

Chapman University

Chapman University Digital Commons

Computational and Data Sciences (MS) Theses

Dissertations and Theses

Fall 12-2021

Identifying Functional Profiles of Challenging Behaviors in Autism Spectrum Disorder with Unsupervised Machine Learning

Emily Daskas

Chapman University, daska102@mail.chapman.edu

Follow this and additional works at: https://digitalcommons.chapman.edu/cads_theses



Part of the [Analytical, Diagnostic and Therapeutic Techniques and Equipment Commons](#), and the [Mental and Social Health Commons](#)

Recommended Citation

E. Daskas "Identifying functional profiles of challenging behaviors in Autism Spectrum Disorder with unsupervised machine learning," M. S. thesis, Chapman University, Orange, CA, Year. <https://doi.org/10.36837/chapman.000330>

This Thesis is brought to you for free and open access by the Dissertations and Theses at Chapman University Digital Commons. It has been accepted for inclusion in Computational and Data Sciences (MS) Theses by an authorized administrator of Chapman University Digital Commons. For more information, please contact laughtin@chapman.edu.

Identifying Functional Profiles of Challenging Behaviors in Autism Spectrum Disorder with
Unsupervised Machine Learning

A Thesis by

Emily Elizabeth Daskas

Chapman University

Orange, CA

Schmid College of Science and Technology

Submitted in partial fulfillment of the requirements for the degree of

Master of Science Computational and Data Sciences

December 2021

Committee in charge:

Erik Linstead, Ph.D., Chair

Dennis Dixon, Ph.D.

Rene German, M.S.

The thesis of Emily Elizabeth Daskas is approved.



Erik Linstead, Ph.D., Chair

DocuSigned by:

F641F861F55D4E6...

Dennis Dixon, Ph.D.



Rene German, M.S.

June 2021

Identifying Functional Profiles of Challenging Behaviors in Autism Spectrum Disorder with
Unsupervised Machine Learning

Copyright © 2021

by Emily Elizabeth Daskas

ACKNOWLEDGEMENTS

I would like to thank my friends, family, and the Machine Learning and Affiliated Technologies (MLAT) lab for all of their support and encouragement throughout my time at Chapman and during the completion of this project. Specifically, I would like to thank Rene German for pushing me and equipping me with the tools to pursue computer science. And of course, I would like to thank Dr. Erik Linstead for his encouragement, his confidence in me, and for all of the opportunities that he has made possible— my time at Chapman has been a wonderful experience, and a huge part of that is because of Dr. Linstead. I am honored to have been able to learn and work with everyone at Chapman and MLAT. Thank you for making my time at Chapman all that it was.

ABSTRACT

Identifying Functional Profiles of Challenging Behaviors in Autism Spectrum Disorder with
Unsupervised Machine Learning

by Emily Elizabeth Daskas

Machine learning and deep learning methods are becoming increasingly used in the understanding, identification, and improvement of the diagnosis and treatment of Autism Spectrum Disorder. People with ASD often exemplify challenging behaviors that can put their safety, education, and general quality of life at risk. Challenging behaviors are driven by one of four functions. The combination of common occurrences of challenging behaviors and their respective behavioral functions are unique to the individual and circumstance, and the most successful therapies account for both challenging behaviors and their respective functions. Therefore, it is important that research is done on these concepts to lead to improvements in therapy and outcomes.

In this thesis, we apply a cluster analysis to a sample of 1,416 individuals with Autism Spectrum Disorder. The aim is to find groupings of patients based on the relative frequency of each unique challenging behavior and function pair. As the first machine learning study to focus on combining the behavioral functions and challenging behaviors of ASD, we find that there are some patterns to be found based on eight identified clusters. The results of the study could impact the way that treatment and therapy plans are paved for children with Autism Spectrum Disorder.

TABLE OF CONTENTS

	<u>Page</u>
ACKNOWLEDGEMENTS.....	IV
ABSTRACT.....	V
LIST OF TABLES.....	VII
LIST OF FIGURES.....	VIII
LIST OF ABBREVIATIONS.....	IX
1 INTRODUCTION	1
1.1 Challenging Behaviors	2
1.2 Functions of Behaviors	3
1.3 Applied Behavioral Analysis	5
2 RELATED WORK.....	6
3 DATA.....	8
4 METHODS	14
4.1 Self-Organizing Maps	15
4.2 K-means	16
5 RESULTS	19
5.1 Self-Organizing Maps Results	19
5.2 K-means Results	22
6 DISCUSSION.....	27
7 FUTURE WORKS.....	29
8 CONCLUSION.....	31
REFERENCES.....	33
APPENDICES.....	39

LIST OF TABLES

	<u>Page</u>
Table I: The Four Functions of Behavior and Their Definitions	4
Table II: Top 15 Most Frequent Function And Behavior Pairs.....	11
Table III: The Top 15 Co-occurring Function And Behavior Pair.....	12

LIST OF FIGURES

	Page
Figure 3.1 Total Count of Each Function	9
Figure 3.2 Total Count of Each Challenging Behavior.....	9
Figure 5.1 SOM Nodes Grid with Weight Difference.....	20
Figure 5.2 Feature Network of Nodes	21
Figure 5.3 Elbow Plot.....	22
Figure 5.4 Polar Bar Charts for Eight K-Means Clusters.....	23
Figure 5.5 Polar Bar Chart Cluster 2.....	24
Figure 5.6 Four Feature Network of Nodes for SIB.....	26
Figure 5.7 Four Feature Network of Notes for Disruptive Behaviors	26

LIST OF ABBREVIATIONS

<u>Abbreviation</u>	<u>Meaning</u>
ABA	Applied Behavioral Analysis
ANN	Artificial Neural Network
ASD	Autism Spectrum Disorder
BMU	Best Matching Unit
ML	Machine Learning
SIB	Self-Injurious Behavior
SOM	Self-Organizing Maps
SSE	Sum of Squared Errors

Chapter 1

Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder. An individual with ASD may have difficulties engaging in behaviors that are of social significance, often showing a hindered ability to partake in everyday social interactions, lacking verbal and nonverbal communication skills, showing patterns of repetitive behaviors, and having restricted interests [1][2][3]. These behaviors are referred to as challenging behaviors. To help minimize the expression of these behaviors, a person with ASD will often undergo Applied Behavioral Analysis (ABA) [1][4].

ABA is a scientific therapy that focuses on respondent and conditioning techniques with the aim of minimizing the expression of challenging behaviors, and thus, limiting their consequences. ABA is highly individualized, based on the patient's expression of challenging behaviors as well the reason that the behaviors are occurring. The reason, or driving force, behind why a behavior occurs is called a function [4]. Therapies that put emphasis on functions of behaviors rather than challenging behaviors alone, produce measurably better outcomes [4][5][6]. Thus, it is essential to get a well rounded understanding of an individual in order for treatment to be most effective.

In this thesis, unsupervised machine learning techniques will be used to analyze the relationships between challenging behaviors and their corresponding functions. Specifically, we will use k-means to make inferences about functional profiles. This paper will serve as a step towards a

more comprehensive and data driven approach to understand challenging behaviors in context with their functions. It will leverage machine learning clustering approaches used in similar past research [7][8] as well as introduce new approaches. The hope is that the results of this research will serve as a mechanism in identifying the best treatment approach for successful early intervention ABA therapy regimens. Benefits of this would be to optimize the treatment outcomes as well as making treatment more effective, timely, and affordable [3][7].

Moving forward, for clarity and consistency purposes, unique challenging behaviors and their respective functions will be referred to as a “function and behavior pair”. When a specific function and behavior pair is discussed, it will be identified first by the function and then by the behavior, such as “automatic stereotypy”.

The rest of the paper will be presented in the following order: Chapter 2 will go over related works, Chapter 3 will discuss the data, Chapter 4 will describe the methods used to conduct this study, and Chapters 5 and 6 will cover the results and discussion, respectively. Finally, Chapters 6 and 7 will discuss the future of this research and draw conclusions.

First, however, Chapter 1 will wrap up by providing a brief overview of challenging behaviors [1.1], functions of behaviors [1.2], and ABA therapy [1.3] in order to provide a more comprehensive understanding of the purpose and applications of this study.

1.1 Challenging Behaviors

Challenging behaviors are defined as behaviors that are not culturally or socially appropriate. These behaviors need to be addressed because they have a wide array of potential consequences [1]. These consequences can include health or safety risks for patients or for others, inhibited learning, social isolation, and limited access to adequate living, education, and social environments [6][9][10].

While challenging behaviors are not exclusive to those with Autism Spectrum Disorder, individuals with ASD are often found to demonstrate them with more severe or frequent expression [1][11]. A person with ASD is more vulnerable to developing challenging behaviors, even in comparison to those with other neurodevelopmental disorders [6]. A study by Jang et al. [12] showed that out of 84 participants, 94% of children with ASD exhibited at least some form of challenging behavior.

While there are many common challenging behaviors, the most common ones are often understood to be related to aggressive, self-injurious, offensive, stereotypic, and destructive behaviors [1][12]. However, these top behaviors vary between studies and are not universally agreed upon.

