Mouse Project

Final Results

Rose, Jessie, Shuting

Agenda —— what we did

- **Identify research subjects: Zero_Maze experiment**

  Mouse 409 w/ 25 neurons and all neurons
  All mice w/ all neurons

- **Apply different models**

  Logistic Model with PCA, Simple Neural Network, RNN, Bidirectional LSTM

- **Outputs**

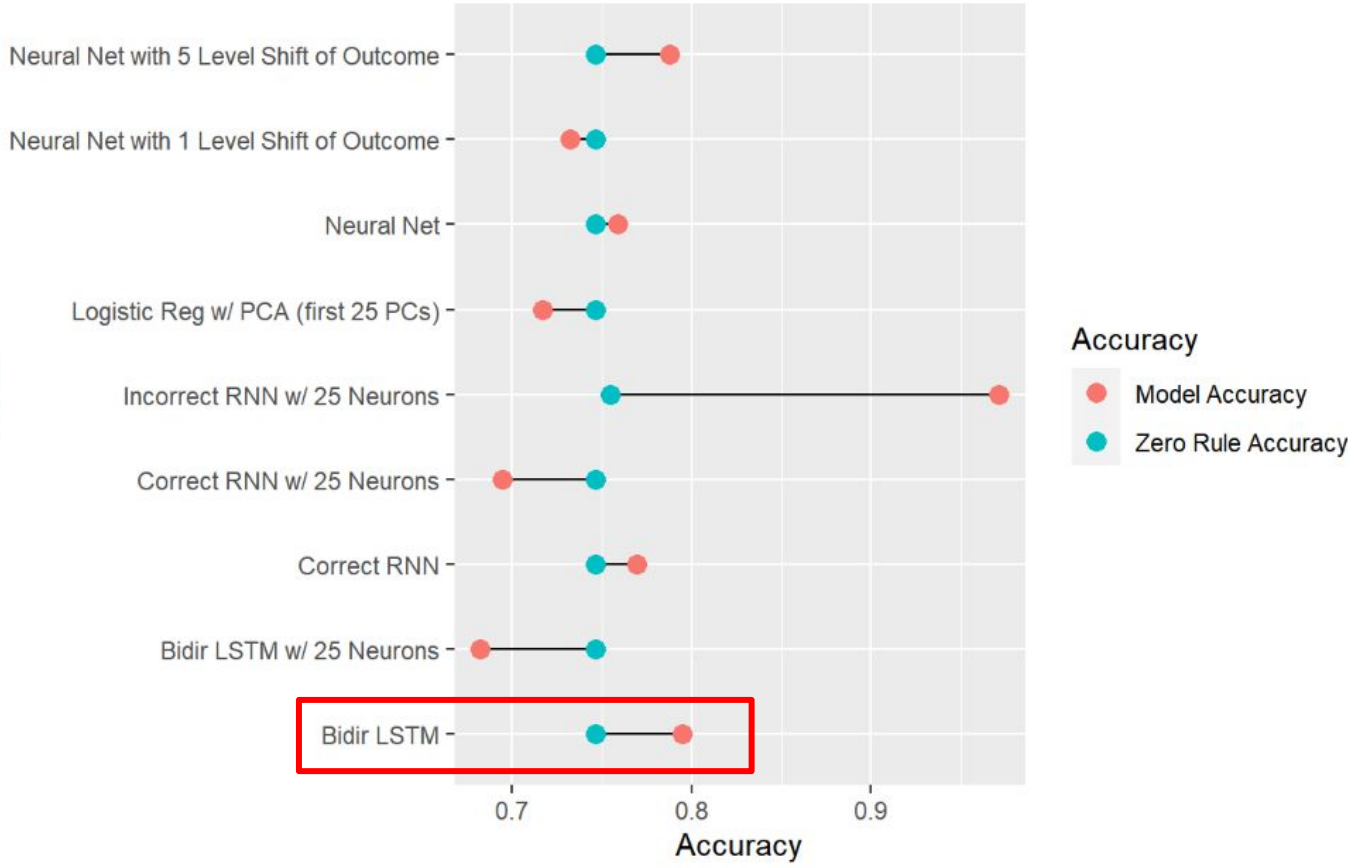
  Confusion Matrix for each model and Accuracy Comparison plot

Conclusions

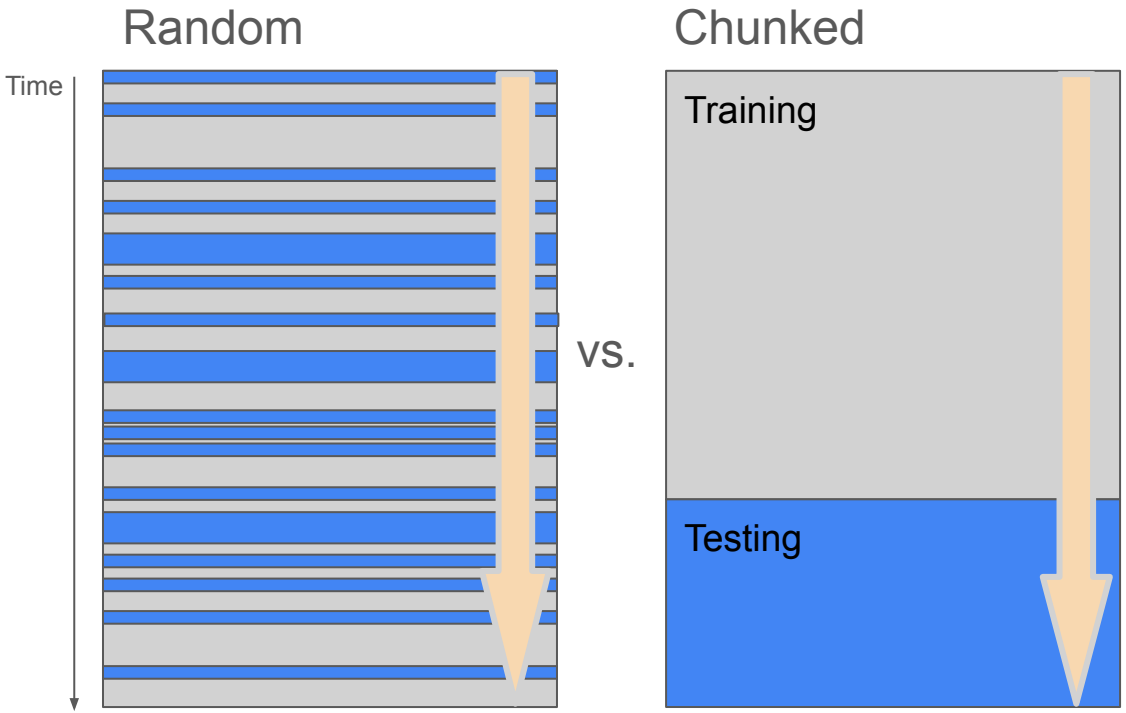
- **Split train / test in chunks** instead of randomly selected to ensure the sequence of time series
- Take **all neurons** into consideration to improve the predicted accuracy
- **Try different lags**, and it turns out that past events / neuron activity may not influence current outcome
- **Shift the behavior** to explore the causal relationship between neuron activation and behaviors

Results —— what we found

- **Model Comparisons with Mouse 409**
- **Zero Rule Accuracy =** Selecting the most common class
- **Goal =** Model Accuracy > Zero Rule Accuracy



Train/Test Split of Data



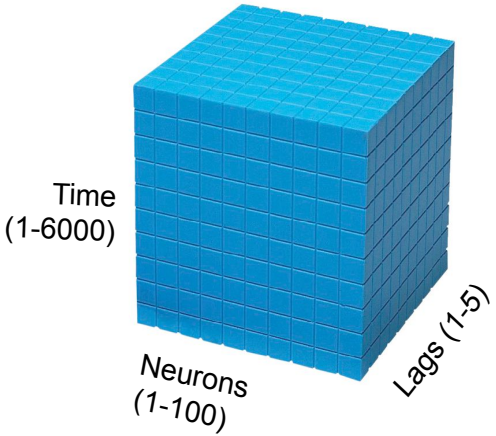
Random - learns a little bit of everything but loose sequence

Chunked - preserves sequence

Does a neuron convey the same information throughout the trial?

