2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA) | 978-1-6654-4337-1/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICMLA52953.2021.00215

Classifying Challenging Behaviors in Autism Spectrum Disorder with Word Embeddings

Abigail Atchison
Fowler School of Engineering
Chapman University
Orange, USA
atchison@chapman.edu

Gabriela Pinto
Fowler School of Engineering
Chapman University
Orange, USA
pinto@chapman.edu

Ali Woodward
Fowler School of Engineering
Chapman University
Orange, USA
alwoodward@chapman.edu

Elizabeth Stevens
Fowler School of Engineering
Chapman University
Orange, USA
estevens@chapman.edu

Dennis Dixon
Center for Autism and Related Disorders
Woodland Hills, USA
d.dixon@centerforautism.com

Erik Linstead
Fowler School of Engineering
Chapman University
Orange, USA
linstead@chapman.edu

***Abstract*—The understanding and treatment of challenging behaviors in individuals with Autism Spectrum Disorder are paramount to enabling the success of behavioral therapy; an essential step in this process is the labeling of challenging behaviors demonstrated in therapy sessions. This paper seeks to add quantitative depth to this otherwise qualitative task of challenging behavior classification. Here we leverage neural document embeddings with Word2Vec to represent clinical notes capturing 1,917 recorded instance of challenging behaviors from therapy sessions conducted by a large autism treatment provider. These embeddings then serve as training data for supervised machine learning algorithms in both binary and multiclass classification tasks to identify challenging behaviors, achieving high classification accuracies ranging from 82.7% to 98.5%. We demonstrate that the semantic queues derived from the language of challenging behavior descriptions, modeled using natural language processing techniques, can be successfully leveraged to extract and identify challenging behaviors from real-world clinical data.**

I. Introduction

Challenging behaviors are defined as behaviors that are not culturally or socially acceptable and can put the physical safety of the individual or others in jeopardy, affect learning, and limit access to community settings [1]–[3]. These behaviors often affect an individual's ability to interact with their environment consistently and positively [4]–[6]. Individuals with Autism Spectrum Disorder (ASD) have been found to demonstrate such challenging behaviors with significant frequency [4]. Further, previous research has not proven the severity and prevalence of challenging behaviors in the lives of individuals with ASD limited to one specific developmental stage such as adolescence or early childhood [2], [7], [8]. In short, improving the efficacy and efficiency of the treatment of ASD means that researchers must look to better understand challenging behaviors exhibited by individuals with ASD to best leverage and target behavioral treatments.

Recently, machine learning has shown great promise in exploring and understanding multiple aspects of ASD [9]–[15]. In this paper, we look to explore the application of natural language processing and supervised machine learning to the problem of classifying challenging behaviors exhibited in ASD. By demonstrating that neural word embeddings can be used to differentiate between challenging behaviors, we present the potential for classification models to be deployed in a clinical environment to aid in establishing the context of behavioral treatment.

A common practice in treating behavioral patterns in ASD is Applied Behavior Analysis (ABA), which employs Behavior Analysis to develop a specific therapy regimen. Behavior Analysis, defined as the scientific study of behavior [16], is leveraged by ABA as a means by which an individual's behavioral patterns can be shaped through the consistent application of reinforcement learning and controlling of environmental factors. During ABA, challenging behaviors exhibited are analyzed and defined on a case-by-case basis. This approach takes into consideration the reality that the same challenging behavior, for example, Aggression, can manifest differently across individuals as well as within a single individual throughout treatment [1], [17]. ABA can be facilitated in various settings, such as home sessions, clinic visits, or even in a school setting. It seeks to deliver a daily therapy regimen targeting both the behavioral strengths and weaknesses of the individual being treated, a goal that has been demonstrated to provide measurable improvements in the outcome of children on the spectrum when early, intensive intervention is facilitated [18].

While ABA represents a demonstrably successful way of addressing challenging behaviors, identifying and tracking these behaviors is highly personalized. These identifications can only be made within the scope and experience of the particular specialist conducting the assessment. We present this paper with the hopes of adding quantitative depth to this process through behavioral classification.

The potential to leverage neural embeddings constructed from behavioral descriptions to classify challenging behaviors

978-1-6654-4337-1/21/$31.00 ©2021 IEEE
DOI 10.1109/ICMLA52953.2021.00215

Authorized licensed use limited to: Chapman University. Downloaded on August 04,2022 at 05:26:41 UTC from IEEE Xplore. Restrictions apply.

holds great value in the understanding of these behaviors and clinical applications. By exploring the models constructed, we demonstrate this potential by outlining how quantitative techniques can be applied to highly qualitative data to improve guidance in identifying behaviors exhibited throughout ABA treatment. We detail the data used in this study, the machine learning methods leveraged in the construction and modeling of word embeddings using Word2Vec, results obtained, and their practical significance.

