國立臺灣大學電機資訊學院生醫電子與資訊學研究所

碩士論文

Graduate Institute of Biomedical Electronics and Bioinformatics
College of Electrical Engineering and Computer Science
National Taiwan University
Master Thesis

使用穿戴式裝置資料運用人工智慧模型預測自閉症類群障礙症病患之情緒徵兆

Using Wearable Device and AI to Predict Mood Symptoms in Autism Spectrum Disorder

呂振均
Chen-Chun Lu

指導教授：賴飛羆 博士、簡意玲 醫師
Advisor: Feipei Lai, Ph.D., Yi-Ling Chien, M.D., Ph.D.

中華民國 112 年 7 月

July 2023

doi:10.6342/NTU202302181

論文口試委員審定書

國立臺灣大學碩士學位論文
口試委員會審定書
MASTER'S THESIS ACCEPTANCE CERTIFICATE
NATIONAL TAIWAN UNIVERSITY

（論文中文題目）(Chinese title of Master's thesis)

使用穿戴式裝置資料運用人工智慧模型預測自閉症譜系障礙症病患之情緒徵兆

（論文英文題目）(English title of Master's thesis)

Using Wearable Device and AI to Predict Mood Symptoms in Autism Spectrum Disorder

本論文係 呂振均 (姓名) R10945049（學號）在國立臺灣大學生醫電子與資訊學研究所 完成之碩士學位論文，於民國 112 年 5 月 30 日承下列考試委員審查通過及口試及格，特此證明。

The undersigned, appointed by the Institute of Biomedical Electronics and Bioinformatics on 30 (date) 5 (month) 2023 (year) have examined a Master's thesis entitled above presented by Chen Chun Lu (name) R10945049 (student ID) candidate and hereby certify that it is worthy of acceptance.

口試委員 Oral examination committee:

（指導教授 Advisor）

系主任/所長 Director:

doi:10.6342/NTU202302181

誌謝

特別感謝賴飛羆老師和共同指導的簡意玲醫師，在每次的實驗中給予許多臨床上與資料處理的建議，覺得受益良多，有了老師與醫師的長時間協助才因此得以完成該篇論文。

同時也特別感謝謝明憲醫師與實驗室的學長們的協助，給予我關於實驗上的建議與指導，一同討論出解決方法，得以在實驗的表現上更好一些。

doi:10.6342/NTU202302181

中文摘要

在這項研究中，我們使用從穿戴式設備（Garmin Vivosmart 4）收集的生理數據，其中包含心率、活動和睡眠時間的長度，通過機器學習演算法構建自閉症譜系失調的預測模型。我們可以根據可解釋模型的結果推薦治療方法或預防方法。

本研究招募了 14 名受試者。參與者是 20 至 55 歲的患者，他們被診斷患有自閉症類群障礙症，該疾病基於美國精神醫學會（American Psychiatric Association, APA）的精神疾病診斷和統計手冊（DSM-5）。使用的數據由兩部分組成。問卷數據包括貝克憂鬱量表（Beck Depression Inventory, BDI）和躁鬱症評定量表（Young Mania Rating Scale, YMRS），分別用於評估憂鬱症和躁症，以及用於模型構建的數位生物標記。我們使用 6 種機器學習演算法來構建預測模型。

預測模型在憂鬱發作的測試集上達到了 79%的準確率，0.88 AUROC 和 0.88 f1 得分。通過可解釋的模型 SHAP 我們發現相對較低的靜息心率、高活動與憂鬱症有關，並可能預測憂鬱症的發作。

總之，我們可以使用從可解釋模型中獲得的資訊來提供早期的臨床評估和一些預防治療。

關鍵詞：自閉症類群障礙症，可穿戴設備，機器學習，可解釋模型，預測

doi:10.6342/NTU202302181

Abstract

In this study, we utilized digital biomarkers collected from wearable devices (Garmin Vivosmart 4), which include heart rate, activity level, and length of sleeping hours, to construct predictive models of Autism Spectrum Disorder using machine learning algorithms. Based on the results obtained from interpretable models, we can recommend preventive methods.

Fourteen participants were recruited for this study. These participants were patients aged 20 to 55 who had received a diagnosis of ASD, according to the criteria outlined in the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders (DSM-5). The data used for analysis consisted of two parts: questionnaire data, including the Beck Depression Inventory (BDI) and the Young Mania Rating Scale (YMRS), used to assess depression and mania respectively, and digital biomarkers for model development. We employed six machine learning algorithms to construct a predictive model.

The predictive model achieved 79% accuracy, 0.88 AUROC, and 0.88 F1 score on the test set of depressive episodes. Using the explainable model SHAP, we discovered that relatively low resting heart rate and high activity were associated with depression and could potentially predict the onset of depressive episodes.

In conclusion, the information obtained from interpretable models can be used to provide earlier clinical evaluations for prevention.

Keywords: autism spectrum disorder, wearable device, machine learning, explainable model, prediction

doi:10.6342/NTU202302181

Contents

doi:10.6342/NTU202302181

doi:10.6342/NTU202302181

LIST OF FIGURES

doi:10.6342/NTU202302181

doi:10.6342/NTU202302181

LIST OF TABLES

doi:10.6342/NTU202302181

doi:10.6342/NTU202302181

Chapter 1 Introduction

1.1 Background

Autism spectrum disorder (ASD) is a developmental and neurological disorder that involves communication and behavior [1]. Repetitive behaviors and restricted interests are also common characteristics of ASD. Due to its increasing prevalence, ASD affects a significant number of individuals worldwide. The abilities of autism patients vary significantly from those of others. For example, some patients need a lot of help in their lives, while others can work and live easily with no support. In most cases, symptoms may appear in the first 2.5 years [2].

The primary cause of autism cannot be clearly identified at this time, but some factors are associated with the development of ASD, including having a sibling with ASD [3], having older parents [4], and having certain genetic conditions (such as fragile X syndrome or Down syndrome) [5]. ASD occurs in all ethnic and racial groups, and it is more common in men than in women, with a ratio of more than 4 to 1 [6]. Thus, autism can manifest differently in different individuals, making it challenging to identify a single definitive cause.

The mood symptoms in ASD can vary widely among individuals and may present differently across the lifespan [7]. While depression and anxiety are the most commonly observed mood symptoms, other conditions such as bipolar disorder and irritability are also prevalent in this population [8]. The presence of mood symptoms in individuals with ASD can significantly impact their overall well-being and functioning [9]. Additionally, mood symptoms in ASD can overlap with the core characteristics of the condition, making diagnosis and treatment more complex. In recent years, various therapies and

doi:10.6342/NTU202302181

interventions aim to address specific difficulties related to communication and behavioral patterns, empowering individuals to reach their full potential and lead fulfilling lives.

1.2 Related Work

The integration of wearables and mobile technologies in autism interventions holds great promise, thanks to their portability, accessibility, and ability to collect real-time data [10]. These devices include fitness trackers, smartwatches, and smart glasses, which measure various physiological and behavioral functions like heart rate, skin conductance, and eye gaze. These mobile solutions have the potential to revolutionize how we approach ASD interventions, from assisting in communication and social skill development to providing sensory regulation tools and facilitating behavioral monitoring and tracking.

Types of wearables used in ASD interventions can be divided into three categories: head-worn technology, body-worn devices, and accessory and clothing-based technology. Head-worn technology [11] includes glasses or devices worn in the form of headbands or hats. Various studies have explored smart or augmented reality glasses that capture facial expressions and eye gaze patterns to provide real-time feedback to adolescents with ASD, supporting their social functioning and communication skills. Some systems also employ machine learning and facial expression recognition software to aid emotion recognition in ASD individuals [12]. Body-worn devices utilize sensors like accelerometers and electromyography (EMG) sensors placed on the body to monitor and measure specific movements and physiological signals. Certain studies focus on using accelerometers to detect stereotypical or self-injurious [13] behaviors, while others utilize EMG sensors to assess smiling behavior during social interactions. Accessory and clothing-based technology involve devices like wristbands or shirts with embedded sensors to measure

doi:10.6342/NTU202302181

physiological signals such as skin conductance, touch, and heart rate. Certain wrist-worn devices aim to quantify touch events between individuals, while shirts with embedded sensors facilitate the assessment of physiological activity during social interactions or therapeutic interventions.

Table 1.1 provides a brief summary of the use of wearable technology to measure mood symptoms in ASD. The studies presented in this table demonstrate the feasibility and potential usefulness of wearables in ASD interventions. The review encompasses a diverse range of wearable and mobile technologies, including smartwatches, smartphone applications, augmented reality, virtual reality, and other emerging tools that have been explored in the context of ASD interventions. By examining the strengths and limitations of these technologies, we aim to provide a comprehensive understanding of their potential to facilitate communication, social skills, sensory processing, and other aspects relevant to individuals on the autism spectrum.

Table 1.1 A brief summary about the use of wearable technology to measure mood symptoms in ASD

Author (year)	Wearable technology	Participant description	Quality
Goodwin et al., (2014)	3 wireless accelerometers worn on the left and right wrists and torso	Children and young adults with ASD, n = 6, 12-20 years	Strong (81%)
Di Palma et al., (2017)	ECG chest belt	Male children with ASD, n = 5, 6-8 years	Adequate (63%)
Magrelli et al., (2013)	WearCam (cameras and mirror mounted on headband)	Children with ASD, n = 14, 2-11 years	Strong (85%)
Takahashi et al., (2016)	Smart Clothe ECG sensor embedded in sleeve cuff	Male children with ASD, n = 4, 3.8 – 5.5 years	Adequate (63%)
Daniels, Schwartz et al., (2018)	SuperPower Glass system (smart glasses)	Children with ASD, n = 14, Mean age: 6-12 years	Quantitative Strong (86%)

doi:10.6342/NTU202302181

Sahin et al., (2018b)	Brain Power Autism System (Smart glasses)	Children with ASD, n = 8 (7 male, 1 female), 6.7 – 17.2 years	Adequate (60%)

The studies discussed in this paragraph demonstrate the feasibility and potential usefulness of wearables in ASD interventions, particularly in improving social engagement and interaction. However, further research is necessary to determine their efficacy and address potential challenges, such as comfort and user acceptability. The wide variety of available wearables emphasizes the growing interest in leveraging technology to support individuals with autism in various aspects of their lives. Overall, these findings point towards the promising role of wearables in ASD interventions, but ongoing investigation is essential to fully understand their effectiveness and address any limitations they may have.

1.3 Objective

In this study, we will utilize digital biomarkers collected by smart wearable devices to develop predictive models for mood symptoms (i.e., manic and depressive episodes) using machine learning algorithms [14]. The features used for modeling will include the length of sleep stages, daily activity, and heartbeats. Additionally, we will employ two questionnaires to supplement the results of the symptoms. Our hypotheses are that lower activity levels, lower heartbeats, and longer sleeping hours may predict the development of depressive symptoms, while higher activity levels, lower heartbeats, and shorter sleeping hours may predict the development of manic symptoms. Early detection is crucial to minimize the negative impact on the individual's well-being.

doi:10.6342/NTU202302181

Chapter 2 Method

2.1 Participants

There were 17 participants in this study, consisting of 8 men and 9 women, as per the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) guidelines (American Psychiatric Association, 2013). The participants were patients aged between 20 and 55 years. Their mean age was 28 years old, and their educational levels ranged from middle school to graduate school, with the majority having a college degree. The participants were recruited from the Psychiatric Department of the National Taiwan University Hospital (NTUH) between October 2020 and June 2023.

Before implementing the study, it was approved by the Research Ethics Committee at the National Taiwan University Hospital. Every participant was thoroughly explained the study procedure, and their data confidentiality was assured. Informed consent was obtained from each participant prior to their involvement in the study.

2.2 Data Collection

The data in this study consisted of two parts: digital biomarkers and questionnaire data. The questionnaire data were used for labeling, while the digital biomarkers were used for training.

2.2.1 Digital biomarkers

Digital biomarkers were collected using wearable devices (Garmin Vivosmart 4). They included three types of features: activity, sleep length, and heart rate. More specifically, the activity features comprised the number of steps and floors climbed, total

doi:10.6342/NTU202302181

sleep hours, three sleep stages, and heart rate.

2.2.2 Questionnaire data

For this experiment, we employed two questionnaires to complement our digital biomarkers. The first questionnaire was based on the Beck Depression Inventory (BDI), used to measure characteristics and symptoms of depression. The second questionnaire was based on the Young Mania Rating Scale (YMRS), utilized for self-report evaluation of manic symptoms. These questionnaires covered various aspects such as activity, sleep quality, emotions, sense of illness, suicidal thoughts, and appetite. Due to the different types of questions, we used these two questionnaires as our labels. To enable early intervention in detecting mood symptoms during depressive and manic episodes, we set the criteria for mild depression and mild mania as having a total score higher than 13 on the BDI and YMRS, respectively.

2.3 Workflow Architecture

The whole workflow architecture can be divided into data preprocessing, model training, model testing, and the interpretable model Shapely Additive exPlanations (SHAP). In the experiment, we will train two different models for manic and depressive episodes in ASD. We utilized the Python, scikit-learn, and SHAP packages.

2.4 Data Preprocessing

Because the questionnaire was filled out once a week, the missing data was backfilled for 7 days [14]. (See Figure 2.1) Second, to prevent overfitting, we removed all the missing values from the dataset. Then, the datasets were split into two parts: the training sets and the testing sets, using a stratified 80-20 random split to avoid overfitting.

doi:10.6342/NTU202302181

Additionally, to consider personalized features, we transformed the original features into differences from personal baseline values. Table 2.1 displays all the features adopted in the prediction model for manic and depressive symptoms.

