STA 141A: Project Report

Zeeva Chaver

2024-03-18

Abstract

In this report, data was analyzed from a study conducted by Steinmetz et al. (2019), which measured the neural activity of mice upon performing a visual discrimination task. With aims to predict the outcome of this visual discrimination task based on neural data, given in the form of neuron spike trains, neuron firing rates were explored in this report in two different ways. To reduce the numerous neural spike trains into comparable forms across experiment trials, spike trains were either averaged for all neurons in a trial or each neuron had their spike trains averaged over all trials they were present in. To capture the neural activity across potential types of neurons, clustering on neurons was performed for both 3 and 4 clusters (for comparison purposes). To capture the differences in neural activity across various decision types (relating to the decision a mouse made in the visual discrimination task), models were trained and tested on data subsetted by each decision. Logistic Regression and XGBoost models were used to predict the outcome of the visual discrimination task, success or failure, and were applied to each decision type subset of the overall data, for each clustering type. This report shows the data exploration, integration, and predictive modeling processes performed with the overall goal of predicting either success or failure in the discrimination task.

Section 1: Introduction

In this study, experiment data was collected during 39 sessions, for 10 different mice. This report analyzes data from 18 sessions (Sessions 1 to 18), for 4 different mice: Cori, Forssmann, Hench, and Lederberg. Each session consists of several hundred trials, in which visual stimuli were randomly presented to the mouse on two screens positioned on both sides of it. The stimuli varied in terms of contrast levels, which took values of 0, 0.25, 0.5, or 1, with the value 0 for the absence of a stimulus. The mice were required to make decisions based on the stimuli, using a wheel controlled by their forepaws. Success per trial (called "feedback" in this report) was allocated in four different ways. The first way, called a trial of type Decision 1 in this report, has the left contrast greater than the right contrast, with success (1) given if the wheel is turned to the right and failure (-1) otherwise. The second way, called a trial of type Decision 2 in this report, has the left contrast smaller than the right contrast, where success (1) is allocated if the wheel is turned to the left and failure (-1) otherwise. The third way, called a trial of type Decision 3 in this report, has both contrasts being of value 0, where success (1) is allocated if the wheel is held still and failure (-1) otherwise. The fourth way, called a trial of type Decision 4 in this report, has both left and right contrasts being equal but non-zero, where the left or right stimuli will be randomly chosen (50%) as the correct choice.

From 18 RDS files, we see information about each respective file's session through eight variables. The structure of data for each trial revolves around eight features. The date the session took place and the mouse's name are features that have one value per session. The left contrast, right contrast, and feedback type features are vectors with the number of elements being the total number of trials in a session. In Session 1, there are 114 trials, so these vectors contain 114 elements. However, the brain area feature is a vector with the number of elements as the total number of neurons with spikes tracked per session. In Session 1, there are 734 individual neurons (so this will be the element number of the brain area vector) with recorded activity throughout the 114 sessions. Each element of the brain area vector is the brain area that the specified neuron belongs to. The neuron spikes feature is in the form of a list of matrices, each element being a matrix for each trial with dimensions of the total number of neurons in a session and 40 time bins, representing 0.01 seconds over the 0.4 seconds that each trial is held over. For Session 1, the neuron spikes feature will be a list of 114 elements, each element being a matrix with the dimensions of 734 neurons by 40 time bins. The time feature is a list of 114 vectors, with each element being a 40-element vector. The elements in each 40 vector represent the 40 time bins, a specific 0.01 second time interval.

Section 2: Exploratory Analysis

This table showcases some of the features available for each session, including the number of unique neurons, the number of trials, the number of trials for each decision type, and the success rate over all trials in that session.

Session	Number of Neurons	Number of Trials	Number of Decision 1 Trials	Number of Decision 2 Trials	Number of Decision 3 Trials	Number of Decision 4 Trials	Success Rate
1	734	114	30	54	24	6	0.6053
2	1070	251	109	139	90	27	0.6335
3	619	228	155	223	166	49	0.6623
4	1769	249	239	303	234	66	0.6667
5	1077	254	315	386	305	90	0.6614
6	1169	290	415	476	384	111	0.7414
7	584	252	486	555	473	124	0.6706
8	1157	250	595	639	514	140	0.6440
9	788	372	766	744	590	160	0.6855
10	1172	447	945	866	703	193	0.6197
11	857	342	1068	978	785	218	0.7953
12	698	340	1171	1089	893	236	0.7382
13	983	300	1283	1180	975	251	0.7967

Session	Number of Neurons	Number of Trials	Number of Decision 1 Trials	Number of Decision 2 Trials	Number of Decision 3 Trials	Number of Decision 4 Trials	Success Rate
14	756	268	1371	1263	1060	263	0.6940
15	743	404	1511	1393	1177	280	0.7649
16	474	280	1618	1475	1255	293	0.7179
17	565	224	1690	1562	1308	305	0.8304
18	1090	216	1762	1633	1371	315	0.8056

The dimensions of the data in this study vary across sessions, with each session containing differing numbers of unique neurons, trials, and trials falling under each decision type. With the interest of summarizing neuron activity across trials and eventually sessions, neuron spike trains and their patterns in trials with different decision types can be explored.

To see what an average neuron firing rate plot looks like, the average spikes across all neurons in one trial, in one session, can be plotted as an example. Below is a plot visualizing the neuron spikes in Trial 1, Session 1, averaged over the 40 time bins.

Trial 1, Session 1
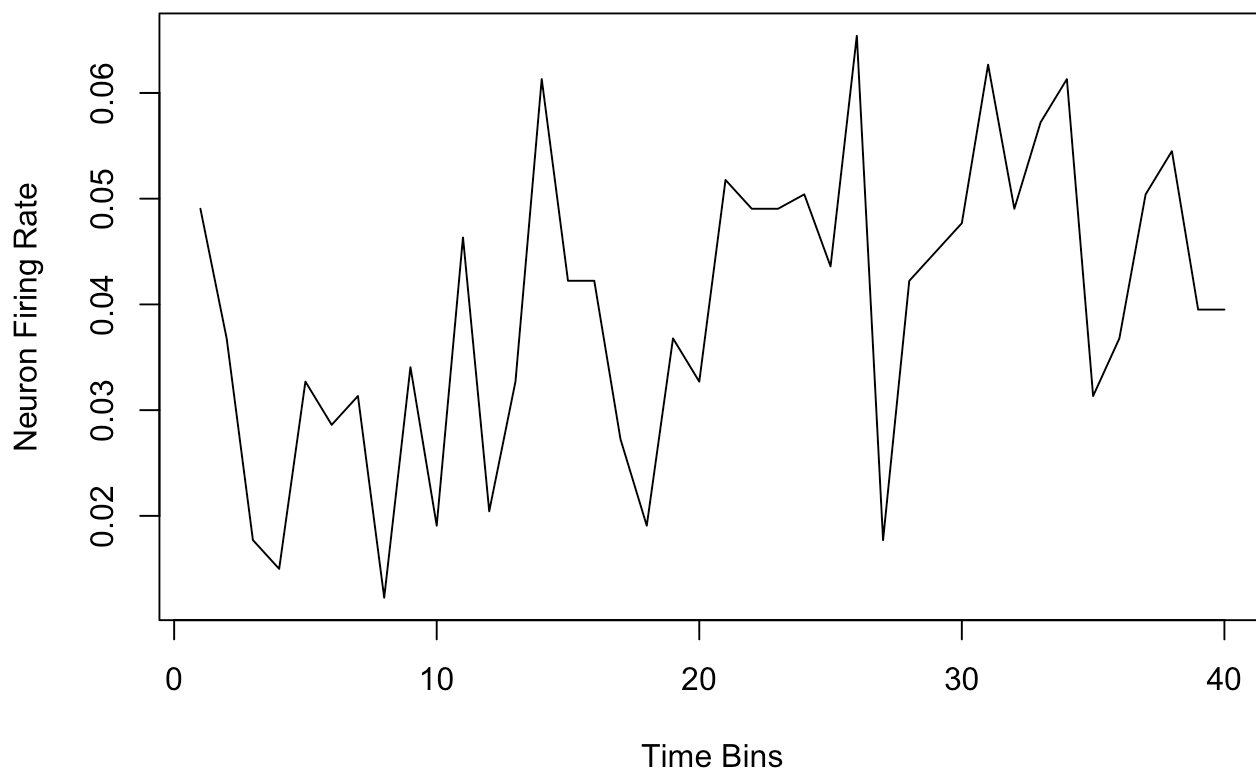
In this plot, the x-axis represents time bins, each bin contributing to the 0.4 second time period in which experiment data is collected per trial. The y-axis represents the neuron firing rate, which is the neuron firing activity averaged over all neurons in a trial (in this plot, trial 1 in session 1). To explore neuron firing rate at a higher level, neuron firing activity can be averaged over all neurons in a session, since neurons are the same across all trials in a given session.

Neuron firing rates can be plotted for each session, separated and averaged over the two feedback outcomes of success (1) and failure (-1).