Bidirectional LSTM (Long Short Term Memory)

Input Data



LSTM Overview

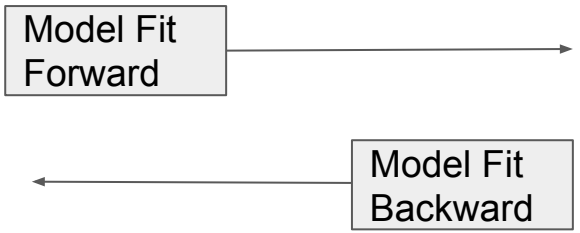
Memory

Model that retains prior sequence information

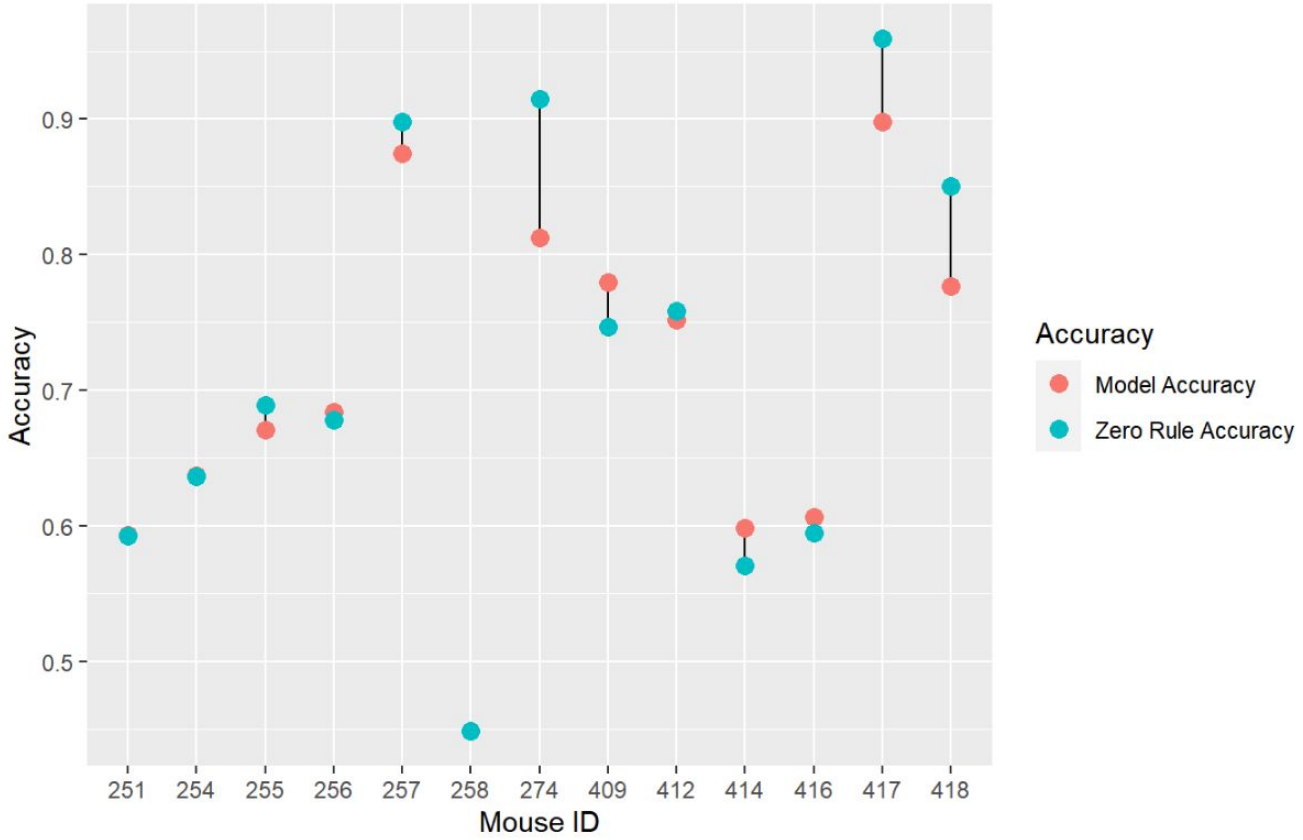
Bidirectional

Two directions

Information from before and after time point of interest

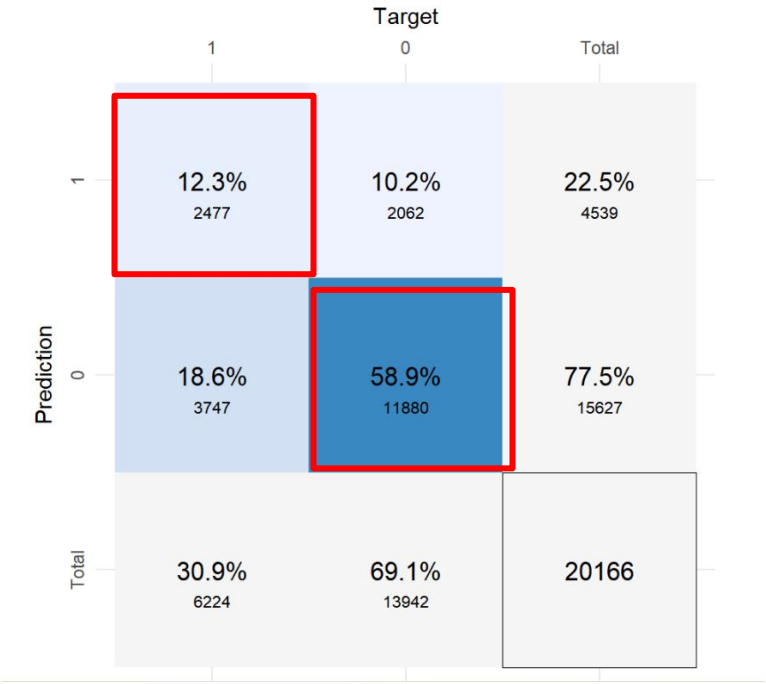


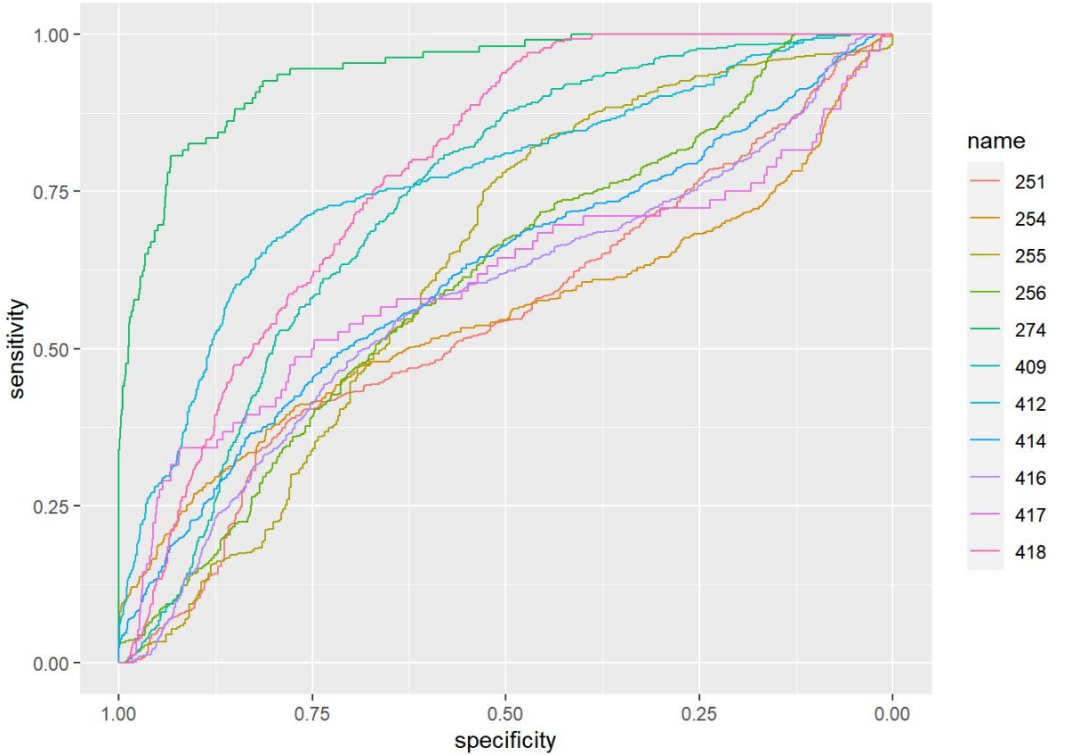
Results —— what we found



Logistic Model for all mice—— other model we tried (1)

- Select neurons (predictors) with PCA=25, make confusion matrix and ROC lines.
- Accuracy is around 71%

