M. Alfi Hasan, PhD

Staff Data Scientist at Walmart Global Tech (Current)

Email : mdalfihasan19@gmail.com
Phone : +1-347-265-5476
Immigration : Permanent resident of USA
Website : www.malfihasan.com

Data scientist, system architect, and technical leader specialized in applied artificial intelligence (AI) models for industry across multiple diverse domains, including computer vision, large language and vision models, remote sensing, climatology, agriculture, and hydrology. Currently working as a staff data scientist and tech lead at the innovation wing of Walmart Global Tech. Possesses a unique strength in adopting multidisciplinary approaches in cross-functional corporate and academic settings, demonstrating excellence in proposing and implementing innovative solutions for the organization. Orchestrated and led collaborative efforts in deploying a series of deep learning frameworks in IoT/edge devices with proper CI/CD pipelines for an innovative multi-camera item recognition system at scale for the largest retailer in the world, Walmart Inc. Demonstrated technical acumen and supported large-scale collaborative projects on deploying satellites and drones for agricultural operations. Mentored coworkers, acted as an advisor, provided corporate mentorship with universities such as Purdue University, and taught classes. The doctoral research focused on predicting the epidemic cycle of infectious diseases using satellite data and machine learning models. Has extensive experience and deep domain knowledge of the latest machine learning/deep learning architectures and techniques such as large language models (LLMs), large vision models (LVMs), GANs, YOLO series, EfficientNet series, CLIP, Grounding DINO, Transformers, etc. Proposed unique architectures for generative AI models, and is an author of several patent applications in the remote sensing and computer vision domains (in process). A passionate programmer with professional working knowledge of several programming languages (e.g., Python, C#Go), and has designed many end-to-end processes from camera image capture and processing backend to uploading images to dashboards. Currently interested in advancing to senior scientist or leadership roles in the tech industry.

Experience

Staff Data Scientist — May 2024 - Present

Project Constellation, Walmart Global Tech, Bentonville, AR, USA

* Leading the data science team to **provide a cost-saving management solution** for store associates by creating an intelligence service for security cameras utilizing the latest large vision models (LVMs). This project leverages various LLM backbones like GPT-4, Vicuna 13b, Mistral 7b, and LLaMA 2, combined with LangChain and Agents, to provide solutions from image to text.
* Orchestrating an in-house benchmarking framework for large language models, which has the potential **to impact 10+ teams** in the development of Generative AI-based products.
* Guiding and coding a multi-barcode reader solution that **enhances the shelf-stocking process in over 1,000 Walmart stores** using deep learning architectures like Grounding DINO and SAM (Segment Anything).
* Designing and leading the technical framework **that save time and cost of the annotation team,** utilizing the latest models like Grounding SAM and multi-modal models.

Senior Data Scientist — May 2022 - April 2024

Project Diamond, Walmart Global Tech, Bentonville, AR, USA

* Led the development of a mobile annotation tool (both Android and iOS) **as a tech lead, significantly reducing the cost and speed of manual-human annotation** using transfer-learned object detection (YOLOv8) and classification (embedding-based EfficientNet) models. Managed cross-functional engineering and product teams.

* Conducted technical orchestration and coded a machine learning-based barcode detector, replacing Walmart's current barcode detection system, estimated **to save about $4 million a year.** The framework consists of a real-time object detection pipeline (both in Android and iOS, EdgeYOLO-inspired) with a Zebra crossing library, providing storage efficiency and high performance.
* Proposed, orchestrated, and directed an automated end-to-end model training and inference pipeline for Walmart's entire Data Science constellation. This **helped scale dynamic product detection for several Walmart stores using 100+ custom-made cameras.**
* Developed and implemented a novel architecture for dimension detection, creating the **potential to save thousands of dollars** using in-house deep learning architecture similar to MiDAS (depth detection) and EfficientNet (classification).
* Orchestrated the deployment of a framework **handling over 1,000 images per second processing requests**, utilized for machine learning model (both object detection & classification) inference in conjunction with embedding vector databases.
* Organized and **led overall innovative research** by hosting bi-weekly Data Science meetups and maintaining a mono-repo to facilitate knowledge sharing and team objectives for the Data Science team.
* Developed a recommendation system that provides **a dynamic product recognition solution for Walmart's inventory management process,** utilizing an embedding-based Siamese model with LightFM.

Senior Data Scientist **Oct 2021 - May 2022**

Research and Development (Breeding), Bayer Crop Science, St. Louis, MO, USA

* Diven roadmaps for quantitative solutions and reduced operational cost [both labor and logistics] ***by 50%***+ for North America and Europe canola operations by proposing & developing an automated UAV-based canola flowering detection model. The development process utilized image segmentation models like **DeepLab3**, and image classification models like **ResNet** Backbone, Conventional Computer Vision (CV) algorithms using **OpenCV**, and cloud computing platforms like **AWS-EC2** instances.
* Automated the cotton UAV-pipeline operation for thousands of fields across ***entire North America*** by co-developing deep learning-based scalable solutions that detect maturity of cotton from images. For the project, utilized **AWS-Sagemaker** in conjunction with **parallelized dockers** with **YOLOv3 object detection** architecture , with multi-year **transfer learning** inferences and collaborated with stakeholders from diverse domain expertise using **code versioning systems such as Git**.
* Created navel and creativity embeddings for temporal models that resulted in ***a patent***. The development involved modeling with architecture families like **GANs(Pix2Pix, StyleGAN) , Transformers** on the multi-band geospatial images using **PyTorch (Python**) and querying languages like SQL
* Provided visibility model-derived results that helped ***several operational teams*** by developing geospatial visualization through **Tableau** dashboard. The project required geospatial expertise using packages like **GDAL or Shapely** which conventional software architects lacked.
* Investigated new possibility of satellite acquisition which resulted in ***the new acquisition of satellite products*** by providing market analysis on latest satellite vendors for the R&D on its 5+ year vision.
* Help the company to recruit data scientists for ***three teams*** by being in the technical interview panel in the **HR interview process**.

