Vincent Roger - PhD

Data and Machine Learning Scientist

@ Vincent.Roger.Pro@proton.me
1 Rue Saint-Dominique; 31000 Toulouse
vincent-roger.fr
+336 33 25 61 97
vroger11
vroger11
Vincent-Roger

Skills

Data science techniques

- Supervised, non-supervised and semi-supervised learning.
- Deep Neural Networks on big data and limited data (few-shot).
- Sequence learning: RNN, GRU, HMM.
- Representation learning using generative models such as GAN, VAE or DPGMM.
- Signal processing (on audio and images).
- Big data analysis and visualization.
- Prototype creation.

Data science tools

- Python, Numpy, Scipy
- PyTorch, Scikit-learn
- Pandas, Modin, Vaex
- Bokeh, Plotly, Pygal
- Streamlit

Management

- Organized (take notes and classify)
- Collaborative work (supervise meetings)
- Respect of due times (to publish)
- Risk Management (measure choices)
- Leadership (can supervise people)

Linguistics

- French
- English

Work Experience

PhD, three years and three months, **IRIT** — **Toulouse** 2018–2022

People with ENT cancers have speech difficulties after surgery or radiation therapy. It is important for the practitioner to have a measure that reflects the severity of speech. I propose two approaches to create an automatic measure, although with little data (about 1h of audio recordings for 128 speakers). The first one is based on "few shot" methods, while the second one is based on entropic measurement of speech features (learned with a self-supervised model on an annexed corpus). Our results on the latter have allowed us to consider a medical application. Thus, I obtained a grant to supervise an engineer to realize an application delivered to the Toulouse University Hospital.

Study Engineer, two years, **LIS** — **Toulon** 2016–2018

Following my previous contract, I created a deep self-supervised model representation of underwater acoustic environments to help categorize the different behaviors of cetaceans within range of buoys. I then created a deep model for the classification of 1500 bird species. For these two problems, I had large volumes of data.

Study Engineer, teen months, **LIS - TVT Innovation** — **Toulon** 2015-2016

Following the installation of buoys in the sea equipped with microphones, we have large quantities of data. My work consisted in modeling the bioacoustic environment using generative models. Thanks to this, I was able to produce a report of narwhal activities in relation to lunar activity.

Study Engineer, five months, **IRIT** — **Toulouse** 2015

Temporal planning by compiling satisfaction of temporal constraints. Theoretical and experimental analysis of temporal problem representation languages. Publication of an article.

Junior Software Engineer, four months, **LAAS** — **Toulouse** 2014

Modernization of the humanoid robot motion management system from GEPETTO team (representation of joints, bodies and position) by accelerating the system initialization and improving the programming interface.

Junior Software Engineer, two months, **IRIT** **Toulouse** 2013

Developed a software for automatic subtitling of audio-video streams (stream manipulation with real-time transcription). The software has become a demonstration tool of the SAMOVA research team.

Junior Software Engineer, two months and a half, **CEICOM** **Toulouse** 2011

Porting a communication application from Windows to Linux. Result: C++ rewriting of Windows routines for Linux; successful porting with a strategic impact.

Teaching

Substitute Teacher, Paul Sabatier University **Toulouse** 2019-2021

- **Model and computer science** (14h of project class): Supervision of Master projects for machine learning on automatic speech recognition.
- **Introduction to TensorFlow** (6h of tutorial class): for Master students.
- **Probability and statistics** (16h of tutorial class): basic use of the statistical functions Scipy and Numpy for Master students.
- **Introduction to Python** (116h of tutorial class): for Bachelor students in computer science, economics and biologist Master students.
- **Introduction to Network** (10h of tutorial class): basic knowledge of the OSI model and use of basic commands to describe the state of the network for undergraduates.

Substitute Teacher, University of Toulon **Toulon** 2017-2018

- **Machine Learning basis** (8h of practical class): teaching and writing of practical works on the use of neural models with TensorFlow for Master students.
- **Basic Algorithmic** (8h of tutorial class): algorithmic proofs and sorting algorithms for undergraduates.
- **Graph Theory** (12h of tutorial class and 57h of practical class): teaching and participating in the writing of graph theory subjects for undergraduates.

Education

PhD, Computer Science, **Paul Sabatier University** **Toulouse** 2022

I improved my communication during presentation and on the radio.

Master Degree, Artificial Intelligence, **Paul Sabatier University** **Toulouse** 2015

Statistical models, Signal Processing, Pattern Recognition, Robot control, and Management.

Bachelor Degree, Computer Science, **Paul Sabatier University** **Toulouse** 2013

Fundamental in development tools, low-level programming, Statistics, Probabilistic and Calculus.

University degree in technology, Computer Science, **IUT Paul Sabatier** **Toulouse** 2011

Two-year degree in technical Computer Science skills and ways to design applications for the industry.

Hobbies

Associative: I am an active member of Toulouse dataviz association since September 2022. More details on the association here: toulouse-dataviz.fr

Blogging: I blog about my projects, especially about visualizations I have done and tips about the Linux environment. It's here: vincent-roger.fr/blog

Sports: Weight training (twice a week) and running (thrice a week). I finished the marathon of Toulouse.

Readings: I read about productivity, health and sport. Also, I read Japanese Shōnens.