Machine Learning Approach to Identifying Neural Features That Predict Rodent Behavior

Washington State University - Department of Psychology

Nunnerson Computing

Aidan Nunn, Charlie Nickerson

9/21/22

TABLE OF CONTENTS

I. Introduction

This document showcases Nunnerson Computing's knowledge thus far regarding our partnership with Dr. Angela Henricks' team and the WSU Department of Psychology. We are assisting Dr. Henricks by developing a machine learning model using lasso (least absolute shrinkage and selection operator) which will predict when to collect data from their Sprague-Dawley rodent test subjects. Their previous method of data collection has been time consuming, and by creating this machine learning model we will streamline the data collection process and save time for the researchers. Our end goal is to present Dr. Henricks' team with a machine learning model which can predict what frames of data they need to study, thereby removing the need for them to manually locate desirable frames of data.

II. Background and Related Work

Among individuals who practice substance abuse, there seem to be neurological processes that underlie behaviors linked to the abuse of those substances. Dr. Henricks and her team hypothesize that LFPs could be used to determine whether substance abuse was occurring or if the male and female subjects were partaking in normal activity.

Currently, the focus of Dr. Henricks' team is on the difference between males and females concerning which parts of the brain are responsible for the decision making process to self administer alcohol. Dr. Henricks' team is not the first to tackle psychology related questions by using rat test subjects. Doucette et al. used the lasso (least absolute shrinkage and selection operator) algorithm with LFP data to predict outcomes of brain stimulation on a rat's desire to binge eat [3]. Dwiel et al. used lasso to predict the amount of food eaten, the increase in consumption following food deprivation, and the type of food eaten by the rats [2]. Additionally, Dwiel et al. managed to predict if a rat was going to consume 42.5 seconds before it would happen [2]. These and previous studies have used male models, however, which work well on other male subjects but do not translate over to female subjects. Due to this, Dr. Henricks' team suspects that a different part of the brain is responsible for the decision to self administer in females and are making female brains a priority in their research.

The main method for gathering behavioral data in experiment subjects, Sprague-Dawley rats, are ventral striatal local field potentials (LFPs). A previous experiment conducted by Dr. Hendricks and her team found that LFP measurements taken from specific regions in the brain can indeed be used in a machine learning model to differentiate LFPs recorded during the consumption of ethanol (EtOH) and sweet-fat foods (SF) from all other behaviors. This is great news, as it means Hendriks' team is already familiar with using machine learning in their experiments. Our machine learning model will still be making new contributions to the project domain, as it's assisting streamline their research.

We need to learn new technical knowledge and skills in order to accomplish this project. Most importantly is understanding the LFP data collected from the rats and understanding the MatLab programming language for use in constructing the machine learning program that will construct our model. We will also need a thorough understanding of the lasso machine learning algorithm. In order to aid us in learning more about the background of this project, Dr. Henricks provided our team with some papers covering previous research that her team referenced when working on their research. These were helpful for understanding how the LFP data was collected from the rats but it is still unclear how the code was implemented to make a predictive model. Later on Hendricks will provide us with the codebase from previous capstone projects. These will be helpful for teaching us how to use MatLab and lasso.

III. Project Overview

Dr. Henricks' and her team are studying which parts of the brain are involved in the decision making process to self administer addictive substances. Their experimental process involves placing a rat in a box for 30 minutes and recording the local field potentials (LFPs) using electrodes in the rat's brain. Inside the box is a lever that dispenses alcohol, and a camera that records the rat for the entire session. Currently, the team watches the video footage and waits for the moment when the rat is making the decision to self-administer alcohol or not; this is when they want to study the LFPs. Our goal is to improve upon a machine learning program which will generate a model that predicts when the rat is making a decision. With this model, the team can isolate the frames of data collected from when the rat is making the decision whether or not to administer alcohol. This will allow them to save time reviewing experiment sessions and avoid accidentally reading data collected from when the rat is not in the decision making process.