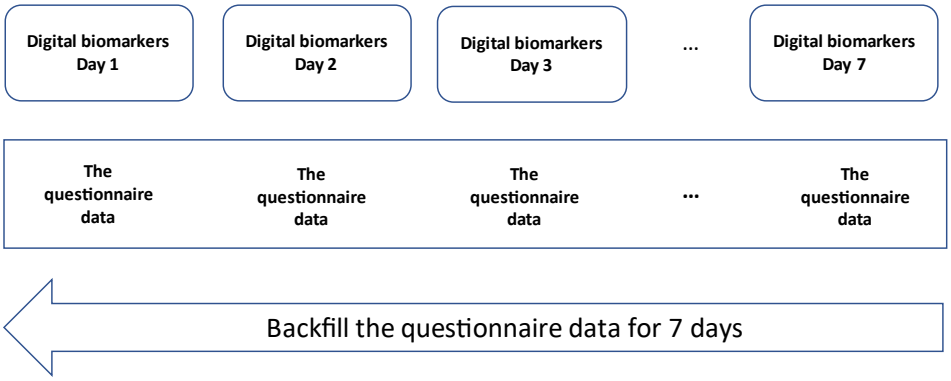

Figure 2.1 Backfill the label form the questionnaire data for 7 days

Table 2.1 Features for model training

	Feature Name	Unit
Activity		
	Steps	/day
	Distance in Meters	/day
	Climbed Floors	/day
Heart Rate		
	Max Heart Rate	/min
	Min Heart Rate	/min
	Average Heart Rate	/min
	Resting Heart Rate	/min
Sleep Status		
	Total Sleep Duration	hr/day
	REM Sleep Duration	hr/day
	Light Sleep Duration	hr/day
	Deep Sleep Duration	hr/day
	Awake Duration	hr/day
Difference from Baseline		
	Difference from Average of Steps	/day
	Difference from Average	/day

doi:10.6342/NTU202302181

	of Distance in Meters	
	Difference from Average of Max Heart Rate	/min
	Difference from Average of Min Heart Rate	/min
	Difference from Average of Average Heart Rate	/min
	Difference from Average of Resting Heart Rate	/min
	Difference from Average of Total Sleep Duration	hr/day

2.5 Imbalanced Data

The data in the dataset is imbalanced, which is a common occurrence in the real world [15], especially in medical data. When dealing with imbalanced data, machine learning classifiers tend to exhibit bias toward the majority class. To address this imbalanced data problem, we employed the under-sampling strategy.

Under-sampling [16] is a technique used to balance datasets by retaining the entire training data from the minority class and reducing the size of the data from the majority class. This approach allows us to create a more balanced training set, ensuring that the machine learning models can learn effectively from both classes and make more accurate predictions for the minority class. By applying under-sampling, we aim to improve the performance and fairness of our predictive models when dealing with imbalanced data, thereby enhancing the reliability and practicality of our findings in medical applications.

2.6 Machine Learning Algorithms

In our study, we employed various machine learning models to develop our predictive models, including logistic regression, decision tree [17], K-Nearest Neighbors [18], Random Forest [19], Adaptive Boosting [20], and Extreme Gradient Boosting (XGBoost). Each of these models offers unique strengths and capabilities, contributing to

doi:10.6342/NTU202302181

a comprehensive approach in analyzing the data and predicting manic and depressive symptoms in ASD individuals. The diversity of these models allowed us to explore different aspects of the data and obtain well-rounded insights into the predictive factors of mood symptoms.

2.7 K-Fold Cross-Validation

In K-fold cross-validation, we would divide the data into k equal parts and do training. For example, if k = 10, which is mean it would train the same model 10 times, in each training, it would pick up 9 equal parts as training set, the only 1 part as validation set. Hence, there would 10 different error of validation in this training. Then, we average the 10 errors as final result. The method can reduce the variation of the model and uses the data efficiently. In this study, 5-fold cross-validation (Figure 2.2) was applied.

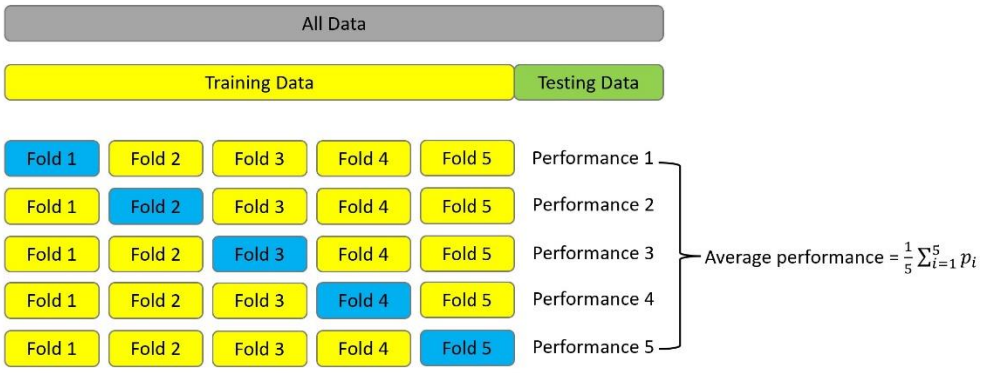

Figure 2.2 Flow chart of 5-fold cross-validation

2.8 Model Assessment

We utilize several metrics for assessing the performance of the prediction model [22], including Accuracy, Sensitivity, Specificity, Precision, F1 score, and Area Under the Receiver Operating Curve (AUROC) [21] (see Table 2.2 and Figure 2.3).

doi:10.6342/NTU202302181

Sensitivity, also known as recall, measures the percentage of correctly recognized correct example data among all actual correct example data. On the other hand, Specificity focuses on the percentage of correctly recognized incorrect example data among all actual incorrect example data. Precision, on the other hand, evaluates the percentage of data marked as positive that are actually positive. The F1 score represents the harmonic mean of precision and recall, combining both metrics to give an overall measure of the model's performance.

By utilizing this comprehensive set of evaluation metrics, we gain a detailed understanding of the prediction model's capabilities and effectiveness in correctly identifying manic and depressive episodes in individuals with ASD. These metrics allow us to make well-informed decisions regarding the model's performance and to optimize it for accurate and reliable predictions, ensuring better clinical applications and improved patient outcomes.

Table 2.2 Confusion matrix

	Actual Positive	Actual Negative
Predicted Positive	TP (True Positive)	FP (False Positive)
Predicted Negative	FN (False Negative)	TN (True Negative)

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad Precision = \frac{TP}{TP + FP}$$

$$Sensitivity = \frac{TP}{TP + FN} \qquad F1\ Score = \frac{2 * Precision * Sensitivity}{Precision + Sensitivity}$$

$$Specificity = \frac{TN}{TN + FP}$$

Figure 2.3 Equations Model assessment metrics

doi:10.6342/NTU202302181

2.9 Explainable Machine Learning Model

In this study, we utilized the Shapely Additive exPlanations (SHAP) technique to comprehensively explain the output of our machine learning models [23]. Through SHAP, we not only gained insights into the inner workings of the model and understood how it made predictions but also acquired valuable knowledge to provide additional suggestions for clinical issues. The transparency offered by SHAP significantly advances the understanding of our machine learning models, making them more interpretable and facilitating their potential application in real-world scenarios. With this enhanced interpretability, we can harness the power of machine learning to better address clinical challenges and improve the overall quality of care for individuals.

doi:10.6342/NTU202302181

Chapter 3 Result

3.1 Data Description

In this study, Table 3.1 provides a comprehensive overview of all patients' case descriptions, including details such as the start date of their participation, the number of digital biomarkers collected, the frequency of questionnaires completed, and the medication used per week. These case descriptions are essential for understanding the data collection process and the characteristics of each participant in the study.

By examining Table 3.1, we can gain valuable insights into the duration and extent of data collection for each patient, the frequency of their questionnaire responses, and the medication regimen they were following on a weekly basis. This information is vital for comprehending the variations in data availability and medication compliance among the study participants, allowing us to factor these elements into our analysis and interpretation of the results.

Table 3.1 The data description of all cases

Case #	Start date	The datapoints of digital biomarkers	The datapoints of questionnaire 1	The datapoints of questionnaire 2	The medicine of the case used per week
Case 1	2021/03/01	826	14	8	0
Case 2	2021/03/15	821	82	51	7
Case 3	2021/04/26	270	19	15	-
Case 4	2021/10/18	573	80	69	7

doi:10.6342/NTU202302181

Case 5	2021/11/22	398	14	9	-
Case 6	2021/11/30	80	-	-	-
Case 7	2021/12/12	433	15	15	7
Case 8	2021/12/20	420	40	19	0
Case 9	2021/12/30	510	46	35	7
Case 10	2022/02/01	491	63	59	7
Case 11	2022/02/22	390	44	41	7
Case 12	2022/05/04	347	32	28	0
Case 13	2022/05/23	324	54	44	7
Case 14	2022/05/30	57	15	9	7
Case 15	2022/10/17	228	28	29	0
Case 16	2022/10/21	225	22	24	7
Case 17	2022/11/08	258	5	5	7

3.2 Patient Characteristics

All the patients' characteristics were categorized as either non-depressive/manic or depressive/manic, as presented in Table 3.2 and Table 3.3. Upon analyzing individual features' data, we observed that heart rate data exhibited a slightly higher frequency in the depressive episode among the non-depressive labels.

Table 3.2 Characteristics of enrolled patients with a manic episode in ASD

	Non-manic (n = 3,202)						Manic (n = 185)				
	mean	Std.	Min.	median	Max.		mean	Std.	Min.	median	Max.
Activity											
Steps	6,898	4,246	101	6,145	29,725		9,841	5,762	519	8,875	30,118
Distance	4,730	2,849	74	4,310	24,841		6,838	4,198	392	5,930	23,435

doi:10.6342/NTU202302181

(m)											
Floors climbed	9	10	0	7	99		11	9	0	10	51
Heart Rate											
Max Heart Rate	128	13	92	127	194		130	13	92	128	194
Min Heart Rate	54	8	33	54	75		50	5	41	50	66
Average Heart Rate	76	9	50	77	100		77	6	58	78	96
Resting Heart Rate	60	8	38	61	80		57	5	45	57	75
Sleep State											
Total Sleep Duration	7.97	2.44	0.81	7.95	17.13		7.64	2.08	2.8	7.63	15.8
Deep Sleep Duration	1.09	0.92	0	0.93	11.90		0868	1.11	0	0.4	5.8
Light Sleep Duration	5.34	2.02	0	5.11	13.53		4.53	1.66	0.85	4.27	12.97
REM Sleep Duration	1.50	0.98	0	1.37	7.30		2.18	1.56	0	2.13	7.63
Awake Duration	0.24	0.31	0	0.13	3.90		0.15	0.19	0	0.1	1.06

Table 3.3 Characteristics of enrolled patients with a depressive episode in ASD

	Non-depressive (n = 698)						Depressive (n = 1,953)				
	mean	Std.	Min.	Median	Max.		mean	Std.	Min.	Median	Max.
Activity											
Steps	8,608	5,239	153	8,091	30,010		6,708	3,936	101	6,070	30,118
Distance (m)	5,632	3,373	108	5,298	21,294		4,721	2,845	74	4,286	24,841
Floors climbed	12	11	0	9	100		7	9	0	4	99
Heart Rate											
Max Heart Rate	130	13	92	129	188		128	13	92	126	194

doi:10.6342/NTU202302181

Min Heart Rate	53	7	36	53	72		53	8	34	53	75
Average Heart Rate	76	7	50	77	98		75	9	55	77	100
Resting Heart Rate	59	8	39	61	80		59	8	40	61	79
Sleep State											
Total Sleep Duration	8.09	1.98	0.81	7.9	14.62		7.7	2.52	1.38	7.83	17.13
Deep Sleep Duration	1.05	0.89	0	0.93	10.11		1.12	0.97	0	0.95	11.90
Light Sleep Duration	5.40	1.74	0.46	5.2	11.58		5.10	2.04	0	4.86	13.53
REM Sleep Duration	1.59	0.96	0	1.53	7.63		1.44	1.05	0	1.23	7.30
Awake Duration	0.17	0.21	0	0.1	1.57		0.24	0.30	0	0.11	2.52

In Figure 3.1 and Figure 3.2, the correlation between each pair of features is displayed. The color gradient represents the strength of the correlation, where shades closer to red indicate a stronger relationship between the features, while lighter colors (closer to white) signify a weaker or non-existent relationship. The features marked with 'Δ' denote personalized features, which are obtained by transforming the original features into differences from personal baseline values.

