SHRIVATSASINGH RATHORE

480-703-7417 srathor9@asu.edu linkedin.com/shrivatsasingh-rathore github.com/sthor07

EDUCATION

Master of Science in Data Science, Analytics and Engineering May 2026
Arizona State University, Tempe, Arizona *GPA: 3.52*

Bachelor of Engineering in Computer Engineering Jun 2024
Savitribai Phule Pune University (University of Pune), Maharashtra, India *GPA: 3.47*

TECHNICAL SKILLS

Programming Languages: Python, Javascript, SQL, C++, HTML, CSS
Tools and Frameworks: Django, MySQL, Snowflake, Databricks, Apache Airflow, Hadoop, Spark, Amazon Web Services (Lambda, S3, EC2), React, MongoDB, Tableau, Tensorflow, Docker, PowerBI, Streamlit, Gradio
Concepts: Object oriented programming, Data Structures and Algorithms, Operating Systems, Machine Learning, Data Processing at Scale, Distributed Data Systems, Web Development, REST API, Natural Language Processing, LLM, Artificial Intelligence, Computer Vision, Data Visualization, Deep Learning

PROJECTS

F1RAGBot: Formula 1 Q&A Chatbot | LangChain.js, Next.js, OpenAI, Astra DB Jun 2025 – Jul 2025

- Engineered a **Retrieval-Augmented Generation (RAG)** chatbot using **LangChain.js**, **Next.js**, and **OpenAI APIs**, enabling real-time Q&A over structured Formula 1 datasets, reducing average query response time by 50%, and a 35% boost in user interaction
- Integrated **DataStax Astra DB** vector embeddings and semantic search pipelines, and operationalized scalable production deployment on **Vercel** with **CI/CD pipelines** and performance monitoring, achieving over 90% retrieval accuracy and improving system reliability by 40%.

Smart-Based Book Recommendation System | LLM, Gradio, LangChain, OpenAI, Chroma May 2025 – Jun 2025

- Developed an interactive book recommendation system using **Gradio**, processing over **7,000** book entries to deliver personalized suggestions based on semantic similarity, achieving **78% recommendation accuracy** and enhancing user experience.
- Integrated **LangChain** and **OpenAI embeddings** to build a vector search engine with **Chroma** and refined retrieval pipelines, reducing query processing time by **40%** compared to traditional keyword search methods.

Augmented Analytics Platform: Retail Sales Insight Dashboard | Streamlit, ydata-profiling, Python Apr 2025 – May 2025

- Built an interactive analytics dashboard using **Streamlit** to automate analysis of over **10,000** retail transactions, surfacing KPI cards and sales/profit trends, and enabling managers to make data-driven decisions.
- Designed interactive filters and real-time charts using **ydata-profiling** and custom **NLG templates**, uncovering **20+ hidden insights** from retail transactions and delivering the most impactful project of the quarter.

CrimeNet: Urban Crime Hotspot Detection with GNNs | Python, PyTorch, SHAP, Scikit-learn Jan 2025 – Mar 2025

- Constructed a crime hotspot prediction system using **500,000+ Phoenix PD incident reports**, achieving **84% accuracy** in classifying high-risk zones and informing law enforcement resource allocation to enhance community safety.
- Leveraged **graph neural networks** to model spatial dependencies between neighborhoods and applied **SHAP** for feature importance analysis, delivering interpretable insights that supported recommendations for optimized law enforcement resource allocation.

EXPERIENCE

Research Assistant Jan 2024 – Jun 2024
Dr. D. Y. Patil Unitech Society *Maharashtra, India*

- Forged a student data management platform on **AWS EC2** and **S3**, designing scalable database systems that streamlined research workflows and reduced maintenance time by **15 hours per week**.
- Built **ML models** to analyze trends in student performance data, improving decision-making accuracy by **20%**.

Engineering Intern Feb 2023 – Mar 2023
Bajaj Auto Ltd *Maharashtra, India*

- Pioneered a **SharePoint** and **Power Apps**–based inventory management system integrated with **Microsoft Cloud**, eliminating **80% of data entry errors** and accelerating asset retrieval times by **65%** across departments.
- Constructed financial performance models in **Excel**, integrating data from **five disparate sources** to identify **three key cost optimization areas**—delivered as the sole intern on the project.
- Engineered category performance dashboards using **Excel macros** and **Power Automate**, reducing manual data processing by **30%** and accelerating strategic decision-making for key stakeholders.