

Pedro Rodriguez | Research Scientist

✉ pedro@entilzha.io • 🌐 www.entilzha.io

Experience

University of Maryland at College Park

Research Assistant to Jordan Boyd-Graber, PhD Candidate

Machine Learning, Deep Learning, NLP

August 2015 - Present

- Research experience in machine learning, deep learning, natural language processing, and question answering.
- Created an AI for a trivia game (Quiz Bowl) that beat the best players in the world. <https://youtu.be/bYFqMINXayc>

Microsoft Research - Information and Data Sciences

Research Intern

Machine Learning, Deep Learning, NLP

Summer 2018

- Conversational NLP for machine guided human learning, a work in progress to be published in 2019.

Riot Games

Data Science Research Intern

Machine Learning, NLP, Bayesian Modeling

Summer 2017

- Improved NLP model effectiveness for automatically sorting and/or answering 12,000 player support tickets per day.
- Created data labeling process for player support tickets to increase effectiveness and enable incremental future improvement.
- Used Machine Learning ablation studies, Bayesian modeling and Monte Carlo simulations to understand behavior of new sportsmanlike play system called Honor 2.0. This also included querying over 56 million records with Spark SQL.

Oracle Data Cloud/Datalogix

Data Science Research Intern

Machine Learning, Data Engineering, DevOps

Summer 2016

- Architected and implemented ETL pipeline with Apache Spark/Airflow which ingests 200GB/day of advertising data.
- Improved throughput from AWS S3 to Apache Spark by 40-60x which resulted in substantial cost savings.

Zillow Group/Trulia

Data Scientist

Machine Learning, Crowdsourcing, DevOps

January 2015 - July 2015

- Designed and iteratively improved crowdsourcing experiments to produce rank order for machine learning generated photo albums. In doing so wrote a Amazon Mechanical Turk library which managed approximately 20,000 tasks.
- Architected and implemented infrastructure for internal data science APIs built on Docker, Swagger, Jenkins, and predominantly Django as a web framework. This code is now called on 100% of visits to trulia.com.

UC Berkeley AMPLab: Undergraduate Research Assistant, Fall 2014: Optimized Gibbs LDA on Spark

Boise State University Geophysics: Cryosphere Scientist, Field Researcher, May 2013-October 2014

UC Berkeley Center for Time Domain Informatics: Undergraduate Researcher in Astrophysics, 2012

Avalanche Safety Educator: Volunteer for National Ski Patrol, Professionally in South America, 2010-Present

Publications

- Eric Wallace, **Pedro Rodriguez**, Shi Feng, and Jordan Boyd-Graber. Trick Me If You Can: Adversarial Writing of Trivia Challenge Questions. arxiv.org.
- Shi Feng, Eric Wallace, Alvin Grissom II, Mohit Iyer, **Pedro Rodriguez**, and Jordan Boyd-Graber. Pathologies of Neural Models Make Interpretations Difficult. EMNLP 2018
- Jordan Boyd-Graber, Shi Feng, and **Pedro Rodriguez**. Human-Computer Question Answering: The Case for Quizbowl. The NIPS '17 Competition: Building Intelligent Systems, 2018

Projects and Code

Qanta.org: An NSF funded project for advancing AI question answering by learning to play quiz bowl.

Tools: Python Ecosystem (PyTorch, pandas...), SQL, Apache Spark, Scala/Java, HTML/CSS/JS, Tex, Rust

Education

University of Maryland at College Park, Computer Science

PhD Candidate in Machine Learning, Deep Learning, and NLP

Expected Spring 2020

University of California at Berkeley, Computer Science

Bachelor of Arts in Computer Science with coursework in Physics and Astrophysics

2010 - 2014