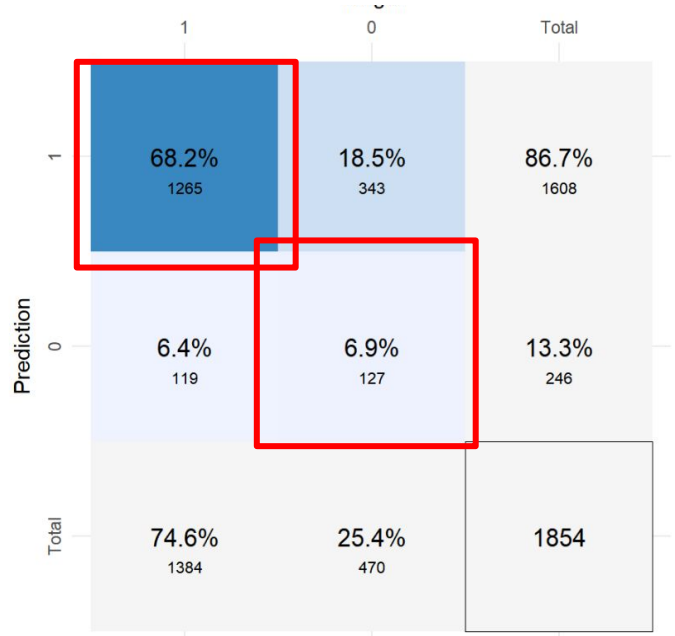


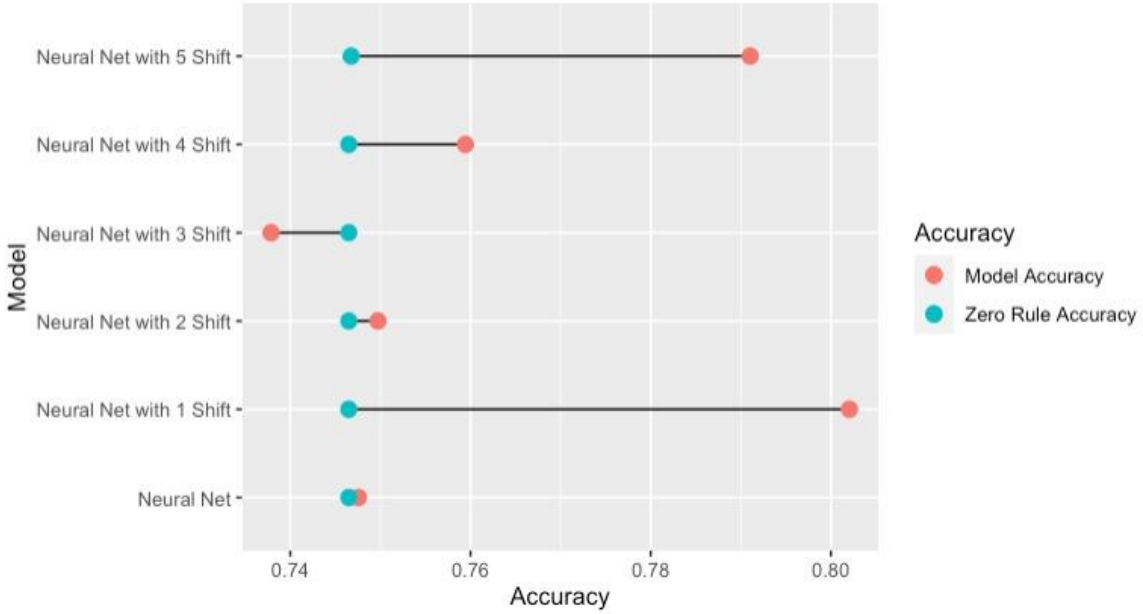


Simple Neural Network —— other model we tried (2)

- Features: not take into account the time series predicts the current location based on the current state of the neurons
- Accuracy = 75%

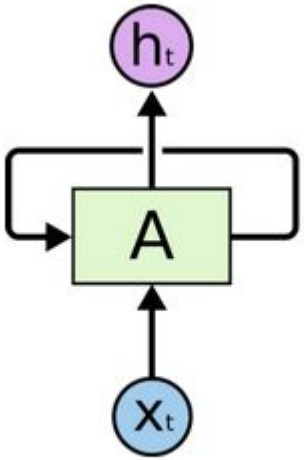
- **Shift the behavior (Y)** with 1 to 5 position forward

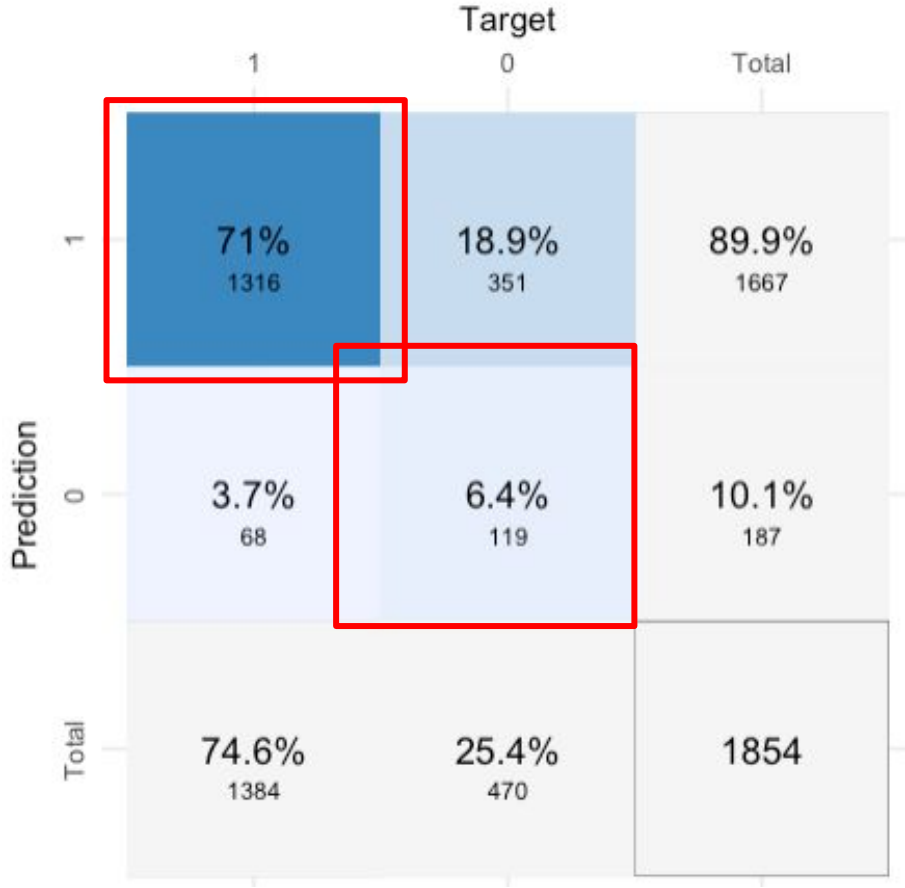




RNN —— other model we tried (3)

- Sequential RNN
- Accuracy: 77%

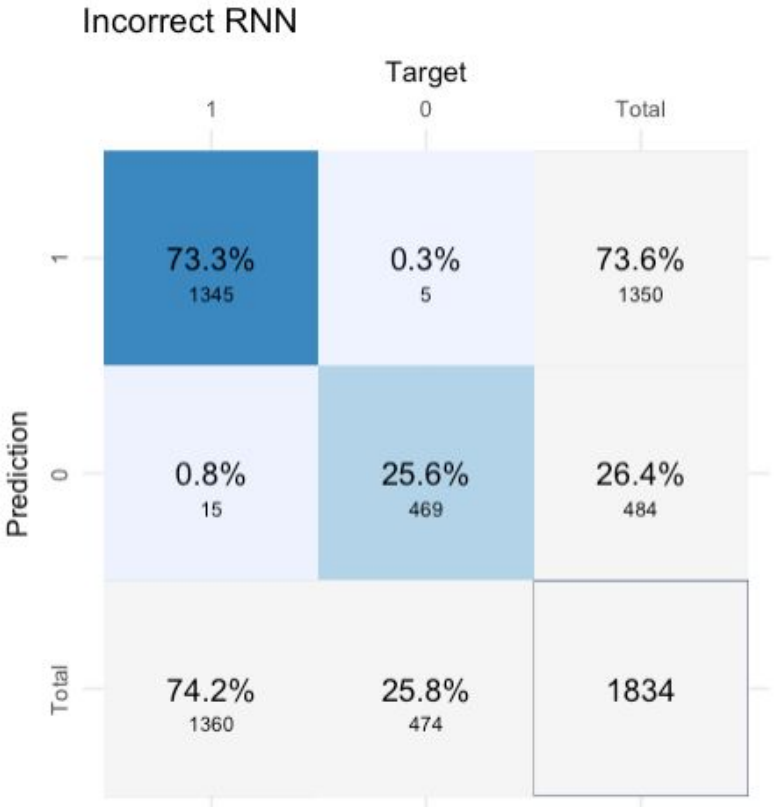


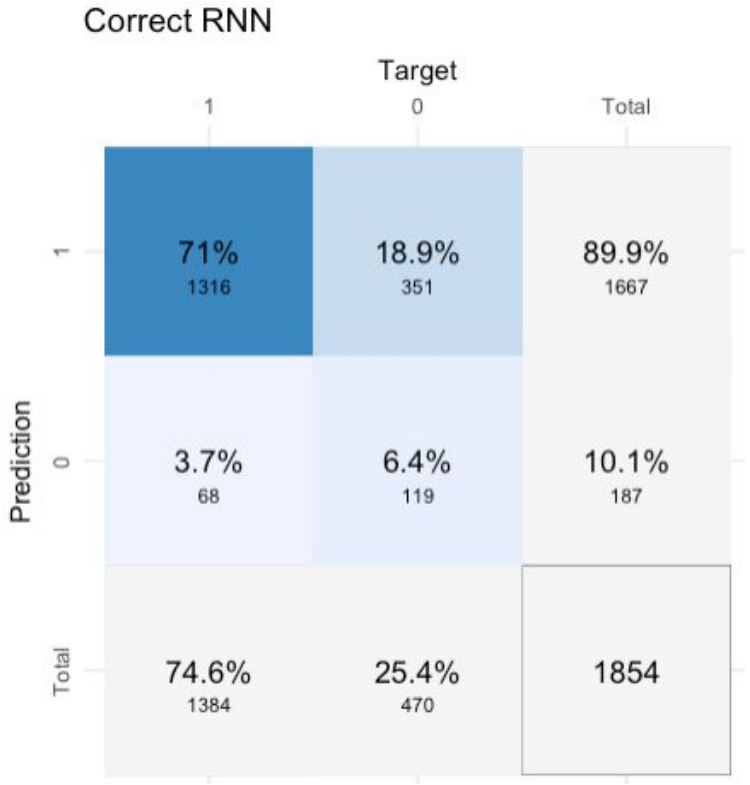


Thank you

Appendix

- Randomly Selected vs. Sequential RNN





Mouse Project

Final Results - Flipgrid Presentation

Rose, Jessie, Shuting

Agenda — what we did

- **Identify research subjects**

  Zero Maze with Mouse 409 (25 neurons and all neurons)

