

# Time-Series Prediction of Proximal Aggression Onset in Minimally-Verbal Youth with Autism Spectrum Disorder Using Physiological Biosignals

Ozan Özdenizci, Catalina Cumpanasoiu, Carla Mazefsky, Matthew Siegel,  
Deniz Erdoğan, Stratis Ioannidis, Matthew S. Goodwin

**Abstract**—It has been suggested that changes in physiological arousal precede potentially dangerous aggressive behavior in youth with autism spectrum disorder (ASD) who are minimally verbal (MV-ASD). The current work tests this hypothesis through time-series analyses on biosignals acquired prior to proximal aggression onset. We implement ridge-regularized logistic regression models on physiological biosensor data wirelessly recorded from 15 MV-ASD youth over 64 independent naturalistic observations in a hospital inpatient unit. Our results demonstrate proof-of-concept, feasibility, and incipient validity predicting aggression onset 1 minute before it occurs using global, person-dependent, and hybrid classifier models.

**Index Terms**—autism, minimally verbal, aggression, physiological arousal, naturalistic observation, logistic regression

## I. INTRODUCTION

A substantial number of youth with autism spectrum disorder (ASD) show unpredictable and potentially dangerous aggressive behavior [1–4]. Aggression is particularly impairing in youth with ASD who are minimally verbal (MV-ASD). Due to their difficulty to adequately self-report increasing distress, aggression in MV-ASD is often unpredictable, which makes it dangerous, difficult to manage, and constitutes a barrier to accessing the community. In typically developing youth, greater ability to regulate physiological arousal is associated with fewer behavior problems [5]. Prior research suggests an association between physiological arousal and problem behavior to alleviate distress in ASD [6–10]. Thus, our current study evaluates whether proximal biomarkers of physiological arousal [11], [12] can be used to predict the onset of aggression before it occurs.

A growing number of researchers have approached this goal using various observational designs [13–17]; however, they all rely on artificial experimental settings and tasks that call into question the ecological validity of obtained results. In the present study, we utilize a novel data set in which physiological biosignals were collected wirelessly from behaviorally unstable MV-ASD youth during unstructured inpatient hospital observations. We employ ridge-regularized linear classifiers with time-series features of cardiovascular and electrodermal activity as physiological biomarkers of imminent aggression. Using global, person-dependent, and

hybrid (i.e., partially global) models, we demonstrate high predictability of aggressive behavior onset in a specific upcoming time interval. A preliminary version of this work was presented in [18].

## II. MATERIALS AND METHODS

### A. Participants and Data Acquisition

Physiological and physical activity data collected from 15 ADOS-2 [19] confirmed, behaviorally unstable MV-ASD youth recorded over 64 unstructured observational sessions in the Developmental Disorders Unit at Spring Harbor Hospital, Portland, ME, United States, were used in this Institutional Review Board approved study. Data were collected during naturalistic observational sessions at this specialized ASD psychiatric inpatient unit using a wrist-worn E4 biosensor (Empatica Inc., United States) as well as time-synchronized and operationally defined (hitting, kicking, biting, scratching, grabbing, pulling) coding of aggression to others by inpatient research staff with at least 90% inter-rater reliability. All MV-ASD youth in our sample tolerated the biosensor after desensitization, usable data were obtained in all cases, and a total observation time of approximately 83 hours was achieved with an average of  $5.5 \pm 4.9$  hours across participants. Aggressive episodes were observed 35.9 (SD = 34.1) times with average durations of  $31.9 \pm 33.2$  seconds.

To capture measures of physiological arousal, the following autonomic nervous system indices were recorded by the wrist-worn E4 biosensor: (1) heart rate and heart rate variability, which is a measure of the variation in beat-to-beat interval, both derived from *blood volume pulse (BVP)* and *inter-beat interval (IBI)* via photoplethysmography [20] at 64 Hz; and (2) *electrodermal activity (EDA)* sampled at 4 Hz, which reflects autonomic innervation of sweat glands and provides a sensitive measure of alterations in physiological arousal. To quantify changes in physical activity, the E4 records *movement acceleration (ACC)* using an embedded 3-axis accelerometer at 32 Hz sampling rate.

