Ryan C. Corley

ryan.corley@berkeley.edu

Summary

I am a high energy senior data scientist with a Masters from UC Berkeley and a proven track record of 6+ years driving high value analytics solutions.

Education

University of California, Berkeley

Masters in Engineering, Industrial Engineering and Operations Research	2019
Bachelor of Science, Industrial Engineering and Operations Research	2016

Certificate in Technology Entrepreneurship

Skills

Python, Spark, SQL, Jira, Scala, R, Java, MATLAB, AMPL, Excel

Experience

Student, Product School   **May 2023 – Present**

- Pursuing Senior Product Manager Certification to expand skill set in product strategy, user research, product analytics, and stakeholder management through a globally recognized program.
- Combining data science expertise with product management skills to drive innovation, lead cross-functional teams, and achieve business objectives.

Senior Quantitative Scientist, Verana Health   **2022 – 2023**

- Developed, validated and scaled algorithms to extract Bladder and Prostate cancer staging information from unstructured clinical notes. Achieved .97 F1 Score. Use PySpark, HuggingFace, Amazon Sagemaker & Ground Truth
- Led workstream for coding workspace and environment, acting as a liaison between the quantitative sciences team and DevOps, ML Ops, and Security teams.
- As workstream lead, conducted presentations to data architecture leads, documented best practices, built a process to triage blockers, and facilitated a smooth transition from EMR Studio to Databricks through an in depth PoC reducing operational costs and increasing team productivity.
- Oversaw growth of quantitative sciences team from 7 to 35 members, improving interviewing process, onboarding documentation and onboarding buddy program.

Quantitative Scientist, Verana Health   **2020 – 2022**

- Lead execution of commercial retrospective clinical studies leveraging real-world data, through creation of statistical analysis plans, implementation of analyses, development of research objectives and writing of publications.
- Collaborated cross-functionally with commercial, product, clinical & data engineering teams to translate clinical questions into detailed data analytics requirements for internal and external projects.
- Delivered and presented large projects, each ~1-2M+, to external life science partners teams. Studies leverage clinical data from both large cohorts of >300k patients and orphan disease cohorts of <15k patients. Techniques include propensity matching, mixed effects models, hypothesis testing, and structuring clinical text via NLP. Use PySpark, Amazon Athena, Amazon EMR & Jupyter notebooks.

Data Scientist, Innovaccer   **2019 – 2020**

- Developed and led the execution of the risk analytics product roadmap of a team composed of engineers, data scientists, and data analysts.

- Recruited and managed by Director of Data Science, a former colleague from Nuna Health. As a lead for risk analytics, I traveled to the office in Delhi, India three times in the four months prior to the pandemic.
- Synthesized product and customer requirements, and deployed the platform for success to drive high value care, measure and demonstrated impact created to customers.
- Developed and deployed models leveraging clinical, claims and social determinants data allowing physicians and care coordinators in inNote and inCare to triage and attend to high risk patients to operationalize value-based care. Algorithms included CMS HCC, HHS HCC, CDPS, and the development of a proprietary cost prediction algorithm. Used PySpark, Amazon Athena, Sagemaker, and Redshift.

Data Science Consultant (part-time), UCSF Institute for Computational Health Sciences   2018 – 2019

- Worked on a team to develop an algorithm to de-identify 58 million clinical notes across 139 categories to make them available for researchers. Achieved 98% retention of medical information.
- Proposed and executed on a research topic for a Capstone Masters project. Used NLP, Machine Learning and Deep Learning techniques to build a Questioning/Answering system allowing researchers and clinicians to query Electronic Health Records (EHRs) and free text Clinical Notes. Used AWS, pySpark, BERT, Flask to develop our prototype.

Data Analyst, Nuna Health   2017 – 2019

- Utilized open source and third-party tools to build and configure a data pipeline to ingest data from various healthcare claims carriers and mental health platforms to translate into digestible insights in employee maternal health, mental health, and utilization for the front-end Looker dashboards. Used AWS, SparkSQL, Apache Airflow, Python, R, Looker, and Google Big Query.
- Effectively communicated with product managers, data engineers, data scientists, and client services to deliver actionable data insights from aggregated healthcare utilization data for self ensured employer clients, such as Facebook, Home Depot or Mitre.
- Transitioned to part time upon returning to graduate school

Data Engineer, Mede/Analytics   2016 – 2017

- Worked as the lead engineer in conjunction with a product manager to prototype and develop a supply chain analytics and surgical supply utilization solution for a system of over 10 hospitals.
- Collaborated with product owners, business analysts to meet business requirements, develop, and deliver an analytics product for HCSC the largest payer conglomerate of BCBS Plans in the US.

Activities

Eagle Scout, Troop 64, CA
Wilderness First Responder
Completed iconic 'Escape from Alcatraz' Triathlon