

# Research

Here lies a brief overview of all my current research interests!

## DeepBCI

The goal of this research project is multifaceted and has plethora of researcher directions. The main focus of the project is to integrate both machine learning (in particular deep learning and reinforcement learning) with brain-computer interfaces (BCIs). Intrinsic feedback is the center piece of this project as this is feedback communicated to agents via brain signals.

### **Learning from intrinsic feedback (DeepBCI advisor)**

Intrinsic feedback communicated via error-related potentials (ErrPs) allows for the approximation of a humans subjective ground truth. ErrPs are a brain signal that is automatically elicited when a human commits or observes an erroneous event. Using this error signal is used to teach an agent how to play a task using reinforcement learning. The end goal of this project is to have a two-way interaction where the human provides intrinsic feedback to the agent and the agent advises the human in order to prevent errors.

### **Classification of intrinsic feedback signals**

One of the defining properties of brain signals is that they follow a non-stationary distribution which means a brain signal's distribution can shift causing the performance of the classification model to degrade. The goal of this project is to further investigate and discover new methods for successfully decoding brain signals, with a focus on deep learning.

## Interactive Reinforcement Learning

Interactive reinforcement learning (IRL) aims to progress RL towards becoming more robust, scalable, and aligned by expanding the RL framework to account for human guidance . IRL can then be seen as incorporating a human-in-the-loop (HITL) to allow for human-to-agent knowledge transfer. This knowledge-transfer inherently helps to alleviate robustness/scalability and alignment challenges. Thus, this work looks at further researching how to integrate human guidance with reinforcement learning.

