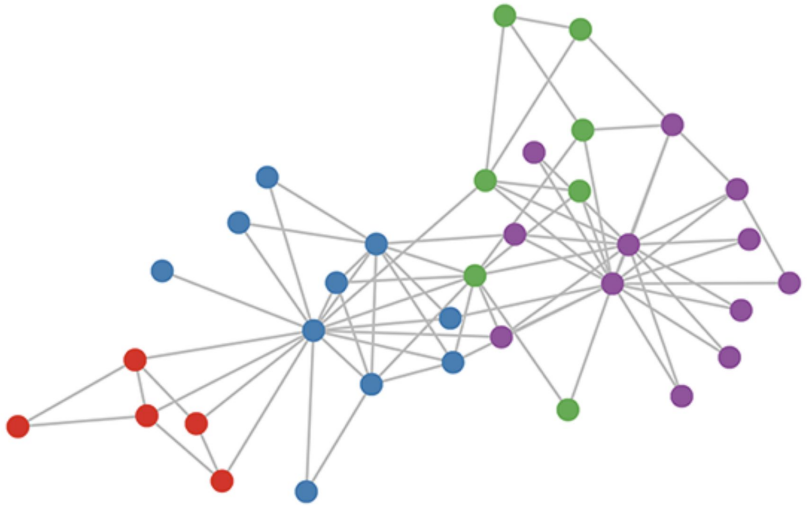


“Connectomics”
Final Presentation

Machine Learning,
Sackler Institute of Graduate Biomedical Sciences,
Fall 2019

Presented by David Coit

Network Reconstruction

```
11,1,1
15,1,1
20,1,1
27,1,-1
28,1,1
37,1,-1
47,1,1
55,1,-1
60,1,1
71,1,1
13,2,1
16,2,1
20,2,1
26,2,1
39,2,1
48,2,1
55,2,-1
70,2,1
77,2,1
83,2,1
88,2,1
8,3,1
9,3,1
22,3,1
52,3,-1
58,3,1
```

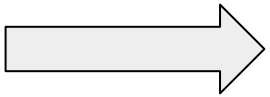
Connections

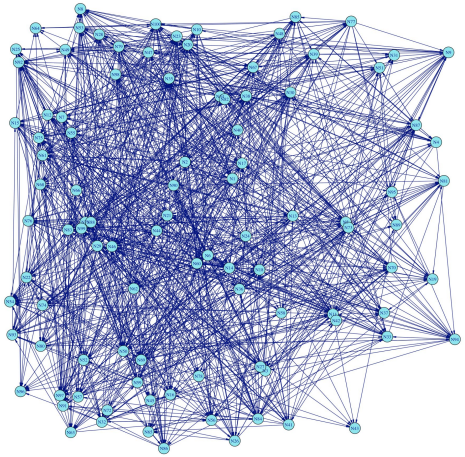
```
0.488,0.619
0.384,0.656
0.560,0.191
0.934,0.700
0.736,0.522
0.434,0.448
0.114,0.756
0.155,0.997
0.961,0.900
0.410,0.951
0.708,0.317
0.083,0.761
0.620,0.536
0.481,0.420
0.012,0.743
0.352,0.138
0.509,0.653
0.547,0.416
0.926,0.395
0.390,0.917
0.345,0.537
0.037,0.399
```

Mapping

```
0.023, 0.087, 0.008, -0.039, 0.043,
0.063, 0.001, 0.047, -0.026, 0.034,
0.056, -0.025, 0.048, 0.049, 0.069,
0.049, 0.027, 0.098, 0.006, 0.057,
-0.009, -0.035, 0.049, -0.048, 0.17
0.080, 0.008, -0.030, 0.146, -0.008
0.095, 0.035, -0.008, 0.069, 0.088,
0.067, -0.006, 0.060, 0.078, -0.010
0.055, -0.049, 0.050, 0.020
0.093, 0.195, -0.011, -0.006, 0.066
0.027, -0.015, 0.059, -0.006, -0.00
0.012, -0.012, 0.061, 0.026, 0.081,
```

