Kartik Raut

Boston, MA | +1 857-465-9236 | raut.kar@northeastern.edu | linkedin.com/in/kartik-raut/ | github.com/kartikraut98 | Portfolio

Summary

Data Scientist with over three years of experience specializing in Machine Learning, NLP, and Generative AI. Successfully designed and deployed end-to-end Machine Learning Operations (MLOps) pipelines, delivering scalable AI solutions that drive measurable business outcomes. Demonstrated leadership abilities through mentoring junior data scientists and ML engineers. Available for full-time opportunities starting May 2025 or later.

Technical Skills

Languages: Python
Frameworks: NumPy, Pandas, Matplotlib, Scikit-learn, PyTorch, Tensorflow, NLTK, HuggingFace, Langchain, Streamlit, Hadoop
Databases: SQL, NoSQL, MySQL, MongoDB, PostgreSQL
Tools: Tableau, Git, Docker, Kubernetes, Google Cloud Platform(GCP), Amazon Web Services(AWS), MlFlow, Airflow, DVC, Spark

Experience

Accenture **Dec 2020 – Jun 2023**
Data Scientist *Pune, MH*

- Engineered a real-time anomaly detection system for Vodafone UK using Isolation Forest, Autoencoders, and Bayesian Changepoint Detection, reducing false alarms by 40% and Mean Time to Response (MTTR) by 10% through automated root cause analysis
- Formulated a churn prediction model with CatBoost and Focal Loss, improving precision to 82% and recall to 75%, enabling A/B-tested retention strategies that cut churn by 15% QoQ
- Implemented 4G/5G network monitoring system using Bayesian Changepoint Detection, Grafana, and InfluxDB that highlighted bottlenecks and anomalies in network logs, resulting in significant improvements and reducing customer complaints by 30%
- Architected an MLOps pipeline with Docker, Jenkins, and MLflow, integrating Evidently AI for drift detection that triggered automated retraining, reducing manual validation effort by 40% through automated testing
- Crafted a customer segmentation framework using K-means for behavioral grouping and DBSCAN for anomaly detection, improving retention targeting and increasing campaign effectiveness by 40%

Projects

ProdGenius AI | *GCP, Airflow, Mlflow, Jenkins, Docker, TensorFlow, DVC* | GitHub **Sept 2024 – Dec 2024**

- Developed and deployed a multi-agent chatbot powered by LlaMA 3.1 models, enabling intelligent product Q&A by leveraging FAISS for vector search and metadata summarization
- Constructed and implemented a complete MLOps pipeline on GCP, automating data ingestion, model training, and deployment to streamline the development workflow
- Established scalable CI/CD and CT/CM pipelines using GitHub Actions and Cloud Run, enabling real-time model retraining and deployment with rollback mechanisms via Artifact Registry

Article Summarizer | *OpenAI GPT 3.5, RAG, Prompt Engineering* | GitHub **Jan 2024 – Apr 2024**

- Created a Streamlit app for article summarization powered by OpenAI's GPT-3.5, processing 1,000+ articles
- Implemented FAISS-based vector search for fast retrieval of 500+ documents, achieving sub-50ms latency
- Attained ROUGE-1: 40, ROUGE-2: 20, ROUGE-L: 35, and BLEU: 2, ensuring effective summarization

AdVision | *BART, Computer Vision, NLP, HuggingFace Transformers, Pytesseract* | GitHub **Jan 2024 – Feb 2024**

- Pioneered an AI-powered advertisement censorship system analyzing video, audio, and transcripts for policy violations
- Built a BART-based model for image classification and video analysis, processing video frames with Vision Transformer and OCR
- Integrated an ML pipeline combining text, audio, and visual analysis, summarizing 10+ hours of transcripts with Facebook MNLI and evaluating ads on 22 criteria

Volunteering Experience

Open Data Science Conference East **Apr 2024 – Apr 2024**
Team Assistant *Boston, MA*

- Supported technical sessions on LLMs, computer vision, and ethical AI, assisting attendees and speakers while gaining exposure to cutting-edge data science and AI topics presented by industry experts from leading tech companies

Education

Northeastern University **Sep 2023 – May 2025 (Expected)**
Master of Science in Data Analytics Engineering — GPA: 3.8/4.0 *Boston, MA*

- **Coursework:** Data Mining, Machine Learning Algorithms, MLOps, Neural Networks and Deep Learning, Natural Language Processing, Statistical Modeling, Operations Research, Product Design & Management, Computation and Data Visualization
- Currently serving as Teaching Assistant for IE 7275 Data Mining in Engineering (Spring 2025)

Nagpur University **Aug 2016 – Oct 2020**
Bachelor of Engineering in Electronics and Telecommunications *Nagpur, India*