Amarsh Pedapati

Data Scientist | AI Engineer | LLMs & Computer Vision

 amarshp | amarsh.pedapati | amarshp | amarsh.pedapati@gmail.com | +91-9959822444

EDUCATION

Indian Institute of Technology, Hyderabad   Jul 2019 – May 2023
*Bachelor of Technology, **Artificial Intelligence**   (GPA - 8.34/10)*   *Hyderabad, India*

WORK EXPERIENCE

Data Scientist | Blend360, Hyderabad, India   Oct 2023 – Present

For Walmart

- Led end-to-end customer segmentation and uplift modeling for a large-scale holiday sales campaign, designing incrementality tests (A/B testing) across multiple customer segments.
- Built a high-performance Spark datamart ($7\times$ faster) and engineered ETL/preprocessing pipelines with SQL optimization and Python data engineering, achieving runtimes over $1000\times$ faster.
- Improved model lift by 100%, applied probability calibration to correct biases from prior experiments, and enabled statistically significant targeting of 13.4M customers from a 130M population through power analysis.

For VISA

- Developed an explainable AI web app in Streamlit enabling clients to explore model behavior via trendline discrepancy analysis, residual-based Shapley explainability, and forecast comparison.
- Engineered advanced LangChain reduction pipelines (multi-core custom implementation) to summarize discrepancies with LLMs, generating both technical insights and executive-level summaries.
- Integrated an LLM-powered chatbot with Agents, combining RAG (ChromaDB) for discrepancy-focused Q&A and data exploration via PandasAI, improving model interpretability and client trust.

Data Scientist Intern | AIBOD, Fukuoka, Japan   Jan 2022 – June 2022

- Enhanced reliability of an AI-powered object detection system for unmanned cash registers by resolving out-of-distribution (OOD) misclassification issues, which were a major blocker for retail automation deployment.
- Optimized classification thresholds through linear, polynomial, ensemble-based, and ML-driven calibration methods (logistic regression, gradient boosting), enhancing confidence scoring.
- Applied advanced data augmentation (synthetic generation, domain randomization, adversarial noise) to boost robustness across diverse real-world conditions.
- Integrated an OOD detection pipeline with uncertainty estimation and Mahalanobis distance, reducing false positives and achieving a 32% drop in misclassification rate with an 18% accuracy gain.

Data Scientist Intern | ManaliSwing, Manali, India   Sept 2023 – Dec 2023

- Designed and deployed a real-time Computer Vision–based safety verification system for adventure sports, detecting carabiner open/closed status to prevent safety failures.
- Built a diverse dataset across challenging environmental conditions, benchmarked state-of-the-art object detection models, and implemented YOLOv5 with live GoPro video streaming optimized for low-latency inference.
- Delivered high reliability in safety-critical detection (95% accuracy, 92% precision, 93% recall), and deployed the model in production at a commercial adventure facility in Manali, significantly strengthening operational safety and compliance standards.

TECHNICAL SKILLS

Programming & AI: Python, C++, SQL | Machine Learning, Deep Learning | LLMs, Generative AI (RAG, Fine-tuning, Prompt Engineering, Agents) | Data Engineering

Frameworks & Libraries: PyTorch, Fastai, Hugging Face, LangChain, Streamlit

Data Science & Visualization: Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, Plotly | Tableau, Power BI | Dashboard Design & Data Visualization Best Practices

Cloud & MLOps: AWS (SageMaker, Lambda), Azure ML, Databricks, MLflow, Docker, Kubernetes

Databases & Storage: Snowflake, PostgreSQL, MySQL | Vector DBs (FAISS, Pinecone, Weaviate)

Certifications: AWS Certified AI Practitioner, Oracle Generative AI Professional