Time Series Fluorescence Data





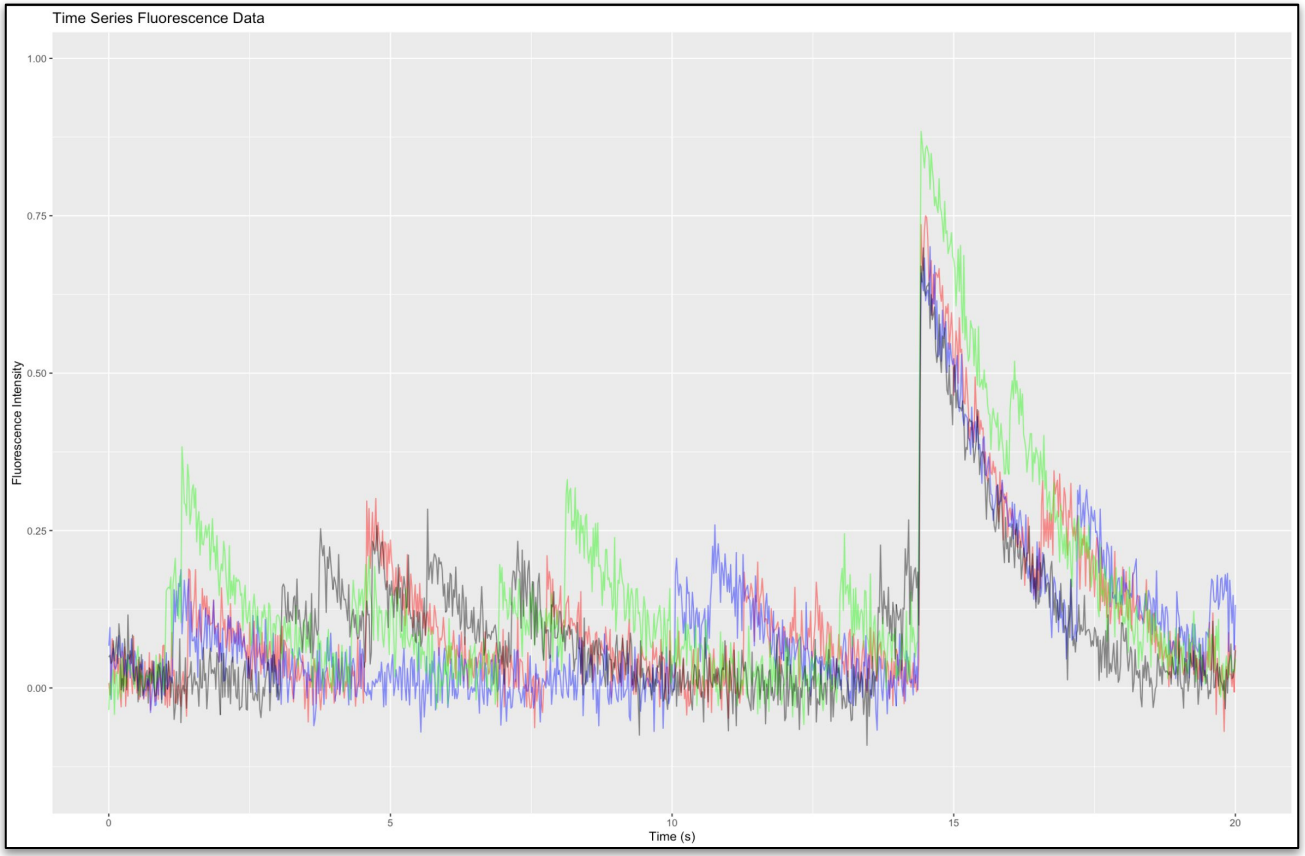
Connected Network

Approaches

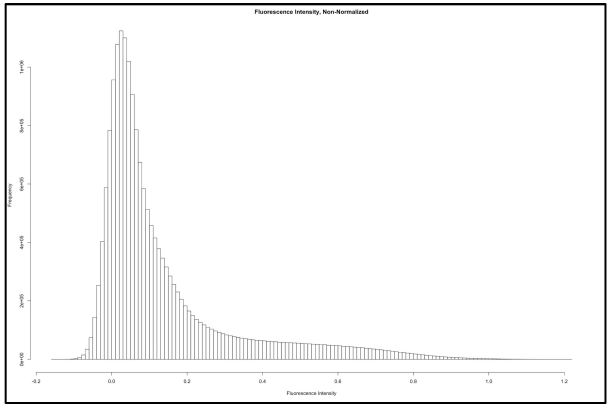
- **Signal statistics (marginal “success”)**
  - Filtering / processing (high pass, low pass, threshold, spike detection)
  - Partial Correlation
  - Mutual Information
  - Transfer Entropy*
  - Support Vector Machine and Random Forest on results