- **Apply different models**

  Logistic Model with PCA, Simple Neural Network, RNN, Bidirectional LSTM

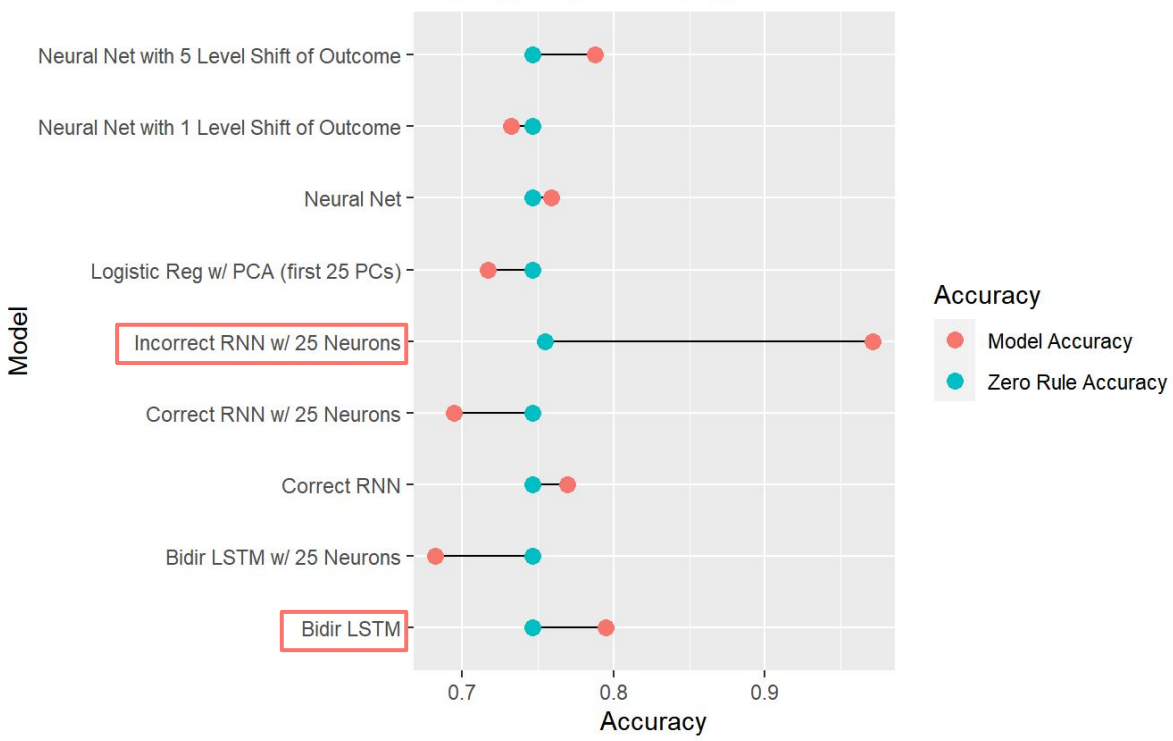
- **Detailed Updated**

  Train/test splitting matters

Model Comparisons with Mouse 409

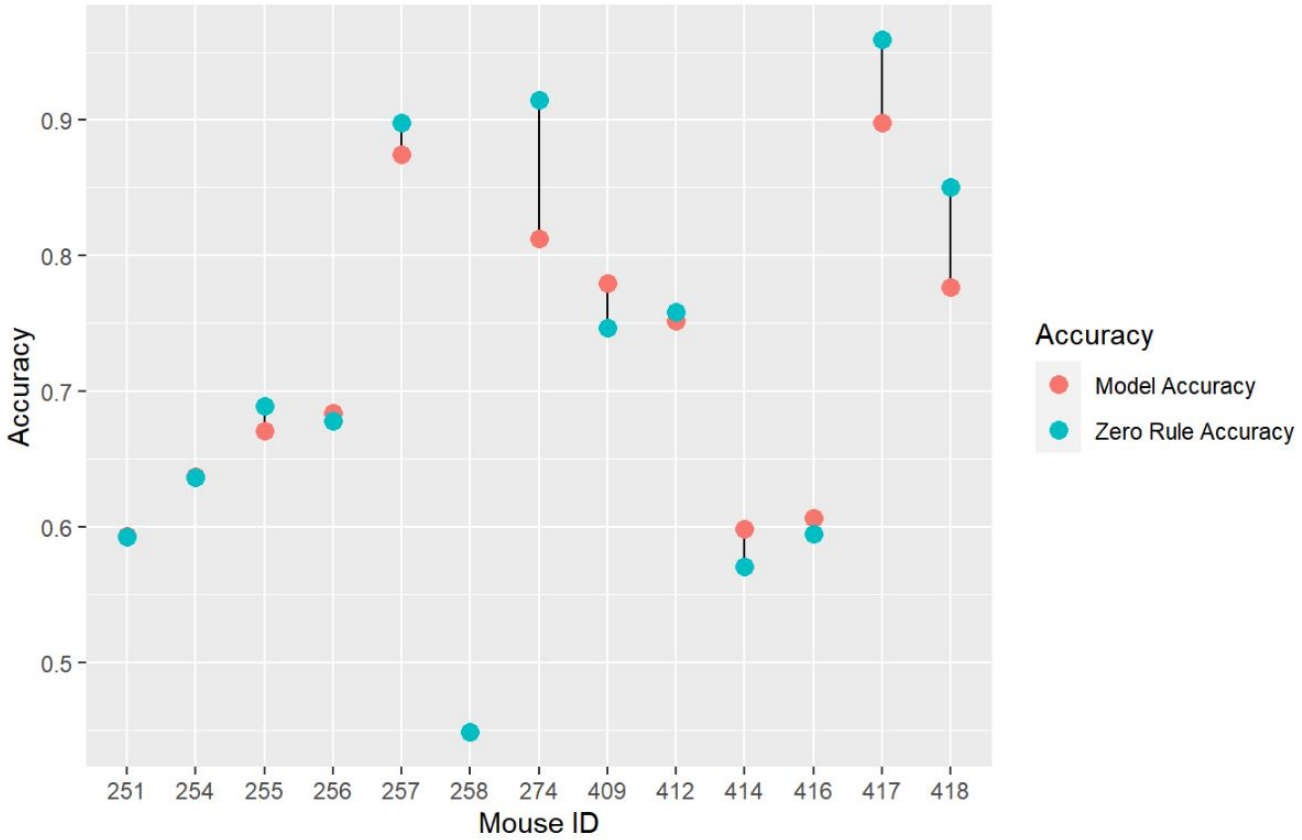
Zero Rule Accuracy =
Selecting the most common class

Goal =
Model Accuracy > Zero Rule Accuracy



Results — what we found

- Accuracy of Test Data with Bidirectional LSTM
- **All Mice w/ all neurons**



Conclusions

- Split train / test in **chunks**
- Take **all neurons** into consideration
- Try different **lags**

- **Key Questions:**

  Does a neuron carry the same information throughout the 10 minute trial?

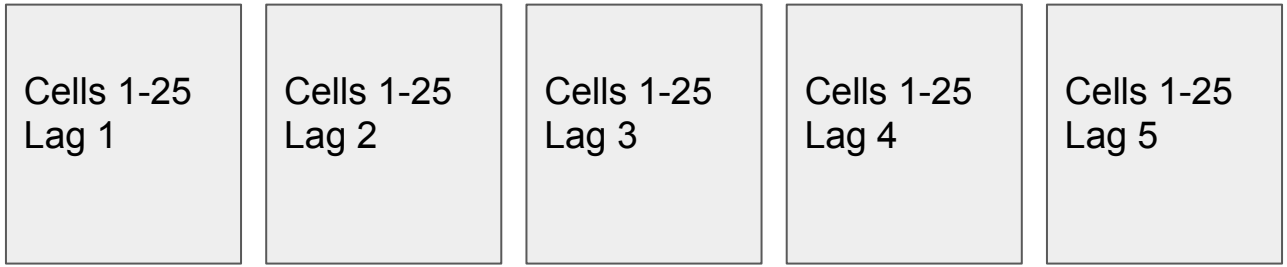
  Does behavior influence neural activity? OR Does neural activity influence behavior?

