MINISTRY OF EDUCATION
TECHNICAL UNIVERSITY
OF CLUJ-NAPOCA, ROMANIA

Electronics, Telecommunications and Informational Technology

# PhD THESIS Summary / Rezumat

## Advanced processing methods and machine learning algorithms applied on neurobiological signals

**PhD Student: Andrei Ciuparu**

**PhD Supervisor: Prof. Dr. Eng. Corneliu RUSU, PhD**

**- Cluj-Napoca -**
**2022**

**Examination committee:**
Chair: Prof. Dr. Eng. **Sorin Hintea**, PhD - Technical University of Cluj-Napoca;
PhD supervisor: Prof. Dr. Eng. **Corneliu Rusu**, PhD - Technical University of Cluj-Napoca
Members:
- Prof. Dr. Eng. **Corneliu Burileanu**, PhD - Polytechnic University of Bucharest;
- Prof. Dr. Eng. **Radu Vasiu**, PhD - Polytechnic University of Timisoara;
- Prof. Dr. Eng. **Tudor Palade**, PhD - Technical University of Cluj-Napoca;

# Contents

# 1. Introduction

Machine learning (ML) techniques are fundamental to artificial intelligence, finding extensive use across commercial and research domains. These methods involve supplying algorithms with data and enabling them to adapt their internal models to make relevant predictions or interpretations. ML algorithms essentially tackle complex optimization problems. In neuroscience, we commonly deal with high-dimensional datasets with complex relationships between input features. In this chapter, we provide an overview of various ML algorithm types and the data collection methods used for this thesis.

## 1.1 Supervised and Unsupervised Learning

Machine learning algorithms can be categorized into two major groups based on their data access: supervised and unsupervised methods. The central distinction lies in the presence of target variables; supervised learning uses them, while unsupervised learning does not.

Supervised learning involves algorithms that use input features and corresponding target variables to adjust their models. In contrast, unsupervised algorithms only have access to input features and aim to group similar objects together based on feature similarity.

Supervised methods are effective when large labeled datasets are available, enabling predictions for new data instances. Applications range from facial recognition to speech-to-text algorithms. Unsupervised methods, though less suited for applications, are valuable in research, handling scenarios with limited labeled data, and have applications in neuroscience and data labeling.

### 1.1.1 Applications of Supervised ML

Supervised ML encompasses various algorithms where the underlying principle involves iteratively refining model parameters based on input data. Examples include linear regression, neural networks, and convolutional networks.

Fitting algorithms focus on optimizing parameters to minimize a cost function (e.g., mean squared error) for accurate predictions. Multi-layer perceptrons (MLPs), or neural networks, extend this concept to nonlinear functions, making them capable of arbitrary mappings, and thus, powerful information detectors. Convolutional neural networks (CNNs) build upon MLPs, excelling in image processing by applying weighted sums over input regions.

### 1.1.2 Applications of Unsupervised ML

Unsupervised learning aims to discover data clusters without target labels. This category includes density-based and distance-based clustering algorithms, as well as Self-Organizing Maps (SOMs).

Density-based algorithms like DBSCAN analyze local point densities to identify clusters and discard outliers, while distance-based algorithms like k-means rely on distances between points. SOMs map high-dimensional data onto a lower-dimensional space, preserving relational information.

These methods have strengths and weaknesses; density-based algorithms excel at non-gaussian clusters and outliers, while distance-based algorithms are not capable of detecting non-linear separation boundaries, and all are sensitive to dimensionality. SOMs offer insight into data relationships, bridging clustering and visualization, but also rely on distance as a measure of similarity.

## 1.2 Neurobiological Data Acquisition

The synergy between machine learning and neurobiology holds immense potential, particularly in dealing with the intricacies of high-dimensional data derived from neuroscientific techniques. These methodologies offer a pathway to extract meaningful insights from complex neural recordings that traditional statistical approaches might struggle to unveil. Notably, machine learning techniques operate directly on data, unlike model-based traditional methods that impose predefined assumptions.

To comprehend the efficacy of machine learning in neuroscience, we must delve into the two main types of recording methods used in animal studies. These methods, shaped by their distinct characteristics, present challenges that make machine learning indispensable in their analysis.