To provide more specific examples, we observed that in the case of steps, it exhibits a higher correlation with distance compared to other features. Additionally, the minimum heart rate demonstrates a stronger relationship with both the average heart rate (beats per minute) and resting heart rate (beats per minute) than other features.

doi:10.6342/NTU202302181

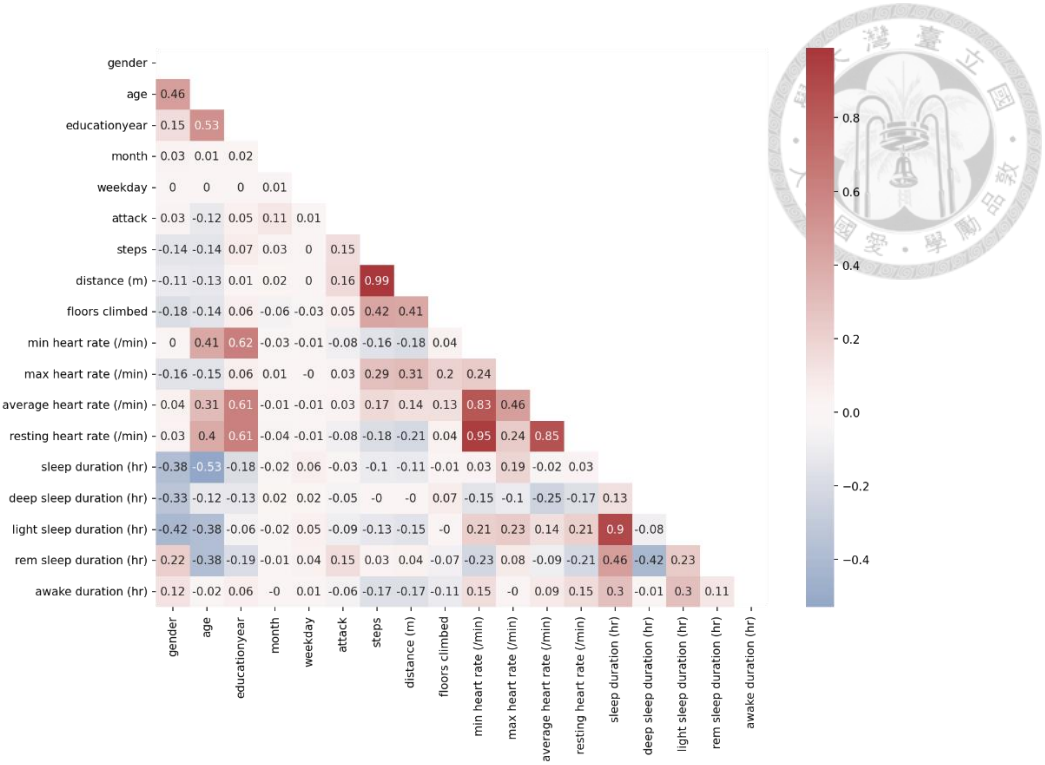

Figure 3.1 The correlation matrix for a manic episode in ASD

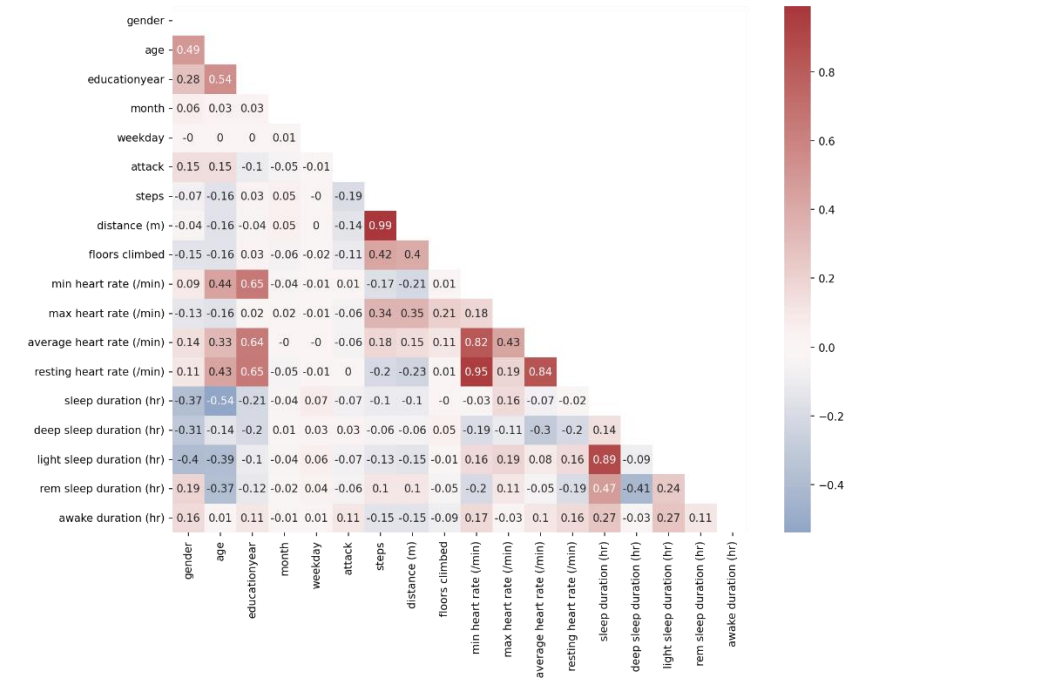

Figure 3.2 The correlation matrix for a depressive episode in ASD

Figure 3.3 illustrates the distribution of YMRS scores, comprising a total of 849 replies in this questionnaire. The YMRS scores range from a minimum of 0 to a maximum

doi:10.6342/NTU202302181

of 56. Furthermore, the average score of all responses is 3.69, and the median score is 2.

This graphical representation allows us to gain a comprehensive understanding of the YMRS scores' distribution among the participants. The range of scores and the central tendency measures (average and median) provide valuable insights into the variability and typical values of manic symptoms reported by the individuals in the study.

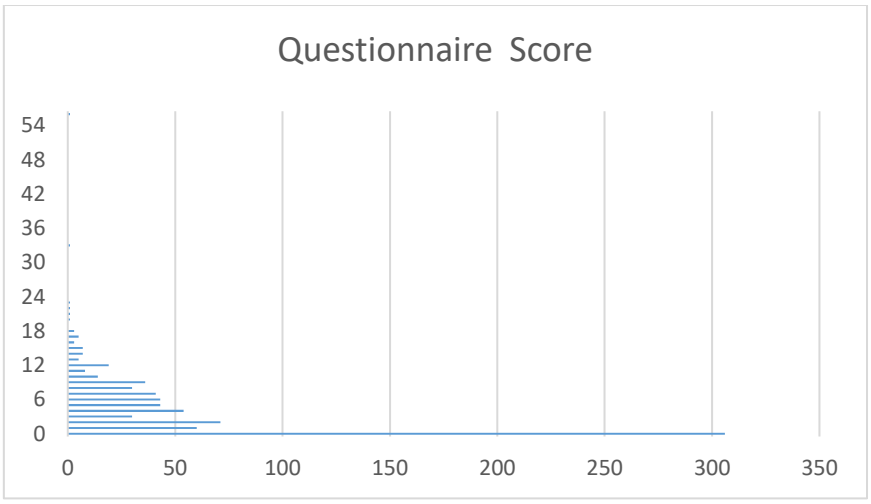

Figure 3.3 The distribution of YMRS scores

Figure 3.4 displays the distribution of BDI scores, comprising a total of 564 replies in this questionnaire. The BDI scores range from a minimum of 0 to a maximum of 53. Additionally, the average score of all responses is 11.10, and the median score is 6. This graphical representation allows us to comprehensively visualize the distribution of BDI scores among the participants. This information is essential for understanding the severity and prevalence of depressive symptoms in the sample population, aiding in the interpretation and implications of our research findings.

doi:10.6342/NTU202302181

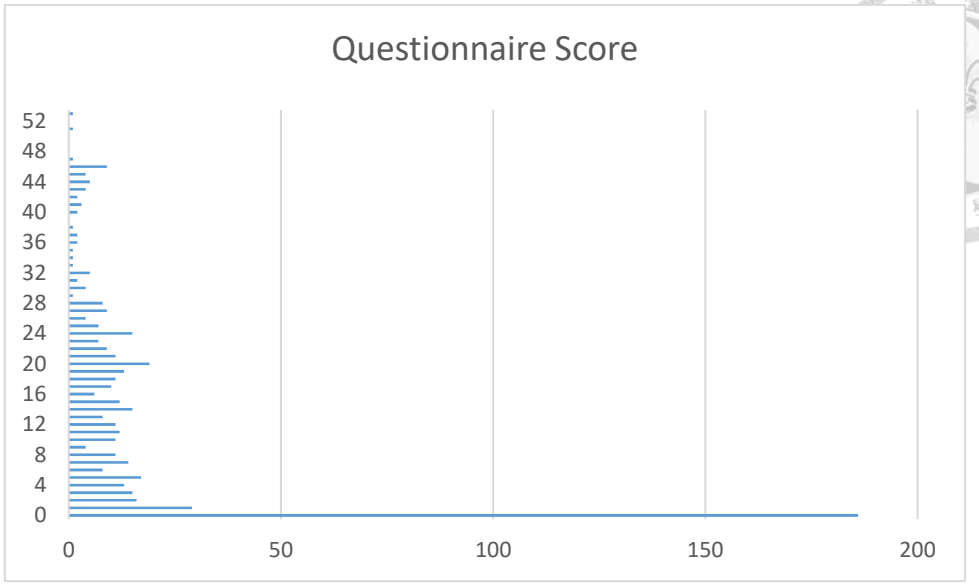

Figure 3.4 The distribution of BDI scores

3.3 Prediction Model

3.3.1 Prediction model for a manic episode in ASD

Based on the clinical condition and correlation matrix of features, we selected 15 digital biomarker features for model training and subsequently removed 5 features: deep sleep duration, light sleep duration, REM sleep duration, awake sleep duration, and minimum heart rate. By considering the clinical context and the relationships among the features, we optimized the model's performance, leading to more accurate predictions of manic episodes in individuals with autism. The performance of the proposed models for manic episodes in autism was evaluated using a 5-fold cross-validation with the YMRS score questionnaire, and the results are presented in Table 3.4.

doi:10.6342/NTU202302181

Table 3.4 Performance of the proposed models for a manic episode in ASD on 5-fold cross-validation with the difference from baseline features

Model	Accuracy	AUROC	Sensitivity	Specificity	Precision	F1 score
Logistic Regression	0.75	0.90	0.96	0.67	0.50	0.65
Decision Tree	0.90	0.92	0.78	0.93	0.80	0.79
KNN	0.82	0.76	0.64	0.88	0.64	0.64
AdaBoost	0.83	0.87	0.61	0.91	0.68	0.64
Random Forest	0.71	0.83	0.86	0.66	0.46	0.60
XGBoost	0.89	0.93	0.86	0.90	0.74	0.80

Then, the performance of the proposed models for manic episodes in autism was evaluated on the testing set using the YMRS score questionnaire, and the results are presented in Table 3.5.

Table 3.5 Performance of the proposed models for a manic episode in ASD on the testing set with the difference from baseline features

Model	Accuracy	AUROC	Sensitivity	Specificity	Precision	F1 score
Logistic Regression	0.73	0.92	0.97	0.71	0.16	0.28
Decision Tree	0.94	0.92	0.70	0.95	0.46	0.55
KNN	0.88	0.72	0.54	0.90	0.24	0.33
AdaBoost	0.93	0.92	0.62	0.95	0.40	0.48
Random Forest	0.76	0.87	0.84	0.75	0.16	0.27
XGBoost	0.91	0.96	0.86	0.91	0.37	0.52

The purpose of the ROC Curve is to provide an overview of the pretrained model's classifier performance. As shown in Figure 3.5, Logistic Regression, AdaBoost, and

doi:10.6342/NTU202302181

XGBoost demonstrate better performance compared to the other three models.

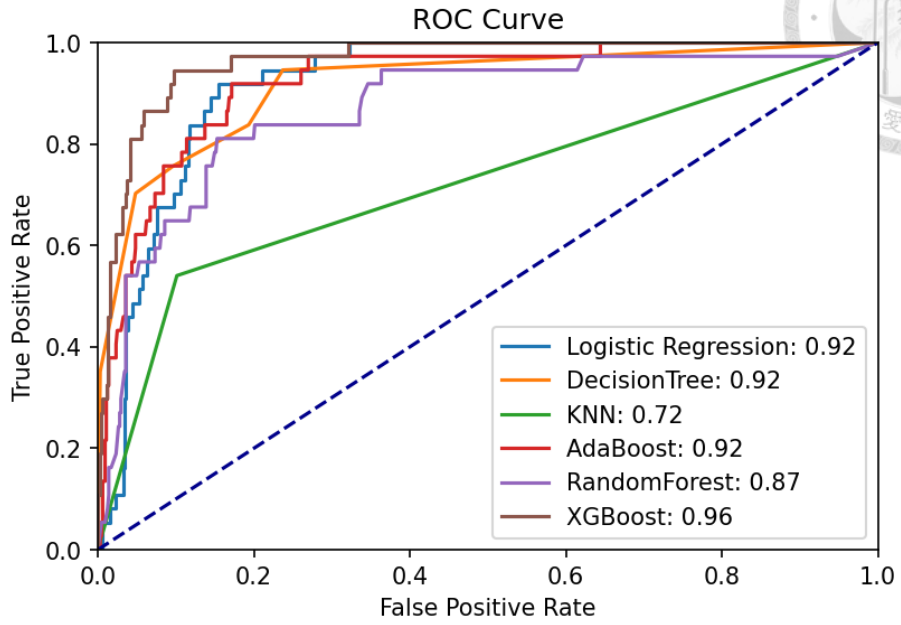

Figure 3.5 The ROC Curve of the proposed models for a manic episode in ASD on the testing set with the difference from baseline features

3.3.2 Prediction model for a manic episode in ASD with no demographic data

Based on the clinical condition and correlation matrix of features, and without considering demographic data, we selected 15 digital biomarker features for model training. Subsequently, we removed 5 features: deep sleep duration, light sleep duration, REM sleep duration, awake sleep duration, and minimum heart rate. The performance of the proposed models, considering no demographic data, for manic episodes in autism was evaluated using a 5-fold cross-validation with the YMRS score questionnaire, and the results are presented in Table 3.6. The feature selection process aimed to emphasize the most relevant and informative biomarkers, enhancing the model's performance.

doi:10.6342/NTU202302181

Table 3.6 Performance of the proposed models with no demographic data for a manic episode in ASD on 5-fold cross-validation with the difference from baseline features

Model	Accuracy	AUROC	Sensitivity	Specificity	Precision	F1 score
Logistic Regression	0.64	0.54	0.94	0.54	0.41	0.57
Decision Tree	0.80	0.85	0.58	0.87	0.62	0.59
KNN	0.82	0.76	0.64	0.88	0.64	0.64
AdaBoost	0.78	0.81	0.43	0.90	0.58	0.49
Random Forest	0.66	0.81	0.87	0.60	0.42	0.57
XGBoost	0.80	0.87	0.77	0.81	0.58	0.66

Then, the performance of the proposed models, without considering demographic data, for manic episodes in autism was evaluated on the testing set using the YMRS score questionnaire, and the results are presented in Table 3.7.