Mouse Project

Simple RNN(Based on Mouse 409 in Zero-Maze)

Rose, Jessie, Shuting

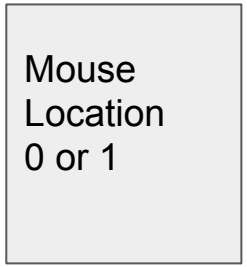
Input Data



Subsample of 25 neurons with 5 time lags.
Total Rows = 6179 (70% training; 30% testing)

Output Data

Total Rows = 6179



Model Details

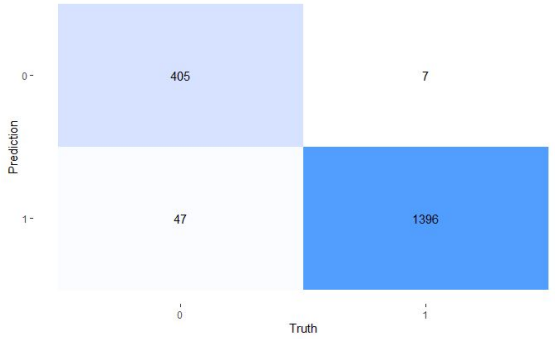
```
model <- keras_model_sequential () %>%
  layer_simple_rnn ( units = 100 ,
                     input_shape = list (n_lags , n_neurons) ,
                     dropout = 0.1 , recurrent_dropout = 0.1) %>%
  layer_dense ( units = 50, activation = "relu") %>%
  layer_dense ( units = 1, activation = "sigmoid")
model %>% compile ( optimizer = optimizer_rmsprop() ,
                    loss = "binary_crossentropy",
                    metrics = c("accuracy") )
```

- Three layers
- Activation Function: ReLu, Sigmoid
- Loss Function: Cross-entropy

Result

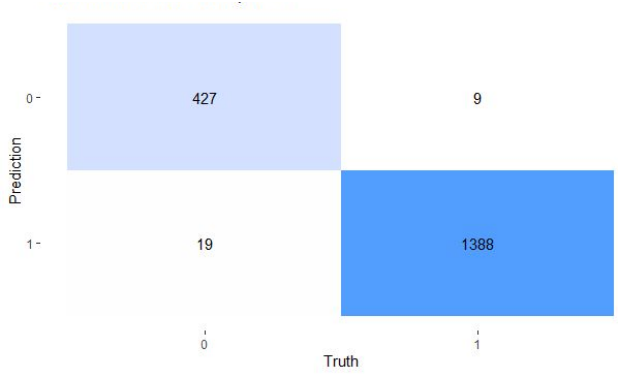
Epochs = 50

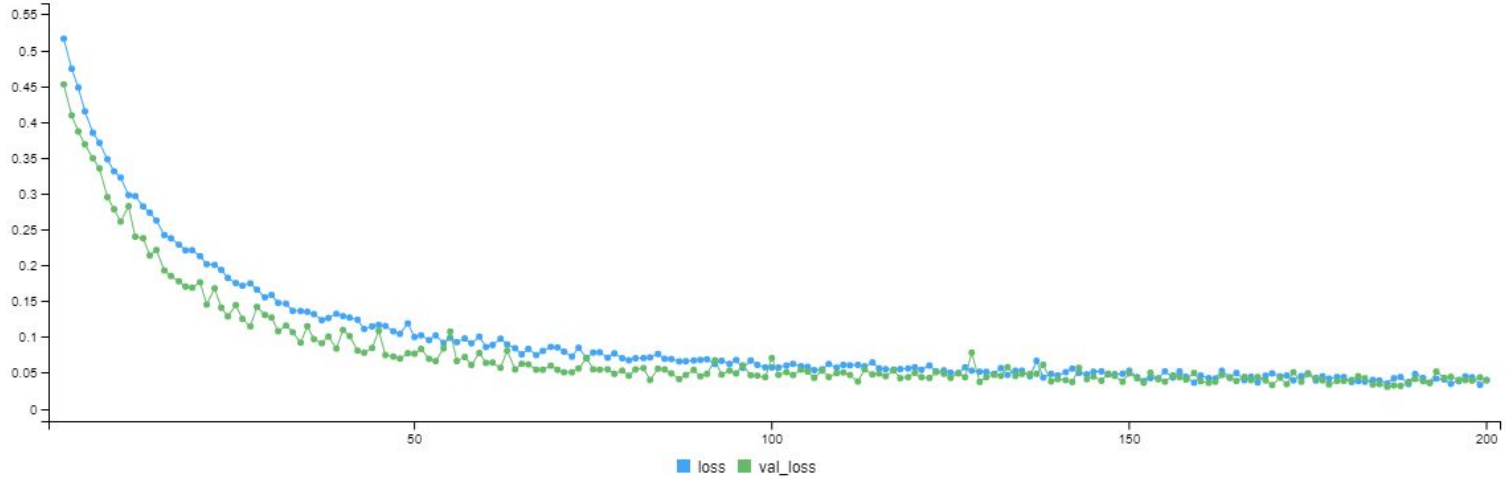
Accuracy: 0.962

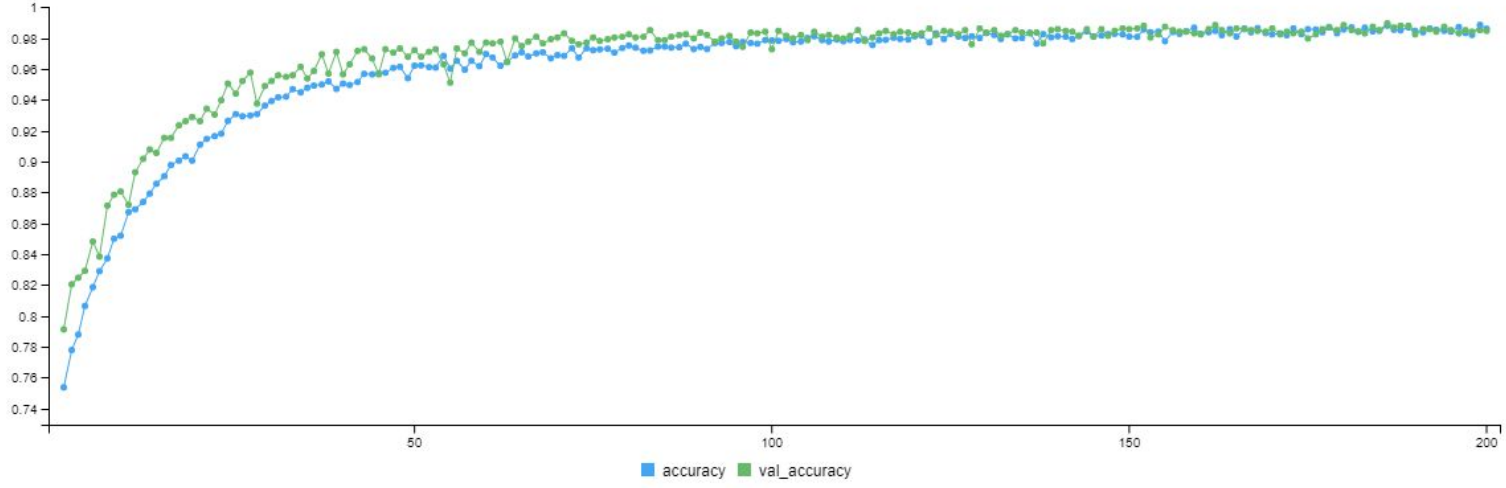


Epochs = 200

Accuracy: 0.985




0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
50
100
150
200
loss
val_loss


1
0.98
0.96
0.94
0.92
0.9
0.88
0.86
0.84
0.82
0.8
0.78
0.76
0.74
50
100
150
200
accuracy
val_accuracy

Next Steps:

- Shuffle the neurons for the RNN model to get a more general result.
  - Try neural network with PCA.
  - Try add some noise (more neurons).
- Try different “windows”: 10 seconds, 25 seconds, etc.
- Make ROC lines for every mouse and try the different number of components for PCA.
- Explore the correlated neurons according to the consistency of correlated and PCA plots. Are they distributed in a cluster or randomly (space or distant)?