Spatial Data Scientist **Mar 2020 - Oct 2021**

Research and Development Breeding, Bayer Crop Science, St. Louis, MO, USA

* Increased breeding pipeline efficiency and introduced bias-free automated information for global cotton stakeholders with ***a cost-saving potential of almost half a million*** by leading and developing a UAV-based multi-flight cotton maturity detection framework to eliminate ground notes. To develop the framework in this multi-year project utilized semantic segmentation models like **UNet** and **RCNN** in conjunction with **Decision tree models**, CV algorithms like **Canny Edge detector** in **Unix OS like Ubuntu** using **Tensorflow**. The project is being implemented in production for 500 + flights each season using computer vision models.
* Developed several automation modules for the first scalable UAV processing pipeline for the company where the project enabled the elimination of field notes for many crops and ***saved millions of dollars on the field operations***. In the development process used geospatial python libraries like **geopandas, shapely, and Kerus** in the **Postgres** server and conducted feasibility studies using **IoT** devices. The project won the breeding excellence awards for 2021.
* Orchestrated components of the global UAV-based soy maturity automation pipeline that ***created several thousand acres o***f new soy maturity operations by developing tools for color calibration, GCP detections, and maturity modeling. The work resulted in another ***patent*** for the company.

Research Data Scientist (Geo-Spatial) **Sep 2018 - Mar 2020**

Bayer Crop Science (Contract: Colaberry Inc), St. Louis, MO, USA

* Eliminated field monitoring operation ***by more than 50% by*** proposing and developing a Satellite-based crop health quality recommendation system. The project conducted time series forecasting and anomaly detection using model families like **GANs** and also utilized high-resolution multi-band structures using the model families like **LSTM.**
* Provided operational placement benefit for a global testing network by generating satellite-based clustered embedding using **word2vec** model framework.
* Increased understanding and helped to ***generate new seeds for corn stakeholders*** by developing lodging & yield prediction models using Satellite in the company. Used traditional machine learning models like **Random Forest, TabNet, and XGBoost.**

Remote Sensing Data Scientist (Summer intern) **Jun 2018 - Aug 2018**

Monsanto (Contract: Colaberry Inc), St. Louis, MO, USA

* Automated open source, cost-free stitching solution for hundreds of UAV flights by developing a scalable stitching pipeline of **Open Drone Map (ODM).**
* Developed code for the **AWS** pipeline, specifically in the automated streamlining of data in Google Big Query and **AWS - S3**.

Graduate Research and Teaching Assistant **Jan 2015 - May 2018**

University of Rhode Island, Kingston, RI 02881, USA

* Developed a model that can predict the epidemics of rotavirus in the developing world 1-month prior using satellite under the project named “Control of Endemic Cholera in Bangladesh", funded by the Bill & Melinda Gates Foundation.
* Developed irrigation App that can recommend irrigation to farmers under summer project with The International Maize and Wheat Improvement Center(CIMMYT)
* Research Fellow in RI-Water Resources Center.

Research Associate **Jul 2011 - Dec 2014**

Institute of Water and Flood Management (IWFM), Dhaka-1000, Bangladesh

* Collaborated and worked on International Research Projects like:
  * High End Climate Impact and Extremes (HELIX)
  * High-resolution Regional Climate Change Information for Bangladesh to inform Impacts assessments

Visiting Scientist **Jul 2012 - Sep 2012**

Met Office, Hadley Center, FitzRoy Road, Exeter, Ex1 3PB, United Kingdom

* Contributed global climate change project that was later included in IPCC by working on the PRECIS - HADCM3 climate model over South Asia.

Lecturer **Jun 2011 - Apr 2012**

World University of Bangladesh, 3A, Road No 4, Dhaka 1205, Bangladesh

Leadership, Awards and Extracurricular activities

* Led the team and won **1st place** in **the Gen AI category** and secured a top 20 position overall in the Techathon 2023, hosted by Walmart Global Tech.
* Received the **'Exceptional Performer'** award within a large team of 50+ employees at Walmart Global Tech in 2023.
* Won the **Digital Innovation Award 2021** from **Bayer Crop Science** for developing an automated UAV pipeline, which is an end-to-end process from capturing images by drones to backend image processing, deploying deep learning models, and visual inference on a dashboard.
* Facilitated novel solutions in remote sensing (RS) in agriculture, **leading to the development of patents** for the organization by connecting individuals and scientists through meetups on satellite and RS disciplines, and creating cultural awareness.
* Collaborated with universities such as **Purdue University, the University of Illinois at Urbana-Champaign, and the University of Missouri** as a corporate mentor, providing mentorship and emotional intelligence to graduate students in RS and agriculture.
* Managed 3,000+ student blood donors as the nominated President of the Blood Donation organization, Badhon-Suhrawardi Hall Unit, in 2008 and 2009.

Education

Ph.D. in Civil and Environmental Engineering **Aug 2018**

University of Rhode Island, Kingston, RI 02881, USA

Master of Science in Water Resources Development **Dec 2014**

Bangladesh University of Engineering and Technology (BUET), Dhaka-1000, Bangladesh

Bachelor of Science in Water Resources Engineering **Feb 2011**

Bangladesh University of Engineering and Technology (BUET), Dhaka-1000, Bangladesh

Publications

Have published **9 journal papers** and **20 conference papers** and reviewed many more. Have more than 700+ citations as of now. Presented talks and posters in various conferences. List : https://scholar.google.com/citations?user=mi0w3fgAAAAJ&hl=en