HAIR ALBEIRO PARRA BARRERA

Montreal, Quebec, Canada. Mobile.: 514-586-8551
Email: jair.parra@outlook.com / hair.parra@gmail.com
GitHub: https://github.com/JairParra
LinkedIn: https://www.linkedin.com/in/hair-parra-526ba19b/
Porfolio : https://jaiparraml.com
Portfolio Blog: https://blog.jairparraml.com/
Medium: https://hair-parra.medium.com/
Coursera: coursera.org/user/cb8fa0e883e97055aa51ec103786dbe0

WORK HISTORY

Data Scientist – DevX Analytics	Cisco - Canada	*May 2020 - Present*

- Technical product development and testing using analytics, data science, machine learning and any other related tools. Implementation of end-to-end **data science pipelines**, from data extraction, engineering, modeling, forecasting and rest API Elastic prediction index setting.
- **Defect Foreshadow:** Directly contributed to the creation of an integrated **pre-customer shipment release recommendation system** based on a multi-layer architecture including classification, regression, KDE and time series forecasting supervised machine learning models and advanced ensemble techniques such as bagging, boosting and stacking and deep learning on product and component level defects, used by thousands of users per month.
- **NextML Ecosystem**: Directly participated in the creation of **an internal machine learning ecosystem** including creating training tutorials on classical machine learning and deep learning, as well as custom ML modules for specific Cisco use cases, adopted by hundreds of engineers per month.
- **Connected-workflow:** Manager and lead engineer-level workflow automatization of systems, releases and error control dashboard including Big-Data aggregation and **automatic document parsing using integrated statistical NLP TFIDF N-gram BOW analysis** and ML inference.

NLP & Data Science Research Assistant	DemoTip Lab	*June 2019 – January 2020*

- **Multilingual NLP Research**: Perform research and tests on NLP tools with libraries such as Python re, nltk, spaCy, genism, translator, spellchecker, and R quanteda, spacyR, etc. for treatment of English, French and Spanish. Perform lab-project-related linguistic research.
- **Lexicoder Lemmatizer** *:* Create and implement NLP algorithms for lemmatizing, extending and improving lexical sentiment dictionaries in English, French, Spanish and Italian; Design and implement evaluation metrics and testing.

EDUCATION

McGill University	Montreal, QC	*September 2016 – May 2020*

- B.A. **Computer Science** with double minor in **Statistics** & **Linguistics.** CGPA: 3.61 – Distinction

Computer Science	*Mathematics*		*Linguistics*
- Object Oriented Programming	- Calculus I, II & III	- Generalized Linear Models	- General Linguistics
- Algorithms and Data Structures	- Honours Applied Linear Algebra	- Time Series Analysis	- Focus on Syntax and Semantics
- Theoretical Computer Science	- Intermediate Probability	- Self study: Abstract algebra, analysis, stochastic processes	- Formal Logic.
- Applied Machine Learning	- Mathematical Statistics		
- Natural Language Processing	- Discrete Mathematics		
- Database Systems	- Honours Regression and Analysis of Variance		
- Functional Programming.			

Vanier College	Montreal , QC	*August 2014 – Jun 2016*

- DCS / D.E.C. Languages and Cultures , **Honour Roll** (2014-2016), **Dean's Honour Roll** (Winter 2014, Autumn 2014).

TECHNICAL SKILLS

Programming Languages, Technologies & Skills:	Human Languages:	
Data Science Python, R/RStudio , SQL (MySQL, PostgreSQL), Bash/Shell, Git/GitHub, Elasticsearch, Tableau. - (NumPy, Pandas, Matplotlib, Seaborn, re, NLTK, SpaCy, genism, Scikit-learn, TensorFlow, etc.) *Other Backend* Java C, C++ , OCaml [some familiarity]. *Web Development* HTML5, CSS3, Bootstrap5, JavaScript (ECS6), jQuery.	1. **Spanish** (Native) 2. **English** (Bilingual) 3. **French** (Bilingual) 4. **Mandarin Chinese** (Upper Intermediate) 5. **Japanese** (Upper Intermediate)	6. **Italian** (Intermediate) 7. **Japanese** (Upper Intermediate) 8. **Russian** (Beginner) 9. **Cantonese** (Beginner)

PROJECTS & TECHNICAL WRITING

- **Data Science & Statistics Technical writing on Medium (+50 articles):** https://hair-parra.medium.com/
- **LDA Centroid-based Summarization:** https://github.com/JairParra/LDA_centroid_based_summarization
- **Modified MNIST: State-of-the-art Exploration** https://github.com/JairParra/Modified_MNIST_state-of-the-art_exploration
- **Kaggle_Reddit_Multiclass_Classification** https://github.com/JairParra/Kaggle_Reddit_Multiclass_Classification
- **ML Reproducibility Challenge:** https://github.com/JairParra/ML_Reproducibility_Challenge
- **Implementing Logistic Regression & LDA from scratch:** https://github.com/JairParra/Logreg_LDA_from_scratch
- **Dating Rental Database Project:** https://github.com/JairParra/Dating_Rental_Database_Project/blob/master/README.md