Mouse Project

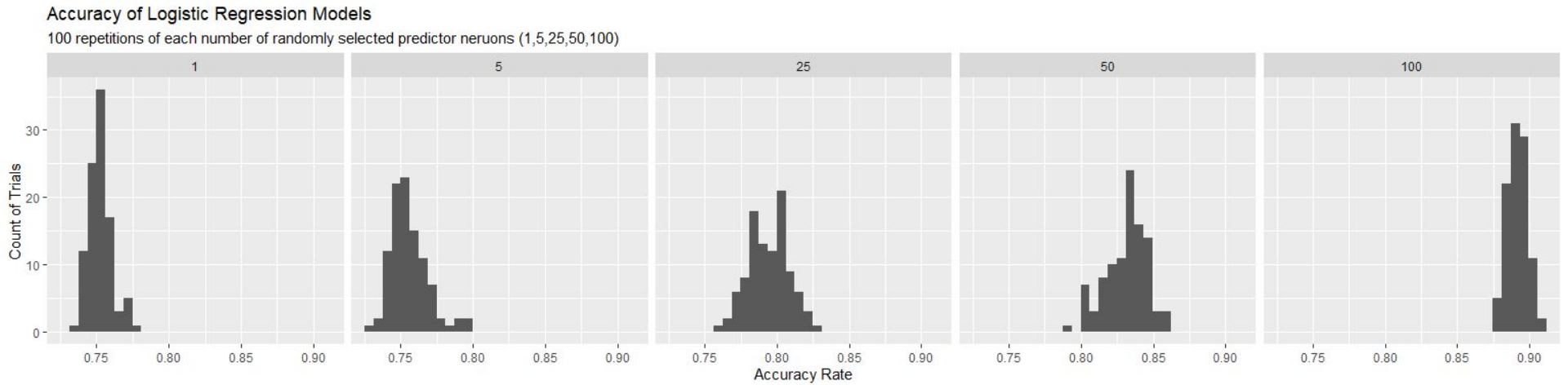
Baseline Model (Based on Mouse 409 in Zero-Maze)

Rose, Jessie, Shuting

Logistic Regression

"The **logistic regression** was trained and tested using an increasing interval of randomly selected cells with 1000 repetitions for each number of cells"

Johnson et al. (2022)



Logistic Regression Considerations

- Does not account for time and autocorrelation

- Correlation between neurons

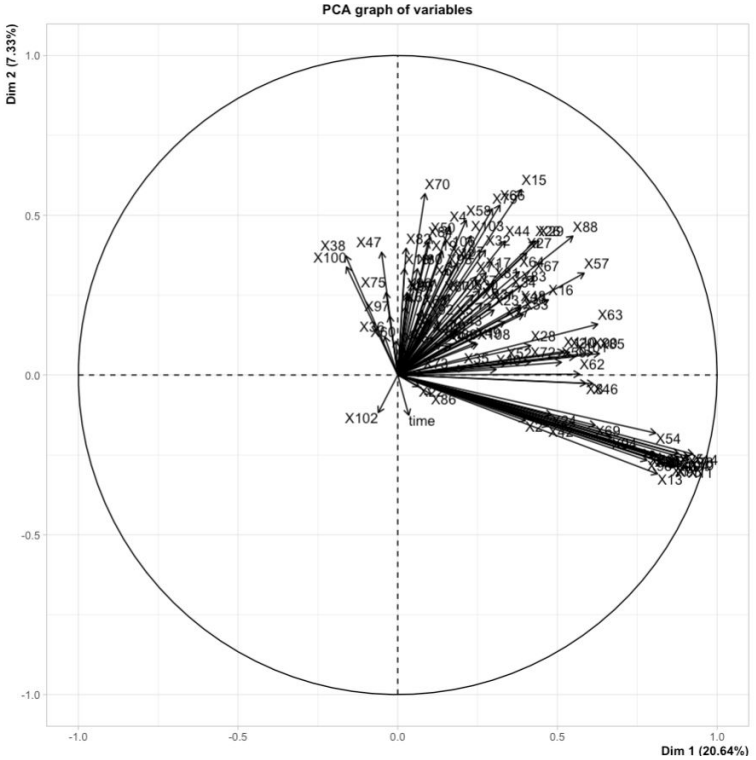
- Unbalanced data (In total: 1 = 4647 rows; 0 = 153)

	Reference	
Prediction	1	0
1	2137	207
0	144	606

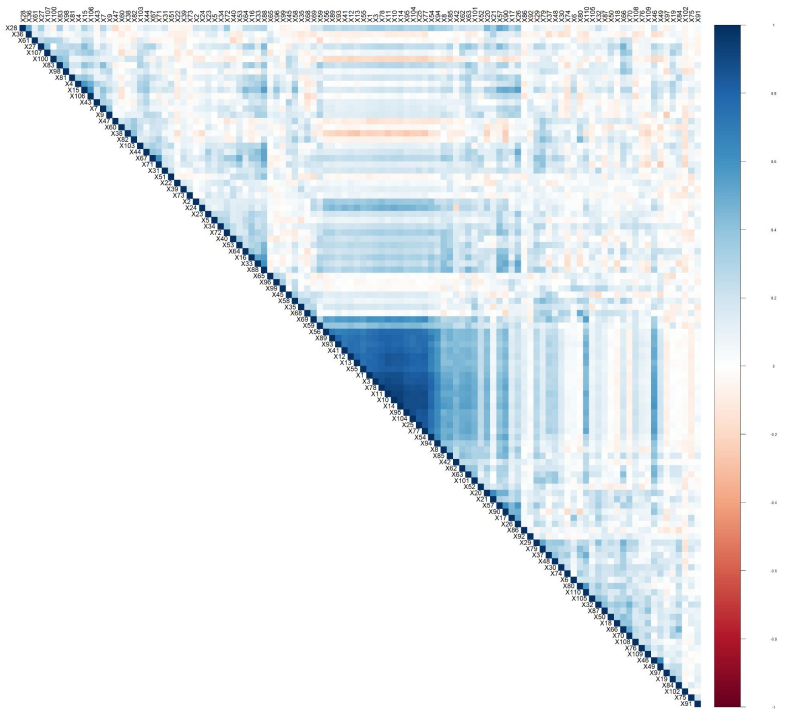
Accuracy : 0.8866

Deal with Correlation between Neurons (PCA)

PCA Plot



Correlation Plot



Logistic Regression and PCA

Eigenvalue & Contribution Format

	eigenvalue	variance.percent	cumulative.variance.percent
Dim.1	22.90938441	20.63908506	20.63909
Dim.2	8.13994788	7.33328637	27.97237
Dim.3	5.94709912	5.35774695	33.33012
Dim.4	4.75210760	4.28117802	37.61130
Dim.5	3.56853359	3.21489513	40.82619
Dim.6	2.89839073	2.61116282	43.43735
Dim.7	2.63542867	2.37426007	45.81161
Dim.8	2.45687416	2.21340015	48.02501
Dim.9	2.03942374	1.83731868	49.86233
Dim.10	1.87910008	1.69288296	51.55522
Dim.11	1.84050439	1.65811207	53.21333
Dim.12	1.71331261	1.54352488	54.75685
Dim.13	1.59762885	1.43930527	56.19616
Dim.14	1.50689683	1.35756472	57.55372

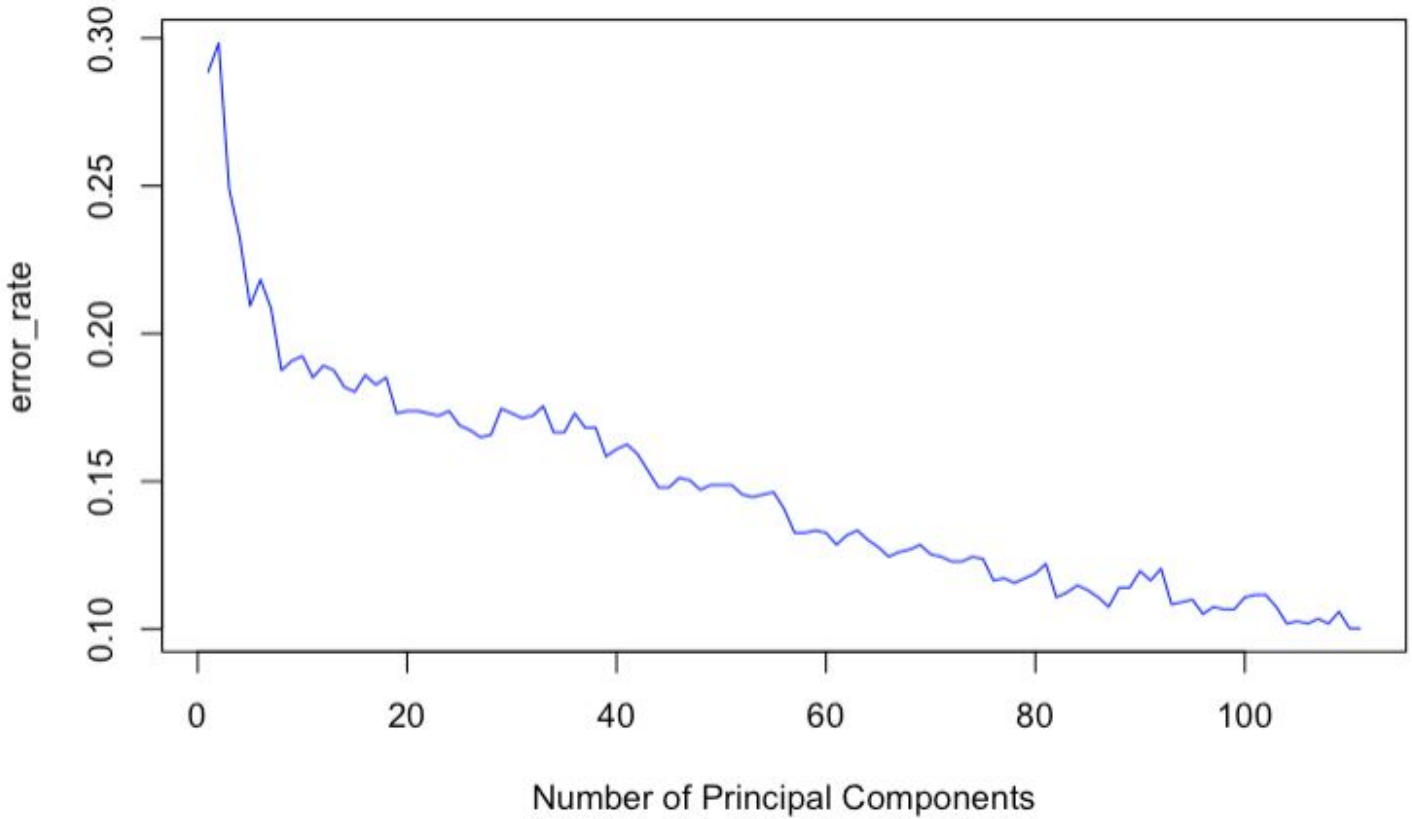
Result when we choose first 40 principal components:

```
Call:
glm(formula = open ~ ., family = binomial("logit"), data = training)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-3.07276  -0.06894   0.24742   0.52941   2.66530

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)  1.951235   0.062801  31.070  < 2e-16 ***
Dim.1       -0.031187   0.010739  -2.904 0.003683 **
Dim.2        0.046791   0.018421   2.540 0.011083 *
Dim.3       -0.611384   0.027581 -22.167  < 2e-16 ***
Dim.4        0.324319   0.024087  13.465  < 2e-16 ***
Dim.5        0.408173   0.029219  13.969  < 2e-16 ***
Dim.6       -0.133940   0.033148  -4.041 5.33e-05 ***
Dim.7        0.152302   0.033419   4.557 5.18e-06 ***
Dim.8       -0.423454   0.035565 -11.907  < 2e-16 ***
Dim.9       -0.006772   0.033604  -0.202 0.840291
Dim.10       0.291025   0.036563   7.959 1.73e-15 ***
```