For this study, we will focus on the eight most prevalent behaviors as described by our dataset: “aggression”, “disruption”, “elopement”, “noncompliance”, “obsessive behaviors”, “self-injurious behaviors” (SIB), “stereotypy”, and “tantrums”. The expression of each behavior manifests in many different ways. This variability can depend on age, social factors, and environmental factors, among other things. While there may be a general consensus on what defines each behavior, there is not a single, standard definition for any one behavior [13][14].

In general, it is believed that challenging behaviors do not have a more severe expression at any particular stage of life [10][15][16]. However, as more research is needed to understand ASD across the entire lifespan, there are underlying patterns of ASD and challenging behaviors that are not adequately understood yet [1].

1.2 Functions of Behaviors

A function is the driving force of a behavior [5]. In Applied Behavior Analysis, it is believed that challenging behaviors do not occur without reason. Instead, they are maintained by a function that triggers the behavior [17].

TABLE I

THE FOUR FUNCTIONS OF BEHAVIOR AND THEIR DEFINITIONS

Function	Definition
Attention	Behaviors that are expressed in an attempt to get attention from others [4].
Automatic Reinforcement	Behaviors that stimulate automatic reinforcement inherent to the behavior itself and not dependent on behaviors of other individuals outside of one's self (e.g. scratching inherently relieves an itch) [4].
Escape	Behaviors that are done with the goal of avoiding or removing one's self from a situation that they do not want to be in [4].
Access to Tangibles	Behaviors conducted to gain access to something tangible, or physical [4].

Lists the four functions of behaviors and their respective definitions.

In ABA therapy, there are four recognized functions, which are listed in our dataset as “escape”, “attention”, “access to tangibles”, and “automatic reinforcement” [5]. Understanding these functions can help an ABA provider accommodate better to an individual's needs and provide a more individualized and focused therapy plan. The recognition of functions is a critical aspect in providing comprehensive and successful therapies [5][18][19]. Table I provides a list of definitions of each of the four functions of behaviors.

Currently, there is not a single, concrete way of identifying a function [12]. Commonly implemented practices of identifying functions range from costly and resource intensive full

functional analysis, more inexpensive tools such as IISCA and QABF [18], or making use of different available rating scales [3][20]. Still, however, with the use of any of these options, it is not with complete certainty that the functions will be correctly identified, especially in young children. Although progress has been made, identification of behavioral functions, especially with early diagnosis and early intervention, is a highly predictive practice. Therefore, the understanding and practice of identifying behavioral functions is an area of ASD research that is in need of improvement [21].

For the remainder of this paper, “access to tangibles” will often be referred to as “tangible”. “Automatic reinforcement” will be referred to as “automatic”.

1.3 Applied Behavior Analysis (ABA) Therapy

Applied Behavior Analysis (ABA) is a common, well-recognized practice in the treatment of Autism Spectrum Disorder. It is empirically validated and, currently, the only treatment with substantial evidence to support its effectiveness, where the evidence comes via documentation, research, and measurable outcomes [1][2]. ABA treats its patients with a highly individualized, well-rounded approach, focusing on behavioral strengths and weaknesses, as well as recognizing the importance of using functional analysis to create optimal therapy plans. It works to control challenging behaviors by using consistent applications of reinforcement learning and controlling environmental factors, excelling at helping patients adapt to social situations [4]. Children will partake in the therapy daily over the span of several years, often for 30 or more hours per week [7][22].

Studies show that, depending on the individual and their severity of symptoms, ABA and similar therapies may have an 80-90% decrease in challenging behaviors [6]. In the best cases, it has demonstrated the ability to mitigate these behaviors completely [14][19].

Chapter 2

Related Works

In recent years, ASD research has become increasingly data driven. As a result, there have been several statistical, machine learning, and deep learning applications of ASD research to help improve both diagnosis and treatment [21]. One example of a machine learning application includes examining the relationship between treatment intensity and outcomes, validating the hypothesis that higher intensity treatments lead to measurably better outcomes [23]. Research also extends to the improvement in quality of life and inclusivity for ASD individuals. For example, one study uses data driven approaches to review how employers implement policies and practices to hire and best support employees with ASD [24]. There have also been several studies utilizing machine learning and deep learning techniques, such as SVMs and neural networks, on MRI and neuroimaging data to improve the diagnosis and early detection of ASD [21][25][26]. Similarly, several eye-tracking experiments have been conducted. To give an example, work done in [27] predicts the gaze and fixation patterns of children with ASD. Research and breakthroughs in the diagnosis of ASD is important because early intervention leads to superior treatment outcomes, yet the average time between the recognition of first symptoms and a clinical diagnosis is about a two year span [28][29][30]. This suggests that with the furthering and implementation of diagnosis-based research, outcomes and treatments could be improved by therapies beginning at even younger ages.

There have also been a number of studies that leverage unsupervised machine learning to make inferences about ASD by subgrouping individuals. These studies include a cluster analysis to discover ASD phenotypic patterns with Gaussian Mixture Models [8], clustering of co-occurring conditions with ASD [31], k-means clustering of challenging behaviors to explore treatment efficacy [32], as well as hierarchical clustering for phenotypic heterogeneity [33].

A similar unsupervised research paper, and perhaps, the most related work to this thesis, comes from [7], where a cluster analysis of challenging behaviors was performed. Here, k-means was used on aggregate patient data and individuals were sub-grouped based on similar challenging behavior patterns. Interestingly, as the study focused on eight challenging behaviors, most of the eight cluster profiles were dominated by a single challenging behavior [7]. [32] found similar clustering results as well. Both of these studies, though, along with the prior studies mentioned, did not include any analysis on functions of behavior.

Other research and literature regarding ASD, outside of data science applications, have covered the topics of challenging behaviors [1][34], ABA therapy [19], and functions of behaviors [3][4]. These studies show the positive outcomes when therapy focuses on functional analysis. Many studies also give attention to the most effective approaches of treating the functions of specific challenging behaviors. For example, work by [35] focused on noncompliance behavior. It showed that responses to consequence-based intervention varied depending on what function was driving the noncompliant behaviors, highlighting the importance of a function-based treatment approach. However, aside from case studies such as the one just mentioned, studies which incorporate functions of behavior are lacking, especially if we hone in on data-driven approaches. There have notably been very few comprehensive studies, or any with machine learning, that put emphasis on functions of behavior [21].

This research aims to serve as an extension of the previous work done in [7] by extending it to include behavioral functions, beginning to fill in the gap of lacking research on challenging behaviors in congruence with their corresponding functions.

Chapter 3

Data

The data in this study was provided by the Center for Autism and Related Disorders (CARD), one of the largest national providers of ABA therapy. Specifically, the data in this research comes from the CARD Skills™ dataset. This Skills™ dataset is a clinical database which holds all information regarding ABA curriculum as well as detailed documentation of each individual's progress for the totality of their time in therapy.

For the purpose of this study, we will focus on the observed challenging behaviors of an individual, the function of each exhibited behavior, and the count of how many times each unique function and behavior pair occurred.

It should be noted that all patients in this dataset were under supervision of BCBAs and other specialists employed by CARD during the full course of their treatment. The data in this dataset was logged whenever a patient was seen to exhibit a challenging behavior.

With this in mind, some limitations of the data should be addressed. We acknowledge the fact that the dataset is limited due to human error and therapy constraints. Counts in this dataset only occur where behaviors were both observed and recorded, and it is recognized that sometimes occurrences of challenging behaviors may not be witnessed or charted. There may also be variance in how an analyst chooses to specify or identify a function since there is no one correct way of making the distinction.

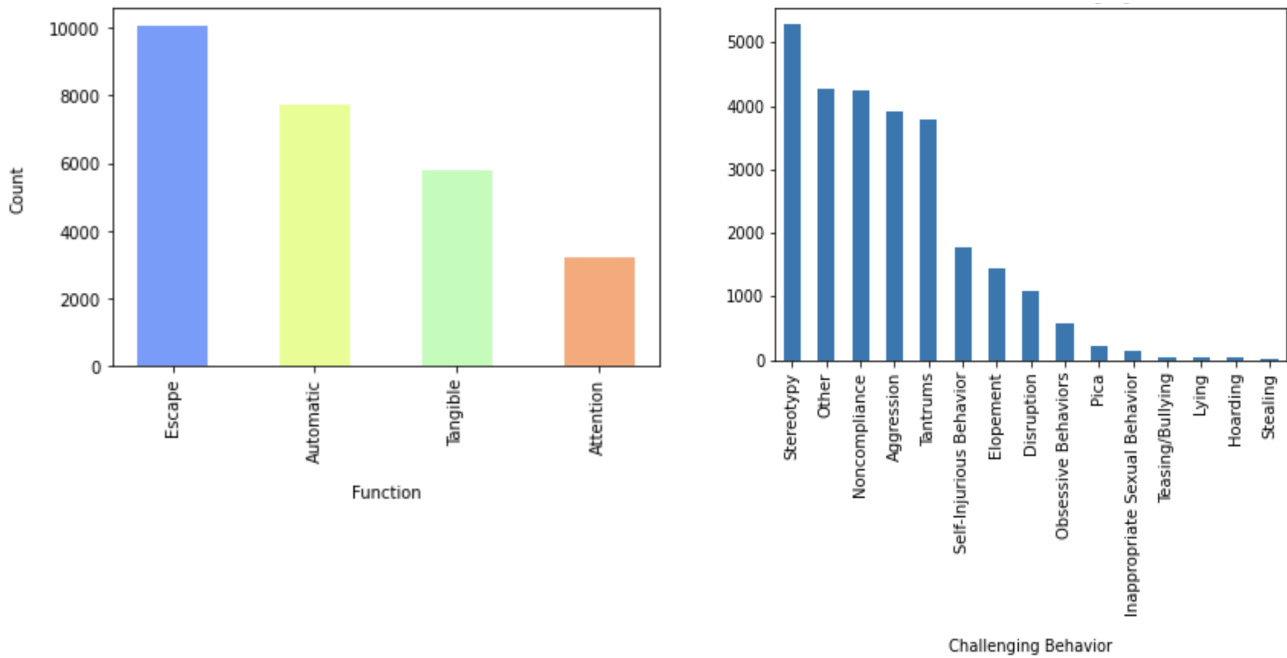


Fig 3.1 Total Count of Each Function: Shows counts of behaviors in the uncleaned dataset.

