

# Investigating the Presence of Self-Organized Criticality in Neural Networks

Ty Brennan, Dr. Ashforth

April 20, 2023

## Abstract

The human brain is an incredibly complex system with unparalleled efficiency and capabilities, but very little is known about how the brain gives rise to cognition. Recent studies have consistently demonstrated that the brain functions near a critical state, a network state that lies between chaotic and ordered structures. Supercritical and subcritical states of the brain may lead to epilepsy or comas. Modern advances in artificial intelligence have become synonymous with the artificial neural network (ANN), a computer program inspired by networks of neurons in the brain. Although ANNs have been capable of many amazing feats, from human-like text generation to facial recognition, they are still a far cry from their biological inspiration. However, due to their powerful computational abilities, the network formed by the ANN may still operate near a critical state. We experimented with the generation of neuronal avalanches in a multilayer perceptron model trained to classify handwritten digits (MNIST) [3] and the presence of critical phenomena during network operation. We show that the activation ratio is very close to one. From these results, we show that criticality is strongly exhibited in the trained multi-perceptron model.

## Introduction

Criticality is often described as the boundary between order and disorder. A common example of a critical system is a sandpile – where adding individual grains of sand can randomly cause an avalanche of activity. Strong

experimental evidence points to the existence of a critical state within the human cortex, with supercritical and subcritical states being associated with epilepsy and comatose states, respectively. Firing neurons can cause neuronal avalanches of activity, leading to a cascade of resultant firings in downstream neurons – like avalanches of sand on a sandpile. The extent of these neuronal avalanches follows a power law, which is indicative of a self-organized critical system, that is, a system that tends towards a point of criticality without external help. Although there does not exist a working model of the human brain, artificial neural networks have recently powered a revolution in artificial intelligence. Additionally, any system near a point of criticality has enhanced information transmission capabilities, and high correlation (that is, the extent to which distant parts of the system interact), properties that are theorized to be valuable to intelligent systems. Artificial neural networks are a leading paradigm that has recently dominated machine learning. The ANN comprises a graph of densely connected nodes, each designed to function like a highly simplified biological neuron. The presence of criticality in trained ANNs can help guide our understanding of the importance of criticality in neuronal systems in general.

## **Methods**

## **Results**

## **Discussion**

## **Acknowledgements**

## **Bibliography**

### **Topic(s):**

1. Enforcing Supercriticality in Hopfield neural networks
2. AIRR
3. Using Reinforcement learning methods to train Energy Based Neural Networks

4. Using Sparse Hopfield networks for Immune Repertoire Classification
5. Sparse Hopfield Neural Networks for Deep learning
6. Boltzmann machines
7. Biologically plausible neural network

$$E = -\text{lse}(\beta, \mathbf{X}^T \boldsymbol{\xi}) + \frac{1}{2} \boldsymbol{\xi}^T \boldsymbol{\xi} + \beta^{-1} \log N + \frac{1}{2} M^2$$

$$E_1(\boldsymbol{\xi}) = \frac{1}{2} \boldsymbol{\xi}^T \boldsymbol{\xi}$$

$$E_2(\boldsymbol{\xi}) = -\text{lse}(\beta, \mathbf{X}^T \boldsymbol{\xi})$$

$$\nabla_{\boldsymbol{\xi}} E_1^{t+1} = -\nabla_{\boldsymbol{\xi}} E_2^t$$

## NOTES

- Hopfield Energy func:  $E = -\frac{1}{2} \boldsymbol{\xi}^T W \boldsymbol{\xi} + \boldsymbol{\xi}^T b$
- Hopfield update rule:  $\boldsymbol{\xi}^{t+1} = \text{sgn}(W \boldsymbol{\xi}^t - b)$
- Modern Hopfield Energy func:  $E = -\sum_{i=1}^N F(x_i^T \boldsymbol{\xi})$  where can be  $F = \exp$
- The above energy can also be written as  $-\exp(\text{lse}(1, X^T \boldsymbol{\xi}))$
- Modern Hopfield update rule:  $\boldsymbol{\xi}^{new}[l] = \text{sgn}[-E(\boldsymbol{\xi}^{(l+)}) + E(\boldsymbol{\xi}^{(l-)})]$
- Continuous Valued Energy function:  $-\text{lse}(\beta, X^T \boldsymbol{\xi}) + \frac{1}{2} \boldsymbol{\xi}^T \boldsymbol{\xi} + \beta^{-1} \log N + \frac{1}{2} M^2$ , which is the log of the negative energy above.
- $\nabla_{\boldsymbol{\xi}} (\frac{1}{2} \boldsymbol{\xi}^T \boldsymbol{\xi}) = \mathbf{1}$  and  $-\nabla_{\boldsymbol{\xi}} \text{lse}(\beta, X^T \boldsymbol{\xi}) = X \text{softmax}(\beta X^T \boldsymbol{\xi})$ , which means that  $\boldsymbol{\xi}^{new} = X \text{softmax}(\beta X^T \boldsymbol{\xi})$ , which is the new update procedure.
- The above update rule can be generalized to  $\boldsymbol{\Xi}^{new} = X \text{softmax}(\beta X^T \boldsymbol{\Xi})$
- 

$$\frac{m_{total} * v_i^2}{2} =$$