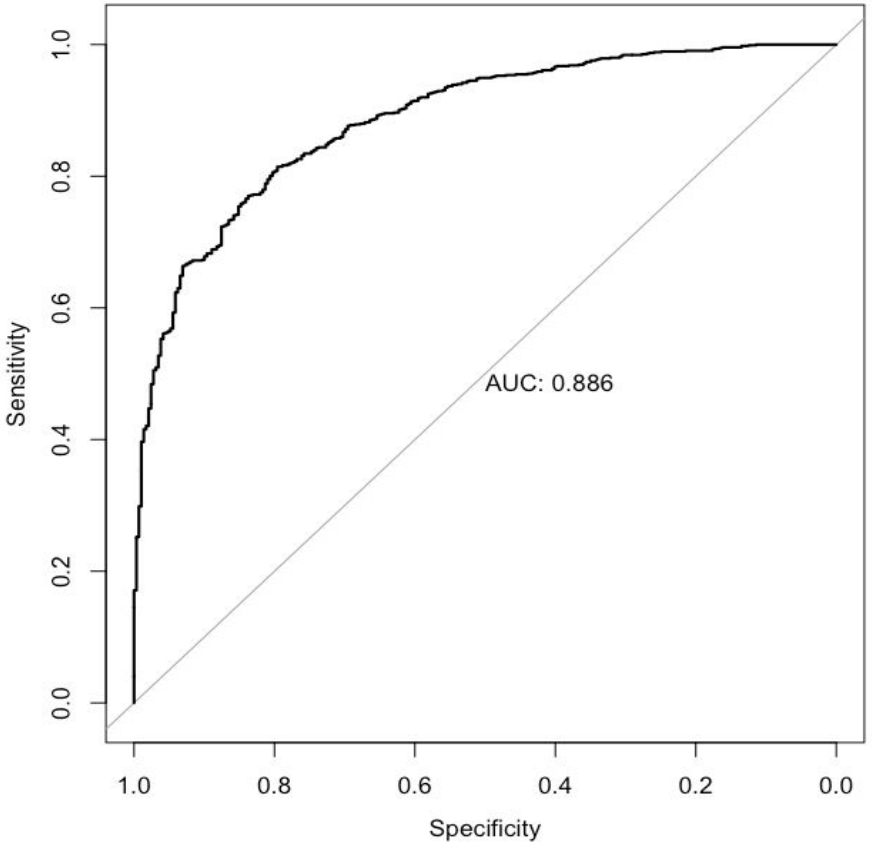
```
error_rate
```

```
[1] 0.1657235
```

Error Rate VS. Number of PC



Deal with Unbalanced Data (ROC Curve)



- The AUC of our logistic model (number of components = 40) is 0.886, which means this classifier performs good even in unbalanced data

