COLIN QUIRK

email: colinquirk1993@gmail.com | website: colinquirk.com | github: @ColinQuirk

SUMMARY

I am a versatile data scientist who enjoys learning new things and tackling whatever is thrown at me. My wide-ranging experience makes me perfect for roles that need someone who can do it all, including experimentation, writing tools, building models, deployment, interacting with stakeholders and managing other data scientists.

I've picked up technologies across the entire ML lifecycle including 10 years of python experience (plus scikit-learn and pandas), pyspark, R, SQL, AWS, docker, and git. I also have extensive experience with many types of models, including regression, gradient boosting (e.g. XGBoost), and deep learning (e.g. CNNs).

WORK EXPERIENCE

SparkNeuro – Director of Machine Learning — 2020-Present
  – Previous roles: Senior Machine Learning Engineer, Data Scientist
  – Managed a team of data scientists working on core algorithm development
  – Performed feature engineering and hyperparameter tuning to improve model performance
  – Assisted with design and creation of the MLOps backend for model training
  – Tools used: python, R, SQL, AWS

Allstate – Data Science Intern — 2019
  – Created a model using xgboost to predict likelihood of a quote binding
  – Assisted with deploying my team's models into production
  – Made multiple contributions to shared machine learning libraries
  – Gave a company wide presentation about data visualization
  – Tools used: SQL, pyspark, R

EDUCATION

University of Chicago
Ph.D. in Psychology (ABD)* — Expected 2023

Hampshire College
B.A. in Cognitive Neuroscience — 2015

RELEVANT ACADEMIC PROJECTS

Classifying Individuals with Eyetracking Data using Deep Learning

  – Organized and cleaned 20GB of eye tracking data
  – Classified individual segments of eye tracking with high accuracy using CNNs
  – Tools used: Python (fastai/PyTorch), R (data cleaning)

Spatial Biases Influence Encoding of Information into Working Memory

  – Reanalyzed experiments with a large number of participants (>470000 trials analyzed)
  – Combined data from multiple, existing sources to address research questions
  – Created open source python libraries which have been used by other researchers to run experiments
  – Used k-means clustering to generate factors in datasets that were missing critical information
  – Tools used: Python (experiment code), R (analyses/visualization)

For a full list of my publications, please see colinquirk.com/publications