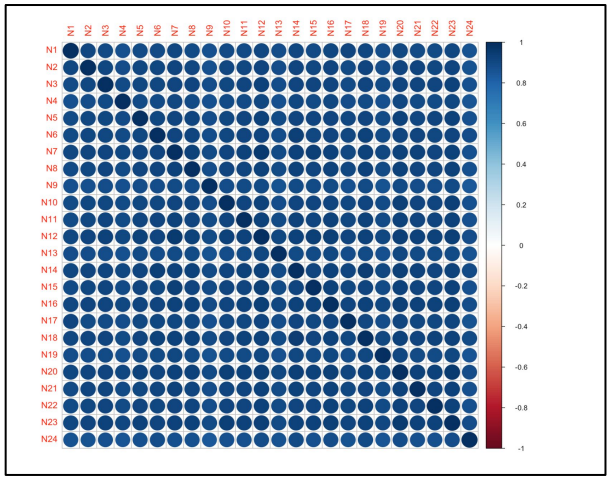
- **Deep learning (utter failure)**
  - Dense ANNs
  - Adapt existing convolutional models

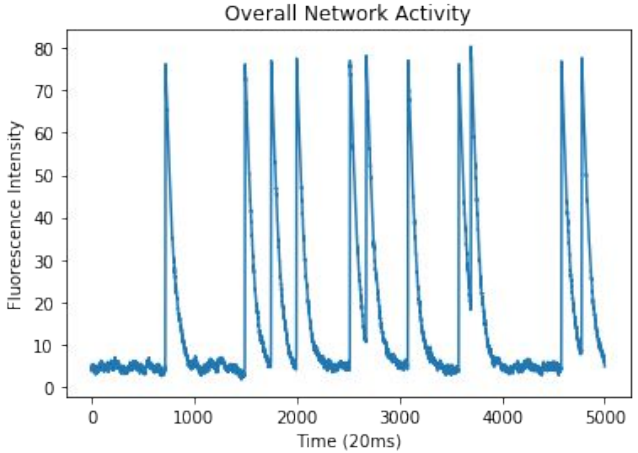
Signal Statistics

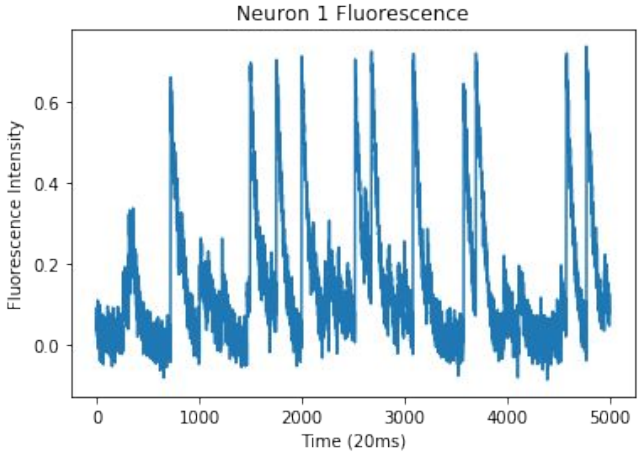


