

# SOHAM CHINCHALKAR

Tempe, AZ, USA | +1 (602) 743-3275 | [schinch8@asu.edu](mailto:schinch8@asu.edu) | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

## EDUCATION

### M.S. Information Systems

*Arizona State University*

**Aug 2024 - May 2026**

### B.E. Computer Engineering

*Savitribai Phule Pune University*

**Aug 2020 - Jun 2024**

## PROFESSIONAL EXPERIENCE

### Data Scientist, Rose AI, New York, US

**May 2025 - Aug 2025**

- Automated 9 daily analytical chart decks using Python and Apache Airflow, streamlining data visualization and supporting rapid prototyping of AI/ML insights to improve operational efficiency and drive a 28% increase in profits.
- Designed and fine-tuned an agentic AI solution using LLMs and Python to autonomously traverse chart decks, verify accuracy, and report discrepancies—achieving 92% efficiency. Leveraged foundation model tuning for domain-specific validation logic.
- Designed a Retrieval-Augmented Generation (RAG) financial assistant using LLaMA 3 and LoRA fine-tuning, enabling contextual querying of internal financial knowledge—successfully deployed for internal use.
- Extracted and cleaned trillion-record datasets via SQL, Automeris APIs, and Python. Supported AI readiness and compliance alignment for potential regulated applications, including finance and clinical trials.
- Engineered robust, scalable data pipelines to process massive datasets from 3+ stock exchanges and 7+ market data vendors, leveraging SQL and Python for high-performance quantitative analysis and establishing a foundation for future AI model development.
- Led strategic consulting and collaborated with two leading investment banks to integrate proprietary financial logic into Rose AI's platform, enhancing compliance and securing key partnerships.

### Graduate Teaching Assistant, & Data Analyst – Applied ML and Visualization, ASU

**Aug 2024 – Present**

- Created real-time, interactive Tableau dashboards linked to live SQL databases, driving a 25% increase in course offerings.
- Delivered 15+ practical sessions on SQL, Tableau, and machine learning to 200+ students and professionals.
- Assisted faculty with assessment design and grading, improving academic operations by 30%.

### Data Engineer, Virtual Galaxy Infotech, Nagpur, India

**Feb 2023 – May 2023**

- Developed Python-based ETL pipelines and optimized complex SQL queries for 5 of the largest banks in India, reducing report generation time by 30% and ensuring data integrity for large-scale banking data, which supports data ingestion for AI initiatives.
- Assisted senior DBAs in resolving 50+ database errors, reducing downtime by 15%, while collaborating with cross-functional teams to streamline workflows and bolster data support for AI/ML initiatives.

## RESEARCH, PATENTS & PROJECTS

### Multi-Agent LLM Query System with LangChain, Gemini & FAISS | [View Project](#)

**June 2025 – July 2025**

- Designed a multi-agent Retrieval-Augmented Generation (RAG) system using LangChain, MiniLM, and FAISS for semantic routing across clustered domain knowledge, enabling real-time NLP query resolution.
- Built a dynamic meta-agent selector using KMeans clustering and similarity scoring, integrating Gemini and OpenAI LLMs for prompt-based, multi-turn reasoning with ranked, explainable outputs.
- Achieved 95% accuracy in response relevance; utilized TensorFlow for confidence scoring and ranking logic, delivering a Colab-based autonomous pipeline handling 150K+ structured/unstructured records.

### Langchain & LLM-Powered Knowledge Graph for Real-Time NLP Queries | [View Project](#)

**Feb 2025 – Apr 2025**

- Developed a Retrieval-Augmented Generation (RAG) system using Langchain, OpenAI APIs, and Neo4j to support real-time semantic querying on 150K+ structured/unstructured records. Deployed on AWS; enabled fine-tuned LLM pipelines for insight generation with 98% accuracy.
- Achieved 98% accuracy in generating relevant, explainable outputs by translating NL queries to Cypher via LLMs and enriching the graph with inferred patterns for actionable insights like product trends and recommendations.

### Financial Fraud Detection System Using Anti-Benford Subgraphs & ML Algorithms (Patented)

**Oct 2023 – Jan 2024**

*Published in Scopus-Indexed Journal, Informatica | Presented at IEEE AIKIE 2023 | DOI: <https://doi.org/10.31449/inf.v498i4.8298>*

- Developed an unsupervised financial fraud detection system using Anti-Benford subgraph mining and Louvain-based community detection, achieving 94.83% anomaly detection accuracy on large-scale transaction networks.
- Engineered scalable ML pipelines using Pandas, NumPy, and Scikit-learn, leveraging CUDA-based parallelism to accelerate inference on high-volume financial data.

**Patent Granted – Anti-Benford & Louvain-Based Financial Fraud Detection, Application No. 202521025821, India, Apr 2025.**

### Keylogger Detection System Using ML Algorithms and Dendritic Cell Algorithm (Patented)

**Jun 2023 – Feb 2024**

*First Author | Published in Scopus-Indexed Journal, Revue d'Intelligence Artificielle (IIETA) | DOI: [10.18280/ria.380128](https://doi.org/10.18280/ria.380128)*

- Designed a hybrid keylogger detection system integrating the Dendritic Cell Algorithm (DCA) with SVM and Naive Bayes, achieving 99.8% detection accuracy on behavior-based threat datasets.
- Built end-to-end ML pipelines in Python using NumPy, Pandas, and Scikit-learn, improving detection precision through custom feature engineering and hyperparameter tuning.

**Patent Granted – Keylogger Detection System Using ML Algorithms & DCA, Application No. 202421040728, India, Nov 2025.**

## SKILLS

- Databases & Cloud:** MySQL, MongoDB, NoSQL, PL/SQL, Hive, Snowflake, Neo4j, Databricks, Microsoft Azure, AWS s3
- AI/ML Stack:** LangChain, Gemini, OpenAI, TensorFlow, PyTorch, FAISS, SentenceTransformers, LLM Prompt Engineering, RAG Architectures, Agent-based Systems, Vector Databases, Scikit-learn, HuggingFace
- Interpersonal Skills:** Leadership, Projects management, Problem-solving, Teamwork, Communication skills, English Proficiency

## ACHIEVEMENTS

- Engineering Graduate Fellowship from Fulton Schools of Engineering.
- 2 Merit based scholarships from Arizona State University.
- Authored 3 peer-reviewed research papers and 2 granted patents by designing production-grade AI/ML systems.
- Earned AWS Data Engineering Foundations and NLP & Machine Learning Foundations digital badges.