II. Data

To determine if neural word embeddings could aid in the classification of challenging behaviors, we started with an analysis of behavioral descriptions across 15 behaviors. The descriptions were provided by the Center for Autism and Related Disorders (CARD), one of the largest national providers of ABA therapy services.

The dataset, known as the CARD Skills™ dataset, is a clinical database that houses the ABA curriculum for patients, as well as a detailed log documenting the evolution of patient progress from the start of services to termination of services, which typically spans several years. In this study, we focus specifically on recorded challenging behaviors and the description associated with that particular demonstration of the behavior recorded. In the case of the CARD Skills™ dataset, all patients were under direct supervision of behavior interventionists (BIs) and board certified behavior analysts (BCBAs) during treatment.

The dataset consists of treatment history for 1,602 individuals. After an initial pre-processing step to remove all challenging behavior labels which did not have descriptions associated with them, the resultant dataset consisted of 1,917 total observations of challenging behaviors during ABA treatments. These descriptions were then tokenized using the tokenize package provided by the Python Natural Language Toolkit [19]. Following tokenization, each row in the dataset corresponds to a single instance of challenging behavior. That instance consists of the label of the challenging behavior being exhibited and an ordered collection of tokens associated with that label, with each token being a word from the original description.

These behavioral labels are one of 15 potential categorizations. The frequency counts of each of these categorizations in the dataset can be found in Table I. One count corresponds to one row in the dataset indicating a description of an action reported during an ABA therapy session that was subsequently labeled as that challenging behavior. Due to the variance of sample size across behaviors, in our model construction, we solely focused on the classification of the 7 most commonly occurring behaviors in the dataset. These behaviors were Aggression, Disruption, Elopement, Noncompliance, Self-Injurious Behaviors, Stereotypy, and Tantrums. A definition for each of these select behaviors derived from previous literature can be found in Table II.

In addition to the behaviors studied in this analysis, it should be noted that behavioral instances listed as "Other" are numerous in the dataset. While we do not seek to analyze these descriptions in this particular study, their prevalence in the dataset further contributes to the potential value of a classification model as a means to reduce the occurrence of these ambiguous labels.

Behavior	**Count**
Aggression	462
Disruption	140
Elopement	64
Hoarding	3
Inappropriate Sexual Behavior	11
Lying	5
Noncompliance	221
Obsessive Behaviors	27
Pica	10
Self-Injurious Behavior	88
Stealing	1
Stereotypy	188
Tantrums	252
Teasing/Bullying	11
Other	438

TABLE I
CHALLENGING BEHAVIOR FREQUENCY COUNTS

Behavior	**Definition**
Aggression	Generalized hitting, kicking, biting, punching,scratching and throwing (i.e. furniture) [20]
Disruption	Hitting, kicking, biting, punching, scratching and throwing (i.e. furniture) in response to something non-specific to the individual [20]
Elopement	Wandering, leaving, and running from safe spaces or adult supervision [21]
Noncompliance	Any behavior other than what has been requested within a specified period of time [22]
Self-Injurious Behavior	Behaviors in which individuals cause physical damage to his or her own body (scratching, biting, head banging, chin hitting, hair pulling, skin picking, eye pressing or gouging [23], [24]
Stereotypy	Repetitive and non-functional behavior [25], [26]
Tantrums	Hitting, kicking, biting, punching, scratching and throwing (i.e. furniture) in response to specific to the individual [20]

TABLE II
CHALLENGING BEHAVIOR DEFINITIONS

III. Methods

To analyze behavioral descriptions in our data, we employ natural language processing techniques to construct neural word embeddings. The vectorization of a natural language

Authorized licensed use limited to: Chapman University. Downloaded on August 04,2022 at 05:26:41 UTC from IEEE Xplore. Restrictions apply.

data collection adds quantitative depth to otherwise flat, textual data, allowing us to apply machine learning algorithms.

The technique used in this paper is Word2Vec. At its core, Word2Vec is the construction of a vector space that models the semantic and syntactic meaning of words found within a textual corpus. These vectors are generated by training a neural network via the Skip-Gram approach or the Continuous Bag of Words (CBOW) approach. Predictions are made given the word and the context the word appears in the text corpus; then, word vectors are generated through a back-propagation over a weight matrix.

Following the construction of this vector space, the term frequency-inverse document frequency (TF-IDF) score for each word in a given document is combined with its corresponding word vector representation to construct weighted document vectors. TF-IDF is a commonly used technique in information retrieval. It considers the overall frequency of a word in the entire corpus, its term frequency, and its frequency in the specific document we are currently generating a vector for, its document frequency. If a word appears more frequently in the corpus, it will lower its TF-IDF score as it is perhaps a more general stop word. Inversely, suppose that word frequently appears in the current document. In that case, this will increase its TF-IDF score as this frequency could indicate the importance of that particular word to the current document. These scores are then used to generate weighted document vectors for each behavior report by taking the TF-IDF weighted sum of the neural embedding of each word that appears in a given document.