Fig. 3.2 Total Count of Each Challenging Behavior: Shows counts of the four functions in the uncleaned dataset.

It is also of relevance to point out that the dataset is zero-inflated; approximately 86% of the dataset has zero counts. Some of these zeros may have been due to the occurrences of particular behavior and function pairs being infrequent and, thus, going undetected. It is also likely that because symptoms of ASD can come in many forms, most individuals will experience a high frequency of some function and behavior pairs and few to none of others. Thus, the zeros will be left as is [36].

Finally, this study does not take into account the longitudinal effects of therapies on challenging behaviors. All data is entered during the span of treatment, and therefore, in the midst of data collection, a behavior that was once observed at a high frequency may end up being expressed to a lower degree over time.

The Skills™ dataset has 15 challenging behaviors: “aggression”, “disruption”, “elopement”, “hoarding”, “inappropriate sexual behavior”, “lying”, “noncompliance”, “obsessive behaviors”, “pica”, “self-injurious behavior”, “stealing”, “stereotypy”, “tantrums”, “teasing or bullying”, and “other”. For this study, “other” will be taken out of the data as this could be a variety of challenging behaviors unrelated to each other and will not allow us the ability to draw clear conclusions. Additionally, this study will specifically target the top eight most frequently occurring challenging behaviors. Data pre-processing removed any individual who exhibited behaviors outside of the top eight challenging behaviors.

Lastly, individuals with only one unique expression of a function and behavior pair were removed. For example, if an individual’s profile exclusively had instances of automatic stereotypy, but did not have any other function and behavior pair recorded, they would be removed from the dataset. This research focused on the co-occurrences of multiple functions and challenging behavior pairs.

After taking these initial preprocessing steps, a sample of $N = 1,416$ individuals was left for the cluster analysis. The finalized subset of challenging behaviors were aggression, disruption, elopement, noncompliance, obsessive behaviors, self-injurious behavior, stereotypy, and tantrums. In addition to the eight behaviors, all four functions- attention, automatic, escape, and tangible- were left in our dataset.

Table II shows the top 15 frequently occurring function and behavior pairs. The percent of occurrence indicates the frequency in which patients exhibited this function and behavior pair at least once. It can be immediately noticed that the dataset is heavily dominated by a handful of pairings such as automatic stereotypy, escape noncompliance, tangible tantrums, and escape tantrums.

TABLE II

TOP 15 MOST FREQUENT UNIQUE FUNCTION AND BEHAVIOR PAIRS AND PERCENT OF OCCURRENCE

Function	Behavior	% Occurrence
Automatic	Stereotypy	61.200%
Escape	Noncompliance	60.700%
Tangible	Tantrums	42.700%
Escape	Tantrums	40.700%
Escape	Aggression	33.000%
Tangible	Aggression	28.700%
Escape	Elopement	17.500%
Tangible	Noncompliance	17.100%
Escape	Self-Injurious Behaviors	10.900%
Tangible	Self-Injurious Behaviors	9.600%
Tangible	Elopement	9.00%
Escape	Disruption	8.400%
Automatic	Obsessive Behaviors	8.100%
Automatic	Self-Injurious Behaviors	7.400%
Attention	Tantrums	6.100%

Lists the top 15 functions of behaviors, “Function”, and challenging behaviors, “Behavior”, of all 32 unique pairs and displays the percent of occurrence among individuals in the dataset. The “% Occurrence” refers to the number of individuals who exhibited this behavior and function pair at least one time.

TABLE III

THE TOP 15 CO-OCCURRING FUNCTION AND BEHAVIOR PAIRS.

Pair 1	Pair 2	% Co-Occurrence
Automatic Stereotypy	Escape Noncompliance	39.55%
Automatic Stereotypy	Tangible Tantrums	25.64%
Automatic Stereotypy	Escape Tantrums	24.93%
Escape Tantrums	Tangible Tantrums	24.58%
Escape Noncompliance	Escape Tantrums	24.58%
Escape Noncompliance	Tangible Tantrums	24.52%
Automatic Stereotypy	Escape Aggression	20.20%
Escape Aggression	Escape Noncompliance	19.99%
Escape Aggression	Tangible Aggression	16.88%
Automatic Stereotypy	Tangible Aggression	16.81%
Tangible Aggression	Escape Noncompliance	16.81%
Escape Aggression	Escape Tantrums	16.17%
Escape Noncompliance	Tangible Noncompliance	15.04%
Tangible Aggression	Tangible Tantrums	14.34%
Escape Aggression	Tangible Tantrums	13.63%

Lists the top 15 unique function and behavior pairs and their co-occurrence with another unique function and behavior pair. % Co-Occurrence refers to the percent of individuals who had this pairing at least once.

Table III shows the top 15 co-occurrences between two unique behavior and function pairs. It is not surprising that we see the first several co-occurrences primarily containing automatic stereotypy, escape noncompliance, and escape and tangible tantrums, as these were the most common function and behavior pairs seen in Table II.

The finalized dataset was an aggregated set of 32 dimensional vectors, where each feature was one of the 32 function and behavior pairs and each vector, or row, represented an individual with ASD. The information stored is a tally, or count, of the number of times each function and behavior pair occurred for an individual. The data was then normalized using a simple approach: dividing each feature count by the sum of all counts in a patient's profile. Normalization left us with a 32 dimensional relative frequency vector for each individual, where the value of all features sum to 1, or 100%.

Other common normalization or transformation techniques for this type of data are L2 normalization of the count data or centred log-ratio [37] or isotropic log-ratio transform [38] over the relative frequencies to bring them from the simplex into Euclidean space. These normalizations and transformations ultimately left us with similar clustering results. Therefore, for simplicity, we will use the relative frequency vectors for the results. This method has the added benefit of applying easily interpretable visualization. On top of this, other studies clustering ASD data have normalized the data in this manner [7]. As this study will serve, in part, as an extension to some of these, it is applicable to normalize the data similarly.

Chapter 4

Methods

This chapter will cover the methods that were taken to perform the cluster analysis.

As a first step, we utilized self-organizing maps (SOMs). SOMs are a type of neural network for unsupervised learning that can be used for dimensionality reduction, data visualization, and clustering [39]. Due to the limited size of the dataset, SOMs were not an ideal fit for clustering analysis as it ultimately led to too much variance and clusters that were too small and specific to our dataset.

Variance, in machine learning, refers to a model being overfit to a dataset and lacking the ability to generalize well to new data. Overfitting can give a false sense of confidence in a model that ends up performing poorly when applied outside of the dataset it was trained on, especially when the dataset is relatively small. For these reasons, the SOM will be used to indicate whether or not there is topology in the data that can provide meaningful sub-groups, but it will not be used for clustering.

Finally, the normalized data will be fed into a k-means model to identify clusters.

Sections 4.1 and 4.2 respectively present a breakdown and mathematical summary of SOMs and k-means. Chapter 5 will then discuss the results of these methods.

4.1 Self-Organizing Maps (SOM)

We can define a self-organizing map as an artificial neural network (ANN) that takes high dimensional data as input and maps it to a two-dimensional grid, organizing it in such a way that preserves the topological structure of a dataset. SOMs differ from other ANNs because they do not apply error-correction learning such as backpropagation with gradient descent. Instead, they use competitive learning where the nodes of neural networks compete with each other to be a part of a subset of data [39]. Thus, the subsetting of data onto a two-dimensional, organized output map makes SOMs a highly effective tool for both data visualization and cluster analysis [39][40][41].

Here, we present the steps of SOMs in the way that they are used in this study:

- 1) Position the grid's neurons randomly within the data space.

A guide for determining the number of neurons, m , is to start with

$m = 5\sqrt{n}$, with n being the sample size, and depending on the application, iteratively increasing until the topology stabilizes [39].

