

ADITI JOSHI

+1 (213) 910-6690 | aajoshi@usc.edu | <https://www.linkedin.com/in/aditi-anant-joshi/> | <https://github.com/aditi-joshi-usc> | LA, CA

SUMMARY

Software Engineer with 1.5 years of experience in fintech and healthcare, building scalable data pipelines and cloud-based platforms using AWS and Snowflake. Strong in Python, Java, and C++, with a foundation in data engineering and real-time processing systems.

EDUCATION

Master's, Applied Data Science | University of Southern California | Aug 2023 – May 2025 | Los Angeles, CA | GPA: 3.66

- Relevant Coursework: Data Mining, AI, Machine Learning, Data Management, UX Design, Information Retrieval

Bachelor's, Computer Engineering | Cummins College of Engineering | Aug 2018 – May 2022 | Pune, India | GPA: 3.78

- Relevant Coursework: Operating Systems, DSA, ML, Big Data, Software Architecture, DB, Programming Languages

SKILLS

- Languages & Scripting: Python, Java, C++, SQL, JavaScript
- Cloud & Big Data: Apache Spark, Hadoop, AWS (Glue, Lambda, S3, EC2), Snowflake, Kafka (familiar)
- Databases: MySQL, PostgreSQL, Oracle, MongoDB
- Machine Learning: XGBoost, Scikit-learn, Spark MLlib, model tuning, feature engineering, evaluation metrics
- Frameworks: React, Angular, Node.js, Django, Streamlit
- DevOps & Workflow: Docker, CI/CD, Airflow (familiar)
- Version Control & Tools: Git, JIRA, Bitbucket, Confluence, Agile, Scrum

PROFESSIONAL EXPERIENCE

Graduate Course Assistant (CSCI-401) | University of Southern California | Aug 2024 – Present | Los Angeles, CA

- Led Scrum for 10 capstone teams, boosting project outcomes by 40% through strategic mentoring, stakeholder coordination, and guidance in Python, SQL, cloud tech, and software development.
- Tripled team productivity by implementing KPIs, 360 peer reviews, and individual mentoring; ensured cross-functional alignment and timely delivery of 10 industry-sponsored projects.

Data Engineering Intern | Sanofi | Jun – Aug 2024 | Bridgewater, NJ

- Built a scalable ETL pipeline using AWS Glue & Athena, boosting data accessibility by 75% and ensuring high-performance SQL execution on Snowflake.
- Enhanced query accuracy by 40% through model tuning and prompt engineering; implemented data governance guardrails to ensure data security and access control.

Software Engineer (Technology Analyst) | Citigroup | Jul 2022 – Jul 2023 | Pune, India

- Optimized and migrated a high-traffic financial reporting pipeline to a scalable microservices architecture using Java (Spring Boot) and SQL, reducing processing latency by 30% and improving platform stability.
- Designed, implemented, and tested a full-stack monitoring application that automated 30% of manual processes and enabled real-time insights for business analysts, accelerating decision-making and operational efficiency.
- Developed robust QA strategies and reusable test cases; built dashboards to track pipeline health and job execution metrics, enhancing system observability and team performance. Recognized by leadership for impact and proactive ownership.

Software Engineering Intern (Summer Analyst) | Citigroup | May – Jul 2021 | Pune, India

- Built and tested a production-grade JavaFX + SQL desktop tool for financial reconciliation using **batch processing**, boosting data validation efficiency by 50% and ensuring 100% data accuracy.
- Demonstrated strong communication skills by presenting to 30+ senior executives, earning a full-time offer for showcasing the tool's business impact and operational value.

TECHNICAL PROJECTS

File System-Based Relational and NoSQL DB | Team Lead & Software Architect | Sep–Dec 2023 | Los Angeles, CA

- Designed and implemented a Python-based file system supporting both relational and NoSQL data models, boosting retrieval speed by 40% through OOP principles, operating systems knowledge, and custom storage logic.
- Enabled scalable, unstructured data handling with a high-performance NoSQL engine, improving accessibility and processing speed by 30%.

Hybrid Recommendation System Using XGBoost and Spark RDD | Individual Contributor | Nov–Dec 2024 | Los Angeles, CA

- **Engineered** a Spark-based data pipeline to process large-scale Yelp data, leveraging RDDs for distributed feature engineering and batch transformations.
- Improved model performance through hyperparameter tuning and error analysis, achieving lower RMSE for personalized recommendations.

IBM Differential Privacy on Genomic Datasets | Project Intern | Sep 2021 – May 2022 | Pune, India

- Developed a secure disease prediction portal using **React (frontend)** and **MongoDB (database)** and **machine learning** to demonstrate privacy-preserving AI in healthcare using genomic data.
- Applied **differential privacy** techniques to enable secure model training and visualization of trade-offs between privacy and accuracy; recognized among 250+ global teams in national and international competitions.