Once these neural embeddings are learned, they can be used as input data in supervised machine learning algorithms to demonstrate whether they contain enough information to differentiate between behaviors. Commonly referred to as classification models, supervised learning learns how to produce a label, given input through exposure to already labeled data. At their core, most classification algorithms operate by minimizing incorrect label assignment by adjusting features used to predict a label given an input. These features can be as complex as a deep weight matrix or as simple as coefficients in a single, linear function. While the specifics to the construction of classification models vary, their overall mission remains consistent when given truth data consisting of inputs and expected labels: learn features that will allow us to classify data we have not seen before correctly.

For this analysis, we consider two algorithms to extract meaningful features from our data. The first, the Support Vector Machine (SVM), is an algorithm that aims to fit a hyperplane to a set of input data by maximizing the margin that separates classes within the dataset in order to differentiate between possible labels [27]. In this study, we used SVMs as a means to produce binary classifications. This baseline algorithmic analysis of the document vectors generated previously demonstrates the potential of studying challenging behaviors using neural embeddings. We utilized a Gaussian Process Classifier to build a classification model that differentiates beyond binary classification and generates a 7-class output. Our Gaussian Process uses Laplace approximation to smooth several one-class-versus-rest Gaussian Distributions, which can then be sampled to determine the final classification of any one input [28].

We provide a brief mathematical overview of SVM and Gaussian Process Classification using Laplace approximation in the following subsections. We then apply these algorithms and analyze the results.

1) Support Vector Machines: Upon construction of vector embeddings, consider a data matrix, D, of dimension $m \times n$. D can then be represented as a collection of vectors, $D = \{X_1, X_2, ..., X_m\}$. Each vector, X_i, represents a unique data instance, in this case a document embedding, and each vector element, $X_{i,j}$, a specific measurement (attribute) for that embedding.

An SVM classifier takes in D as input and outputs a set of weights $W = \{w_1, w_2, ..., w_n\}$, one for each feature, m, in the input such that for each vector X_i its corresponding label y_i will be 1 for positive samples and -1 for negative samples. Thus, we can state that

$$y_i(W \cdot X_i + b) - 1 \geq 0$$

The linear combination of X_i and W predict the class label, y_i for any given data point X_i. The approximation of the best hyperplane, which maximizes the margin between two classifications, determines these weights. This margin, L, can be defined as:

$$L = \sum_i \alpha_i - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j X_i \times X_j$$

Upon maximizing L for the training inputs, our weight vector, W, remain. Given these weights, W, and an unknown data points, u, if $W \cdot u + b \geq 0$ then u is a positive classification.

While SVMs provide an algorithmically simple way to differentiate between two classes in a convex space that avoids local maxima, the mechanism struggles when data points are not linearly separable. This can be addressed by introducing a Kernel method K which represents the dot product between two input vectors into a new space such that $K_{X_i, X_j} = \phi(X_i) \cdot \phi(X_j)$. Some common Kernels include linear and radial basis kernels. In this paper, a radial basis function (RBF), as the kernel, is used to fit each SVM model.

2) Gaussian Process Classifier: Again, consider a data matrix, D, of dimension $m \times n$. D can then be represented as a collection of vectors, $D = \{X_1, X_2, ..., X_m\}$. Each vector, X_i, represents a unique data instance, and each vector element, $X_{i,j}$, a specific measurement (attribute) for that point.

A Gaussian Process defines a non-parametric distribution over functions, $p(f)$, such that for any finite subset $f_{s-w}\{f_s, f_{s+1}, ..., f_{w-1}, f_w\} \subset D$ the marginal distribution over that subset has a multivariate Gaussian distribution.

In the Gaussian Process Classification (GPC) used in this paper, during training, a one versus rest Gaussian Distribution for each of the C classes, $C_1, C_2, ..., C_C$ is constructed. These distributions model the trend of the data, D, while maximizing the space between classes, C, and, once trained,

Authorized licensed use limited to: Chapman University. Downloaded on August 04,2022 at 05:26:41 UTC from IEEE Xplore. Restrictions apply.

can be leveraged in classification. Given a previously unseen input vector X_i we determine the most likely classification by sampling our constructed Gaussians and choosing the closest fit. While computationally expensive to train due to GPC's foundation in Bayesian inference, when applied to small, unbalanced datasets, like the one studied in this paper, this method models multiple classes with great success. For an in-depth treatment of GPC, see work outlined by Rasmussen et al. [28].

3) Laplace approximation in GPC: In this function approximation, the approximate log marginal likelihood across multiple classes combined with the Gaussian prior is non-Gaussian. Thus, the resultant non-Gaussian posterior process must be smoothed using Laplace approximation [28].

