ATHARVA PARIKH

• Portfolio-aaparikh.github.io • LinkedIn-aaparikh • Github-aaparikh • West Lafayette, IN

EDUCATION

Purdue University, West Lafayette, IN *August 2023 – May 2025*
Master of Science, Statistics and Computer Science

Savitribai Phule Pune University, Pune, India *August 2019 – May 2023*
Bachelor of Technology, Information Technology **(Gold medalist, CGPA 9.89/10)**

Courses: Design of Experiments, Stat Methods, Algorithms, Stat Inference, Databases, Data Mining, Statistical Machine Learning
Achievements: SIH 2022 Winner (Top 1% / 20,000 teams nationwide) | Techgium Hackathon Finalist (Top 12 / 7,000 teams)

EXPERIENCE

Research Assistant @ChopraLab | Purdue West Lafayette, IN, USA | *Jan 2024 – Present*
- Engineered AI healthcare assistants for Cleveland Clinic using Langchain and OpenAI API. Implemented predictive models for brain haemorrhage patients, enabling early intervention and personalised care through AI-driven support systems.
- Contributing to NCATS and NIH collaboration on AI-driven smart labs, building an AI assistant for autonomously managing diverse laboratory machines thus assisting scientists in expediting repetitive setup and experimental processes.
- Co-authored a paper on an LLM architecture built for doing scientific discovery (on arXiv soon)

Machine Learning Engineer Intern | Cook Medical Bloomington, IN, USA | *Jan 2024 – Aug 2024*
- Engineered a "Complaints Forecasting" model with an R^2 score of 0.92 using Azure AutoML and Azure ML pipelines. Led the entire process from data mining using SQL, data cleaning using Pandas, to batch deployment and hosting, and created a PowerBI dashboard for real-time monitoring.
- Developed "CookGPT" using Azure AI (Search Service, Datalake, OpenAI API) for Cook Medical's internal knowledge base. This greatly reduced the search time of employees for finding relevant information thus reducing Information localisation.

Research Assistant @StaigerLab | Purdue West Lafayette, IN, USA | *Sept 2023 – Dec 2023*
- Engineered algorithms for analysing actin microfilaments using SOAX, OpenCV, NumPy, and Pandas, optimising data processing by 35% and creating a phenotype-specific registry for plant types.

Machine Learning Engineer Intern | Labellerr Chandigarh, India | *Jan 2023 – May 2023*
- Developed "Auto-label Jobs", leveraging cutting-edge research (DINO, CLIP, ImageBind, yolov8) for automated labelling of large datasets leveraging domain-specific custom trained or zero-shot fine-tuned models. Increased labelling efficiency by 60% for image datasets. Studied and analysed competitive products, conducted technical sessions, and contributed to technical blogs.

PROJECTS

AbbVie: Marketing Investment Analytics | The Data Mine Purdue < Video >

- Engineering multi-touch sequence-based attribution models using XGBoost methods with Shapely Values for interpretability, tuning models to optimise spend and maximise ROI for marketing investments. Built Streamlit UI to provide interpretability.

Few-shot Defect Classification in Manufacturing Setup (PoC) | Techgium (L&T Technology Services) < Video >
- Developed a scalable defect detection and classification solution for manufacturing, integrating few-shot learning for new products. Deployed the solution on an embedded device with TensorRT. Selected as top 12 finalists in Techgium hackathon

ArmyPose : Validating drills using pose estimation | Simulator Development Division, Government of India < Link >
- Developed and deployed an offline human pose estimation model using yolov8, Mediapipe, and OpenCV, improving real-time drill validation process at scale with personalised feedback. Currently submitted to IEEE conference.

Pairup: Find a perfect project partner | VIIT < Link >
- Designed recommendation algorithm for a **copyrighted** React Native app that allows students and professionals to find peers in their organisation for doing group-based activities. Conducted a product analysis on applications such as Tinder and LinkedIn

AutoFis: Automatic fish detection and recognition | Smart India Hackathon 2022 – Winning solution < Link >
- Innovated an end-to-end system to classify fish species with over 93% accuracy with the help of a one-shot learning-based classifier by harnessing the SWIN transformer. Designed the UX of the application and overlooked product development

SKILLS

Programming Languages: Python, C++, R, SQL, Shell Scripting | ML/AI Frameworks: Tensorflow, PyTorch, scikit-learn, langchain | Cloud & DevOps: Azure, Docker, Git, Linux | Data Analysis/Visualisation: Power BI, Tableau, Matplotlib, ggplot2 | Tools & Techniques: NLP, Time Series Analysis, Computer Vision, A/B Testing, Experiment Design | Software Development: Flask, Django, REST APIs, TensorRT

PUBLICATIONS

Suresh, R., et al. (2024). **Revolutionising physics: a comprehensive survey of machine learning applications**. Frontiers in Physics, 12. https://doi.org/10.3389/fphy.2024.1322162

Parikh, A., et al. (2022). **Offer and Deal-Quality Prediction using Machine learning and Fuzzy approach: A Shark Tank India Case Study (ACM Conference)**. Proceedings of the 4th International Conference on Information Management & Machine Intelligence. https://doi.org/10.1145/3590837.3590891

Parikh, A., et al. (2022). **Comparison of machine learning algorithms for Twitter sentiment analysis (IEEE Conference)** 2022 International Conference on Augmented Intelligence and Sustainable Systems (ICAISS), 209–215. https://doi.org/10.1109/ICAISS55157.2022.10010781