DOI: 10.3217/978-3-99161-050-2-235

Neural Symbolic Regression for Interpretable & Efficient Brain-Computer Interfaces

A. Minhas[1,2]*, E. Kinney-Lang[1,3], A. Kirton[1,2,3]

[1]*Cumming School of Medicine,* [2]*Neurosciences;* [3]*Biomedical Engineering. University of Calgary, Canada.*

*2500 University Drive NW, Calgary, AB, Canada, T2N 1N4. E-mail: araz.minhas @ucalgary.ca

***Introduction**:* Brain-computer interfaces (BCIs) offer transformative potential for underserved populations like children with neurological disabilities, by enabling environmental interaction via direct neural signals[1]. Scalable deep learning (DL) algorithms needed to learn from these signals effectively are central to improving BCI performance. However, the black-box nature and computational demands of DL pose many challenges to clinical deployment, mechanistic research, patient accessibility and do not provide the adaptability needed for pediatric BCI applications, wherein each child's unique neurodevelopmental profile necessitates continually-learning algorithms that can support personalized solutions.

***Material, Methods and Results**:* We present a novel DL architecture that addresses these challenges by learning concise, interpretable mathematical expressions, while serving as a drop-in replacement for the most common DL module in DL/BCI pipelines– multilayer perceptrons (MLPs). Unlike traditional MLPs that learn opaque representations, our symbolic regression (SR) model aims to discover simple equations that describe the underlying relationships in data, and builds on past neuro-symbolic methods[2] by introducing: (1) self-compressive training via adaptive pruning, quantization and hyperparameter selection; (2) single-phase training compatible with standard DL workflows; (3) support for discontinuous operators via neural arithmetic logic units[3]; and (4) automatic feature selection for high-dimensional datasets. Preliminary evaluation on BCI Competition IV Dataset 2a [4] with standard training features demonstrate that the symbolic expressions obtained achieve higher classification accuracy than MLPs (52% vs. 47%), reduce model size by ~91%, inference time by ~83% and yield readable equations to compute classification class probabilities for motor imagery (Fig. 1c). Such interpretability may enable neurophysiological hypothesis generation, and the reduced computational overhead can facilitate deployment on resource-constrained devices common in clinical settings.

***Conclusion**:* These improvements in efficiency and interpretability provide a foundation for more accessible and personalized BCI systems– aspiring to benefit pediatric applications where understanding individual variation and enabling widespread deployment are crucial for clinical success.

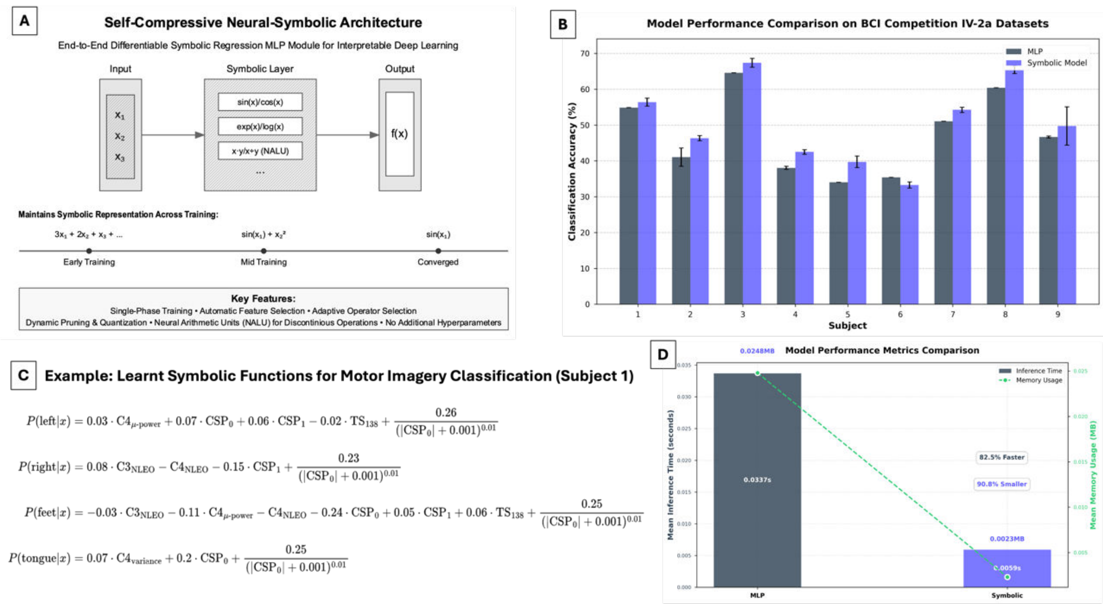

Figure 1: ***A)** Proposed self-compressive architecture that maintains concise symbolic representations across training. **B)** Classification performance across subjects compared to standard multilayer perceptron (MLP). **C)** Example of learnt symbolic classifiers for Subject 1 (C3/C4 ≈ motor cortex electrodes; CSP = Common Spatial Pattern features; NLEO = Nonlinear Energy Operator; μ-power ≈ motor-related mu-band power). **D)** Model compression and inference time comparisons against standard MLP (n = 5).*

References:

[1] Kinney-Lang, E., Floreani, E. D.., ... & Kirton, A. (2023). Pediatric Brain–Computer Interfaces: An Unmet Need. Handbook of Human-Machine Systems, 35-48.
[2] Martius, G., & Lampert, C. H. (2016). Extrapolation and learning equations. arXiv preprint arXiv:1610.02995.
[3] Trask, Andrew, et al. "Neural arithmetic logic units." Advances in neural information processing systems 31 (2018).
[4] Brunner, C., Leeb, R., Müller-Putz, G., Schlögl, A., & Pfurtscheller, G. (2008). BCI Competition 2008–Graz data set A. Institute for knowledge discovery (laboratory of brain-computer interfaces), Graz University of Technology, 16, 1-6.

Published by Verlag der Technischen Universität Graz