

EXECUTIVE SUMMARY

Award-winning strategic GenAI leader (11 honors, 1 patent pending), securing \$7 million in business funding for GenAI-powered chatbot and other GenAI use-cases. Driving enterprise-scale, cutting-edge GenAI, Agentic AI, and ML solutions from concept to production at Citi and Bank of America, combining innovation, business use-case ideation, and high-impact executive-level presentations. Scaled MRM (Model Risk Mgmt.) compliant LLM/RAG to 1.6k+ users, fine-tuned in-house LLMs on Citi data, and implemented guardrails - raising data quality from 36% → 84% and reducing RCA from 60 → 24 days. Built & mentored high-performing teams, accurate estimates, established Generative AI processes, deployed solutions, and scalable infrastructure from the ground up.

TECHNICAL EXPERTISE

Generative AI & LLMs: Google Gemini, ChatGPT, Agentic AI, MCP (Model Context Protocol), Mistral, Llama, Hugging Face, LangChain, Hybrid RAG (BM25 + embeddings), Prompt Engineering, Zero/Few-Shot Learning, AI Safety Guardrails
ML/DL & NLP: BERT, DistilBERT, Sentence Transformers, TF-IDF, Cosine Similarity, Logistic Regression, Random Forest
Fine-Tuning: Quantization (BitsAndBytes), LoRA/QLoRA, PEFT, Multi-GPU, Hyperparameter Tuning, Latency Optimization
Infrastructure & Deployment: Python, Accelerators, Flask, Microservices, OpenShift, NGINX, REST APIs, CI/CD Pipelines
Cloud & Data Systems: AWS, pgvector (Postgres), Hadoop, Hive, MySQL, Oracle, MongoDB, Docker, Kafka

PATENT PENDING (CO-AUTHOR WITH CITI) – RECEIVED CITI GOLD AWARD

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

AWARDS & RECOGNITION (11 AWARDS)

Citi – 4x Awards – Gold, Silver, Copper & Bronze Awards (2023–2025)

Gold Award – GenAI Hackathon Winner (2023), Silver Award – GenAI Playground Launch for DCRM (2024) – by Managing Director, Copper Award – Exceptional Technical Leadership (2025), Bronze Award (2024)

Bank of America – 7x Awards (2017–2020)

2× Silver – Exceptional Performance (2019–2020), 3× Bronze – Client Delight & Excellent Delivery (2018), High Five Award – ML Paper at Devcon (2017), Special Recognition – Strategic Relationship Rebuilder (2019)

WORK EXPERIENCE (Total 10+ years)

Citi (~3 Years)	Dallas, Texas	VP, Technology, Special Projects	Jan 2023 – Till Date
– GenAI Intelligent Chatbot (Technology Leader, Solution Architect, Application Manager): (Copper Award)			
Solution Owner responsible for end-to-end concept, POC, Architecture, system design, infrastructure setup, deployment, and Model Risk Management (MRM) of global GenAI chatbot. Secured project funding of 1 million \$ by compelling business case presentations to executive leadership. Established best practices in ethical AI, implemented guardrails to minimize hallucinations, and optimized model performance for production scalability. Streamlined data concern filing process for 1600+ users. The data quality has increased from 36% to 84%.			
– In-house trained Open Source LLM GenAI RCA & Invalid\Valid Detection: (Silver Award)			
After a \$136M regulatory fine in July 2024, improving data quality became Citi’s top priority. I led the special projects team that conceptualized, built a POC, and secured \$7M in business funding for two GenAI-powered Projects - Root Cause Analysis (RCA) and Valid \ Invalid detection for data concerns created by thousands of systems and users across different LOB (Line of Business). Owned the full lifecycle from POC to enterprise delivery, including in-house LLM model selection, fine-tuning, and deployment . Impact: Average Root Cause Analysis time required reduced from 60 days to 24 days, increasing data quality and reducing turnaround time.			
– Gen-AI powered Duplicate Detection/Similarity Grouping: (Bronze Award)			
Delivered a GenAI-based semantic search project to identify and group potential duplicate data concerns, enabling faster resolution. Impact: Improved resolution speed and reduced analyst workload by eliminating repeat investigations, ~20% of the workload for a team of 1500 global LOBs.			
– NLP (Natural Language Processing) models to predict Q&A for ARCM – Activity, Risk, Control, Monitoring:			
Led a team to develop, fine-tune NLP (Natural Language Processing) models to predict answers to multiple questions from input descriptions.			