- 2) Randomly select a data point and find the Best Matching Unit (BMU). The BMU is the closest neuron to the data point.
- 3) The BMU moves some distance closer to the currently selected data point. The distance it moves is determined by the learning rate.
- 4) Move the BMUs neighbors, as defined by a radius around the BMU, closer to the data point, where the closest neighbors move the most and the farthest neighbors move the least.

- 5) Iterate through Steps 1 to 4 until positions of neurons have been stabilized, continuously updating the radius and the learning rate.

The suggested number of iterations is, at minimum, ten times the size of the sample size.

4.2 K-means

K-means is an unsupervised machine learning algorithm used to find meaningful relationships in data via natural groupings, or clusters. The data is grouped into a pre-defined number of clusters, k , where the data in a single cluster will have similar features.

The algorithm works by randomly selecting a centroid as a starting point for each k and then iteratively performing calculations until the position of the centroids are optimized. Optimization occurs when clusters converge, or otherwise, stop changing with iterations [41].

If we have n samples in an m -dimensional space, $\vec{x}_i \in \mathcal{R}^m, i = 1, \dots, n$

$$X = \{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n\} \{C_1, \dots, C_k\}$$

K-means distinguishes the space, X , into a set of k clusters.

$$\{C_1, \dots, C_k\}$$

So that each sample belongs to a cluster $C_1 \cup C_2 \cup \dots \cup C_k = X$,

and clusters do not intersect.

The goal is to solve the following optimization problem, where we find the division into k clusters that minimizes the overall within-cluster distances across all clusters:

$$\min_{C_1, \dots, C_k} \left\{ \sum_{p=1}^k \frac{1}{|C_p|} \sum_{i, j \in C_p} (\vec{x}_i - \vec{x}_j)^2 \right\}$$

The steps of k-means goes as follows:

- 1) Choose k , the total number clusters in the data.
- 2) Randomly assign each sample to a cluster, or centroid.
- 3) Compute each cluster's centroid:

$$\vec{x}_{C_p}^- = \frac{1}{|C_p|} \sum_{i \in C_p} \vec{x}_i$$

- 4) Assign each sample to a cluster using Euclidean distance to find the closest centroid to the point.
- 5) Repeat steps 3 and 4 until the clusters stop changing or the maximum number of iterations is reached.

K-means has many advantages, including being conceptually easy to understand, having simple and regularly available implementations, good scalability, and the need to only tune a single hyper-parameter, k . However, finding the optimal k is one of the drawbacks. Due to the fact that k-means is an unsupervised algorithm, we must determine the value of k without a priori knowledge of how many clusters there may be [42]. Here, we will utilize the elbow method to make this distinction.

4.2.1 Elbow Method

The elbow method is one of the most common techniques for choosing the number of clusters. The method works by iteratively running k-means over a set of possible k values. We plot the number of clusters, k , along the x-axis, and the inertia along the y-axis. The inertia is the sum of squared distances of each data point to the center of its nearest cluster.

The plot forms what we call an “elbow”, where the curve bends similar to the shape of an elbow when an arm is bent. As k increases, the SSE gets smaller, and this “elbow” on the plot is where the SSE begins to show diminishing returns, representing the optimal number of clusters.

It is important to note that selecting k according to the elbow plot, as well as all other methods for selecting k , is subjective. Different analysts may look at the plot and choose a similar but different k value, especially when the elbow on the plot is not that distinguishable [42].

Chapter 5

Results

The self-organizing maps visualizations and k-means clustering results are presented in this chapter.

In Section 5.1, we display results from the SOM, and in Section 5.2, we show visualizations and results from our k-means cluster analysis. Section 5.2 will also present additional visualizations from the SOM to provide a more in-depth look at the k-means generated clusters.

5.1 Self-Organizing Maps Results

The self-organizing map was used to visualize if the data could be divided into meaningful subgroups as well as provide more context into the way individual features drive those groupings.

For our implementation, the number of neurons were chosen using the formula $m = 5\sqrt{n}$, where n is the number of samples and m is the number of neurons, a guideline initially suggested by Kohonen [39]. Ultimately, the number of neurons was increased to a 32x32 grid space, when the topology of the node grid had stabilized. We ran the SOM over 30,000 iterations. This was

per suggestion of [39], who recommends running the algorithm iteratively at least 10 times the sample size. We also set periodic boundary conditions when initializing the SOM.

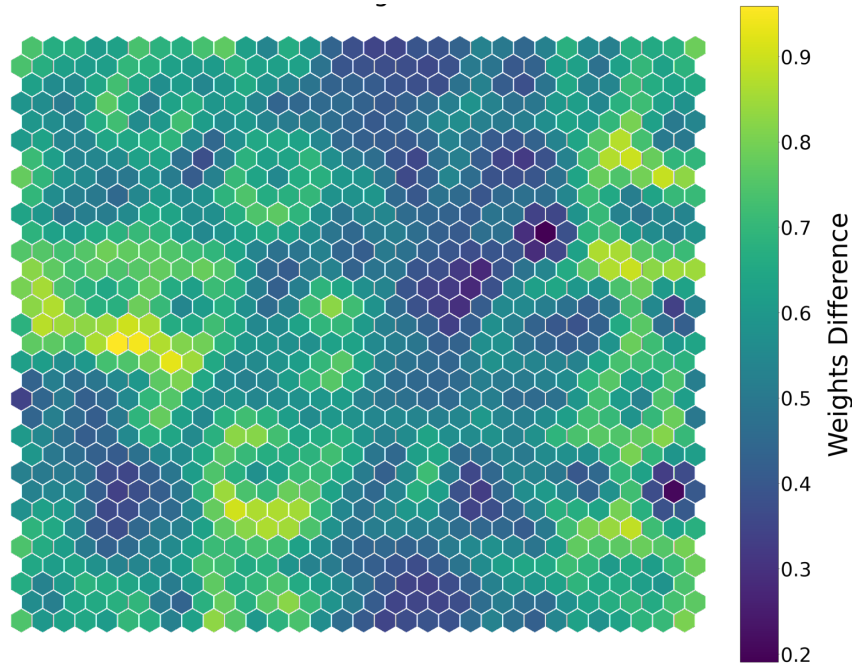


Fig. 5.1 SOM Nodes Grid with Weight Difference: Graph showing the nodes grid with weight difference to visualize the topology of the entire sample space.

Figure 5.1 is a hexagonal nodes grid map with weight differences. The darker nodes indicate similarities of surrounding nodes, or otherwise, what is likely a cluster. The light nodes indicate large differences or separation between surrounding nodes. The fact that we see darker structures surrounded by lighter indicates that there are close similarities between certain vectors, or patients, as well as large enough differences to separate them from other groupings.

The variability in what appears to be cluster size is likely due to the fact that SOMs are able to give a highly detailed view of the topology, picking up on the smallest differences in data as SOMs are sensitive to both small differences and outliers. Even a single node can be recognized as a different cluster entirely. Thus, this result is not unexpected.

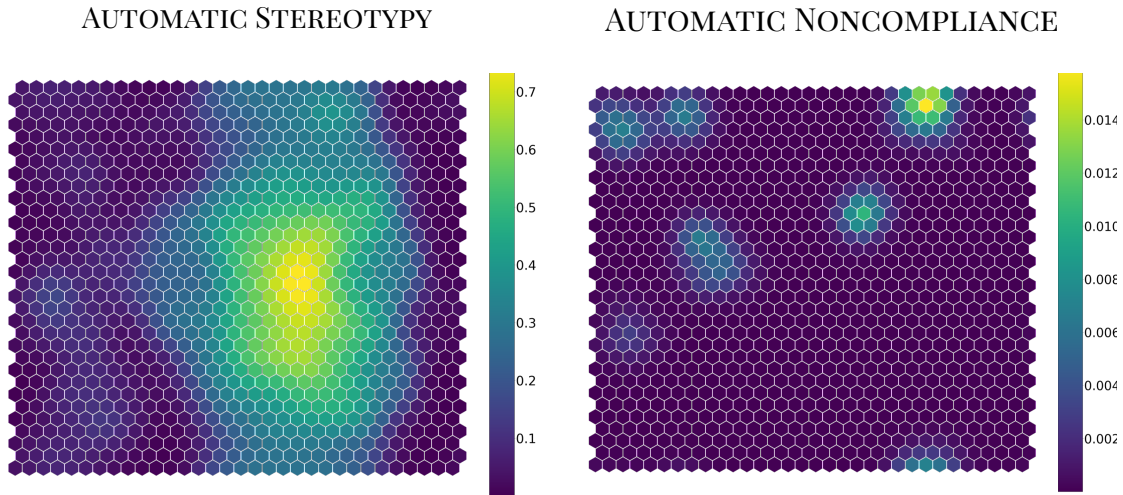


Fig 5.2 Feature Network of Nodes: Two examples of features (function and behavior pairs) as a network of nodes. Maps are colored according to a single feature and organized according to the distance between each node and its neighbors. The feature weights are on a different scale in each graph.

The plots in Figure 5.2 show individual features as a network of nodes. When the mapping of nodes is highlighted, it means that the feature is prominent in that region. The features that are shown on these maps are automatic stereotypy and automatic noncompliance.