### B. Time-Series Feature Extraction

Statistical analyses of physiological and physical activity signals were performed through extracted time-series features offline. For each one of the 6 signal sources (i.e., *BVP*, *IBI*, *EDA*, *ACC<sub>x</sub>*, *ACC<sub>y</sub>*, *ACC<sub>z</sub>*) the following features were calculated in bins of 15 seconds: first, last, maximum, minimum, mean and median value, amount of unique values,

O. Özdenizci, C. Cumpanasoiu, D. Erdoğan, S. Ioannidis and M. S. Goodwin are with Northeastern University, Boston, MA, USA. M. Siegel is with Maine Medical Center Research Institute, Portland, ME, USA. C. Mazefsky is with University of Pittsburgh, Pittsburgh, PA, USA. Corresponding author: oozdenizci@ece.neu.edu.

and the sum, standard deviation, and variance of values within a bin. To exploit temporal information of aggression episodes, we extracted two more features using time-synchronized binary aggression labels offline; *time since past aggression (TPA)*, which indicates the amount of time elapsed since the last observation of an aggression episode; and a binary *aggression observation flag (AOF)* feature, indicating whether an aggression episode has so far occurred within that recording session. The standard deviation of each calculated feature across time-series bins was included in all prediction models.

### C. Ridge-Regularized Logistic Regression Classifier

Binary decision making on aggressive behavior onset was performed through a ridge-regularized logistic regression model. Among all extracted features (see Section II-B), the following five different feature subsets were used as predictor variables in our analyses: (1) only temporal information (*TPA, AOF*); (2) only physical activity (*ACC*); (3) only physiological activity (*BVP, IBI, EDA*); (4) physical and physiological activity features combined (*BVP, IBI, EDA, ACC*); and (5) all features combined (*BVP, IBI, EDA, ACC, TPA, AOF*).

At every time point  $t$ , using features extracted in a previous time range  $[t - \tau_p, t)$ , a classifier was used to predict a binary dependent variable, estimating whether aggression will be observed or not in an upcoming time range  $(t, t + \tau_f]$ . We adopted a 5-fold cross-validation protocol to generate training and testing data splits, repeated five times to produce confidence intervals. Predictions with particular values of  $\tau_p$  and  $\tau_f$  were performed via the following schemes:

- 1) *Global Prediction Models*: Time-series data across all participants and observational sessions were concatenated. Using the training data split obtained, a single classifier was constructed as the global model.
- 2) *Person-Dependent Prediction Models*: Data were pooled across all observational sessions within each one of the 15 participants individually. Obtained training data splits were used to construct 15 distinct person-dependent prediction models.
- 3) *k-Hybrid Prediction Models*: In this framework, person-dependent prediction models were trained in a partially global scheme. Specifically,  $k$  most significant physiological or physical biomarker features were identified via their regression weights in the global models (i.e.,  $k$  features with the highest weight magnitudes). Using each participant’s pooled data, the corresponding  $k$  feature weights were trained globally, whereas the other feature weights were trained solely on person-dependent data. In this scheme, 15  $k$ -hybrid prediction models are constructed for each participant.

Considering a predictive variable feature space dimensionality of  $d$  which varies with  $\tau_p$ , iterative training of logistic regression weights at every cross-validation fold requires estimation of  $d$  hyperparameters for global models, and  $d \times S$  hyperparameters for person-dependent models where  $S$  denotes the number of participants. In  $k$ -hybrid

models, the number of hyperparameters to be estimated are  $((d - k) \times S) + k$ . This becomes more computationally parsimonious with respect to person-dependent models when the number of participants  $S$  is high.

