CAIJUN OIN

Gainesville, FL

Contact: (352) – 872 – 6633 ♦ qcaijun2013@gmail.com

Websites: https://caijungin.wixsite.com/portfolio ♦ https://devpost.com/Fennec2000/

Linkedin: https://linkedin.com/in/cq-profile GitHub: https://github.com/Fennec2000GH

EDUCATION

Aug 2019 -UNIVERSITY OF FLORIDA ♦ GAINESVILLE, FL

May 2022 Major: Computer Science B.S. & Statistics B.A. Major GPA: 3.79/4.00

> Services: Treasurer of ACM, Society of Software Developers, and UF Hackathoners; Project Lead in Open Source Club and Digital Science and Informatics; Tech Lead in Google DSC; Webmaster of SWE Club Courses: Machine Learning, Statistical Learning, Competitive Prog., Data Science, Prog. with Data in R

Aug 2018 – FLORIDA STATE UNIVERSITY ♦ TALLAHASSE, FL

May 2019 Major: Computer Science B.S. Major GPA: 4.00/4.00 ♦ University Honors Program

Services: Treasurer of ACM

Courses: Programming I, Object-Oriented Programming, Linear Algebra, Calculus III

RESEARCH & PROJECT EXPERIENCE

Jun 2021 -Using AI to Trace the History of Race and Inequality ◆ Undergraduate Research, Dept. of Classics, UF

Present

- Engineered NLP pipeline to retrieve, transform, and index Latin and Greek texts for querying Leads student team to extract raw text sections from digital collections of classics in XML format
- Leveraged high performance computing with Apache Spark and Dask for parallel document retrieval

Data Analytics and Info Retrieval at Univ. of North Texas ♦ NSFREU, Dept. of Information Science, UNT Jun 2021 -

Aug 2021

- Comparison study of traditional and deep learning algorithms on legal text classification (JURISIN 2021)
- Investigated dataset-related factors explaining why traditional ML generally performed more accurately
- Performed transfer learning with BERT and Sentence-BERT to classify corporate contracts and clauses
- **Open Ocean** ♦ Digital Ocean Hackathon Aug 2021 –
- Individually created web app for public sharing of articles centered around technology and science Aug 2021
 - Integrated Fetch API with MongoDB database to manage submitted articles
- May 2021 -**PowerSearch** ♦ Ryerson University Hackathon
- May 2021 Programmed Discord bot that can summarize, analyze sentiment, and recommend topics from the web
 - Applied Datastax Astra to store bot replies with high a vailability and won company's sponsored prize
- Jan 2021-Society of Software Engineers Website ♦ UF Software Engineering Course
- Primarily developed blog section of new website with Angular for on-campus CS student organization Apr 2021
 - Assisted in various tasks, including designing wireframes and testing new features, as part of Agile team
- May 2020 -Predictive Sampling Method for Spread Models in Networks ♦ University Scholars Program, UF
- April 2021 Developed new sampling method for large networks by setting quotas for high-degree nodes
- Authored paper published to the UF Journal of Undergraduate Research
- May 2020-**Backend Engineer for OCR Note-taking Application ◆** UF Performant Programming Course
- Aug 2020 Integrated machine learning and database functionalities for optical character recognition (OCR) app
 - Created pipeline to pre-process text images with OpenCV and Pillow to boost text prediction accuracy
- Jan 2020-American Sign Language Image-to-Letter Translator ♦ UF Machine Learning Course
- Apr 2020 Collectively optimized ASL translation with supervised KNN model (~90% accuracy) using Scikit-Learn
 - Engineered pipeline that preprocesses image, trains classifier, predicts letter, and evaluates accuracy

AWARDS, CERTIFICATIONS, & RECOGNITION

Mar 2022	NVIDIA Deep Learning Institute: Building Transformer-Based NLP Applications
----------	---

- Feb 2022 NVIDIA Deep Learning Institute: Applications of AI for Anomaly Detection
- Apr 2021 MATLAB Onramp & MATLAB Machine Learning Onramp
- Jul 2020 M001: MongoDB Basics
- Mar 2019 1st Place in Lower Division of FSU Spring Programming Competition

PROFESSIONAL SKILLS

Programming and Markup Languages

Intermediate to Proficient: Python, C++, Java, R, Rust, Julia

Beginner to Working Knowledge: MATLAB, Go, Solidity, C#, HTML/CSS/JS, XML, YAML

Databases, Frameworks, and Tools

Working knowledge of CircleCI, Apache Spark, Dask, MongoDB, SQLite, MySQL, CockroachDB, GCP APIs, Node.js, React

Articulate communicator and writer with fluency in English and Mandarin Chinese