IV. Results

Previous analyses of the topography of challenging behaviors have manifested as largely qualitative in nature [3]. These studies effectively underline the difficulty and the necessity of the consistent and reliable labeling of challenging behaviors exhibited by individuals with ASD. In this labeling, one must distinguish between behaviors which, at times, can only be separated based on the implied function of these actions as is the case with Disruptions and Tantrums [3]. In this paper, we look to diverge from a qualitative analysis into the quantitative, allowing for the discovery of patterns and features in behavioral data that could only be extracted through machine learning algorithms. Below we will discuss the results of applying the methods presented to the challenging behavior descriptions and their labels. We present these findings with the hope that in the future, clinicians can use feedback from similar models to identify the most likely behavioral label given actions exhibited in therapy sessions.

A. Construction of Neural Document Embeddings

In analyzing the Skills™ data of challenging behavior descriptions and their labels, we constructed a neural embedding for each challenging behavior report using a TF-IDF weighted sum of Word2Vec word vectors. Each set of embeddings serve as input to the Support Vector Machines and Gaussian Process Classifier. The computing software, Python, performed the computations.

Gensim's "Word2Vec" package was used for the creation of the Word2Vec embeddings on behavior description using the Skip-Gram architecture [29]. The optimal embedding size to allow for sufficient information capture was empirically chosen to be 50. In order to ensure that instance-specific words do not affect behavioral classification, we set a minimum frequency of 5, meaning that a vector embedding was only constructed for a word if it appeared at least 5 times across the corpus. From scikit-learn, the "TfidfVectorizer" was used to generate TF-IDF scores for each of the tokenized words in the corpus [30]. To generate the weighted Word2Vec embeddings, each word in the description we queried its word vector generated by the Word2Vec model and multiplied by its corresponding TF-IDF weight. The calculated product was then summed across all of the words in the specific description, and the final sum was the resultant document vector. We performed this process for each description in the dataset.

B. Analysis of Weighted Word Embeddings

After constructing the weighted Word2Vec vectors document, the information captured by these embeddings was evaluated to determine if this method successfully differentiates between challenging behaviors in ASD. The evaluation took in two forms, binary classification using SVMs and multi-class classification model using GPC. The test accuracy resultant from classification models built on an 80/20 training/testing split was used as a metric for the effectiveness of the embeddings. Ultimately, the model's performance ranged from 84.3% and 98.5% using the SVMs, and 82.8% using the GPC.

1) Binary Classification using Support Vector Machines: The first classification model applied to the Word2Vec embeddings was a Support Vector Machine (SVM) using a Radial Basis Function kernel. To apply the SVM classifier to the word vectors for binary classification, we use the "SVC" class in the "sklearn" Python package [30]. Table III outlines the results of the unweighted, binary SVM classification. The model was trained on an 80/20 stratified split of the weighted W2V.

Class A	Class B	Test Accuracy
Elopement	Self-Injurious Behavior	97.4
Aggression	Noncompliance	96.5
Aggression	Elopement	97.9
Noncompliance	Self-Injurious Behavior	**98.5**
Sterotypy	Self-Injurious Behavior	95.0
Elopement	Stereotypy	96.8
Aggression	Disruption	94.6
Elopement	Tantrums	97.3
Aggression	Stereotypy	96.8
Sterotypy	Tantrums	96.4
Noncompliance	Stereotypy	94.9
Disruption	Self-Injurious Behavior	96.9
Disruption	Noncompliance	91.6
Aggression	Tantrums	93.9
Noncompliance	Elopement	95.0
Aggression	Self-Injurious Behavior	94.3
Disruption	Elopement	95.6
Disruption	Stereotypy	89.4
Noncompliance	Tantrums	89.4
Disruption	Tantrums	84.3

TABLE III
SVM Classifier Results sorted by test accuracy

Test Accuracy displayed is for inputs of TF-IDF Weighted Vec. The highest classification accuracy has been bolded.

The accuracy metrics shown in Table III demonstrate a separation between embeddings given any two behavioral classes in our dataset, bearing particular promise as previous qualitative analysis of these descriptions has highlighted the difficulty of separating challenging behaviors when considering the topography of these descriptions.

Hong et al. used a part-of-speech analysis on bag-of-words representations of challenging behavior descriptions to identify the top 20 verb terms used to define the topography of each challenging behavior [3]. Their findings, particularly the qualitative overlap between challenging behaviors when

Authorized licensed use limited to: Chapman University. Downloaded on August 04,2022 at 05:26:41 UTC from IEEE Xplore. Restrictions apply.

studying the verbiage used in clinical descriptions, can be used as a counterpoint to our quantitative analysis of the vector space created by neural embeddings constructed from similar descriptions.