Session 13

Session 14

Session 15

Session 16

Session 17
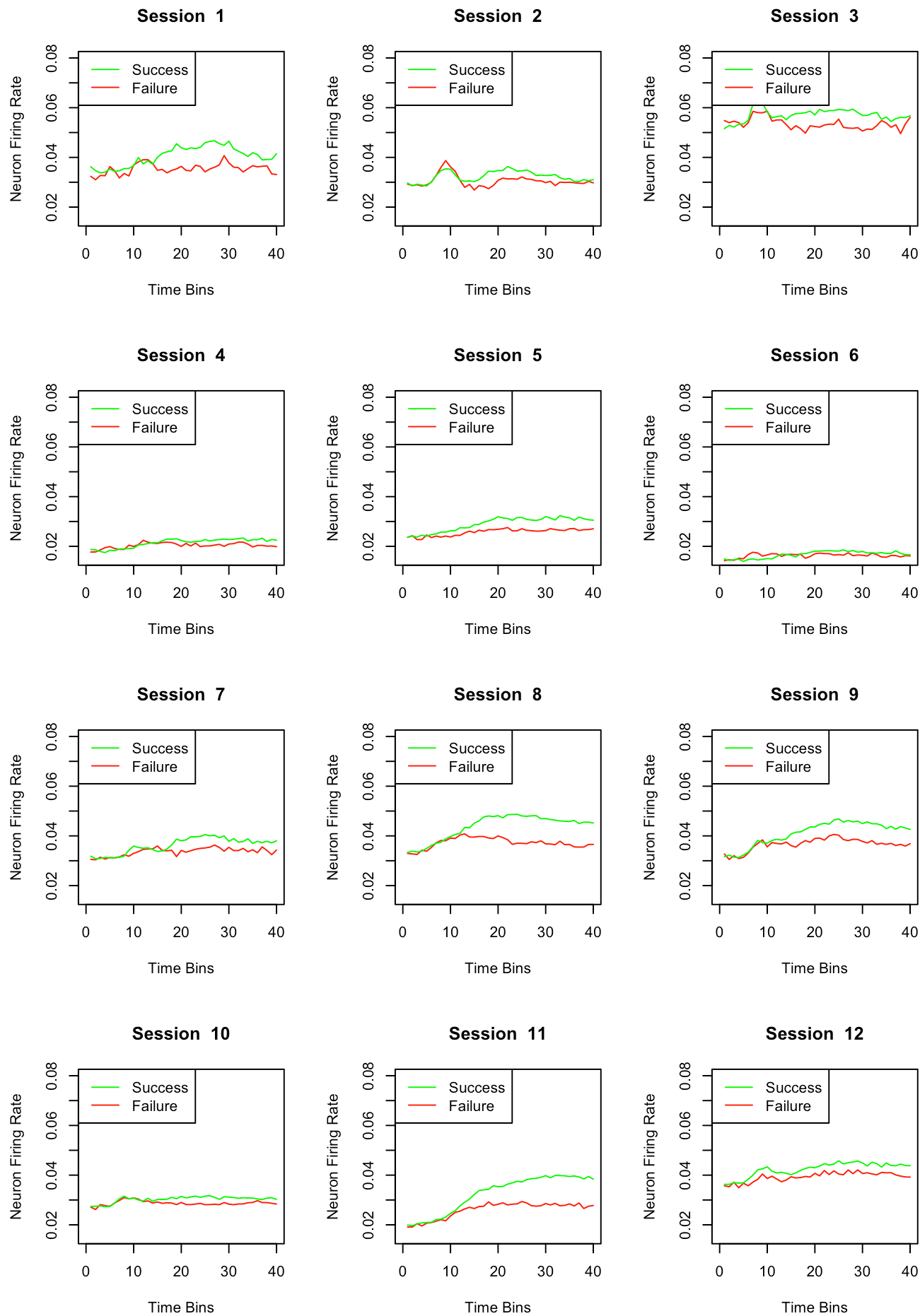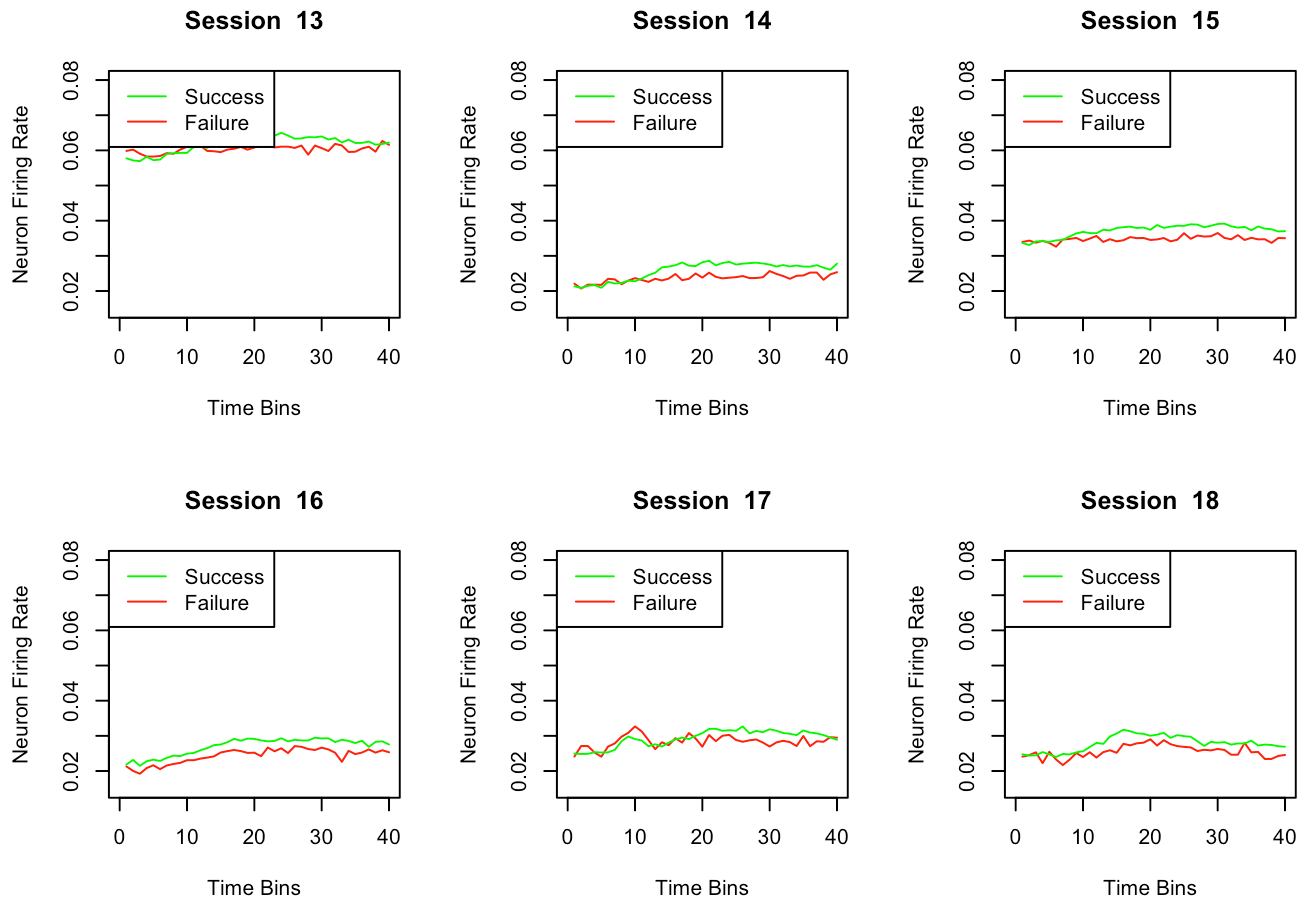
Session 18

From these plots, it is seen that the neuron firing rate averaged over trials resulting in success is higher than of those resulting in failure. Some of the plots with greater average neuron firing rates can be seen for Session 3, 13, and 15. However, the success rates for each of these sessions fall under different ranges, being between 0.6-0.7 for Session 3, between 0.7-0.8 for Session 15, and almost 0.8 for Session 13. To look more into success rate and neuron firing rate levels, vectors that are the difference of the average neuron firing rates for success and for failure can be graphed, with colors relating to different success rate ranges.
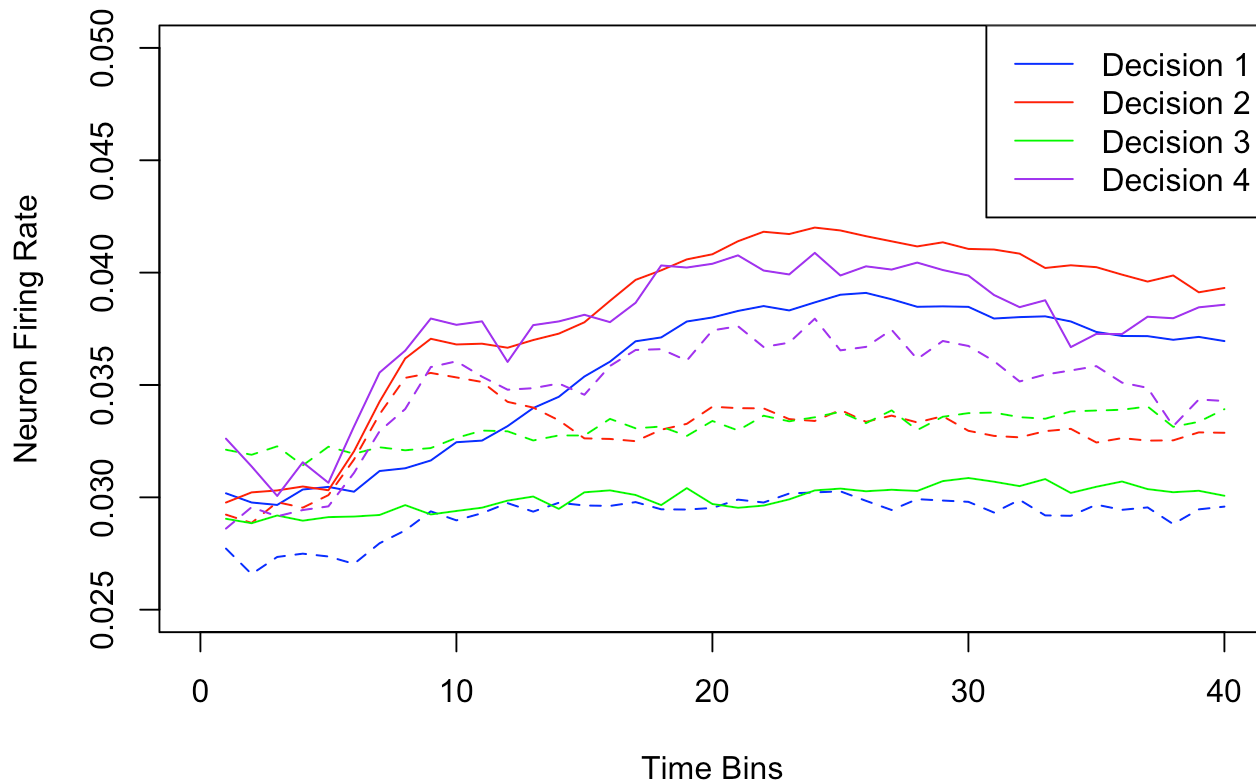
Session Neuron Firing Rate for Success-Failure Differences
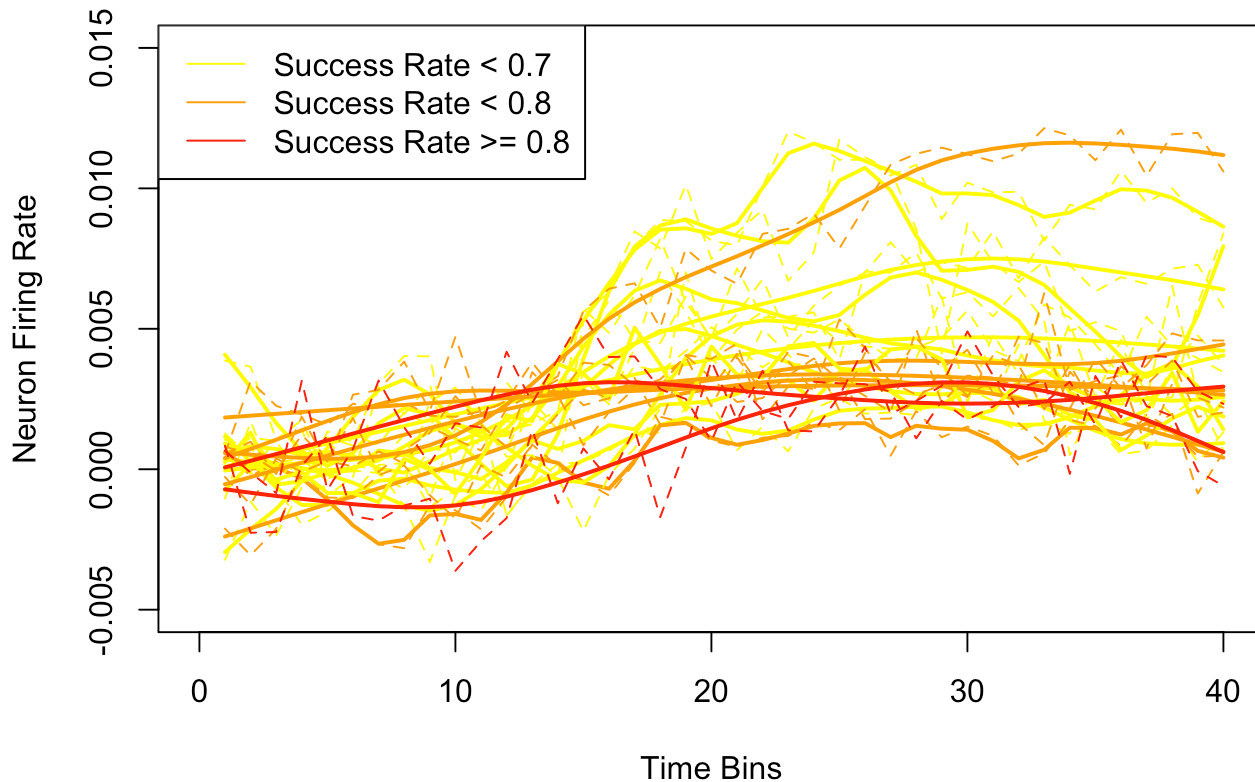
From this plot, it seems that sessions with greater success rates had generally lower differences in neural spike activity between success and failure. Sessions with the greatest success rate appear in red and show near 0, and generally more constant, differences between success and failure firing rates. Sessions with the smallest success rate appear in yellow, some of which show much greater differences between success and failure firing rates than those of other success ranges, as well as non-constant differences with upward sloping tending to start after the 10th time bin. Sessions with success rates in between the other two ranges generally show lower and more constant differences, similar to those of sessions with success rates at or above 0.8, except for one session which shows similar difference activity to sessions with lower success rates.

