Deep Learning - Stringer

Julia Suzuki, Golnaz Baghdadi, Arman Charan, Amit Kumar
Megapod: Cotopaxi / Pod: Othnielia Zydeco

neuromatch
academy

Content

1. Introduction
2. Dataset
3. Deep Learning Model Process
4. Analysis Methods & Results
5. Summary & Future Direction

Introduction

1

Conducted **LITERATURE REVIEW** to understand the previous research conducted using the Stringer dataset

RESEARCH ARTICLE SUMMARY

NEUROSCIENCE

Spontaneous behaviors drive multidimensional, brainwide activity

Carsen Stringer*†, Marius Pachitariu*†, Nicholas Steinmetz, Charu Bai Reddy, Matteo Carandini‡, Kenneth D. Harris‡‡

Can we predict spontaneous behavior (e.g., running speed) of a mouse from its neural activity?

2

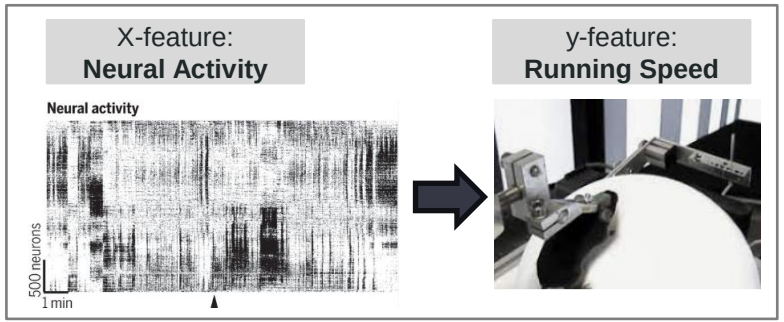
Defined a QUESTION

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_r x_r + \varepsilon$$

X-feature: Neural Activity
y-feature: Running Speed

3

Defined a HYPOTHESIS



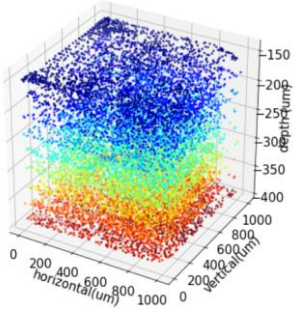
Stringer Dataset

Numpy dataset

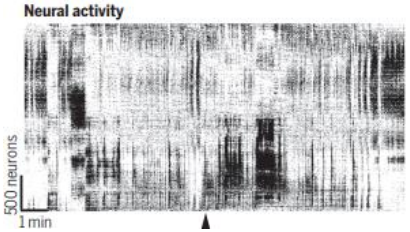
8 Features

```
dat = np.load('stringer_spontaneous.npy', allow_pickle=True).item()
print(dat.keys())
```

```
dict_keys(['sresp', 'run', 'beh_svd_time', 'beh_svd_mask', 'stat', 'pupilArea', 'pupilCOM', 'xyz'])
```



3-dimensional view
of approx. **12,000 Neurons**
Data[‘xyz’]



X-feature: Neural Activity
Data[‘sresp’]

y-feature: Running Speed
Data[‘run’]

Stringer dataset: https://github.com/NeuromatchAcademy/course-content/blob/main/projects/neurons/load_stringer_spontaneous.ipynb

Deep Learning Model Process

Preprocessing

- Running Speed
  shape: (7018, 1)
- Neural Activities
  shape : (11983, 7018)

- Dimension reduction:
- Principal Component Analysis
- Number of PCs: 100

- Neural Activities
  new shape : (100, 7018)

- Normalization: Z-score

Model Architecture

1. Fully-connected Multilayer Models:
   a. one ReLU layer
      i. 10 : 400 step=50
   b. two ReLU layers
      i. (400 , 10)
      ii. (400, 100)
   c. three ReLU layers
      i. (400, 10, 10)
   d. one Sigmoid layer
   e. two Sigmoid layers
   f. three Sigmoid layers

Model Training

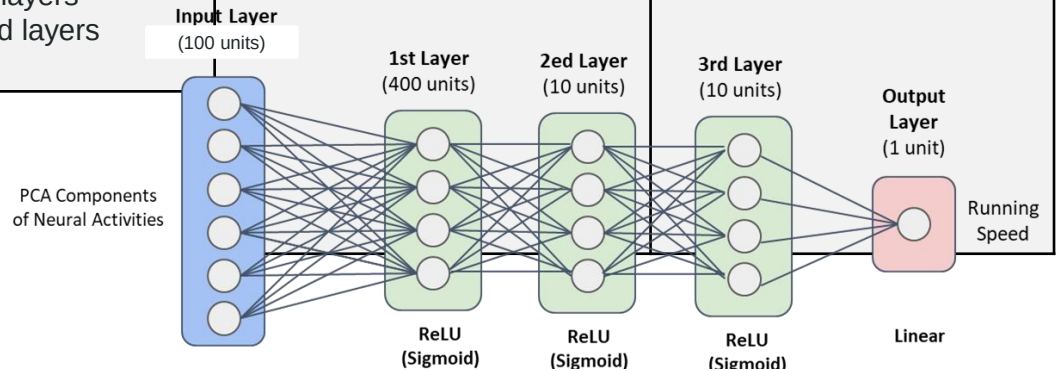
- Train/Test split:
  Train: 70%
  Test: 30%

