HARSHIT JAIN

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

EDUCATION

The Pennsylvania State University, University Park PA Graduation: Dec 2024
Bachelor of Science in Computer Science
GPA: 3.7/4.0 | Teaching Assistant @CodePath | AlgoPSU Captain @ACM | Recommendations | Projects | Resident Assistant
Course: Data Structures & Algorithms, Operating Systems, Object-Oriented Programming, Linear Programs, Database Systems
Certifications: Microsoft: Azure AI Engineer Associate, AWS, Machine Learning, CodePath: Advanced Software Engineering

TECHNICAL SKILLS

Programming Languages: Python, Java, C/C++, JavaScript/TypeScript, SQL, Shell scripting
Frameworks: Streamlit, Node.js, React.js, Next.js, LangChain, TensorFlow, Scikit-Learn, GraphQL, RESTful APIs, Git
Softwares: AWS, GCP, Azure, MS SQL Server/PostgreSQL, Linux/UNIX, Docker, SonarQube, Postman, Swagger, JIRA

WORK EXPERIENCE

Software Engineer Jan 2025 - Present
Hughes Network Systems, LLC (Aeronautical Team) Washington, DC

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

- Engineered an AI-powered video platform (Next.js, AWS) enabling 10+ U.S. institutes to streamline research video analysis
- Reduced video processing latency by 20% by optimizing GraphQL APIs and integrating DynamoDB for faster video retrieval
- Designed a secure authentication flow (AWS Cognito, API Gateway) to enable controlled CRUD access for research datasets
- Deployed a Python AI pipeline (LangChain, Pinecone) to enable intelligent search across 1k+ videos using vector embeddings
- Technologies Used: Python, Amazon Web Services (AWS), LangChain, Pinecone, Next.js, GraphQL, REST APIs

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

- Automated an AWS-to-BigQuery pipeline (Python) with strict row-level validation and reduced manual intervention by 95%
- Delivered SQL views and a Python-based analytics pipeline to power Delta Airlines' flight insights for 61k+ flights monthly
- Built a Python-based automated invoicing system (GCP) to drive multi-stream revenue capture for flight operations at scale
- Technologies Used: Python, Google Cloud Platform (GCP: Cloud Functions, BigQuery, Pub/Sub), SQL, Agile/Scrum

Software Engineer Intern Jan 2024 - May 2024
Materials Research Institute (2DCC-MIP Team) University Park, PA

- Delivered an AI-driven video pipeline using AWS Lambda to automate transcription and tagging for 1,000+ hours of content
- Integrated LLMs (GPT-4, Claude 3) into a LangChain chatbot to enable RAG-based retrieval for 800+ queries monthly
- Built a video classification system with AWS Rekognition to automate speaker identification and cut manual tagging by 75%
- Developed an automated CI/CD pipeline for serverless research apps, ensuring rapid deployments and 99.9% uptime reliability
- Technologies Used: Python, Amazon Web Services (AWS), LangChain, Next.js, RESTful APIs, LLMs

Software Engineer Co-op May 2023 - Dec 2023
VIAVI Solutions Inc. (PNT Team) Washington, DC

- Built a distributed Python/C++ test framework for PNT products to cut testing time by 70% and speed up market delivery
- Redesigned a multi-tiered system to microservices architecture to resolve SCPI bottlenecks and boost response time by 40%
- Executed 35+ SCPI compliance tests to reduce product validation time from 2 weeks to 3 days for critical product releases
- 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) University Park, PA

- Built scalable React.js application with 50+ features and Jest test coverage to accelerate data analysis for 200+ scientists
- Integrated MS SQL Server to manage data from 900+ instruments and structured it across 18+ BCNF-normalized tables
- Built a Python-based ETL pipeline to automate data extraction, transform logs, and sync research data to a cloud warehouse
- Research: Re-architected Python library for Raman Fitting Model (try here) to implement spectra deconvolution, automate data preprocessing, and optimize export functions, reducing analysis time by 40% and improving model accuracy
- 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