Self-Injurious Behavior (SIB) and Aggression shared 38.5% overlap in their descriptional verbiage. This is unsurprising when one considers these behaviors only truly diverge regarding their target, SIB targeting oneself while Aggression targets an outside entity. Our model distinguishes between these behaviors with a testing accuracy of 94.3%.

Noncompliance and Tantrums also bore many similarities in the verbiage used to describe them in a clinical setting. Specifically, 37.9% of the 20 most common verb terms used in describing them were shared between these two behaviors. Our SVM model distinguishes between the two behaviors with a test accuracy of 84.3%.

Additionally, the percentage of overlap presented difficulty qualitatively distinguishing between Aggression, Disruption, and Tantrums. They overlapped by 33%, sharing 2 of 20 most common verb terms used to describe them, including crying, screaming, throwing, and protesting. In our results, the classification accuracy achieved when fitting a hyperplane to an Aggression vs. Tantrum and Aggression vs. Disruption vector space was 93.5% and 94.6%, respectively. The worst accuracy of the pairwise binary classification between these three classes was 84.3% when classifying between disruption and tantrums. This classification accuracy provides further insight into the shared action between Disruptions and Tantrums. The process of demonstrating disruptive behavior is often the same process followed in the manifestation of a tantrum, an empirical truth reinforced by the interchangeable use of these terms in the behavioral analysis of ASD [3]. Furthermore, of the top 10 verbs used to describe these two behaviors in a clinical setting, four of them were shared between disruption and tantrums (crying, screaming, throwing, and yelling). Thus, our hypothesis holds. A quantitative analysis of challenging behaviors, even those found to be qualitatively indistinguishable at times, will produce deeper insight and perhaps provide a mechanism to aid clinicians in analyzing behavioral manifestations during treatment.

To further demonstrate the linear separability of the embeddings constructed in this analysis, a pairwise comparison of skills has been visualized for the two most successful classifications (see Figure 1 and 2) and the two least successful classifications (see Figure 3 and 4) in our SVM model. Given the weighted word embeddings of two challenging behaviors, t-Distributed Stochastic Neighbor Embedding (t-SNE) was used to reduce the 50-D vector space constructed to a 2-D, graphically-feasible entity. The Word2Vec model generated the word vector space used in this reduction and visualization.

Per these visualizations, we can see that challenging behaviors are distinguishable in our constructed vector space, even when the frequency of classes is unbalanced in our dataset. This separation is universal across each binary classification presented, except for Disruption vs. Tantrums as shown in 3, Noncompliance vs. Tantrums as shown in 4, and Disruption

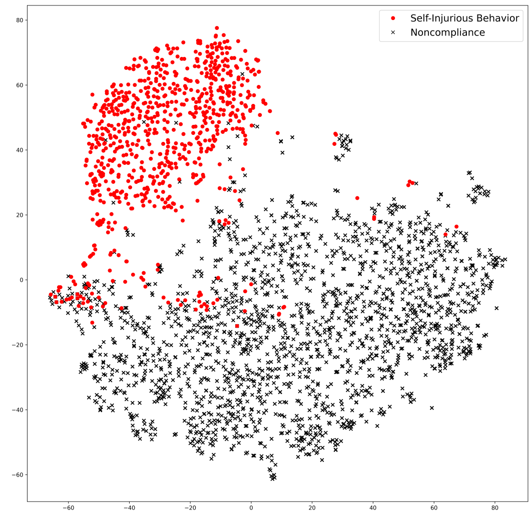

Fig. 1. Self-Injurious Behavior vs Noncompliance - 98.5%

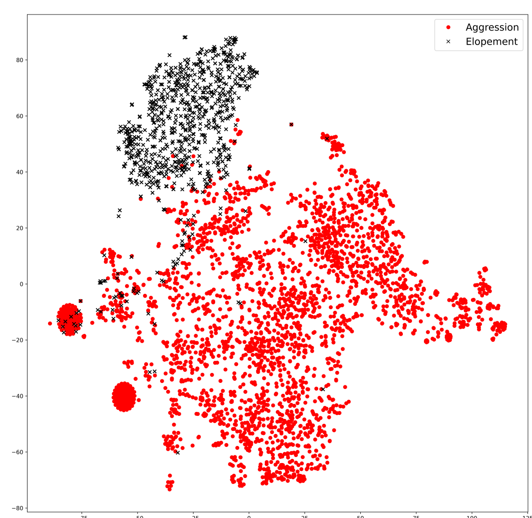

Fig. 2. Aggression vs Elopement - 97.9%

Authorized licensed use limited to: Chapman University. Downloaded on August 04,2022 at 05:26:41 UTC from IEEE Xplore. Restrictions apply.