Signal Statistics - Correlation and Network Behavior

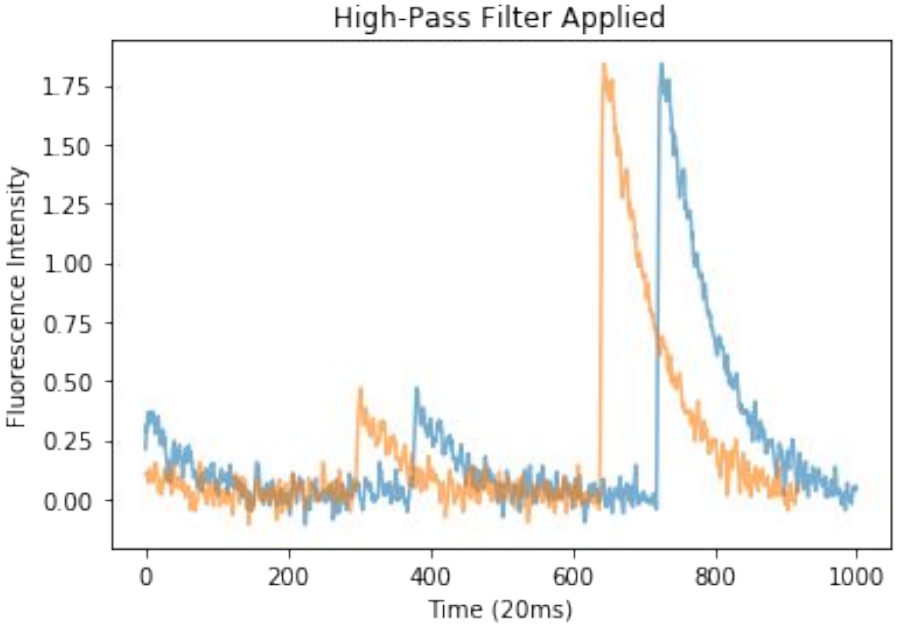


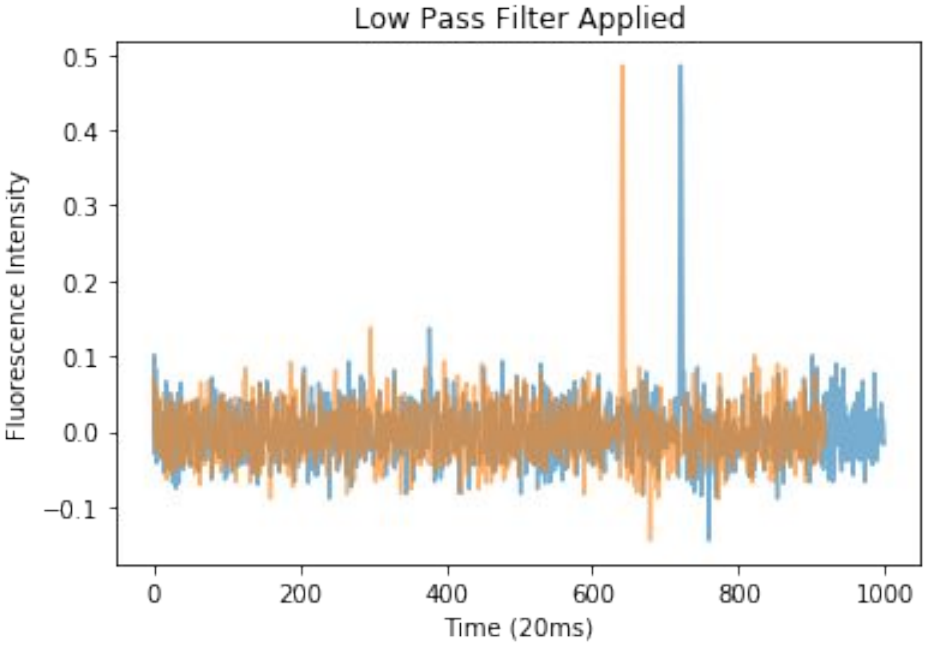






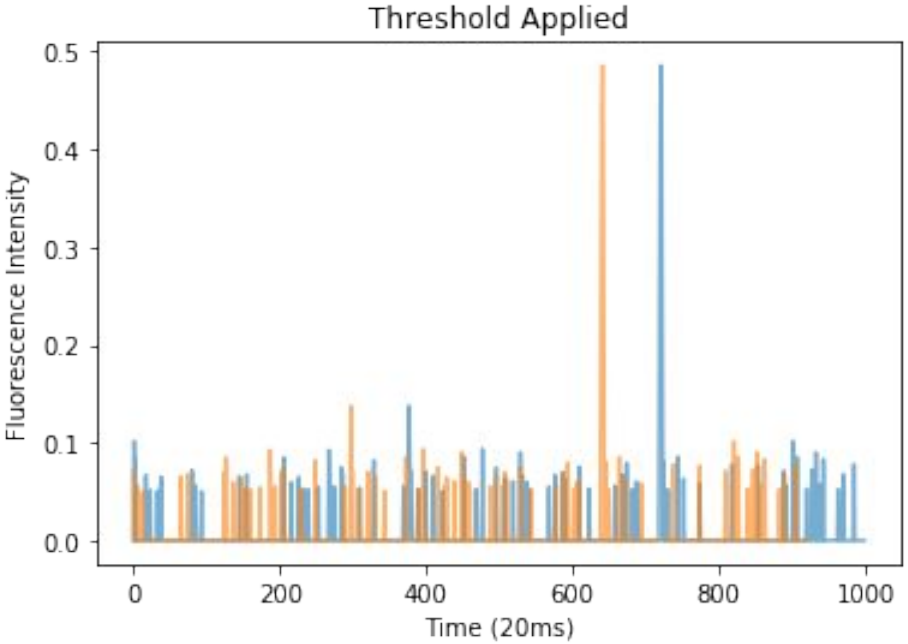
Signal Statistics - Filtering / Processing

