

# Caijun Qin

**Location:** Gainesville, FL ♦ **Phone:** (352) – 872 – 6633

**Email:** qcaijun2013@gmail.com ♦ **Website:** <https://caijunqin.wixsite.com/portfolio>

**Linkedin:** <https://linkedin.com/in/cq-profile> ♦ **GitHub:** <https://github.com/Fennec2000GH>

---

## EDUCATION

- May 2022 **UNIVERSITY OF FLORIDA** ♦ Gainesville, FL, USA  
**Major:** Computer Science B.S. & Statistics B.A. **Major GPA:** 3.79 / 4.00  
**Relevant Coursework:** Software Engineering, Data Structures and Algorithms, Machine Learning, Performant Programming in Python, Competitive Programming, Principles of Programming Language, Data Science, Programming with Data in R, Statistical Learning
- May 2019 **FLORIDA STATE UNIVERSITY** ♦ Tallahassee, FL, USA  
**Major:** Computer Science B.S. **Major GPA:** 4.00 / 4.00  
**Relevant Coursework:** Intro Programming I (C++), OO Programming, Linear Algebra, Calculus III

---

## RESEARCH INTERESTS

Machine Learning, Natural Language Processing, Cognitive Computing, Graph Theory, Network Science

---

## UNIVERSITY SERVICES

- Nov 2021 – **Data Science and Informatics (DSI)**  
May 2022 *Project Coordinator* ♦ *University of Florida*
- Aug 2021 – **UF Chapter of Association of Computing Machinery (ACM)**  
May 2022 *Treasurer* ♦ *University of Florida*
- Aug 2021 – **Open Source Club (OSC)**  
May 2022 *Project Lead* ♦ *University of Florida*
- Aug 2021 – **Google Developer Student Club (GDSC)**  
May 2022 *Tech Lead* ♦ *University of Florida*
- Aug 2021 – **Software Engineering Club (SEC)**  
May 2022 *Webmaster* ♦ *University of Florida*
- Aug 2021 – **Society of Software Developers (SSD)**  
May 2022 *Treasurer* ♦ *University of Florida*
- Jan 2021 – **UF Hackathoners**  
May 2022 *Treasurer* ♦ *University of Florida*
- Apr 2019 – **FSU Chapter of Association of Computing Machinery (ACM)**  
Aug 2019 *Treasurer* ♦ *Florida State University*

---

## RESEARCH EXPERIENCE

- Jan 2022 – **Lake Kivu Plant Pathology Analysis** ♦ Undergraduate Research, Dept. of Plant Pathology, IFAS, UF  
Present *Undergraduate Research Assistant* ♦ *Department of Plant Pathology, IFAS, University of Florida*  
*Advisor: Dr. Karen Garrett*
- Constructed ML models for crop disease prediction from weather patterns provided by NASA Power API
  - Upped efficiency by ~10-fold in training, evaluating, and ranking models via GPU-enabled H2O AutoML
  - Standardized data cleaning with pipelines built using Apache Spark extensions in R (sparklyr)
- Jun 2021 – **Using AI to Trace the History of Race and Inequality**  
Present *Undergraduate Research Assistant* ♦ *Department of Classics, College of Liberal Arts & Sciences, University of Florida*  
*Advisor: Dr. Eleni Bozia*
- 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
- Jun 2021 – **“Data Analytics and Information Retrieval” NSF Research Experience for Undergraduates**  
Aug 2021 *NSF REU Recipient* ♦ *Department of Information Science, College of Information, University of North Texas*  
*Advisor: Dr. Junhua Ding*
- Quantitatively compared traditional machine and deep learning algorithms on legal text classification
  - Analyzed factors of dataset to explain why traditional ML, especially boosting algorithms, performed best
  - Performed transfer learning with BERT and Sentence-BERT to classify corporate contracts and clauses
  - First author of research paper accepted to the JURISIN 2021 workshop and proceedings

- May 2020 – **Predictive Sampling Method for Spread Models in Networks**  
April 2021 *Undergraduate Researcher* ♦ *University Scholars program, University of Florida*  
*Advisor: Dr. Peter Dobbins*
- Developed new sampling method for large networks based on quota sampling of high-degree nodes
  - Authored paper published to the UF Journal of Undergraduate Research

---

## PROJECT EXPERIENCE

---

- Jan 2021 – **List of Hackathons**  
Present *Participant* ♦ *Organizer Varies*
- Frequently compete in hackathons hosted by various institutions and organizations
  - Selected list of projects: <https://github.com/Fennec2000GH/Hackathon-Repository-Hub/blob/main/README.md>
  - Online full portfolio of projects: <https://www.devpost.com/Fennec2000>
- Aug 2021 – **Plant Root Analysis Using Machine Learning**  
Dec 2021 *Student / Group Member* ♦ *Senior Project Course, University of Florida*  
*Advisor: Dr. Alina Zare*
- Performed image segmentation of minirhizotron (MR) root images using U-Net deep learning architecture
  - Improved original model by tweaking hyperparameters and training with only subset of original training data
  - Programmed framework to apply multiple cumulative learning paradigms on models created with PyTorch
- Mar 2021 – **Theoretical Modeling of Dynamic Vegetation in Agricultural Terrains for Active Passive Microwave Retrieval of Soil and Crop Parameters**  
Oct 2021 *Undergraduate Research Assistant* ♦ *Institute of Food and Agricultural Sciences, University of Florida*  
*Advisor: Dr. Jasmeet Judge*
- Developing functional-structural plant model (FSPM) in Blender and SpaceClaim to model crops across growth stages
- May 2020 – **OCR Note-taking Application**  
Aug 2020 *Backend Engineer* ♦ *Performant Programming Course, University of Florida*
- Integrated machine learning and database functionalities for optical character recognition (OCR) [app](#)
  - Maintained customizable pipeline to preprocess text images with OpenCV and Pillow
  - Experimentally optimized parameter selection and preprocessing steps to boost text prediction accuracy
- Jan 2020 – **American Sign Language Image-to-Letter Translator**  
Apr 2020 *Student / Group Member* ♦ *Intro to Machine Learning Course, University of Florida*
- Collectively built classification system for ASL translation with supervised KNN model (~90% accuracy)
  - Engineered pipeline that preprocesses image, trains classifier, predicts letter, and evaluates accuracy
- Jan 2019 – **Cost Minimization and Optimization of Criteria-based Matchings**  
May 2019 *Honors Project Student* ♦ *University Honors Program, Florida State University*  
*Advisor: Dr. Peixiang Zhao*
- Implemented Hungarian Algorithm in C++ to optimally choose pair from bipartite graph representing costs
  - Designed similar algorithm but which optimally chooses pairs based on sum of positive attribute values

