# Satvik Kishore

satvik.kishore@gmail.com | (984) 377-1121
github.com/satvikk | linkedin.com/in/satvik-kishore | www.satvikkishore.com

## EDUCATION

**Duke University** | *Master's in Interdisciplinary Data Science.* GPA: 4.0 *Expected May 2023*
Coursework: Computer Vision and Medical Imaging, Natural Language Processing (NLP), Machine Learning, Statistics, Causal Inference, Data Engineering, Deep Learning
**Indian Institute of Technology Madras** | *Bachelor of Technology in Materials Engineering.* *Jul 2013 – May 2017*

## SKILLS

- R, SQL, Python, PyTorch, TensorFlow, Docker, Tableau, Amazon Web Services, A/B Testing
- **Certification:** AWS Machine Learning Specialty

## PROFESSIONAL EXPERIENCE

**Proofpoint Inc. & Duke University** *Aug 2022 – Present*
*Part Time Machine Learning Engineer*

- Benchmarked methods to reduce variance in ML model predictions when models are retrained with fresh data.
- Developed novel methods based on knowledge distillation to reduce model variance metrics (Prediction Churn) by 35%.

**Interpretable Machine Learning Lab, Duke University** under Dr. Cynthia Rudin *Jun 2022 – Present*
*Graduate Research Assistant*

- Designed computer vision breast cancer detection model based on case based reasoning architecture in PyTorch.
- Achieved SOTA AUC scores of 0.95 while improving the interpretability of the model by making it more sensitive to fine-grained features in images.

**Data+, Rhodes Information Initiative** *May 2022 – Aug 2022*
*Machine Learning Engineer Intern*

- Developed Gaussian Process Regression Machine Learning models for earthquake early warning systems in Nepal.
- Implemented pipeline that takes in input the seismograph wave data, transforms the data to wave features, then uses the trained model to predict intensity of oncoming earthquakes to trigger alarms.
- Improved upon the SOTA $R^2$ scores by 30%. The model is set to be installed in earthquake detection facilities in Nepal.

**St. John's National Academy of Health Sciences** *Jul 2017 – May 2021*
*Data Scientist*

- Standardized and centralized health data from multiple sources into a single repository to be used by multiple institutions.
- Secured $200,000 in funding from the Bill and Melinda Gates Foundation to lead a project on analyzing the impact of improved household income on child undernutrition in India.
- Modelled child undernutrition in the country through Geospatial Machine Learning Models using PyTorch. Identified three key regions with potential for 15-25% reduction in undernutrition through improved monthly income.
- **Publication**: Kishore S, Thomas T, Sachdev H, et al. Modeling the potential impacts of improved monthly income on child stunting in India: a subnational geospatial perspective. BMJ Open 2022;12:e055098.

**Lighthouse Datalabs** *May 2016 – Dec 2016*
*Data Scientist Intern*

- Developed machine learning models to help a healthcare blog client perform demographic based ad targeting.
- Trained tree-based ensemble models to classify user's age and sex, improving AUC by 10%, ad click-through rate by 5%.

## PROJECTS

- Brain Tumour image segmentation using U-nets. *Spring 2022*
  - Developed an AI tool for brain tumor image segmentation for precise neurosurgery, using TensorFlow.
  - Implemented a "U-Net" Deep Neural network architecture to achieve a DICE similiarity score of 81%.
- CT Scan Imaging Optimization. *Spring 2022*
  - Conducted a study on optimization of radiation exposure in CT scans to balance image quality and lesion detection.
  - Trained YOLO models to reduce radiation exposure by 50% while maintaining lesion detection capability.
- AWS Powered Tweet Generator. *Fall 2021*
  - Implemented and dockerized NLP model to retrain on fresh daily data and deployed to generate trend-relevant tweets.
  - Engineered AWS cloud solution to schedule NLP docker containers to run everyday. Brought down costs to < $1/month.