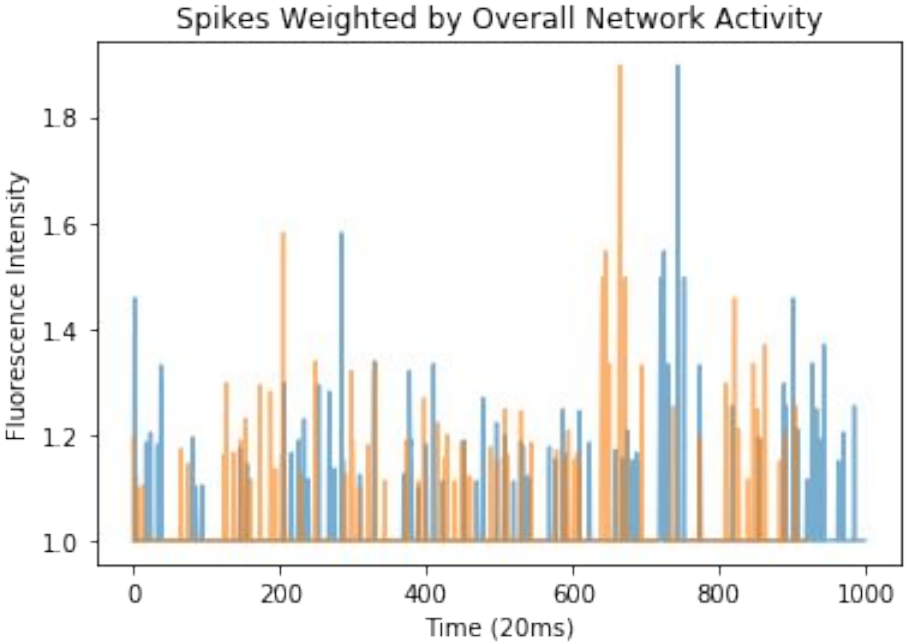




recreated from Sutera et al 2015

Signal Statistics - Filtering / Processing





recreated from Sutera et al 2015

Signal Statistics

- Partial-correlation
- Conditional mutual information
- Granger causality
- Transfer entropy*
-
- See if we can build a set of features based on time series behavior...

Signal Statistics - Data Frame

Neuron_Pair	Dataset	Weighted_Spike_Partial_Correlation	Mutual_Information_Conditional	GrangerF	ConnectionPresent	ConnectionType	Distance
01-000-001	1	0.075237	0.000289	8.260917	absent	none	0.110386
01-000-002	1	0.004265	0.000296	5.346523	absent	none	0.434014
01-000-003	1	0.002911	0.000235	1.806921	absent	none	0.453296
01-000-004	1	0.003555	0.000319	0.870090	absent	none	0.266295
01-000-005	1	0.013011	0.000305	4.199372	absent	none	0.179324
...	...	...	...	...	...	...	...
01-000-096	1	0.009193	0.000251	0.603130	present	excitatory	0.613329
01-000-097	1	0.004091	0.000275	0.724274	absent	none	0.345485
01-000-098	1	-0.002298	0.000179	0.582041	absent	none	0.349116
01-000-099	1	0.003481	0.000293	1.222376	present	excitatory	0.382832
01-001-000	1	0.075237	0.000289	7.141375	absent	none	0.110386

Signal Statistics Random Forest - Connection Detection

AUC = 0.518

	Connection	No Connection
Connection	39	279
No Connection	205	1977

	precision	recall	f1-score	support
No Connection	0.88	0.91	0.89	2182
Connection	0.16	0.12	0.14	318
accuracy			0.81	2500
macro avg	0.52	0.51	0.51	2500
weighted avg	0.79	0.81	0.80	2500

