Roza Gunes Bayrak, PhD

rgbayrak@gmail.com linkedin.com/in/rgbayrak rgbayrak.github.io

SUMMARY

Data Scientist and Applied ML Researcher with expertise in multimodal time-series modeling, representation learning, and large-scale neuroimaging analysis. 7+ years building deep learning pipelines and scalable data workflows for high-dimensional biomedical data. Author of 20+ peer-reviewed publications (including NeurIPS, MICCAI, Imaging Neuroscience) and developer of widely used open-source neuroimaging tools. Experienced in production ML pipelines, distributed training, and reproducible ML systems.

SKILLS

Programming: Python, MATLAB, D3.js, SQL (PostgreSQL)
Machine Learning / AI: Deep learning, transformers, self-supervised learning, representation learning, graph neural networks
Libraries / Frameworks: PyTorch, PyTorch Lightning, HuggingFace Transformers, Pandas, NumPy, Scikit-learn, SciPy, Nilearn, MNE, Matplotlib, Seaborn, Plotly
Infrastructure / Experimentation: Docker, SLURM, GitHub Actions, Hydra, Weights & Biases
Data: Large-scale time-series data, multimodal data integration, feature engineering, statistical modeling

PROJECTS

Neuroimaging and Brain Dynamics (NEURDY) Lab **Nashville, TN**
Senior Research Engineer (*Aug. 2024 – Present)*

- Leading R&D for benchmarking state-of-the-art (SOTA) learning models across heterogeneous datasets**.**
- Evaluated 20+ general purpose time series and domain-specific large pretrained and foundation models. Ran hundreds of large-scale experiments to compare model performance across datasets and modalities. Built reproducible workflows for experiment configuration, training, and tracking
- Designed a software workflow to automate assessment and provide model-based suggestions for correcting common artifacts. Leading a team of undergraduate and graduate students to develop a full-stack application for visualizing physiological signals, reviewing data quality, detecting and correcting artifacts using SOTA technology
- Built reusable data ingestion and preprocessing pipelines for fMRI, EEG, and physiological time-series datasets across 15+ independent datasets. Developed modular I/O layers to standardize heterogeneous dataset formats, enabling consistent data access across datasets with different structures. Processed and harmonized large multimodal time-series datasets containing millions of signal measurements
- Architected a prompt-driven workflow for using Claude Code to automatically generate dataset adapters for machine learning pipelines. Built LLM powered structured prompt protocols that extract model input schemas, dataset specifications, and feature mappings. Automated generation of adapter code, configuration files, and test suites to validate dataset compatibility. Reduced engineering overhead for integrating new datasets into benchmarking pipelines
- Built a research platform for physiological signal quality control and analysis, *physAI* is a centralized, user-facing platform to improve accessibility and reuse of internal research outputs including analytical tools, published papers and quality assessed datasets

Vanderbilt University **Nashville, TN**
Assistant Research Professor (*Aug. 2023 – Present)*

- Principal investigator for an NIH U24 grant submission (in revision) and co-investigator on an RO1 as co-investigator (resubmitted)
- Organized and leading a global collaborative effort with 7 labs 24 contributors to standardize physiological signal quality assessment
- Leading the BIDS-ification of two in-house datasets to organize the data collected in a standardized, reproducible format that makes sharing, analysis, and use with automated tools much easier.

- Supervised a student to design and evaluate foundation model backboned pipelines for cross-modal translation between EEG and fMRI signals. Published and presented at NEURIPS.
- Supervised a student on development of DL pipelines for classifying vigilance states from neuroimaging data using pretrained models. Published and presented as second senior author at SPIE.
- Contributed to design and evaluation of benchmarking experiments for graph neural network models applied to brain graph data. Created the preprocessing pipeline, curated dataset and guided evaluation metric selection. Published and presented at NEURIPS.

Graduate Researcher, Computer Science PhD (*Jan. 2018 – Jul. 2023)*
Advisor: Catie Chang

- Built deep learning pipelines to reconstruct physiological signals from temporal brain imaging data.
- Developed ML models for individualized brain mapping using graph neural networks and spherical convolutional networks.
- Co-developed a d3-based interactive visualization platform for functional brain parcellations.
- Conducted large-scale data analyses experiments linking brain signals, physiological measurements, and cognitive and behavioral measures.
- Designed and evaluated reproducible workflow protocols for manual tractography.
- Contributed to open-source scientific software tools used in neuroimaging research pipelines.

EDUCATION

Ph.D., Computer Science, Vanderbilt University — *2023*
Dissertation: *Computational Methods to Advance Individual Precision in Brain Mapping*
M.S., Electrical Engineering, Tufts University — *2016*
B.S., Electrical & Communication Engineering, Çankaya University — *2010*

COMMUNITY & LEADERSHIP

Open-Source Contributor, Physiopy, Nilearn, Datalad
Founder, Chair, BrainHack Vanderbilt
Co-Chair, CVPR Workshop on Medical Computer Vision (2025)
Reviewer, NeurIPS, ICML, ICLR, MICCAI, MiDL

SELECT PUBLICATIONS (out of 30+)

(Full list: https://scholar.google.com/citations?user=QHN1CZsAAAAJ&hl=en)

- **Bayrak RG**, Hansen CB, Salas JA, Ahmed N, Lyu I, Mather M, Huo Y, Chang C. DeepPhysioRecon: Tracing peripheral physiology in low frequency fMRI dynamics. Imaging Neuroscience. 2025 Sep 25;3:IMAG-a.
- Li Y, Lou A, Xu Z, Zhang S, Wang S, Englot DJ, Kolouri S, Moyer D, **Bayrak RG**, Chang C. Neurobolt: Resting-state eeg-to-fmri synthesis with multi-dimensional feature mapping. Advances in NEURIPS. 2024 Dec 16;37:23378-405.
- Said A, **Bayrak RG,** Derr T, Shabbir M, Moyer D, Chang C, Koutsoukos X. Neurograph: Benchmarks for graph machine learning in brain connectomics. Advances in NEURIPS. 2023 Dec 15;36:6509-31.
- **Bayrak RG,** Hoang N, Hansen CB, Chang C, Berger M. PRAGMA: Interactively constructing functional brain parcellations. In 2020 IEEE Visualization Conference (VIS) 2020 Oct 25 (pp. 46-50). IEEE.