Sourav Kumar Khan

Portfolio: sourav-kumar-khan.github.io/portfolio/
Github: github.com/Sourav-Kumar-Khan

Email: khansourav1999@gmail.com
Mobile: +91-7908341191
LinkedIn: Sourav Kumar Khan

EDUCATION

- **Defense Institute of Advanced Technology (DIAT), DU, DRDO** — Pune, India
  Master of Technology - MTech, Data Science; GPA: 8.34 — *July 2021 - June 2023*
  Courses: *Operating Systems, Data Structures, Statistics, Artificial Intelligence, Machine Learning, NLP, Computer Vision, BigData*

TECHNICAL SKILLS:

- **Machine Learning** - DL, RL, RAG, NLP, PEFT, LoRA, CV, Neural Networks
- **Programming** - Python, C, C++, LATEX, (Familiar with JavaScript, MySQL)
- **Frameworks/Libraries** - PyTorch, TensorFlow, Transformers, Keras, Flask, FastAPI, Pandas, YoLo, Matplotlib, OpenCV, Spacy, HuggingFace
- **Cloud and MLOps** - AWS, Microsoft Azure, CI/CD, Jenkins, Deployment using AWS ECS and Azure container app
- **Other** - LLMs, LangChain, Text-To-Speech, Object-Oriented Programming, SQL, MongoDB, VectorDB, Docker, Git & Github

EXPERIENCE

- **SproutsAI** — United States (Remote)
  Machine Learning Engineer - II — *Sep 2022 - Present*
  - **AI Model Development & Deployment**: Developed and deployed task-specific AI models, including fine-tuned LLMs (Mistral-7B, Zephyr-7B, LLaMA3), in scalable environments using AWS SageMaker and Azure ML.
  - **NLP & Recruitment Automation**: Implemented AI-driven solutions for parsing resumes, candidate matching, and optimizing an end-to-end autonomous recruitment process.
  - **Cloud Integration & API Optimization**: Leveraged cloud platforms for model hosting, developed API services, and optimized usage to reduce reliance on external APIs.
  - **Advanced AI Techniques**: Incorporated in-context learning (ICL) and feedback mechanisms with various models to address specific business requirements and improve system performance.

- **Tata Advanced Systems Limited (TASL)** — Mumbai, India
  Machine Learning Engineer — *July 2023 - Nov 2023*
  - **Radar Data Analysis**: Processed complex radar data including reflectivity, refraction, and velocity parameters of fighter jets to track flight trajectories.
  - **Data Preprocessing & Visualization**: Cleaned, analyzed, and visualized complex radar datasets, effectively removing outliers to improve model accuracy.
  - **ML Model Development**: Developed and implemented efficient ML/DL models for trajectory prediction, optimized for low computational resources.
  - **Performance Optimization**: Created models that outperformed traditional numerical approaches in defense instrumentation while maintaining deployability on various systems.

- **Indian Institute of Tropical Meteorology Pune** — Pune, Maharashtra, India
  Project Intern — *June 2022 - May 2023*
  - **Geospatial Time Series Analysis**: Analyzed complex geospatial time series data for master's thesis using Python modules such as NetCDF, xarray, and dask.
  - **Data Visualization**: Utilized libraries like Seaborn and Matplotlib to extract insights into Mixed Layer Depth (MLD) over a 20-year period in the Bay of Bengal.
  - **Advanced Modeling**: Leveraged the Pratyush supercomputer to train over 100 complex ConvLSTM and ConvGRU models for studying seasonal variations.

PROJECTS

- **Construction Site Safety Detection using YOLOv5**: Trained and validated a state-of-the-art YOLOv5 model on a custom construction dataset to detect helmets on workers. Implemented real-time object detection for automated monitoring of basic safety measures on construction sites using video footage.

- **WhatsApp Group Chat Analysis and Visualization**: Exported WhatsApp group messages and performed data cleaning and preprocessing using NLTK and spaCy. Generated visualizations including word clouds, highly active members, and most discussed topics using Matplotlib and Seaborn.

CERTIFICATIONS

- Machine Learning A-Z: AI, Python & R + ChatGPT
- Learn C++ Programming - Beginner to Advance - Deep Dive in C++
- Ultimate AWS Certified Solutions Architect Associate SAA-C03
- The Complete Web Development Bootcamp

HONORS AND AWARDS

- **First Position in Poster Presentation**: Annual Monsoon Workshop, IMSP Pune, 2023
- **Gold Badge**: NPTEL course (National Programme on Technology Enhanced Learning)