LEA GAUTHIER, PHD

Data Scientist

@ lea.gauthier.ai@gmail.com

1 +1 438 933 4654

% https://LeaGauthier.github.io

in linkedin.com/in/lea-gauthier-ai

With more than 10 years of experience in the field of particle physics and data science, I have demonstrated an ability to unravel complex problems and share their solutions comprehensively to collegues and business partners. Within high-performance computing environments, I have cultivated a deep knowhow in data mining, ETLs and data visualization.

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



PROFESSIONAL EXPERIENCE

Now January 2018

Senior Data Scientist, ARIANN SOLUTION/GUILLEMOT R&D, Montréal

- ✓ Developped algorithms and machine learning techniques in the Finance, Gaming and Advertising sectors (Deep Neural Networks with TensorFlow and Keras, Random Forest with Scikit-Learn, Timeseries analysis, Bayesian statistical modelling and probabilistic machine learning):
 - > Lead developer in a sentiment analysis: analyzed the impact of tweets on the market using topic modeling (Latent Dirichlet Allocation model) and study of news impact on the market using Embedding and LSTM model of Keras
 - > Developped a new market prediction model for stock price and fundamentals using market datas (economics, consensus, ...) and linked it with Interactive Broker for investment
 - > Analyzed Real Time Bidding data to optimize Online Advertisement Revenues with One-Hot encoder, RandomForest model and Dense Neural Network
 - > Lead developer in analysis of music datas to predict the next trending song using Bayesian statistical modelling. Developped an API to inferred results as a JSON according to the user's options.
 - > Developed model to extract 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

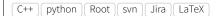
- ✓ Worked on pattern recognition and machine learning techniques:
 - > Empowered the production and business by studying the sales impact from advertisements with Time Series Analysis (SARIMAX method)
 - > Developped a Geolocalisation model (C++ programming) for better ad matching
 - > Developped algorithms for RTB on different platforms (Appnexus, Spotx) to increase the impression of online advertisement banners and videos
 - > Worked on topic modelling and natural language processing to define new categories of mobile's games
- ✓ Web interface development for data visualization of the sales' impact
- Utilization of 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

- ✓ Work on the ATLAS experiement at the LHC
 - > Optimized and expanded signal extraction strategies using Boosted Decision Trees
 - > Developped a Likelihood method to analyse the reconstruction error of the electron's charge identification and correct for it
 - > Coordinator of the analysis Chargino-Neutralino pair production in same-sign dilepton events
 - > Hardware work on the development of a new particle subdetector
- ✓ Teaching PHY1902L Électricité et optique at Université de Montréal
- ✓ Popularizing science for Cegep sudents
- ✓ Publications of four scientific articles



September 2012 October 2009

Doctorant-Enseignant, CEA-Saclay, Paris/Genève

- ✓ Thesis on fundamental physics at the ATLAS experiment at the LHC (work on hardware, experimental and theoretical physics)
 - > Developped a model for data mining and pattern recognition of big data (Terabytes) (Bayesian statistical modelling and probabilistic machine learning with Boosted Decision Tree) with the goal of affirm or refute a new theory
 - > 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)
 - > Development of a new particle physics model to explain the origin of the dark matter
- ✓ Teaching Activities at Université Paris-Sud (Electromagnetic and Optic)
- ✓ Publications of five scientific articles



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 for team work and communication/presentation

Language French and English

♣ Interests

- Rowing
- Cross-country skiing
- o Yoga