

Samit Watve

DATA SCIENTIST

samitwatve@gmail.com
linkedin.com/in/samitwatve
github.com/samitwatve

Profile:

I am a fundamentally curious person who enjoys discovering interesting answers to complex questions. I have extensive training in scientific research, hypothesis testing, statistics, and data analysis. Several years as an educator have taught me how to excel at condensing complex information into simple ideas for audiences of different backgrounds. I am excited to apply my skill set to solving intriguing business problems in the field of Data Science.

SKILLS

Python, Jupyter, Scikit-Learn, TensorFlow, R (ggplot2), SQL, Spark, Hadoop, GitHub, Tableau, C, Visual Basic, Excel, GraphPad Prism, Command line, Bowtie, DESeq2

EXPERIENCE

Postdoctoral Fellow | Tufts University

FEB 2017- DEC 2020, BOSTON, MA

- Performed A/B testing on 400 molecular interactions to determine best ligand fit with CqsR receptor
- Conducted hypothesis testing, data analysis and data visualization
- Published 2 first author papers in prominent journals in microbiology

Graduate Research Assistant | Georgia Institute of Technology

AUG 2011 - 2016, ATLANTA, GA

- Performed RNA-sequencing and differential gene expression analysis to identify factors regulating genetics circuits in human pathogen *Vibrio cholerae*
- Discovered novel genetic interactions between quorum sensing and natural transformation
- Published 5 first author papers in prominent journals in microbiology

Head Teaching Assistant | Georgia Institute of Technology

JAN 2011- 2016, ATLANTA, GA

- Led hands-on training in lab techniques to a class of 40 undergraduates for Introductory Microbiology Lab
- Developed novel teaching materials for a class of ~300 students in Cell and Molecular Biology

EDUCATION

BrainStation | Diploma, Data Science

SEP- DEC 2020, TORONTO, ON

Georgia Institute of Technology | PhD in Biology, Double Minor Microbiology and Entrepreneurship

AUG 2011 - 2016, ATLANTA, GA

Outstanding research award 2016, Graduate Student Excellence Award 2015, Broadening Experiences outside Scientific Training (BEST) Award 2014

University of Pune | Masters in Bioinformatics and Biotechnology

2005 - 2010, PUNE, INDIA

Bronze medal in International Genetically Engineered Machines (iGEM '09)

PROJECTS

Predicting IBD risk from genomic data using machine learning | Data Scientist

NOV 2020 - DEC 2020, BrainStation Data Science Capstone project

- Developed a machine learning pipeline to identify genomic risk predictors for Inflammatory bowel disease
- Achieved industry leading accuracy score of 0.83 /ROC-AUC of 0.91

Tools used: Scikit-learn, numpy, pandas, matplotlib, tensorflow, jupyter notebook