Atharva Yeola

+1 8582203508 | ayeola@ucsd.edu | in/atharvayeola | /atharvayeola /atharva-yeola

Education

University of California San Diego (UCSD) **September 2023 – Present**
Master of Science (M.S) – Machine Learning & Data Science *San Diego, USA*

Pune Institute of Computer Technology (PICT) **July 2019 – June 2023**
Bachelor of Engineering (B.E.) – Electronics & Telecommunication – ***GPA: 3.9/4.0*** *Pune, India*

Technical Skills

Programming Languages – Python, R, C++, Matlab, Javascript
Tools & Frameworks – Docker, AWS, CI/CD, Tableau, Linux, Git
Databases – MySQL, ETL, MongoDB, NoSQL, PostgreSQL
Machine Learning – PyTorch, TensorFlow, MLflow, Keras, LLMs, CUDA, Scikit-Learn, OpenCV, SciPy

Work Experience

Machine Learning Intern **July 2024 – Present**
Qualcomm Institute - UC San Diego Division of Calit2 *San Diego, USA*

- Co-developed "**Waldo**", an **open-source** NLP tool with 97.9% test accuracy in detecting adverse events from over 360,000 Reddit posts, advancing **automated pharmacovigilance**.
- Authored multiple analyses using **data mining** and **forecasting** to identify public health trends. Key findings include a 23% increase in gambling addiction searches post-Murphy v. NCAA and a predictive model for HIV-related search trends, with predictions accurate to within ±5.2%.

Machine Learning Researcher **March 2024 – June 2024**
University of California San Diego *San Diego, USA*

- Developed and implemented a **novel RNA image segmentation technique** that achieved an Intersection over Union **(IoU) greater than 70%** across nearly 300 individual cells, facilitating more precise biological analyses and insights.
- Accelerated the segmentation algorithm by leveraging **parallel GPU processing**, decreasing the processing time by 300% and **reduced CPU overhead** by 19.5 hours for a dataset comprising 2.2 million molecules

Research Intern *(paper accepted at CVMI 2024)* **December 2022 – August 2023**
Indian Institute of Technology (IIT) Patna *Remote*

- Created a new dataset by utilizing advanced **image augmentation** techniques to generate high-quality synthetic data, thereby expanding the dataset size to 100,000 images (**30% increase**), improving robustness of the model.
- Tailored a **Vision Transformer** architecture to address edge-case scenarios in traffic sign detection, beating the **state-of-the-art** F1 score by 9%.

Machine Learning Intern **August 2022 – April 2023**
RhythmFlows Solutions Pvt. Ltd *Pune, India*

- Conducted **A/B testing** to compare the effectiveness of content-based filtering against collaborative filtering in a music recommendation engine, resulting in a **33% boost** in relevant recommendations. **Deployed** the content-based model on **AWS Sagemaker** with continuous weight updation for improved performance.
- Reduced document processing time by 7 seconds per document by utilizing **Pytesseract OCR** for electronic receipt scanning, and revamped the data analysis pipeline, **minimizing human effort by 30%**.

Machine Learning Intern **January 2022 – June 2022**
Atomic Loops *Pune, India*

- Spearheaded a team of three in training a **YOLOv5** food detection model achieving **98% precision** and then smoothly integrated it with a **food train system**.
- Collaborated with the data platform team to construct preprocessing pipelines, facilitating a notable **20% acceleration in project delivery** rate.

Projects

End-to-End Text Summarizer | *Transformers, NLP, GitHub Actions* **July 2024 – July 2024**

- Built a text summarizer for concise summaries, managing the entire workflow from configuration to deployment.
- Automated CI/CD with AWS and GitHub Actions, handling Docker, EC2, ECR, and runner setup.

Retail Vision Enhancement | *YOLOv8, SuperGlue, Docker* **February 2024 – February 2024**

- Implemented YOLOv8 for object detection, achieving over 90% accuracy in labeling on-shelf retail products.
- Utilized SuperGlue for precise product identification, and created a Dockerfile to ensure cross-platform reproducibility.