## III. RESULTS

Classification analyses for particular values of  $\tau_p$  and  $\tau_f$  were performed through global, person-dependent, and  $k$ -hybrid models. To evaluate performance, Receiver Operator Characteristic (ROC) curves and corresponding Area Under the Curve (AUC) values were calculated through probabilities that logistic regression classifiers generate two classes: aggression and non-aggression. Data were processed for decision making every 15 seconds, resulting in 18,869 samples in global models.

### A. Global Prediction Models

Global prediction models of aggression in the upcoming one minute using all extracted features from the past  $\tau_p = 60$  seconds resulted in the blue solid ROC curve presented in Figure 1(a) with a corresponding AUC value of 0.69. Figure 1(b) depicts increases in AUC when E4 biosensor data is included, compared to using any other subset of features to predict aggression episodes.

Regarding the relationship between past time range ( $\tau_p$ ) and future time range used to make aggression onset predictions ( $\tau_f$ ), Figure 1(c) depicts stationary performance in global models using all features from various past  $\tau_p$  durations. However, when using physical and physiological biosensor data from the past (c.f. bold dashed red curve), we observe relative AUC increases compared to temporal data of past aggression only (c.f. light dotted dark blue curve).

### B. Person-Dependent Prediction Models

We repeat the above analysis for  $\tau_p = \tau_f = 60$  seconds in person-dependent models; Table I contains the corresponding mean AUCs. Across participants, we observe an average 0.16 increase in AUC values up to 0.80 using E4 biosensor data compared to using temporal aggression information only. Figure 1(a) shows ROC curves corresponding to person-dependent model prediction performance using all features compared to the global prediction model with the same parameters. We observe a higher mean AUC and a much favored behavior of steep increases of sensitivity for low false positive rates in most person-dependent models compared to global prediction models.

### C. k-Hybrid Prediction Models

Table II demonstrates results obtained by  $k$ -hybrid models with  $\tau_p = \tau_f = 60$  seconds for varying values of  $k$ . For parameter  $\tau_p = 60$ , dimensionality of feature vectors became  $d = 310$  when all extracted features in bins of 15 seconds were concatenated to be used as predictor variables. In this case, by definition (see Section II-C), a  $k$ -hybrid model with  $k = 310$  is equivalent to a global prediction model. Mean AUC values of the  $k$ -hybrid models yielded stationary performance between person-dependent and global models

