Akash Kharita

Seattle, Wa | +1-206-694-3206 | akharita1999@gmail.com | Linkedin | GitHub | Website

Summary

Data scientist skilled in machine learning and Python with experience building scalable ML pipelines using deep learning frameworks like PyTorch. Demonstrates strong instructional design capabilities by transforming complex technical insights into clear, engaging learning materials. Proven track record in developing and delivering ML models for real-world applications, making technical concepts accessible for engineering teams.

EDUCATION

University of Washington, Seattle **Sep 2022 - Dec 2025**
Masters of Science, Seismology, Minor in Data Science *Seattle*
- **Coursework:** Data Visualization, Machine Learning in Geosciences, Advanced Geospatial Sciences, Data Science Seminar

Indian Institute of Technology, Roorkee **Jul 2017 - Jul 2022**
Bachelors & Masters of Technology, Geophysical Technology *Roorkee India*
- **Achievements:** Department Rank: 2
- **Coursework:** Computer Programming, Probability and Statistics, Field Theory, Geophysical Prospecting, Signal Processing, Numerical Modelling, Geophysical Inversion

WORK EXPERIENCE

University of Washington | *Graduate Researcher* **Sep 2022 - Present**
- Designed and optimized CNN and machine learning models for multi-class seismic event classification (EQ, explosion, noise, surface) for over 100k seismic events, supporting internal technical discussions.
- Enhanced label quality and model interpretability through a human-in-the-loop system, improving performance by over 20% and clarifying model insights for technical audiences.
- Engineered scalable pipelines using Python, PyTorch, and AWS S3 to deploy models on a global seismic dataset (1 PB stored in AWS S3), demonstrating robust ML framework proficiency.
- Collaborated with interdisciplinary teams and contributed to peer-reviewed publications in the GJI journal and AGU conference, reinforcing effective technical communication.

University of Alaska Fairbanks | *Data Science intern* **Jul 2021 - Sep 2022**
- Developed a Random Forest classifier to differentiate icequakes from tectonic earthquakes using 50+ waveform-derived features, achieving over 95% f1 score across multiple stations and showcasing core ML model proficiency.
- Created a feature extraction pipeline using ObsPy and NumPy for multi-station seismic data (~10k+ events), highlighting strong technical and data processing skills.
- Published research findings in a peer-reviewed high-impact journal, underscoring the ability to convey technical insights effectively.

Data Science Projects

Food Image classifier with Hugging Face | Github
- Fine-tuned Hugging Face Vision Transformer (ViT) on 10-class food image dataset using PyTorch and torchvision.

Analyzing car reviews with LLMs | Github
- Used OpenAI GPT-3.5 + LangChain to build an LLM pipeline that extracts sentiment and emotion tags from car reviews.

Building Airline chat assistant by fine tuning TinyLlama | Github
- Fine-tuned TinyLLaMA on labeled airline chat data for intent classification and response generation using Hugging Face Transformers.

SKILLS

- **Languages:** Python, SQL, Bash, LaTeX
- **ML/DL:** scikit-learn, PyTorch, Transformers, CNNs, Transfer Learning, Machine Learning, ML Frameworks
- **NLP:** Tokenization, Sentiment Analysis, Summarization, Prompt Engineering
- **Data:** Pandas, NumPy, Time Series, Feature Engineering, EDA
- **Visualization:** Tableau, matplotlib, seaborn, Vega-Lite
- **Tools:** Git, Jupyter, Google Colab, VSCode, Linux CLI
- **Scientific:** Seismology, ObsPy, Signal Processing, Geophysical Inversion
- **Training & Development:** Instructional Design, Training Development

Research Publications

- Kharita, A., Denolle, M., West, M.. Discrimination between icequakes and earthquakes in Southern Alaska using Random Forests. Geophysical Journal International. https://doi.org/10.1093/gji/ggae106
- Kharita, A., Gilligan, A.. Cluster analysis of velocity models around the Hudson Bay region. GJI. https://doi.org/10.1093/gji/ggac456