Saibala Sundram

Machine Learning & AI Engineer

Bangalore, KA | saibalasundram1@gmail.com |+91 8073022760 | linkedin.com/in/saibala-sundram-0a1931183/

SKILLS

LANGUAGES: Python, JavaScript
DATABASE: MongoDB, PostgreSQL, Pinecone, Chroma
LLM APPLICATION: Ollama, LangGraph, LamaIndex, Hugging Face Transformers
EDA: Superset, Pandas, Tableau, Bokeh, Plotly, Seaborn, Grafana, Streamlit
ETL: Apache Airflow, Flow Fuse
TRAINING FRAMEWORK: TensorFlow, Pytorch, Scikit, Keras
ML OPS: MLFlow, Weights & Biases
PACKAGING & DEPLOYMENT, Docker Container, Hyper-V

EXPERIENCE

Cymbeline Innovation (Collaborating with Panasonic India Innovation Center)| *Machine Learning & AI Engineer*
Bangalore, KA | **06/2023 – Present**

- Was selected from a startup challenge to collaborate with Panasonic India Innovation Center on the Miraie smart factory solution, recognized for exemplary software engineering skills and extensive domain knowledge. This collaboration is ongoing and continues to leverage cutting-edge innovations in AI and smart factory solutions.
- Researched, developed, and deployed on-prem LLMs applications using production-grade servers, enhancing system implementation capabilities.
- Implemented AI models (Gemini, Phi3, Llama3) to enhance user accessibility and sustainability via domain-specific chatbot solutions, employing advanced artificial intelligence techniques.
- Led the development, staging, and production deployment of agentic chat solutions using the LanGraph framework, focusing on robust DevOps practices.
- Developed predictive models using TensorFlow and PyTorch to analyze cycle times and historical data, enabling accurate MTTF and MTTR predictions, thereby reducing machine breakdowns by 15%.
- Trained machine learning models with Scikit-Learn and Keras for production trend analysis and count forecasting, improving quality assurance measures and forecast accuracy by 25%.
- Deployed predictive maintenance models through MLFlow, enabling continuous monitoring of model performance and facilitating rapid iterations based on real-time machine data.
- Tracked model experiments and hyperparameters with Weights & Biases, optimizing models for production trend analysis and ensuring scalable deployment for count forecasting solutions, incorporating best practices in CI/CD.
- Optimized CI/CD pipelines, reducing process time by 50% with ML Ops frameworks (ML Flow, WB) and Docker for Ubuntu and Hyper-V for Windows builds.
- Upgraded backend systems to microlithic architecture with flow-based programming tools (Flowise), improving deployment speed and AI solution integration, enhancing software design and agile development methodologies.
- Spearheaded a team of 7 in deploying and conducting UAT in a live production environment, collaborating closely with DevOps and field engineers, showcasing leadership in Scrum and product engineering.

Cymbeline Innovation | *IoT Engineer* *Bangalore, KA* | ***10/2021 – 05/2023***

- Collaborated with Tier 1 manufacturing plants across India, identifying key machinery pain points through cross-functional analysis and enhancing business technology insights.
- Resolved production quality issues by implementing real-time machine data acquisition and streaming using MQTT and Kafka, improving operational efficiency.

- Upgraded backend systems to a microlithic architecture with Flow Fuse, reducing deployment time and enabling real-time adjustments for IoT systems using cloud-native strategies.
- Availed one year of AWS Cloud support through the startup program, implemented GitHub best practices, and led feature development for DataAct and TabAct, enhancing version control across AWS, GCP, and Azure.
- Integrated SAP functionalities into the backend for real-time production tracking and reporting, utilizing JavaScript modules for shopfloor insights.
- Improved security by integrating OAuth and JWT authentication, ensuring compliance with best practices, and securing deployments with NGINX reverse proxy.
- Deployed Node.js applications using Docker and automated scaling, ensuring consistent performance and demonstrating strong DevOps and CI/CD methodologies.

ISRO | *Data Analyst Apprenticeship at Laboratory for Electro Optics Systems* *Bangalore, Ka* | **03/2021 – 07 /2021**

- Developed an automated program and GUI for real-time qualitative analysis of laser-induced plasma spectra, utilizing Python and MATLAB with advanced data visualization libraries, optimizing software design and application development.
- Conducted exhaustive literature reviews and applied statistical techniques to optimize the analysis of multi-element spectral data under low-pressure conditions, enhancing computer science methodologies.
- Contributed to and published research in a high-performance conference, IEEE CONNECT 2021, with the paper titled Denoising and Baseline Correction Prospects of Laser Induced Plasma Spectra Acquired Under Low Pressure Conditions: Comparison of Algorithms and Optimization Perspectives, featured in IEEE Xplore, demonstrating scientific research capabilities.
- Collaborated with senior engineers and scientists, enhancing the accuracy and efficiency of spectral analysis, driving significant research outcomes for the institution and bolstering quality assurance and product engineering processes.
- Continued to support the tool's development with advanced features even after the completion of the apprenticeship. This tool was successfully utilized in the Chandrayaan 3 mission, a proud milestone in my career, demonstrating the long-term impact and relevance of my work.
- Translated critical software components from Java to Python, improving the system's maintainability and performance while integrating software engineering best practices and leveraging coding development skills.