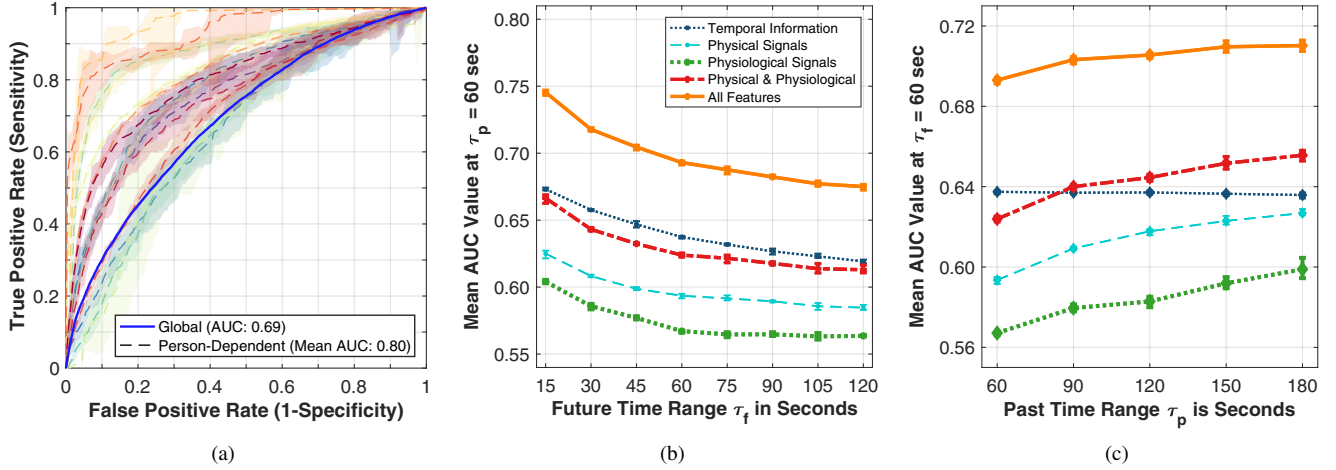


Fig. 1. (a) ROC curves with 90% confidence intervals to predict onset of aggression in the next minute, using all features from the past minute. The blue solid line represents the global prediction model, and each curve with dashed lines represents one of the person-dependent models. (b) Mean AUC values of the global models varying as a function of  $\tau_f$  using features from the past minute; and (c) as a function of  $\tau_p$  at  $\tau_f = 60$  seconds. For (b) and (c), each color represents one of the five feature subsets as indicated. Error bars represent minimum and maximum values across cross-validation repetitions.

TABLE I

AUC VALUES AVERAGED ACROSS CROSS-VALIDATION REPETITIONS IN PERSON-DEPENDENT PREDICTION MODELS WITH  $\tau_p = \tau_f = 60$  SECONDS (I.E., PREDICTING AGGRESSION ONSET FOR THE NEXT MINUTE, USING ACCUMULATED DATA FROM THE PAST ONE MINUTE).

Features	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	Mean	SD
Temporal	0.67	0.58	0.63	0.66	0.77	0.54	0.67	0.68	0.73	0.59	0.65	0.51	0.61	0.65	0.63	<b>0.64</b>	<b>0.07</b>
Physical	0.63	0.69	0.69	0.62	0.76	0.53	0.76	0.71	0.80	0.76	0.77	0.92	0.66	0.55	0.75	<b>0.71</b>	<b>0.10</b>
Physiological	0.76	0.57	0.76	0.56	0.82	0.61	0.80	0.61	0.83	0.90	0.77	0.86	0.58	0.65	0.76	<b>0.72</b>	<b>0.11</b>
Physical and Physiological	0.75	0.67	0.76	0.63	0.84	0.63	0.83	0.73	0.85	0.95	0.81	0.93	0.69	0.74	0.80	<b>0.77</b>	<b>0.10</b>
All Features Combined	0.76	0.67	0.79	0.67	0.88	0.67	0.85	0.79	0.89	0.96	0.82	0.93	0.70	0.75	0.80	<b>0.80</b>	<b>0.09</b>

TABLE II

AUC VALUES AVERAGED ACROSS CROSS-VALIDATION REPETITIONS IN PREDICTION MODELS WITH  $\tau_p = \tau_f = 60$  SECONDS. BY DEFINITION OF THE  $k$ -HYBRID MODELS,  $k = 0$  CORRESPONDS TO PERSON-DEPENDENT MODELS, WHEREAS  $k = 310$  CORRESPONDS TO THE GLOBAL MODEL.

Model	$k$	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	Mean	SD
Person-Dependent	0	0.76	0.67	0.79	0.67	0.88	0.67	0.85	0.79	0.89	0.96	0.82	0.93	0.70	0.75	0.80	<b>0.80</b>	<b>0.09</b>
$k$ -Hybrid	1	0.73	0.58	0.74	0.60	0.86	0.57	0.83	0.76	0.86	0.87	0.81	0.87	0.67	0.69	0.75	<b>0.75</b>	<b>0.11</b>
$k$ -Hybrid	10	0.74	0.57	0.75	0.62	0.86	0.54	0.83	0.76	0.87	0.86	0.81	0.87	0.67	0.69	0.74	<b>0.75</b>	<b>0.11</b>
$k$ -Hybrid	100	0.75	0.57	0.76	0.67	0.87	0.53	0.83	0.77	0.86	0.85	0.81	0.83	0.69	0.71	0.73	<b>0.75</b>	<b>0.10</b>
$k$ -Hybrid	200	0.76	0.59	0.73	0.70	0.87	0.52	0.83	0.77	0.87	0.84	0.80	0.83	0.68	0.71	0.72	<b>0.75</b>	<b>0.10</b>
$k$ -Hybrid	300	0.74	0.60	0.70	0.69	0.85	0.54	0.81	0.75	0.83	0.70	0.73	0.77	0.66	0.67	0.62	<b>0.71</b>	<b>0.09</b>
Global	310	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	<b>0.69</b>	—

