APRIL J. HARRY

STATISTICAL DATA SCIENTIST

Austin, TX (337)-326-2838
harry.april@gmail.com – apriljharry.com - github.com/ajharry - linkedin.com/in/april-j-harry/

SKILLS

LANGUAGES AND TOOLS: Python, R, SQL, Spark, dbt, Sisense for Cloud Data Teams

EXPERIENCE

MAY 2019 –NOW

ROVER
SENIOR DATA SCIENTIST (MAR 2022 - NOW)
DATA SCIENTIST (MAY 2019 - MAR 2022)
- Developed a survival model approach for improving customer LTV through search ranking.
- Established team processes for monitoring and triaging issues with production-facing pipelines.
- Was technical lead for a team of 5 data scientists in developing a PySpark feature library and set of machine learning models for search result ranking, significantly increasing conversion.
- Created a framework for describing the customer lifecycle in daytime pet services (dog walking and daycare), and formalized metrics to facilitate timely marketplace health monitoring based on transition rates.

APR 2018 –MAY 2019

DATA SCIENTIST, REALSELF
- Engineered SQL-based email recommendations, resulting in ~20% lift in clickthrough rate.
- Collaborated cross-functionally to build a microservice using fuzzy matching to extract review snippets mentioning sponsoring businesses. Used a similar approach to link >10,000 previously unattributed reviews to medical providers.
- Built a front-end site feature to collect data on duplicative user-generated questions.

2018–2019

TECHNICAL ADVISOR AND FELLOW, INSIGHT DATA SCIENCE
TECHNICAL ADVISOR (2018-2019)
- Conducted admissions interviews and participated in alumni panels for recruiting.
- Mentored data science fellows through project advice and mock interviews.
FELLOW (2018)
- Built a Python-based Support Vector Machine classification algorithm for identifying sponsored content in YouTube videos based on Natural Language Processing.
- Deployed algorithm through a Google Chrome extension and Flask/AWS web app

2015–2018

GRADUATE ASSISTANT/RESEARCH TECHNICIAN, NORTHEASTERN UNIV and PURDUE UNIV
- Developed a Bayesian statistical modeling and testing framework for molecular imaging data, identifying differences in the composition of biological tissues.
- Prototyped the Markov Chain Monte Carlo model estimation as an R package (msiCompare).
- Full listing of peer-reviewed publications on my LinkedIn page.

EDUCATION

2012–2017   **PHD IN STATISTICS,** Purdue University, West Lafayette, IN

2010–2012   **MS IN APPLIED STATISTICS,** Purdue University, West Lafayette, IN

2006–2010   **BS IN MATHEMATICS,** Xavier University of Louisiana, New Orleans, LA