Sagar Yadav

Data Scientist, AI/ML, Generative AI Engineer

|Mob: 9104371519 | sagar.aidata@gmail.com | GitHub | LinkedIn |

Professional Summary:

Skilled AI/ML Engineer and Data Scientist with 4+ years of experience building scalable machine learning and generative AI solutions. Eager to contribute expertise in Python, NLP, and transformer models to drive innovation in real-world applications. Passionate about advancing AI by integrating emerging technologies, including future-ready innovations like next-gen computational paradigms, to solve complex, cross-domain challenges.

Work Experience

AI-Data Scientist, Topia Life Sciences, Ahmadabad. (Dec 2023 – Present)

- Architected and deployed end-to-end AI/ML pipelines for prediction, optimization, and intelligent automation.
- Designed agent-based AI systems capable of autonomous task execution, reasoning, and workflow orchestration.
- Applied Generative AI techniques for model augmentation, feature synthesis, and knowledge extraction.
- Collaborated with cross-functional teams to align AI solutions with enterprise goals and operational needs.
- Led the development of predictive analytics tools to support decision-making and process optimization.
- Collaborated with cross-fun Integrated AI outputs into production systems, dashboards, and business intelligence platforms.
- Improved model accuracy, scalability, and computational efficiency across multiple product pipelines.
- Improved predictive accuracy and computational efficiency, enabling scalable insights across product pipelines.

Mentor, AI-Adventures, Pune. (Oct 2022 – Nov 2023)

- Mentored students in data science and AI, guiding them through complex concepts and practical applications.
- Supported their academic and career development, helping them build skills for real world challenges.

Sr. Research Associate, Sai Life Sciences Ltd, Hyderabad (Jun 2021 – Sep 2022)

- Applied machine learning models to scientific datasets to extract patterns, predict outcomes, and automate analysis workflows.
- Managed end-to-end data workflows including cleaning, transformation, and modeling for large datasets.
- Developed reproducible ML pipelines for pattern classification, property prediction, and data clustering.

Career Transition to AI/ML

Prior to 2021, worked for approximately 9 years as a Research Associate in R&D and Analytical Chemistry across pharmaceutical domains. This experience laid a strong foundation in experimental analysis, data integrity, structured problem-solving, and computational tool usage skills now translated effectively into AI/ML applications.

Professional Skills

- **Programming:** Python, R
- **Databases:** SQL, MySQL, MongoDB
- **Machine Learning:** Supervised/Unsupervised learning, AutoML, Regression, Classification, Clustering, PCA, Hyperparameter Tuning
- **Libraries:** TensorFlow, Keras, Scikit-learn, NumPy, Pandas, SciPy, Seaborn, Matplotlib, Yellow- brick, PYSCF, Q-Chem, Hugging Face, Psi4
- **Generative AI:** GPT, GANs, VAEs, LLMs – Fine-tuning and application, AIAgentic
- **NLP:** Text classification, summarization, named entity recognition, sentiment analysis
- **Transformer Models:** BERT, GPT family, LLaMA – for diverse NLP tasks
- **Visualization & Analysis:** Streamlit, EDA, Feature Engineering, Interactive Dashboards
- **DevTools:** GitHub, Jupyter, Streamlit
- **(Optional domain tools):** Familiarity with computational tools like AutoDock & PyRx (for scientific data projects)

Key Projects

- **AI-Driven Drug Discovery (Company Projects):** Improved QSAR and ADMET toxicology predictions by 15% using ensemble methods and domain-specific feature engineering, additionally integrating novel quantum chemistry derived parameters to enhance molecular property representation and achieve higher predictive accuracy.
- **Personal Open-Source Research Projects (GitHub)**
    - **3D Conformer-Aware Drug Interaction Modelling:** Developed molecular agents that evaluate drug interactions using full conformational ensembles.
    - **Design-Thinking AI Innovation Framework:** Created a structured AI-driven framework supporting ideation, experimentation, and systematic knowledge creation.
    - **Self-Driving Laboratory Automations:** Built agentic automation pipelines for autonomous experiment planning, execution, and analysis.
    - **Robust Literature Review Generator:** Developed an AI tool for comprehensive retrieval, synthesis, and structured scientific literature reporting.

Education

- Post Graduate Diploma in Machine Learning and AI, Amity University Online
- M.Sc. in Analytical Chemistry, Nowrosjee Wadia College, Pune University

---

Certifications

- Sensors and Actuators – IISc Bangalore
- Bioinformatics: Algorithms and Applications – IIT Madras
- Design Thinking - A Primer – IIT Madras

Languages

- English, Hindi, Marathi