William Thyer, PhD

thyerwilliam@gmail.com | williamthyer.github.io

Education

PhD Integrative Neuroscience, Psychology — 2023
University of Chicago, Institute for Mind and Biology — Chicago, IL

BS Psychology, Minor in Statistics — 2017
Florida State University — Tallahassee, FL

Work Experience

Senior Data Scientist (*formerly Data Scientist I, Data Scientist II*) — August 2023 - Current
Cohere Health — Remote

- Oversaw maintenance and enhancement of 19 production ML models automating prior authorization, preventing over $300K in monthly operational costs and saving thousands of patient hours per day.
- Led and mentored a co-op data scientist to build integrated monitoring solutions (Airflow, PySpark, Slack alerts, Tableau) for ML model performance, data health, and business impact—significantly improving decision transparency, monitoring capabilities, and response times to production issues.
- Coordinated several cross-functional teams to expand ML model coverage to a new clinical area, resulting in 1000s of auto-approvals and preventing >$100,000 in operational expenses per month.
- Owned and developed a *patient profile* data pipeline, establishing a centralized, rich, and accessible source of patient information that simplified analyses and enabled more informed decision-making across teams.

Graduate Researcher — July 2018 - July 2023
Awh & Vogel Lab, University of Chicago — Chicago, IL

- Collaborated with researchers on projects resulting in publications in high-impact peer-reviewed journals.
- Lead development of an open-source machine learning library for neural time series analysis with Python, Scikit-Learn, and Matplotlib, which is now used by more than 5 researchers.
- Built, tested, and deployed a data processing pipeline that improved useability and overall infrastructure which is now used by 2 research labs.

Data Scientist Intern (Full-Time) — June 2022 - September 2022
Intuitive Surgical — Atlanta, GA

- Researched machine learning-based feedback systems for surgeons during robotic surgeries, leading to improved understanding of potential future applications.
- Trained machine learning models on multivariate time series data with Scikit-Learn and TSFresh.
- Applied unsupervised clustering methods to improve model interpretability, providing insights into model performance and improving feedback systems.

Data Scientist Intern (Full-Time) — July 2021 - December 2021
Spark Neuro — Remote

- Trained XGBoost models on biomedical time series data for the diagnosis of neurological disorders, improving accuracy and specificity of the diagnosis processes.
- Created 3D interactive visualizations of neural data which increased interpretability of medical reports.
- Utilized AWS cloud services and parallel processing techniques to efficiently process large datasets, reducing computation time and improving data analysis efficiency.

Skills

Languages/Platforms: Python, SQL, MATLAB, SAS, pyspark, Airflow, AWS, git, Tableau

Libraries: Scikit-Learn, Fastai, PyTorch, TSFresh, TSAI, Matplotlib, Seaborn, Numpy, Pandas, Scipy, Flask

Skills: Machine learning, statistical analysis, ML production and monitoring, project management