Martin Skarzynski

BIOSTATISTICS · DATA SCIENCE · EPIDEMIOLOGY · GENOMICS · MACHINE LEARNING · PHARMACOMETRICS · PUBLIC HEALTH · SOFTWARE ENGINEERING

240-595-3460 | mskar@dr.com | mskar.github.io | mskar | mskar | marskar

Summary

My primary research interest is in understanding health risk factors by combining scientific expertise from diverse fields with machine intelligence. I believe I am uniquely equipped to bridge the gaps between scientific disciplines and deliver on the promise of data science in health research. My preferred tools are R and Python, open source programming languages kept on the cutting edge by their active and supportive communities. Through research and teaching, I am constantly improving my ability to obtain, tidy, explore, transform, visualize, model, and communicate data. I aim to utilize my technical skills and science background to become a leader among the next generation of multidisciplinary data scientists.

Recent Experience

Lead Scientist

BOOZ ALLEN HAMILTON   *October 2019 - Present*

- Translates Subject Matter Expertise into Python and R code
- Engineers machine learning algorithms, probabilistic models, and statistical inference tools
- Writes tests to ensure model validity and code reliability
- Uses git version control to track and manage changes to code, tests, and documentation
- Mines data by webscraping and via Application Programming Interfaces
- Engineers data transformation pipelines
- Determines optimal algorithmic approach through iteration and experimentation
- Uses Docker and Conda to ensure reproducibility and reliable deployments
- Designs SQL queries for improved database querying efficiency
- Prototypes new applications using Django, Flask, Plotly Dash, and R Shiny
- Builds dashboards products to display data and model outputs
- Develops, maintains, reviews, and refactors Python, Shell, R, and JavaScript code
- Coordinates model and dashboard code deployment via Amazon Web Services
- Communicates results using compelling static and interactive plots
- Leverages broad scientific and technical expertise to deliver rigorous analyses
- Employs the latest technologies and best practices in data science and software engineering projects
- Keeps stakeholders informed with parametrized reports written using Jupyter notebooks and R Markdown
- Creates applications to make analyses accessible to non-technical audiences
- Leads multidisciplinary COVID-19 statistical modeling effort

Lead Instructor

DATA SOCIETY   *October 2018 - Present*

- Teaches courses such as:
    - Classification
    - R Programming
    - Data Science Bootcamp
    - Data Science Academy
    - Text mining
    - Python Programming

GENERAL ASSEMBLY   *July 2019 - Present*

- Teaches courses such as:
    - Data Science 5000
    - Python Programming
    - Tech Excellence Data Science
    - Data for Leaders

Treasurer

DATA COMMUNITY DC   *April 2020 - Present*

- Manages the finances of a non-profit organization that supports eleven Meetup groups

Cancer Prevention Fellow

BIOSTATISTICS BRANCH, DIVISION OF CANCER EPIDEMIOLOGY AND GENETICS, NATIONAL CANCER INSTITUTE   *June 2017 - April 2021*

- Integrated clinical, laboratory, epidemiologic, genomic, and medical imaging data
- Combined deep learning and statistical inference using stacked ensembles (super learning)

Bioinformatics and Data Science Department Co-Chair

FOUNDATION FOR ADVANCED EDUCATION IN THE SCIENCES   *January 2014 - May 2020*

- Co-administered an academic program with over twenty faculty members
- Taught three graduate data science courses:
    - Introduction to Python
    - Introduction to Text Mining
    - Applied Machine Learning
- Taught ten graduate biotechnology workshops:
    - Pharmacometric Analyses using R
    - Junior Scientist Training Program
    - Immunochemistry and Antibodies
    - Methods in Molecular Biology
    - Recombinant DNA Methodology
    - Research Tools for Studying Disease
    - Cancer Prevention
    - Cellular Immunology
    - Genetics Institute
    - Flow Cytometry

Online Course Developer

DataCamp *October 2018 - Present*

- Developed an online course called Creating Robust Python Workflows, part of Coding Best Practices with Python track

Postdoctoral Fellow

National Cancer Institute *July 2015 - June 2017*

- Conducted genomic analysis of immune and cancer cells
- Developed and tested therapeutic bispecific antibodies
- Mentored a trainee from the NIH High School Scientific and Training Enrichment Program

Adjunct Professor

George Washington University *January 2015 - May 2015*

- Taught two undergraduate courses:
    - Biology of Organisms
    - Women and Leadership

Predoctoral Fellow

National Institues of Health *July 2010 - October 2015*

- Tested kinase inhibitor and monoclonal antibody therapeutic efficacy and drug interactions
- Engineered and assayed therapeutic immunotoxins
- Quantified cancer cell signaling pathways
- Mentored trainees from three different NIH training programs:
    - Summer Internship Program
    - Medical Research Scholars Program
    - Biomedical Research Training Program for Individuals from Underrepresented Groups

Education

MPH, Epidemiologic and Biostatistical Methods for Public Health and Clinical Research

Johns Hopkins University School of Public Health *May 2018*

- Data Science Certificate

PhD, Tumor Biology

Georgetown University-NIH Graduate Partnerships Program *May 2015*

- Teaching Certificate

MS, Biotechnology

Jagiellonian University *June 2009*

BA, Biology

St. Mary's College of Maryland *May 2007*

Awards

Trans-Fellowship Research Award

Cancer Prevention Fellowship Program *June 2019*

Summer School on Modeling Immunology Scholarship

Emory University *May 2017*

Cancer Research Training Award

National Cancer Institute *November 2015*

Fellows Award for Research Excellence

National Institutes of Health *June 2015*

Director's Award

National Heart Lung Blood Institute *October 2014*

Certificate of Appreciation

US State Department *July 2014*

Best Poster Award

Targeted Anticancer Therapies Congress *March 2014*

Certificate of Appreciation

Rockville Science Center *January 2014*

Orloff Science Award

National Heart Lung Blood Institute *January 2014*

Intramural Research Training Award

National Institutes of Health *February 2012*