Akash Pawar

Jersey City, New Jersey

Phone: +1 (929) 724-7345 | **Email**: akashpawar9619@gmail.com

LinkedIn: linkedin.com/in/akash-pawar | **Portfolio**: https://akashsp7.github.io | **GitHub**: github.com/akashsp7

Summary

Dynamic Machine Learning professional with a passion for transforming complex data challenges into actionable insights. Excels at developing innovative AI solutions across diverse domains while maintaining exceptional accuracy and performance. Combines technical expertise with strategic thinking to deliver production-ready applications that drive meaningful results. Adept at working independently and collaborating across teams to break down complex concepts for diverse audiences.

Technical Skills

Python, SQL, pandas, numpy, seaborn, Tableau, Plotly, matplotlib, NetworkX, Neo4j, PIL, DuckDB, Dask, OpenCV, Streamlit | PyTorch, TensorFlow, Keras, scikit-learn, XGBoost, LightGBM, Hugging Face, stable-baselines, Wandb, MLflow | OpenAI, LangChain, Unsloth, NLTK, Chroma, spaCy | Docker, AWS (S3, Lambda, SageMaker), Flask, Git, Gradio, and FastAPI

Projects

Protein Subcellular Localization Predictor using ESM2 | GitHub

Developed a high-accuracy protein localization predictor achieving 84.79% top-3 and 92.09% top-5 accuracy across 12 cellular locations.

Designed an end-to-end pipeline from data acquisition via UniProt's API to model deployment, optimizing with gradient checkpointing and mixed precision.

Tools/languages used: Python, PyTorch, ESM2, Mixed Precision Training, UniProt API

Deepseek for Advanced Mathematical Reasoning | GitHub

Fine-tuned DeepSeek-R1-Distill-Qwen-1.5B for enhanced mathematical reasoning, outperforming Claude-3.5 Sonnet on specific reasoning tasks.

Optimized training pipeline using Unsloth framework and LoRA adaptation, achieving 2x faster inference while reducing parameters by 98.8%.

Tools/languages used: PyTorch, Unsloth, LoRA, Pandas, Hugging Face

Dependency Chain Analysis | GitHub

Processed 500,000+ relationships across 442,275 nodes in a large-scale dependency graph, extracting key structural patterns in under 25 seconds.

Implemented Node2Vec embeddings with custom features, achieving 100% precision, recall, and F1-score on critical node identification.

Led a team of 5 developers, managing role assignments and providing technical guidance throughout the project lifecycle while personally handling all substantive implementation work.

Tools/languages used: Neo4j, NetworkX, Node2Vec, Random Forest, Python

RAG for Document QA | GitHub

Engineered a full-stack document QA system using LangChain, Chroma vector store, and OpenAI embeddings with 100% precision on retrieval evaluation.

Built with React frontend and FastAPI backend featuring comprehensive PyTest coverage and deployed on AWS using Docker.

Tools/languages used: LangChain, Chroma, OpenAI, React, FastAPI, Docker, AWS, PyTest

Experience

ElevateMe Bootcamp, Columbus, Ohio **Mar 2025 - Present**

Data Analytics & Machine Learning Fellow Trainee

- Engaging in 150+ hours of hands-on learning and project work, including active participation in a live project and two optional capstone projects.
- Applied data analytics and machine learning techniques to address real-world problems.
- Utilized advanced tools such as Python, Jupyter Notebooks, Pandas, Numpy, Matplotlib, Seaborn, and Scikit-learn to build and evaluate machine learning models.
- Deployed machine learning models on Microsoft Azure, integrating Azure ML with Azure SQL databases to enable real-time analytics and predictions.
- Performed data preprocessing, feature engineering, and model selection, and evaluated regression, classification, and clustering algorithms to optimize performance.
- Collaborated in a professional team environment using tools like Jira, Confluence, and Slack for project management and effective communication.

Education

Master of Science, Machine Learning : Stevens Institute of Technology, Hoboken, New Jersey
Notable Coursework: DL/ML, NLP, Computer Vision, LLMs, Text Mining, Statistical Machine Learning.

Bachelor of Engineering, Computer Science : Mumbai University, Mumbai, India
Notable Coursework: Data Science, Tableau, Linux, Big Data Analytics, Cloud Computing, Distributed Computing.

Certifications

Data Analytics and Machine Learning Certificate of Completion Mar 2025 - Present
ElevateMe Bootcamp
150+ hours of hands-on coursework, having completed a Data Analytics project and Machine Learning project using Tableau, Python, Machine Learning and SQL

PyTorch for Deep Learning Bootcamp | Link

Modern Natural Language Processing in Python | Link

TensorFlow Developer Certificate in 2023: Zero to Mastery | Link

Google IT Automation with Python | Link