Our team has been recommended to use the lasso (least absolute shrinkage and selection operator) algorithm to generate our model. We will use LFP data collected from experiment sessions using plex to train the model. Lasso is a regression analyst method that calculates regularization and variable selection in order to increase prediction accuracy and help with the interpretability of the regression model. Lasso was originally created for linear regression models, but has other generalizations as well including generalized linear models, generalized estimating equations, proportional hazards models, and M-estimators. Lasso works by forcing the sum of the absolute values of the regression coefficients to be less than a certain value. By doing this some coefficients are forced to be zero and will be excluded from impacting the prediction. The lasso method is similar to ridge regression because it shrinks the size of coefficients but it does not set any coefficients to zero making the method unable to perform variable selection. Doucette et al. used the MatLab package *Glmnet* to assist in constructing their model [3], so this is something we will look into as well.

Based on the work performed by last semester's computer science capstone team, Dr. Henricks has recommended that we spend our first semester on the project learning about their research and studying the codebase left behind by the previous students. We will also have the opportunity to play with small LFP datasets and learn lasso by constructing prototypes of our project. This time period will also involve planning out our second semester on the project. Our team's second semester will involve coding our algorithm and deriving an accurate model that will be used by Dr. Henricks' team. At this point we will also begin presenting our findings to the Hendricks lab and potentially to conferences such as SURCA, as outlined in the abstract provided to our team.

We will be utilizing a node in the Kamiak computing cluster on the WSU campus to run our algorithm. We will be working with a great deal of data in order for our model to be accurate for many scenarios, the additional computer power of Kamiak will ensure that we can create test models without wasting too much time. The previous round of capstone students assisting Dr. Henricks created a protocol for using Kamiak. They used Python while working on this problem; however, Dr. Henricks has recommended we try using MatLab instead, as it runs faster than Python and she has a better understanding of that language. MatLab has built in tools for data analysis, like the aforementioned *Glmnet*, and tools for graphics that will help us present our data.

IV. Client and Stakeholder Identification and Preferences

Our client is the WSU Department of Psychology with Dr. Angela Henricks as our primary contact for the project. The model we build will be used by Dr. Henricks' team, but could potentially be used by other teams if they take over the research. Other stakeholders include the rest of Dr. Henrick's team. Our team mentor is Ananth Jillepalli, an assistant professor at WSU.

The model we build could be used by future research teams to better understand the link between neural processes in the brain to real life behavior. Furthermore, if researchers come to understand these mechanisms they could attempt to manipulate them to better help people who suffer from addictive behavior (i.e. drinking alcohol, eating, smoking, etc…), so people helped by this research are potential stakeholders as well.

V. Glossary

Kamiak: A high performance computer on the Washington State University campus. Built out of a cluster of local computers connected via a high speed network.

Lasso: A machine learning algorithm used to construct a machine learning model.

Local Field Potential (LFP): The result of electrical activity in cells (mainly neurons).

Machine Learning: The use of computer programs that can learn based on input data and make predictions based on future data.

Machine Learning Model: A data file that has been trained to recognize specific data patterns.

Sprague-Dawley Rats: A species of rats ubiquitous to biomedical research.

SURCA (Showcase for Undergraduate Research and Creative Acts): A WSU event showcasing undergraduate research.

SF: Sweet Fat

EtOH: Ethanol

VI. References

[1] György Buzsáki, Costas A. Anastassiou, and Christof Koch, “The origin of extracellular fields and currents — EEG, ECoG, LFP and spikes,” *Nat Rev Neurosci*, June, 2016. [Accessed Sep. 20, 2022]

[2] Lucas L. Dwiel, Jibran Y. Khokhar, Michael A. Connerney, Alan I. Green, Wilder T. Doucette, “Finding the balance between model complexity and performance: Using ventral striatal oscillations to classify feeding behavior in rats,” *PLOS Computational Biology*, April, 2019. [Accessed Sep. 20, 2022]

[3] Wilder T. Doucette, Lucas Dwiel, Jared E. Boyce, Amanda A. Simon, Jibran Y. Khokhar, Alan I. Green, “Machine learning based classifications of deep brain stimulation outcomes in a rat model of binge eating using ventral striatal oscillations,” *Frontiers in Psychiatry*, August, 2018. [Accessed Sep. 20, 2022]