Table 3.7 Performance of the proposed models with no demographic data for a manic episode in ASD on the testing set with the difference from baseline features

Model	Accuracy	AUROC	Sensitivity	Specificity	Precision	F1 score
Logistic Regression	0.58	0.85	0.97	0.56	0.11	0.20
Decision Tree	0.85	0.83	0.65	0.87	0.22	0.33
KNN	0.88	0.72	0.54	0.90	0.24	0.33
AdaBoost	0.91	0.83	0.49	0.93	0.30	0.37
Random Forest	0.69	0.85	0.78	0.68	0.12	0.22
XGBoost	0.88	0.92	0.84	0.88	0.29	0.43

The purpose of the ROC Curve is to provide an overview of the pretrained model's classifier performance. As shown in Figure 3.6, Logistic Regression, Random Forest, and XGBoost demonstrate better performance compared to the other three models.

doi:10.6342/NTU202302181

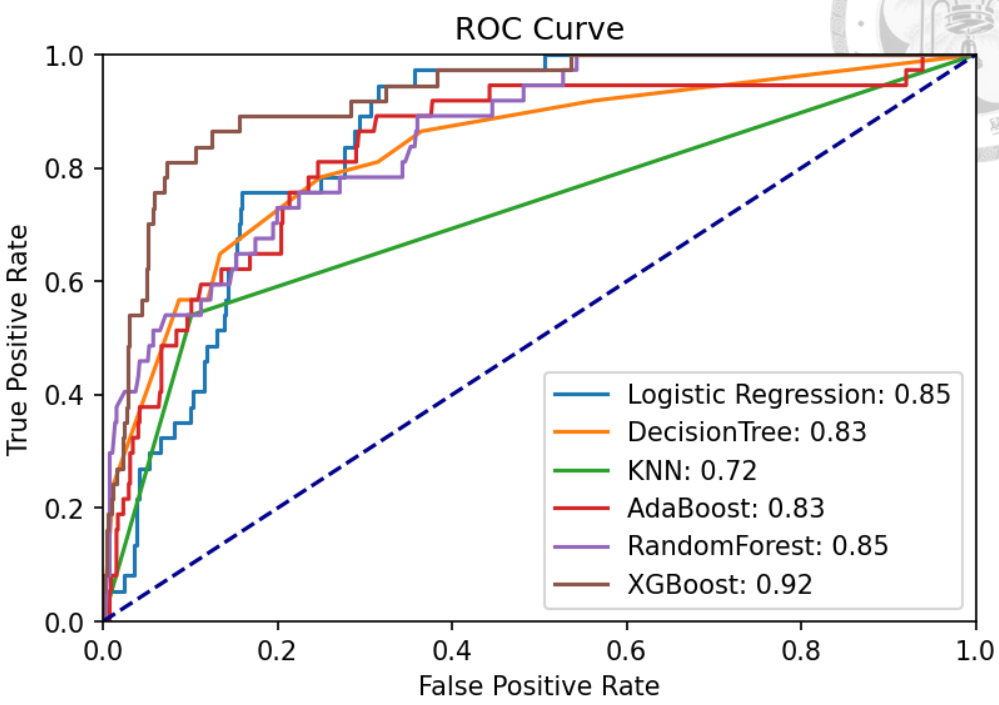

Figure 3.6 The ROC Curve of the proposed models with no demographic data for a manic episode in ASD on the testing set with the difference from baseline features

3.3.3 Prediction model for a depressive episode in ASD

In another experiment, considering the clinical condition and correlation matrix of features, we selected 15 digital biomarker features for model training. Additionally, we removed 4 features: light sleep duration, REM sleep duration, awake sleep duration, and minimum heart rate. The performance of the proposed models for manic episodes in autism was evaluated using a 5-fold cross-validation with the BDI score questionnaire, and the results are presented in Table 3.8.

Table 3.8 Performance of the proposed models for a depressive episode in ASD on 5-fold cross-validation with the difference from baseline features

Model	Accuracy	AUROC	Sensitivity	Specificity	Precision	F1 score

doi:10.6342/NTU202302181

Logistic Regression	0.78	0.75	0.99	0.34	0.75	0.86
Decision Tree	0.87	0.94	0.96	0.70	0.87	0.91
KNN	0.82	0.80	0.86	0.74	0.87	0.86
Random Forest	0.86	0.95	0.91	0.77	0.89	0.90
AdaBoost	0.72	0.87	0.99	0.19	0.71	0.83
XGBoost	0.88	0.96	0.99	0.66	0.85	0.91

Then, the performance of the proposed models for manic episodes in autism was evaluated on the testing set using the YMRS score questionnaire, and the results are presented in Table 3.9.

Table 3.9 Performance of the proposed models for a depressive episode in ASD on the testing set with the difference from baseline features

Model	Accuracy	AUROC	Sensitivity	Specificity	Precision	F1 score
Logistic Regression	0.84	0.76	0.99	0.38	0.82	0.90
Decision Tree	0.87	0.93	0.92	0.72	0.90	0.91
KNN	0.80	0.76	0.84	0.69	0.88	0.86
Random Forest	0.85	0.92	0.90	0.70	0.90	0.90
AdaBoost	0.77	0.85	0.99	0.15	0.77	0.87
XGBoost	0.89	0.94	0.99	0.63	0.88	0.93

The purpose of the ROC Curve is to provide an overview of the pretrained model's classifier situation. As shown in Figure 3.7, Logistic Regression, AdaBoost, and XGBoost exhibit better performance compared to the other three models.

doi:10.6342/NTU202302181

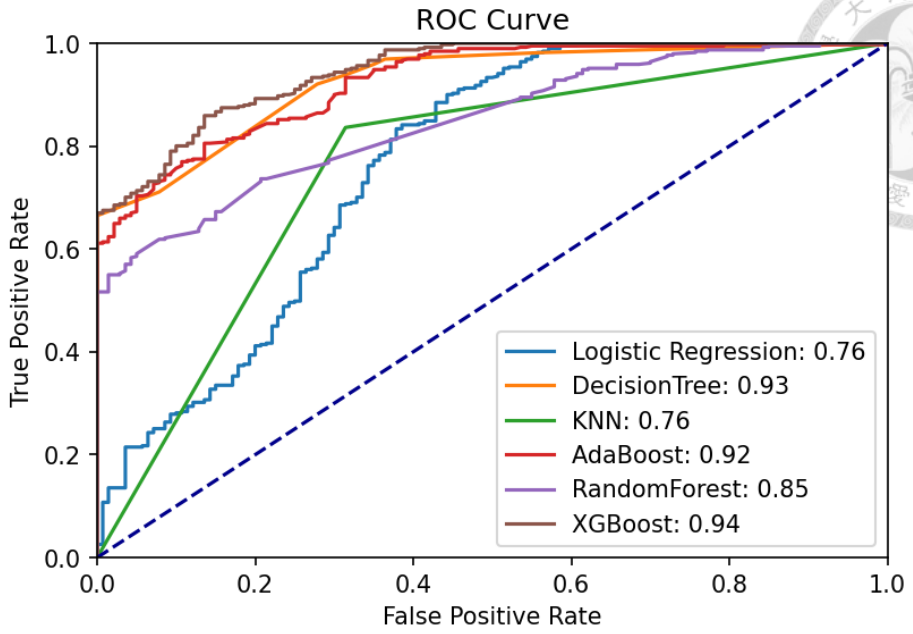

Figure 3.7 The ROC Curve of the proposed models for a depressive episode in ASD on the testing set with the difference from baseline features

3.3.4 Prediction model for a depressive episode in ASD with no demographic data

In another experiment, based on the clinical condition and correlation matrix of features, we used 15 digital biomarker features for model training and removed 4 features: light sleep duration, REM sleep duration, awake sleep duration, and minimum heart rate. The performance of the proposed models for manic episodes in autism was evaluated using a 5-fold cross-validation with the BDI score questionnaire, and the results are shown in Table 3.10.

doi:10.6342/NTU202302181

Table 3.10 Performance of the proposed models with no demographic data for a depressive episode in ASD on 5-fold cross-validation with the difference from baseline features

Model	Accuracy	AUROC	Sensitivity	Specificity	Precision	F1 score
Logistic Regression	0.78	0.74	0.99	0.34	0.75	0.86
Decision Tree	0.71	0.75	0.92	0.30	0.73	0.81
KNN	0.86	0.87	0.74	0.86	0.80	0.82
AdaBoost	0.75	0.81	0.86	0.53	0.79	0.82
Random Forest	0.70	0.73	0.99	0.11	0.69	0.82
XGBoost	0.78	0.91	0.98	0.37	0.76	0.85

Then, the performance of the proposed models for manic episodes in autism on the testing set using the YMRS score questionnaire is shown in Table 3.11.

Table 3.11 Performance of the proposed models with no demographic data for a depressive episode in ASD on the testing set with the difference from baseline features

Model	Accuracy	AUROC	Sensitivity	Specificity	Precision	F1 score
Logistic Regression	0.84	0.77	0.99	0.38	0.82	0.90
Decision Tree	0.78	0.75	0.95	0.31	0.79	0.86
KNN	0.80	0.76	0.84	0.69	0.88	0.86
AdaBoost	0.75	0.79	0.83	0.51	0.83	0.83
Random Forest	0.76	0.76	0.99	0.10	0.75	0.86
XGBoost	0.84	0.91	0.98	0.43	0.83	0.90

The purpose of the ROC Curve is to provide an overview of the pretrained model's classifier situation. As shown in Figure 3.8, Logistic Regression, AdaBoost, and XGBoost

doi:10.6342/NTU202302181

exhibit better performance compared to the other three models.

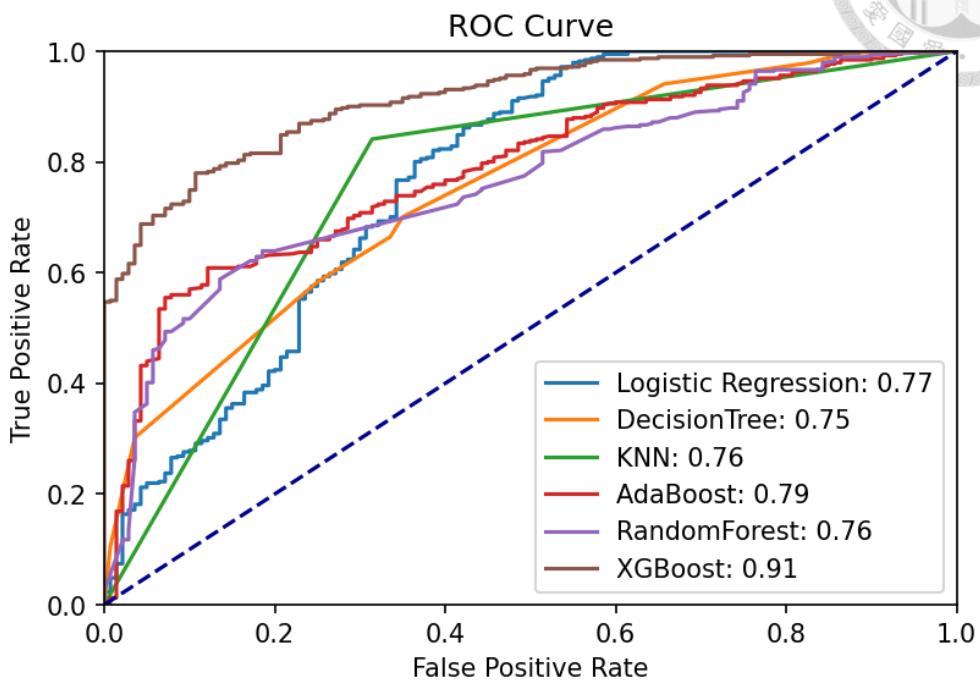

Figure 3.8 The ROC Curve of the proposed models with no demographic data for a depressive episode in ASD on the testing set with the difference from baseline features

3.4 Explanation of Prediction Model

3.4.1 Explanation for a manic episode in ASD

In Figure 3.9, the summary plot is designed to display an information-dense summary of how the top features in a dataset impact the model's output. The plot consists of scatter points with different colors. The gathering of red points indicates higher values for the corresponding feature, which suggests a higher likelihood of the predicted event occurring, and vice versa.

doi:10.6342/NTU202302181

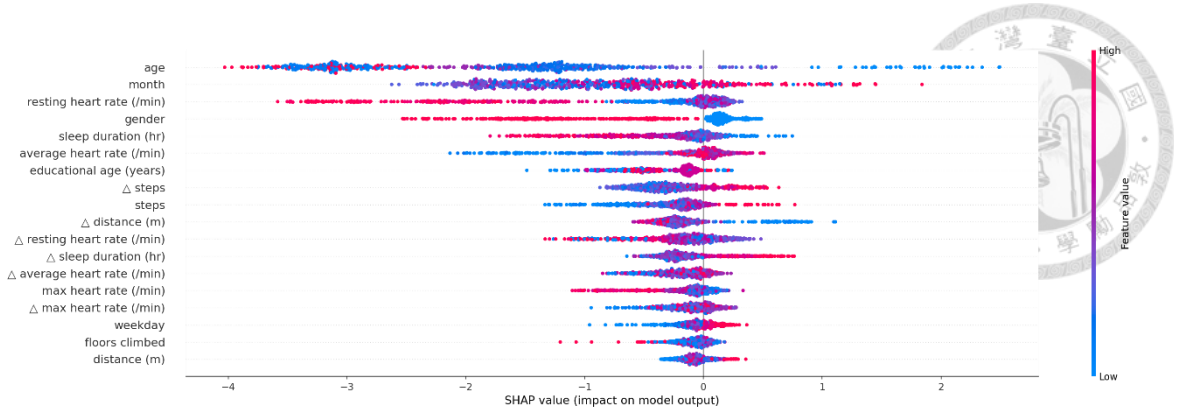

Figure 3.9 Feature importance of XGBoost with the summary plot

By analyzing the patterns and distribution of the scatter points, we can identify which features have a more significant impact on the model's output, enabling us to gain a deeper understanding of the underlying relationships between the features and the predicted outcomes.

