

Yiming Shen, Ph.D.

EEG/BCI Research Scientist | Neural Signal Machine Learning | Scientific Software

yiming.shen001@umb.edu | +1 774-578-4554 | Boston, MA | Open to Relocate

yimings.com.cn | github.com/Yiming-S | linkedin.com/in/yiming-shen

Ph.D. in Computational Science specializing in reliable EEG/BCI decoding under session drift. Develops leakage-controlled cross-session evaluation, confidence-gated source-selection methods, drift diagnostics, and reproducible Python/R research software. Experience includes motor-imagery EEG, domain adaptation, Riemannian methods, and sample-paced pseudo-online inference.

Core BCI and Technical Expertise

- **EEG/BCI:** Motor-imagery decoding, cross-session evaluation, CSP/FBCSP, covariance and tangent-space features, Euclidean alignment, Riemannian geometry.
- **Machine Learning and Statistics:** Domain adaptation, transfer learning, nested cross-validation, bootstrap uncertainty, mixed-effects models, drift analysis.
- **Programming:** Python (MNE-Python, MOABB, pyRiemann, scikit-learn, PyTorch, NumPy/SciPy, Streamlit); R (package development, Rcpp); SQL and Bash.
- **Research Engineering:** Reproducible experiment pipelines, CLI/config-driven workflows, testing and CI, Docker, Linux, Slurm/HPC, Git/GitHub.

Research Experience

Doctoral Researcher - EEG Transfer and Decoding Reliability

Sep 2019 - Aug 2026

University of Massachusetts Boston

Boston, MA

- Developed confidence-gated source-selection strategies for within-subject, cross-session motor-imagery EEG transfer, comparing full pooling (MAP), distance-weighted pooling (DWP), minimum-distance selection (MMP), and bridge-domain selection (BDP) on Stieger2021 and BNCI2014-004.
- Designed matched evaluation across feature families, classifiers, domain-adaptation methods, and three reporting endpoints. On Stieger2021, MAP, DWP, and BDP were statistically tied as the top pipelines after data-driven configuration selection.
- Showed that BDP remained in the MAP/DWP selected-performance range while its shared CI gate retained 74.29% of source sessions and the pipeline used 37% of MAP's complete benchmark-search runtime on Stieger2021; framed this as an efficiency result rather than an accuracy improvement.
- Built leakage-controlled leave-one-session-out workflows with target labels reserved for final testing, auditable source-session roles, configuration scoring, final evaluation outputs, and reproducible Linux/Slurm sweeps.
- Developed complementary analyses of signal- and feature-space drift using covariance and Riemannian metrics, mixed-effects models, and fixed-reference comparisons of no adaptation, domain adaptation, and retraining.

Selected Publications and Manuscripts

- **Shen, Y.**, and Degras, D. "A Confidence-Gated Source Selection Strategy for Cross-Session Transfer in Brain-Computer Interfaces." *Frontiers in Human Neuroscience*, Brain-Computer Interfaces. **Accepted for publication**, Aug 2026. Article 1895016. [doi:10.3389/fnhum.2026.1895016](https://doi.org/10.3389/fnhum.2026.1895016)
- **Shen, Y.** "Small calibration cohorts do not reliably select motor-imagery EEG systems for new participants." *Under Review*.

Selected BCI Research Software

CrossDA

[GitHub](#)

Python, MOABB, MNE-Python, pyRiemann, scikit-learn

- Built a CLI and Python API for leave-one-session-out EEG experiments across MAP, DWP, MMP, and BDP source-utilization pipelines, with configurable feature, classifier, domain-adaptation, and distance-method banks.
- Produced per-subject summaries, detailed run records, and role tables documenting source, bridge, far, and target-session assignments for downstream audit and analysis.

ShiftDx

[GitHub](#)

Python, Streamlit, Plotly, statsmodels, DA4BCI-Python

- Developed an offline fixed-reference MI-EEG drift-analysis dashboard: models are fit on session 0, applied unchanged to later sessions, and compared across no-adaptation, domain-adaptation, and retraining conditions.
- Integrated five shift metrics, subject trajectories, adaptation gain/harm views, retraining-gap analysis, statistical reports, and synthetic Page-Hinkley demonstrations.

NeuroStream

[GitHub](#)

Python, MOABB, MNE-Python, pyRiemann, tkinter

- Built a sample-paced pseudo-online motor-imagery replay application that trains on MOABB data and releases held-out trials through a replay source and ring buffer.
- Implemented CSP, FBCSP, tangent-space LDA/SVM, Euclidean alignment, background inference, and progressive confidence and performance visualization.

Supporting Method Libraries

[GitHub](#)

DA4BCI, DA4BCI-Python, BCIFeatR, eegwhiten

- Developed train/test-separated EEG feature extraction and whitening APIs together with reusable domain-adaptation methods, Riemannian transport, optimal transport, and source-target shift metrics in R and Python.

Prior Industry Experience

Data Engineer

Jul 2014 - Aug 2016

China Mobile IoT Company Limited

Chongqing, China

- Managed GB-scale time-series sensor logs in MySQL and improved aggregation workflows through indexing and partitioning.
- Built automated network-health dashboards and visual reports for operational monitoring.

Education

University of Massachusetts Boston

Aug 2026

Ph.D. in Computational Science - Data Analytics Track

Boston, MA

Valparaiso University

May 2018

M.S. in Analytics and Modelling

Valparaiso, IN

Chongqing University

Jun 2014

B.Eng. in Electrical and Electronics Engineering

Chongqing, China

Selected Presentations and Recognition

- **Poster:** “Cross-Session BCI Transfer: A Matched Comparison of Global and Selective Pooling.” ENAR, Indianapolis, IN, Mar 2026.
- **Invited Talk:** “Benchmarking Classification Pipelines Within and Across Sessions on the PhysioNet EEG Dataset.” MIND Seminar, Inria, France, Jun 2025.
- **Silver Medal:** Kaggle HMS - Harmful Brain Activity Classification, rank 98 of 2,767 teams (top 4%), 2024.