PROJECTS & HACKATHONS

Project| *Optimizing LLaMA-2 Fine-Tuning on Google Colab*

- Description: Developed an optimized process for fine-tuning LLaMA-2 models on Google Colab , focusing on efficient resource utilization and improved training times.
- Article: Optimizing Llama-2 Fine-Tuning in Google Colab for Efficient GPU Usage

Project| *Simplifying NoSQL Queries Using Natural Language*

- Description: Created a framework for simplifying NoSQL queries by translating natural language inputs into query operations by using T5 foundational transformer encoder decoder model.
- Article: Simplifying NoSQL Queries Using Natural Language

Product| *Arachnospec - Chrome Extension (Google DevPost Generative AI Hackathon)*

- Description: Spearheaded the development of Arachnospec, an innovative Chrome extension that augments YouTube viewing with an AI-powered chat interface. Engineered with React, Flask, and the Gemini AI Pro model, Arachnospec facilitates application development by generating timestamped transcripts and responding to user queries based on video content.
- Key Contributions:
    - Led a team of six, all dynamic members of the AI Tamil Nadu community, to develop a pioneering Chrome extension for YouTube, demonstrating leadership in agile development methodologies.

- Employed Gemini AI Pro for precise transcript generation and content-driven user queries, showcasing artificial intelligence expertise.
- Constructed using React and Flask, ensuring a seamless and responsive user experience on YouTube, highlighting skills in software engineering.
- Overcame GCP technical challenges, including API limitations and Google Cloud intricacies, to deliver a functional MVP, demonstrating problem-solving skills and DevOps capabilities.
- Promoted AI learning and collaboration within the team, reinforcing a culture of innovation and active community engagement, pivotal in team management and community building.
- LinkedIn Post: Arachnospec

Project| *Tamil Dataset Creation Tool (AI Tamilnadu Community Project)*

- Description: Contributed to the AI community by performing custom Tamil language articles web scraping and tokenization, enhancing the data science initiatives within the AI Tamil Nadu community.
- Hugging Face Link: AI Tamilnadu

Project| *MTech Final year dissertation*

- Advancing Manufacturing Efficiency: Integrative Data Analytics and Machine Learning for Predictive Maintenance.

EDUCATION & CERTIFICATES

Education	Degree/Certification	Dates
PM Accelerator	Product Manager Accelerator - AI PM Bootcamp and AI Engineering Certification Program	01/7/2024 to 15/9/2024.
WILP BITS, Pilani, Rajasthan	M.Tech in Data Science	12/11/2024
M.S. Ramaiah University of Applied Sciences, Bangalore	B.Tech in Electronics and Communication	12/08/2021
Kendriya Vidyalaya ASC, Bangalore	12th, CBSE	29/05/2018
Kendriya Vidyalaya AFS, Kasauli	10th, CBSE	29/05/2015

Certifications & Achievements	Details	Dates
Code Vipasana Season 6	Data to AI with Gemini & Gemma	30/07/2024
IBM Skills Network	Build a chatbot to analyze PDF with LLM	28/02/2024
GUVI Geek Networks, IITM Research Park	AI for India, The Guinness World Record Event	31/05/2021
GUVI Geek Networks, IITM Research Park	Build a Face Recognition Application using Python	30/04/2021
IEEE CONNECT 2021	Denoising and Baseline Correction Prospects	31/05/2021
Coursera	Time Series Analysis and Prediction	30/06/2020
Smart India Hackathon	6th position at university level	01/02/2020 - 31/03/2020
NPTEL, IIT Madras	Machine Learning Using TensorFlow	01/01/2020 - 31/03/2020
NPTEL, IIT Madras	Python Data Structures and Algorithms	01/01/2019 - 31/03/2019

EVENTS & COHORTS

- **Active Member of AI Tamil Nadu & Cohere Community**: Currently engaged in the AI Alignment Cohort to stay abreast of the latest developments in emerging AI technology solutions.

DEVELOPER PROFILE & LINKS

- Github: https://github.com/saibala2905
- GDG : https://g.dev/saibalasundram , Hugging Face: https://huggingface.co/saibala29