

SAGAR FADATARE

Python Developer | Generative AI Engineer | AI/ML Engineer

Pune, Maharashtra, India | +91-8668687998 | sfadatare10@gmail.com

sagarfadatare.github.io | linkedin.com/in/sagar-fadatare | github.com/sagarfadatare | huggingface.co/sagarfadatare

SUMMARY

Python and GenAI engineer with an MCA (2026) who builds practical AI systems end to end. I work with LLM agents (LangChain, Anthropic Claude, Text-to-SQL, tool calling), Stable Diffusion apps deployed on Hugging Face, and the data pipelines (PySpark, Databricks, dbt) that feed them, backed by a foundation in machine learning and SQL. Seeking an entry-level Python / GenAI / AI-ML engineering role where I can ship practical AI solutions.

EXPERIENCE

Data Engineer Intern | GAOTek Inc. (Remote)

Jan 2026 – Apr 2026

- Completed hands-on data engineering tasks on Databricks, building and running PySpark pipelines to ingest, transform, and load datasets.
- Cleaned and standardized raw data by handling nulls, duplicates, inconsistent formats, and invalid values, then validated the output before use.
- Built multi-stage pipelines (raw to cleaned to analytics-ready) following data engineering best practices taught in the program.
- Completed a structured three-month internship program covering Databricks and data cleaning concepts (certified April 2026).

PROJECTS

Text-to-SQL Agent

Sep 2026 – Oct 2026

Python, SQL, LLM Agents

- Built an LLM agent that converts plain-English questions into SQL queries and runs them on an e-commerce dataset.
- Returns results as answers to business questions, so non-technical users can get sales and order insights without writing SQL.
- Extended my earlier e-commerce data engineering work into a question-answering tool.

Python AI Research Agent

Jul 2026 – Aug 2026

Python, LangChain, Claude, Pydantic

- Built a tool-calling research agent with LangChain and Anthropic Claude that searches the web (DuckDuckGo) and Wikipedia for a given topic.
- Added a custom tool that saves the agent's findings to a text file, and enforced structured output (topic, summary, sources, tools used) with Pydantic.
- Secured API keys with python-dotenv and documented setup and usage in a detailed README.

Text-to-Image Generation App

Aug 2025 – Sep 2025

Python, Stable Diffusion, Hugging Face, Gradio

- Built a text-to-image pipeline on a pretrained Stable Diffusion model from Hugging Face with an interactive Gradio interface.
- Deployed it as a live public demo on Hugging Face Spaces, usable through a shareable link with no local setup.

End-to-End Data Engineering Pipeline

Nov 2025 – Dec 2025

PySpark, Databricks, Delta Lake, SQL

- Designed a Bronze-Silver-Gold Medallion pipeline on Databricks for an e-commerce dataset of 100,000+ transactions.
- Cleaned and validated the data with PySpark, stored it in Delta Lake, and modeled it into a star schema for analytics.
- Wrote SQL (CTEs, window functions) to produce business KPIs by region, category, brand, and time.

TECHNICAL SKILLS

Languages	Python, SQL, Java
Generative AI	LLMs, Prompt Engineering, AI Agents, Tool Calling, RAG, Stable Diffusion
ML & Deep Learning	Machine Learning, Deep Learning, Scikit-learn, PyTorch
Frameworks	LangChain, LangGraph, Pydantic, Hugging Face, Gradio
Data Engineering	PySpark, Databricks, Delta Lake, dbt, ETL, Medallion Architecture, Unity Catalog, Star Schema
Analysis & Visualization	Pandas, NumPy, Matplotlib, Seaborn, Power BI, Tableau
Databases	MySQL, PostgreSQL, MongoDB
Tools & Cloud	Git, GitHub, Docker, Jupyter Notebook, Google Colab, VS Code, AWS, Azure

EDUCATION

MCA, Savitribai Phule Pune University, Pune | CGPA 8.22

2024 – 2026

BCA, Punyashlok Ahilyadevi Holkar Solapur University | 83.97%

2021 – 2024

CERTIFICATIONS

- Azure Data Engineering End-to-End, Udemy (77.5 hours)
- Programming in Java, NPTEL, IIT Kharagpur (Elite, 84%)
- Python for Data Science, NPTEL

RESEARCH PUBLICATION

- The Rise and Dominance of Cricket Over Other Sports in India, IRJMETs, October 2025. Analyzed survey data from 331 respondents.

Languages: English, Hindi, Marathi, Kannada