



Summary

Results-oriented Data Scientist with 5+ years of experience in developing and deploying machine learning solutions, specializing in NLP, computer vision, Gen-AI, and Agentic-AI. Proven ability to build scalable models, derive actionable insights from complex datasets, and mentor junior data scientists. Expertise in Python, R, and cloud platforms (AWS, Azure, GCP). Seeking a challenging Senior Data Scientist/GenAI role to leverage AI and machine learning expertise to drive innovation and business impact.

Education

Thakur College of Engineering & Technology

Bachelor of Engineering in Electronics & Telecommunication, CGPA: 9.04/10

- Relevant Courses included: Signals & Systems, Discrete Time Signal Processing, Image Processing and Machine Vision, and Database Management Systems, providing a strong foundation for data science applications.

Experience

Tata Technologies

DataScientist

Mar 2025

- Led the development and deployment of a No-Code ML Platform, enabling users to build, train, and deploy machine learning models through a drag-and-drop interface with integrated AutoML.
- Automated the end-to-end machine learning pipeline, streamlining data preprocessing, model selection, training, and deployment, resulting in a 70% reduction in model development time.
- Empowered non-technical users to leverage AI-driven insights, leading to significant improvements in organizational decision-making and innovation speed.

Trident Services (Data Society Group)

Data Scientist

Jul 2022 - Jul 2024

- Implemented a Local RAG System using a Mistral-7B quantized model, reducing assessment creation time by 2-3 man-days per course and improving team productivity by 15%.
- Developed and fine-tuned NLP models for sentiment analysis and information extraction using BERT, achieving a 20% improvement in prediction accuracy.
- Enhanced product reliability and improved user experience through accurate sentiment analysis and information extraction.

Employtech Center India Private Limited

Data Scientist

Aug 2024 - Jan 2025

- Developed and implemented machine learning, deep learning, computer vision, and generative AI solutions for the medical and retail domains.
- Created the "assessment-creator" Python package, automating course assessment creation and streamlining question-and-answer generation.

SACHINMISHRA



Portfolio Link

-  8928-450-584
-  sachin19566@gmail.com
-  linkedin.com/in/sachin-mishra19566/
-  <https://sachinmishra-ux.github.io/sachin-portfolio/>
-  Mumbai, Maharashtra, 400064
-  <https://github.com/SachinMishra-ux>



Skills

- Python
- R
- SQL
- PySpark
- Spark SQL
- Deep Learning
- Machine Learning
- NLP
- Computer Vision
- Generative AI
- Regression & Classification
- Time Series Analysis
- Hypothesis Testing
- Ensemble Methods
- BERT
- Transformers
- Langchain
- LangGraph
- CrewAI
- MLFlow
- DVC
- AWS Sagemaker

- Google A2A
- Google ADK
- AWS Bedrock AgentCore
- MCP
- GCP
- Vertex AI
- Azure
- AWS
- GCP
- Tableau
- Power BI
- MySQL
- MongoDB
- Snowflake
- ChromaDB
- Pinecone



Extra Curricular

- Created LocdataMAC Python Package & released on [PyPI Page](#):
- Created a 15+ Hour long Content for LangGraph on [YouTube](#) & [GitHub](#)
- Gold Medalist for IC the future Competition

- Designed and deployed a serverless data analysis platform using AWS Lambda for real-time monitoring of key performance metrics for a Social Network Application.
- Provided actionable insights from the platform's data analysis pipeline, enhancing decision-making capabilities for the client.

Tata Consultancy Services (TCS)

Assistant System Engineer

Jul 2021 - Jul 2022

- Managed large-scale SAP HANA databases for production, development, and quality environments, ensuring data integrity and reducing data downtime by 20%.
- Developed a Tableau BI dashboard to track monthly incidents and tasks, reducing SLA time by 30% and optimizing task allocation.
- Collaborated with cross-functional teams using Agile methodologies to improve project execution efficiency.

Projects

Agentic RAG Platform (MetaMech)

- Architected an enterprise-scale RAG platform with hybrid retrieval, reranking, query decomposition, and multi-agent reasoning, enabling knowledge discovery across hundreds of training courses with sub-3-second response times.
- Integrated vector databases, knowledge graphs, and evaluation frameworks, improving answer accuracy by **30%+ and significantly reducing hallucinations** in production environments.
- Built scalable retrieval pipelines and production-grade APIs supporting thousands of monthly queries, **contributing to ₹15–25 Lakhs** in business value through improved learner productivity and faster information access.
- Implemented end-to-end observability, evaluation, and monitoring pipelines, reducing issue detection and troubleshooting effort by 50%+.
- Automated **CI/CD and cloud deployments using Azure DevOps**, reducing release cycles from hours to under 15 minutes and enabling reliable production rollouts.

Tech: LangGraph, LangChain, Neo4j, Pinecone/Weaviate, FastAPI, Azure DevOps, Azure Cloud.

AutoNPI Intelligence Engine

- Developed an AI-powered workflow automation platform for end-to-end New Product Introduction (NPI) processes, reducing approval and onboarding cycles from 5–7 days to under 24 hours.
- Leveraged LLMs, document intelligence, and predictive ML models to automate compliance validation, approval workflows, and decision-making, reducing manual effort by 60%+.
- Built scalable microservices and AI-driven workflow orchestration supporting multiple business units and hundreds of NPI requests annually, improving process visibility and governance.
- Implemented CI/CD pipelines using Jenkins and deployed production workloads on AWS EC2, reducing deployment time from hours to less than 15 minutes while ensuring reliable releases.

Tech: Python, FastAPI, AWS EC2, PostgreSQL, LangChain, Machine Learning, Jenkins, REST APIs.