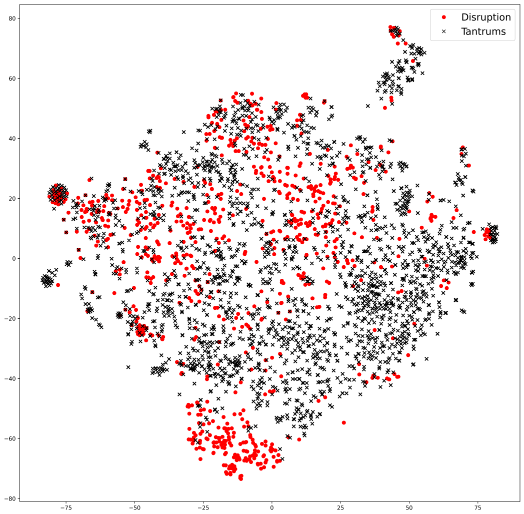

Fig. 3. Disruption vs Tantrums - 84.3%

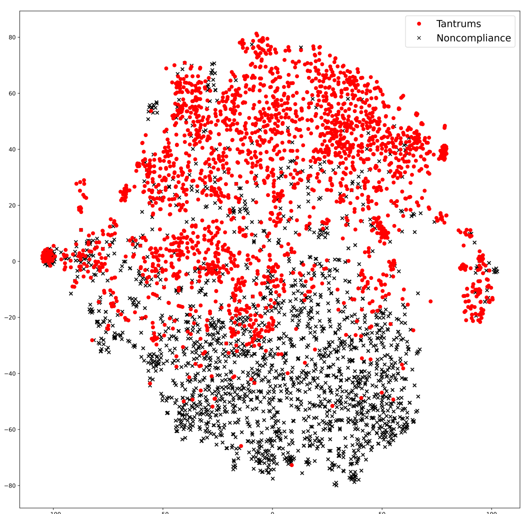

Fig. 4. Tantrums vs Noncompliance - 89.4%

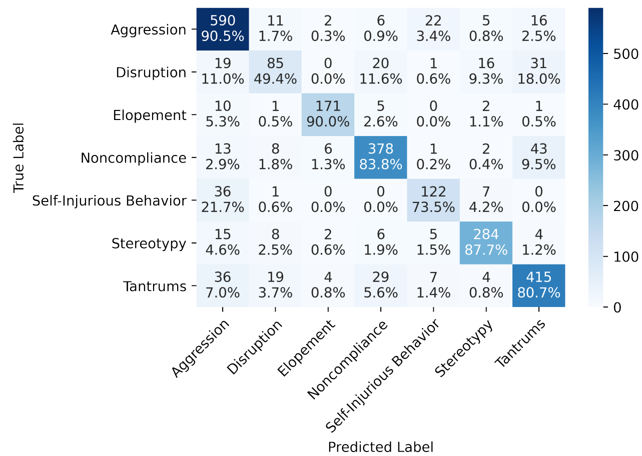

Fig. 5. TF-IDF Weighted Word2Vec

vs. Stereotypy.

2) Multi-Class Classification using Gaussian Processes: Binary classification within the behavioral dataset provides a promising first analysis of challenging behaviors and a counterpoint to previous qualitative analysis. Nonetheless, to be leveraged in practice, a multi-class classification model is needed to differentiate between the many behaviors that could manifest during behavioral treatment.

To achieve this, a Gaussian Process Classifier (GPC) was built to generate classifications across seven challenging behaviors. The “GaussianProcessClassifier” function in the scikit learn Python package [30] constructed the model, then trained on an 80/20 stratified split with the weighted word embeddings as input. Figure 5 displays the confusion matrices produced when using Weighted Word2Vec embeddings as input. The GPC performed with a training accuracy of 87.9% and testing accuracy of 82.8%.

A weakness of the classifier to identify Disruption, Tantrums, and Noncompliance Behavior is revealed in the confusion matrix in Figure 5. These exemplars were classified with 49.42%, 80.74%, and 83.81% accuracy, respectively. In the case of Disruption, this class was misclassified as Tantrums 31 times and Aggression 19 times, a reasonable failing when considering these behaviors overlap in their topographic definition and share common descriptive verbiage [3], [20]. The misclassification of Disruption could be attributed to the ambiguous nature of challenging behavior labeling and the overlap that exists across these behaviors. Elopement and SIB were most commonly misclassified as Aggression, 10 times and 36 times respectively. Again, both of these true behavioral classifications could be seen to manifest as Aggression as well when one considers their generalized definitions.

Analyzing the Word2Vec neural embeddings via two classification model architectures demonstrates that the semantic and syntactic queues extracted during neural embedding construction allow for the differentiation between challenging behaviors. Often, these differentiations have been made between

Authorized licensed use limited to: Chapman University. Downloaded on August 04,2022 at 05:26:41 UTC from IEEE Xplore. Restrictions apply.

classes that qualitatively are nearly indistinguishable.

