模型来捕捉药物与细胞基因表达之间的关系。

自然语言处理实习生 2019年2月 — 2019年5月
广州天鹏计算机科技有限公司
主要进行预处理以及分类医学相关文本数据。所有代码都以**Python**完成

- 构建了一个基于BERT的多标签模型来执行症状描述到医学诊断的分类任务。
- 完成了多个网络爬虫来提取药品/医学文本数据。
- 建立了一个基于大量正则表达式的预处理脚本，来清理和重构混乱的原始文本数据。

参与项目（部分）

量化交易策略开发 2018年4月 — 2018年11月
广发证券
本科实验室课题。参与开发涉及小额交易的量化交易策略。所有代码以**Python**的形式完成

- 开发了一个递归强化学习模型来生成交易策略，同时设计了一个LSTM模型来预测股票价格波动。
- 建立了一个CNN模型来验证利用K线图预测股票价格波动的可行性。

微信数据集成应用开发 2017年五月 — 2017年11月
医学图像处理实验室, 中山大学
应用旨在自动帮助用户收集微信群聊信息，并过滤展示群聊当中的通知、文件等。代码主要以**Python**和**Vue.js**实现

- 使用了SQLCipher来破解微信的本地数据库系统. 同时采用了网络爬虫的方式来通过网络版微信更好地提取群聊数据。
- 利用Vue.js参与了部分前端开发
- 构建了一个简单的自然语言处理模型来对收集的文本信息重要性进行分类。

论文

Causality Extraction based on Self-Attentive BiLSTM-CRF with transferred Embeddings

- 第三作者，主要贡献于数据预处理部分。
- 文章当前链接 https://arxiv.org/pdf/1904.07629.pdf.

专业技能

编程语言	Python, Javascript/ Node.js, C++, C#, Java, R, HTML5, CSS, SQL
网络开发	Vue.js, Django, MongoDB, Express, MySQL
机器学习	Tensorflow 2, Keras, scikit-learn
其它工具	Linux, Postman, Slurm, SSH, Git

ZOU XIAOTIAN

Email: xtzou@ucdavis.edu Github: https://github.com/XiaotianZou Linkedin: Xiaotian Zou

EDUCATION

University of California, Davis Sep 2019 — Mar 2021(expected)

Master in Computer Science (GPA: 3.74/4.0)

Core modules include *Theory of Computation, Computer Architecture, Networks, Distributed database system, Computer Graphics, Information Visualization, etc.*

Course Projects during study:

- Developed a VR map visualization application using **Unity**.
- Re-implementation of a visualization paper. I mainly worked on building the back-end system with **Django**.

Sun Yat-sen University Aug 2015 — Jun 2019

Bachelor of Software Engineering (GPA: 3.7/4.0)

Core modules include *Data Structures and Algorithms, Operating System, Computer Networks, Database system,Linear Algebra, Web 2.0 Programming, Modern Operating Systems Application Development, Multi-core Programming, etc.*

Course Projects during study(selected):

- Participated in developing a check-in application with QR code with **Vue.js**
- An android application named "Sports Expert". Worked on back-end development with **Express**, also finished part of the front end design with **Java**
- A dormitory management helper system running on UWP. Worked on the front end part with **C#**
- A mini game "Tank glory" developed with **cocos2d**
- A simple voting system based on **Ethereum** framework. The front-end is implemented with **Vue.js**

PROFESSIONAL EXPERIENCE

Research Assistant Feb 2020 — Current

University of California, Davis, US

In Dr.Gerald Quon's lab, I mainly worked on bio-related machine learning projects, involving single cell data and brain signal processing. Implemented in **Tensorflow2** and **R**

- Implemented Transformer model to capture the inner relations within stimulus and the brain responses.
- Designed a pre-processing pipeline that could better preserve the information from the single cell data set.
- Built a multitask auto encoder to be trained on finding the relationships across drugs and the corresponding gene expressions.

Natural language processing Intern Feb 2019 — May 2019

Tianpeng Computer Technology Co.ltd, Guangzhou, China

Worked on processing and classifying drug and medical text data. All codes were implemented in Python.

- Built a BERT based multi-label model to classify the symptom descriptions with the diagnosis labels.
- Set up several **web crawlers** to extract drug/medical information for further building of knowledge map.
- Established a **regular expression** based pipeline that can clean and restructure the muddled raw text data set.

SELECTED PROJECTS

Quantitative Trading Strategy Development Apr 2018 — Nov 2018

GF Securities, Guangzhou, China

Aimed at developing a quantitative trading strategy that can advise customers to buy/sell stocks based on small amount transaction. Implemented in **Python.**

- Developed a recurrent reinforcement learning model to generate transaction advice, along with a LSTM model to predict stock price.
- Established a CNN to test the feasibility of predicting stock price from the K-line diagram.

WeChat Information Integration Application May 2017 — Nov 2017

Laboratory of Computerized Medical Imaging and Graphics, Sun Yat-sen University, China

Aimed at developing an application that can help users to automatically filter and present the important information in group chat. Implemented in **Python** and **Vue.js**

- Used SQLCipher to crack WeChat local database. Also proposed to apply web crawlers to gather group chat information through the web version of WeChat.
- Participated in part of the front-end design.
- Established a simple natural language processing model to classify the information importance.

PUBLICATIONS

Causality Extraction based on Self-Attentive BiLSTM-CRF with transferred Embeddings

- The third author, mainly worked on data preprocessing.
- The paper is currently available at https://arxiv.org/pdf/1904.07629.pdf.

PROFESSIONAL SKILLS

Programming Languages	Python, Javascript/ Node.js, C++, C#, Java, R, HTML5, CSS, SQL
Web Development	Vue.js, Django, MongoDB, Express, MySQL
Machine learning library	Tensorflow 2, Keras, scikit-learn
Other tools	Linux, Postman, Slurm, SSH, Git