---

## PUBLICATIONS

---

- Qin, C.,** Yang, Y., Chen, H., & Ding, J. (2021). *A Comparison Study of Machine Learning and Deep Learning for Legal Contract Understanding* [Manuscript submitted for publication], *Department of Computer & Information Science & Engineering (CISE), University of Florida.*
- Qin, C.** (2021). Predictive Sampling Method for Spread Models in Networks. *UF Journal of Undergraduate Research*, 23(Fall 2021). <https://doi.org/10.32473/ufjur.v23i.128429>

---

## PRESENTATIONS

---

- Qin, C.,** Yang, Y., Chen, H., Ding, J. (2021, November). A Comparison Study of Machine Learning and Deep Learning for Legal Contract Understanding [Paper presentation]. In *International Workshop on Juris-Informatics 2021 (JURISIN 2021)* (pp. 110–123), Keio University, Yokohama, Kanagawa, Japan.
- Qin, C.** (2021, March). Predictive Sampling Method for Spread Models in Networks. In *2021 Virtual Spring Undergraduate Research Symposium*, University of Florida, Gainesville, Florida, USA.

---

## AWARDS & HONORS

---

July 2022	<b>MLH Top 50</b> <i>MLH Top 50 Hackers of 2022 ♦ Major League Hacking, NY, USA</i>
May 2021	<b>Best Use of Datastax Astra</b> <i>RU Hacks 2021 ♦ Ryerson University, Toronto, ON, CAN</i>
Feb 2021	<b>Best Covid-19 Hack</b> <i>BrickHack 7 ♦ Rochester Institute of Technology, Rochester, NY, USA</i>
Feb 2021	<b>4<sup>th</sup> Place</b> <i>EconHacks 2021 ♦ Virtual Hackathon</i>
Jan 2021	<b>InfoTech Challenge</b> for most innovative use of a public dataset for the public good <i>SwampHacks VII ♦ University of Florida, Gainesville, FL, USA</i>
Mar 2019	<b>1<sup>st</sup> Place in Lower Division</b> <i>FSU Spring 2019 Programming Competition ♦ Florida State University, Tallahassee, FL, USA</i>
Jan 2019	<b>University Honors Program</b> <i>Lateral Admission into Honors Program ♦ Florida State University, Tallahassee, FL, USA</i>

---

## FUNDING & SCHOLARSHIPS

---

Jun 2021	<b>NSF REU: College of Information at UNT</b> <i>\$7000 ♦ University of North Texas, Denton, TX, USA</i>
Mar 2021	<b>Gartner Group Information Technology Fund</b> <i>\$1000 ♦ University of Florida, Gainesville, FL, USA</i>
May 2020	<b>Russell and Mary Hyatt McCaughan Scholarship</b> <i>\$1000 ♦ University of Florida, Gainesville, FL, USA</i>
Feb 2020	<b>University Scholars Program Stipend</b> <i>\$1750 ♦ University of Florida, Gainesville, FL, USA</i>
May 2018	<b>University Freshman Scholarship</b> <i>\$1200 / Semester ♦ Florida State University, Tallahassee, FL, USA</i>

---

## CERTIFICATIONS

---

Mar 2022	<b>NVIDIA Deep Learning Institute: Building Transformer-Based NLP Applications</b>
Feb 2022	<b>NVIDIA DLI: Applications of AI for Anomaly Detection</b>
Apr 2021	<b>MATLAB Machine Learning Onramp</b>
Apr 2021	<b>MATLAB Onramp</b>
Jul 2020	<b>M001: MongoDB Basics</b>
May 2015	<b>Microsoft Office Specialist: Word 2013</b>
May 2015	<b>Microsoft Office Specialist: Excel 2013</b>
May 2015	<b>Microsoft Office Specialist: PowerPoint 2013</b>

---

## PROFESSIONAL SKILLS

---

<b>Languages</b>	English, Mandarin Chinese
<b>Programming Languages</b>	<i>Intermediate / Proficient:</i> Python, C++, Java, R, Rust, Julia <i>Beginner / Working Knowledge:</i> JavaScript, Matlab, C#, Solidity, Go
<b>Markup Languages</b>	HTML, CSS, XML, YAML
<b>Databases</b>	MySQL, SQLite, CockroachDB, MongoDB, Cloud Firestore
<b>Frameworks &amp; Tools</b>	CircleCI, Apache Spark, Dask, Node.js, GCP APIs, React