To explore the differences between neuron firing rates of different decisions, neuron firing rates across all sessions can be averaged over decision and feedback type combinations. To do this, average neuron firing rates calculated for each trial can be added to an overall dataframe, which then has averages taken over each decision type and feedback combination. This results in eight 40-element vectors, one for each decision type and feedback outcome combination (e.g. one combination being Decision 1 - Failure), which when plotted can aid in visualizing neuron firing rate patterns across decisions.

## Average Neuron Firing Rate for Decision Type (Success and Failure)

This plot shows the neuron firing rate across all trials and sessions for the four decisions, separated by the feedback outcome of success (solid line) and failure (dashed line) for each decision. Decision 1 represents the case where the left contrast (the contrast of the left stimuli) is greater than the right contrast (the contrast of the right stimuli). Decision 2 is the case where the right contrast is greater than the left contrast. In these two cases, success occurs when the mouse in a trial turns towards the greater contrast. Decision 3 is the case where the left contrast and right contrast is 0. Success occurs when the mouse does not turn towards either contrast. Decision 4 is the case where the left and right contrasts are equal but non-zero. Success occurs when the mouse turns towards the randomly chosen correct stimulus.

From this plot, it can be observed that for Decisions 1, 2, and 4, the neuron firing rate for success is greater than that of failure. However for Decision 3, the neuron firing rate for failure is greater than that of success. Since success for trials of Decision type 3 is determined by no movement by the mouse, higher neural activity for failure may be seen due to mice trying to choose one stimuli over another, as needed in every other decision type in this visual discrimination task, rather than correctly assess that both the stimuli are zero.

To further understand the difference between the success and failure neuron firing rates for each decision, the difference of the success and failure vectors for each decision can be graphed.

Average Neuron Firing Rate per Time Bin
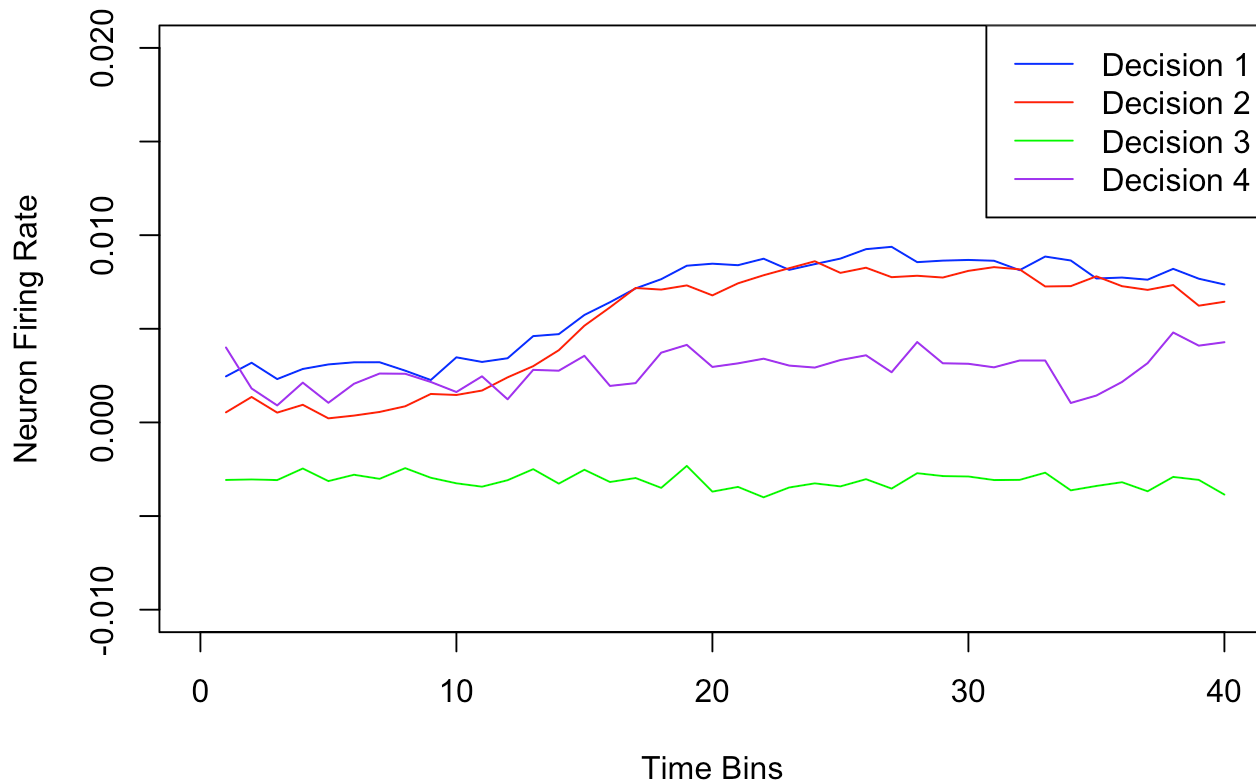
From this plot, the difference between the neuron firing rates resulting in success and failure are near constant for Decision 3 and Decision 4. Decision 3 and 4 are cases where both stimuli have equal contrast levels, but the decisions are differentiated by either both stimuli being zero or non-zero. In Decision 3, both stimuli have contrast 0 and the lowest difference in neuron firing rates, a slightly negative difference, can be seen but as a relatively flat line. In Decision 4, both stimuli have equal contrast levels but are non-zero, and we see a similarly flat line for the difference in neuron firing rates, but with a higher level of neural activity.

The difference between the firing rates resulting in success and failure is not constant for Decisions 1 and 2, however the difference lines are very similar for these two decisions, following a similar curve pattern of increase followed by a leveling out. Feedback for these two decisions types is awarded in a similar way, based on turning the wheel towards the greater contrast level.

We see similarities between the difference curve patterns and firing rates, nearly lining up entirely, for Decisions 1 and 2, two decision types which have a similar method of determining success and failure. We see similarities between the line patterns for Decisions 3 and 4, both being relatively flat, but different firing rates, with Decision 4 having higher levels than Decision 3.

For the purpose of predicting feedback, experiment data can be subsetted according to decisions, allowing for three separate models to be created. One model will contain data from Decisions 1 and 2, due to the differences in average neuron firing rate for success and failure being very similar, but from time bins 20 to 40 (to capture only the constant differences between success and failure for each decision). The other two models will contain data from Decision 3 and Decision 4, one model for each decision type in order to capture the differences in firing rate levels between the two decisions. However, as these two decisions show constant differences between success and failure neuron firing rates across all 40 time bins, these two models will be trained and tested on data from all time bins.

However, due to the current dataframe that has been created, where each row represents one trial in one session with all neuron activity for each trial averaged over the 40 time bins, there is the possibility that specific neuron data cannot be captured in the modeling process. As each session contains multiple (the number of trials worth of) spike trains for each neuron, the firing rates of each neuron in a session over 40 time bins can be averaged over all trials in that session.

After seeing that decisions types affect neuron spike activity and patterns, the spike activity for each neuron in this study (each individual neuron in each session) can be averaged over trials. These individual neuron firing rates can be averaged specifically over trials with each decision and feedback type combination, with the goal of being incorporated into predictive feedback modeling according to decision types. With over 16,000 individual neurons having recorded activity in this study, the firing rate for each neuron, averaged over trials with particular decision types can be explored. Individual neurons will be clustered according to their firing rates across each decision type and feedback combination using K-Means clustering, carried out for both 3 and 4 clusters.

