Mayur Kishor Kumar

Oakland, CA | Open to Relocation | P: +1 857.296.2366 | mayurakash1999@gmail.com | LinkedIn | Github | Medium | Portfolio

EDUCATION

NORTHEASTERN UNIVERSITY May 2024
Master of Science in Artificial Intelligence Boston, MA

DAYANANDA SAGAR COLLEGE OF ENGINEERING June 2022
Bachelor of Engineering in Computer Science Bangalore, India

SKILLS

Languages & Web: Python, SQL, Java, JavaScript (Node.js, React), C/C++, HTML/CSS
AI/ML: PyTorch, TensorFlow, Scikit-Learn, XGBoost, Hugging Face, NLTK, SpaCy, OpenCV
Cloud & DevOps: AWS, GCP(Google Cloud), Docker, Kubernetes, Git, CI/CD, FastAPI, Flask, Django
Data & MLOps: Apache Spark, Airflow, MLflow, Snowflake, Tableau, Matplotlib, Seaborn, Plotly

WORK EXPERIENCE

AFTERQUERY EXPERTS March 2025 – Present
Software Engineer San Francisco, CA

- Diagnosed 20+ critical bugs across Numpy, PyTorch, Pandas by navigating complex repos and implementing targeted patches.
- Constructed end-to-end test suites that increase code coverage from 68% to 92%, ensuring robust integration with CI pipelines.
- Orchestrated Dockerized environments to reduce onboarding time by 40%, enabling seamless cross-platform development.
- Collaborated with OSS maintainers to deliver high-impact features adopted across 3 core libraries with minimal regression.

HUMANITARIANS AI September 2024 – Present
AI Engineer and Researcher Boston, MA

- Boosted model training by 1.5× by reengineering ETL pipelines processing over 10TB of multimodal data using Spark and Pandas.
- Minimized batch processing delay from 3 hours to 10 minutes by deploying Kafka streams and Airflow-based orchestration.
- Automated data validation workflows that flagged 60% of edge-case labeling errors, reducing manual QA dependencies.
- Conducted large-scale hallucination benchmarking across 1M+ LLM outputs, guiding dataset refinement and tuning decisions.

IG GROUP August 2021 – November 2021
ML Intern Bangalore, India

- Reduced monthly fraud impact by ₹12L by applying XGBoost ensembles on high-frequency transactional datasets.
- Built a real-time trading insight engine parsing 500+ events/min, improving trader decision speed by 25%.
- Enhanced internal tool discoverability by 45% using custom transformer-based search optimization.
- Deployed automated Airflow DAGs for data refresh, saving 10 hours/week in manual pipeline management.

INDIAN SPACE RESEARCH ORGANIZATION(ISRO) March 2021 – June 2021
Project Trainee Bangalore, India

- Forecasted server overheating risks with ML models trained on 10 years of telemetry, ensuring zero downtime during missions.
- Designed a D3.js-based dashboard reducing EMS access time from 120s to under 30s across operational teams.
- Implemented a real-time alerting system in Node.js, accelerating incident response by 30% for remote monitoring.
- Partnered with systems engineers to embed predictive logic into legacy infrastructure without disrupting workflows.

PROJECTS

AUTOMATED STOCK FORECASTING PIPELINE WITH AWS AND AIRFLOW

- Predicted stock prices with 93% accuracy by training a CNN-LSTM model on historical data using AWS Glue and S3 pipelines.
- Eliminated 10+ hours/week of manual work by automating 8+ tasks with Apache Airflow and QuickSight dashboards.

BRAIN TUMOR CLASSIFICATION AND SEGMENTATION PIPELINE

- Processed 5K+ MRI scans to classify tumors with ResNet50 and segment regions using ResUNet with 98% pixel accuracy.
- Enhanced segmentation precision by optimizing Tversky loss and added visual interpretability using Grad-CAM heatmaps.

AI-POWERED DOCUMENT QUERY SYSTEM WITH VECTOR DATABASE INTEGRATION

- Designed an LLM-based semantic search tool using Dolphin-Mistral, GenAI embeddings, and Chroma across 1,200+ documents.
- Accelerated query speed from 3.5s to <1s by streamlining embedding pipelines and recursive text splitting in Streamlit.

MULTI-HOP REASONING AGENT FOR BUSINESS INTELLIGENCE

- Engineered a reasoning agent with Groq API + Ollama to analyze SQL, PDFs, and APIs in parallel using LLaMA 3-70B.
- Extracted insights across 4+ data streams by chaining LLM outputs into subqueries, reducing BI response time to <10s.