V. Discussion

Currently, there is no cure for ASD; however, early intervention and treatment have been shown to facilitate the future success of a child diagnosed with ASD [31]. In the treatment of ASD, behavioral intervention is essential since studies demonstrated a reduction in the manifestation of challenging behaviors in individuals by 80-90% [4]. In some cases, these treatments succeed in virtually eliminating all appearances of challenging behaviors [32], [33], as such, aiding in the reliability of this treatment by providing a mechanism that lends additional insight into behavioral classifications holds great value.

A topographical analysis of challenging behaviors has highlighted the ambiguity of behavioral descriptions due to shared similarities in verbiage [3]. Nonetheless, in our analysis of the neural embeddings created from this textual data, we effectively and consistently differentiated between challenging behaviors. The importance of our work can thus be highlighted in the provision of a quantifiable mechanism for differentiating between behaviors that have qualitative overlap.

It should be noted that in providing quantifiable separation between behaviors, we have also added greater depth to the ongoing discourse surrounding the existence of Disruption and Tantrums as two separate behavioral labels. These labels are believed to describe very similar actions, which their implied function can only differentiate. Consistently labeling these behaviors calls for a functional analysis, a utility that is separate from the behavioral analysis being conducted during ABA therapy. To this end, we present our findings, specifically the lack of quantifiable separation between the semantic structures used to describe Tantrums and Disruption, as further evidence that these behaviors should be revisited and a single label should be offered to cover this subset of behavioral patterns.

In addition to providing additional insight to therapists labeling demonstrated behaviors in treatment, the prevalence of behaviors labeled as "Other" within the dataset, which represents the second highest occurring label at 438 instances, provides an interesting aside to the classification presented in this paper. The number of "Other" instances demonstrated an ambiguity in challenging behavior representation only further reinforces the need for an additional, quantitative metric when analyzing behaviors. Additionally, it provides an opportunity to give more depth to existing files by giving a sub-label to behavioral reports already labeled as "Other".

In short, the classification success achieved in this study not only demonstrates the quantifiable differences in observational descriptions of challenging behaviors during behavioral therapy but also provides the potential for increasing the efficacy of this crucial form of intervention.

VI. Conclusion

Despite the significant likelihood that individuals with ASD demonstrate challenging behaviors [4], there is still substantial variability in the definitions of these behaviors, [34], potentially contributing to degrees of uncertainty when facilitating behavioral interventions. This variability can be attributed to the tendency of challenging behaviors to manifest differently across individuals with ASD and even within individuals over contexts or time [1], [17]. This reality is taken into account during ABA therapy in which behavioral definitions are constructed on a case-by-case basis [16]. Despite this demonstrated variability, in this paper, we show that weighted word embeddings can synthesize clinical descriptions into highly classifiable entities.

This study takes the clinical descriptions of challenging behaviors and their corresponding labels and generates embeddings via TF-IDF weighted Word2Vec. The resultant embeddings are then independently analyzed through both binary and multi-class classification via SVM models and multi-class GPC, respectively. The latter of which achieves a classification accuracy of 82.8% over seven behavioral labels.

It should be noted that even as the application of Machine Learning to clinical ASD data continues to gain momentum [35], the lack of an Internet-scale, public repository of longitudinal data remains. This dataset could serve as a baseline for more exploratory research in big data. In understanding this, it becomes evident that while this paper lays the foundation for other studies to follow, the most obvious of these future works lies in the continued analysis of challenging behavior descriptions on a larger scale data repository.

In a more general sense, the analysis of textual data curated in the clinical treatment of ASD using word embeddings is still a largely unexplored direction of ASD research. The quantitative representations of otherwise highly qualitative information in ASD present the opportunity to analyze a new dimension of data and, when possible, leverage this analysis to make informed decisions about treatment. The work in this study represents a modest contribution toward that goal. In a future study, we plan to implement document vectors, Doc2Vec, as another means of classification through natural language processing.

References

[1] E. Emerson, *Challenging behaviour: Analysis and intervention in people with learning disabilities.* ERIC, 1995.

[2] J. L. Matson, S. Mahan, J. A. Hess, J. C. Fodstad, and D. Neal, "Progression of challenging behaviors in children and adolescents with autism spectrum disorders as measured by the autism spectrum disorders-problem behaviors for children (asd-pbc)," *Research in Autism Spectrum Disorders*, vol. 4, no. 3, pp. 400–404, 2010.

[3] E. Hong, D. R. Dixon, E. Stevens, C. O. Burns, and E. Linstead, "Topography and function of challenging behaviors in individuals with autism spectrum disorder," *Advances in Neurodevelopmental Disorders*, vol. 2, pp. 206–215, 2018.

[4] 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. 423–446, 2002.

[5] F. C. 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.

Authorized licensed use limited to: Chapman University. Downloaded on August 04,2022 at 05:26:41 UTC from IEEE Xplore. Restrictions apply.