In order to perform clustering over all neurons, a dataframe with neuron firing rate information is first created. In this dataset, each row represents one specific neuron that had activity recorded in one of the study's 18 sessions. Summing all of the unique neurons from all sessions, this dataset has 16,305 rows, one row for each neuron. This dataset has 321 columns. The first 320 columns include the average firing rate for the neuron across 40 time bins, for each decision and feedback outcome combination. The first 40 time bins in each row have the average firing rate for that neuron over the trials that fell under Decision 1 and had an outcome of success. The next 280 time bins have the average firing rate for that neuron over 40 time bins for each decision and feedback outcome combination, in the order of Decision 1 and failure, Decision 2 and success, Decision 2 and failure, Decision 3 and success, Decision 3 and failure, Decision 4 and success, and Decision 4 and failure. The last column includes the session number that each neuron comes from, to keep track of each neuron's origin in this overall dataframe.

K-Means: 3 Clusters

K-means clustering will be performed on a subset of this dataset, not including the session ID column, as the average firing rates for all decision and feedback outcome combinations are numerical data. This subset will be one-hot encoded and scaled before performing K-means clustering, initially for three clusters.

After performing K-means clustering with a center of 3, 3 clusters were determined with sizes 1782, 189, 14334.

```
## [1]  1782   189 14334
```

By looking at the cluster means for the first 6 time bins (out of 320 time bins), it appears that the third cluster of neurons shows the lowest average firing rates across time bins, as compared to the average firing rates for the two other clusters. The second cluster of neurons shows the highest average firing rates across time bins, being much higher than either of the other two cluster means. The first cluster of neurons shows average firing rates across time bins being in the middle of the other two clusters, but still much lower than those the second cluster. The actual values of the cluster means should not be interpreted directly, as the cluster means are computed from the scaled data.

```
##             X1          X2          X3          X4          X5          X6
## 1   1.2795528   1.3102090   1.2953579   1.3134553   1.3329491   1.3086736
## 2   5.9872106   5.9280873   5.9460802   5.9763226   5.9801259   5.9742164
## 3  -0.2380177  -0.2410493  -0.2394403  -0.2420889  -0.2445625  -0.2414667
```

The cluster vector (containing the cluster number for each neuron in this dataframe) can be extracted and binded to the neuron dataframe. This will allow for further analysis of neuron firing rates in clusters, within sessions themselves and eventually individual trials (with the neuron firing rates still split between decision and feedback
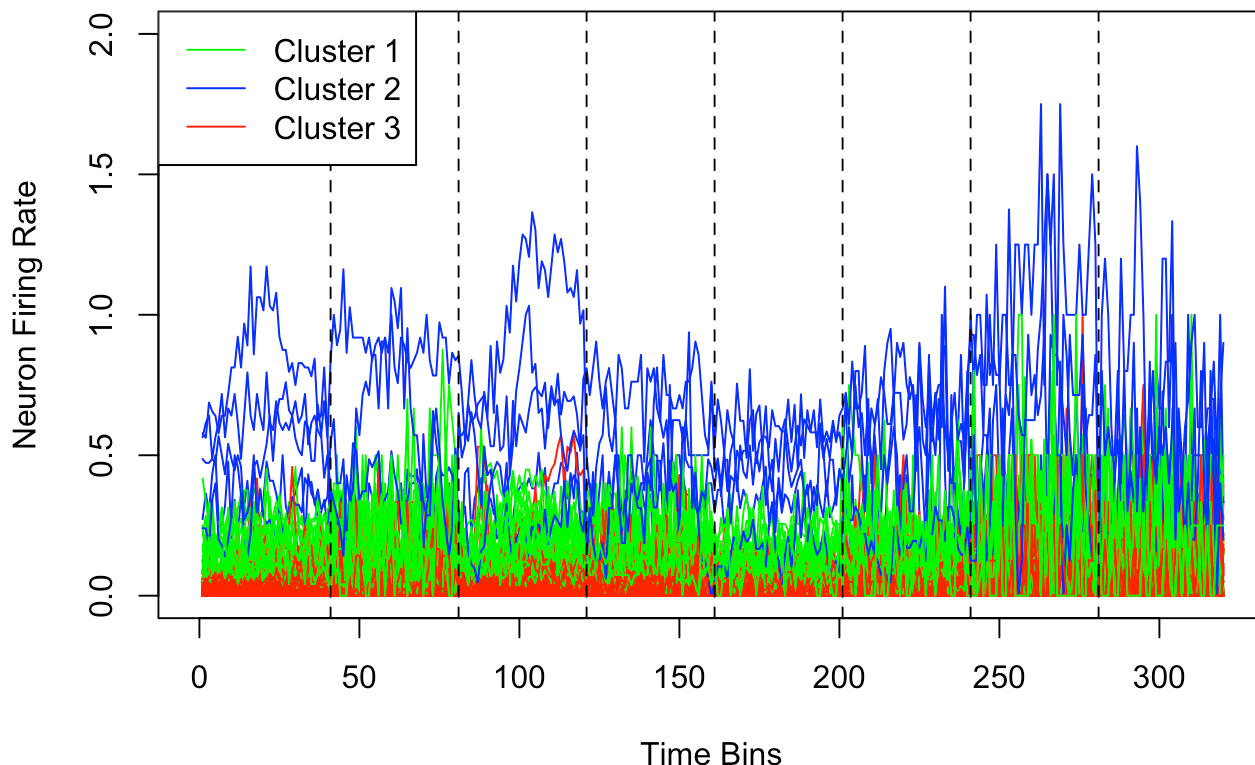
type combinations).

It's important to note for later data integration that cluster 2 has a small size, meaning some sessions may not have neurons in that cluster, for example, in session 16. As more numbers of clusters are explored, cluster sizes may also continue to shrink.

```
## [1] 0
```

To visualize the similarities in firing rates among neurons in the same clusters, 3000 random neurons can be picked to be graphed over the 320 time bins.

Average Neuron Firing Rate with 3 Clusters

In this plot, neurons in the same cluster have similar neural activity and trends across decision and feedback combinations. Neurons in Cluster 3 have the lowest activity, with their average neural spikes (in red) found to be very low on the graph. Cluster 1 neurons, seen in green, have higher activity than those in Cluster 3 with their average neural spikes being in the middle of the graph (and in the middle of the activity of the other two clusters). Cluster 2 has neurons with the highest neural activity, with average firing rates for those neurons being much higher than those of the other clusters. Before beginning data integration with this cluster information to build models for prediction, K-Means clustering can be performed and visualized once again for 4 clusters. This will allow for comparison of prediction models based on the different number of clusters later in this report.

K-Means: 4 Clusters

After performing K-means clustering with a center of 4, 4 clusters were determined with sizes 531, 60, 2444, and 13270.

```
## [1]    531     60   2444 13270
```

The cluster means for the first 6 time bins (out of 320 time bins) for the K-Means 4 Cluster results can be looked into again. It appears that the fourth cluster of neurons shows the lowest average firing rates across time bins for each decision type and feedback outcome combination, as compared to the average firing rates for the three other clusters. The second cluster of neurons in these results also shows the highest average firing rates across time bins for the decision type and feedback outcome combinations, being much higher than any of the other cluster means. The third cluster of neurons shows average firing rates being similar but slightly higher than those of the fourth cluster. The first cluster of neurons shows average firing rates across time bins being in the middle of all the other clusters, but still much lower than those the second cluster. Again, the actual values of the cluster means should not be interpreted directly, as the cluster means are computed from the scaled data.

```
##              X1          X2         X3          X4          X5          X6
## 1    2.9665340   3.0065799   2.986688   2.9304440   2.9499863   2.9797427
## 2    9.0113711   9.1827707   9.115610   9.1979164   9.1446840   9.1682971
## 3    0.6933076   0.6823037   0.682142   0.7090842   0.7246632   0.6892985
## 4   -0.2871406  -0.2874914  -0.286362  -0.2894456  -0.2928561  -0.2876403
```

The cluster vector (containing the cluster number for each neuron in this dataset) can be extracted and binded to the neuron dataframe again. This will allow for further analysis of neuron firing rates in the 4 clusters.

Once again, to visualize the similarities in firing rates among neurons in the same clusters, 3000 random neurons can be picked to be graphed over the 320 time bins.

Average Neuron Firing Rate with 4 Clusters
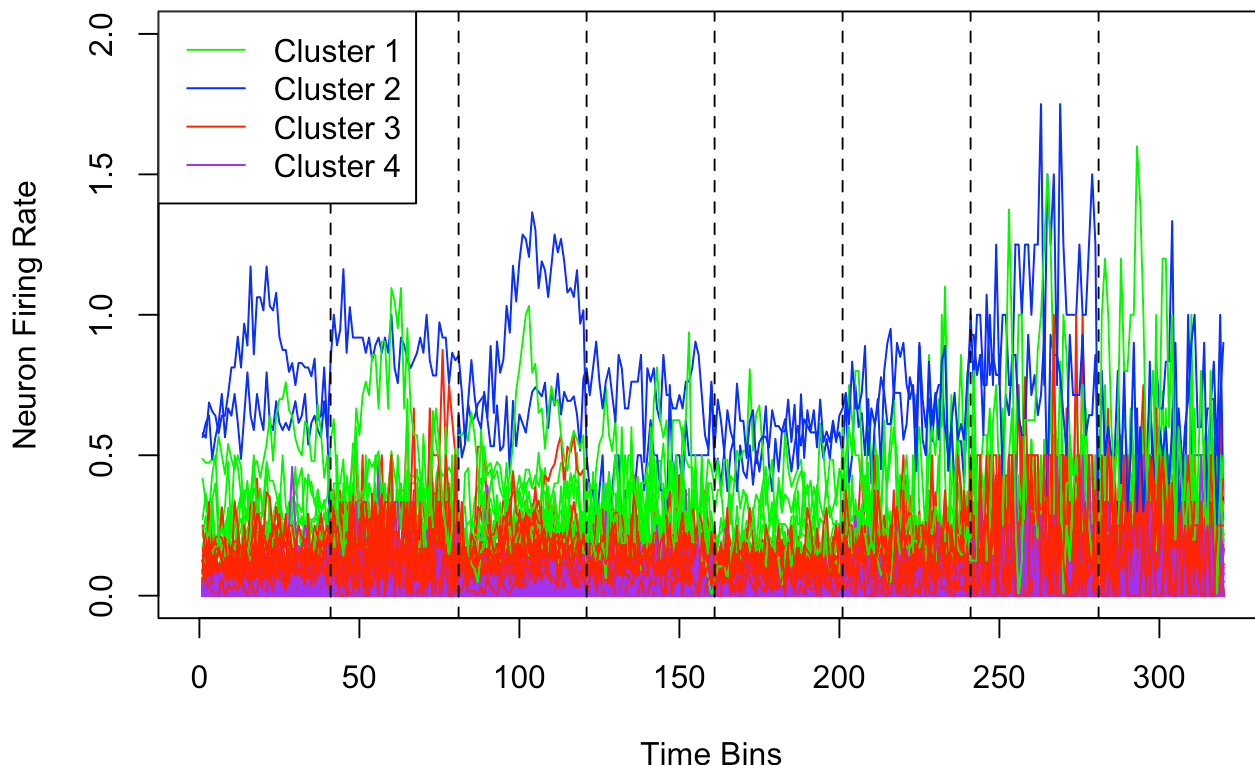
In this plot, neurons in the same cluster also have similar neural activity and trends across decision and feedback combinations. Neurons in Cluster 4 have the lowest activity, with their average neural spikes (in purple) found to be very low on the graph. Cluster 3 neurons, seen in red, have higher neural activity than those in Cluster 4, then followed by Cluster 1 neurons, seen in green. Cluster 2 has neurons with the highest neural activity, with average neural spikes being higher than those of the other clusters. However, it seems that the gap between firing rates for neurons in Cluster 2 and Cluster 1 are smaller than in the K-Means results for 3 clusters.

