MARINA VABISTSEVITS

Doctoral student at the University of Bristol, working in the interdisciplinary research of applying data mining methods to answer epidemiological questions

EDUCATION

2023 | 2019

PhD, Data mining in Epidemiology

University of Bristol — Bristol, UK

- Thesis: "Data mining breast cancer epidemiological relationships"

2017 | 2015

MSc, Bioinformatics

University of Copenhagen — Copenhagen, Denmark

- Master's thesis: "Identification of autophagy signatures in breast cancer using The Cancer Genome Atlas data"

2015 | 2011

BSc, Biochemistry

University of Bath — Bath, UK

- Final year project: "Structural analysis of substrate binding in Sulfolobus solfataricus 2-keto-3-deoxygluconate aldolase variants"

RESEARCH EXPERIENCE

2023 | 2019

Doctoral Student

MRC Intergative Epidemiology Unit — University of Bristol

- Mini-project 1: Implemented a generalisable R workflow of performing multivariate correlation analysis (metaCCA) on several GWAS, using summary statistics from MR-Base, in order to find pleiotropic variants across the measured traits
- Mini-project 2: Carried out a Mendelian Randomisation study to investigate the mechanism mediating the effect of early life BMI on breast cancer risk
- Main PhD project: Working with EpiGraphDB, a Neo4j graph database, to answer causal relationship questions in breast cancer and build a comprehensive model of the disease aetiology, by applying data mining and machine learning methods

2017 | 2016

Visiting Researcher / Master's Thesis Student

Danish Cancer Research Centre — Copenhagen, Denmark

- For my master's project, I was invited to the DCRC as a visiting researcher after building positive working relationships during collaborations earlier in my master's degree.
- Explored TCGA breast cancer gene expression RNA-Seq data to identify the involvement of autophagy-related genes in certain disease subtypes.
- Performed an extensive EDA, followed by differential expression and enrichment analyses, allowing me to find over-represented autophagy genes
- Built an efficient workflow in R, and identified and applied a new method to this field, which was enthusiastically adopted by other group members

CONTACT

marina.vabistsevits@bristol.ac.uk
marina_vab
github.com/mvab
marinalearning.netlify.com
linkedin.com/marina-vabistsevits-82590443

TECH SKILLS

- R / tidyverse
- Python
- Bash / awk
- Git
- SQL
- Neo4j / Cypher
- Docker
- AWS
- Tableau

OTHER SKILLS

- Data visualisation
- Reproducible research
- Documentation writing
- Remote collaboration

CV made with the R package pagedown.

Code available at github.com/mvab/cv.

Last updated on 2021-02-07.

INDUSTRY EXPERIENCE

2021 | 2020

Data Scientist (Senior Executive Officer)

Public Health England — UK

- Currently undertaking a 6-month part-time secondment at PHE Data Surveillance and Management team, working within the Contact Tracing cell to assist with COVID-19 transmission data analysis
- Working with a large group of epidemiologists and R developers on processing and analysing the daily cases and contacts data to extract useful insights into the pandemic development in the UK
- Contributed to the investigation of the new variant identified in December 2020, which was published as a technical briefing for the government

2020 | 2017

Bioinformatician

Living DNA — Frome, UK

- Led the research work on improving the ancestry reference panels used by the core pipeline behind the company's direct-to-consumer ancestry genetics test, bringing considerable improvement to results accuracy
- Gained experience in working with a legacy codebase through maintaining and contributing to the in-house pipelines (Python)
- Honed my R programming skills by switching to tidyverse approach and advanced my data visualisation skills

2017 | 2016

Student research assistant in the Big Data group

3Shape — Copenhagen, Denmark

- Performed data preparation and visualisation tasks in Python, gaining practical experience of programming in a professional environment
- Used deep learning framework Caffe2 to develop a neural network training pipeline for scan image classification tasks

2014 | 2013

Placement Student in Bioinformatics team

Oxford Gene Technology — Oxford, UK

- Became responsible for a multitude of exome- and RNA-seq projects, running in-house data analysis pipelines and performing custom analysis for different projects

EXTRA CONTRIBUTIONS

2020

COVID-19 data analytics Hackathon

by TrueCue and Women in Data

2020

Food Standards Agency Data Visualisation Challenge

by Jean Golding Institute and FSA — Runner-up prize

2018

Air pollution in Bath hackathon with Bath:Hacked group

by BreATHe project by B&NES Council

2018

Hackathon for Dogs Trust charity

by Microsoft and R-Ladies London

2017

Hackathon for Danish Cancer Society

by Deloitte and Danish Cancer Society — Best Technical solution Award