



Passionate about statistics and machine learning, I have been developing models for four years in order to solve problems in various industries.

## EXPERIENCE

### Data Scientist

CM-CIC Leasing Solutions, Paris | 06/2019 - .../....

- Did residual value forecasting using panel data analysis
- Evaluated business introducers with a random forest classifier
- Estimated the cost of material failures with a model for recurrent events
- Developed an anomaly detection module for the compliance department using LOF

### Math Teacher

Secondary school Grandmont, Tours | 09/2017 - 03/2019

### Data Scientist

Lazada, Hanoï | 02/2014 - 11/2016

- Did customer churn prediction using PCA and SVM
- Developed a fraud detection model and worked with imbalanced datasets
- Assessed customer satisfaction analysis using Natural Language Analysis

### Data Scientist Intern

Hardis, Nantes | 02/2013 - 09/2013

- Developed the blood sugar tracker app ©MobiDiab
- Oversaw the project from start to finish and managed an undergraduate intern

## EDUCATION

### M. Sc. in Statistics & Data Science

ISUP | 2019 - 2020

- Theory : deep learning, machine learning, time series, survival analysis...
- Practice : projects available on GitHub

### Math Education Degree

Tours | 2017

### M. Sc. in Engineering

Institut Mines Télécom | 2010 - 2013

- Probability theory, statistics, computer Science : Python/R/C++/SQL
- Internship abroad in Statistical Process Control

### Bachelor of Mathematics

Rennes | 2008 - 2010

## VOLUNTEER WORK & INTERESTS

- Descriptive / inferential statistics for the French Society of Professional Home Birth Care (APAAD)
- GNU/Linux, Free and Open Source Software, Functional programming (Haskell)

## REFERENCES & DIPLOMAS

- My **references** and copies of **diplomas** are available on request.

## CONTACT

- 📍 Paris, France
- ✉️ French citizenship
- ✉️ contact@datatrigger.org
- 👉 www.datatrigger.org
- 👉 github.com/datatrigger
- 👉 @vincentgoualher

## SKILLS

- ▶ Deep Learning
- ▶ Supervised Learning
- ▶ Clustering
- ▶ Feature Selection / Dimensionality Reduction
- ▶ Advanced Statistics
- ▶ Time Series, Reliability

## TOOLS

### 🐍 Python

- ▶ Pandas/Spark
- ▶ Plotly/GraphX/Dash
- ▶ sklearn/TensorFlow/MLLib

### QR

- ▶ The Tidyverse
- ▶ caret/ggplot2/shiny...

### leftrightarrow Others

- leftrightarrow SQL/C++
- leftrightarrow Git/Bash
- leftrightarrow Markdown/LATEX

## LANGUAGES

### French



### English



TOEIC 2020 : 990/990

### Vietnamese



### German

