ABHISHEK SURESH

📞 551-229-8434 ✉ as18757@nyu.edu in [LinkedIn](#) ⌨ [GitHub](#)

EDUCATION

Master of Science, Information Systems   **May 2025**
New York University, New York, USA

Bachelor of Technology, Computer Science & Engineering   **May 2019**
SRM Institute of Science and Technology, Chennai, India

EXPERIENCE

AI/ML Intern   **Aug 2025 – Present**
Archemy Inc, New York, USA

- Developed a **SQL-to-vector RAG pipeline** syncing hierarchical taxonomy data from **MySQL** to **ChromaDB Vector Store**, enabling real-time semantic search via **MiniLM-L6-v2** embeddings and improving data search term matches by **70%**

AI Engineer   **Jan 2023 – Aug 2023**
Bread Financial, Bengaluru, India

- Deployed an **Isolation Forest** model on **AWS SageMaker** for **credit card fraud detection**, integrating predictions into **RPA workflows** via REST APIs—boosting recall by **18%** over baseline models
- Developed a **RoBERTa-Large**-powered **ticket classification bot** on **UiPath** using **SageMaker endpoints** to assess sentiment and urgency, improving response prioritization and cutting resolution time by **20%**
- Automated root-cause detection and process recovery across **UiPath Orchestrator**, reducing downtime by **25%** and improving system reliability through data-driven anomaly resolution

Systems Engineer   **Sep 2019 – Jan 2022**
Infosys Ltd, Pune, India

- Optimized **Automation** workflows by integrating **API triggers** and automating ticket resolutions, reducing average processing time by **30 minutes per task** and improving CRM data accuracy by **80%**
- Developed **RPA bots** to execute **SQL queries** and trigger **API calls** for ticket and user data retrieval, minimizing UI automation failures from dynamic layouts and improving query efficiency by **3 seconds**
- Implemented dynamic **workload balancing** and auto-recovery in **Automation Anywhere**, optimizing bot distribution and cutting recovery time by **20 seconds per failure** while collaborating across **4+ teams** to standardize PDD/SDD documentation

PROJECTS

FinSearch AI | Python, LangChain, Cassandra, BM25, TF-IDF, Pandas, NumPy

- Built a **hybrid RAG** system (**BM25** + vector embeddings) with **Cassandra** for real-time financial document retrieval, achieving 94% retrieval accuracy via metadata filtering and **MMR reranking** outperforming single-method approaches by **18%** in precision

Deep Research Agentic AI Workflow | Python, OpenAI Agents, LangChain, AsyncIO, Gradio, Pydantic, OpenAI Tracing

- Built a **multi-agent research platform** using **OpenAI's Agents SDK framework**, coordinating 4 specialized agents (Planner, Search, Writer, Email) with **asynchronous concurrency** to execute parallel web searches, synthesize comprehensive reports, and automate email delivery reducing research turnaround time from hours to minutes.
- Designed a **scalable agentic workflow** with **structured output validation (Pydantic)**, **real-time progress streaming via Gradio**, and **OpenAI trace instrumentation**, enabling intelligent query decomposition and parallel task execution across 20+ concurrent research threads with robust observability.

Political Sentiment and Trend Analysis | Python, Apache Kafka, NLP, Reddit API (PRAW), NY Times API, Plotly

- Developed a real-time data pipeline with **Apache Kafka** to collect Reddit data, extract keywords using TF-IDF, and fetch relevant articles from the **NY Times API**, applying **LDA** for **topic modeling** to track U.S. 2024 election trends
- Built an XGBoost machine learning model (**81.19% accuracy**) to predict trending election topics, and aid to influence political **campaign strategies**

Product Price Prediction using LLMs | Python, Hugging Face, Ollama, LoRA, QLoRA, Pandas, NumPy

- Engineered a price prediction system fine-tuning **LLMs (GPT-4, Claude, LLaMA 3.1)** on a **400K sample Hugging Face dataset consisting of Amazon product details**, outperforming baselines (**linear regression, random forest**) by **20%** with MSLE of **0.80**
- Performed data munging by removing uninformative tokens, extracting metadata (brand, category) as auxiliary features, and leveraging **LoRA/QLoRA** for parameter-efficient fine-tuning

SKILLS

- **Programming Languages/ Databases :** Python, C/C++, JavaScript, MySQL, MongoDB, PostgresSQL, ChromaDB, Cassandra
- **Data Analysis/Machine Learning :** Altair RapidMiner, Apache Airflow, NumPy, Pandas, Tensorboard, TensorFlow/Pytorch, Scikit-Learn, Tableau
- **Cloud & DevOps :** AWS MSK, Jenkins, GCP, Heroku, Azure ML Studio, AWS Sagemaker, AWS S3, UI Path, Automation Anywhere, SocketIO
- **Certifications :** Automation Anywhere Advanced RPA Professional V11.0 , Google Data Analytics