

Contact

swati.mittal.indore@gmail.com

www.linkedin.com/in/swatisjsu
(LinkedIn)

Top Skills

Prompt Engineering

Retrieval-Augmented Generation (RAG)

Model Training

Languages

English (Native or Bilingual)

Hindi (Native or Bilingual)

Certifications

MongoDB

Patents

Applying cascading machine learning models to command prompts

Swati Mittal

Vice President, Technology Special Projects - Cutting edge GenAI, Agentic AI, MCP solutions
Frisco, Texas, United States

Summary

GenAI Technology Leader with 10 years of hands on experience in building teams, solutions and deploying latest technology at scale in agile fast paced setup.

Technology leader:

Shaping the GenAI future at Citigroup (US Personal Banking Division). Currently driving innovative AI solutions - including GenAI (ChatGPT, Gemini, Claude), Agentic AI, Anthropic MCP (Model Context Protocol), and in-house open source LLM fine-tuning (Llama, Mistral, hyperparameter optimization).

Solution Architect:

A seasoned leader, building high-impact AI solutions from concept POC ground up, encompassing business case development, estimation, benefit summary, securing funding, and end-to-end enterprise delivery. Adept at building and mentoring innovative cross-functional technical teams, system design, infrastructure & process setup, implementing best practices in ethical AI, and Model Risk Management (MRM).

Subject Matter Expert:

Exceptional professional passionate about using technology leadership to drive real life changes. Driving force in navigating cross business presentations, decisions, serving technology guide on model selection & performance, hallucinations, guardrails, and leading technical team as an SME.

Patent application pending with Citi:

19/017,003 with USPTO on 17-Jan-2025 – Titled: “Applying cascading machine learning models to command prompts”.

Experience

Citi

Vice President, Technology Special Projects - Cutting edge GenAI, Agentic AI, MCP solutions – USPB

January 2023 - Present (2 years 8 months)

Irving, Texas, United States

GenAI Intelligent Chatbot (Application Manager, Technology Leader)

Solution Owner - Architecture, POC, system design, infrastructure setup, deployment, and Model Risk Management (MRM) of global GenAI chatbot. Secured project funding - compelling business case presentations to executive leaders. Established best practices in ethical AI, implemented guardrails to minimize hallucinations, & optimized model performance for scalability.

Streamlined data concern filing for 1600+ users. The data quality has increased from 36% to 84%.

Tech: GenAI (Google Gemini, ChatGpt), Hybrid RAG (Retrieval-Augmented Generation), Embeddings, Prompt Engineering, Zero-shot / Few-shot Learning, Python, AI safety guardrails, Inference time (latency) optimization, Transformer, Multi-threading, Scalable microservices

Data Concern and Remediation – Root Cause Analysis and Invalid\Valid Detection:

After a \$136M fine in July 2024, improving data quality became a Citi priority. I led special projects team to design & secure \$6M for two GenAI POCs, delivering full lifecycle from POC to deployment (in-house LLM selection, fine-tuning).

Cut RCA time from 60 to 24 days.

Tech: Open Source LLM Model (Mistral, Llama), Quantization using Bits and Bytes, LoRA/QLoRA, PEFT, Multiple GPU, Hyper parameter tuning (Rank, Alpha), Accelerator

Duplicate Detection/Similarity Grouping: Delivered GenAI semantic search project to identify and group potential duplicate data concerns, enabling faster resolution.

ARCM – Activity, Risk, Control, Monitoring- Q&A: Led a team to develop and fine-tune NLP(Natural language processing) models to predict answers to multiple questions from input descriptions.

Tech:

Tokenizer, vectorization, hugging face, sentence transformer, all-mpnet-basev2, BM25, TF-IDF, cosine similarity, DistilBERT, BERT, RoBERTa, logistic regression, Random Forest, Binary Classification, Open-shift, Lightspeed, CI/CD pipelines, NGNIX Load balancer, Flask, Pandas, REST API

Bank of America

Software Engineer

February 2017 - February 2020 (3 years 1 month)

Plano, Texas, United States

Lead Software Developer for fast paced high visibility and impact project in Visual Analytics Team

- Real Time Payment Reporting – PySpark, Python, Hadoop, Shell Script
Extremely complex calculations on multiple data sources and multiple rows and columns achieved using PySpark to generate required reports. First project in my team to use PySpark.

- BoA Customer Survey Review Sentiment Analysis - Machine Learning, Python, PySpark

Building machine learning Algorithm (SPM Logistic Regression) to find if the sentiment was positive or negative

- Medallia Projects (PMMR Incentive, PMMR Home Office, Voices Checs) - Hadoop, Hive

Completed 3 complex end to end projects involving data ingestion, data cleansing, data retrieval & data processing

- High Risk Financial Center Automation - Hadoop, Hive

Complex Query to harmonize data from different sources with numerous sanity checks to ensure data consistency

- Anti-Money Laundering (AML) Report - Tableau, Hive, Hadoop

Learned and Created Tableau Dashboard in 1 month. LOBs delighted with the dashboard output.

- Develop end to end scalable application in ADI – Groovy, Apache Camel, Java 8

Built module to read 128 million records in a 15 GB file on server in < 3 minutes by using Java8 stream functionality

- Value Add - Python Scripts (Erica & Others)

To reduce routine work, to optimize data processing time (reduced processing time of 8 hours for a file to 10 min)

PlayStation Network

Software Engineer Intern

June 2016 - August 2016 (3 months)

- Agile Scrum Setup DevNet Platform Team- Fixed highly integrated text editor
- one of the biggest pain points
- Worked on PHP, JavaScript, jQuery, HTML, CSS, XML, SQL, Docker, Git, Server Client Architecture, Selenium

CSC

Software Engineer

November 2009 - March 2013 (3 years 5 months)

- Back-end Developer for life insurance client for portal used to issue and update policies.
- Worked on various projects involving the fund management application for investment plan module.
- Performed requirement analysis, design, coding and testing for various projects.
- Designed and prepared detailed workflow charts and diagrams describing the inputs, outputs and logical operation steps to achieve business objectives.
- Performed integration testing, code retrofitting and version management.
- Consistently rated as the top performers in the team. Primary Technologies – Java, HTML, JSP, SQL

Education

San Jose State University

Master's Degree, Computer Software Engineering · (2015 - 2016)

Shri G S Institute of Technology & Science

Bachelor of Engineering (BE), Engineering · (2004 - 2008)