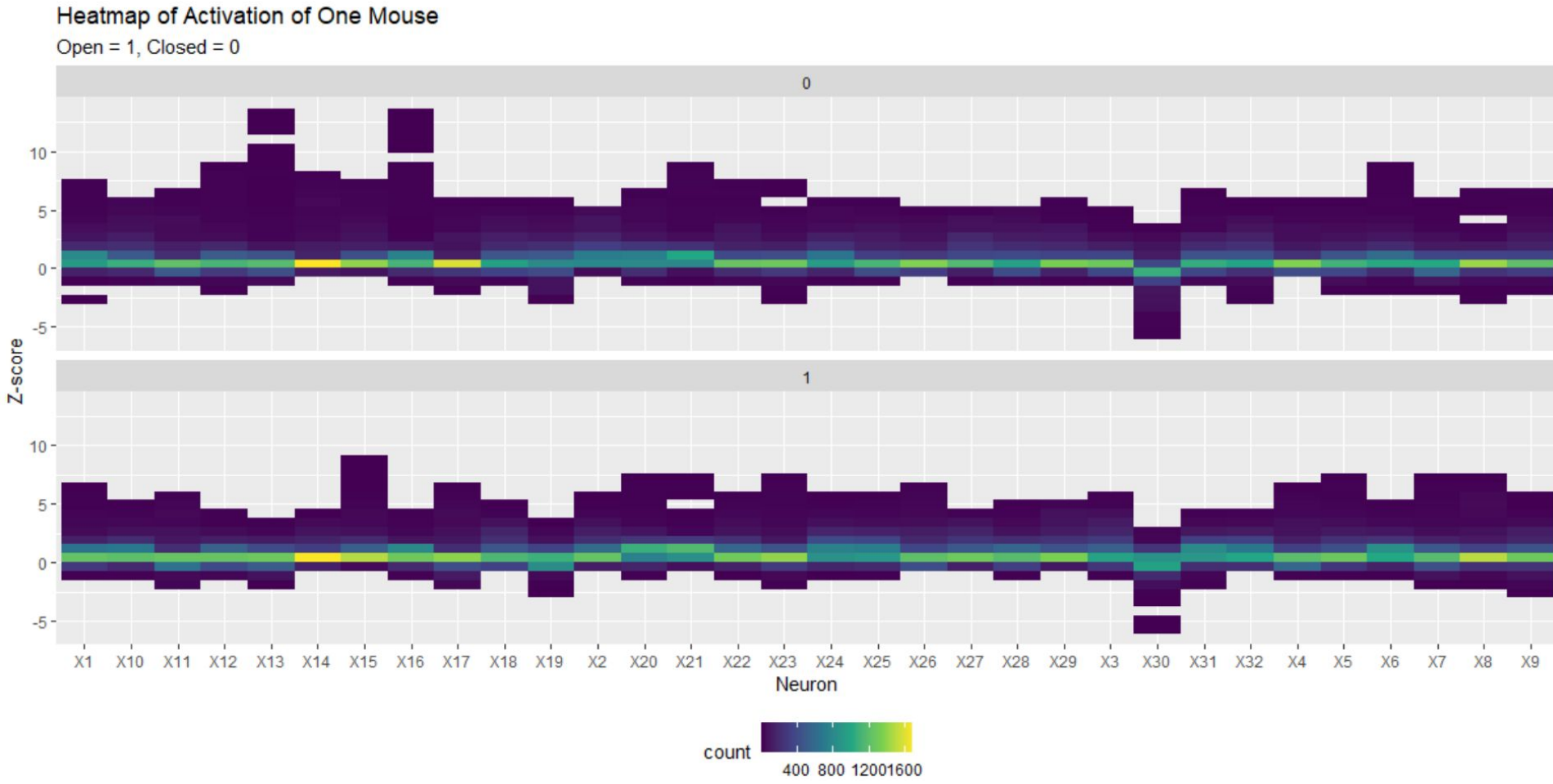
Mouse Project

EDA

Rose, Jessie, Shuting

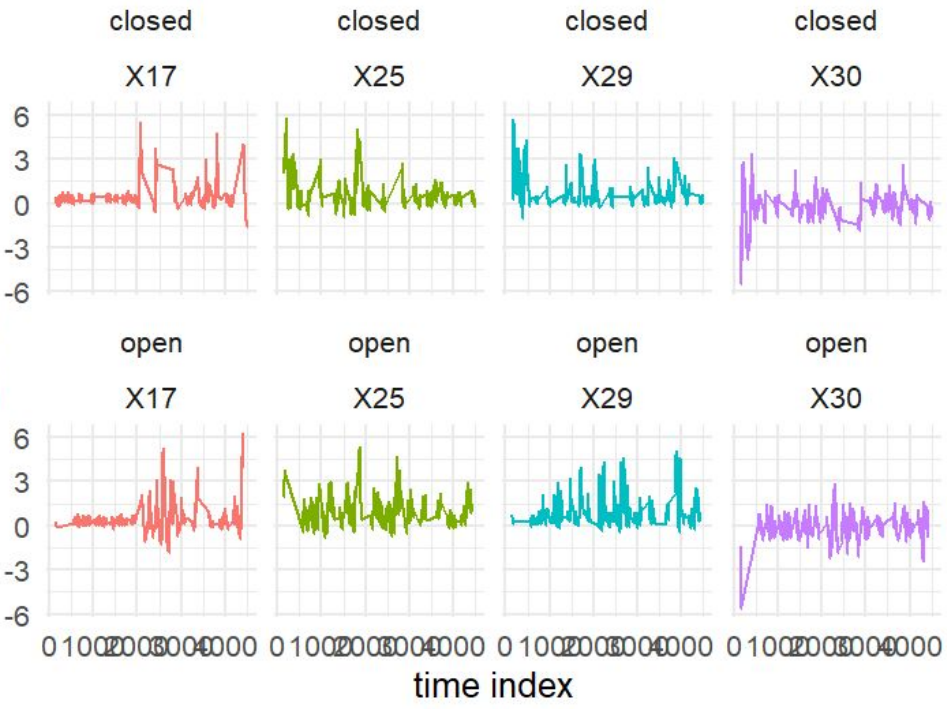
Background

- Objective: decoding neural performance to analyse mice’s social behavior
- Method: use machine learning skills to predict mouses’ behavior based on their neural recordings
- This Week: use mouse 251 in Zero_Maze experiment as example, to do exploratory data analysis
- Next Steps: Continue EDA. Understand “Machine Learning for Neural Decoding” paper. Consider potential models.

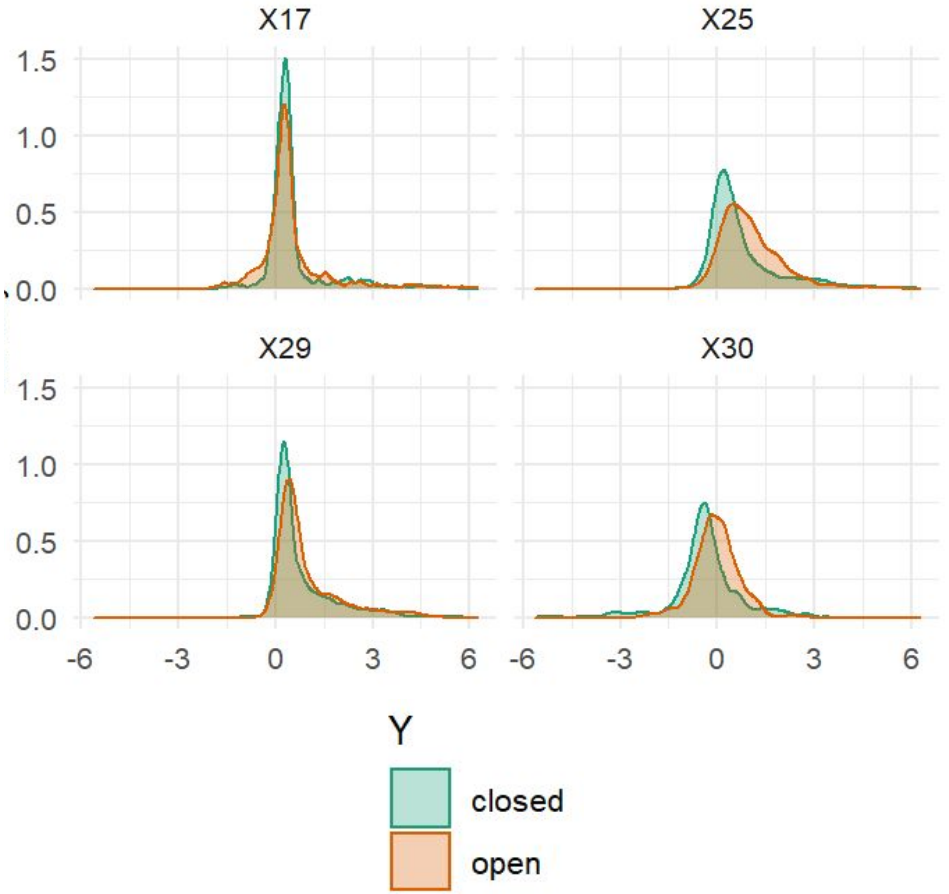

Heatmap of Activation of One Mouse
Open = 1, Closed = 0
0
1
Z-score
10
5
0
-5
X1 X10 X11 X12 X13 X14 X15 X16 X17 X18 X19 X2 X20 X21 X22 X23 X24 X25 X26 X27 X28 X29 X3 X30 X31 X32 X4 X5 X6 X7 X8 X9
Neuron
count
400 800 12001600

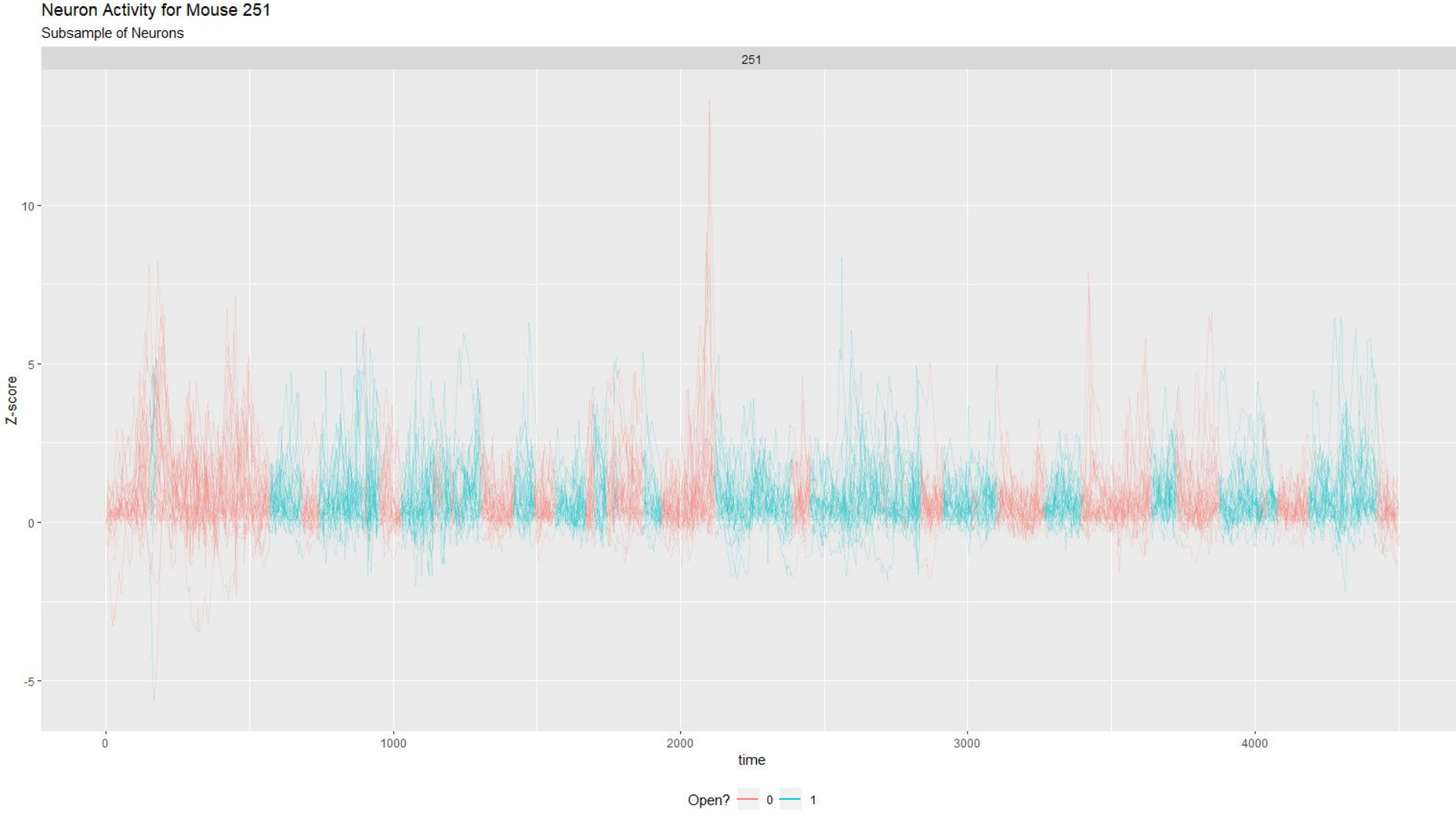

mouse 251 in Zero_Maze experiment
average zscore
1.0
0.5
0.0
X1 X10 X11 X12 X13 X14 X15 X16 X17 X18 X19 X2 X20 X21 X22 X23 X24 X25 X26 X27 X28 X29 X3 X30 X31 X32 X4 X5 X6 X7 X8 X9
neurons
behavior
closed
open

Time series plot of z-score



Density plot of z-score




Neuron Activity for Mouse 251
Subsample of Neurons
251
10
5
0
-5
Z-score
0
1000
2000
3000
4000
time
Open?
0
1