Santa Barbara, CA
GitHub: adrian-pardo
Portfolio: adrian-pardo.github.io

Adrian Pardo

C: 562.716.0932
in: adrian-pardo
E: pardo.adrian@outlook.com

Data Scientist

"My passion is to leverage big data to drive strategic decisions that amplify scientific discovery for community health."

SUMMARY: Results-oriented, highly self-driven Data Scientist and clinical research professional with a background rooted in life sciences. Over eight years experience working with multi-disciplinary research teams across various health industries. I thrive solving complex problems by applying actionable, data-driven solutions.

TECHNICAL SKILLS

- Python, R, Excel, SQL
- Pandas, NumPy, Matplotlib
- ArcGIS, Tableau, Power BI
- Scikit-Learn, Tensorflow
- OpenCV, NLTK, PyTorch
- Jupyter Notebooks, R Studio
- Bash/Unix, Git, AWS

HARD SKILLS

- Natural Sciences & Economics
- Data / Quantitative Analysis
- Programming / Debugging
- Mathematics / Statistics / Modeling
- Machine Learning / Deep Learning
- Computer Vision, NLP
- Structured & Unstructured Data

SOFT SKILLS

- Communication / Interpersonal Skills
- Multi-disciplinary Team Collaboration
- Translating Technical Language
- Creative & Critical Thinking
- Problem Solving
- Self-motived Learner
- Curiosity for Discovery

PROFESSIONAL EXPERIENCE

Consultant JUN 2020 - Present

Data Scientist

- Engineered computer vision algorithms to efficiently query 10TB+ database and identify duplicate business documents with ~97% accuracy.
- Automated time-intensive data analysis processes with Python, saving client ~$5,000 per week in manual labor costs.
- Introduced clients to cloud computing tools (e.g. S3 & EC2 AWS) to reduce business costs and improve feasibility of computationally expensive tasks.
- Built interactive dashboards with Power BI to visualize company trends and provide support for future business decisions.

Cottage Health Research Institute, Santa Barbara, CA JUN 2019 - Present

Research Data Analyst

- Guided decisions of C-Level executives by using Python, R, and ArcGIS to identify epidemiological trends (e.g. COVID-19) within Santa Barbara community.
- Developed Boosted Tree algorithms to determine the relationship between patient data and *C. Difficile* status, achieving a 5% more accurate prediction of *C. Difficile* status than previous year's model.
- Served as REDCap Database Administrator for entire hospital.
- Taught educational workshops for hospital employees seeking to improve their data analysis and visualization skills.
- Created reusable, timesaving data analysis pipelines for over one million electronic health records.
- Used Power BI, Excel, and Matplotlib to illustrate data visualizations that were ultimately presented at nationwide health conferences.

PIH Health Hospital, Whittier, CA OCT 2016 - JAN 2018

Research Data Coordinator

- Managed collaborative (physicians, nurses, analysts, coordinators) nationwide project by instigating Clinical Trial Phases II-III.
- Integrated algorithms with shared-decision making to promote health literacy among underrepresented groups afflicted with osteoarthritis.
- Utilized EHR data and predictive modeling tool (e.g. Hidden Markov Model) to generate health information for each patient.
- Facilitated clean data collection by communicating with patients in both English / Spanish, translating patient feedback into discrete data, and inputting data into REDCAP database.

PROFESSIONAL EXPERIENCE (cont.)

University of California, Riverside: Department of Psychology SEP 2015 - JUN 2016

Research Assistant

- Researched iconic memory in humans by working with over 100 study participants over a 10-month period.
- Operated MATLAB software to administer computational neuroscience memory tests, electrophysiological (EEG) recordings, and eye-tracking experiments to collect quantitative information regarding human memory.

University of California, Riverside: Department of Botany & Plant Sciences MAR 2014 - JUN 2015

Research Assistant

- Performed statistical significance testing (e.g. ANOVA) to determine efficiency of biofuels from transgenic tobacco plants.
- Collected quantitative and qualitative data by completing over twenty wet lab procedures to prepare glycosyl composition of whole cell wall fraction, highly purified tobacco cell walls, and alcohol-insoluble residues.

EDUCATION

M.S. ANALYTICS
Georgia Institute of Technology

B.A. NEUROSCIENCE
University of California, Riverside

PROJECTS & COURSEWORK

- **Work & Academic Projects:**
  - M.S. Capstone Project — Dog Diseases Classification and Prediction though Genetic Composition
  - Analysis of Countywide Community-Acquired and Hospital-Associated *Clostridium Difficile* Infections
  - Developing a Machine Learning Framework to Predict Severe COVID Disease
  - Improving Developmental Positioning in a Level III NICU using Evidence-Based Teaching and a Standardized Tool: An Evidence-Based Quality Improvement Project
  - Visualization of the Opioid Crisis in the USA
- **Graduate & Upper-level Data Science/Machine Learning Coursework**:
  - Analytics & Financial Modeling (Python, R)
  - Statistical Computing (Python, R, Excel)
  - Data & Visual Analytics (Python, R, Excel, Javascript, Power BI, Tableau)
  - Convolutional Neural Networks for Visual Recognition (Python)
  - Practical Deep Learning For Coders (Python)
  - Regression Analysis (R)
  - Data Analytics for Business (R)
  - Simulation (Python & ARENA)
  - Biostatistics (Python & Excel)
  - Computational Neuroscience: Artificial Neural Networks (MATLAB)
  - Neuroscience of Learning & Memory: Biological Neural Networks (MATLAB)
  - Digital Marketing (Excel)

LANGUAGE & CERTIFICATIONS

- Spanish (Read, Write, and Speak – Fluent/Native)
- "Protecting Human Research Participants" certified by NIH
- "Essentials of Research Administration" by CITI