Section 3: Data Integration

For Data Integration, two separate dataframes will be built to incorporate the results from the different K-Means clustering methods into trial information.
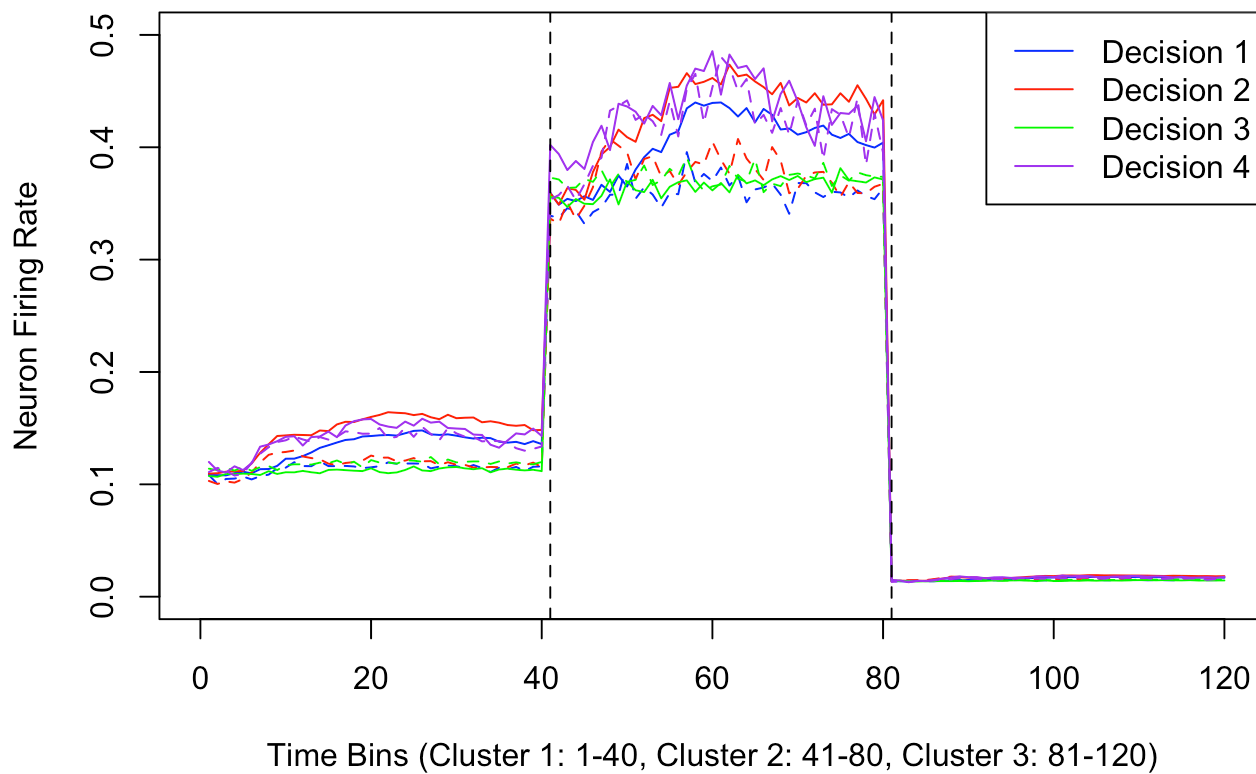
3 Clusters

The first dataframe will be created with neuron activity incorporated based on the results of the K-Means clustering for 3 clusters. With a cluster type assigned to each neuron, it is now possible to compute the average firing rate for neurons in a cluster, in each trial. In the first new dataframe, each row will contain information for each trial in each session. The trial information contained in each row will include trial number, session number, feedback type, and decision type. Each row will also contain the average firing rate over 40 time bins for the neurons in each cluster that are present in that trial. As this dataset incorporates neuron information based on 3 clusters, a vector of 120 time bins will be included in the row for each trial, with 40 time bins for each neuron cluster firing rate averages. This new dataframe will have the dimensions of 5081 rows, the total number of trials in these experiments, by 124 columns (with the four key trial identifiers listed above, as well as the 120 time bins explained above).

As this dataset averages firing rates over neurons in each cluster for every trial, it requires neurons from every cluster to be present in each trial (and each session). To take into account the possibility that some sessions may not have neurons present in each cluster type, due to small cluster size, NA values generated by this occurrence will be filled with the value 0. If there are no neurons present in a cluster type for a session, then the average firing rate for neurons in that cluster will be 0 across the 40 time bins.

The neuron firing rates averaged over each decision and feedback type can be visualized again, but now split between cluster types.

Average Neuron Firing Rate per Time Bin

Time Bins (Cluster 1: 1-40, Cluster 2: 41-80, Cluster 3: 81-120)

As seen in the K-means clustering results for 3 centers, particularly in the cluster means, Cluster 2 shows very high neural activity as compared to Cluster 1 and 3. It can still be seen that Cluster 1 has higher firing rate levels than Cluster 3, which has firing rate levels near 0.

The difference of the success and failure vectors can be plotted again for each decision type.

Average Neuron Firing Rate per Time Bin

Time Bins (Cluster 1: 1-40, Cluster 2: 41-80, Cluster 3: 81-120)
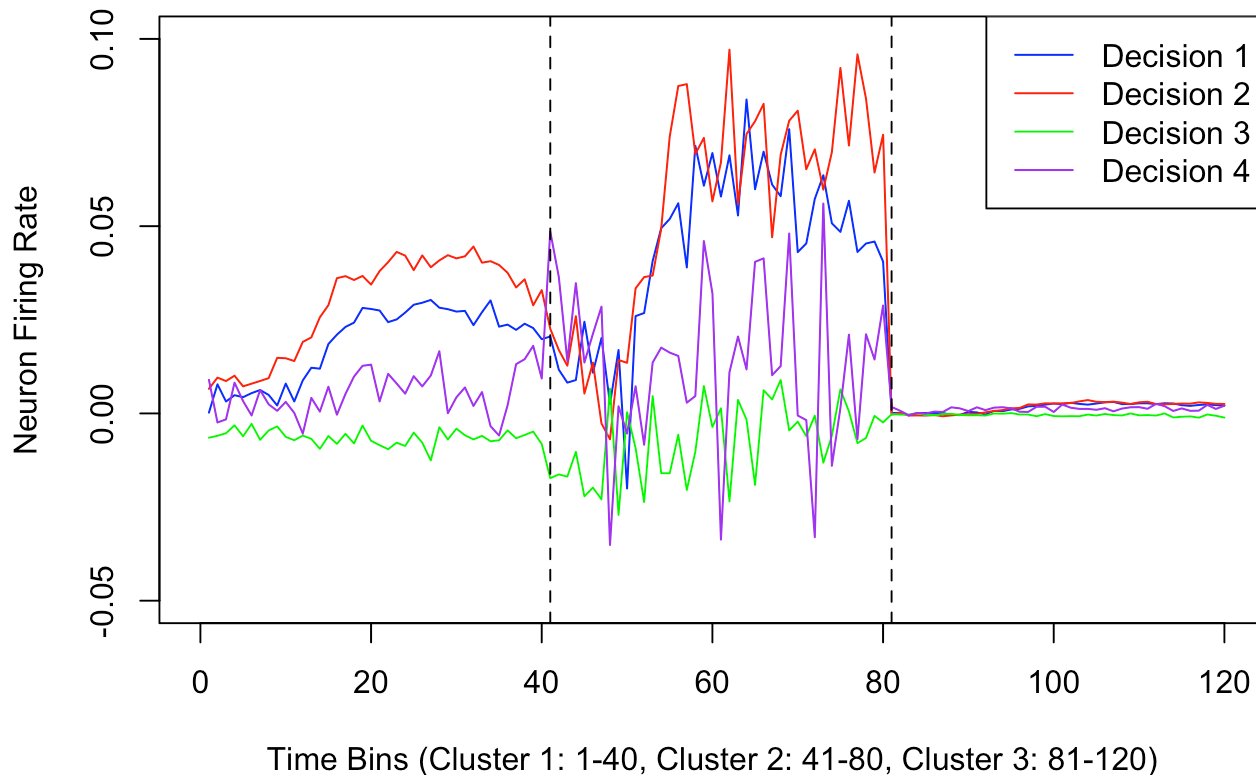
In this plot, we can look at the difference in neural activity in each cluster, for 3 clusters.

In Cluster 1, we see fairly similar patterns as the initial decision success and failure difference plot (before clustering). Decision 3 and Decision 4 saw the "flattest" and lowest differences (Decision 3 having the lowest) across the 40 time bins. However, though we see Decision 1 and 2 still following similar difference patterns, there is a larger gap between the difference curves with Decision 2 having greater firing rate levels. The curves for Decisions 1 and 2 appear to flatten after about 20 time bins, with slight sloping down from Decision 2.

In Cluster 2, there are much higher levels of neural activity, as well as more spiking in activity, having higher peaks and lower troughs than any curves in the other clusters. We can still see similar overall trends in these curves, with Decision 3 and Decision 4 having generally flatter trends in their difference curves across the 40 time bins, with Decision 3 still having lower neural activity than Decision 4. For Decision 1 and 2, both curves follow a similar pattern as seen previously, however with a larger dip around the 10 time bin mark (time bin 50 in the graph). The two curves again appear to flatten after 20 time bins, still with spiking, though Decision 1 sees slight sloping down in this cluster.

