Mustaffa Hussain

mustaffahussain4734@gmail.com
+91 9057260535
Hyderabad, Telangana, India

Portfolio
LinkedIn
Medium Blogs
Github

EDUCATION & CERTIFICATE

Qualification	Institution	CGPA/Percentage	Year Of Pass
M.Sc Computer Science	South Asian University	6.85 / 75%	2019
B.Sc Computer Science	Central University of Rajasthan	4.80 / 68%	2017
HuBMAP + HPA - Hacking the Human Body	Kaggle Competition	Bronze medal	2022

SKILLS

- **Programming Languages:** Python, C, C++
- **Frameworks:** Pytorch, Tensorflow, Pytorch lightning, Keras
- **Computer Vision:** Opencv, Classification, Object Detection, Segmentation, Image Processing
- **Dev/ML Ops:** GraphDBs, SQL, Flask Restful APIs, Git, Docker, AWS, GCP
- **Libraries:** NumPy, SciPy, scikit-learn, pandas, matplotlib, openCV, transformers, Hugging Face
- **Relevant Coursework**: Machine Learning, Statistics, Optimisation & Data Structures and Algorithms, Theory of computations, Operating systems
- **Research**: Multimodal vision language models, LLMs, Contrastive learning, Foundational models

PROFESSIONAL EXPERIENCE

Onward Assist : Hyderabad,IN

Senior Machine Learning Scientist Dec 2020 - Till Now

- Breast Cancer Detection Tool (IHC): There are several Immunohistochemical (IHC) markers to diagnose cancer such as ER, PR, KI67, HER2. Developed individual AI algorithms for quantification of these markers on large Whole Slide Images [100k X 100k pixels]. Implemented a novel Instance segmentation model for dense pixel-level predictions. Successful clinical study and technical paper in writing.
- Lung Cancer Detection Tool (IHC): Led the development of ML models for assessing PDL1 prognostic biomarkers in lung cancer. Novel model inspired by Unet, Hibou foundational model for dense pixel-level prediction on Whole Slide Images. Ongoing study.
- Generative Modeling: WSIs scanned from multiple scanners and labs have a lot of variability. Utilised Unpaired Image to Image generative modeling such as pix2pix, cycleGAN, and Schrodinger bridge to learn target distribution to standardize data pipelines. Successfully generated large-scale medical images of 100k X 100k pixels. Accepted at USCAP 2025.
- Breast Cancer Detection Tool (H&E): Nottingham Grading Score quantifies how the cancer cells are growing compared with normal cells in tissue stained with Hematoxylin and eosin(H&E). Led the development of object detection and segmentation models for Mitosis, Tubules, and Pleomorphism detections in WSIs. Nottingham Scoring algorithm boosting breast cancer diagnosis accuracy by 6% compared to our previous baseline. Technical Paper
- Image Registration: Implemented coregistration of 3D CT and 3D PET scans using linear and affine transformations. Segmented tumor regions from the above and reduced the turnaround time of the radiologist by 60% on a single scan of 4 test subjects.
- Deployment: Integrated ML models into the web platform.Built backend Flask APIs, containerised applications using Docker, and utilised GCP and Kubernetes for deployment at scale.

Myways : New Delhi,IN

Machine Learning Trainee May 2020 - Dec 2020

- Concepted Employability Positioning System(EPS) where predefined career paths are mapped to graphs and the journey of a job seeker is tracked. Developed a recommendation engine using knowledge graphs for next skills to learn, online courses to upskill, and book recommendations.
- Developed analytics dashboard from 100GB of textual data for trending skills, trending books, trending papers etc.
- Developed a module for the analysis of customer feedback. Implemented custom NER and sentiment analysis pipeline using NLP techniques.
- Set up an automated data scraping pipeline on LinkedIn, Naukri, and Internshala. Cleaned the scraped data to generate leads and integrated a live dashboard.

Machine Learning and Statistical Inference Lab (MLSI-LAB) New Delhi,IN

Graduate Research Assistant Jul 2018 - Mar 2019

- The project investigates Multi-Task learning in SVMs.The project involved reading research articles, and papers and devising a solution to improve the existing models. Successfully developed a novel Multi-Task Learning framework based on Twin support vector machines under the supervision of Dr. Reshma Rastogi.

PAPERS AND POSTERS

- Generating IHC to H&E Whole Slide Images, USCAP 2025.
- Generating Digital Stains Via Neural Schrodinger Bridge in Pathology Images, ICGA 2024.
- Switched auxiliary loss for robust training of transformer models for histopathological image segmentation.
- Robust Multi-Domain Mitosis Detection via Unpaired Domain Generalisation- MICCAI 2021 proceedings.
- Robust Multi-task Least Squares Twin Support Vector Machines for Classification. Advanced Machine Intelligence and Signal Processing. Singapore: Springer Nature Singapore, 2022. 393-405.

PROJECTS AND LEADERSHIP

Technical Project: Computer Vision

- **KEYE-** developed real-time video analytics safety monitoring system. The work detected PPE kit violations, fire, spillage of liquids, crowds, wrong lanes, etc on real-time video feed from multiple IP cameras.
- **Deep learning Tutorial**- Set up tutorial notebooks for various computer vision models from classification, segmentation, pose estimation, etc. I also explored model optimisation at inference with quantization.

Blogging: TheCyPhy@Medium

- I started a publication on Medium named TheCyPhy. We are 20+ writers and have 35+ published articles in the domain of ML and DL.
- Some of the articles written/moderated by me on TheCyPhy are ‘Impression Generation From Medical Imaging Report’, ‘Deploy Deep Learning Model Using Docker, Tensorflow Serving, Nginx and Flask’,’ Let's Understand Object Recognition’,’ Test drive with YOLO v4’ etc.

OTHER ACHIEVEMENTS

- Table Tennis and Badminton: Singles and Doubles winner in Intra University tournament.
- NCC: certificate A,B,C holder.
- Represented school at CBSE zonal meet for Hockey.