HARSHIT JAIN

+1 (582) 203-9755 ◇ harshitj.cs@gmail.com ◇ linkedin.com/in/harshitjain17 ◇ github.com/harshitjain17

EDUCATION

The Pennsylvania State University, University Park PA Expected Graduation: Dec 2024
Bachelor of Science in Computer Science
GPA: 3.7/4.0 | Dean's List (5/5) | Technical Coach @ CodePath | AlgoPSU Captain @ ACM | Recommendations | Projects | RA
Relevant Coursework: Data Structures and Algorithms, Math of Machine Learning, Operating Systems Design, Systems Programming, Supervised Machine Learning, Advanced Learning Algorithms (Deep Learning), Generative AI with LLMs, Theory of Computation, CodePath: (Intermediate+Advanced) Software Engineering, Database Management Systems, Financial Engineering

TECHNICAL SKILLS

Programming Languages: Python, C/C++, JavaScript, Java, HTML/CSS, MATLAB, Verilog, Assembly (64/32-bit x86)
Frameworks & Tools: TensorFlow, Scikit-Learn, NumPy, Node.js, React.js, Next.js, GraphQL, RESTful APIs, LaTeX, Git
Softwares: AWS, GCP, LLMs, MS SQL Server/MySQL/PostgreSQL, Linux/UNIX, SonarQube, Postman, Bitbucket, JIRA

WORK EXPERIENCE

Software Engineer Intern Aug 2024 - Present
Materials Research Institute (2DCC-MIP Team), Penn State University University Park, PA

Software Engineer Intern May 2024 - Aug 2024
Hughes Network Systems, LLC (Aeronautical Team) Germantown, MD

- Automated AWS-to-BigQuery pipeline with strict row-level validation & GCS audit trails, cutting manual processes by 95%
- Developed Python frameworks and BigQuery SQL views for monthly KPI evaluation of aircraft and fleet performance, processing data from 61k+ Delta Airlines flights with O(1) time complexity and sub-10-second execution
- Designed an automated script for flight performance billing and dynamic performance adjustments for accurate invoicing
- Technologies Used: Python, Google Cloud Platform (GCP: Cloud Functions, BigQuery, Pub/Sub), SQL, Agile/Scrum

Machine Learning Engineer Intern Jan 2024 - May 2024
Materials Research Institute (2DCC-MIP Team), Penn State University University Park, PA

- Integrated GPT-4, Jurassic2, Llama2 LLMs into a chatbot framework (LangChain) for answer retrieval via an RAG model
- Built a serverless Next.js app with AWS integration, optimizing RESTful APIs for high scalability and CI/CD workflows
- Developed a fully automated AWS Lambda pipeline for video processing, including transcription and AI-generated metadata
- Configured AWS Rekognition to boost speaker identification, improve transcription, and automate chapter generation
- Technologies Used: Python, LangChain, Amazon Web Services (AWS), Next.js, LLMs, Deep Learning Models

Software Engineer Co-op May 2023 - Dec 2023
VIAVI Solutions Inc. Germantown, MD

- Engineered a Python-based automated test suite for PNT instruments on Linux, in collaboration with a 6-member R&D team, to validate system performance and ensure strict adherence to the SCPI protocol
- Troubleshot PNT unit bugs, reducing by 55% and increasing code coverage by 30% (SonarQube) through C/C++ debugging
- Executed 35+ SCPI-driven long-term tests on core devices to ensure compliance with release-level quality standards
- Technologies Used: Python, C/C++, Linux, SCPI Protocol, Bitbucket, Confluence, SonarQube, Git, JIRA, Agile/Scrum

Software Engineer Intern May 2022 - May 2023
Materials Research Institute (2DCC-MIP Team), Penn State University University Park, PA

- Engineered a scalable React.js app with 50+ user-facing features across 20+ components, and employed Jest for unit testing
- Integrated MS SQL Server to manage data from 500+ instruments, and organized it across 18+ BCNF-normalized tables
- Built a Python script for automated data retrieval and integration with a cloud-based data warehouse using ETL processes
- Research: Re-architected a Python library for the Raman Fitting Model (try here) to implement precise spectra deconvolution algorithms with automated data preprocessing, fitting, and export functions, reducing analysis time by 40% (validated)
- Technologies Used: Python, JavaScript, React.js, Node.js, MS SQL Server, RESTful APIs, Git, HTML/CSS, JIRA

PROJECTS

Dynamic Memory Allocator [C/C++] Jan 2024 - Feb 2024

- Designed custom malloc, free, realloc; segregated free lists and footer optimization to improve memory management
- Achieved a utilization score of 69% and benchmark throughput at 100% across diverse computing environments

mdadm Linear Device [C/C++, Linux] Feb 2023 - May 2023

- Configured 16 disks of size 64 KB as a 1MB linear device, providing users with a unified address space for data access
- Implemented mount/unmount operations to the linear device, mitigating potential data loss and system crashes
- Designed the read/write functions and engineered data caching solutions reducing I/O wait time by 60%