Sauham Vyas

405, Rajalaxmi West Village, 14 Mamtanagar, West Ring road, Indore, 452012
+91-7000202271 sauhamv28@gmail.com LinkedIn-https://www.linkedin.com/in/sauham-vyas/
GitHub https://github.com/Sauham https://sauham.github.io/Final-Portfolio/

EDUCATION

Sage University Indore, Dept-Institue of Advance Computing Indore **Sep 2022 - Sep 2026**
Bachelors for Technology (B.Tech), Artificial Intelligence (GPA: 8.14(till 5th Semester)) *Indore, M.P*

Alpine Academy **Apr 2021 - Jul 2022**
PCM (GPA: 80%) *Indore*

EXPERIENCE

Horeca Store | *AI Developer* **Apr 2025 - Present**

- Developed and deployed a bulk order recommendation system using customer purchase history, product categories, and seasonal trends to optimize B2B order volumes and inventory planning.
- Collaborated cross-functionally with product and engineering teams to integrate AI-driven recommendations seamlessly into the Horeca Store platform, improving client retention and upsell opportunities.
- Researched and implemented advanced machine learning techniques (e.g., collaborative filtering, RAG pipelines) to power various personalization, forecasting, and decision-support features across the platform.

AI CERTS | *Prompt Engineer* **Mar 2025 - Present**

- Engineered and optimized complex prompts for Large Language Models (LLMs) like GPT-4 and Claude, enhancing response accuracy and contextual relevance by 30% for AI-driven automation solutions.
- Developed and fine-tuned prompt pipelines using LangChain and LLamaIndex, integrating vector databases (FAISS/Pinecone) to improve retrieval-augmented generation (RAG) workflows for dynamic AI applications.
- Conducted systematic A/B testing and prompt evaluations across various use cases, refining zero-shot, few-shot, and chain-of-thought (CoT) prompting techniques, reducing AI hallucination rates by 25%.

SETV Global | *AI-ML Engineer* **Nov 2024 - Feb 2025**

- Implemented advanced computer vision techniques with CNN models to accurately detect tumors and Alzheimer's in MRI scans, significantly improving diagnostic outcomes.
- Generated detailed diagnostic reports by identifying sizes, locations, and abnormalities in brain diseases and third-trimester pregnancy complications, leveraging YOLO, PyTorch, TensorFlow, and OpenCV.
- Applied NLP techniques to analyze physicians' inputs, enhancing system understanding and boosting response accuracy by 20%.
- Optimized language model development efforts, resulting in a 30% increase in model performance and a 40% improvement in user satisfaction.

Outlier AI | *Prompt Engineer* **Oct 2024 - Jan 2025**

- Designed and optimized AI-generated responses by crafting effective prompts for various business applications.
- Conducted A/B testing on different prompt structures to enhance accuracy and coherence in AI model outputs.
- Worked closely with AI researchers and developers to fine-tune prompt strategies, improving contextual understanding and reducing biases.
- Created structured prompt libraries to standardize AI interactions across multiple domains.

Codsoft Infotech | *AI and Data Science Intern* **Jul 2024 - Sep 2024**

- Developed AI projects using Flask and CNN, enhancing project efficiency and accuracy.
- Developed automated ETL pipelines for structured and unstructured data, reducing data processing time and improving data accuracy for analysis
- Built predictive models to improve customer engagement strategies, leading to more personalized marketing efforts and increased customer retention

Cosmic 365 AI | *Web Developer Intern* **Mar 2024 - Jun 2024**

- Assisted in designing and developing the company's website using Figma and JavaScript, contributing to a user-friendly interface
- Developed front-end components using JavaScript, enhancing the website's responsiveness and visual appeal
- Collaborated with team members to integrate front-end and back-end functionalities, ensuring seamless website performance

Devaditya Technocrats LLP | *Operations and Data Analyst* **May 2023 - Feb 2024**

- Created interactive dashboards and reports using Matplotlib to visualize key business metrics and trends.
- Conducted in-depth data analysis using Python (Pandas, NumPy) and SQL to identify insights and optimize operational strategies.
- Managed data entry, performed financial modeling, automated reports using Excel formulas, and streamlined documentation processes.

SKILLS

- **Programming Languages**: Python, C++, Javascript, C, HTML, CSS, R, Scala
- **Libraries/ FRAMEWORKS**: Scikit-Learn, Numpy, Pandas, PyTorch, Keras, TensorFlow, Hugging Face, scipy, NLMK, Spacy, CNN, RNN, NLP, LLM, Computer Vision, YOLO, Machine Learning, Deep Learning, PySpark, MATLAB, MERN Stack, Tailwind CSS, BeautifulSoup, ReactJS, Matplotlib
- **Tools / Platforms**: Git, Github, VS Code, Chaturst, API, Restful API's, FAST APIs, Figma
- **Databases**: SQL, MongoDB, MySQL, AWS, Google Cloud, Azure

PROJECTS / OPEN-SOURCE

Document Chat Application using RAG

- Developed an AI-powered RAG-based Chatbot using LangChain, LlamaIndex, and OpenAI APIs, enabling efficient retrieval-augmented generation (RAG) for answering document-based queries.
- Implemented a vector database (ChromaDB) to store and retrieve document embeddings, optimizing search accuracy and response relevance.
- Built a scalable web interface using Streamlit, integrating user authentication and seamless document upload for interactive AI-driven conversations.

AI Recommendation System **Mar 2025 - Apr 2025**

- Built a modular AI Recommendation System leveraging collaborative filtering and content-based techniques to deliver personalized product suggestions across diverse user profiles and behaviors.
- Integrated Python-based data pipelines and machine learning models to process user interaction logs and product metadata, improving recommendation accuracy and relevance.
- Designed a scalable architecture using Flask for API deployment and ensured real-time response handling for seamless integration into eCommerce or SaaS environments.

Fetus Location and Organ Detection

- Designed an object detection pipeline to identify fetal organs using YOLOv8 and Roboflow datasets.
- Developed an automated data handling system to manage empty label files and organize training data efficiently.
- Optimized the detection model's performance using Ultralytics and OpenCV, enhancing accuracy and reducing misclassification.

Data Extraction and Sentiment Analysis

- Automated Web Data Retrieval: Developed a Python-based tool using BeautifulSoup to extract and process article text from multiple URLs.
- Sentiment Analysis Integration: Implemented TextBlob for sentiment analysis, computing polarity and subjectivity scores from extracted content.
- Data Processing & Storage: Structured extracted data using Pandas, saved results in Excel format, and optimized text cleaning with regex-based preprocessing.

CERTIFICATIONS

- **Advance Machine Learning**: https://verify.netcredential.com/roy8LbfRjD
- **Deep Learning**: https://verify.netcredential.com/roy8Pzl90d
- **AWS Cloud Foundation**: https://www.credly.com/badges/39ff689b-ec87-4443-87f4-0ee916cc245b/linked_in?t=rwcsoa
- **Machine Learning**: https://verify.netcredential.com/roy8Ui0SIi
- **Technology for PM**: https://hellopm.co/app/api/certi/S2CIXP