In Figure 3.10, the features listed at the top are deemed important, indicating their significant impact on the model. The higher the features are placed on the list, the greater their influence on the model's predictions. In this figure, resting heart rate, steps, month, age, and gender are the top 5 influential features in the model. This information is invaluable for developing targeted interventions and personalized approaches to support individuals with autism experiencing manic episodes, ultimately leading to improved outcomes and better management of their condition.

doi:10.6342/NTU202302181

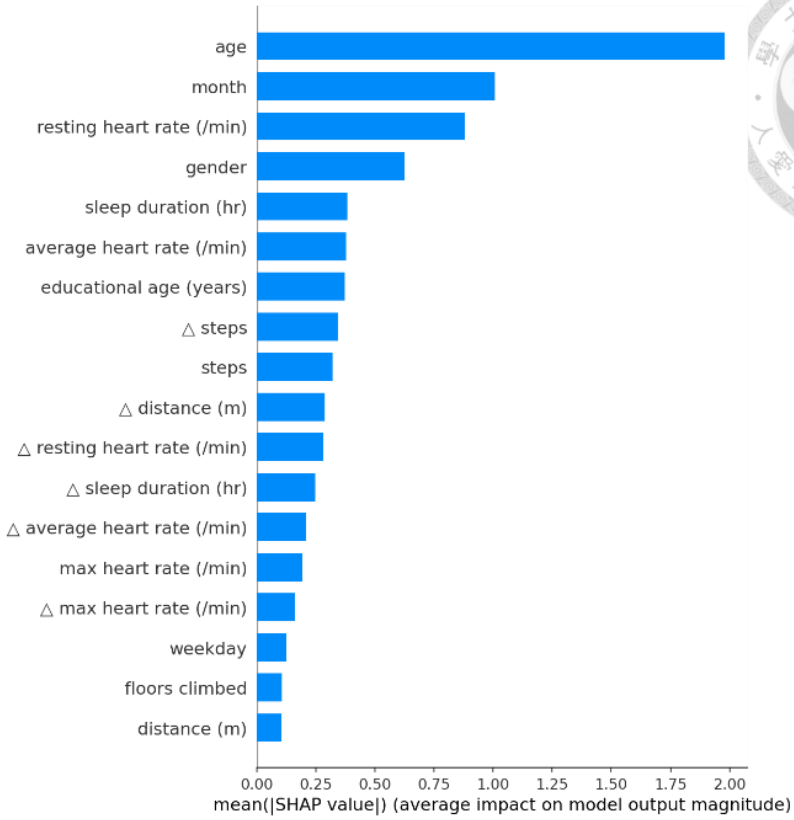

Figure 3.10 Feature importance of XGBoost with the summary plot

The following 5 charts (Figures 3.11~3.15) illustrate how each single feature affects the model. The X-axis in each figure represents the number of beats per minute at resting heart rate, while the Y-axis indicates the magnitude of the impact on the model. In the charts, a positive value is represented by the red region, indicating a tendency to predict the presence of mood symptoms of depression. On the other hand, a negative value is depicted by the blue region, suggesting the opposite, i.e., a tendency to predict the absence of mood symptoms of depression.

Figure 3.11 displays the resting heart rate in beats per minute (/min). The data points falling between 45 to 60 are indicative of predicting the likelihood of mood symptoms related to depression.

doi:10.6342/NTU202302181

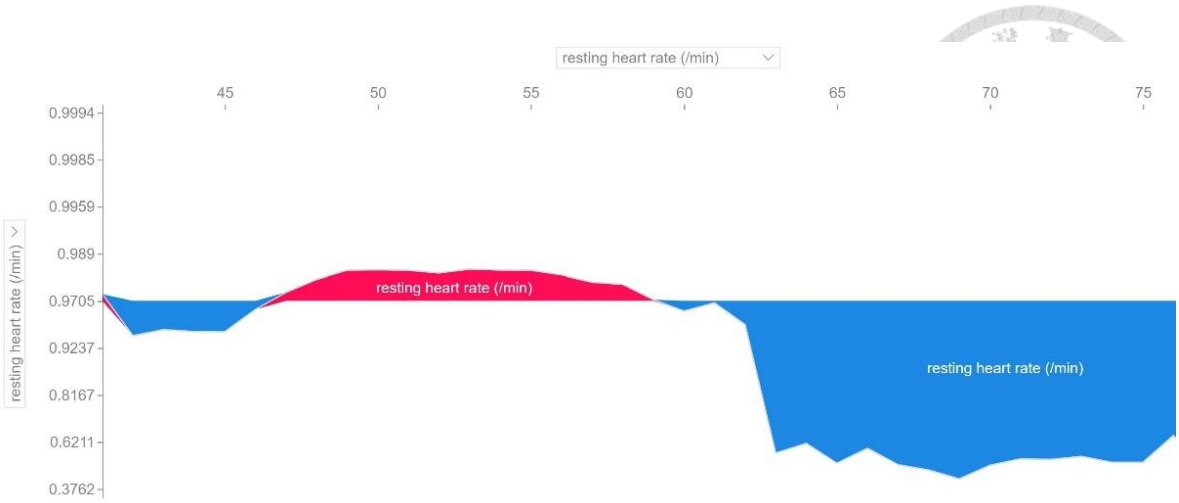

Figure 3.11 Impact of resting heart rate on the model with the force plot

In Figure 3.12 and Figure 3.13, the number of steps in activity and total sleep duration in hours during sleep state are depicted, respectively. The figures suggest that taking more than 8,000 steps and sleeping fewer than 7 hours might also be associated with the likelihood of experiencing mood symptoms related to mania.

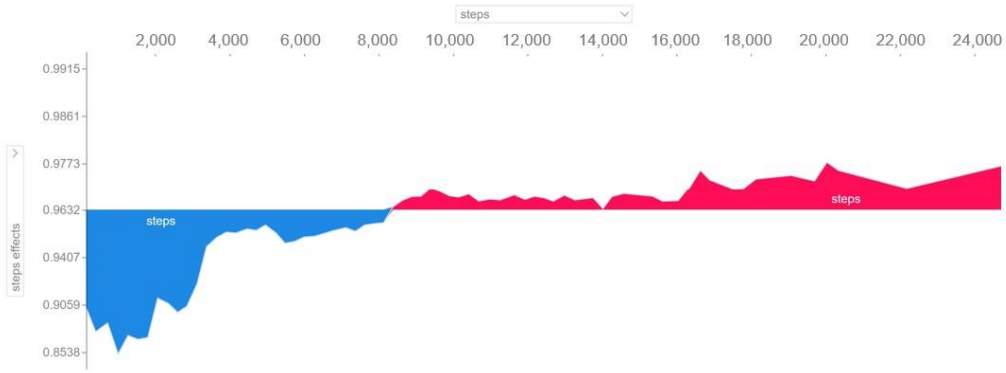

Figure 3.12 Impact of steps on the model with the force plot

doi:10.6342/NTU202302181

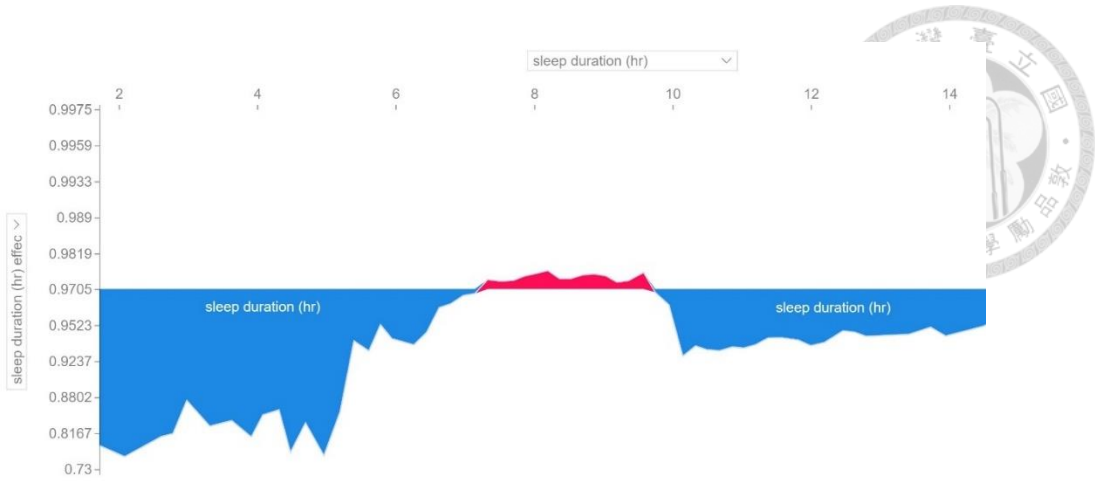

Figure 3.13 Impact of total sleep duration on the model with the force plot

In Figure 3.14 and Figure 3.15, the average heart rate (/min) and the personalized total sleep duration in hours during the sleep state are depicted, respectively. The figures suggest that an average heart rate between 75 and 85 (/min) and personally sleeping less than 2 hours might also be associated with the likelihood of experiencing mood symptoms related to depression.

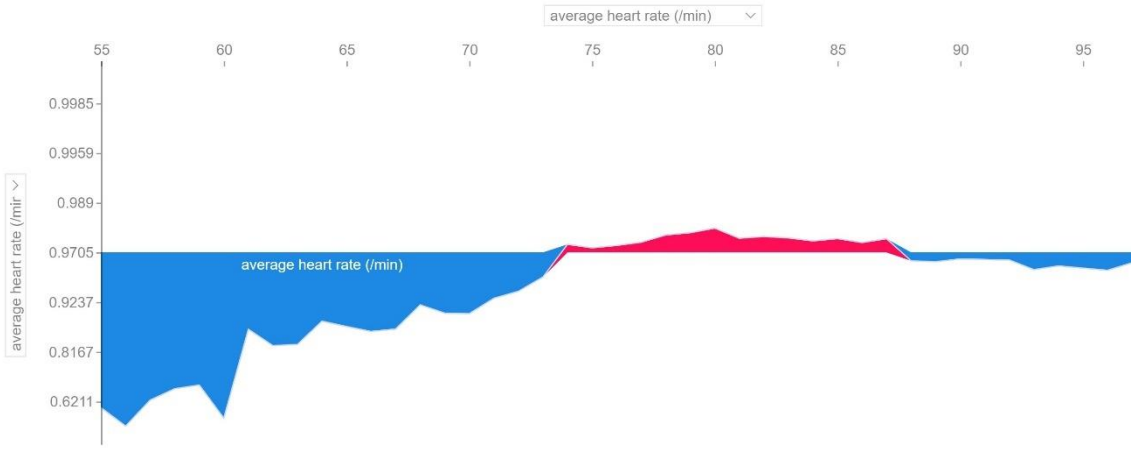

Figure 3.14 Impact of average heart rate on the model with the force plot

doi:10.6342/NTU202302181

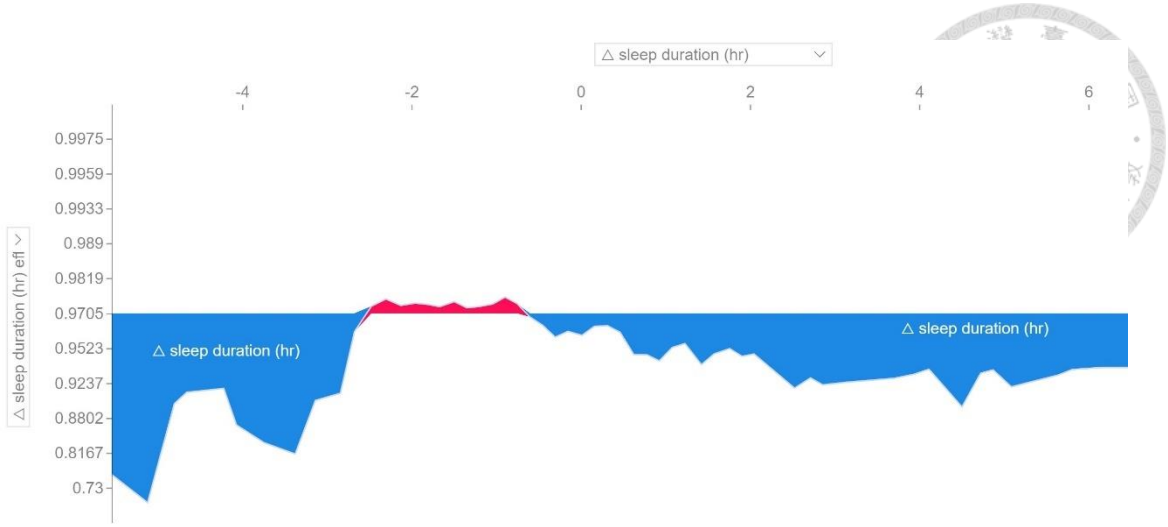

Figure 3.15 Impact of difference sleep duration on the model with the force plot

3.4.2 Explanation for a manic episode in ASD with no demographic data

In Figure 3.16, the summary plot is specifically designed to present an information-dense summary of how the top features in a dataset impact the model's output, specifically when no demographic data is included. The plot showcases points scattered with different colors. The gathering of red points indicates higher feature values, suggesting a higher likelihood of the event being predicted, and vice versa for the blue points. This visualization provides valuable insights into the influence of key features on the model's predictions and helps in understanding the significance of each feature in relation to the occurrence of the predicted event.

Figure 3.16 Feature importance of XGBoost with the summary plot

doi:10.6342/NTU202302181

In Figure 3.17, the features at the top are considered important, indicating their higher impact on the model. Resting heart rate, steps, month, distance, and average heart rate are the top 5 influential features in this figure. These findings suggest that these specific features play a crucial role in the model's predictions, making them significant contributors to understanding and predicting mood symptoms in the context of our study.

