307-1177 Hornby St.
Vancouver, BC V6Z 2E9 CANADA
+1 (778) 819-8256
http://duboue.net/

pduboue@textualization.ca
pablo.duboue@gmail.com
https://linkedin.com/in/pabloduboue
https://github.com/DrDub

Pablo Duboue, PhD

Full stack senior ML Engineer / Data Scientist. I can design experiments, build prototypes and productize ML code for deployment within an existing infrastructure.

Highlights

- Member of the team that built IBM Watson (and defeated the Jeopardy! champions). Contributions to the ML architecture.

- ML Book ("The Art of Feature Engineering") published by Cambridge University press.

- Ten years of consulting experience, including consulting for a startup purchased by Intel Corp. Proven trajectory architecting ML production systems.

- 5 patents, 40 peer-reviewed research papers written with 50 different coauthors and cited over 1,700 times in the fields of Text Generation, Information Retrieval and Question Answering.

- Strong open source ties, including having organized the Debian Developers conference and having contributed more than 100k lines of code in Scala, Java, Python and other languages.

- PhD in Computer Science from Columbia University in the City of New York. Thesis in Machine Learning for Text Generation applied to Medical and Intelligence Analysis domains.

Focus

Consulting for projects with technical uncertainty, helping reduce the risk and secure maintainable ML solutions that evolve well over time.

Works well in larger, heterogeneous teams. Takes the time to mentor junior team members and prepare the team to outgrow his services.

Skills

Technology

Scala, Java, Spark, Python, Scikit-Learn, NumPy, OpenCV, Tensorflow, Keras, OpenAI Gym, Jupyter Notebooks, Golang, Haskell, Node.js, Elm, Bootstrap, Flask, Scalatra, Django, Git, Bash, Linux (Debian and derivatives), Docker, Weka, Mallet, UIMA, spaCy, NLTK, Knowledge Graphs, SPARQL, RDF, JSON, XML, XSLT, Rule-based systems, RuTA, ClearTk, PHP, Perl, Latex, Lua, SQLite3, PostgreSQL, JNI, Delphi, VB, Win32

Interpersonal

Experienced teacher (taught 20 college-level and above courses), experienced presenter (more than 50 public presentations), basic management experience (interim manager for a 4-people team plus managed own contractors), familiar with agile methodologies, with experience assessing bids for RFPs and assessing job candidates for ML positions

Spanish, English, French (workplace), Portuguese (conversational), Italian (basic), Chinese Mandarin (beginner), American Sign Language (basic)