In Cluster 3, the lowest neural activity is seen, with all difference curves lying near 0. It seems that Decision 3 may still have the lowest and flattest difference curve, followed by Decision 3, and also seeing similar patterns in the curves for Decision 1 and 2.
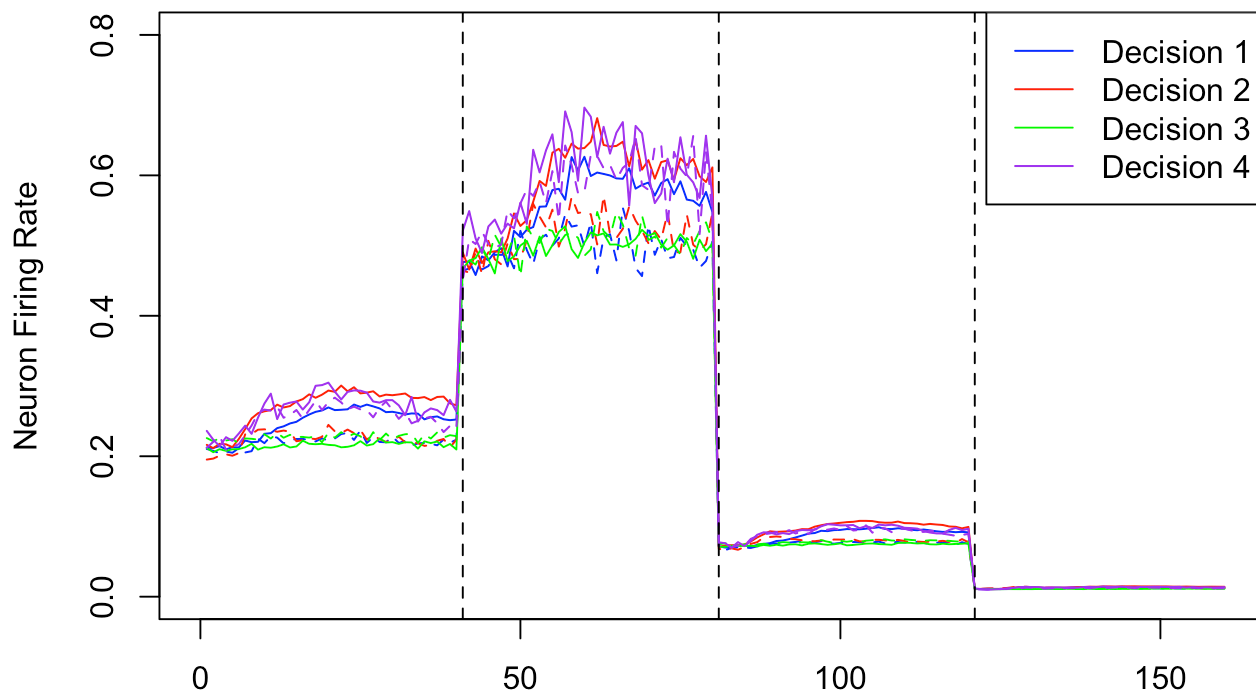
4 Clusters

The second dataframe will be created with neuron activity incorporated based on the results of the K-Means clustering for 4 clusters. In the second new dataframe, each row will still contain information for each trial in each session, including trial number, session number, feedback type, and decision type. Each row will also contain the average firing rate over 40 time bins for the neurons in each cluster that are present in that trial. As this dataset

incorporates neuron information based on 4 clusters, a vector of 160 time bins will be included in the row for each trial, with 40 time bins for each neuron cluster firing rate average. This new dataframe will have the dimensions of 5081 rows, the total number of trials in these experiments, by 164 columns (with the four key trial identifiers listed above, as well as the 160 time bins explained above).

The neuron firing rates averaged over each decision and feedback type can be visualized again, but now split between cluster types for 4 clusters.

## Average Neuron Firing Rate per Time Bin

Time Bins (Cluster 1: 1-40, Cluster 2: 41-80, Cluster 3: 81-120, Cluster 4: 121-160)

As seen in the K-means clustering results for 4 centers, particularly in the cluster means, Cluster 2 shows very high neural activity as compared to all other clusters. It can still be seen that Cluster 1 has higher firing rate levels than Clusters 3 and 4. Cluster 3 and 4 have firing rate levels near 0, with those of Cluster 4 being slightly higher.

The difference of the success and failure firing rates can be plotted again for each decision type.

Average Neuron Firing Rate per Time Bin

Time Bins (Cluster 1: 1-40, Cluster 2: 41-80, Cluster 3: 81-120, Cluster 4: 121-160)
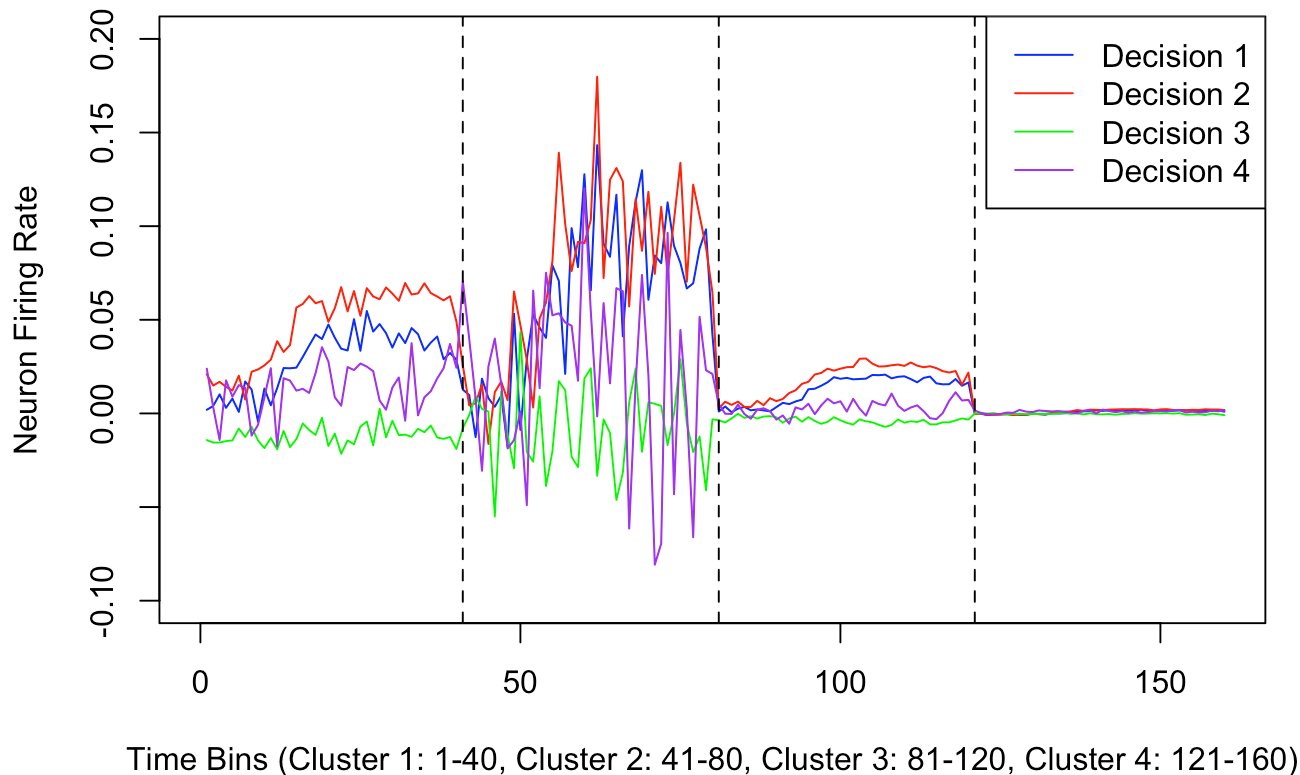
In this plot, we can look at the differences in neural activity between each cluster for 4 clusters.

In Cluster 1, there are fairly similar patterns as the initial decision success and failure difference plot (before any clustering). Decision 3 and Decision 4 saw the most constant and lowest differences (Decision 3 having the lowest) across the 40 time bins. Though we see Decision 1 and 2 still following similar difference patterns, there is again a larger gap between the difference curves and Decision 2 continues to have greater levels. The curves for Decisions 1 and 2 generally flatten after the first 20 time bins.

In Cluster 2, there are much higher levels of neural activity, as well as greater spiking in activity, similar to what is seen in this plot for 3 clusters. Similar trends can be seen in these curves, Decision 3 and Decision 4 having generally flatter trends in their difference curves across the 40 time bins, with Decision 3 still having lower neural activity than Decision 4. For Decision 1 and 2, both curves follow a similar pattern as seen previously, this time without a larger dip around the 10 time bin mark as seen in this plot for 3 clusters. The two curves again appear to relatively flatten after 20 time bins, still with spiking.

In Cluster 3, one can see similar patterns to the initial decision success and failure difference plot (before any clustering), like for Cluster 1, but at overall lower neuron firing rate levels. In Cluster 4, the lowest neural activity is seen, with all difference curves near 0 and the differences for each decision difficult to interpret.

Section 4: Predictive Modeling

3 Clusters

From the dataset that incorporates K-Means clustering with 3 clusters, logistic regression models can be created to predict feedback outcome. Three models will be created, with one model trained and tested on trials that fall under Decision 1 and 2, a second model for Decision 3 trials, and a third model for Decision 4 trials.

Since similarities were seen between the difference curve patterns and firing rates of Decisions 1 and 2, and these two decision types have a similar method of determining success and failure, a model will be created that is trained and tested on trial data from both decisions. However, the data will be subsetted according to time bins, leaving off the first 20 time bins to capture the constant differences between success and failure neuron firing rates.

Decisions 1 and 2:

```
## [1] 0.2017673
```

The Logistic Regression model for Decision 1 and 2 data has a 0.2018 error rate on the test data.

Decision 3:

```
## [1] 0.4072727
```

The Logistic Regression model has a 0.4073 error rate for Decision 3 test data.

Decision 4:

```
## [1] 0.4603175
```

The Logistic Regression model has a 0.4603 error rate for Decision 4 test data.

XGBoost can also be used for the same predictive modeling purposes as Logistic Regression.

Decision 1 and 2:

```
## [1]   train-logloss:0.433785
## [2]   train-logloss:0.367326
## [3]   train-logloss:0.336252
## [4]   train-logloss:0.302558
## [5]   train-logloss:0.279016
```

```
## [1] 0.2724595
```

The XGBoost model shows an error rate of 0.2725 on the Decision 1 and 2 test data.

Decision 3:

```
## [1]   train-logloss:0.501482
## [2]   train-logloss:0.440051
## [3]   train-logloss:0.372060
## [4]   train-logloss:0.314412
## [5]   train-logloss:0.272672
```

```
## [1] 0.4
```

The XGBoost model has an error rate of 0.4 for Decision 3 test data.