[6] 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.

[7] 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.

[8] 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.

[9] 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," in *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA)*. IEEE, 2015, pp. 414–418.

[10] E. Linstead, R. Burns, D. Nguyen, and D. Tyler, "Amp: A platform for managing and mining data in the treatment of autism spectrum disorder," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2016, pp. 2545–2549.

[11] 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," in *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*. IEEE, 2017, pp. 661–666.

[12] K. Hyde, A.-J. Griffiths, C. Giannantonio, A. Hurley-Hanson, and E. Linstead, "Predicting employer recruitment of individuals with autism spectrum disorders with decision trees," in *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*. IEEE, 2018, pp. 1366–1370.

[13] R. Anden and E. Linstead, "Predicting eye movement and fixation patterns on scenic images using machine learning for children with autism spectrum disorder," in *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2020, pp. 2563–2569.

[14] 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.

[15] J. Gardner-Hoag, M. Novack, C. Parlett-Pelleriti, E. Stevens, D. Dixon, E. Linstead *et al.*, "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, p. e27793, 2021.

[16] C. for Autism and R. Disorders, "ABA resources: What is ABA?" http://www.centerforautism.com/aba-therapy.aspx, accessed: 2016-05-26.

[17] J. L. Matson, D. E. Kuhn, D. R. Dixon, S. B. Mayville, R. B. Laud, C. L. Cooper, C. J. Malone, N. F. Minshawi, A. N. Singh, M. A. Luke *et al.*, "The development and factor structure of the functional assessment for multiple causality (fact)," *Research in Developmental Disabilities*, vol. 24, no. 6, pp. 485–495, 2003.

[18] O. I. Lovaas, "Behavioral treatment and normal educational and intellectual functioning in young autistic children." *Journal of consulting and clinical psychology*, vol. 55, no. 1, p. 3, 1987.

[19] S. Bird, E. Klein, and E. Loper, *Natural Language Process with Python*. O'Reilly Media Inc., 2009.

[20] J. Matson, "Aggression and tantrums in children with autism: A review of behavioral treatments and maintaining variables," *Journal of Mental Health Research in Intellectual Disabilities*, vol. 2, no. 3, pp. 169–187, 2009.

[21] C. C. Piazza, G. P. Hanley, L. G. Bowman, J. M. Ruyter, S. E. Lindauer, and D. M. Saiontz, "Functional analysis and treatment of elopement," *Journal of applied behavior analysis*, vol. 30, no. 4, pp. 653–672, 1997.

[22] 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.

[23] E. G. Carr, "The motivation of self-injurious behavior: a review of some hypotheses." *Psychological bulletin*, vol. 84, no. 4, p. 800, 1977.

[24] J. L. Matson and S. V. LoVullo, "A review of behavioral treatments for self-injurious behaviors of persons with autism spectrum disorders," *Behavior Modification*, vol. 32, no. 1, pp. 61–76, 2008.

[25] A. P. Association *et al.*, *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub, 2013.

[26] J. T. Rapp and T. R. Vollmer, "Stereotypy i: A review of behavioral assessment and treatment," *Research in developmental disabilities*, vol. 26, no. 6, pp. 527–547, 2005.

[27] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[28] C. E. Rasmussen, "Gaussian processes in machine learning," in *Summer School on Machine Learning*. Springer, 2003, pp. 63–71.

[29] R. Řehůřek and P. Sojka, "Software Framework for Topic Modelling with Large Corpora," in *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*. Valletta, Malta: ELRA, May 2010, pp. 45–50, http://is.muni.cz/publication/884893/en.

[30] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[31] A. Speaks, "CDC update on autism shows gap between early concerns and evaluation," https://www.autismspeaks.org/news/news-item/cdc-update-autism-shows-gap-between-early-concerns-and-evaluation, accessed: 2016-05-14.

[32] 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, no. 2, pp. 229–232, 2001.

[33] J. L. Matson, D. R. Dixon, and M. L. Matson, "Assessing and treating aggression in children and adolescents with developmental disabilities: a 20-year overview," *Educational Psychology*, vol. 25, no. 2-3, pp. 151–181, 2005.

[34] J. L. Matson and M. L. Matson, *Comorbid conditions in individuals with intellectual disabilities*. Springer, 2015.

[35] K. K. Hyde, M. N. Novack, N. LaHaye, C. Parlett-Pelleriti, R. Anden, D. R. Dixon, and E. Linstead, "Applications of supervised machine learning in autism spectrum disorder research: a review," *Review Journal of Autism and Developmental Disorders*, vol. 6, no. 2, pp. 128–146, 2019.

Authorized licensed use limited to: Chapman University. Downloaded on August 04,2022 at 05:26:41 UTC from IEEE Xplore. Restrictions apply.