Ben Yang

CONTACT INFORMATION

Home Address:
6 Brunelleschi Court
Monmouth Junction, NJ 08852

+1 (609) 356-8275
by2297@columbia.edu
benjyang.com

RESEARCH INTERESTS

My research interests are in augmented reality and virtual reality. More specifically, I'm interested in multimodal perception and biometrics in VR/AR, and how ML/AI supports these intersections.

EDUCATION

Columbia University, New York, NY — GPA: 3.95/4.00 — Jan 2021 - Dec 2022
M.S. in Computer Science: Vision, Graphics, Interaction, and Robotics Track

Rutgers University, New Brunswick, NJ — GPA: 3.89/4.00 — Sept 2016 - May 2020
B.S. in Computer Science, Math, Cognitive Science, Minor in Philosophy

HONORS AND AWARDS

Dean's List, Columbia University — Jan 2021 - Dec 2022

Honors College Scholar, Rutgers University — Sept 2016 - May 2020

Cap & Skull Senior Honors Society, Rutgers University — 2020

hackNY Fellow, New York, NY — 2018

Dean's List, Rutgers University — Sept 2016 - May 2020

National Merit Finalist Scholarship, Rutgers University — Sept 2016 - May 2020

RESEARCH EXPERIENCE

Computer Graphics & UI Lab, Columbia University, *Research Assistant* — Sept 2021 - Present

Worked with Professor Steven Feiner for the AntARctica & MercuryMessaging projects:

- Manipulated and cleaned ~ terabytes of radar data from Antarctica ecological projects to construct a Unity VR application for easy data viewing and visualization, including networking
- Ported application from Hololens to Quest 2, reconfigured all manipulation interactions and reduced lag of application from 30FPS to 80FPS
- Built a radar to OBJ converter that compresses the OBJ to avoid Unity loading errors
- Released Version 2.0 of a network routing / event handling library in Unity for MercuryMessaging, which includes networking, APIs for study design, and updated tutorials

Software Systems Lab, Columbia University, *Research Assistant* — Jan 2022 - Present

Worked with Professor Junfeng Yang for the Cerberus project:

- Modified Google Electra to take source code instead of natural language data and pretrained / finetuned models for bug detection
- Pretrained various BERT, Grace, SPT-Code, etc. models from scratch and then finetuned on Java Defects4J dataset to compare against Cerberus for an ablation study
- Achieved a 70% accuracy on detecting security vulnerabilities in C code (the Devign dataset) using NLP based off Google Electra

AWS CodeGuru, Amazon, *Applied Scientist Intern* — May 2022 - Sept 2022

- Created AST parser for Java, Javascript, Python to be used in ML CodeGuru models, used parser to generated gigabytes of data from private Amazon and public Github codebases
- Wrote commit history dataset builder for any Github / Amazon repository, reduced dataset creation time from minutes to seconds (100x improvement)
- Built graph-based bug fixing / detection ML models using custom AST parser

Alexa AI, Amazon, *Applied Scientist Intern* May 2021 - Sept 2021

- Leveraged multilingual language models such as mBART and XLM-r for semantic parsing tasks
- Annotated and parsed 2.5 million data utterances, used machine translation on French, Hindi, and Italian data for training
- Achieved an 80% accuracy on English, and 60% accuracy on French, Hindi, and Italian data

Combra Lab, Rutgers University, *Research Assistant* Spring 2019

Worked with Professor Konstantinos Michmizos with the Astrocytes project:

- Recreated SNN with CNN-inspired architecture following Diehl et al. 2015 STDP Paper in Tensorflow, built Astrocyte units in SNN for MNIST digit recognition

Working Experience

Datadog, Software Engineering Intern May 2019 - Sept 2019

- Developed microservice to aggregate millions of notifications processed per second for Slack & HipChat, worked with Kubernetes, Cassandra, and containerization / data wrangling tools

MakeSpace, Software Engineering Intern May 2018 - Sept 2018

- Developed Item Tagging Widget using Computer Vision and ML APIs, in-house ML model used to tag 1 million photographs of stored items, wrote a data analysis report on the volume and type of stored items using data visualization software (D3.js, etc.)

XLP Capital, Software Engineer March 2018 - April 2019

- Implemented data source integrations (loading and analysis) for Interactive Brokers, Bloomberg, CCXT, Quandl finance APIs
- Developed trading platform GUI in J, allowing for resizing, tabbing, Excel functionality and handling of 2GB sized data using JQT Framework

Teaching Experience

COMS 4771: Machine Learning, Columbia University, *Teaching Assistant* Summer 2021, Fall 2021

- Held weekly office hours, graded exams and homework

iLab Assistant, Rutgers University, *Tutor* Spring 2019, Fall 2019, Spring 2020

- Tutor at CS Lab for Data Structures, Algorithms, Computer Architecture, Operating Systems

CS First-Year Interest Group Seminar, Rutgers University, *Instructor* Fall 2018

- One of 80 undergraduates teaching an accredited college course, taught 10 activity-based lesson plans to a diverse group of 25 first-years as the sole instructor for Exploring Computer Science

Math 477: Probability Theory, Rutgers University, *Grader* Fall 2017

Leadership & Service

Undergraduate Student Alliance of Computer Scientists
Rutgers University, *Outreach Director / Vice President* Sept 2016 - May 2020

- Built a strong Rutgers tech community by establishing monthly speaker series with 10+ speakers reaching a 300+ student audience, acquired thousands of dollars of sponsorship for HackRU
- Personally mentored 5 students for career and academic advice, ran RUnite CS Ambassadors with 15 student developers, visiting 20+ middle and high schools, reaching 1500 students

Quantitative Finance Club, Rutgers University, *President* Sept 2017 - May 2020

- Orchestrated and taught 14 advanced technical presentations involving trading models, finance, Black-Scholes & computer science theory
- Mentored 100+ members regarding technical and practical knowledge for interviews and career search in addition to weekly mock interviews