with 0.75 for a  $k$  value as high as 200. While partially globalizing the model slightly reduced AUC values compared to person-dependent models, model training becomes computationally more efficient. Thus, the  $k$ -hybrid framework serves as a partially global model alternative that approximates person-dependent accuracies with shorter model training run time. Furthermore, although person-dependent models provide better accuracies, they cannot be used to predict aggression onset for a new patient for whom no prior data is available. From that perspective, hybrid models are desirable in that predictive power can be maintained by having some parameters trained globally, and then combined and updated as person-dependent training data becomes available.

#### IV. DISCUSSION

In the present study we demonstrate that naturalistically observed aggressive behavior in MV-ASD youth in a hospital inpatient setting can be predicted with high accuracy using preceding physiological and physical activity biosensor data and temporal information on recently observed aggressive episodes. We implemented linear classifier models with regularization over relevant time-series features during training. Our results demonstrate proof-of-concept, feasibility, and incipient validity predicting proximal aggression in this population and setting using biosignal data in global, person-dependent, and hybrid models.

As communicated to us by inpatient clinical staff, the potential benefit of avoiding or reducing a dangerous aggressive event is likely to outweigh the potential harm associated with a false positive in clinical practice. Predicting aggression onset during naturalistic observation in the upcoming 1 minute with at least 80% sensitivity is of high clinical value and could create new opportunities for preventing or mitigating aggression emergence, occurrence, and impact in MV-ASD [18].

Regarding future work, we will next model aggression onsets as a non-homogeneous Poisson process to assess whether they can yield more robust global prediction models. Using this framework, we will also evaluate whether regressing hazard rates from past observations can be performed through maximum likelihood estimation, comprising an iterative solution of several weighted linear regressions [21], and lead to improved predictive performance across upcoming time ranges.

#### ACKNOWLEDGMENT

This work was supported by grants from the Simons Foundation (SFARI 296318), the Nancy Lurie Marks Family Foundation, the National Institute on Deafness and Other Communication Disorders (P50 DC013027), and NSF (SCH-1622536, IIS-1118061).