### 1.2.1 Electrophysiological Recordings

Electrophysiological recordings provide valuable insights into neural dynamics. Techniques like electroencephalography (EEG) and magnetoencephalography (MEG) capture neural activity externally, albeit with spatial distortions. In contrast, invasive recordings within animals' brains, such as patch-clamp measurements, offer precise insights into individual cell behavior. However, this method limits contextual understanding. Extra-cellular electrode arrays mitigate this by capturing neural signals from multiple neurons, though challenges remain in data interpretation.

Preprocessing electrophysiological data is crucial. Techniques like spike sorting, aided by unsupervised machine learning, help identify individual neuron activity. Machine learning tackles the daunting task of deciphering spiking patterns across high-dimensional datasets. It contributes significantly to understanding the complex dynamics within neural circuits.

### 1.2.2 Calcium Imaging

Incorporating genetically engineered proteins into animal cells allows for calcium imaging, offering dynamic insights into neural activity. Fluorescence microscopy captures changes in fluorescence, revealing calcium influx. Post-acquisition, motion correction addresses movement-induced distortions. The selection of regions of interest (ROIs) and subsequent data processing further refine the acquired signals. Machine learning assists in correcting fluorescence changes, extracting spike times, and addressing the multidimensional nature of the data.

## 1.3 Thesis Objectives

This thesis seeks to harness machine learning to enhance neurobiological data analysis. The primary goal is to innovate and adapt machine learning methods for comprehensive neurobiological insights. Deep learning, although often viewed as a 'black box,' holds promise in this realm. The thesis aims to explore the use of deep learning while striving to unravel its outcomes. The intention is to develop transparent and interpretable strategies to bridge the gap between complex algorithmic outputs and meaningful biological understanding. Ultimately, the thesis endeavors to employ machine learning to address pertinent neuroscience queries and surmount challenges related to model interpretability.

# 2. Implementation and Testing of a Novel Activation Function

## 2.1 Introduction

Activation functions play a crucial role in shaping the behavior of neural networks. This section introduces the significance of activation functions, their functioning, and their impact on network behavior. It then outlines classical activation functions (sigmoid, hyperbolic tangent, rectified linear units (ReLU), softplus) and introduces a new activation function, Soft++, highlighting its unique features. The chapter focuses on the testing of this function on benchmark datasets and presents the corresponding results.

### 2.1.1 Activation Function Overview:

Neural networks apply activation functions to a weighted sum of inputs before passing them on to the next layer. These functions greatly influence network behavior by affecting input transformations and error gradient propagation. Activation functions have evolved from neuron-inspired models to designs with properties that facilitate faster and robust learning. The choice of activation function significantly impacts network performance.

**Biologically Inspired Functions:**

Early activation functions like sigmoid and hyperbolic tangent were based on neuron models. These functions have on and off states, continuous domains, and suffer from the vanishing gradient problem. In deeper networks, during the backpropagation step, applying the chain rule multiple times results in the gradient becoming insignificantly small at early layers.

**Rectified Linear Units (ReLU) and Variants:**

ReLU emerged as a solution to the vanishing gradient problem, with distinct on and off states. It became the most widely used activation function in deep networks. Variants like leaky ReLU (LReLU) and parametric ReLU (PReLU) introduced slopes to address dying units, allowing negative gradients. Softplus, GELU, and Swish were created as smooth approximations of ReLU. However, ReLU functions can cause units to become inactive or "die."

**Exponential-Linear Units (ELU):**

ELU was developed as an alternative to ReLU, attempting to mitigate its issues. It has continuity and exponential behavior for negative inputs. ELU and scaled ELU (SELU) attempt to solve the vanishing gradient problem but can still saturate in the negative domain.

## 2.2 Introducing and Testing Soft++:

Soft++ is designed as a continuous version of PReLU. It combines the smoothness of the softplus function with a non-zero negative gradient. Parameters in Soft++ control gradient saturation in both positive and negative domains, enabling diverse shapes and behaviors. The chapter explores various parameter settings and evaluates Soft++ on benchmark datasets.

### 2.2.1 Test Datasets:

Benchmark datasets, including MNIST, CIFAR-10, and CIFAR-100, are used to evaluate Soft++ performance. A synthetic dataset with low signal-to-noise ratio and a neurobiological dataset from mouse visual cortex recordings are also tested. These datasets cover diverse scenarios to assess the generalizability of Soft++.

### 2.2.2 Comparative Testing Procedure:

Different network architectures are used for each dataset, and multiple runs assess sensitivity to initial conditions. Identical networks are employed with varying activation functions. Soft++ parameters are tested extensively to optimize performance.