We can see in Figure 5.2 that automatic stereotypy is a frequently expressed behavior and function pair, evident by the large highlighted regions on its respective map. This is also shown to be true in Table II. The map also indicates that the presence of this feature is driving the nodes in figure 5.2 topologically closer to one another when present in a patient's profile and, thus, having a heavy influence on which cluster a patient belongs to. We see this by how highly concentrated the highlighted region is and the wide space that it spans.

In contrast, automatic noncompliance appears less often and more dispersed on its nodes map. The rare appearance on the grid and the large distance between highlighted regions indicate that the presence of the feature holds little weight in determining the cluster that a patient belongs to.

Figure 5.2 is solely meant to provide some additional visual detail into how self-organizing maps use features to drive the topological mapping. Now that an organized topological structure has been confirmed in the data, Section 5.2 will not present the results of the more generalized clusters produced by k-means. Additionally, a figure similar to Figure 5.2 will be displayed to provide more contextual depth to the presented k-means clustering results.

5.2 K-Means Results

The first step in implementing k-means is determining the number of clusters, k . Figure 5.3 shows the elbow plot with which we utilize for finding k . The optimal number of clusters based on this plot is approximately eight, where we start to see diminishing returns.

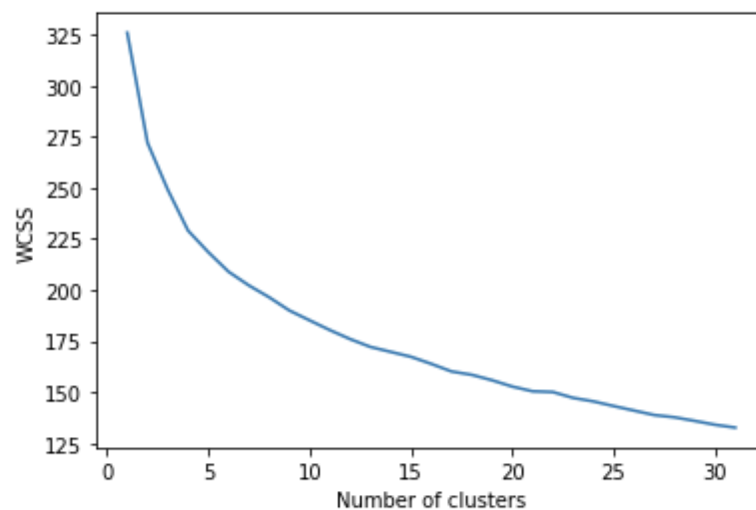


Fig 5.3 Elbow Plot: Elbow plot to determine the number of k-means clusters.

Figure 5.4 displays polar bar plots as a visualization tool to see the contents of each cluster. Figure 5.5 shows Cluster 2 individually in order to provide a better view of its unique and more complex portfolio.

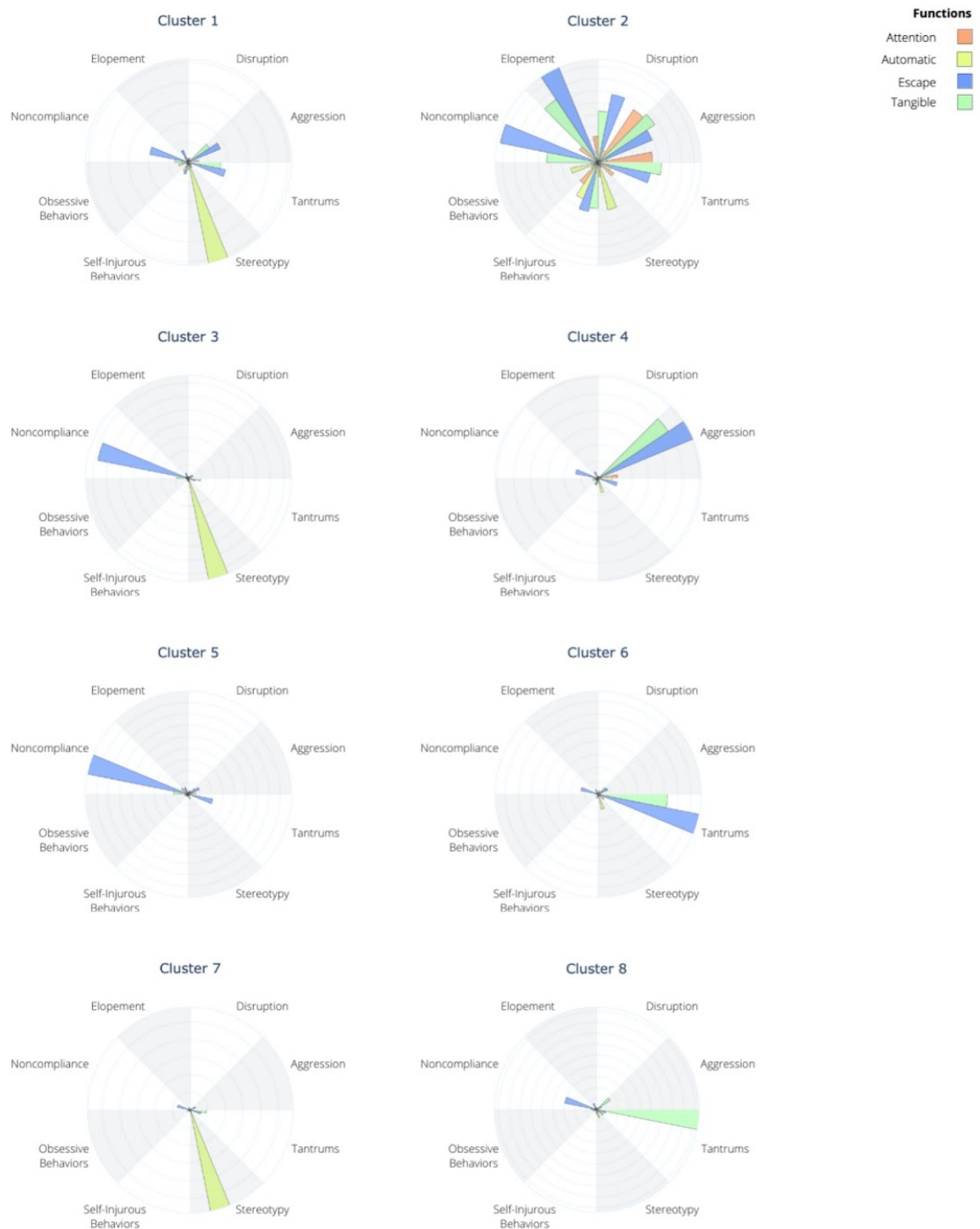


Fig 5.4 Polar Bar Charts for Eight K-Means Clusters: Polar bar plots that display the average frequency of each behavior and function pair for 8 clusters. Functions are color coordinated as described by the legend. Behaviors are labeled around each plot. The maximum value of a bar plot is equal to the frequency of patients in the most common feature of a specific cluster.

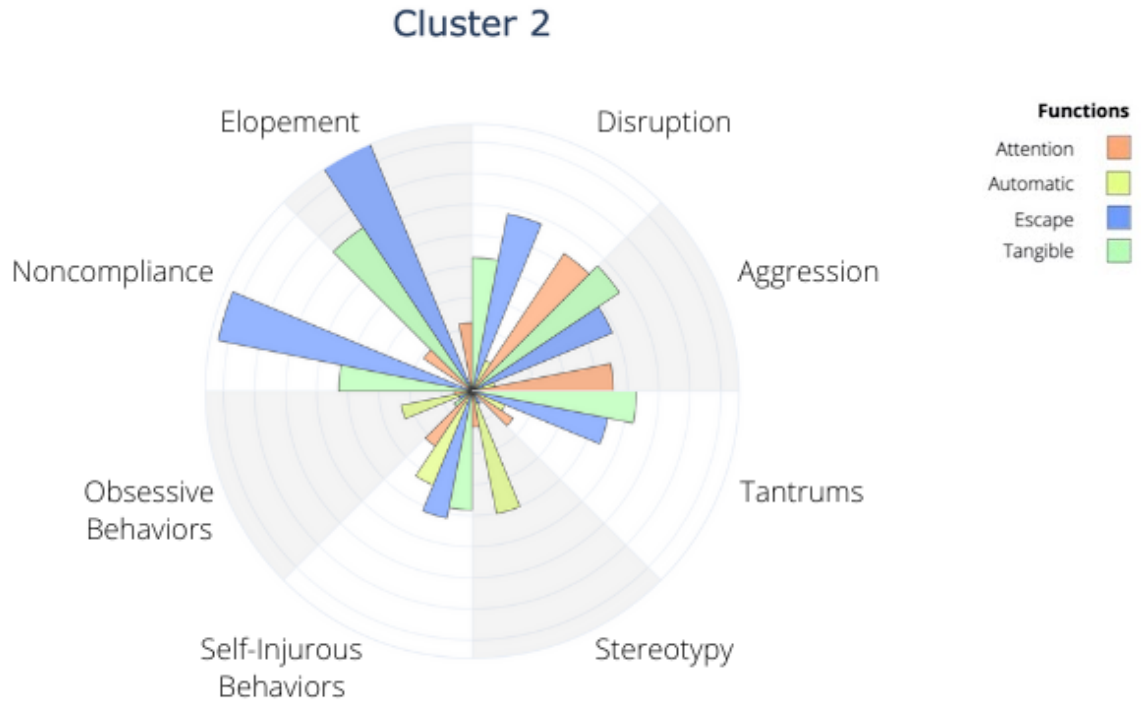


Fig 5.5 Polar Bar Chart for Cluster 2: Functions are color coordinated as described by the legend. Behaviors are labeled around each plot.