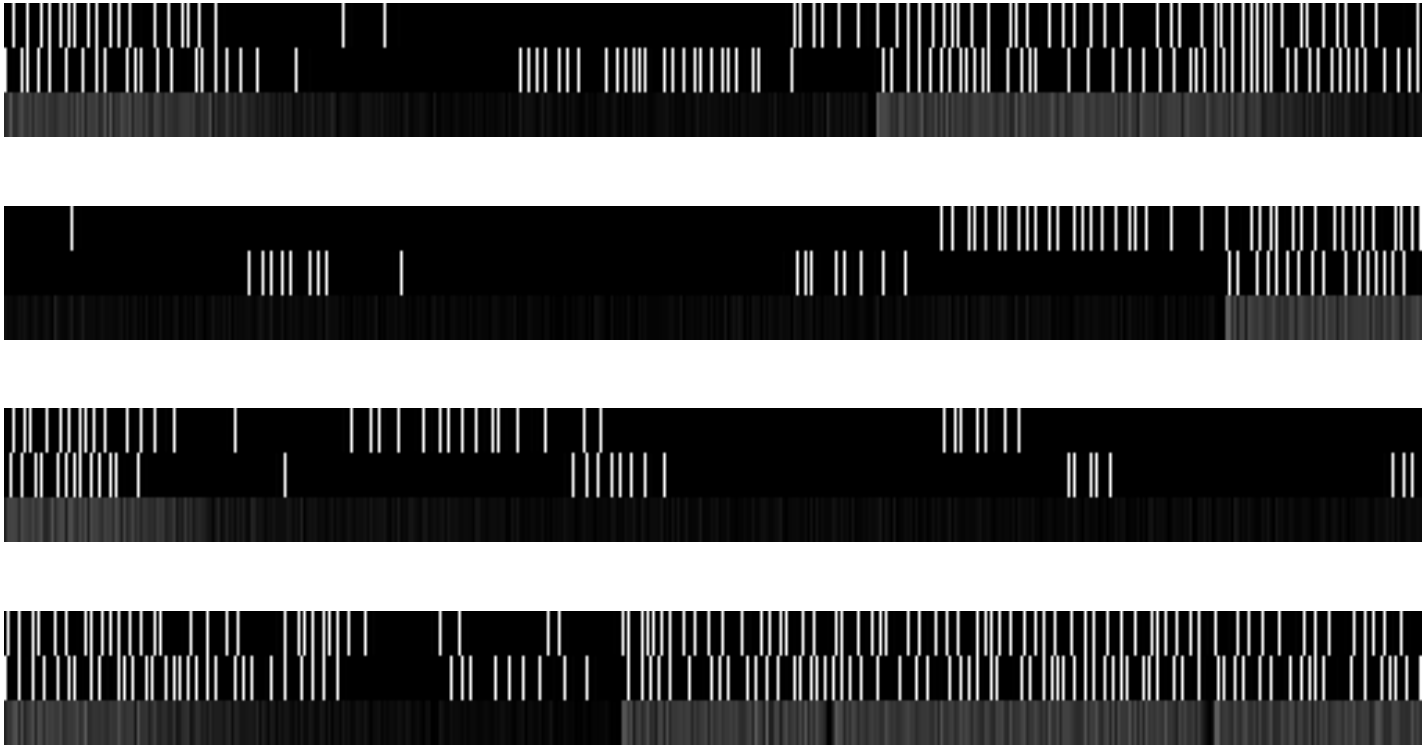
Signal Statistics Random Forest - Connection Type*

