

Shivani Singh

USA | +1 571 337 0488 | shsingh844@gmail.com | [linkedin.com/in/shivani-singh13](https://www.linkedin.com/in/shivani-singh13) | github.com/shsingh844 | [Portfolio](#)

SUMMARY

Software Engineer with 3+ years building production applications and scalable infrastructure. AI-native full-stack developer specializing in Python, LLM integration, and real-time data systems. Expert in end-to-end AI product development, from React frontends to distributed ML backends, with hands-on experience in production AI system deployment and optimization. Currently pursuing doctoral research in predictive analytics, combining academic rigor with production engineering excellence.

Available for full-time AI engineering opportunities with immediate relocation. Currently in California, USA, with US work authorisation. Actively seeking UK Skilled Worker Visa sponsorship for London-based AI/fintech roles.

SKILLS

AI & ML: LLM Integration & Deployment, Agent Systems, Model Serving & Orchestration, Prompt Engineering, Claude API, OpenAI, LangChain, RAG, Context Handling, Embedding Systems, Statistical Analysis, Predictive Analytics

Full-Stack Development: Python, React, TypeScript, JavaScript, Flask, NodeJS, RESTful APIs, Real-time Systems, WebRTC, Frontend Libraries, Responsive Design, User Experience

Data & Infrastructure: R, SQL, PostgreSQL, ETL/ELT Pipelines, Airflow, Real-time Processing, Data Versioning, Docker, Kubernetes, CI/CD, AWS, GCP, Terraform, Monitoring (Prometheus/Grafana), Tableau, Power BI, BigQuery

Research & Analytics: Statistical Modelling, Business Intelligence, Geospatial Analysis, Academic Research

WORK EXPERIENCE

Mixhalo | San Francisco, CA

Full Stack Engineer

Apr 2022 - Nov 2022

- Architected and prototyped real-time audio translation system (Mixhalo Translate), building foundational demo to demonstrate technical feasibility and market vision to stakeholders and investors
- Conducted comprehensive technical research and proof-of-concept development for WebRTC-based audio processing pipeline, establishing scalable architecture patterns and integration strategies for ML translation services in live events
- Built production-ready infrastructure components using Kubernetes and Docker, improving cost-efficiency by 67% and creating deployment frameworks that enabled seamless integration of real-time processing tools and analytics dashboards
- Designed data pipeline architecture combining audio streaming, metadata processing, and user interaction patterns, laying technical foundation for personalised content delivery, and establishing monitoring systems (Loki, Prometheus) for operational insights

Extreme Networks | San Jose, CA

Software Applications Engineer

Mar 2021 - Apr 2022

- Architected data consolidation systems integrating multiple sources with SQL/NoSQL databases, building scalable analytics infrastructure supporting real-time decision-making across teams
- Developed production-grade visualisation dashboards in Power BI and Tableau, creating a self-service analytics platform used by multiple stakeholders for KPI tracking and business intelligence
- Implemented comprehensive testing framework, reducing production incidents by 45%, establishing quality standards and deployment practices for rapid iteration cycles
- Mentored engineering team members on network management and software architecture, demonstrating technical leadership and knowledge transfer capabilities that are essential for founding team environments

Extreme Networks | San Jose, CA

Frontend Engineering Intern

Mar 2020 - Mar 2021

- Collaborated with the UX designer to revamp the product's user interface, enhancing user experience and visual appeal
- Developed frontend libraries using Dojo and Honeycomb frameworks, improving code reusability and efficiency

Questrom School of Business, Boston University | Boston, MA

Teaching Assistant

Jul 2019 – Oct 2019

- Taught data analytics and business intelligence using R, Python, and SQL, leveraging BigQuery for analysis & visualisation and foundational Data Science concepts, including algorithms, statistical analysis, and predictive modelling
- Mentored students in research methodology and proposal development, resulting in improved proposal quality and increased approval rates

Disability & Access Services, Boston University | Boston, MA

Assistive Technology Consultant

Sep 2018 – Jan 2020

- Analysed concussion data using R, Scala, and Python, reducing incorrect assistive tool distributions by 97% through predictive modelling and data-driven decision making
- Automated student profile display system with HIPAA-compliant PII redaction/encoding practices, streamlining accommodations and reducing data redundancy by 80%
- Launched "Library Scan Project," optimising accommodation request workflows and reducing service request processing times by 50%

RESEARCH

Westcliff University | Doctoral Research

Oct 2022 - Aug 2026 (Expected)

Spearheading a quantitative research study tackling the issue of smaller banks facing reputational contagion when a larger bank collapses. The robust conceptual framework utilises advanced machine learning models to analyse multimodal data with high precision, thereby establishing a strong foundation for future predictive analysis.

NOTABLE PROJECTS

Windborne Constellation Tracker

Live Demo: <https://windborne-dashboard.onrender.com>

- Built real-time analytics platform with Claude API integration for intelligent balloon trajectory analysis and operational insights
- Developed context-aware prompt engineering system that processes multimodal data (telemetry, weather, historical patterns) for actionable insights
- Created robust data processing pipeline with intelligent caching and error handling for 24-hour continuous operation

Sustainable Fashion Analytics Platform

Live Demo: <https://sustainability-fashion-tracker.onrender.com>

- Built production AI recommendation system using OpenAI GPT-4 and LangChain for personalised sustainability insights based on user behaviour patterns and interaction history
- Implemented intelligent prompt engineering strategies for context-aware recommendations, achieving strong user engagement through context-aware AI-generated recommendations
- Developed PostgreSQL-backed user interaction tracking system for continuous learning and model improvement with real-time context composition
- Created browser-native analytics platform with interactive visualisations and location-based intelligent filtering using geospatial algorithms

EDUCATION

Westcliff University, Irvine, CA

Oct 2022 - Aug 2026 (Expected)

Doctor of Business Administration (concentration: Business Intelligence & Data Analytics) | **Online**

Boston University, Boston, MA

Sep 2018 – Jan 2020

M.S., Computer Information Systems (concentration: Database Management & Business Intelligence) | **GPA 3.60**

Indira Gandhi Delhi Technical University, New Delhi, India

Aug 2014 – Jun 2018

Bachelor of Technology, Electronics & Communication Engineering | **CGPA 8.9**