Classifier Performance in Primary Somatosensory Cortex Towards Implementation of a Reinforcement Learning Based Brain Machine Interface

David McNiel[2], Mohammad Bataineh[1*], John Choi[2], John Hessburg[2], Joseph Francis[1, 2]
[1] Department of Biomedical Engineering, Cullen College of Engineering, University of Houston, Houston, TX 77204-6022
[2]Department of Physiology and Pharmacology, SUNY Downstate Medical Center, Brooklyn, NY-11203
*Corresponding author: *dave.b.mcniel@gmail.com*

Abstract- Increasingly accurate control of prosthetic limbs has been made possible by a series of advancements in brain machine interface (BMI) control theory. One promising control technique for future BMI applications is reinforcement learning (RL). RL based BMIs require a reinforcing signal to inform the controller whether or not a given movement was intended by the user. This signal has been shown to exist in cortical structures simultaneously used for BMI control. This work evaluates the ability of several common classifiers to detect impending reward delivery within primary somatosensory (S1) cortex during a grip force match to sample task performed by a nonhuman primate. The accuracy of these classifiers was further evaluated over a range of conditions to identify parameters that provide maximum classification accuracy. S1 cortex was found to provide highly accurate classification of the reinforcement signal across many classifiers and a wide variety of data input parameters. The classification accuracy in S1 cortex between rewarding and non-rewarding trials was apparent when the animal was expecting an impending delivery or an impending withholding of reward following trial completion. The high accuracy of classification in S1 cortex can be used to adapt an RL based BMI towards a user's intent. Real-time implementation of these classifiers in an RL based BMI could be used to adapt control of a prosthesis dynamically to match the intent of its user.

I. Introduction

Brain machine interfaces (BMI) have been used to control a variety of devices including computer cursors [1], robotic arms [2], and assisted spelling devices [3]. While these BMIs are quite different, they all rely on the same basic concept: neural activity in some form is decoded and used to control the device. One means of generating an additional stream of useful data towards increasing the accuracy of a BMI is decoding a reward-associated signal in parallel to decoding the variables of direct interest to the BMI, for example position and velocity [4]. This reward-associated signal could summarily be utilized as a reinforcing signal to increase the accuracy of the BMI's output without direct knowledge of the user's intent. The implementation of this type of system is made possible via reinforcement learning algorithms that are capable of aligning an arbitrary system, such as a robotic arm, with an arbitrary goal, such as reaching and grasping an object [5]. While such a reinforcing signal has been observed in several areas of the brain [4,6,7], this work will focus on decoding this signal from primary somatosensory cortex (S1). This is the first time to our knowledge that such a signal has been observed in S1.

II. Methods

One Bonnet macaque (Macaca Radiata) was trained to grip with its right hand and hold to match a visually displayed target force. S1 cortex was implanted with an array of electrodes (96 channel Utah array, Blackrock Microsystems) [8]. Peristimulus time histograms (PSTH) of neural spiking data were generated for two time intervals by aligning spike times either with the timing of a visual cue relaying the impending result of a given trial (reward will be delivered or reward will be withheld) or with the timing of the trial's outcome (reward delivered or reward withheld) and extended 500 milliseconds after the stimulus of interest using a 100 millisecond bin width. PSTH data was then reduced in dimensionality first by transforming the data via principal component analysis (PCA) and subsequently eliminating all but the first two components (PC1 and PC2) as features for classification [9]. A variety of classifiers (Naïve Bayes, Nearest Neighbor, Linear Support Vector Classification (SVC), AdaBoost, Quadratic Discriminant Analysis (QDA), and Linear Discriminant Analysis (LDA)) were then trained on data obtained from 60% of the trials (82 trials) and then tested on the remaining 40% of trials (55 trials) to determine the most accurate method of decoding a trial's outcome as rewarding or non-rewarding. Trials used in training and testing of each classifier were randomly assigned to either the training or testing group using the "train_test_split" function found in the sklearn python package utilized in this work [9].

III. RESULTS

A. *Robust Classification Accuracy is Possible with a Variety of Classifiers*

All six classifiers provided robust classification within both intervals of interest, with the nearest neighbors classifier providing the highest classification accuracy in the post-cue period (92% accurate classification of test dataset). The output of the nearest neighbors classifier is shown in Fig. 1.

