

# ISHAN SRIVASTAVA

in [ishansrivastava96](#) • [+ \(91\) 9517222283](#) • [ishan.all@gmail.com](#) • [Ishan-phys](#) • [portfolio](#)

## EXPERIENCE

### Intellithink Industrial IoT

Dec 2022 - Present

Data Scientist

Bengaluru, India

- Led the seamless integration of machine learning models into the product lineup, yielding significant benefits such as a 20% reduction in costs for clients and a noteworthy 10% enhancement in heavy machinery efficiency achieved through predictive maintenance strategies.
- Built a scalable AWS Sagemaker infrastructure for 10,000 daily requests, realizing a 25% cost reduction through dockerized serverless inference tailored for real-time IoT streaming data. Implemented MLOps best practices, including CI/CD pipelines with Bitbucket, automating endpoint deployment.
- Optimized ML model training by over 100% through the strategic utilization of Apache Spark and datalakes, facilitating terabyte-scale ETL operations on IoT sensor data.
- Accelerated inference time by up to 50% through parallel processing of features.

### Intellithink Industrial IoT

Sep 2022 – Dec 2022

Data Scientist - Intern

Bengaluru, India

- Enhanced Anomaly Detection in Sensor Time Series: Enhanced accuracy by up to 2% by utilizing LSTM autoencoder based modelling for time series data.

### Chennai Mathematical Institute

Jul 2019 – July 2021

Research Assistant

Chennai, India

- Conducted in-depth research on specific facets of Quantum Field Theory, leveraging problem-solving prowess to formulate and validate mathematical proofs. Contributed to the field by publishing findings in the [Journal of High Energy Physics, Springer](#).

## TECHNICAL SKILLS

- **Programming:** Python, SQL, Bash, C++, LaTeX
- **Dev Tools:** Github, VS Code, Docker, AWS
- **Databases:** MongoDB, SQL, AWS based Databases
- **MLOps:** MLFlow, AWS Sagemaker, DVC, Apache AirFlow, Apache Kafka

## EDUCATION

### M.Sc. in Scientific Computing and Data Analysis

Sep 2021 - Sep 2022

University of Durham, First Class Honours (71%)

Durham, U.K.

### Bachelors in Theoretical Physics

Oct 2015 - Jul 2019

University of Durham, First Class Honours (80%)

Durham, U.K.

## AWARDS AND ACTIVITIES

CIUK Cluster Challenge - Second Position – STFC, UK

2021

Outstanding Achievement Award – University of Durham, U.K.

2017, 2018 & 2019

J.A. Chalmers Prize for Masters thesis – University of Durham, U.K.

2019

## PROJECTS

- **ML Based Predictive Maintenance Solution (Anomaly Detection)** | *Streamlit, MongoDB, Docker* [Streamlit](#)  
Developed a robust three-tier predictive maintenance demo application integrating Streamlit for real-time monitoring, an unsupervised anomaly-detection ML model, and MongoDB for efficient data storage. Utilized IMS Bearing Data and cloud deployment to enhance operational efficiency, offering real-time anomaly detection and user-friendly predictions via a web interface. Demonstrated expertise in statistical methods, ML algorithms, and thoughtful UI design for proactive equipment maintenance.
- **Image-to-Image Translation for Medical Datasets (Computer Vision)** | *PyTorch, Deep Learning* [PyTorch](#)  
Automated the labour-intensive task of analysing the vasculature of retinal images using CycleGANs, cutting the costs by more than 50% to do so. The CycleGAN, trained from scratch using PyTorch, performed unpaired image-to-image translation from retinal images to the vasculature annotation of the retinal image.
- **Generating New Examples for Image Dataset (Computer Vision)** | *Generative AI, Distributed GPU Training* [Generative AI](#)  
Extended the existing medical datasets using state-of-the-art generative models by generating new data, cutting costs and time to collect new data. Created a python package 'diffusion-sde' for easy adaptation of the PyTorch code for distinct datasets.
- **SDE based Generative AI Models (Computer Vision)** | *Generative AI, Distributed GPU Training* [Generative AI](#)  
Researched state-of-the-art models to perform unpaired image-to-image translations, potentially improving other existing generative models for the task. The SDE-based generative modelling was used for the task, and the code base was implemented from scratch using PyTorch. Article on [Medium](#).