Rafael Lacerda

rafaelplmelo@gmail.com
Toronto, ON, Canada
(647) 878 9650

Overview

Adaptable professional with over 3 years of experience. Proven knowledge in data science, natural language processing and machine learning engineering. Seeking opportunities to further develop my craft.

Expertise

- Creating solutions to business problems using automation and machine learning.
- Research and development of ML and NLP systems.
- Developing all the infrastructure required to host machine learning applications in the cloud.

Work Experience

Machine Learning Engineer *May 2017 — Feb 2019*

Ross Intelligence Inc., Toronto, ON

- Increased the performance of our search engine by developing a document ranking system that surpassed IBM Watson's.
- Explored advances in Deep Learning and NLP literature to implement document ranking prototypes. Used Python, Spacy, Tensorflow, Keras.
- Confirmed our model's superiority by developing blind statistical tests using human assessments.
- Improved search performance by developing better data cleaning and machine learning features.
- Ensured the platform ran quickly by developing pipelines for deep learning, ML and NLP models.
- Ensured the system was stable and scalable, developing microservices for the pipeline and the CI/CD infrastructure to deploy them in the cloud.
- Further improved search performance by creating specialized models to mimic human experts and extract entities in documents.
- Decreased latency by 96% by restructuring the retriever's distributed index architecture and creating a caching microservice.

Data Analyst, Treasury *Aug 2014 — April 2016*

OR Investimentos SA, São Paulo, Brazil

- Reduced treasury risk by developing regression models to estimate risk for financial institutions where CDS spreads were unavailable.
- Pioneered report automation in the department, increasing data resolution from monthly to daily, saving hundreds of analyst hours monthly:
  - Created an Extract Transform Load (ETL) task manager for Excel to enable automated reports with 1000+ dependencies. Modeled tasks using Directed Acyclical Graphs. Used VBA and Python.
- Increased visibility of cash balances, saving over 3M USD monthly:
  - Created an aggregate visualization of excess cash in over 300+ regional branches' balances, allowing us to invest every end of day. Used VBA.

Education

MSc. in Applied Computing *Sep 2016 — Dec 2017*

University of Toronto, *Department of Computer Science*

- Courses of note: Natural Language Processing, Computational Neuroscience, Probabilistic Learning and Reasoning.

- Recipient of 30k Mitacs scholarship (2x terms).

Data Science Bootcamp *July 2016 — Aug 2016*
Signal Data Science, *Berkeley, CA*
- 500 hours reviewing a wide array of data science theory and practice.
- Developed a recommendation system for movies based on temporal flow of Ekman emotions.

BSc. in Economics *Aug 2011 — Jul 2014*
Universidade Presbiteriana Mackenzie, *Department of Applied Social Sciences*
- Previously a student of Computer Science at the same university:
  - Transitioned into Economics to focus on modeling human activity. Since then I have used my CS skills as tools for modeling and automation.
  - Completed the first three semesters of CS fundamentals. *Courses of note*: Optimization Algorithms, Data Structures, Linear Algebra, Statistics, Calculus, C Programming.

Technical Skills

Languages & Frameworks: Python, R, Tensorflow, Keras, Numpy, Pandas, Scikit-learn, Spacy, NLTK, Matplotlib.
Databases, Caching: SQL, NoSQL, MongoDB, Redis, Solr.
Cloud computing: AWS ecosystem, Microservices architecture, REST APIs.
Continuous Integration, Continuous Delivery: Github, Travis CI, Docker, Kubernetes.
Machine Learning: Regression, Classification, Clustering, Dimensionality Reduction, Model Validation, Model Selection.
Deep Learning: NN, CNN, RNN, LSTM, GRU, Autoencoders.
Natural Language Processing: Embeddings, Information Retrieval, Natural Language Understanding, Text Classification, Topic Modeling, Question Answering, Learning to Rank, Summarization.

Projects

- Computational graph framework to chain video filters. Used OpenCV, Python, Numpy (2019).
- Tensorflow-like API to chain image filters that operate spatial and temporal dimensions.
- Emotion based movie recommendation system. Used Python, BeautifulSoup, Numpy (2016).
  - Scraped movie scripts annotated with Ekman emotions throughout the timeline. Recommendations through similarity of emotional progression.
- BSc. Economics thesis on a time series model to measure the effects of uncertainty brought upon by Federal intervention on oil policy. Used the Gretl stats package for the regression model (2014).
- 6 month internship in real estate management (OR Investimentos) on financial modeling. Automated processes using VBA (2014).
- 6 month internship at a hedge fund (Kondor Invest, risk office) clearing daily trades and automating processes. Used Python, SQL and VBA (2012).