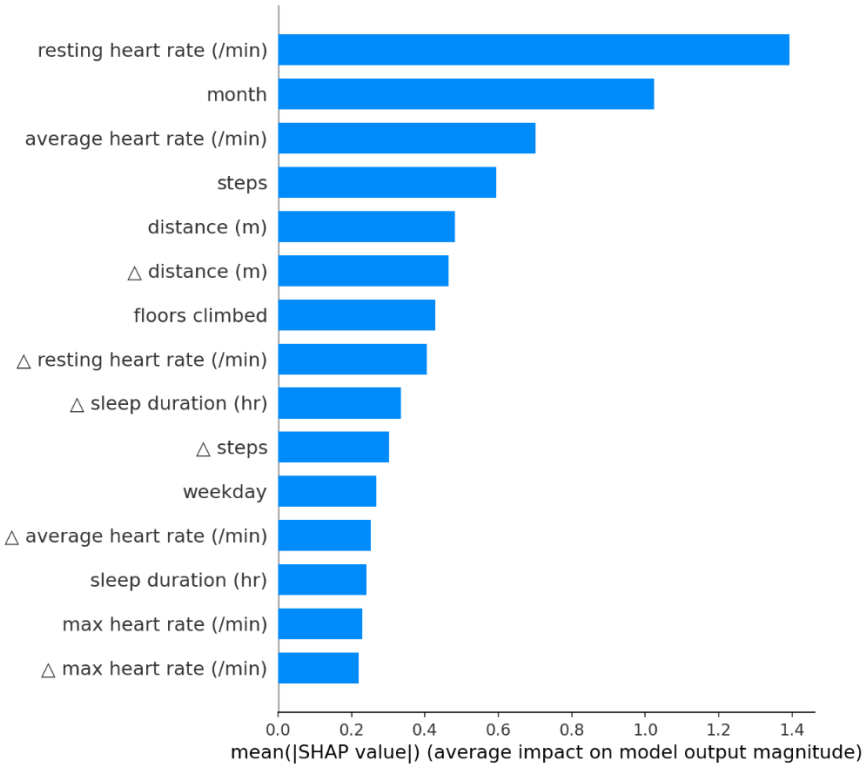

Figure 3.17 Feature importance of XGBoost with the summary plot

In Figure 3.18 and Figure 3.19, the resting heart rate and average heart rate are displayed, respectively. The figures suggest that if the resting heart rate falls between 50 to 60, it may lead to mood symptoms in mania. Additionally, if the average heart rate is over 75, it may also be associated with mood symptoms in mania. These findings indicate that heart rate measurements can serve as important indicators for predicting manic episodes in individuals with autism.

doi:10.6342/NTU202302181

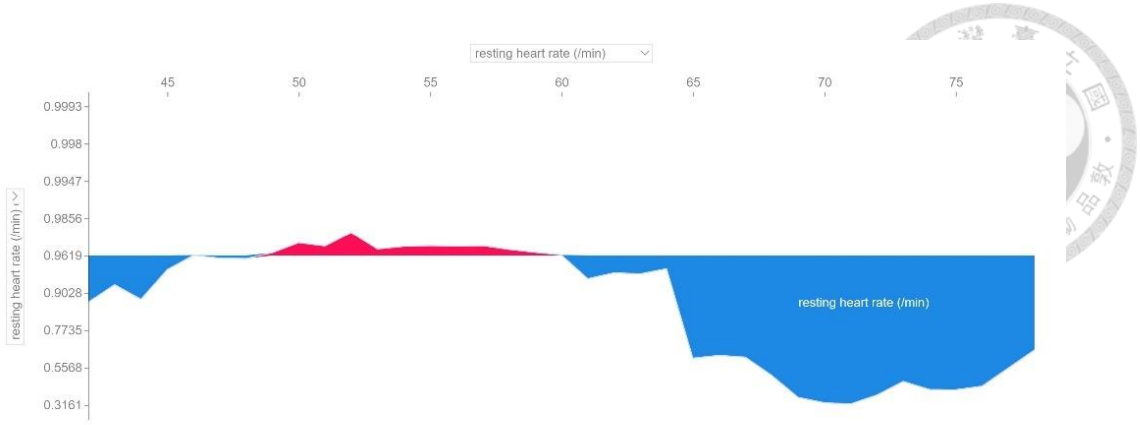

Figure 3.18 Impact of resting heart rate on the model with the force plot

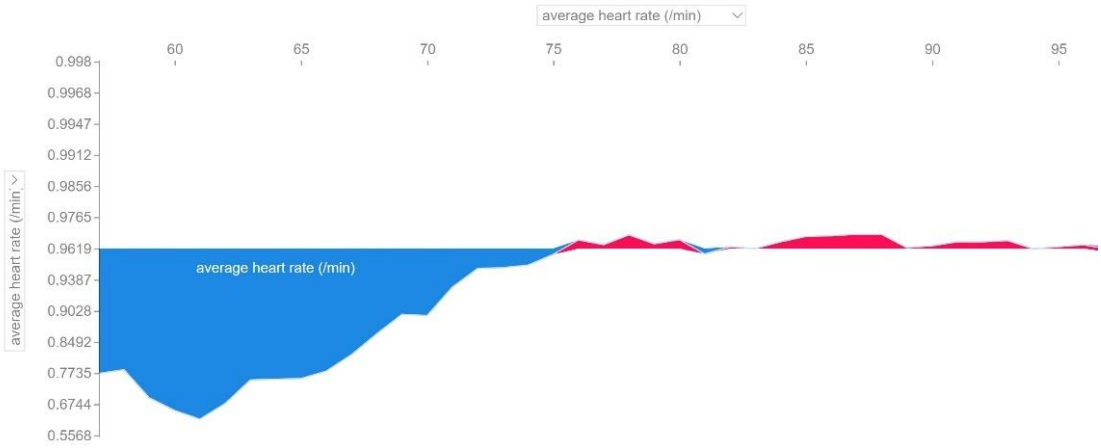

Figure 3.19 Impact of average heart rate on the model with the force plot

In Figure 3.20 and Figure 3.21, the activity data is depicted. The figures suggest that if the number of steps is over 14,000 and the distance covered exceeds 12,000 meters, it may lead to mood symptoms in mania. It implies that both excessive steps and significant distance covered may be associated with the onset of manic symptoms.

doi:10.6342/NTU202302181

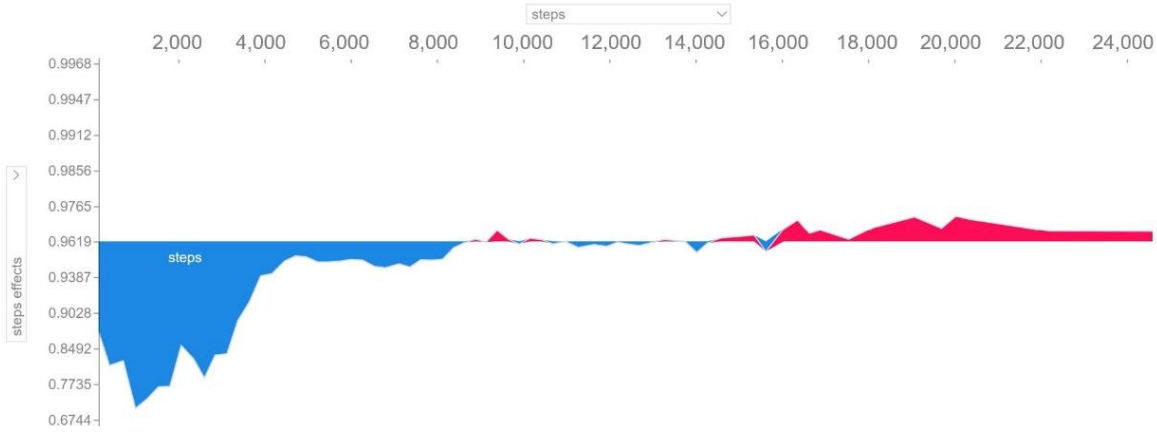

Figure 3.20 Impact of steps on the model with the force plot

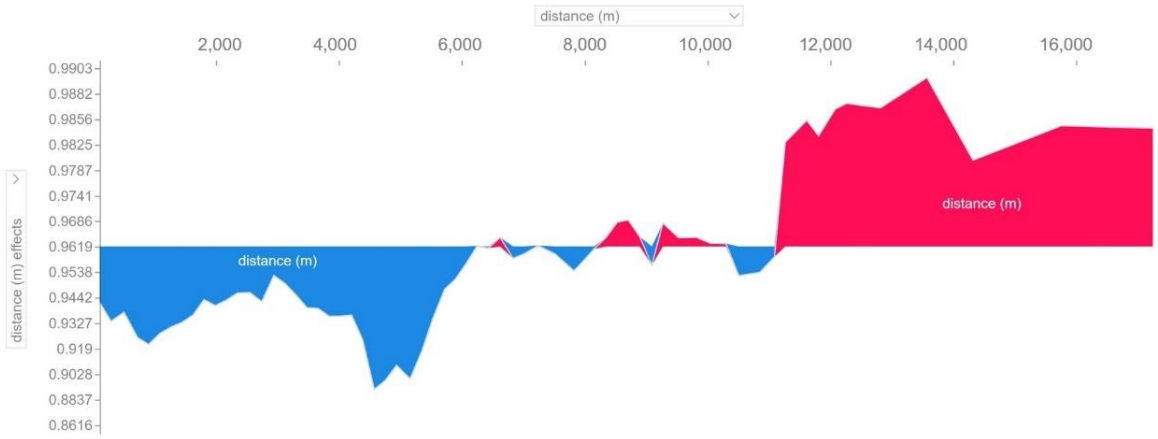

Figure 3.21 Impact of distance on the model with the force plot

3.4.3 Explanation for a depressive episode in ASD

Figure 3.22 presents a summary plot designed to provide a concise and information-rich overview of how the top features in the dataset influence the model's output.

doi:10.6342/NTU202302181

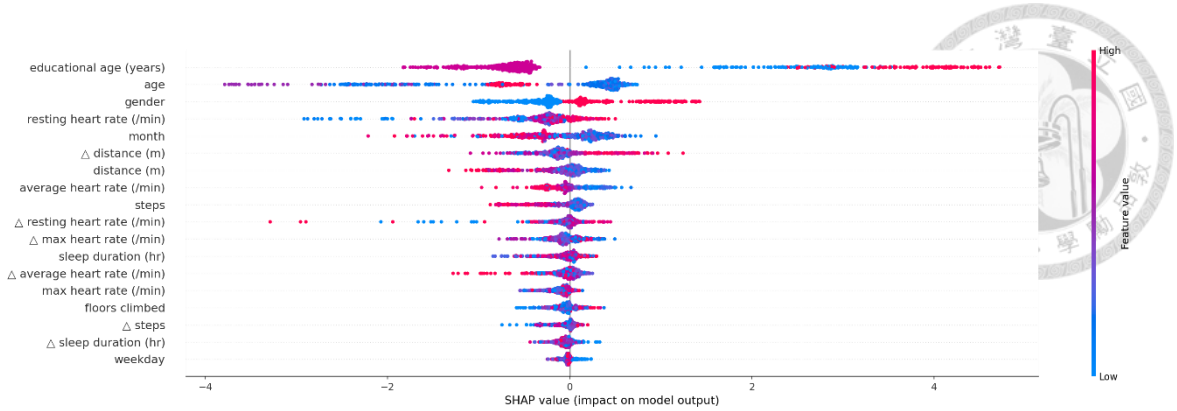

Figure 3.22 Feature importance of XGBoost with the summary plot

In Figure 3.23, the features located at the top are considered important, indicating their significant impact on the model's output. These findings suggest that variations in resting heart rate and the specific month have a considerable effect on the model's ability to predict mood symptoms in the given context of the study.

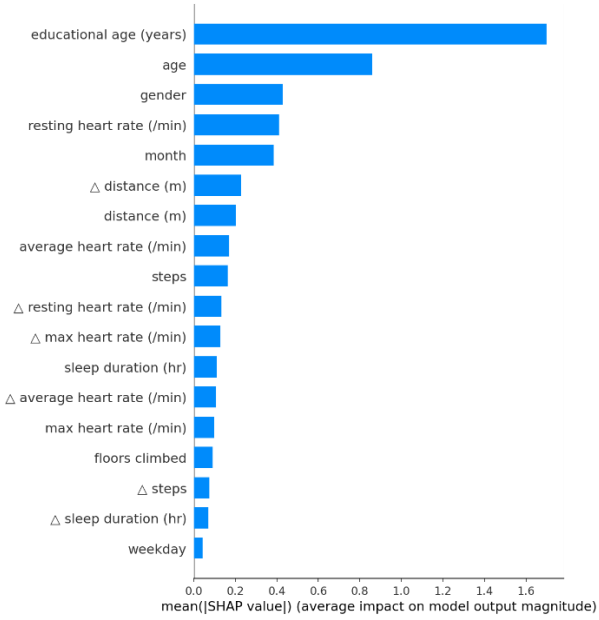

Figure 3.23 Feature importance of XGBoost with the summary plot

Figure 3.24 and Figure 3.25 display the distance in meters during activity and the total sleep duration in hours during sleep state, respectively. The figures suggest that moving less than 6,000 meters and sleeping less than 9 hours may also be associated with

doi:10.6342/NTU202302181

mood symptoms in depression.