## 2.3 Results:

Results show Soft++ performing exceptionally well. On MNIST, it reaches comparable performance with modern functions. On CIFAR-10, Soft++ outperforms all other functions in terms of convergence speed and final performance. Soft++ also surpasses ELU and SELU on CIFAR-100, showcasing improved performance with longer training.

## 2.4 Conclusions:

Soft++ offers benefits of other activation functions without their drawbacks. Its smooth gradient and non-saturating nature make it a powerful tool. Tests demonstrate its faster learning, better generalization, and potential to balance learning speed and power. Soft++ excels in scenarios with limited samples and features. Its potential impact is underscored by its success on neurobiological data. The results underscore its potential as a valuable tool in neural network applications, particularly in scenarios with low signal-to-noise ratio.

# 3. Gradient-K Clustering: An Improved K-Means Algorithm

## 3.1 Introduction

In this chapter, we present a novel approach, known as the Gradient-k algorithm, which addresses various limitations of the traditional k-means clustering algorithm. Our goal was to design an algorithm that can handle clusters with varying densities, shapes, and sizes by using additional information derived from the gradient of the density function for more accurate results.

The primary challenge with distance-based algorithms, such as k-means, is their reliance on distance metrics as the sole similarity measure. These algorithms tend to create linear separation barriers between clusters, leading to issues like the tessellation problem. Density-based algorithms, on the other hand, overcome this problem by designing non-linear separation boundaries and using local density thresholds to assign points to clusters. We aimed to fuse the benefits of both approaches by using the gradient of an approximate density function to modify the distances used in k-means.

## 3.2 Methods

### 3.2.1 Gradient-k Algorithm

The Gradient-k algorithm, although generalizable to N dimensions, is discussed here for two-dimensional datasets or datasets reduced via Principal Component Analysis (PCA). To avoid the curse of dimensionality, parameters of this algorithm must be adapted for higher dimensions. The algorithm is similar to k-means with two key modifications.

Discretization: The point space is divided into evenly spaced boxes along each dimension. This transforms the problem into box space, offering computational advantages and normalization. A density function is created by counting the points in each box, providing a density distribution.

Gradient Calculation: The density function's gradient is computed after applying a smoothing kernel. The gradient is utilized to adjust the distance metric for clustering.

Gradient-k starts by selecting K initial boxes using the K++ initialization algorithm. The distance from each box to the cluster centers is corrected using a factor calculated from the angle between the density gradient and the direction to the cluster center. The corrected distance guides the assignment of boxes to clusters, and cluster centers are updated iteratively until convergence.

### 3.2.2 Parameter Search

Optimizing algorithm parameters is crucial for fair comparisons between clustering methods. We employed the Optuna framework to automatically sample parameter combinations. Parameters for Gradient-k and DBSCAN were optimized using Optuna's TPE algorithm, ensuring balanced comparisons with k-means.

### 3.2.3 Comparative Analysis

Our comparative study involved classic k-means, DBSCAN, and Gradient-k algorithms on various benchmark and synthetic datasets. We conducted 1000 trials for each algorithm and dataset combination, evaluating iteration count until convergence and clustering accuracy against ground truth.

## 3.3 Results

In terms of accuracy and convergence, Gradient-k consistently outperformed or matched k-means on all tested datasets. Additionally, Gradient-k exhibited better or comparable performance compared to DBSCAN in most cases, except for the Aggregate dataset. Notably, Gradient-k converged faster than k-means on the spike dataset. Significance testing supported the observed differences between algorithms.

Qualitative analysis revealed that Gradient-k's clustering results closely resembled ground truth, particularly when dealing with non-linear and non-Gaussian clusters. The algorithm demonstrated superior performance on datasets with complex separation boundaries, where traditional methods struggled.

## 3.4 Discussion

### 3.4.1 Advantages and Disadvantages

Gradient-k addresses several k-means limitations by leveraging density gradient information. It overcomes issues like linear separation barriers and the tessellation problem, thanks to its non-linear separation boundaries and handling of varying densities. The algorithm's computational complexity remains constant with respect to the number of input points due to its operation in box space. Moreover, Gradient-k often converges as fast or faster than k-means.

However, Gradient-k has drawbacks, including the need for a pre-determined number of clusters and parameter adjustments. Furthermore, its computational complexity escalates exponentially with dataset dimensionality.

