Summary

I am a Data Engineer at the end of a Ph. D. thesis in Biostatistics.

Passionate by Data Analysis and web development, I enjoy programming and automating routine tasks on web interfaces. Last 03 years, I have worked on spatial analysis of seafloor data. I have proposed Machine Learning based methodology and algorithms for geomorphological mapping.

Education

2020 | Master's Degree

Applied Mathematics and Computer Science Option: Statistics

UNIVERSITY OF LILLE (FRANCE)

2017 | Master's Degree

Econometrics and Quantitative Finance
UNIVERSITY CHEIKH ANTA DIOP (SENEGAL)

2011 | Baccalaureate

Scientific option

COURS SACRE-COEUR (SENEGAL)

Languages

FRENCH ENGLISH (**B1** – CECR) • • • • • •

SPANISH

• • • • • •

Work Experience

Ph. D. student in Biostatistics

Institut Montpelliérain Alexander Grothendieck (IMAG)

Montpellier University | France

Oct. 2020 to present

- Propose a Machine Learning based methodology to automate seafloor geomorphology mapping.
- Communication through seminars, international conferences, colloquia, ..., to share the results of my research.
- Develop skills around ethics and reproducibility in research, Data management, High Performance Computing (HPC), R packages development, etc.
- Teach Statistics to university students.

Spatial Data Analyst (intern)

Lille University | France

Feb. - Aug. 2020

- Data wrangling: Data cleaning, Missing values handling, Feature engineering.
- Explanatory analysis: Univariate and bivariate and correlation analysis.
- Dimensionality reduction: Factor Analysis (PCA, NMDS, etc.).
- Spatial data analysis: Spatial interpolation (Kriging, IDW, Regression Kriging, Clustering, etc.).

Sales Advisor

Premium Contact Center International | Senegal

Jun. 2016 - Jun. 2017

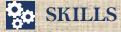
- Improve brand awareness.
- Support for the use of the brand's products and services.

Key projects

- Creation of a website for event equipment leasing.
- Creation of RShiny application for quantitative data analysis.
- Development of a R package for spatial data analysis and visualization.
- Creation of VBA macros to customize cover letters templates generation.

Hobbies

I like to draw, write short narrative texts/poems, read, go for a



Machine Learning

- <u>Data wrangling</u>
- Cleaning, Missing data and outliers handling, feature engineering.
 - R: tidyverse, dyplr, stringr, etc. | Python: Numpy, Pandas, SciPy.
- Explanatory analysis
- Univariate / Bivariate / Mutivariate and Correlation analysis, Visualization.
 - R: FactoMiner, ade4, ggplot2, plotly, etc. | Python: Matplotlib, Seaborn.
- <u>Supervised / Unsupervised Learning</u> Regression, Classification,
 - Dimensionality reduction, Clustering

 R: caret | Python: Scikit Learn,
 TensorFlow.

❖ Big data

- High Performance Computing (HPC)
 - Cluster MESO@LR.
- Parallel computing
 - > Parallel, Multidyplr.

❖ Programming Languages

• R, Python, SAS, SQL, VBA, LaTeX, HTML, CSS, PHP, Java Script, Bootstrap.

❖ Softwares / IDE

- RStudio, Jupyter Notebook, Azure ML, Anaconda, Git, Github.
- Tableau, Gretl, Eviews, Stata, GeoDa
- Excel, PowerPoint.

❖ Databases Systems

- MySQL, DB Browser, PostgreSQL, Access, SQL Server.
- Xampp, PhpMyAdmin.



- SQL pour la Data Science de A à Z (analyse de data réelles)
 SAS PROGRAMMING 1 : ESSENTIALS
 - Python 3: Formation intensive à la programmation
 - * Machine Learning with Scikit Learn
 - * R Package development