Yunlan Li

yl4387@columbia.edu +86 18651113325 https://yunlanli.com yunlanli

Education

Columbia University in the City of New York Expected May 2023
B.S in Computer Science GPA: 4.03 / 4.33
Relevant Coursework:
Operating System, Database System, Data Structures in Java, Fundamentals of Computer Science, Computer Science Theory, Advanced Programming in Unix/C, Computational Linear Algebra, Ordinary Differential Equation

Skills

Programming Languages: Python, C, Java, JavaScript
Technologies: Git, Docker, React.js, AWS (S3, DynamoDB, EMR), Hadoop, Spark
Tools: Linux, Vim, VSCode, Project Jupyter, LaTex, Confluence, JIRA

Experience

PatSnap May 2021 – Aug 2021
Data Engineering Intern Suzhou, China

- Developed an end-to-end CRF-based NER model for extracting nucleotide and protein sequences from patent text
- Pinpointed bugs in data platform's database synchronization program and synchronized billions of data using Apache Spark
- Implemented antibody CDR matching algorithms with >90% accuracy and recall against known CDR databases (SAbDab and abYsis) to allow IP strategists view CDR regions under 5 different numbering schemes and CDR definitions
- Extracted sequence metadata from EPO sequence listings from 1989 to 2018

Internet Real-time Lab @ Columbia University Jan 2021 – Jun 2021
Research Assistant New York, US

- Designed and implemented an architecture for the MCV testbed to allow researchers run Jupyter Notebooks on experiment data stored on remote servers directly from their web browsers, using JupyterHub and Docker
- Published on PyPI NISTAuthenticator – a HTTP header based JupyterHub authenticator

Spectator Publishing Company Aug 2019 - present
Senior Web Developer New York, US

- Migrated the landing page of columbiaspectator.com from JSP to React.js on Arch XP
- Implemented the user authentication and management system for Clubs@CU with Firebase

Projects

Linux Kernel v5.4 Hacking Jan 2021 – Apr 2021
In a team of 3, implemented a weighted round robin scheduler, an in-memory filesystem and miscellaneous system calls for exposing process page tables in user space, tracing state changes of processes, and traversing process trees, all in the Linux v5.4 kernel for the x86 architecture.

Event @ CU Jan 2021 – Apr 2021
An Event Management System for clubs and students at Columbia University, built with Python Flask backed by a PostgreSQL database.

Sorting Visualization May 2020 – Aug 2020
A website that visualizes various sorting algorithms built with the Canvas API and React.js.