

ADITI JOSHI

Los Angeles, CA | (213) 910-6690 | aditi.jo2000@gmail.com | <https://www.linkedin.com/in/aditi-anant-joshi/>

SUMMARY

Passionate data science professional with hands-on experience as a Data Engineering Intern at Sanofi and Technology Analyst (full stack developer) at Citigroup, now pursuing a master's in applied data science at USC. Enthusiastic about **problem-solving** and building AI-driven solutions, and scalable data systems, leveraging machine learning, cloud computing, and Python to tackle real-world challenges in fast-paced environments of both **fintech and healthcare**. Possesses strong **communication skills**, enabling effective collaboration and stakeholder engagement.

EDUCATION

University of Southern California Los Angeles, California
Master's in Applied Data Science: GPA-3.65/4 August 2023-May 2025

Courses: Foundations and Applications of Data Mining, Data Visualization, Foundation to Artificial Intelligence, Machine Learning for Data Science, Foundation to Data Management, Data Science at Scale

Cummins College of Engineering, Pune (Pune University) Pune, India
Bachelor of Technology: GPA-3.78/4 August 2018-May 2022

Coursework: Artificial Intelligence and Machine Learning, Statistics, Big Data Analytics, Data Structures and Algorithms

SKILLS AND CERTIFICATIONS

- **Python**, Java, Scala, JavaScript, HTML, CSS, Angular, ReactJS, Spring Boot, Django, JavaFX, Machine Learning, Apache Spark
- SQL, Oracle, MongoDB, Firebase, Snowflake, AWS Bedrock, Lambda, Glue, Athena, S3, Power BI, Tableau
- Git, Bitbucket, JIRA, Agile, Data Science Specialization (University of Michigan) - Coursera (5 Courses)

WORK EXPERIENCE

Sanofi Bridgewater, New Jersey
Data Engineering Intern, R&D digital products (Data) team June 2024-August 2024

- Designed wireframes and prototyped a **Generative AI** Q&A application to enhance comprehension of complex datasets stored in Snowflake to achieve efficient conversion of natural language user questions into contextualized SQL queries
- Executed experiments on Large Language Models (**LLMs**) leveraging **AWS** cloud computing services Bedrock, Athena, S3 and Glue enhancing model performance through advanced prompt engineering techniques

Citigroup Pune, India
Full stack developer July 2022-July 2023

- Collaborated with the Foreign Exchange and Local Markets Application Stability team and worked on Java, SQL, and Angular, CI/CD development for migrating a legacy application code to micro-services in Agile life cycle and ceremonies
- Built a stand-alone report monitoring tool - Leveraged by business analysts to automate 30% of manual work

Citigroup Pune, India
Summer Analyst May 2021-July 2021

- Engineered a new desktop reconciliation tool, to compare existing business logic to new logic being adopted; application is in production and operated by business analysts for integrity testing of data with increased efficiency by 50%
- Accessed financial data from three **APIs**, processing it in batches, displaying results for each batch, and downloading results. Demonstrated project to over 30 senior members of the organization and received a full-time job offer

LEADERSHIP AND INVOLVEMENT

- **Course Assistant** for the Design and Construction of Large Software Systems under Prof. Victor Adamchik at USC. Mentored five teams, facilitated stakeholder communication, and led weekly **Agile** scrum meetings for project updates and guidance
- Won second prize in IBM hackathon organized Innovation 2022- college technical fest, third place in international project competition, fourth place in a national project competition
- Mentored at the AthenaHacks Hackathon 2024, member of the Society of Women Engineers (SWE)
- Ranked 1st among 200 students in Principles of Programming Languages, earning a prestigious academic scholarship

PROJECTS

Research Project - Using ML for Early Prediction of Conduct April 2024-September 2024

- Leading advanced data preprocessing on the ABCD MRI dataset (11,000+ adolescents) and ABIDE dataset with Prof. Anita Penkova at USC, developing cutting-edge ML models for early diagnosis and intervention in conduct disorder

File System-Based Relational and a non-relational database September 2023-December 2023

- Spearheaded development of an efficient file system-based relational database using Python and **MySQL**
- Engineered a high-performance **NoSQL** system with MongoDB and Python, elevating data management to the next level

IBM Differential Privacy on Genomic Datasets September 2021-May 2022

- Headed a genetic data privacy research project on differential privacy. Developed a user-friendly web portal with Python and Django. Implemented a disease prediction Machine Learning model with and without **differential privacy**
- Won awards in state-level project competitions, and **international project competitions** among 250+ Global teams