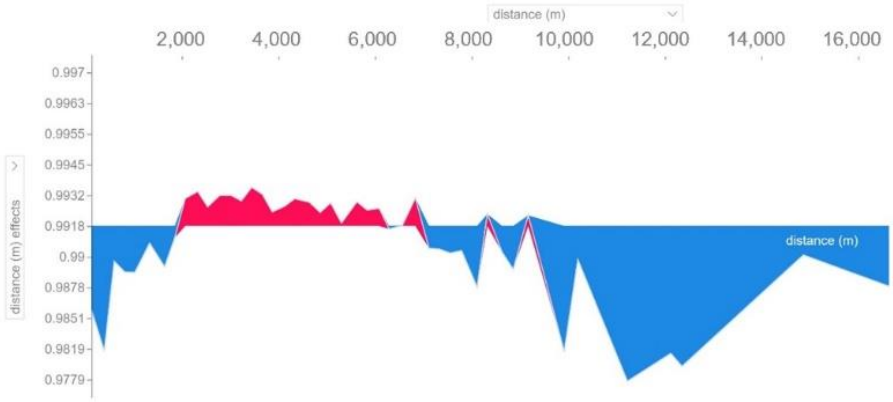

Figure 3.24 Impact of total distance on the model with the force plot

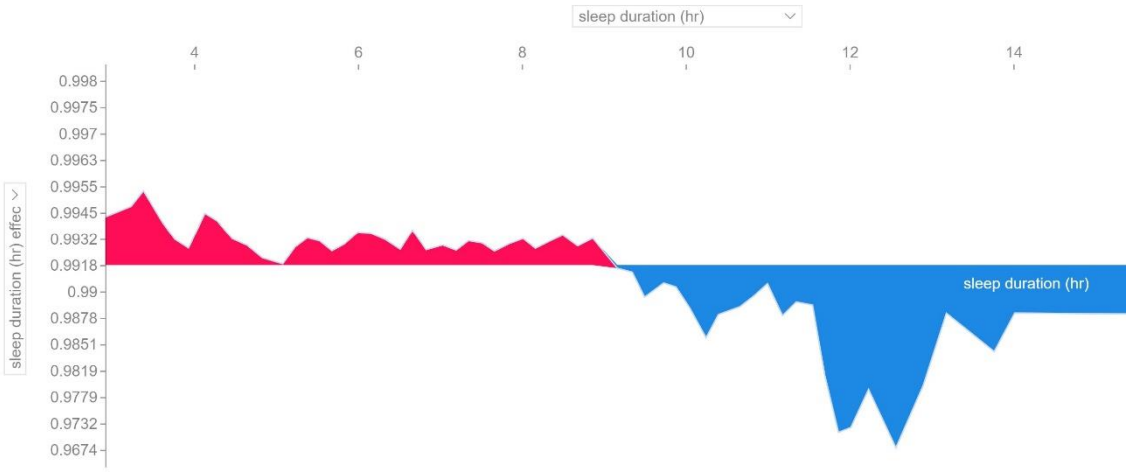

Figure 3.25 Impact of total sleep duration on the model with the force plot

Figure 3.26 and Figure 3.27 display the heart rate in beats per minute, respectively. The figures suggest that an average heart rate (/min) over 80 and a resting heart rate lower than 55 may be associated with mood symptoms in depression.

doi:10.6342/NTU202302181

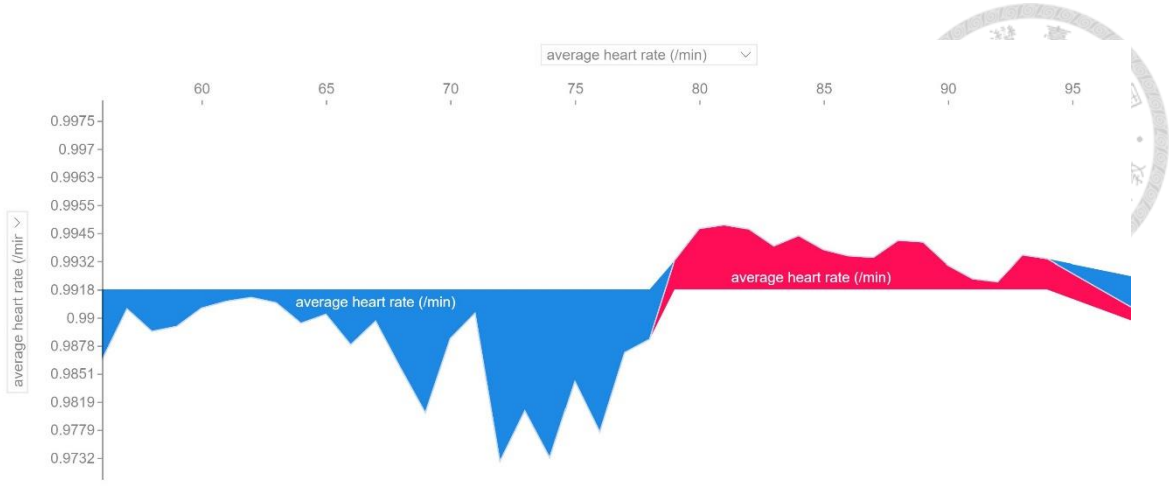

Figure 3.26 Impact of resting heart rate on the model with the force plot

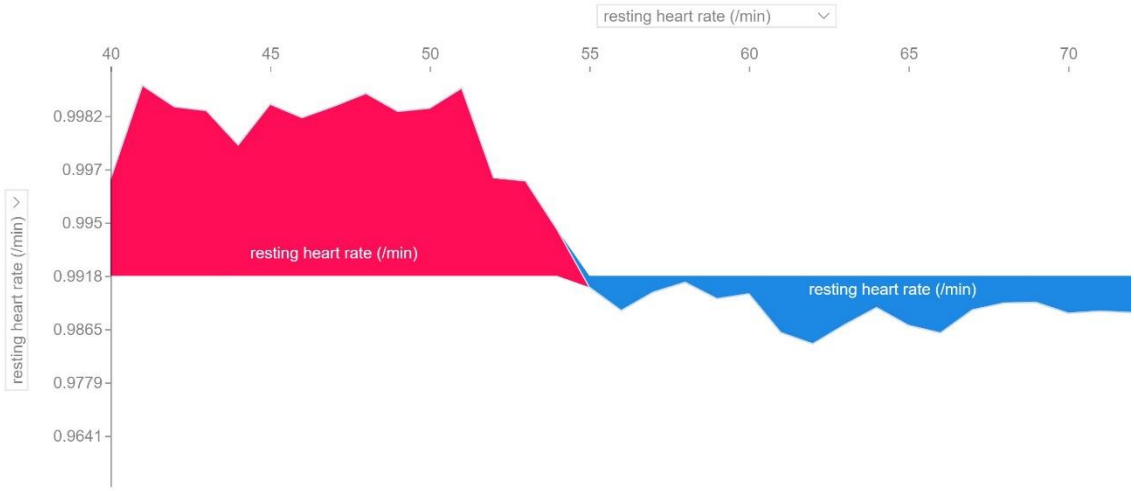

Figure 3.27 Impact of resting heart rate on the model with the force plot

3.4.4 Explanation for a depressive episode in ASD no demographic data

In Figure 3.28, the beeswarm plot is created to present a concise and information-rich summary of how the top features in the dataset influence the model's output.

doi:10.6342/NTU202302181

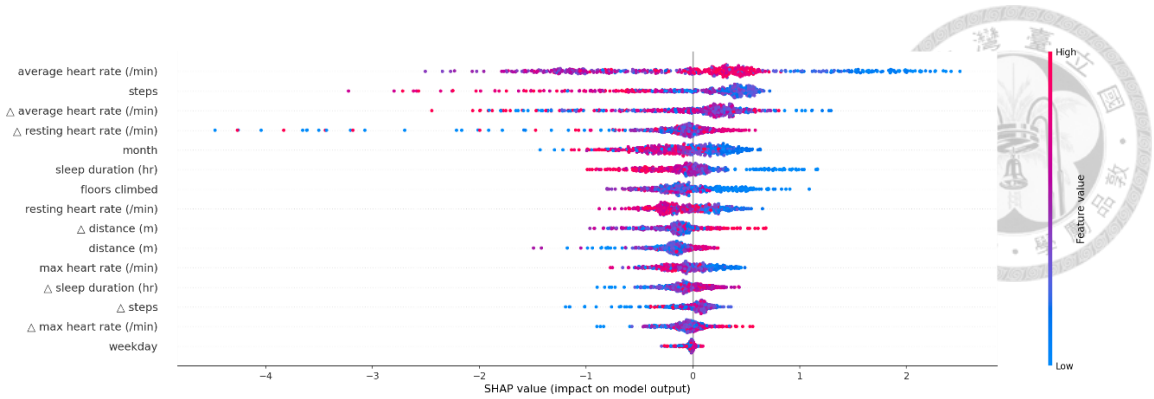

Figure 3.28 Feature importance of XGBoost with the beeswarm plot

In Figure 3.29, the features positioned at the top indicate their significance in influencing the model's output. Among these features, resting heart rate and month stand out as particularly important contributors to the model's predictions. This finding suggests that variations in resting heart rate and the specific month play crucial roles in predicting mood symptoms in the context of our study.

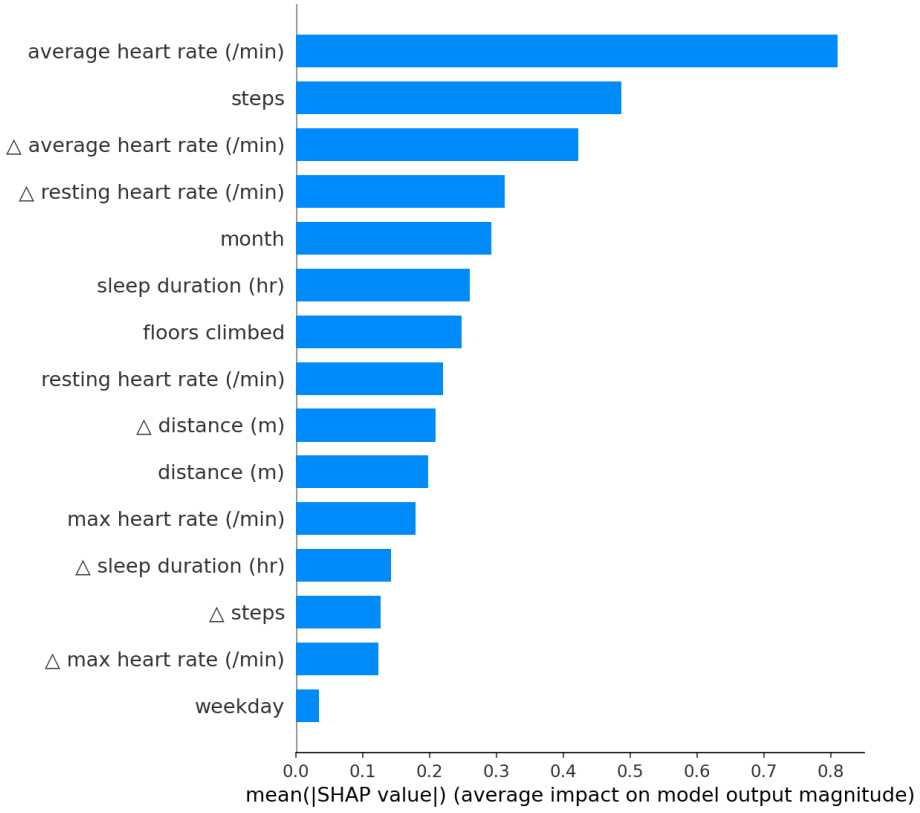

Figure 3.29 Feature importance of XGBoost with the summary plot

doi:10.6342/NTU202302181

In Figure 3.30 and Figure 3.31, the distance in meters during activity and the total sleep duration in hours during the sleep state are depicted, respectively. The figures indicate that moving less than 8,000 steps and sleeping less than 8 hours might also be associated with mood symptoms in depressions.

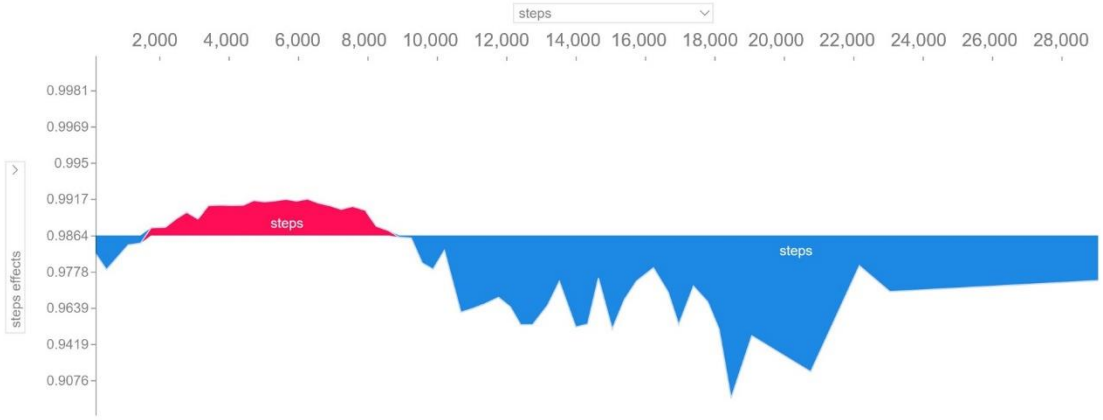

Figure 3.30 Impact of steps on the model with the force plot

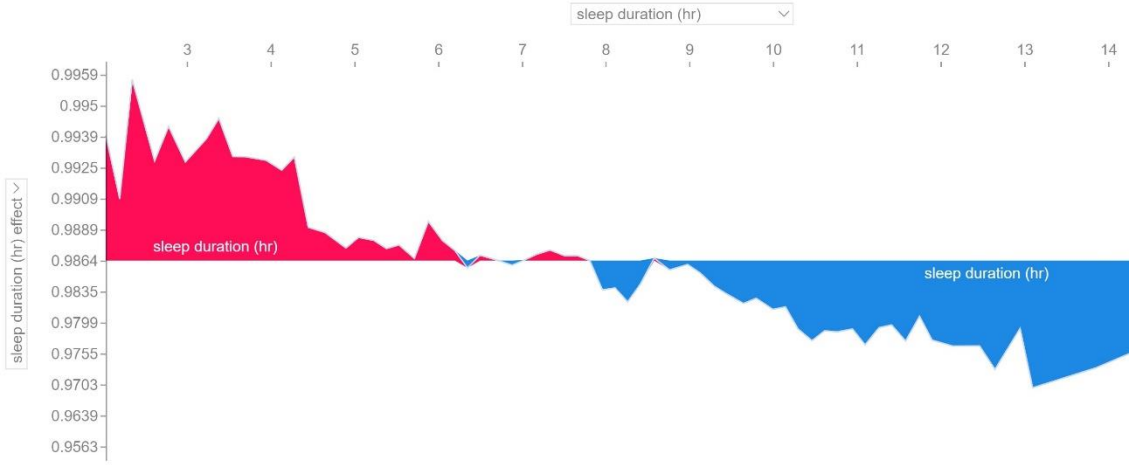

Figure 3.31 Impact of total sleep duration on the model with the force plot

In Figure 3.32 and Figure 3.33, the heart rate in beats per minute is displayed, respectively. The figures suggest that an average heart rate (/min) between 55 to 70, and a resting heart rate lower than 55 may be associated with mood symptoms in depression. These findings highlight the potential impact of heart rate measurements on predicting

doi:10.6342/NTU202302181

depressive symptoms in individuals with autism.

