Prathamesh Kulkarni

Pune, India | +919404120428 | kprathamesh67@gmail.com | https://www.linkedin.com/in/kprathamesh

PROFILE SUMMARY

Senior AI Engineer with **2+ years** of experience providing end-to-end ML solutions, spanning forecasting, fraud detection, and document processing across industries. Skilled in feature engineering, model optimization, backend architecture, and microservices, driving both operational efficiency and data infrastructure improvements with strategic insight and technical autonomy.

SKILLS

- **AI/ML**: Python, SQL, Nixtla, Pandas, NumPy, Sklearn, TensorFlow, PyTorch, MLflow, Time Series Forecasting, Computer Vision, NLP,
- **LLMs**: Langchain, CrewAI, Vector Stores, LLM Agent Development, Bedrock, Azure OpenAI, Ollama, VLLM, Hugginface
- **Data Visualization and Front-End**: Matplotlib, Seaborn, Plotly, Dash
- **Cloud & DevOps**: AWS (Advanced), Azure & GCP (Intermediate), Docker, Jenkins, CI/CD Pipelines
- **Project Management and Collaboration**: Agile Methodology, Cross-functional Team Leadership, Client Engagement, Training Delivery

PROFESSIONAL EXPERIENCE

bizAmica Software   Pune, India

Senior AI Engineer   Feb 2023 - Present

- Engineered distributed, GPU-accelerated AI/ML solutions capable of processing **200,000+ documents daily**, achieving an **87% reduction in model inference** time through advanced computational optimization.
- Pioneered **LLM-driven solutions** across logistics and banking, developing **four new product offerings** that expanded the company's technical capabilities and client services.
- Established strategic leadership in AI technologies, becoming the primary point of contact for LLM expertise after developing domain knowledge, directly interfacing with executive leadership including VPs and CFOs.
- Delivered **product demonstrations** and developed **pricing models,** gathering **technical requirements** to align solutions with client needs.
- Implemented robust DevOps infrastructure using Ansible, Docker, and Jenkins, streamlining deployment processes and **reducing deployment times from 5 to 1 day.**
- Conducted comprehensive **technical training for 25+ employees and interns**, fostering knowledge sharing across API development, AWS, LLMs, and DevOps methodologies.
- Managed **cloud and on-premises infrastructures** with rigorous security protocols, ensuring data integrity in sensitive banking environments.

Sky Betting and Gaming   *Leeds, United Kingdom*

Data Scientist   *Jun 2022 − Oct 2022*

- Developed an end-to-end machine learning solution for bonus abuse fraud detection, processing approximately **100,000 monthly transactions with 89-95% accuracy** in production environments.
- Engineered **comprehensive data pipelines** for data preprocessing, feature selection, and model development, ensuring robust and scalable fraud detection infrastructure.
- Deployed **production-ready models** using Docker and Apache Airflow, creating automated workflows that streamlined the fraud detection process and reduced manual intervention for the risk management team.
- Collaborated closely with domain experts and senior data scientists to refine model performance, **translating complex technical insights into actionable business strategies.**

- Created **custom prediction threshold optimization techniques** that aligned technical outcomes with strategic business priorities, significantly enhancing fraud detection capabilities.
- Proactively developed and presented a strategic business case for a novel data feature, demonstrating through **comparative modelling** that the proposed enhancement could potentially increase fraud detection accuracy to **96-98%.**

PROJECTS

- **Air Cargo Tonnage Prediction:** Led the development of data-driven forecasting models for air cargo tonnage, addressing multiple levels of granularity (airport, cost centre, customer, and commodity) to enhance decision-making and resource allocation in logistics. This solution reduced the time required for the financial team to generate forecasts from **6 weeks to just a few minutes**, significantly improving efficiency and decision-making.
- **Logistics Chatbot for Airport and Dock Management:** Created chatbot to provide **real-time and up-to-date information** on airport facilities and docks, improving cargo management for freight forwarders and airport authorities.
- **High-Volume FASTag Processing for Banks:** Architected and deployed AWS infrastructure for processing **14,000 FASTags daily** and managing a backlog of 7.5 million FASTags. The architecture handled networking, load balancing, scaling, security, and deployment to support the new KYV (Know Your Vehicle) process.
- **HR, Legal, and RFP Chatbots:** Engineered LLM-powered chatbots using **Retrieval-Augmented Generation (RAG)** for efficient document retrieval, content generation, and updates. Enhanced operational efficiency across HR, Legal, and RFP teams.
- **LLM-Powered IVR System for Logistics Client:** Implemented an IVR system integrated with LLMs to automate routine customer inquiries, **reducing support call volume by 30-40%** and enhancing service delivery.
- **izDox AI Platform Enhancements:** Optimized document processing pipelines in the izDOX platform, achieving a **25% improvement in extraction accuracy** by integrating LLMs for financial and logistics document processing.
- **Fault Detection using Computer Vision in Assembly Lines:** Created an advanced fault detection system using computer vision and pattern matching to identify silent faults in PLC programming, resulting in a **96% improvement** in efficiency and minimising assembly line disruptions in the automotive industry.

EDUCATION

Degree	University	Duration
Master of Science in Data Science	Lancaster University, Lancaster, United Kingdom	Oct 2021 - Sept 2022
Bachelor of Engineering in Electronics and Telecommunication	Sinhgad Institute of Technology and Science, Pune, India	Aug 2016 - May 2020

LANGUAGES

English(C2), Hindi(C1), Marathi(Native), French(A2)