Decision 4:

```
## [1]  train-logloss:0.392550
## [2]  train-logloss:0.227868
## [3]  train-logloss:0.146582
## [4]  train-logloss:0.100851
## [5]  train-logloss:0.075864
```

```
## [1] 0.5079365
```

The XGBoost model shows an error rate of 0.5079 for Decision 4 test data.

4 Clusters

From the dataset that incorporates K-Means clustering with 4 clusters, logistic regression models can be created again to predict feedback outcome. Three models will be created similar to previously, with one model trained and tested on trials that fall under Decision 1 and 2, a second model for Decision 3 trials, and a third model for Decision 4 trials.

The model for Decisions 1 and 2 will have the first 20 time bins left off Clusters 1 and 2.

Decisions 1 and 2:

```
## [1] 0.2120766
```

The Logistic Regression model for 4 clusters, for Decision 1 and 2 test data, has a 0.2121 error rate.

Decision 3:

```
## [1] 0.4
```

The Logistic Regression model has a 0.4 error rate for Decision 3 test data.

Decision 4:

```
## [1] 0.4444444
```

The Logistic Regression model has a 0.4444 error rate for Decision 4 test data.

It is also possible to create XGBoost models on the dataset that takes 4 clusters into account.

Decision 1 and 2:

```
## [1]  train-logloss:0.434621
## [2]  train-logloss:0.370445
## [3]  train-logloss:0.318073
## [4]  train-logloss:0.289172
## [5]  train-logloss:0.257943
```

```
## [1] 0.2488954
```

The XGBoost model for 4 clusters shows an error rate of 0.2489 for Decisions 1 and 2 test data.

Decision 3:

```
## [1]  train-logloss:0.489696
## [2]  train-logloss:0.393715
## [3]  train-logloss:0.323306
## [4]  train-logloss:0.262039
## [5]  train-logloss:0.210098
```

```
## [1] 0.3636364
```

The XGBoost model for 4 clusters has an error rate of 0.3636 for Decision 3 test data.

Decision 4:

```
## [1]  train-logloss:0.407377
## [2]  train-logloss:0.227736
## [3]  train-logloss:0.146532
## [4]  train-logloss:0.095851
## [5]  train-logloss:0.074035
```

```
## [1] 0.5555556
```

The XGBoost model for 4 clusters on Decision 4 test data has an error rate of 0.5556.

To compare the various models created for the data subsets (according to decision), the ROC curves can be graphed for each model.

Decisions 1 and 2:
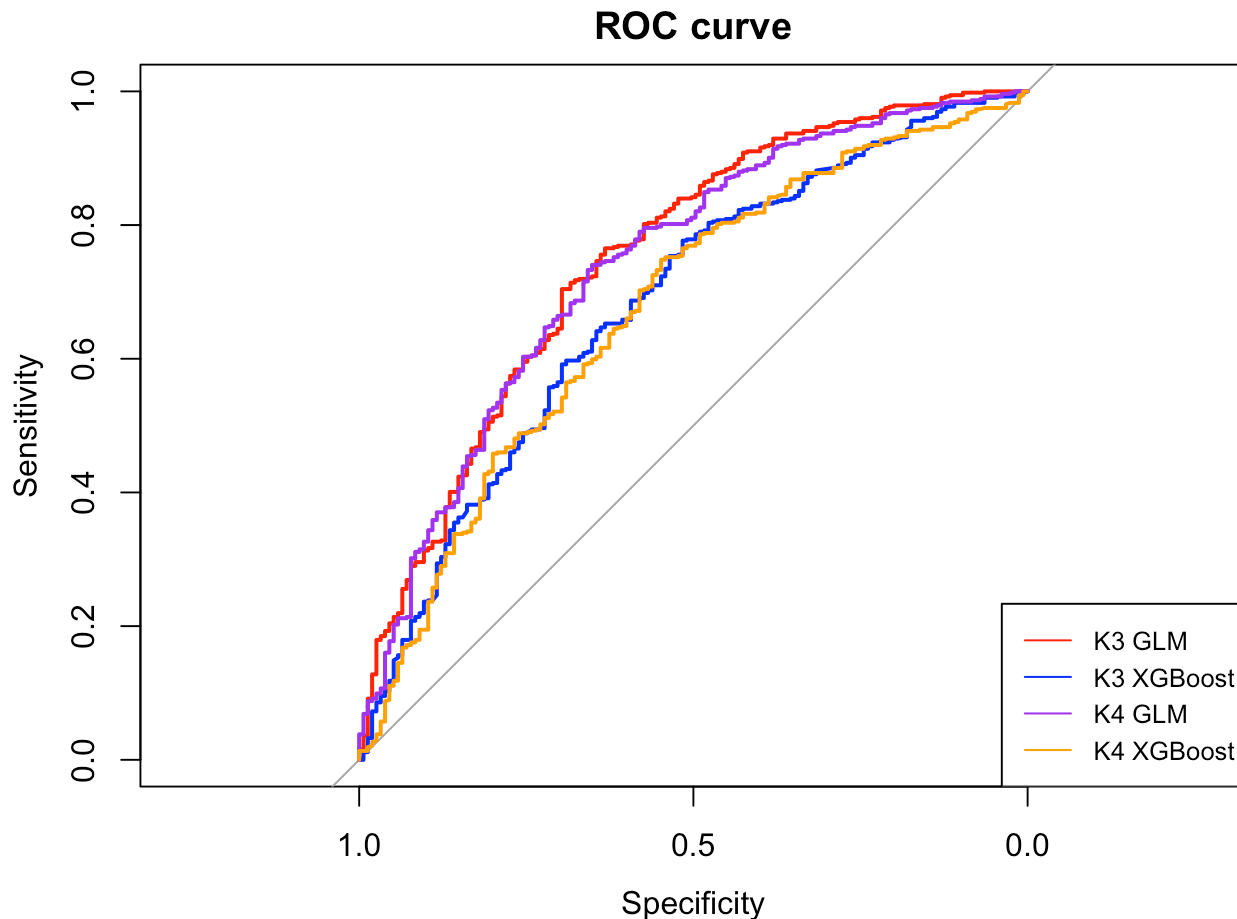
ROC curve

From the ROC curves, similar performance is seen by the XGBoost models and by Logistic Regression models, across the different clustered datasets. The Logistic Regression models for either the dataset with 3 or 4 clusters performed better than the XGBoost models for these datasets. The performance between the Logistic Regression models was very similar, as with the performance between the XGBoost models.

Model	AUC
K3 GLM	0.7497
K3 XGBoost	0.6813
K4 GLM	0.7403
K4 XGBoost	0.6736

Similar results are seen from the AUC values. The Logistic Regression models have very similar performance, with the model for 3 clusters dataset performing slightly better than the dataset with 4 clusters. The XGBoost models also have very similar performance, but are both lower than the performances of the Logistic Regression models, and again the model for 3 clusters dataset performs slightly better than the dataset with 4 clusters.

The best performing model for Decisions 1 and 2 was the Logistic Regression model trained on a dataset that incorporates K-Means clustering for 3 clusters.

Decision 3:

ROC curve
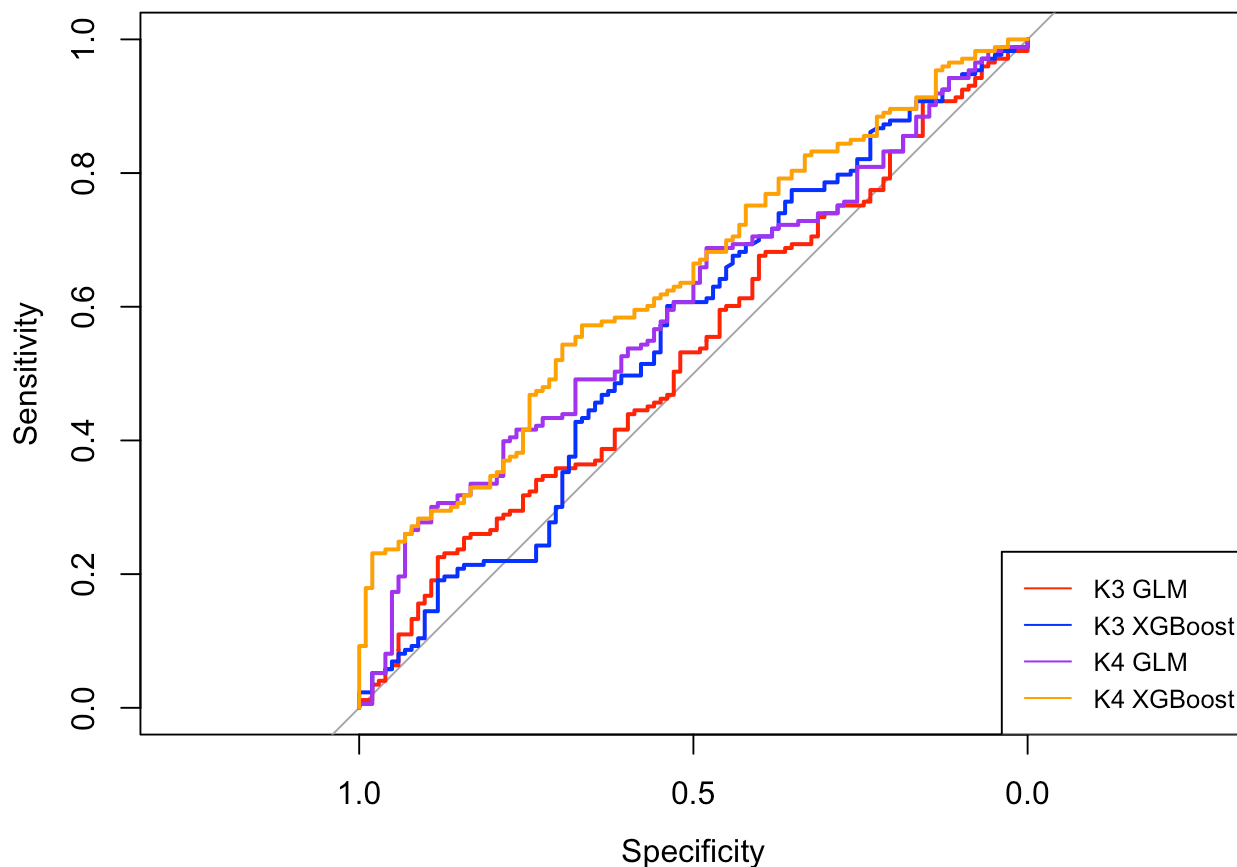
From the ROC curves, performance by the different models seems to vary more in predicting data for Decision 3 than data for Decision 1 and 2. The Logistic Regression and XGBoost models for the dataset with 4 clusters seemed to perform better than those for the dataset with 3 clusters. It seems that the XGBoost models performed better than their Logistic Regression counterparts for each clustered dataset.

