

Naman Lalit

📞 732-554-9961

✉ naman.lalit@gmail.com

🌐 [linkedin.com/in/namanlalit](https://www.linkedin.com/in/namanlalit)

🐙 github.com/namanlalitnyu

Education

New York University, Courant

Master of Science in Computer Science (GPA: 3.95 / 4.00)

Sept 2023 - May 2025

New York City, New York

- **Graduate Grading Assistant:** Operating Systems and Multicore Processors
- **Relevant Coursework:** Deep Learning, Multicore Processors, Predictive Analytics, Cloud & Machine Learning

National Institute of Technology Hamirpur

Bachelor of Engineering in Computer Science (GPA: 4.00 / 4.00)

Jul 2017 - Jun 2021

Hamirpur, Himachal Pradesh

- **Relevant Coursework:** Data Structures & Algorithms, Computer Networks, Prob. & Statistics, Software Engineering

Technical Skills

Languages: C, C++, Java, Python, Golang, JavaScript, Rust

Databases: PostgreSQL, MongoDB, AWS DynamoDB, Google BigQuery

Software: Node.js, React.js, Typescript, React-Native, Express, Flask, Taipy, Spring, TensorFlow, PyTorch, LangChain, HTML

Cloud & Devops: AWS, GCP, CI/CD, Docker, Kubernetes, Git, Linux

Others: Gen AI, LLM (Large Language Models), Redis, Salesforce, SnapAR, Lenstudio

Experience

Skillz

Jan 2025 – Present

Software Engineer Co-op

Las Vegas, U.S.A

- Working in the Monetization team and building Fast Withdrawals and Deposits using ACH payment method by leveraging Golang (Backend) and React Native (Mobile SDK), leading to an increase in revenue of \$5 million per year.
- Orchestrated the successful delivery of an important feature of showcasing different Advertisements to game players, using Typescript, and ReactNative, positively impacting the user conversion rate by 60%.

Salesforce

Feb 2022 – Aug 2023

Software Engineer (AMTS)

Bangalore, Karnataka

- Pioneered the development of the Benefits Management Portal in the Public Sector Cloud using Salesforce's Omnistudio, Java, JavaScript, and React.js, generating over \$20 million in Annual Contract Value (ACV).
- Collaborated with cross-functional teams and delivered a feature of automated PDF uploads on the Salesforce platform on license activation using Java, Spring, and JUnit, increasing efficiency for over 100 developers.
- Led the development of the Guided Setup product for more than 2 releases involving shipping multiple new features, fixing existing bugs, and collaborating with the product management team to develop product roadmaps.
- Spearheaded feature discussions, successfully shipped multiple key features, authored comprehensive documentation, and ensured quality with rigorous unit and functional testing.

Infoedge

Jul 2021 – Feb 2022

Software Engineer

New Delhi, Delhi

- Led the back-end development of Report's product in Python and Flask, used across the organization by more than 80 members.
- Boosted revenue by 50% by developing and deploying interactive dashboards and enhanced features using HTML, CSS, Node.js, React.js (frontend) and Flask, Python, MySQL (backend).
- Automated the processing of over 5000 daily AWS SQS events using a Python script, storing data in a MongoDB database.
- Integrated AWS services including SES, DynamoDB, and S3 into existing products using Node.js and Express.js.
- Mentored interns and led company-wide onboarding sessions, gave product demonstrations to new hires and created comprehensive documentation.

CRED

Jan 2021 – Jul 2021

Backend Engineer Intern

Bangalore, Karnataka

- Reduced Redis memory usage by 50% by designing a Python script to delete unused game tickets.
- Streamlined and managed the migration of 10 REST APIs from one service to another while overseeing the entire end-to-end deployment process using Java/Spring.
- Implemented SQL queries on Google Big Query and reduced REST API query response time from 200 ms to 20 ms by using Celery for managing asynchronous tasks.

Academic Research

AI and Predictive Analytics Lab

AI & ML Researcher

Aug 2024 – Present

New York City, New York

- Built an ensemble Retrieval Augmented Generation framework on LENR (Low Nuclear Energy), leveraging over 3000 preprocessed articles and integrating outputs from four LLMs (ChatGPT, Gemini, Claude, Llama 3). [Link](#)
- Led NYU's AI and Predictive Analytics team on the Politics Project, utilizing NLP and predictive analytics to forecast the 2024 presidential election, featured on CNN News, and published at NYU. [Link](#)

NYU Secure Systems Lab

Research Assistant

May 2024 – Aug 2024

New York City, New York

- Developed core functionalities like system calls and integrated Rust with the C++ compiler of an operating system (RustPosix) in Rust and C++, maintaining 100% code coverage and improving system reliability and performance.
- Designed the CI/CD workflow backed by GitHub Actions and Docker, reducing deployment time by 80% and increasing efficiency.

Projects

TickerTeller - Stock Price prediction platform | LLM, Python, Tensorflow, Taipy, MongoDB | [Link](#)

- Engineered a platform using Taipy and Python, integrating an ensemble machine learning model, leveraging LSTM outputs and sentiment analysis of Wall Street Journal articles, to predict stock prices for 20 companies.
- Leveraged OpenAI ChatGPT-3.5, Lang Chain, and Prompt Engineering to generate news article sentiments, evidence extraction, and LLM explanations of more than 4000 news articles.
- Designed a Python script to retrieve and store the companies' stock prices for 2 years using Yahoo Finance's Vantage API.

Stock.ai - Retrieval Augmented Generation Application | RAG, Python, AngularJS, Javascript, Flask, Docker, GCP | [Link](#)

- Engineered a RAG application tailored for academic research on a dataset of 1,000 stock prediction research articles with technologies including React.js, LLM, Prompt Engineering, and Docker.
- Developed an intuitive query-processing system using Flask on Google Kubernetes Engine; ensured users received prompt access to the most pertinent academic resources while leveraging ChromaDB for enhanced content relevance.

RapidEdit - Real-time instruction-based image editing | Python, SAM CLIP, LCM Pipeline | [Link](#)

- Developed an advanced image editing pipeline integrating the Segment Anything Model (SAM) for precise mask generation, enabling dynamic image segmentation and object isolation based on user prompts.
- Engineered a robust framework leveraging SDXL and LCM.LoRA weights to seamlessly apply targeted modifications to segmented regions, delivering high-quality, customized output images.

Topic Modeling - Research Articles | Python, K-Means, Principal Component Analysis, OpenAI | [Link](#)

- Employed K-Means clustering to organize 200+ research papers into thematic groups utilizing advanced NLP methods and embeddings, enhancing understanding through Prompt Engineering and OpenAI API.

Rust and C++ - Qualitative Analysis of Multi-Threaded Languages | Rust, C++, GPU | [Link](#)

- Conducted a thorough qualitative analysis using the NVIDIA V100 GPU on OpenMP (C++) and Rust across 4 benchmark programs, evaluating performance based on scalability, programmability, speedup, and efficiency metrics.

Achievements

- Awarded Best Overall Hack, Best Healthcare Project, and two prizes in the Snap and Google challenge at the Columbia DivHacks '24 hackathon among 300 participants. [Link](#)
- Published a research paper, Qualitative Analysis of Text Summarization Techniques, in Hindawi Journal that performed analysis of 5 text summarization techniques using Python and Machine Learning with over 4000+ views and around 2500+ downloads. [Link](#)
- Secured 30th rank out of 200 teams in the regional competition at ACM ICPC (Competitive Coding) in IIT Kharagpur.