Biostatistician

Romane LE GOFF
+33 608831762 | romane.le-goff@laposte.net | romanelgff.github.io

Skills

Programming: R (RShiny, caret, keras), Python (tensorflow, sklearn, pandas), SAS, SQL
Data Science Algorithms: Regression (Multiple Linear, Ridge, Lasso), Classification (PPV, SVM, Decision Trees, Random Forests), Text Analysis (NLP), Clustering (K-Means, Hierarchical), Deep Learning (CNN)
Epidemiology: Analysis and modeling of epidemiological data, Risk factor assessment, Evaluation of the impact of health interventions, Interpretation of research results

Education

Rennes 1 University - ENSAI Rennes, France
Master in Applied Mathematics, Statistics track Public Evaluation and Decision (Merit) *Sept 2020 – Sept 2022*

- European Master in Official Statistics

University of Plymouth Plymouth, United Kingdom
Year 3 Applied Mathematics, Economics - Erasmus+ (Merit) *Sept 2019 – May 2020*

University of Western Brittany Brest, France
Degree in Mathematics and Computer Science Applied to Social and Human Sciences (L1-L2) *Sept 2017 – May 2019*

Experience

Biostatistician Sept 2022 – Present
IQVIA *Courbevoie, France*

- Experience with the National Health Data System (SNDS) and IQVIA databases (Electronic Medical Records, pharmacy dispensing).

Data scientist intern March 2022 – Aug 2022
IQVIA *Courbevoie, France*

- 6-month research internship aimed at providing and comparing methods for grouping similar treatment trajectories in order to identify "typical patient paths". Presentation of the internship at the 2022 Biostatistics days (SFdS).
- *Programming: SAS, R*

Data visualization intern April 2021 – July 2021
CECLANT - Maritime Prefecture of the Atlantic *Brest, France*

- Creation of automated dashboards for the Commandant of the Brest-Lorient Defense Base and for the Social Action of the Armies (ASA) with R Shiny.

Projects

Academic | *R* 2020-2021

- Fully interactive web application: mapping, graphical representation, and informative statistical tables on COVID-19 vaccine deliveries/stocks in early 2021. (website). *Done with RShiny*
- Modeling of coronary artery disease using Bayesian networks, with the aim of better understanding diagnoses related to this disease. (GitHub). *Done with R*

Personal | *Python* 2021

- Creation of conventional neural networks from 16,000 images of clothing, with the aim of classification and labeling into 3 categories and 17 subcategories. (GitHub) This project was carried out as part of an online course (365DataScience). *Done with Python (Tensorflow)*

About me

Languages: French (Native), English (C1), Spanish (A2), Italian (A2)
Unique Experiences: 8km BUCS Cross Country Championships (Edinburgh 2020); Paris Marathon 2022