### 3.4.2 Further Research and Improvements

Future research should focus on extending Gradient-k to higher dimensions and devising an automatic initialization process. This would eliminate the need to specify the number of clusters and enhance algorithm speed and accuracy. Additionally, exploring alternative ways to leverage density information could provide new avenues for advanced clustering techniques.

## 3.5 Conclusions

The Gradient-k algorithm presents a promising solution to k-means limitations, yielding accurate and efficient results. It bridges the gap between distance-based and density-based algorithms, offering advantages in various scenarios. While challenges remain, such as high-dimensional scalability and automated initialization, Gradient-k showcases the potential of density gradient-based information in enhancing clustering algorithms.

# 4. Leveraging Irregular Sampling to Capture Fast Brain Oscillations

## 4.1 Introduction

Calcium imaging has revolutionized research in biology, neuroscience, and medicine by enabling simultaneous measurements of neuronal activity. However, its effectiveness is constrained by the protein's behavior and hardware limitations, leading to restricted observable frequencies. We propose a solution using irregular (jittered) sampling to detect fast brain activity like gamma oscillations (30-150Hz) using GCaMP.

### 4.1.1 GCaMP Recording Constraints

GCaMP imaging offers advantages like tracking neuron identity and measuring multiple neuronal activations. Yet, limitations arise from slower calcium signals and hardware constraints, leading to lower-pass filtered signals. This restricts the observable bandwidth, currently around 30-60 Hz. Hardware factors like sample scanning speed also impose limitations, resulting in lower sampling rates (10-20 Hz).

### 4.1.2 Jittered Sampling Across Fields

Jittered sampling involves capturing samples at random intervals, refining the sampling grid and capturing oscillations at multiple phases. Crucially, this does not imply decreasing the inter-sample interval below what would be possible with a given sampling rate, but sampling at longer intervals. While discussed in theory, practical applications have been limited. This method potentially circumvents bandwidth limitations by introducing randomness to the sampling process.

## 4.2 Methods

Synthetic datasets were generated to test irregular sampling's effectiveness in extracting high-frequency information from undersampled signals. Signals included target and distractor oscillations along with noise. We performed analyses using a simple sine function model to predict target oscillation amplitude, using optimization algorithms for fitting.

## 4.3 Results

Under regular sampling, aliases caused challenges in predicting target oscillation amplitude, especially when distractor oscillations interfered. Irregular (jittered) sampling demonstrated significant improvements, enabling accurate prediction of the target oscillation amplitude, even in the presence of distractor frequencies.

## 4.4 Discussion

Our study demonstrates how incorporating random delays between frames through irregular sampling can enhance the extraction of high-frequency information from GCaMP recordings. The approach successfully predicted oscillation amplitude, frequency, and timing. The main limitation of this research is the fact that the tested signal is markedly different than real brain data. Implementing irregular sampling in recording hardware and applying this methodology to real data could validate its potential for capturing fast brain activity.

## 4.5 Conclusion

Irregular sampling emerges as a promising technique to overcome limitations in capturing fast brain oscillations using calcium imaging. By introducing a jittered sampling scheme, it becomes possible to accurately estimate high-frequency oscillation properties that are typically challenging to observe with traditional techniques. This novel approach holds potential for advancing our understanding of brain activity dynamics and their implications in various neurological disorders.

# 5. ML Applications in GcAMP Analysis Pipeline

In this section, we present our collaboration with Cold Spring Harbor Laboratory (CSHL) to analyze fluorescence data obtained from mouse brains using two-photon microscopy. The study aimed to understand the role of feedback from the piriform cortex (PC) on the olfactory bulb (OB) activity using a rule-reversal go/no-go task. We improved data preprocessing, developed novel interpolation methods, and applied machine learning techniques for analysis.

## 5.1 Data Preprocessing improvements

### 5.1.1 Data-driven Interpolation

The recorded data consisted of GCaMP (calcium imaging) signals from mice performing a rule-reversal task. To process the data, we addressed challenges in motion correction and ROI signal extraction. We introduced a novel empirically derived interpolation procedure that considered signal characteristics and the length of missing data windows, resulting in improved data retention compared to standard methods. Our method allowed for more lenient rejection criteria, retaining around twice as many trials.

### 5.1.2 Bleaching Correction and Normalization

We redesigned the preprocessing pipeline to address bleaching correction and normalization separately. Exponential detrending was used for bleaching correction, effectively removing bleaching effects while preserving signal integrity. We also implemented z-scoring for normalization, providing a more accurate representation of the signal's underlying dynamics.

