Jacob Reinhold

New York, NY | jcreinhold@gmail.com | linkedin.com/in/jcreinhold | github.com/jcreinhold

Summary

Data scientist with the proven ability to work in a variety of domains. History of success using rigorous statistics, machine learning, and software development skills to improve products and decision-making.

Technical skills

Python (PyTorch | scikit-learn | numpy | Pyro) | R | C | OCaml | SQL | Git | deep learning | computer vision | machine learning (ML) | causal inference | experimental design | A/B testing | statistics | AWS

Experience

Meta, New York, NY — June 2022 – May 2023
Data scientist

Used rigorous statistics, ML, and causal inference to evaluate new products and discover new revenue opportunities.

- Independently investigated and conducted rigorous statistical analyses on petabyte-sized datasets to discover a $20M and a $80M revenue opportunity for two lead generation ads products.
- Developed Python package to streamline causal inference analyses with matching, doubly-robust IPW, and double ML; used by more than 10 other data scientists to deliver analyses within tight timeframes.
- Designed experiments to evaluate efficacy of products combining observational and interventional data; analyses used for marketing statistics and to greenlight various business messaging ads products.
- Developed Python package to analytically solve for experimental design parameters in meta-analyses; used by various teams to design meta-analytic studies composed of multiple independent experiments.

Memorial Sloan Kettering Cancer Center, New York, NY — July 2021 – June 2022
Data scientist

Researched, developed, and deployed neural networks for computer vision to improve patient outcomes.

- Developed and deployed a DNN-based tumor segmentation pipeline to provide stats to clinicians; first of its kind at the institution; created pipeline infrastructure for training and deploying similar models.
- Researched and developed ML monitoring frameworks to evaluate model performance in deployment.
- Built infrastructure for and deployed MLOps tools to coordinate team of data scientists.

Johns Hopkins University, Baltimore, MD — August 2017 – May 2021
Graduate research assistant

Investigated novel machine learning and causal inference methods for computer vision and speech processing.

- Used probabilistic programming language to implement a novel causal model of disease for multiple sclerosis in MR images; resulted in peer-reviewed conference paper at top conference (MICCAI).
- Developed novel unsupervised anomaly detection technique in CT and MR images by quantifying uncertainty in an image-to-image translation task for an industry partner; resulted in two papers.
- Collected a novel emotion-in-speech dataset and investigated ways to computationally alter emotional affect; resulted in peer-reviewed conference paper at a top speech-processing conference (Interspeech).

Applied Research Laboratories, Austin, TX — November 2014 – June 2017
Engineering scientist associate

Data scientist responsible for analyzing experimental data and implementing statistical software.

- Initiated the development of a new Python package for geolocation using statistical array processing techniques on high-dimensional radio data; used in production to deliver stats to external stakeholders.
- Analyzed and visualized large scientific datasets by creating statistical software tools (with, e.g., pandas, numpy, scipy, matplotlib); results were used in a peer-reviewed scientific conference presentation.

Education

Master of Science (MSE), Electrical and Computer Engineering, Johns Hopkins University
Bachelor of Science (BS), Electrical Engineering, University of Texas at Austin

Peer-reviewed publications

1. P. Tohidi, S. Remedios, D. Greenman, M. Shao, S. Han, B.E. Dewey, **J. Reinhold** et al., "Multiple Sclerosis brain lesion segmentation with different architecture ensembles," in Medical Imaging 2022: Biomedical Applications in Molecular, Structural, and Functional Imaging, 2022.
2. **J. Reinhold**, A. Carass, and J. L. Prince, "A Structural Causal Model for MR Images of Multiple Sclerosis," in International Conference on Medical Image Computing and Computer-Assisted Intervention, 2021.
3. A. Carass, S. Roy, A. Gherman, **J. Reinhold** et al., "Evaluating white matter lesion segmentations with refined Sørensen-Dice analysis," Scientific reports, 2020.
4. **J. Reinhold** et al., "Validating uncertainty in medical image translation," in 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), 2020.
5. **J. Reinhold** et al., "Finding novelty with uncertainty," in Medical Imaging 2020: Image Processing, 2020.
6. J. Sager, R. Shankar, **J. Reinhold**, and A. Venkataraman, "VESUS: A crowd-annotated database to study emotion production and perception in spoken English.," in Interspeech, 2019.
7. B. E. Dewey, C. Zhao, **J. Reinhold** et al., "DeepHarmony: A deep learning approach to contrast harmonization across scanner changes," Magnetic resonance imaging, 2019.
8. **J. Reinhold**, B. E. Dewey, A. Carass, and J. L. Prince, "Evaluating the impact of intensity normalization on MR image synthesis," in Medical Imaging 2019: Image Processing, 2019.
9. G. Wen, H. Chang, **J. Reinhold**, J. Y. Lo, and M. K. Markey, "Virtual assessment of stereoscopic viewing of digital breast tomosynthesis projection images," Journal of Medical Imaging, 2018.
10. **J. Reinhold**, G. Wen, J. Y. Lo, and M. K. Markey, "Lesion detectability in stereoscopically viewed digital breast tomosynthesis projection images: a model observer study with anthropomorphic computational breast phantoms," in Medical Imaging 2017: Image Perception, Observer Performance, and Technology Assessment, 2017.

Published writing

1. **J. Reinhold** et al., "A technical and regulatory perspective on GANs in medical devices", 2021.
2. **J. Reinhold**, "3D Medical Image Analysis with PyTorch", Manning, 2020.
3. **J. Reinhold** et al., "How to choose a neural net architecture for medical image segmentation", 2020.
4. **J. Reinhold** et al., "Get more out of your annotated medical images with self-supervised learning", 2020.
5. **J. Reinhold,** "Knowing known unknowns with deep neural networks" in Towards Data Science, 2020.
6. **J. Reinhold**, "Dropout on convolutional layers is weird" in Towards Data Science, 2019.
7. **J. Reinhold**, "Deep learning with magnetic resonance and computed tomography images" in TDS, 2019.

Projects

- intensity-normalization – Python package for preprocessing medical images for ML (280+ stars).
- counterfactualms – Python package implementing a structural causal model for MR images of multiple sclerosis using the probabilistic programming language Pyro (2021 MICCAI paper).
- tiramisu-brulee – PyTorch implementation of the Tiramisu deep neural network for segmentation.
- pymedio – Python package to flexibly read arbitrary medical image formats.
- selfsupervised3d – PyTorch-based self-supervised learning for 3D images.
- lesion-metrics – Python package to compute various medical image ML model performance metrics.

Honors and Awards

Ferdinand Hamburger Jr. Fellowship, Raytheon-SVA Scholarship, Frederic and Julia Weigl Scholarship, Jean Perkins Combat Veteran Scholarship, Jerry A. and Martha Lel Hawkins Endowed Scholarship.

Meritoriously promoted to Sergeant in the US Marine Corps (led 20+ person team).