#### REFERENCES

- [1] S. M. Kanne and M. O. Mazurek, "Aggression in children and adolescents with ASD: Prevalence and risk factors," *Journal of Autism and Developmental Disorders*, vol. 41, no. 7, pp. 926–937, 2011.
- [2] C. A. Farmer and M. G. Aman, "Aggressive behavior in a sample of children with autism spectrum disorders," *Research in Autism Spectrum Disorders*, vol. 5, no. 1, pp. 317–323, 2011.
- [3] K. Gray, C. Keating, J. Taffe, A. Brereton, S. Einfeld, and B. Tonge, "Trajectory of behavior and emotional problems in autism," *American Journal on Intellectual and Developmental Disabilities*, vol. 117, no. 2, pp. 121–133, 2012.
- [4] A. C. Woodman, M. R. Mailick, and J. S. Greenberg, "Trajectories of internalizing and externalizing symptoms among adults with autism spectrum disorders," *Development and Psychopathology*, vol. 28, no. 2, pp. 565–581, 2016.
- [5] S. W. Porges, "Physiological regulation in high-risk infants: A model for assessment and potential intervention," *Development and Psychopathology*, vol. 8, no. 1, pp. 43–58, 1996.
- [6] J. Groden, J. Cautela, S. Prince, and J. Berryman, "The impact of stress and anxiety on individuals with autism and developmental disabilities," in *Behavioral Issues in Autism*. Springer, 1994, pp. 177–194.
- [7] J. Groden, M. S. Goodwin, M. G. Baron, G. Groden, W. F. Velicer, L. P. Lipsitt, S. G. Hofmann, and B. Plummer, "Assessing cardiovascular responses to stressors in individuals with autism spectrum disorders," *Focus on Autism and Other Developmental Disabilities*, vol. 20, no. 4, pp. 244–252, 2005.
- [8] I. L. Cohen, J. H. Yoo, M. S. Goodwin, and L. Moskowitz, "Assessing challenging behaviors in autism spectrum disorders: Prevalence, rating scales, and autonomic indicators," in *International Handbook of Autism and Pervasive Developmental Disorders*. Springer, 2011, pp. 247–270.
- [9] R. G. Romanczyk, S. Lockshin, and J. O'Connor, "Psychophysiology and issues of anxiety and arousal," in *Self-injurious Behavior*. Springer, 1992, pp. 93–121.
- [10] J. Groden, M. G. Baron, and G. Groden, "Assessment and coping strategies," in *Stress and Coping in Autism*. Oxford University Press, 2006.
- [11] M. S. Goodwin, J. Groden, W. F. Velicer, L. P. Lipsitt, M. G. Baron, S. G. Hofmann, and G. Groden, "Cardiovascular arousal in individuals with autism," *Focus on Autism and Other Developmental Disabilities*, vol. 21, no. 2, pp. 100–123, 2006.
- [12] T. W. Benevides and S. J. Lane, "A review of cardiac autonomic measures: considerations for examination of physiological response in children with autism spectrum disorder," *Journal of Autism and Developmental Disorders*, vol. 45, no. 2, pp. 560–575, 2015.
- [13] C. Liu, K. Conn, N. Sarkar, and W. Stone, "Physiology-based affect recognition for computer-assisted intervention of children with autism spectrum disorder," *International Journal of Human-Computer Studies*, vol. 66, no. 9, pp. 662–677, 2008.
- [14] A. Kushki, E. Drumm, M. P. Mobarak, N. Tanel, A. Dupuis, T. Chau, and E. Anagnostou, "Investigating the autonomic nervous system response to anxiety in children with autism spectrum disorders," *PLoS one*, vol. 8, no. 4, p. e59730, 2013.
- [15] A. Kushki, A. Khan, J. Brian, and E. Anagnostou, "A kalman filtering framework for physiological detection of anxiety-related arousal in children with autism spectrum disorder," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 3, pp. 990–1000, 2015.
- [16] K. C. Welch, "Physiological signals of autistic children can be useful," *IEEE Instrumentation & Measurement Magazine*, vol. 15, no. 1, 2012.
- [17] S. Lydon, O. Healy, and M. Dwyer, "An examination of heart rate during challenging behavior in autism spectrum disorder," *Journal of Developmental and Physical Disabilities*, vol. 25, no. 1, pp. 149–170, 2013.
- [18] M. S. Goodwin, O. Özdenizci, C. Cumanasoiu, P. Tian, Y. Guo, A. Stedman, C. Peura, C. Mazefsky, M. Siegel, D. Erdoğmuş, and S. Ioannidis, "Predicting imminent aggression onset in minimally-verbal youth with autism spectrum disorder using preceding physiological signals," in *Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare*. ACM, 2018.
- [19] C. Lord, M. Rutter, P. DiLavore, S. Risi, K. Gotham, and S. Bishop, "Autism diagnostic observation schedule - second edition (ADOS-2)," *Los Angeles, CA: Western Psychological Corporation*, 2012.
- [20] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiological Measurement*, vol. 28, no. 3, p. R1, 2007.
- [21] W. A. Massey, G. A. Parker, and W. Whitt, "Estimating the parameters of a nonhomogeneous poisson process with linear rate," *Telecommunication Systems*, vol. 5, no. 2, pp. 361–388, 1996.