

Abi Asok (Mr. Abhijith Asok)

Boston, MA 02120 | +1 (857)-289-7825 | aasok@hsph.harvard.edu | [LinkedIn](#) | [GitHub](#) | [Website](#)

Data Scientist

Work eligibility in the US: May 2019 – May 2022 under OPT

Experienced professional with pending **Master of Science degree** in **Health Data Science** from **Harvard University** and **3 years of past experience** with data. Extensive background as data scientist across public health, pharmaceuticals, real estate, communications and more across corporate, social, and research.

EDUCATION & SKILLS

- **Master of Science in Health Data Science** | Harvard University, Boston, MA (Anticipated May 2019)
Coursework: Inference, Big Data, Deep Learning, Machine Learning, Networks, Linear Algebra
- **Master of Science (Honors) in Mathematics**, 2015
- **Bachelor Science (Honors) in Electrical and Electronics Engineering**, 2015
Birla Institute of Technology and Science, Pilani, India

Technical Skills

- Most experienced in R (tidyverse, caret, Shiny), Python (keras, NumPy, pandas), Google Cloud, AWS
- Some experience with Tableau, SQL, MS Excel
- Dabbled with SAS, d3.js; R connections to Spark, h2o, MongoDB

AREAS OF FOCUS

- | | | |
|-----------------------------|---------------------------|------------------------|
| ▪ Machine Learning | ▪ Deep Learning | ▪ Data Wrangling |
| ▪ Tree-based Modeling | ▪ Data Visualization | ▪ Linear Algebra |
| ▪ Organizational Leadership | ▪ Network Science | ▪ Statistical Modeling |
| ▪ Data Science Mentorship | ▪ Collaborative Filtering | ▪ Algorithms |

PROFESSIONAL EXPERIENCE

RESEARCH ASSISTANT | Dr. Xihong Lin's group (Statistical Genetics), Harvard University, Boston, MA 2018
Developing a unique functional annotation tool for vector representation and querying of genetic variants.

DATA SCIENCE INTERN | Mount Sinai Health System, New York, NY 2018
Developed a Computer Vision deep learning solution with convolutional neural networks to detect abandoned tires in aerial drone shots of Guatemala to reduce mosquito-borne diseases.

- Mathematically derived optimum bounding-box dimensions using drone altitude and camera specs.
- Modeled a deep network based on YOLOv3 to achieve a usable base model.

DATA SCIENTIST | Global Eagle Entertainment, Dubai, UAE/Trivandrum, India 2017
Designed unique network intelligence data visualization dashboard focused on maximizing network & signal transmission efficiency and long-range organizational cost reductions.

- Initiated history-based bandwidth prediction ML project aimed at company savings of \$7M.

SENIOR DATA SCIENTIST | Omaxe, New Delhi, India 2016 to 2017
Worked with medium-sized real estate firm's COO to optimize procurement through price tracking.

- Decreased daily dashboarding time from 4-5 hours to 2 minutes through end-to-end automation.

HEAD DATA SCIENTIST | Safecity, Mumbai, India

2015 to 2017

Formed and supervised team of 30+ data volunteers working towards women's safety through text analytics for social media, safest route prediction, and dynamic dashboarding.

- Collaborated with Stanford University's CS 50 Tech for Good course in fall semester 2016 to mentor undergrad students and Evelyn Jacobs Ortner Center on Family Violence at University of Pennsylvania.
- Impacted the general society by creating dashboards that were used by police, administration in multiple countries to alter patrol timings, charter special routes, advance communities etc.

CONSULTANT | GSG, Pune, India

2016

Created successful non-parametric resampling-based K-Means clustering prototype based on telecom device usage for company to prepare multiple incentive plans.

BUSINESS ANALYST | ZS Associates, Pune, India

2015 to 2016

Strategized and successfully created real-impact modeling projects in healthcare.

- Implemented extreme gradient boosting on extensively feature engineered 1.5 million row dataset to determine if women had taken Pap Smear test to detect cervical cancer in the past five years. The two-week project reached 91% detection accuracy on 2 million test patients.
- Devised model to predict optimum medical insurance premiums based on data of 270,000 people resulting in third place among 49 other competing intra-company efforts.
- Earned first place in 51-team company-wide challenge by creating ARIMA-based forecasting model predicting 7-week future sales of various pharmaceutical products based on past sales.

ANALYTICAL PRODUCT RESEARCH INTERN | Mu Sigma, Bangalore, India

2014 to 2015

Administered requirements and wireframes for hybrid clustering and distance matching modules of organization's primary data science workbench product.

- Served on product research group reviewing, arranging, and implementing the modules
- Investigated, established requirements and recorded reference documents for modules including theory and application and helped developer group implement requirements as product features.

PUBLICATIONS

- Lea S G, DSilva E, **Asok A**. Women's strategies addressing sexual harassment and assault on public buses: an analysis of crowdsourced data. *Crime Prevention and Community Safety*. 2017; 19 (227): doi: 10.1057/s41300-017-0028-1
- **Asok A**. Generalized approach to linear data transformation. *2016 IEEE International Conference on Data Science and Engineering (ICDSE), Cochin*. 2017; 1-6 doi: 10.1109/ICDSE.2016.7823937

ACADEMIC PROJECTS

- Modeled a Recurrent Neural Network (GRU)-based deep network as a team to predict ICD-9 codes from doctor's discharge notes using FastText word embeddings.
- Predicted the virality of online media articles using gradient boosting on article metadata.