## 5.2 Analyses and results

### 5.2.1 Kohonen Mapping

To explore the data, we applied Kohonen mapping (Self-Organizing Maps) analysis, converting the data into a color representation. This technique revealed intriguing patterns, including baseline state shifts, multiphasic stimulus responses, and representational leakage across task blocks. We introduced an auxiliary analysis to identify informative patterns and distinguish them from non-specific ones.

### 5.2.2 MLP Analysis

We employed multi-layer perceptrons (MLPs) to classify stimulus and behavioral characteristics across trials. MLPs were trained on segmented data and used to forecast different features related to the task. We conducted control analyses to validate our results and assess the significance of the obtained information. We also investigated how stimulus representations evolved across experimental blocks, analyzing same-block and across-block scenarios.

### 5.2.3 MLP Results

The GFP control results demonstrate an increase in classification performance of 30% compared to the baseline on average. Notably, the classification accuracy enhancement was more pronounced for contingency, instruction, and behavior classifications compared to stimulus classification. This outcome aligns with the fact that these variables are behaviorally linked or correlated, indicating that the algorithm's performance might be driven by non-neural activity-related factors. These results serve as a valuable baseline for future analyses aiming to identify activity-related information.

Contrastingly, the GCaMP datasets exhibited a more consistent increase in performance across datasets and classification tasks. Most datasets showed performance improvements exceeding 50% compared to the baseline for all classification tasks post-stimulus. Stimulus classification performance declined more rapidly than other types, reflecting potential processing dynamics within the mouse olfactory bulb. These findings confirm the presence of rich and complex feedback from the piriform cortex to the bulb, conveying task-relevant information even about non-olfactory stimuli.

The cross-stimulus classification analysis provided insights into task-related patterns. If piriform cortex feedback represents the given task (instruction), one might expect some patterns to flip across stimulus blocks. This effect was observed across most datasets but not consistently, likely due to variations in the measured boutons across sessions and the reduced information available for this analysis. The Kohonen analysis aligns with these new insights, indicating that some stimulus-related patterns persist across blocks and take several trials to update to new rules. By measuring post-stimulus trajectories' similarity to previous block models, the evolving task parameter information's encoding in a changing environment was further elucidated.

### 5.2.4 Multi-Dimensional Analyses

In addition to the machine learning techniques discussed earlier, high-dimensional trajectory analysis offers insights into data features. This analysis treats activity combinations at each time point as high-dimensional points, forming trajectories over time. Symbolic entropy, angular coherence, and trajectory matching measures were developed to explore relationships between experimental conditions.

The trajectory matching analysis employed distance metrics to compare model trajectories for each condition against test trajectories. Angular coherence further distinguished patterns from scaled versions, while symbolic entropy measured system states exhibited across trials. These analyses revealed that sound trials displayed greater self-consistency, odor trajectories exhibited larger deviations upon stimulus onset, and entropy increased for both stimuli. The results suggested that the observed increase in entropy was due to distance scaling rather than pattern changes.

Furthermore, a demixed PCA-based approach provided insights into the variance associated with task parameters across ROIs. This analysis demonstrated higher task-relevant information in odor blocks, consistency in stimulus-related variance across blocks, and a substantial role of condition-independent components in explaining ROI variance.

## 5.3 Conclusions

This chapter showcased innovative machine learning methods applied to fluorescence data analysis. Improvements in preprocessing techniques and algorithm adaptation were discussed. Unsupervised (Kohonen mapping) and supervised (MLP classification) methods revealed top-down encoding of task variables and multi-sensory information in the olfactory bulb. High-dimensional statistical analyses elucidated trajectory dynamics and encoding stability across different experimental conditions. Finally, the dPCA approach provided information profiles for individual ROIs, highlighting their associations with task variables. Collectively, these analyses provided strong evidence for task-related top-down encoding and multi-sensory information processing within the olfactory bulb.

# 6. Measuring Mutual Information Across Brain Areas

Understanding the mechanisms underlying decision-making and information flow in the brain is crucial in neuroscience. In a collaborative effort funded by the H2020 NEUROTWIN grant, the Mrsic-Flogel Lab at the Sainsburry Wellcome Centre, University College London, provided data collected during experiments conducted to explore how information is transmitted across different brain regions.