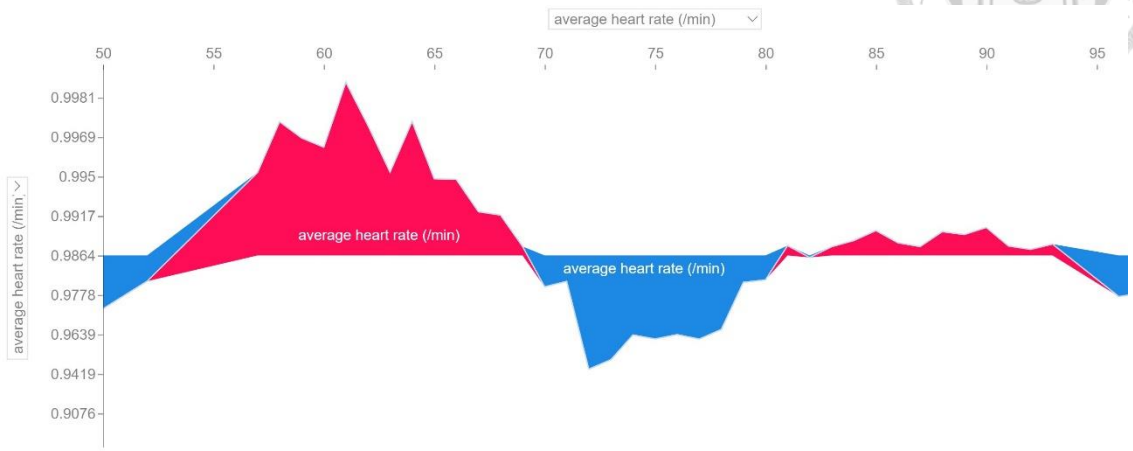

Figure 3.32 Impact of average heart rate on the model with the force plot

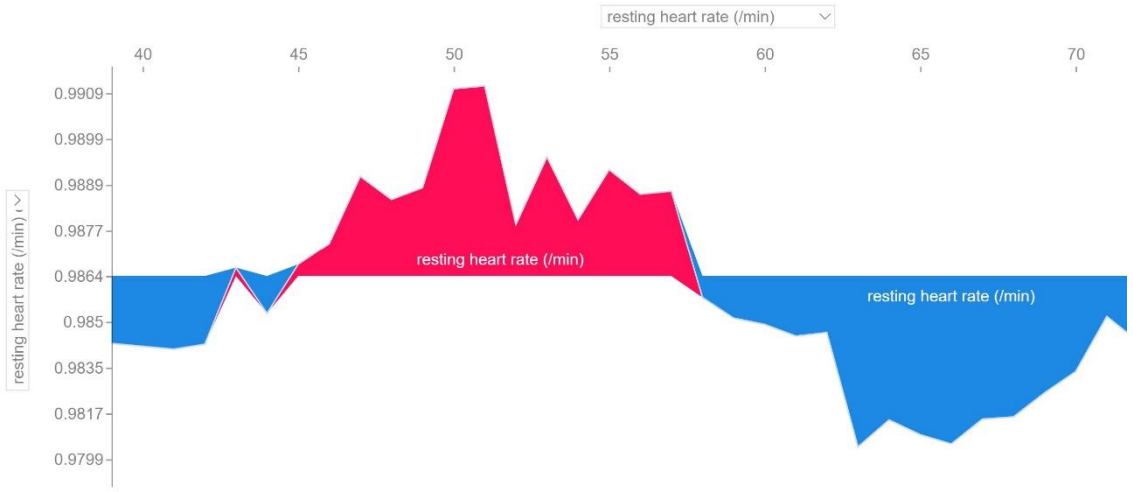

Figure 3.33 Impact of resting heart rate on the model with the force plot

doi:10.6342/NTU202302181

Chapter 4 Discussion

4.1 Major Findings

In this study, we successfully utilized digital biomarkers collected from wearable devices to predict mood symptoms several days later. The results revealed that resting heart rate [24, 25], average heart rate [26], and sleep duration [27-28] were the most critical features, regardless of the questionnaire used. Notably, we observed that a lower resting heart rate (below 55) was associated with a higher likelihood of causing depressive symptoms.

Moreover, our findings indicated that lower resting heart rate, age, gender, and the month were significant predictors of both depression and mania episodes. Demographic data, such as age and gender, played an essential role in the results for all mood symptoms [6].

Furthermore, we found compelling evidence suggesting that a resting heart rate below 55 might be a contributing factor in predicting depression. Additionally, activity levels had a notable impact, indicating that walking fewer than 8,000 steps might be associated with a soothing effect on depressive symptoms [9]. As for sleep, a total duration of fewer than 7 hours might also contribute to predicting depressive symptoms [29]. With these valuable insights, we can predict depressive symptoms by measuring the patient's digital biomarkers and offer relevant health advice to prevent depressive episodes effectively.

In terms of model performance, our depression model outperformed the mania model, despite the imbalanced data. The depression model achieved an impressive accuracy of 89%, an AUC of 0.94, and an f1 score of 0.93 using the XGBoost model. In contrast, the manic model achieved an accuracy of 91%, an AUC of 0.96, and an f1 score

doi:10.6342/NTU202302181

of 0.52 using the XGBoost model.

These results highlight the potential of using digital biomarkers for predicting mood symptoms and emphasize the importance of considering resting heart rate, activity levels, and sleep duration in understanding and managing depression in individuals with autism. By employing machine learning models and leveraging wearable technology, we pave the way for early intervention and personalized approaches to support mental health and well-being in this population.

4.2 Limitations

In the beginning, due to the differences among patients, not all the models are suitable for every individual. Additionally, the amount of data available cannot fully represent all autism spectrum disorder groups.

One of the research methods involved asking our patients to fill out the questionnaire once a week and then backfill for 7 days. However, this approach may lead to recall bias since the data might not accurately reflect their immediate experiences. To improve data reliability, it would be beneficial to have the patients fill out the questionnaire immediately upon awareness of manic and depressive symptoms related to autism. Furthermore, there is a concern that patients completing the questionnaire themselves may not fully convey their real situation for some questions. To address this issue, doctors should be involved to ensure a more accurate understanding of the patients' condition.

Regarding the sleep duration biomarkers, some patients may use medications that could impact the precision of the experimental data. However, restricting patients from using medication for the sake of research would not be ethical.

Navigating these challenges is essential to enhance the validity and applicability of

doi:10.6342/NTU202302181

our research findings. By acknowledging these limitations and working to minimize potential biases, we can improve the reliability of our results and gain deeper insights into mood symptoms in autism, ultimately contributing to more effective interventions and treatments for individuals with this condition.

4.3 Future Work

In the future, there are several strategies we can adopt to enhance our research. Firstly, we aim to collect more detailed patients' biomarker data and questionnaire responses, including heart rate values (in bps) and accelerometer data recorded while they wear the watch. By incorporating this additional data, we can gain deeper insights into the patients' physiological responses and activity levels, which could lead to more robust and accurate predictions of mood symptoms.

Secondly, we plan to integrate weather data [30] into our experiments. Gathering data from the nearest weather monitoring site to the patients' homes and study sites can help us identify potential influences of weather on mood symptoms. This could potentially uncover valuable features that affect mood variations, further enhancing the predictive power of our models.

By combining these strategies and obtaining more valuable features, our models will become more powerful and capable of providing meaningful insights into mood symptom prediction in autism. These improvements will enable us to better understand and support individuals with autism experiencing mood fluctuations, contributing to more effective and personalized interventions for their well-being.

doi:10.6342/NTU202302181

Chapter 5 Conclusion

In this study, digital biomarkers and questionnaire were collected from wearable devices and the patients fill out once a week. And then, we construct a machine model which could predict manic and depressive symptoms. After data preprocessing, 1,696 and 1,191 data points were used to build the machine learning model. Although the data was imbalanced, we achieved 89% accuracy, 0.94 AUROC, and 0.93 F1 score for depressive symptoms prediction on testing data. On the other model, we achieved 91% accuracy, 0.96 AUROC, and 0.52 F1 score for manic symptoms prediction on testing data. Moreover, we utilized the interpretable model SHAP to see how the model to make the prediction. We found that relatively resting heart rate and month were associated with and may predict depressive symptoms. Based on these findings, we may provide clinical assessment earlier to prevent a risk.

doi:10.6342/NTU202302181

References

[1] Bolton, M.J., et al., *Initial evidence for increased weather salience in autism spectrum conditions.* Weather, Climate, and Society, 2020. **12**(2): p. 293-307.

[2] Hodges, H., C. Fealko, and N. Soares, *Autism spectrum disorder: definition, epidemiology, causes, and clinical evaluation.* Translational pediatrics, 2020. **9**(Suppl 1): p. S55.

[3] Miller, L.E., et al., *Characteristics of toddlers with early versus later diagnosis of autism spectrum disorder.* Autism, 2021. **25**(2): p. 416-428.

[4] Watson, L., P. Hanna, and C.J. Jones, *A systematic review of the experience of being a sibling of a child with an autism spectrum disorder.* Clinical Child Psychology and Psychiatry, 2021. **26**(3): p. 734-749.

[5] Parner, E.T., et al., *Parental age and autism spectrum disorders.* Annals of epidemiology, 2012. **22**(3): p. 143-150.

[6] York, A., et al., *Fragile-X syndrome, Down's syndrome and autism: awareness and knowledge amongst special educators.* Journal of Intellectual Disability Research, 1999. **43**(4): p. 314-324.

[7] Werling, D.M. and D.H. Geschwind, *Sex differences in autism spectrum disorders.* Current opinion in neurology, 2013. **26**(2): p. 146.

[8] Hudson, C.C., L. Hall, and K.L. Harkness, *Prevalence of depressive disorders in individuals with autism spectrum disorder: A meta-analysis.* Journal of Abnormal Child Psychology, 2019. **47**: p. 165-175.

[9] Spoorthy, M.S., S. Chakrabarti, and S. Grover, *Comorbidity of bipolar and anxiety disorders: An overview of trends in research.* World journal of psychiatry, 2019. **9**(1): p. 7.

doi:10.6342/NTU202302181

[10] Morie, K.P., et al., *Mood disorders in high-functioning autism: The importance of alexithymia and emotional regulation.* Journal of Autism and Developmental Disorders, 2019. **49**: p. 2935-2945.

[11] Black, M.H., et al., *The use of wearable technology to measure and support abilities, disabilities and functional skills in autistic youth: a scoping review.* Scandinavian Journal of Child and Adolescent Psychiatry and Psychology, 2020. **8**(1): p. 48-69.

[12] Koumpouros, Y. and T. Kafazis, *Wearables and mobile technologies in Autism Spectrum Disorder interventions: A systematic literature review.* Research in Autism Spectrum Disorders, 2019. **66**: p. 101405.

[13] Schmidt, M., et al., *A process-model for minimizing adverse effects when using head mounted display-based virtual reality for individuals with autism.* Frontiers in Virtual Reality, 2021. **2**: p. 611740.

[14] Li, H., et al., *Associations of emotional/behavioral problems with accelerometer-measured sedentary behavior, physical activity and step counts in children with autism spectrum disorder.* Frontiers in Public Health, 2022. **10**: p. 981128.

[15] Priya, A., S. Garg, and N.P. Tigga, *Predicting anxiety, depression and stress in modern life using machine learning algorithms.* Procedia Computer Science, 2020. **167**: p. 1258-1267.

[16] Gaussier, E., et al. *Improving backfilling by using machine learning to predict running times*. in *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. 2015.

[17] Li, D.-C., C.-W. Liu, and S.C. Hu, *A learning method for the class imbalance problem with medical data sets.* Computers in biology and medicine, 2010. **40**(5): p. 509-518.

doi:10.6342/NTU202302181

[18] Sharma, H. and S. Kumar, *A survey on decision tree algorithms of classification in data mining.* International Journal of Science and Research (IJSR), 2016. **5**(4): p. 2094-2097.

[19] Altman, N.S., *An introduction to kernel and nearest-neighbor nonparametric regression.* The American Statistician, 1992. **46**(3): p. 175-185.

[20] Breiman, L., *Random forests.* Machine learning, 2001. **45**: p. 5-32.

[21] Freund, Y., R. Schapire, and N. Abe, *A short introduction to boosting.* Journal-Japanese Society For Artificial Intelligence, 1999. **14**(771-780): p. 1612.

[22] Bradley, A.P., *The use of the area under the ROC curve in the evaluation of machine learning algorithms.* Pattern recognition, 1997. **30**(7): p. 1145-1159.

[23] Raschka, S., *Model evaluation, model selection, and algorithm selection in machine learning.* arXiv preprint arXiv:1811.12808, 2018.

[24] Lundberg, S.M. and S.-I. Lee, *A unified approach to interpreting model predictions.* Advances in neural information processing systems, 2017. **30**.

[25] Cai, R.Y., et al., *Resting heart rate variability, emotion regulation, psychological wellbeing and autism symptomatology in adults with and without autism.* International Journal of Psychophysiology, 2019. **137**: p. 54-62.

[26] Cheng, Y.-C., Y.-C. Huang, and W.-L. Huang, *Heart rate variability in individuals with autism spectrum disorders: A meta-analysis.* Neuroscience & Biobehavioral Reviews, 2020. **118**: p. 463-471.

[27] Bazelmans, T., et al., *Heart rate mean and variability as a biomarker for phenotypic variation in preschoolers with autism spectrum disorder.* Autism Research, 2019. **12**(1): p. 39-52.

[28] Souders, M.C., et al., *Sleep behaviors and sleep quality in children with autism spectrum disorders.* Sleep, 2009. **32**(12): p. 1566-1578.

doi:10.6342/NTU202302181

[29] Dong, L., Y. Xie, and X. Zou, *Association between sleep duration and depression in US adults: A cross-sectional study.* Journal of Affective Disorders, 2022. **296**: p. 183-188.

[30] Jovevska, S., et al., *Sleep quality in autism from adolescence to old age.* Autism in Adulthood, 2020. **2**(2): p. 152-162.

doi:10.6342/NTU202302181