Resume

Yi-Hua Guo 郭益華

g11350006@thu.edu.tw
jerry7776112@gmail.com
Linkedin
Github

Education

02/2022-Present GPA: 4.12/4.30
Tunghai University, Taichung, Taiwan
Master of Engineering, Computer Science
06/2016-06/2020 GPA: 3.17/4.30
Tunghai University, Taichung, Taiwan
Bachelor of Engineering, Environmental Science and Engineering
Courses taken include:
C++ Programming, Decision Support System, Internet of Things (IOT) for healthcare, Information Assurance, Study and Implementation of Big Data and Industry Applications, Advanced Software Engineering Special Topic, Information Management, 5G Network Systems and Applications, The Business Value Creation of Digital Technology, New Economy and Business Model Innovation.
Certificates:

- TUV NORD - Sustainability Management Manger(企業永續管理師)
- 2023 AIS3
- 2021 企業數據競賽
- Career Essentials in Data Analysis by Microsoft and LinkedIn
- Career Essentials in Generative AI by Microsoft and LinkedIn
- AWS Academy Graduate - AWS Academy Cloud Foundations
- Great Learning - Bascis of EDA with Python

Experience

2/2022-Present Research Assistant, Secure Social Computing Laboratory

- Collect public data from social networking platforms for big data analysis and disinformation tracking.
- For Cyber security, build a "Penetration Testing" and "Vulnerability Assessment" environment using VMware virtual machine, and use a second virtual machine, Kali Linux, as the attacking party for testing; paper review, paper writing, and presentation skills.

Skill Sets: Virtual Machine, Penetration Testing Concepts, Vulnerability Assessment Concepts, Cyber Security Implementation, Big Data Analysis, Big Data Processing, Social Network Analysis, paper review, paper writing, presentation skills

1/2023-6/2023 Research Assistant-資安卓越中心規劃建置計畫, National Institute of Cyber Security(國家資通安全研究院)

- Based on the developed API & ETL pipeline, about 30 million pieces of public data of social networks were obtained.
- A social network analysis study was conducted using approximately 30 million pieces of public data.
- Based on the research questions, we visualized the social network for the social network message

dissemination phenomenon study.

Skill Sets: Big Data Analysis, Big Data Processing, Data Visualization, Social Network Analysis

6/2022-12/2022 Research Assistant-資安卓越中心規劃建置計畫, National Institute of Cyber Security(國家資通安全研究院)

- The VMware virtual machine build uses a combination of deep learning semantic similarity analysis and artificial intelligence to develop an online social network disinformation tracking API.
- Packages VMware for delivery as a VMDK file.
- The development process uses Git for version control and facilitates collaboration with other engineers.
- About 30 million social network public data were built into the PostgreSQL database, and design the database index to improve the search performance from the original 30 seconds per search to 0.77 seconds.
- Using Python Flask web application framework was combined with PostgreSQL to develop the API
- About 6 million social network public data were built into the MongoDB database, and design the database index to improve the search performance
- Using Python Flask web application framework was combined with MongoDB to develop the API

Skill Sets: Virtual Machine, Natural Language Processing, Deep Learning, API Development, SQL, Git, Team-Work

2/2022-6/2022 Teaching Assistant, Tunghai University, Computer Science Department

- Instructing discussion sessions in the course of Artificial Intelligence.

Skill Sets: Basic Artificial Intelligence Concepts, Basic Computer Science Concepts, Communication & Team-Work

06/2021-10/2021 Member, Workforce Development Agency, MOL, AI Big Data Enterprise Practice Application Class

- Learn basic concepts of Python and R programming, data processing skills, as well as data visualization and artificial intelligence machine learning concepts.
- Participated in the "2021 Data Station" competition, using the "PChome product purchase list" dataset provided by the organizer.
- Based on the dataset, I processed and added new features to the data, and used third-party data to overlay it, and used Python machine learning algorithms such as decision trees and random forests to analyze the frequency of customers' purchases.

Skill Sets: Python Programming, R Programming, Data Processing Skills, Data Visualization, Data Analysis Based on Enterprise Data, Basic Artificial Intelligence Concepts

Projects

- DockerTutorial(Zero to Hero)-NOTE
- ETL 語意相似度分析
- Python & Kafka 餐點訂單分散式系統設計
- SQL to Power BI Pizza 銷售資料視覺化
- 使用 pgAdmin 操作 Docker 上的 PostgreSQL
- 系統 CPU & Memory 監控 Flask & Docker 實作
- 利用 AWS 建立 Open Weather ETL 並整合至 Apache Airflow
- 資料模型設計實作
- AIS3-信用卡盜刷偵測
- 新經濟創與創新營運(創新提案)_穆斯林 SuperApp
- 數位科技的商業價值創造(創新提案)_虛擬生命園區平台

Computer Skills

Languages: Python, R, C++, Go
AI Library: Sentence-Transformers, Scikit-learn
Data Processing: Pandas, Numpy
ETL Tool: Airflow, SSIS
Data Visualization: PowerBI, Seaborn, Matplotlib, Plotly
Social Network Visualization: Gephi
Python Web Framework: Flask

SQL: MySQL, PostgreSQL, MongoDB
Virtual Machine: VirtualBox, VMware, Hyper-V
OS: Windows, Linux
Version Control: Git
Editor: VScode, Nano, Vim
Scripts: Bash
Cloud: AWS
Container: Docker, Kubernetes

Publications

- Tracking of Disinformation Sources Based on Online Social Media: Examining Pages and URLs with BFS Evolution | CISC2023 第三十三屆全國資訊安全會議
- 基於區塊鏈建立去中心化社群平台之共識演算法研析 | CISC2022 第三十二屆全國資訊安全會

Awards

Session C4 : Social Media Security-Best Session Award
Tracking of Disinformation Sources Based on Online Social Media: Examining Pages and URLs with BFS Evolution | CISC2023 第三十三屆全國資訊安全會議