## 6.1 Experimental Setup

The experiments aimed to investigate decision-making processes and information transfer from sensory input areas (specifically visual) to motor planning and execution areas. Mice were exposed to drifting gratings with varying speeds or temporal frequencies (TFs) selected from a normal distribution. The mice were trained to detect an increase in the TF distribution and respond by licking.

## 6.2 Data, Preprocessing, and Analysis

Data was collected from various cortical areas, including simultaneous recordings from primary visual and motor areas using neuropixels probes. The Kilosort algorithm was employed to extract firing times of individual units . Activity vectors were generated by convolving a delta train with these firing times, with an exponential decay kernel.

The Kohonen Self-Organizing Map (SOM) technique was used to visualize neuron patterns around change onset and false alarms. A multi-layer perceptron (MLP) regression approach was developed to infer information transfer between areas. This algorithm acted as a mapping from input to output area, and the comparison between predictions and actual data generated an error signal (MSE) for training.

## 6.3 Results

The Kohonen analysis revealed patterns related to change onset and false alarms, showing similar activity when the animal was fooled. A communication profile was established using MLP regression, highlighting preferred communication lags and salient events. This profile contained significant information beyond population firing rates.

## 6.4 Discussion

This analysis offered insights into common information shared by motor and visual areas. However, caution is needed in interpreting the results. The analysis did not provide directionality or complete mutual information understanding. Control analyses are required for stronger conclusions. The study also emphasized the ability of MLPs to map complex spaces, highlighting the simplicity and potential online control applications.

### 6.4.1 Explainable AI

The field of explainable AI was considered for better understanding machine learning insights. Feature perturbation analysis can be employed, allowing for precise dPCA-like interpretations of importance. And decisions regarding which features to shuffle and how to shuffle them can allow for the seperation of different encoding schemes (encoding in: population firing rate, temporal fluctuations in population firing rate, etc.). Such methods were useful for separating relevant features in input spaces.

## 6.5 Conclusions

This study demonstrated the utility of Kohonen SOM and MLP regression for uncovering patterns and communication profiles in neural data. Despite limitations, these methods provided valuable insights into information transfer. Further research involving explainable AI techniques could enhance our understanding of complex brain processes.

# 7. Multiple Regression for Behavior Analysis

In this chapter, we delve into the applications of multiple regression for behavioral analysis. Regression involves predicting non-categorical output variables from input data. While established models can be employed, we explore the use of Multi-Layer Perceptrons (MLPs) as self-trained models due to their universal approximation capabilities. We present practical examples to showcase their potential, acknowledging their strengths and limitations. This analysis draws from data provided by Koen Vervaeke's team at the Laboratory for Neural Computation, University of Oslo, as part of the CIRCUITGENUS grant collaboration.

## 7.1 Experimental Setup

Our experiment focuses on the retrosplenial cortex's (RSC) response to tactile cues during mouse locomotion on a simulated corridor. The goal is to observe RSC activity with respect to the animal's position and tactile cues. By examining whether the RSC stores information about position and responds to tactile input, we aim to uncover neural mechanisms related to navigation and external input correction.

## 7.2 Data, Preprocessing, and Analysis

RSC activity was measured via widefield microscopy at a rate of 31Hz. After motion correction and ROI selection, time traces were obtained for each ROI. To maintain data integrity, we used raw data, as MLPs were anticipated to extract the signal from noise.

We initially employed Kohonen Self-Organizing Maps (SOM) to visualize neuron response patterns to stimuli. By assigning colors to fluorescence patterns and synchronizing them with temporal events, we identified patterns associated with circuit completion, reward administration, and new run anticipation. To analyze position-related patterns, we aligned patterns with animal position using interpolation, creating continuous plots.

Furthermore, we utilized a three-layer MLP with soft++ activation for classification and regression tasks. We attempted to separate stimulus identity, stimulus occurrence number, and all task events (including reward).

## 7.3 Results

SOM visualization revealed distinct patterns related to circuit completion, reward, and anticipation of new runs. Spatial analysis uncovered events triggering responses at specific locations along the track. Classification tasks demonstrated the predictability of texture type, first vs. second encounters, and location onset. Moreover, regression analysis accurately decoded position with minimal error.

## 7.4 Conclusions

From a neuroscience perspective, we reaffirm that RSC encodes navigation-related information. Future exploration involves analyzing regression performance variations elicited by stimulus encounters, potentially identifying internal updates. The engineering viewpoint emphasizes the potential of relatively simple MLPs to decipher multi-neuronal encoding of continuous variables. Overall, this chapter showcases the utility of multiple regression and MLPs in uncovering neural mechanisms underlying behavior.