Swati Mittal – Vice President, Technology (10+Yrs Exp., Dallas, TX)
Phone +1 (682) 228-7776 Email – swati.mittal.indore@gmail.com

Master of Science, Software Engg., SJSU, CA
LinkedIn – <https://www.linkedin.com/in/swatisisu>

WORK EXPERIENCE (CONTINUED)

– **GenAI-based classification for Regulatory Report Tagging**

Delivered a GenAI-based classification system solution to predict the regulatory report linked to each data concern, improving accuracy and speed in compliance workflows.

Career Break

Mar 2020 – Dec 2022

Family relocation to Switzerland; maintained skills via independent Machine Learning projects.

Bank of America (3 Years)

Dallas, Texas

Software Engineer

Feb 2017 – Feb 2020

– **Real Time Payment Reporting – PySpark, Python, Hadoop, Shell Script (Silver Award)**

Delivered PySpark-based solution to process high-volume data across multiple sources using DataFrames for complex arithmetic operations. PySpark evangelist - First project within Bank of America to use PySpark.

– **BoA Customer Survey Review Sentiment Analysis - Machine Learning, Python**

Led a team to deliver a Machine Learning POC (Logistic Regression) to classify customer feedback sentiment. Trained and refined the model on high-volume masked data, presenting results to business stakeholders

– **Medallia Projects** (PMMR Incentive, PMMR Home Office) - **Hadoop, Hive (Bronze Award)**

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 (Silver Award)**

Developed Tableau dashboards using advanced calculations to deliver actionable insights. Accelerated LOB decision-making and earned recognition for high-quality delivery under tight deadlines.

– **Develop an end-to-end, highly scalable application to read a large dataset for report generation**

Built a module to read 128 million records in a 15 GB file on the server in < 3 minutes by using 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 mins).

Sony PlayStation (0.5 Years)

San Mateo, California

Software Engineer

June 2016 – Aug 2016

– Agile Scrum Setup DevNet Platform Team- Fixed a highly integrated text editor in the Agile Scrum environment.

– Worked on PHP, JavaScript, jQuery, HTML, CSS, XML, SQL, Docker, Git, Server Client Architecture, and Selenium

CSC (now DXC Technology) (3 Years)

India

Backend Software Engineer

Nov 2009 – March 2013

– Java Developer in the back-end system development and maintenance team for a life insurance client

CERTIFICATION

– LLM Engineering: Master AI, Large Language Models & Agents (2025) - by Ed Donner, Length 8 weeks - Udemy

– The Complete Agentic AI, MCP Engineering Course (2025) - by Ed Donner, Length: 6 weeks - Udemy

EDUCATION

- Master of Science – Software Engineering

GPA 3.60 / 4 San Jose State University, San Jose, CA

Specialization: Machine Learning, Cloud Computing, Data Science, Algorithms, Neural Networks, Internet of Things

- Bachelor of Engineering

GPA 3.71 / 4 SGSITS, Indore, India

PERSONAL PROJECTS

– **MCP Project – Details to be filled**

2025

– **Agentic AI Project – Details to be filled**

2025

– **UnMute (Android, Python, Image Processing, Machine Learning – CNN, RNN, TensorFlow, AWS)**

2016

PROJECT DEMO VIDEO: https://www.youtube.com/watch?v=LAUP0NfLz9U&list=PLe7_BymgrilyuZh0XFsyv1No9KtLJBj4

- Developed an app to help the visually impaired navigate surroundings - the app uses machine learning algorithms to analyze pictures taken of the surroundings, and recites the list of things, location, and text in it.