AUC = 0.567

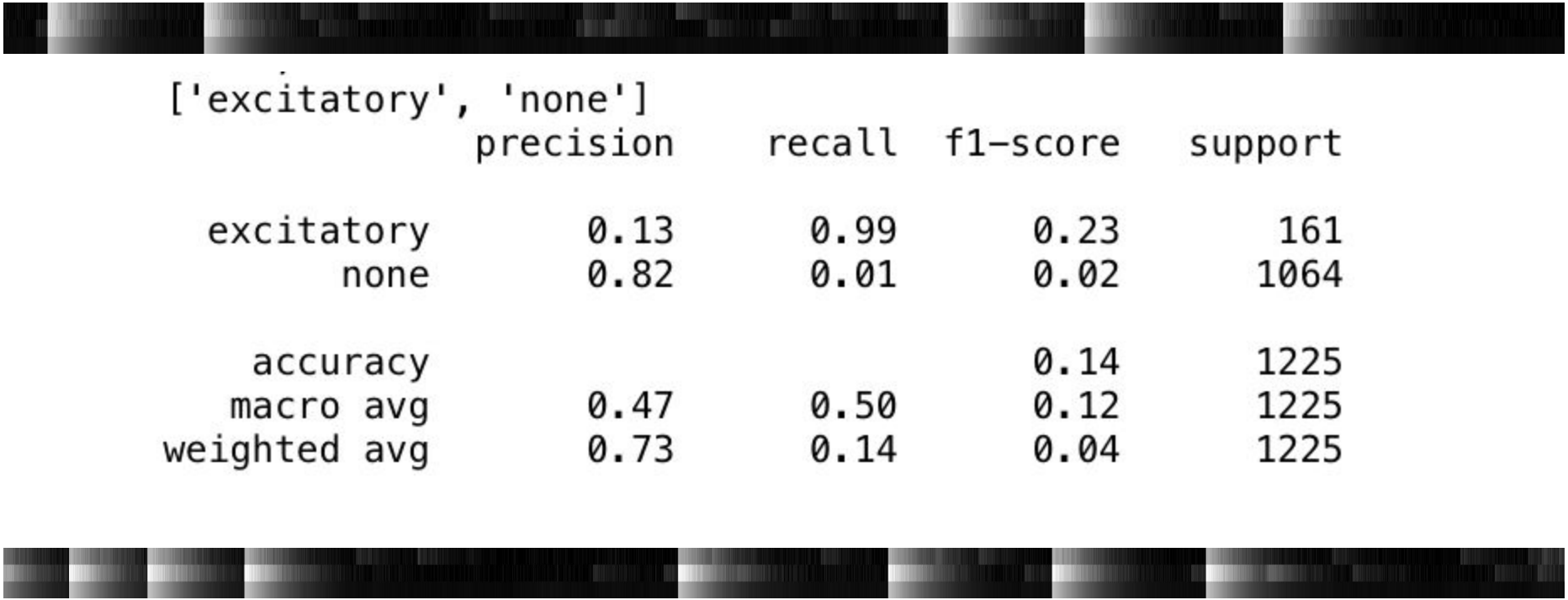
	Inhibitory	Excitatory
Inhibitory	12	52
Excitatory	13	236

	precision	recall	f1-score	support
Inhibitory	0.48	0.19	0.27	64
Excitatory	0.82	0.95	0.88	249
accuracy			0.79	313
macro avg	0.65	0.57	0.57	313
weighted avg	0.75	0.79	0.75	313

Deep Learning with Keras



Deep Learning with Keras



```
['excitatory', 'none']
              precision    recall  f1-score   support

  excitatory       0.13      0.99      0.23       161
        none       0.82      0.01      0.02      1064

    accuracy                           0.14      1225
   macro avg       0.47      0.50      0.12      1225
weighted avg       0.73      0.14      0.04      1225
```

Next Steps

- **Apply model to data examples with more neurons / different parameters**

- **Accelerate / improve cross-signal statistic calculation for further exploration**
  - CPU / GPU parallelization
  - Mutual information calculation is current rate-limiting step in RF approach
  - Transfer entropy R implementation takes > 1hr per neuron pair (9,900 in “small” examples)
  - Find a good estimator for these parameters (confidence interval?)
  - Search over signal filtering / spike detection parameters

- **Deep learning improvements - still hopeful that this can be a useful approach**
  - LSTM is a natural choice to explore, encountered PEBCAK errors during attempts

Thank You!
Questions / Comments?

References:

Sutera, A., Joly, A., François-Lavet, V., Qiu, Z. A., Louppe, G., & Ernst, D. (2014). Simple connectome inference from partial correlation statistics in calcium imaging. In *ECML workshop - "Neural connectomics: From imaging to connectivity"*. Nancy, France.

I. M. de Abril, J. Yoshimoto, and K. Doya, "Connectivity inference from neural recording data:Challenges, mathematical bases and research directions,"Neural Networks, 2018.

Title slide graphic source: https://tkipf.github.io/graph-convolutional-networks/