(selected by trial and error)

- Loss function:
  mean square error (MSE)
- Optimizer: ADAM
- Number of epochs: 100
- Batch size: 10
- Library: Pytorch

Evaluation Metric

- RMSE



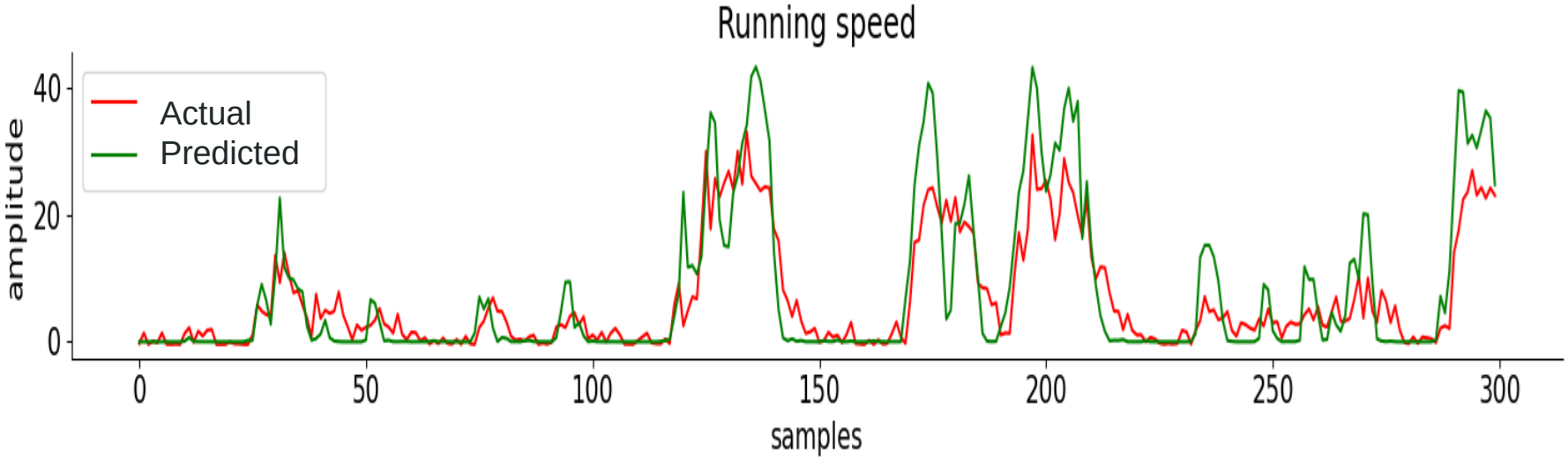
Analysis Methods & Results

Methods & Techniques:

- Deep Linear Neural Network Models
- Conventional Principal Component Analysis (PCA)
- Activation Functions (ReLU, Sigmoid, Tanh)

				Model Architecture		Run Configuration	Results	
	Input	**Output**	**PCA**	**Activation Function**	**No. Layers**	**Running Parameters**	**Train**	**Test**
1	Neural Activity	Running Speed	500	**ReLU**	3 Layers (Layer 1: 400 Layer 2: 10 Layer 3: 10)	No. Epochs: 100 No. Batch: 10	MSE: 59.68 RMSE: 7.73	MSE: **131.31** RMSE: **11.46**
2	Neural Activity	Running Speed	500	**Sigmoid**	1 Layer (Layer 1: 400)	No. Epochs: 100 No. Batch: 10	MSE: **58.01** RMSE: **7.62**	MSE: 132.44 RMSE: 11.51

Predicted vs. Actual Outcome (i.e. Running Speed)



Summary & Future Research

Summary

Our main goal is to predict a spontaneous behavior of a mouse (i.e., running speed) from its neural activity by fully-connected multilayer models.

Our analysis shows that ReLU three-layer model predicted running speed with acceptable error (RMSE: 11.46 (measured for running speed)).

Future Research

Our analysis result is based on the Stringer data with one mouse; therefore, our research findings are limited and do not generalize to all mice. By increasing the size of dataset with more mice data, we will be able to provide more robust models.

For future research, we would like to experiment with the following options to improve model performance:

- Option 1: Increate the size of dataset by data generation methods (e.g., augmentation)
- Option 2: Improve the generalization of models by methods (e.g., dropout)