# 8. Conclusions

This thesis has achieved its primary goals by bridging the gap between machine learning (ML) and neuroscience, yielding advancements in both fields. Through innovative adaptations and improved algorithms, this research demonstrates the potential to address complex questions and challenges. The methods developed have significant implications for future scientific inquiries.

The first studies included in the thesis focused on the development of novel techniques, and the adaptation of existing ones, for use in neuroscience. The first chapter discusses the design, implementation, and assessment of a new activation function named Soft++. When incorporated into multi-layer perceptrons and convolutional neural networks, this function outperformed modern counterparts, especially under low signal-to-noise ratio conditions. Another noteworthy advancement is the conception of a clustering algorithm, Gradient-k. Unlike conventional methods, this algorithm leverages an estimated density function to enhance the accuracy and speed of the k-means algorithm, even recognizing non-gaussian clusters of varied shapes. Furthermore, the research revisits the Jittered sampling technique by applying it on signals resembling a calcium imaging recording. This study suggests that a minor software modification in the recording process could lead to a twofold increase in the maximum measurable frequency, thereby expanding its applicability.

The second half of the thesis is focused on applications of such techniques for answering neuroscience questions directly, as well as tackling the black box problem. First, we tackle the question of the role of piriform cortex feedback to the olfactory bulb. By amalgamating machine learning algorithms with high-dimensional trajectory analyses, the research elucidates how this feedback multiplexes distinct task-related variables. Then, the interplay between sensory and motor regions during decision-making tasks was examined. Through the application of multi-layer perceptron regression, the communication dynamics between these areas were detailed, followed by an exploration of varying control analyses and their potential data implications. Lastly, a preliminary analysis spotlighted the retrosplenial cortex's role in navigation, hinting at the boundaries of machine learning analyses due to experimental design, especially when illustrating how the RSC processes spatial information.

In essence, this research underscores the vast possibility afforded by applying machine learning methodologies on neuroscientific inquiries, paving the way for improved and novel tools and perspectives for upcoming research endeavors.

# 9. Publications:


- Ciuparu, Andrei, and Raul C. Mureşan. ”Sources of bias in single-trial normalization procedures.” European Journal of Neuroscience 43, no. 7 (2016): 861-869. WOS:000373773100001
- Ciuparu, Andrei, Adriana Nagy-Dăbâcan, and Raul C. Mureşan. ”Soft++, a multi-parametric non-saturating non-linearity that improves convergence in deep neural architectures.” Neurocomputing 384 (2020): 376-388. WOS:000513853600032
- Ciuparu, Andrei, and Raul C. Mureşan. ”Jittered sampling-a potential solution for detecting high frequencies in GCaMP recordings.” In 2021 IEEE 17th International Conference on Intelligent Computer Communication and Processing (ICCP), pp. 469-475. IEEE, 2021. IEEE code: https://ieeexplore.ieee.org/document/9733598
- Ciuparu, Andrei, and Raul Cristian Muresan. ”Gradient-k: Improving the performance of K-Means using the density gradient.” bioRxiv (2022): 2022-03.
- Grosu, George F., Alexander V. Hopp, Vasile V. Moca, Harald Bârzan, Andrei Ciuparu, Maria Ercsey-Ravasz, Mathias Winkel, Helmut Linde, and Raul C. Mureșan. ”The fractal brain: scale-invariance in structure and dynamics.” Cerebral Cortex 33, no. 8 (2023): 4574-4605. WOS:000859057800001
- Gheorghiu, Medorian, Andrei Ciuparu, Bartul Mimica, Jonathan R. Whitlock, and Raul C. Mureşan. ”A Machine Learning Approach to Investigate Fronto-Parietal Neural Ensemble Dynamics During Complex Behavior.” In 2020 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR), pp. 1-6. IEEE, 2020. https://ieeexplore.ieee.org/document/9129986
- Trejo, Diego E. Hernandez, Andrei Ciuparu, Pedro Garcia da Silva, Cristina M. Velasquez, Martin B. Davis, Raul C. Muresan, and Dinu F. Albeanu. ”Cortical-bulbar feedback supports behavioral flexibility during rule-reversal.” Chemical Senses 47 (2022). WOS:000936532400078