Model	AUC
K3 GLM	0.5361
K3 XGBoost	0.5594
K4 GLM	0.5975
K4 XGBoost	0.6412

Similar results are seen from the AUC values. The best performing model was the XGBoost model trained and tested on the dataset with 4 clusters, followed by the Logistic Regression model trained and tested on the dataset with 4 clusters. For the models trained and tested on the dataset with 3 clusters, the XGBoost model also performed better than the Logistic Regression model.

The best performing model for Decision 3 was the XGBoost model trained on a dataset that incorporates K-Means clustering for 4 clusters.

Decision 4:

ROC curve
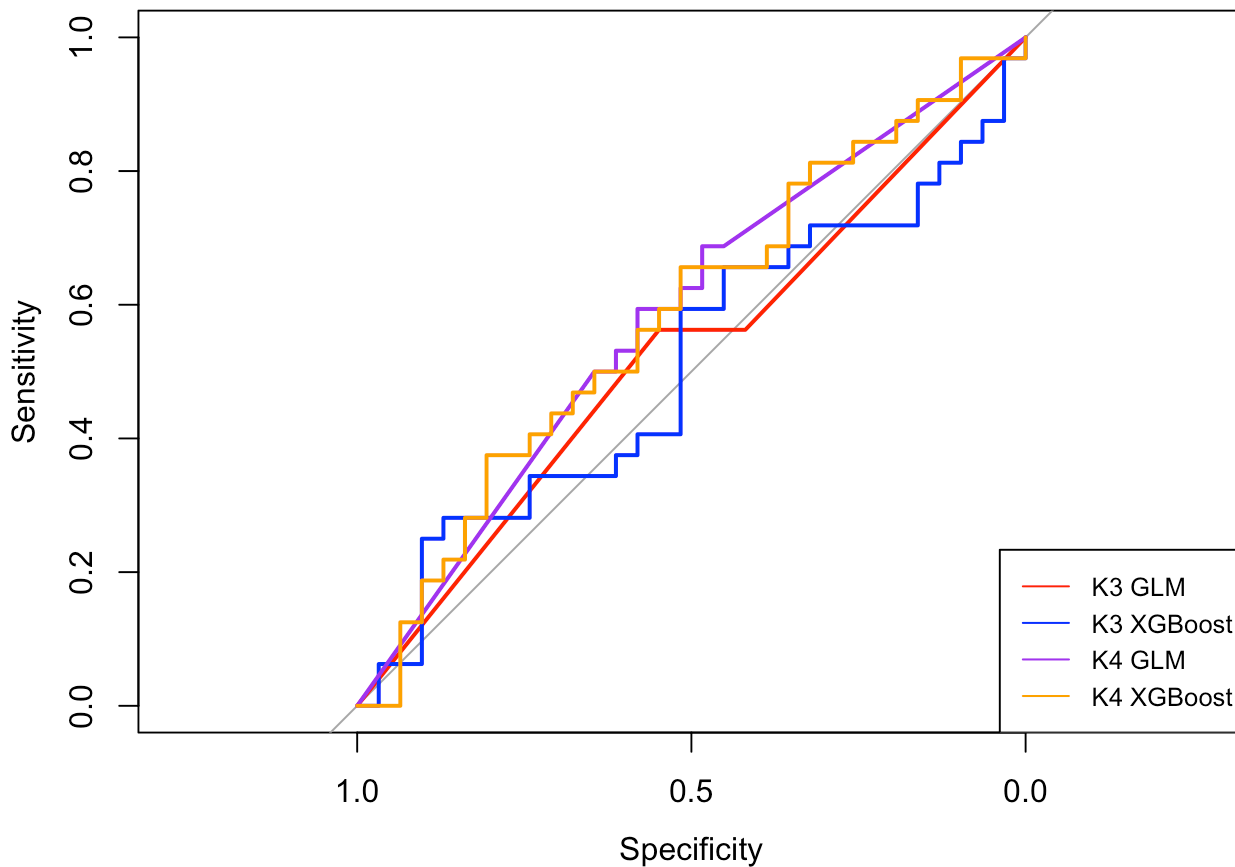
From the ROC curves, performance by the different models again varies more in predicting data for Decision 4 than data for Decision 1 and 2. Models for Decision 4 also perform the lowest of models for all decision types. The Logistic Regression and XGBoost models for the dataset with 4 clusters seemed to perform better than those for the dataset with 3 clusters. However, the XGBoost models performed either similarly (4 clusters dataset) or worse (3 clusters dataset) than their Logistic Regression counterparts for each clustered dataset.

Model	AUC
K3 GLM	0.5272
K3 XGBoost	0.5091
K4 GLM	0.5837
K4 XGBoost	0.5817

Similar results are seen from the AUC values. The best performing models were the XGBoost and Logistic Regression models trained and tested on the dataset with 4 clusters, followed by the Logistic Regression model trained and tested on the dataset with 3 clusters. For the models trained and tested on the dataset with 3 clusters, the XGBoost model performed worse than the Logistic Regression model.

The best performing model for Decision 4 was the Logistic Regression model trained on a dataset that incorporates K-Means clustering for 4 clusters.

Section 5: Prediction Performance on Test Sets

As clustering needs to be performed, first the dataset with each row representing a neuron will be created.

K-means clustering, first for 3 centers, will be performed on a subset of this dataset, not including the session ID column. This subset will be one-hot encoded and scaled before performing K-means clustering.

The cluster vector for 3 clusters (containing the cluster number for each neuron in this dataset) will be extracted and binded to the test neuron dataset.

K-means clustering, for 4 centers, will be performed on a subset of this dataset, not including the session ID column. This subset will be one-hot encoded and scaled before performing K-means clustering.

The cluster vector for 4 clusters (containing the cluster number for each neuron in this dataset) will be extracted and binded to the test neuron dataset.

The same data integration processes will be carried out to create two dataframes for the test data, one dataframe including neuron firing rate averages for 3 neuron clusters for each trial and another dataframe including neuron firing rate averages for 4 neuron clusters for each trial.

Modeling

The best performing models trained and tested in this report's earlier prediction modeling will be applied to this test data, subsetted by decision types.

The best performing model for Decisions 1 and 2 was the Logistic Regression model trained on a dataset that incorporates K-Means clustering for 3 clusters.

Decision 1 and 2 Trials Error Rate:

```
## [1] 0.3333333
```

The Logistic Regression model for Decisions 1 and 2 performs with a 0.3333 error rate on the test data.

The best performing model for Decision 3 was the XGBoost model trained on a dataset that incorporates K-Means clustering for 4 clusters.

Decision 3 Trials Error Rate:

```
## [1] 0.3492063
```

The XGBoost model for Decision 3 performs with a 0.3492 error rate on the test data.

The best performing model for Decision 4 was the Logistic Regression model trained on a dataset that incorporates K-Means clustering for 4 clusters.

Decision 4 Trials Error Rate:

```
## [1] 0.4444444
```

The XGBoost model for Decision 4 performs with a 0.4444 error rate on the test data.

Section 6: Discussion

Through data exploration, neuron firing rates in trials were seen to have different patterns between trials that resulted in success and trials that resulted in failure. Neuron firing rates were also seen to differ between success and failure in trials with different decision types. With this knowledge, performing two K-Means clustering methods resulted in 3 and 4 potential clusters for all neurons in these sessions. In predictive modeling, neuron firing rates averaged over neurons from 3 clusters proved to be better at predicting feedback outcomes for trials that fell under Decisions 1 and 2. Neuron firing rates averaged over 4 clusters proved to be better at predicting feedback outcomes for trials that fell under Decisions 3 and 4. However, prediction errors for the best models for Decision 3 and 4 trials were still higher than those for the best model for Decision 1 and 2 trials. It may be possible that the differences in how success and failure were determined between decision types contributed to these differences in the best performing models and their error rates.

Trials that fall under the type of Decision 3 have the left contrast and right contrast both with the value of 0, with success occurring when the mouse does not turn towards either contrast. Neuron firing rates for Decision 3 saw greater neuron firing rate levels in trials that resulted in failure than in success. As this may be due to mice trying to choose one stimuli over another, as needed in every other decision type for this visual discrimination task, neuron activity may not be a good way of predicting feedback success. If the neuron firing rates associated with making a choice of turning towards one of the stimuli (which results in failure for Decision 3 trials) are higher than the neuron firing rates associated with no movement of the wheel (success for Decision 3 trials) then neuron firing rates may make it difficult to predict correctly for success and failure.

Trials that fall under the type of Decision 4 have the left and right contrasts equal but non-zero, with success allocated for a mouse turning towards a randomly chosen correct stimulus. As success is determined randomly in Decision 4 trials, it would be difficult to predict the correct feedback outcome based on any variables, including neuron firing rates. This makes building predictive models for Decision 4 trials likely to have low prediction accuracy.

The prediction error rates for the three models ran on the test data, Logistic Regression for Decisions 1 and 2 trials with 3 neuron clusters, XGBoost for Decision 3 trials with 4 neuron clusters, and Logistic Regression for Decision 4 trials with 4 neuron clusters, increase with each decision type (lowest error rate for Decisions 1 and 2 and highest for Decision 4). As Decisions 3 and 4 are more difficult to predict success for using neuron firing rate data, we see higher error rates for models predicting these decisions. Decisions 1 and 2 had the highest prediction accuracy for the test data, which is consistent with the results from the initial prediction model building.

It can be concluded that average neuron firing rates are better at predicting the outcome for trials of Decision types 1 and 2 than for Decision types 3 and 4. There are likely to be three different types of neurons in this dataset, as seen in the K-Means clustering results and modeling results for three clusters. The models based on three neuron clusters best predicted the outcomes for Decision types 1 and 2, while the models based on four neuron clusters best predicted the outcomes for Decision types 3 and 4. As Decision types 3 and 4 are difficult to predict the feedback outcome of (as discussed earlier) based on neuron firing rates, this doesn't necessarily make four neuron clusters as likely in this study. Since models based on neuron firing rates averaged over three neuron clusters better predicted the feedback outcome of Decision types 1 and 2, which are easier to predict using neuron firing rates, there may be three neuron clusters (perhaps related to neuron types) recorded for in the 18 sessions analyzed of this study.