The size of the clusters range from 128 to 300 patients. The most prominent bar in each polar bar plot represents the most commonly occurring feature in the cluster. More specifically, all clusters, except for Clusters 2 and 4, have at least one feature that is present in every vector in the cluster. In the case of Cluster 3, the top two most frequent features, automatic stereotypy and escape noncompliance, are both present in every patient in the cluster. It should also be noted that there is no crossover between clusters in k-means. Every patient appears in exactly one cluster.

This cluster analysis shows some similar results to prior research [7][8]. In particular, we see that Clusters 4 through 8 are heavily driven by one challenging behavior. However, although driven by one behavior, Clusters 4 and 6 have two frequent functions driving the behaviors. Those functions are escape and tangible. We also note that in any cluster where there is more than one unique dominant function and behavior pair, there is more than one dominant function. Additionally, we note that there is a common co-occurrence of automatic stereotypy and escape

noncompliance, with the pair of them together dominating Cluster 3. Referring back to the results in Table II and Table III, we see that this clustering outcome is not unexpected.

Next, it can be seen that Cluster 1, while driven primarily by stereotypy, has a semi-frequent expression of other behaviors as well, specifically with escape and tangible functions.

Most notable is Cluster 2. Not only does Cluster 2 stand out because of the presence of multiple challenging behaviors, but also by the variety of functions in the cluster. There is a much higher frequency of attention as a function in this cluster than there is in any other cluster. In fact, this is the only cluster in which attention was a prominent function. This functional variety may indicate that the more behaviors that are in a patient's profile, the more full an individual's functional profile will be, or vice versa.

We also note that Cluster 2 is the only cluster where frequent expressions of SIB or disruptive behaviors are observed. SIB is expressed with all four of the behavioral functions, and disruption is prominently expressed with three functions. Thus, the presence of one of these behaviors in an individual with ASD may be able to serve as a predictive factor of a patient having a full profile of ASD functions.

Cluster 2 stands out among the rest, showing almost all challenging behaviors as being expressed to a noticeable degree, as well as having several functions of each behavior. To have a better look at the contents of this cluster, Figure 5.6 and 5.7 display where the SOMs mapped the four functions of SIB and the four functions of disruption onto a network of nodes. We focus on these two challenging behaviors because of their uniqueness to Cluster 2 as well as their many co-expressed functions.

Figure 5.6 shows that all four functions of SIB occur in a similar location, mapped to the same nodes. This validates what Cluster 2 tells us, which is that all four functions seem to be expressed together when an individual displays SIB, aside from automatic which appears to have some independence. We see similar patterns with disruptive behavior in Figure 5.7. All four

functions of disruptive behavior are mapped to the same location on the node grid, indicating that all four co-occur in individuals when their expression of disruptive behaviors is frequent.

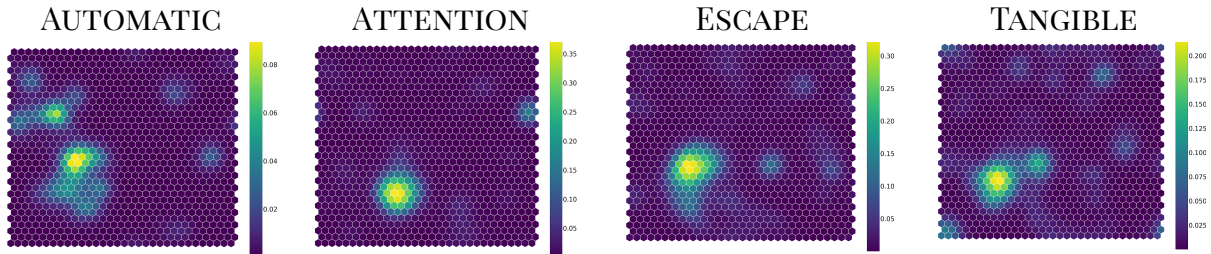


Fig. 5.6 Four Feature Network of Nodes for SIB:

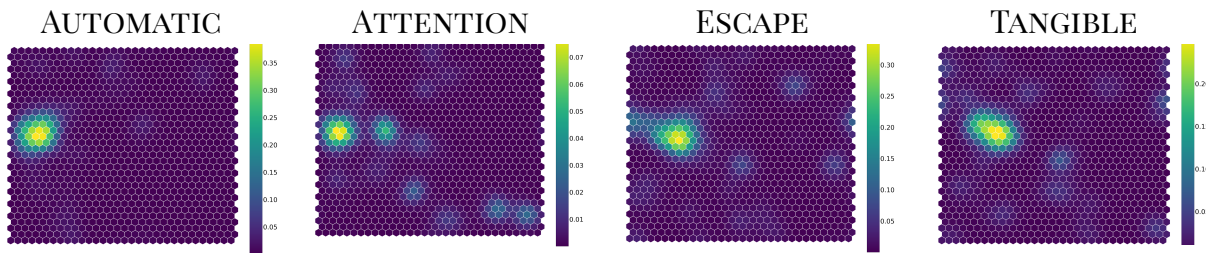


Fig. 5.7 Four Feature Network of Nodes for Disruptive Behaviors: A SOM node grid for each of the 4 functions of SIB and disruptive behaviors, respectively. Maps are colored according to feature prominence and organized according to the distance between each node and its neighbors. Weight scales are unique to each graph.

We present these node maps to serve as validation of what we observed with k-means, being that the presence of SIB or disruption is likely an indicator that a person will have all four functions as driving forces of these specific behaviors. In the following chapter, we will discuss further what these findings mean in terms of application.

Chapter 6

Discussion

Anywhere between 45 and 54 individuals are affected by ASD, and the diagnosis is becoming increasingly common [3]. It is imperative, then, to understand what to do with the diagnosis once it has been made. One of the most effective steps in this process, especially in the diagnosis of children, is to undergo early intervention and ABA therapy [12]. It is recognized that symptoms of ASD can be present as early as before the age of two; therefore, not only is early intervention effective when applicable, but it is also very possible [30].

People who go through functional based therapy have better outcomes in comparison to those whose therapy does not include functional analysis [5][6]. When therapy is targeted at early ages, these outcomes can, in the best cases, lead to challenging behaviors going virtually undetected [14][19]. Therefore, correctly identifying functions of behavior is crucial.

Identifying a behavioral function is not always obvious, though. It often takes many tools and resources to make the distinction efficient and accurate [12][18][20]. Furthermore, the younger a child is, the harder the identification of behavioral functions becomes [28][29].

Overall, it is clear that methods of identifying a patient's functional profile needs improvement. The function and behavior frequency clusters found in this research could serve, then, as a contributing factor in helping determine an individual's functional profile.

Results of this research find eight different clusters. Many patients were grouped together based on a single prominent challenging behavior with either one or two subsequent functions. Then, we saw that as a patient's behavioral profile grows larger, so does their functional profile. This indicates that if a person's behavioral profile is large, more functions will likely need to be targeted in therapy.

Next, we see that the presence of attention emerges in a person's profile who has a large variety of challenging behaviors. Additionally, we see that when attention is present in a profile, it is a prominent function of several behaviors and, likely, not exclusive to a single behavior. Thus, if there is evidence of attention emerging in a patient's functional profile, therapy regimens can be adjusted to target it more aggressively and as a function of several challenging behaviors.

Lastly, we visualize that while challenging behaviors such as stereotypy, noncompliance, and tantrums often appear with only one or two underlying functions, disruptive behaviors and SIB usually coincide with the presence of three or four prominent functions. This means that if one of these behaviors is evident in an individual with ASD, all four functions may need to be addressed in therapy in order to minimize the behavior most effectively.

The results of this research may be able to serve as its own technique in identifying likely occurrences of function and behavior pairs and, ultimately, lead to better treatment plans. The potential could extend to allowing therapy to target these functions early on and before the expression of certain challenging behaviors become evident.

Chapter 7

Future Works

This research involved employing SOMs as a means of data visualization to view the topological structure of the data. We saw in figure 5.1 that the structure of this specific data is partitioned into several large and small groupings. With access to ASD data becoming increasingly thorough, we could use different clustering techniques and deep learning to find more specific clusters without introducing too much variance to our model, ultimately, providing more insight into behavioral and functional profiles. To further support this as a next step, results from [43] suggest that ASD research should expand more into deep learning and neural networks.

Now that structure and groupings among patients based on their function and behavior profile have been identified, a natural next step would be to use supervised ML models to predict the presence of particular functions in a patient's profile based on the combination of their challenging behaviors. This could also be extended to predicting the frequency of expected function and behavior pairs.

