



Om Prakash Jakhar

Lead Data Scientist

✉ opjakhar95@gmail.com ☎ +91 8078656085 📍 Bengaluru, India

🌐 opjakhar 📺 op_jakhar 🇮🇳 Indian

👤 PROFILE

Lead Data Scientist with expertise in Generative AI, LLMs, Multi-Agent RAG, and Model Context Protocol (MCP), with a strong track record of building and deploying intelligent, context-aware systems at scale. I bring deep expertise in NLP, deep learning, and multimodal data (text, image, video), combined with robust engineering skills across Docker, CI/CD, and cloud platforms like AWS and GCP. From architecting scalable AI pipelines to leading high-performing teams, I focus on transforming cutting-edge AI research into real-world applications that automate decisions, enhance user experiences, and drive measurable business outcomes.

🎓 EDUCATION

2017 – 2019
Chennai, India

Master Degree (Computer Science and Engineering)

Indian Institute of Technology Madras (IIT Madras) ☑

Deep Learning, • Natural Language Processing, • Machine Learning, • Computer vision, • Linear Algebra

Teaching Assistant: Advanced Programing Lab

2012 – 2016
Raipur, India

Bachelor Degree (Computer Science and Engineering)

National Institute of Technology Raipur (NIT Raipur) ☑

Artificial Intelligence & Expert System • DataBase Management System(DBMS) • Operating System • Programming Languages

👜 PROFESSIONAL EXPERIENCE

04/2023 – present
Bengaluru, India

Cyble Infosec India Pvt Ltd ☑

Lead Data Scientist

In my role, I design and implement advanced algorithms and models for detecting and mitigating security threats, consistently integrating the latest advancements in AI to enhance our offerings. This includes leveraging LLM capabilities and Multi-Agent Systems (RAG), along with Model Context Protocol (MCP), to accelerate product development and fuel innovation. Whether developing intelligent systems or creating sophisticated models, I ensure that our solutions are both cutting-edge and highly effective in safeguarding against emerging threats.

1. **Automated Threat Reporting:** Built LLM-powered system to generate country-, sector-, and time-based threat reports, cutting generation time from **hours to under 10 minutes** and boosting efficiency.

2. **Blaze AI Layering:** Designed AI layer to reduce false positives, cut alert noise, and streamline threat identification for improved customer experience.

3. **Threat Advisory Automation:** Built AI-driven system to auto-generate threat reports from incident blogs, reducing manual research and enabling instant report delivery.

4. **In-house Logo Detection:** Built advanced image analysis tool to detect logos in complex backgrounds, boosting visual data processing without keyword reliance.

5. **AI Assistant (MutliAgent RAG Based Co-pilot):** Developed a multi-agent CTI co-pilot using RAG to deliver real-time, knowledge-based responses, enhancing customer interaction and query resolution.

6. **Unstructured Document Processing:** Developed system using LLMs to identify document types and extract sensitive info, improving data processing speed and accuracy.

7. **Alert Management System:** Streamlined alert categorization, grouping, and prioritization, reducing workload and focusing on high-severity threats by filtering out false positives.

2022 – 03/2023
Bengaluru, India

Cyble Infosec India Pvt Ltd 

Senior Data Scientist

Developed AI solutions for identifying cyber threats, sensitive data leaks, and phishing domains. Expertise in algorithm development, performance evaluation, solution deployment, load testing, and containerization.

1. **Logo Identification:** Developed and deployed a high-accuracy B&W and color logo detection tool, enhancing brand monitoring in production.
2. **File Config & Header Identification:** Led data prep and model training for a semi-automated solution, achieving **84% accuracy** and reducing manual header tagging efforts.
3. **Noisy Alert Reduction:** Reduced false positives by **60–90%** for GitHub services through optimized keyword matching and an API powered by ML and LLMs to filter suspicious and phishing domain alerts.
4. **Executive Impersonation (Face Recognition):** Deployed a face recognition model with **92%** accuracy to detect and prevent executive impersonation threats.
5. **Advanced Threat Detection (LLM):** Leveraged LLMs to boost threat detection and prevention capabilities.

07/2019 – 07/2022
Bengaluru, India

SAP LAB INDIA PVT LTD 

Data Scientist

Part of a team dealing with industrial problems related to Artificial Intelligence and Machine learning/Deep Learning. Working on various projects and proof of concepts that involved the domains of **deep learning, computer vision, natural language processing(NLP)**, Forecasting and optimisation. Helping customer to find the requirements that best fits their unique product and infrastructure, and producing high-quality results that meet their objectives.

1. **Intelligent Invoice Processing:** Designed end-to-end solution using deep learning to extract, classify, and predict key information, automating the P2P process in SAP Data Intelligence.
2. **Relief Request Classification System** :Developed system to classify relief requests (healthcare, food, travel) by analyzing and categorizing raw text from incoming requests using Language Models and NLP techniques.
3. **Optimization & Analysis Module:** Developed module to prioritize work assignments based on skill matching and implemented analysis tool to assess the impact of configuration changes, aiding decision-making with As-Is and To-Be scenario comparisons.
4. **Demand-Driven MRP (SAP S/4 HANA):** Proposed and integrated MRP solution for optimal stock management, predicting order dates and quantities using mathematical modeling and time series forecasting techniques.
5. **Cloud Migration & Custom Functionality:** Developed caching module (Postgres, Serialization, Locking) as part of on-premise to cloud migration, deployed as Docker image on SAP Cloud with Kyma runtime.

05/2015 – 07/2015
Jodhpur, India

Defence Research and Development Organisation(DRDO) 

Trainee

Project Involvement  : *Printer Automation Module & Library Management System*

- **Printer Automation Module:** Developed a system to automate the generation and printing of stock verification sheets, eliminating the need for manual intervention.
- **Library Management System:** Contributed to managing DRDO's extensive library database. Focused on SQL query processing for tracking issued and available records, ensuring efficient data handling and retrieval.

SKILLS

GenAI & Advanced ML

- LLMs (Large Language Models)
- RAG | Agents | MCP
- Hugging Face
- LangChain
- Vector Databases

● ● ● ● ●

Classical & Deep Learning

- Deep Learning
- NLP (Natural Language Processing)
- Computer Vision(Image & Video Processing)
- Time Series Forecasting
- Classical Machine Learning(Regression, XGboost, random Forest etc)

● ● ● ● ●

Programming & Frameworks

- Python | C/C++ | Java | GoLang
- SQL
- PyTorch
- TensorFlow
- FastAPI
- Streamlit

● ● ● ● ●

MLOps & Deployment

- Docker
- Kubernetes
- MLflow
- AWS/GCP
- Model Monitoring
- Microservices/REST APIs
- S3, EC2, SageMaker, Bedrock

● ● ● ● ●

Version Control & Collaboration

- GitHub
- Bitbucket

● ● ● ● ●

AWARDS

SAP Innovation Service & Solution Honours(Outstanding Performance Award)) [↗](#)

SAP Labs India Pvt Ltd

Recognized for exceptional execution and delivery of a high-impact project under tight deadlines, achieving flawless implementation.

Performance Appreciation Award

SAP Labs India Pvt Ltd

Multiple-time recipient of Appreciation Awards for excellence in performance

CERTIFICATES

- Natural Language Processing(NLP) [↗](#)

- Python Programming [↗](#)

- SQL [↗](#)

- Problem Solving (Basic) Certificate [↗](#)