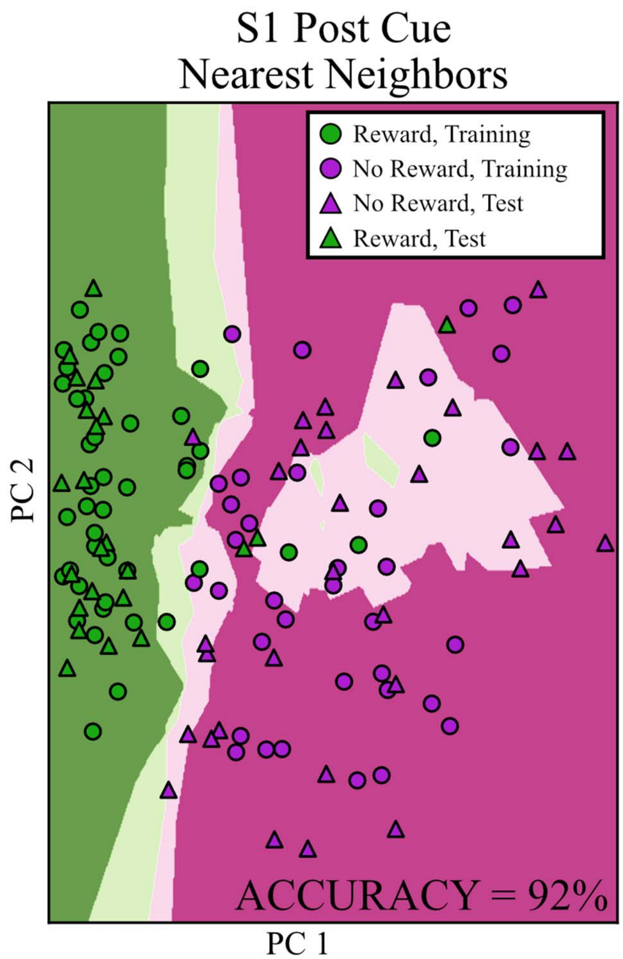

Figure 1: Nearest-neighbor classification of trial outcome decoded from S1.

B. *Classification Accuracy is Greatest Following Cue Presentation and Before Reward Delivery*

Across all classifiers utilized in this work, classification accuracy was higher during the post-cue period as compared to the post-reward/ post-no-reward period. These results are summarized in Table 1.

TABLE I
PREDICTION ACCURACY OF SIX COMMON CLASSIFIERS ON DISCRIMINATING REWARDING TRIALS FROM NON-REWARDING TRIALS

Classifier Name	Post-Cue Accuracy	Post-Reward Accuracy
Naïve Bayes	89%	74%
Nearest Neighbors	92%	69%
Linear SVM	90%	72%
AdaBoost	89%	62%
QDA	89%	72%
LDA	86%	72%

IV. DISCUSSION

Increasing the accuracy of a BMI requires either increased information or a superior algorithmic approach. The utilization of reinforcement learning algorithms can form a basis for utilizing the underlying reward associated data stream described here towards increased BMI accuracy. Interestingly, the decoding accuracy of trial outcome was higher following the visual cue than the trial's outcome. This result is not entirely surprising given that previous work has demonstrated that the presence of conditioned stimuli in reward prediction tasks shift reward modulated activity in the brain to the earliest stimulus predictive of impending reward delivery. Future work will incorporate this signal into a pipeline that can be used for dynamically increasing the accuracy of a BMI without increasing the size or complexity of the implant itself.

REFERENCES

[1] Serruya, Mijail D., et al. "Brain-machine interface: Instant neural control of a movement signal." Nature 416.6877 (2002): 141-142.
[2] Velliste, Meel, et al. "Cortical control of a prosthetic arm for self-feeding." Nature 453.7198 (2008): 1098-1101.
[3] Krusienski, Dean J., et al. "A comparison of classification techniques for the P300 Speller." Journal of Neural Engineering 3.4 (2006): 299.
[4] Marsh, Brandi T., et al. "Toward an Autonomous Brain Machine Interface: Integrating Sensorimotor Reward Modulation and Reinforcement Learning." The Journal of Neuroscience 35.19 (2015): 7374-7387.
[5] Sutton, Richard S., and Andrew G. Barto. Reinforcement learning: An introduction. Vol. 1. No. 1. Cambridge: MIT press, 1998.
[6] Schultz, Wolfram. "Multiple reward signals in the brain." Nature Reviews Neuroscience 1.3 (2000): 199-207.
[7] Elliott, Rebecca, et al. "Differential response patterns in the striatum and orbitofrontal cortex to financial reward in humans: a parametric functional magnetic resonance imaging study." The Journal of Neuroscience 23.1 (2003): 303-307.
[8] Chhatbar, Pratik Y., et al. "A bio-friendly and economical technique for chronic implantation of multiple microelectrode arrays." Journal of Neuroscience Methods 188.2 (2010): 187-194.
[9] Pedregosa, Fabian, et al. "Scikit-learn: Machine learning in Python." The Journal of Machine Learning Research 12 (2011): 2825-2830.