Finally, another continuation would be to perform a cluster analysis on behavioral functions alone. Research by [7] showed a cluster analysis of challenging behaviors. This paper demonstrates a cluster analysis of function and behavior pairs. Thus, a cluster analysis of functions alone would serve as a completion of these studies.

The purpose of this study was to see what insight unsupervised machine learning techniques could provide on functional profiles. As data becomes increasingly available [21], we hope that the results of this research will be used as a catalyst for future studies to leverage machine learning for the understanding of the relationship between functions and challenging behaviors commonly seen in ASD.

Chapter 8

Conclusion

Autism Spectrum Disorder treatment is highly individualized. Treatment geared towards both challenging behaviors and functions of behaviors has proven to be highly effective. Many machine learning studies have leveraged clinical data to improve the diagnosis of ASD, and now, ML research is being done to help pave a path in understanding what to do with this diagnosis once it has been made. This research serves as part of the latter.

This thesis aimed to find general groupings of patients based on their frequent expression of function and behavior pairs. To do this, we first used self-organizing maps to identify possible structure amongst over one thousand ASD patient profiles. Once structure was confirmed, patient profiles were put into a k-means model where eight clusters were identified. Here, we saw that, in general, patients with a higher number of challenging behaviors generally have a higher number of functions. In addition, some challenging behaviors tend to be expressed in congruence with all four challenging behaviors while others are usually expressed as a product of only a single function.

Developing a comprehensive understanding of individuals with ASD's functional and behavioral profile is imperative to the continual improvement of therapy and patient outcomes. Research such as the work done in this thesis provides a further step into the much needed understanding of and identification of behavioral functions. This type of work paves the way into understanding how an individual's entire functional and behavioral profile may develop over time, allowing for the improvement of early intervention and therapy. It is hoped that the results of this study will provide benefit to functional analysis and a more successful therapy plan.

For those working in ASD research, it should be of utmost importance to do work that improves the lives of individuals living with autism. Data driven and ML approaches provide a new, modern way to do so. Applying this type of work specifically to functions and challenging behaviors is crucial, as this can lead to improvement of therapy techniques, individualization, and, ultimately, better outcomes. This all leads to the ultimate goal-- a better quality of life for individuals and families living with ASD.

References

- [1] E. Emerson, *Challenging behaviour: Analysis and intervention in people with learning disabilities*. Cambridge, UK: Cambridge University Press, 1995.
- [2] *Diagnostic and statistical manual of mental disorders (5th ed)*, 5th ed. Washington DC: American Psychiatric Association, 2013.
- [3] J. Hoag, “Exploring the impact of challenging behaviors on treatment efficacy in autism spectrum disorder,” *Advances in Neurodevelopmental Disorders*, vol. 2.2 , pp. 206–215, 2018.
- [4] D. R. Dixon, T. Vogel, and J. Tarbox, “A brief history of functional analysis and applied behavior analysis,” *Functional Assessment for Challenging Behaviors*, pp. 3–24, 2012.
- [5] R. Didden, H. Korzilius, W. van Oorsouw, and P. Sturmey, “Behavioral treatment of challenging behaviors in individuals with mild mental retardation: Meta-analysis of single-subject research,” *American Journal on Mental Retardation*, vol. 111, no. 4, pp. 290–298, 2006.
- [6] R. H. Horner, E. G. Carr, P. S. Strain, A. W. Todd, and H. K. Reed, “Problem behavior interventions for young children with autism: A research synthesis. ,” *Journal of Autism and Developmental Disorders*, vol. 32, no. 5, pp. 424–446, 2002.
- [7] E. Stevens, A. Atchison, L. Stevens, E. Hong, D. Granpeesheh, D. Dixon, and E. Linstead, “A cluster analysis of challenging behaviors in autism spectrum disorder,” *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2017.
- [8] E. Stevens, D. R. Dixon, M. N. Novack, D. Granpeesheh, T. Smith, and E. Linstead,

“Identification and analysis of behavioral phenotypes in autism spectrum disorder via unsupervised machine learning,” *International Journal of Medical Informatics*, vol. 129, pp. 29–36, 2019.

[9] F. Charles Mace, T. J. Page, M. T. Ivancic, and S. O'Brien, “Analysis of environmental determinants of aggression and disruption in mentally retarded children,” *Applied Research in Mental Retardation*, vol. 7, no. 2, pp. 203–221, 1986.

[10] J. L. Matson and M. Nebel-Schwalm, “Assessing challenging behaviors in children with autism spectrum disorders: A review,” *Research in Developmental Disabilities*, vol. 28, pp. 567–579, 2007.

[11] R. Hiebert, G. L. Martin, C. T. Yu, J. R. Thorsteinsson, and T. L. Martin, “Predicting compliance of children with and without developmental delay,” *Canadian Journal of Behavioural Science / Revue canadienne des sciences du comportement*, vol. 41, no. 1, pp. 31–36, 2009.

[12] J. Jang, D. R. Dixon, J. Tarbox, and D. Granpeesheh, “Symptom severity and challenging behavior in children with asd,” *Research in Autism Spectrum Disorders*, vol. 5, no. 3, pp. 1028–1032, 2011.

[13] J. O. Cooper, T. E. Heron, and W. L. Heward, *Applied behavior analysis*. Harlow, England: Pearson, 2020.

[14] M. Horovitz, “Challenging behaviors,” *Comorbid Conditions in Individuals with Intellectual Disabilities*, pp. 27–53, 2015.

[15] J. McCarthy, C. Hemmings, E. Kravariti, K. Dworzynski, G. Holt, N. Bouras, and E. Tsakanikos, “Challenging behavior and co-morbid psychopathology in adults with intellectual disability and autism spectrum disorders,” *Research in Developmental Disabilities*, vol. 31, no. 2, pp. 362–366, 2010.

[16] O. Murphy, O. Healy, and G. Leader, “Risk factors for CHALLENGING behaviors AMONG 157 children with autism spectrum disorder in Ireland,” *Research in Autism Spectrum*

Disorders, vol. 3, no. 2, pp. 474–482, 2009.

[17] T. R. Paclawskyj, J. L. Matson, K. S. Rush, Y. Smalls, and T. R. Vollmer, “Questions about behavioral function (QABF):” *Research in Developmental Disabilities*, vol. 21, no. 3, pp. 223–229, 2000.

[18] J. L. Matson, J. W. Bamburg, K. E. Cherry, and T. R. Paclawskyj, “A validity study on the questions about behavioral function (qabf) scale: Predicting treatment success for self-injury, aggression, and stereotypies,” *Research in Developmental Disabilities*, vol. 20, no. 2, pp. 163–175, 1999.

[19] L. P. Hagopian, D. M. Wilson, and D. A. Wilder, “Assessment and treatment of problem behavior maintained by escape from attention and access to tangible items,” *Journal of Applied Behavior Analysis*, vol. 34, pp. 229–232, 2001.

[20] S. S. Hall, “Comparing descriptive, experimental and informant-based assessments of problem behaviors,” *Research in Developmental Disabilities*, vol. 26, no. 6, pp. 514–526, 2005.

[21] E. Linstead, D. R. Dixon, R. French, D. Granpeesheh, H. Adams, R. German, A. Powell, E. Stevens, J. Tarbox, and J. Kornack, “Intensity and learning outcomes in the treatment of children with autism spectrum disorder,” *Behavior Modification*, vol. 41, no. 2, pp. 229–252, 2016.

[22] E. Linstead, D. R. Dixon, E. Hong, C. O. Burns, R. French, M. N. Novack, and D. Granpeesheh, “An evaluation of the effects of intensity and duration on outcomes across treatment domains for children with autism spectrum disorder,” *Translational Psychiatry*, vol. 7, no. 9, 2017.

[23] E. Linstead, D. R. Dixon, R. French, D. Granpeesheh, H. Adams, R. German, A. Powell, E. Stevens, J. Tarbox, and J. Kornack, “Intensity and learning outcomes in the treatment of children with autism spectrum disorder,” *Behavior Modification*, vol. 41, no. 2, pp. 229–252, 2016.

[24] A. J. Griffiths, A. H. Hanson, C. M. Giannantonio, S. K. Mathur, K. Hyde, and E. Linstead, “Developing employment environments where individuals with ASD thrive: Using machine

learning to Explore Employer policies and practices,” *Brain Sciences*, vol. 10, no. 9, p. 632, 2020.

[25] M. Khodatars, A. Shoeibi, N. Ghassemi, M. Jafari, A. Khadem, and D. Sadeghi, “Deep Learning for Neuroimaging-based Diagnosis and Rehabilitation of Autism Spectrum Disorder: A Review.,” *arXiv preprint*, vol. 2007, 2020.

[26] Y. Zhou, F. Yu, and T. Duong, “Multiparametric MRI characterization and prediction in autism spectrum disorder using graph theory and machine learning,” *PLoS ONE*, vol. 9, no. 6, 2014.

