Rafael Lacerda

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

Overview

Adaptable professional with ~5 years of experience and 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.

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).

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 to assist in modeling and automation.
  - Completed the first three semesters of CS fundamentals. *Courses of note*: Optimization Algorithms, Data Structures, Linear Algebra, Statistics, Calculus, C Programming.

Work Experience

Machine Learning Engineer *May 2017 — Feb 2019*
Ross Intelligence Inc., Toronto, ON

- Developed a state-of-the-art legal research system that surpassed IBM Watson's performance, leveraging deep learning and NLP.
- Coordinated new research that further improved the research system, by creating specialized models to extract entities in documents.
- NLP projects include information retrieval, document ranking, summarization, topic modeling, document classification and data augmentation in the legal domain.
- Deep Learning tasks included designing networks from scratch, customizing existing networks, training models and word embeddings, model selection and productionizing models.
- Built fast and scalable NLP and Deep Learning pipelines.
- Sped up models for production by using caching and precomputation.
- Developed statistical tests with human assessments to ensure model superiority before deployment into production.
- Used Python, Tensorflow, Keras, Solr/Lucene, Pandas, Spacy, NLTK, Scikit-Learn, MongoDB, Redis, Docker, Travis CI, Kubernetes, AWS.

Data Analyst, Treasury *Aug 2014 — April 2016*

OR Investimentos SA, São Paulo, Brazil

- Developed Business Intelligence systems to monitor aggregate treasury risk daily. Used Python and VBA.
- Developed regression models to estimate bank risk when CDS spreads were unavailable.
- Created a system to manage 300+ regional branches' bank balances, saving over 3MM USD monthly by optimizing cash efficiency. Used Python and VBA.
- Created a task manager based on Directed Acyclical Graphs (DAGs) to assist in the computation of our monthly reports, allowing resume on failed tasks and greatly reducing analyst hours. Used Python and VBA.

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, Question Answering, Learning to Rank, Summarization.

Other 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).