Shivani Singh

New Delhi, India | +91 9818155639 | shsingh844@gmail.com | LinkedIn | GitHub

SUMMARY

Full-stack software engineer with 3+ years of experience building production AI/ML systems, NLP applications, and scalable cloud infrastructure. Proven expertise in LLM integration (Claude API, GPT-4, LangChain), full-stack web development, and secure on-premise deployment. Currently pursuing doctoral research in predictive analytics with strong foundation in statistical analysis and data-driven research methodologies. Demonstrated ability to deliver end-to-end solutions from architecture design to production deployment with focus on system reliability, security, and performance optimisation.

SKILLS

AI/ML & NLP: LLM Integration & Deployment, Production ML Systems, NLP/Text Analysis, Claude API, OpenAI GPT-4, LangChain, RAG Architecture, Model Serving & Orchestration, Prompt Engineering, TensorFlow, PyTorch, scikit-learn, Statistical Analysis

Development & Architecture: Python, Full-Stack Web Development, React, TypeScript, JavaScript, Flask, Node.js, Ruby, RESTful APIs, WebRTC, Microservices, Event-Driven Architecture, API Integration, Version Control (Git)

Data Engineering: SQL, PostgreSQL, ETL/ELT Pipelines, R, Statistical Computing, Data Visualisation, Airflow, Databricks, BigQuery, Tableau, Power BI

Infrastructure & DevOps: Docker, Kubernetes, CI/CD Pipelines, AWS, Azure, Terraform, Application Server Deployment, On-Premise Deployment, Prometheus, Grafana, System Monitoring

Security & Compliance: Secure Coding Practices, Data Confidentiality, Compliance-Driven Environments, NDA/Sensitive Data Handling

EXPERIENCE

Independent Software Consultant + DBA Scholar | Remote   Nov 2022 - Present

- Built real-time weather balloon analytics platform with Claude API integration, implementing multi-modal prompt engineering for atmospheric pattern analysis and intelligent caching for 24-hour continuous operation, dramatically reducing telemetry analysis time while enabling data-driven operational decisions
- Developed AI-powered fashion platform using GPT-4 and LangChain with PostgreSQL behavioural tracking and geospatial filtering, achieving strong user engagement through personalised recommendations
- Architected production cloud infrastructure with automated CI/CD pipelines for React/Flask applications, implementing monitoring and database optimisation achieving sub-second query response times
- Conducting doctoral research analysing reputational contagion from 2,180 major bank failures on small regional banks using event study methodology and data triangulation, developing regulatory intervention framework and early-warning indicators to prevent secondary collapses

Mixhalo | San Francisco, California, USA   Apr 2022 - Nov 2022
Full Stack Engineer

- Led 0→1 development of real-time translation prototype that became company's flagship monetisation product, directly influencing future investor presentations and product roadmap pivot toward translation services
- Designed scalable WebRTC architecture through systematic evaluation of multiple frameworks, reducing future product R&D cycles and establishing reusable patterns for all streaming features
- Reduced infrastructure costs by 67%, architecting Kubernetes-based deployment system that improved deployment frequency from weekly to daily, eliminating DevOps bottlenecks for engineering team
- Built end-to-end observability platform capturing user interactions and system metrics, dramatically reducing incident response time and enabling data-driven product decisions that shaped company's roadmap

Extreme Networks | San Jose, California, USA   Mar 2020 - Apr 2022
Software Engineer
- Architected company-wide data platform consolidating multiple siloed systems into unified warehouse processing millions of events monthly, dramatically reducing executive reporting cycles across business units
- Delivered self-service analytics platform with multiple Tableau/Power BI dashboards heavily used across organisation, dramatically reducing ad-hoc reporting requests and freeing capacity for strategic initiatives
- Delivered multiple production features for ExtremeIQ cloud networking platform managing thousands of enterprise devices, gaining deep expertise in TCP/IP, DNS, VPNs, which accelerated customer issue resolution
- Established comprehensive testing framework integrated into CI/CD pipeline, reducing production incidents by 45% and deployment rollbacks significantly, enabling team to accelerate feature velocity

Frontend Engineering Intern
- Led UI redesign that significantly increased customer satisfaction through user research and iterative testing, becoming key differentiator in sales demos and customer retention
- Created reusable component library using Dojo framework, significantly reducing code duplication and cutting feature development time across multiple product teams

EDUCATION & RESEARCH

Westcliff University | Irvine, California, USA   Oct 2022 - Present
Doctor of Business Administration (Business Intelligence & Data Analytics) | **Remote**

- Developing ML-based financial contagion model analysing historical bank failures with ensemble methods, achieving strong predictive accuracy; research progressing toward peer-reviewed publication

Boston University | Boston, Massachusetts, USA   Sep 2018 - Jan 2020
M.S., Computer Information Systems (Database Management & Business Intelligence)

- GPA: 3.60

Indira Gandhi Delhi Technical University for Women | New Delhi, India   Aug 2014 - Jun 2018
Bachelor of Technology, Electronics & Communication Engineering

- CPI: 64.37

NOTABLE PROJECTS

Windborne Constellation Tracker | Live Demo
- Production AI analytics platform processing real-time telemetry data with Claude API integration
- Implemented intelligent caching, multi-modal analysis, and automated reporting reducing manual analysis time
- Technologies: Python, Flask, React, PostgreSQL, Claude API, Docker

Sustainable Fashion Analytics Platform | Live Demo
- AI-powered platform with GPT-4-driven personalised recommendations and geospatial brand matching
- Achieved 78% user engagement through behavioural tracking and intelligent recommendation engine
- Technologies: Python, GPT-4, LangChain, React, PostgreSQL, Geospatial APIs