[27] R. Anden and E. Linstead, “Predicting eye movement and FIXATION patterns on scenic images using machine learning for children with autism spectrum disorder,” *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2020.

[28] S. Olof Dahlgren and C. Gillberg, “Symptoms in the first two years of life,” *European Archives of Psychiatry and Neurological Sciences*, vol. 238, no. 3, pp. 169–174, 1989.

[29] G. Dawson, J. Osterling, A. N. Meltzoff, and P. Kuhl, “Case study of the development of an infant with autism from birth to two years of age,” *Journal of Applied Developmental Psychology*, vol. 21, no. 3, pp. 299–313, 2000.

[30] L. Zwaigenbaum, S. Bryson, and N. Garon, “Early identification of autism spectrum disorders,” *Behavioural Brain Research*, vol. 251, pp. 133–146, 2013.

[31] T. Vargason, R. E. Frye, D. L. McGuinness, and J. Hahn, “Clustering of co-occurring conditions in autism spectrum disorder during early childhood: A retrospective analysis of medical claims data,” *Autism Research*, vol. 12, no. 8, pp. 1272–1285, 2019.

[32] J. Gardner-Hoag, M. Novack, C. Parlett-Pelleriti, E. Stevens, D. Dixon, and E. Linstead, “Unsupervised machine learning for identifying challenging behavior profiles to explore cluster-based treatment efficacy in children with autism spectrum disorder: Retrospective data analysis study,” *JMIR Medical Informatics*, vol. 9, no. 6, 2021.

[33] T. Obafemi-Ajayi, D. Lam, T. N. Takahashi, S. Kanne, and D. Wunsch, “Sorting the

phenotypic heterogeneity of autism spectrum disorders: A hierarchical clustering model,” *2015 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, 2015.

[34] J. L. Matson and M. Nebel-Schwalm, “Assessing challenging behaviors in children with autism spectrum disorders: A review,” *Research in Developmental Disabilities*, vol. 28, no. 6, pp. 567–579, 2007.

[35] J. L. Lipschultz and D. A. Wilder, “Behavioral assessment and treatment Of noncompliance: A review of the literature,” *Education and Treatment of Children*, vol. 40, no. 2, pp. 263–298, 2017.

[36] Z.-Z. Tang and G. Chen, “Zero-inflated generalized Dirichlet MULTINOMIAL regression model for MICROBIOME Compositional data analysis,” *Biostatistics*, vol. 20, no. 4, pp. 698–713, 2018.

[37] J. Aitchison, “The statistical analysis of compositional data,” *Journal of the Royal Statistical Society*, vol. 44, no. 2, pp. 139–160, 1982.

[38] J. J. Egozcue, V. Pawlowsky-Glahn, G. Mateu-Figueras, and C. Barceló-Vidal, “Isometric logratio transformations for compositional data analysis,” *Mathematical Geology*, vol. 35, no. 3, pp. 279–300, 2003.

[39] T. Kohonen, “Self-organizing maps,” *Springer Series in Information Sciences*, 2001.

[40] S. Kaski, “Data exploration using self-organizing maps.,” *Mathematics, computing and management in engineering*, 1997.

[41] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, “An efficient k-means clustering algorithm: Analysis and implementation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 881–892, 2002.

[42] Pham, Duc Truong, Stefan S. Dimov, and Chi D. Nguyen. "Selection of K in K-means clustering." *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of*

Mechanical Engineering Science 219.1 (2005): 103-119.

[43] E. Linstead, R. German, D. Dixon, D. Granpeesheh, M. Novack, and A. Powell, “An application of neural networks to predicting mastery of learning outcomes in the treatment of autism spectrum disorder,” *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA)*, 2015.

Appendix A

Source Code Samples

```

# Preprocessing, Aggregating, and Normalizing Data from CARDS Skills Dataset

# 8 most common behaviors that we'll use for the research
behaviors = ['Noncompliance', 'Tantrums', 'Aggression', 'Stereotypy', 'Self-Injurious Behavior',
'Elopement', 'Disruption', 'Obsessive Behaviors']

#keeping only ids with top 8 behaviors
keep_id = np.setdiff1d(np.unique(data['id']),np.unique(ids_to_remove))
data = data[data['id'].isin(np.unique(keep_id))].reset_index()

#creating feautre columns for behavior and function pair and getting the counts
data['behavior_function'] = data.behavior_description + " " + data.function_description
y = pd.get_dummies(data.behavior_function)
data = data.join(y)

data_function_count = data.function_count
data = data.drop(columns = ['index', 'behavior_function'])
#we'll have a 32 vector matrix of counts
for index, row in data.iterrows():
    for j in range(len(row)):
        if data.iloc[index,j] == 1:
            data.iloc[index,j] = data_function_count[index]
data = data.drop(columns = [ 'behavior_description', 'function_description', 'function_count'])

# aggregating data sum counts for 1 row per id, each id is a 32 dim vector
data = data.groupby("id").sum()

#removing subjects with 0 or 1 unique features
non_zero_count = data.astype(bool).sum(axis=1)
single_entry_ids = []
for i, v in pd.DataFrame(non_zero_count).iterrows():
    if v[0] == 1:
        single_entry_ids.append(i)
data = data.drop(single_entry_ids)

#normalize data
data_norm = data.div(data.sum(axis=1), axis=0)

```

Fig. A.1: Code for gathering SKILLS data, data preprocessing, aggregating counts, and data normalization

```

# Code for First Iteration of SOM with m, n, and epochs set according to Kohonen's suggestion
import SimpSOM as sps

#Finding initial starting points for base SOM model
m = 5*math.sqrt(data_norm.shape[0])
m = math.sqrt(m)
math.ceil(m)

#Build a network 20x20 with a weights format taken from the raw_data and activate Periodic Boundary
Conditions.
net = sps.somNet(math.ceil(m), math.ceil(m), data_norm, PBC=True)

#Train the network for 10000 epochs and with initial Learning rate of 0.01.
net.train(0.2, 10000)

#Print a map of the network nodes and colour them according to the first feature (column number 0) of
the dataset and then according to the distance between each node and its neighbours.
for i in range(0, 32):
    net.nodes_graph(colnum=i)

net.diff_graph()

#Cluster the datapoints according to the Quality Threshold algorithm.
clusters = net.cluster(data_norm, type='qthresh')

```

Fig. A.2: Code to generate SOM with parameters set to baseline parameters based on Kohonen recommendations.

```
#K-Means and Labeled Dataframe creation

dn_clusters = data_norm

# n_clusters chosen from sil plot
kmeans = KMeans(n_clusters=8)

#add 1 for cluster label to start at 1 (for viz purposes_)
dn_clusters['cluster'] = kmeans.fit_predict(data_norm) + 1

labels = dn_clusters['cluster'].tolist()

clustered_df = df_norm

# new df with normalized data and labels attached
# frequencies can now be visualized with polar bar plots
clustered_df = clustered_df['labels'] = labels
```

Fig. A.3: Code for clustering with k-means and adding cluster labels to dataset

```

# Code to Generate Cluster Polar Bar Plots

import plotly.graph_objects as go
from plotly.subplots import make_subplots

# setting variables to properly set rotation and size of polar plot
num_slices = 32
theta = [(i + .5) * 360 / num_slices for i in range(num_slices)]
bg_theta = [(i + 5.5) * 360 / 8 for i in range(8)]
width = [360 / num_slices for _ in range(num_slices)]
bg_width = [360 / 8 for _ in range(8)]

layout = go.Layout(
    margin=go.layout.Margin( l=1, r=1, b=40, t=120))

# create fig
fig = dict(data=data, layout=layout)

# set layouts
layout_options = {
    "title_font_size": 24, "title_x": 0.5,
    "legend_x": 0.85, "legend_y": 0.5, "polar_radialaxis_ticks": "", "polar_radialaxis_showticklabels":
    False, "polar_angularaxis_ticks": "", "polar_angularaxis_showticklabels": False}

# generating polar bar plots
for i in range(0, len(set(np.unique(data_norm_clusters['cluster'])))):
    fig = go.Figure(go.Barpolar(
        r=data_norm_clustered_means[data_norm_clustered_means.index == i].values[0],
        theta = theta, width=width,
        marker_line_color="black", marker_color=colors_arr,
        marker_line_width=.5, opacity=1
    ), layout = layout)

    for t in range(0, len(background_cols)):
        fig.add_trace(go.Barpolar(
            r=[1,1,1,1,1,1,1,1],
            theta = bg_theta,
            width=bg_width,
            marker_color=background_cols,
            opacity=0.25,
            name = 'Range ' + str(t+1)
        ))

    title = str(i + 1)
    fig.update_layout(
        title_x=0.4,
        title = 'Cluster ' + title,
        template='plotly_white',
        polar = dict(radialaxis = dict(range=[0,
max(data_norm_clustered_means[data_norm_clustered_means.index == i].values[0])), showticklabels=False),
        angularaxis = dict(showticklabels=False) ))
    fig.update_traces(text=data_norm_clustered_means.columns)
    fig.update_layout(**layout_options)
    fig.show()

```

Fig. A.4: Code for generating polar bar plots to show frequencies of features within clusters