Lea Gauthier, PhD
Data Scientist

@ lea.gauthier.ai@gmail.com
+1 438 933 4654
https://LeaGauthier.github.io
linkedin.com/in/lea-gauthier-ai

With more than 10 years of experience in the field of particle physics and data science, I have acquired an ability to unravel complex problems and share their solutions comprehensively to colleagues and business partners. Furthermore, within various high-performance computing environments, I have cultivated a profound knowledge in data mining, ETLs and data visualization.

In my current position, I am working on mining, pre-processing data and extracting knowledge using deep learning, Bayesian statistical modelling and various machine learning algorithms. I am skilled in BigData Analytic, Machine Learning, Statistical Data Analysis, etc. More details on my experience follow.

Professional Experience

Now – August 2016 — **Senior Data Scientist,** Ariann Solution/Guillemot R&D, Montréal

- ✓ Developed algorithms and machine learning techniques in the Finance, Gaming and Advertising sectors :
  - › Lead developer in a sentiment analysis : analyzed the impact of tweets on the market using topic modelling (Latent Dirichlet Allocation model) and study of news impact on the market using Embedding and LSTM model of Keras
  - › Developed a new market prediction model for stock price and fundamentals using market data (economics, consensus estimate, ...) and applied it with Interactive Broker for real-time investment
  - › Analyzed Real Time Bidding (RTB) data to optimize online advertisement revenues with One-Hot encoder, RandomForest model with Scikit-Learn and dense neural network with Keras
  - › Lead developer in an analysis of music data to predict the next trending song using Bayesian statistical modelling. Developed an API to infer results as a JSON according to the user's options.
  - › Developed a model to extract the beat-per-minute in songs in order to automatically mix music for DJ platforms
- ✓ Developed a C++ library to quickly extract Real Time Bidding data from Aerospike and created a python wrapper for it with PyBind11
- ✓ Developed a web interface for data visualization with Dash to present the algorithm's result to clients

TensorFlow | Keras | Scikit-Learn | Gensim | python | C++ | PyBind | HTML | PHP | JavaScript | Dash (Plotly) | Bokeh | MySql | Aerospike | git

August 2016 – August 2015 — **Researcher-Developer,** Gameloft, Montréal

- ✓ Worked on pattern recognition and machine learning techniques :
  - › Empowered businesses by studying the sales impact from advertisements with Time Series Analysis (SARIMAX method)
  - › Developed a geolocalization model (C++ programming) for better ad matching
  - › Developed algorithms for RTB on different platforms (Appnexus, Spotx) to increase the number of click-per-minute of online advertisement banners and videos
  - › Worked on topic modelling and natural language processing to define new categories of mobile games
- ✓ Developed a web interface for data visualization of the sales' impact
- ✓ Utilize database management and collaborative Tools

Scikit-Learn | python | C++ | PHP | JS | HTML | MySQL | svn | Jira | Apache Cassandra

July 2015 – October 2012 | **Post doctoral Fellow,** Université de Montréal, Montréal/Genève

- ✓ Worked on the ATLAS experiment at the LHC (Large Hadron Collider)
  - › Optimized and expanded signal extraction strategies using Boosted Decision Trees
  - › Developed a Likelihood method to spin down the reconstruction error of the electron's charge identification and correct for it
  - › Coordinated the analysis of Chargino-Neutralino pair production in same-sign dilepton events
  - › Contributed on the development of a new particle subdetector
- ✓ Taught PHY1902L - Électricité et optique at Université de Montréal
- ✓ Popularized science for Cegep students (organising conferences and exercises sessions)
- ✓ Published four scientific articles

C++ | python | Root | svn | Jira | LaTeX

September 2012 – October 2009 | **Doctorant-Enseignant,** CEA-Saclay, Paris/Genève

- ✓ Thesis on fundamental physics at the ATLAS experiment at the LHC (worked on hardware, experimental and theoretical physics)
  - › Developed a model for data mining and pattern recognition with the goal of affirming or refuting a new theory. This involved PB of data (big data) and the use of Bayesian statistical modelling and probabilistic machine learning with Boosted Decision Tree.
  - › Empowered the trigger system of the Level 1 Electromagnetic Calorimeter by improving the calculation of the energy reconstruction of the particles (data clustering at scale)
  - › Developed of a new particle physics model to explain the origin of Dark Matter
- ✓ Taught Activities at Université Paris-Sud (Electromagnetic and Optic)
- ✓ Published five scientific articles

C++ | python | Root | svn | Jira

Education

2012 PhD in physics CEA-Saclay
2009 Master's degree in fundamental and applied physics Université Paris-Sud
2007 Bachelors in physics Université Paris-Sud/Nice

Skills

Programming Python, C++, Root, LaTeX, Plotly, Dash, Bokeh, PyBind
Machine Learning TensorFlow, Keras, Scikit-Learn, Gensim
Web HTML, PHP, JavaScript
Database phpMyAdmin, MySQL, Aerospike, Parquet
Tools git, svn, Jira
Communication Strong skills in team work and communication/presentation
Language French and English

Interests

- Rowing
- Cross-country skiing
- Yoga