

Pratyush Agarwal

hi@pratyushagarwal.com | +1 (217) 305-3172 | [LinkedIn](#) | [GitHub](#) | [pratyushagarwal.com](#)

TECHNICAL SKILLS

Programming Languages: Java, Python, TypeScript, JavaScript, SQL

Frontend: React.js, Next.js, HTML, CSS

Backend and APIs: Node.js, FastAPI, REST APIs, SQLAlchemy, Postman

Databases: PostgreSQL, MySQL, SQL Server, DynamoDB, Redis, BigQuery

Cloud and DevOps: AWS (EC2, S3, Lambda, API Gateway, IAM), Azure, GCP, Docker, Kubernetes, Git/GitHub, CI/CD, Microservices, Agile/Scrum

CS Foundations: Data Structures, Algorithms, Object-Oriented Programming, System Design

AI: LLM Applications (Claude API, AWS Bedrock, Cursor), Prompt Engineering

PROJECTS

TickerSense | *React, Next.js, Node.js, Claude API, PostgreSQL*

[\[GitHub\]](#)

- Built a full-stack LLM platform with React, Next.js, and Node.js unifying SEC filings, fundamentals, and disclosures into one interface; placed 2nd at the Claude Builder Club Hackathon (UIUC).
- Developed a source-grounded conversational research assistant using the Claude API with exportable PDF reports.

TripTok | *FastAPI, PySpark, AWS, Python*

[\[GitHub\]](#)

- Built an end-to-end pipeline processing 25GB+ of NYC Taxi data with PySpark and AWS S3, serving ML models via a FastAPI MCP server with AWS Bedrock and Strands Agents.
- Deployed XGBoost, Random Forest, and LSTM models behind a conversational LLM analytics layer for real-time mobility insights.

Mnemo: Personal Knowledge Engine | *TypeScript, NestJS, Python, FastAPI, PostgreSQL, Docker*

[\[GitHub\]](#)

- Building a two-service "second brain" app: a TypeScript/NestJS API for notes, tasks, auth, and kanban, and a Python/FastAPI AI service communicating over an HTTP contract, orchestrated with Docker Compose.
- Architected with clean layering (thin routers, service logic, repository DB access) on Postgres and pgvector, evolving toward semantic search, GraphRAG over Neo4j, and an MCP-exposed agent.

EXPERIENCE

Kohler Co. [Digital Transformation Team]

Jun 2025 – May 2026

Data Engineer Intern

Champaign, IL

- Designed and implemented production ETL pipelines with Azure Data Factory and Databricks, ingesting SharePoint and on-prem SQL Server data into Parquet-based curated layers for downstream systems.
- Integrated on-prem operational databases with Azure cloud platforms and introduced Git-based version control to the team, improving pipeline governance across the org.

Apna [Trust and Safety Team]

Oct 2022 – Aug 2024

Data Engineer

Bangalore, India

- Built an API-driven user verification pipeline using Python, REST APIs, and Postman against Government of India identity services, halving processing time for ~1,200 users weekly.
- Optimized high-cost BigQuery workloads through partition-filtering and scope-limiting, cutting warehouse spend by ~\$2,500 monthly and improving query execution time by ~20%.

Google [Channel Sales Activation Team]

Oct 2021 – Apr 2022

Strategy and Analytics Intern

Singapore

- Engineered web scraping automation with Python, ChromeDriver, and Selenium, plus the database and supporting infrastructure to capture daily website metrics; extended the role from 1 to 7 months.
- Collaborated with developers to build scalable scripts and recurring data workflows supporting Pixel 6 market intelligence across US, UK, Germany, and APAC.

EDUCATION

University of Illinois at Urbana-Champaign

Aug 2024 – May 2026

Master of Science in Information Management; GPA: 3.75/4.0

Champaign, IL

Manipal University Jaipur

Jul 2018 – Jul 2022

Bachelor of Technology in Computer and Communication Engineering; GPA: 8.53/10.0

Jaipur, India