An Information System for Symptom Diagnosis and Improvement of Attention Deficit Hyperactivity Disorder: The ADHD360 Project

Niki Pandria, Vasileia Petronikolou, Aristotelis Lazaridis, Christos Karapiperis, Eleutherios Kouloumpris, Dimitrios Spachos, Anestis Fachantidis, Dimitrios Vasileiou, Ioannis Vlahavas, Panagiotis Bamidis

Submitted to: JMIR Research Protocols
on: June 09, 2022

Table of Contents

An Information System for Symptom Diagnosis and Improvement of Attention Deficit Hyperactivity Disorder: The ADHD360 Project

Niki Pandria[1*] BSc, MSc, PhD; Vasileia Petronikolou[1*] BSc; Aristotelis Lazaridis[2*] BSc, MSc; Christos Karapiperis[3] BSc, MSc; Eleutherios Kouloumpris[2] BSc, MSc; Dimitrios Spachos[1] BSc, MSc, PhD; Anestis Fachantidis[2] BSc, MSc, PhD; Dimitrios Vasileiou[3] MEng; Ioannis Vlahavas[2*] BSc, PhD; Panagiotis Bamidis[1*] BSc, MSc, PhD

[1]Medical Physics and Digital Innovation Lab, School of Medicine, Faculty of Health Sciences, Aristotle University of Thessaloniki Thessaloniki GR
[2]Intelligent Systems Lab, Dept. of Informatics Aristotle University of Thessaloniki Thessaloniki GR
[3]The Second Method S.A. Thessaloniki GR
[*]these authors contributed equally

Corresponding Author:
Panagiotis Bamidis BSc, MSc, PhD
Medical Physics and Digital Innovation Lab, School of Medicine, Faculty of Health Sciences, Aristotle University of Thessaloniki
Campus of Aristotle University of Thessaloniki, Building D, Entrance 8, 3rd floor
Thessaloniki
GR

Abstract

Background: Attention Deficit Hyperactivity Disorder (ADHD) is one of the most common neurodevelopmental disorders during childhood, however the diagnosis procedure remains challenging as it is non-standardized, multi-parametric and highly dependent on subjective evaluation of the perceived behavior.

Objective: To address the challenges of existing procedures for ADHD diagnosis, the ADHD360 project aims to develop a platform for (a) early detection of ADHD by assessing the user's likelihood of having ADHD characteristics and (b) providing complementary training for ADHD management.

Methods: A two-phase pilot study was designed to evaluate the ADHD360 platform, including ADHD and non-ADHD participants aged 7-16 years. Machine Learning methods were used to detect discriminative gameplay patterns among the two groups (ADHD, non-ADHD) and estimate a player's likelihood of having ADHD characteristics.

Results: A preliminary analysis of collected data showed that the trained models achieve high performance in correctly predicting a user's label (ADHD or non-ADHD) from his gameplay session in the ADHD360 platform.

Conclusions: ADHD360 is characterized by notable capacity to discriminate player gameplay behavior as either ADHD or non-ADHD. Therefore, the ADHD360 platform could be a valuable complementary tool for early ADHD detection. Clinical Trial: ClinicalTrials.gov NCT04362982

(JMIR Preprints 09/06/2022:40189)
DOI: https://doi.org/10.2196/preprints.40189

Original Manuscript

Research Protocol

Pandria Niki PhD[‡1], Petronikolou Vasileia BSc[‡1], Lazaridis Aristotelis[‡2] MSc, Karapiperis Christos MSc[3], Kouloumpris Eleftherios MSc[2], Spachos Dimitrios PhD[1], Fachantidis Anestis PhD[2], Vasileiou Dimitrios MEng[3], Vlahavas Ioannis PhD[*2], Bamidis D. Panagiotis PhD[*†1]

[1]Medical Physics and Digital Innovation Lab, School of Medicine, Faculty of Health Sciences, Aristotle University of Thessaloniki, Thessaloniki, Greece.

[2]Intelligent Systems Lab, Dept. of Informatics Aristotle University of Thessaloniki, Thessaloniki, Greece.

[3]The Second Method S.A., Thessaloniki, Greece.

† Corresponding author

‡ Equally contributing first authors

* Equally contributing senior authors

Corresponding author

Panagiotis D. Bamidis

Faculty of Health Sciences, Medical Physics and Digital Innovation Lab

Aristotle University of Thessaloniki

P.C. 54124

Thessaloniki, Greece

Email: bamidis@med.auth.gr

Tel: +30 2310 999310

An Information System for Symptom Diagnosis and Improvement of Attention Deficit Hyperactivity Disorder: The ADHD360 Project

Abstract

Background: Attention Deficit Hyperactivity Disorder (ADHD) is one of the most common neurodevelopmental disorders during childhood, however the diagnosis procedure remains challenging as it is non-standardized, multi-parametric and highly dependent on subjective evaluation of the perceived behavior.

Objective: To address the challenges of existing procedures for ADHD diagnosis, the ADHD360 project aims to develop a platform for (a) early detection of ADHD by assessing the user's likelihood of having ADHD characteristics and (b) providing complementary training for ADHD management.

Methods: A two-phase non-randomized controlled pilot study was designed to evaluate the ADHD360 platform, including ADHD and non-ADHD participants aged from 7 to16 years. At the first stage, an initial neuropsychological evaluation along with an interaction with the serious game developed ("PIZZA ON TIME") for about 30-45 minutes is performed. Subsequently, a two-week behavioral monitoring, through the mADHD360 app, is planned after a telephone conversation between the participants' parents and the psychologist where the existence of any behaviors characteristic of ADHD that affect the daily functioning is assessed. Once the behavior monitoring is complete, the research team invites the participants to progress the second stage playing the game for a mean duration of ten (10) weeks (2 times per week). Once the serious game is finished, a second round of behavioral monitoring is performed following the same procedures as the initial one. During the study, gameplay data were collected and preprocessed. The protocol of the pilot trials was initially designed for in-person participation, but after the COVID-19 outbreak, it was adjusted providing remote participation. State-of-the-art machine learning (ML) algorithms are used to analyze gameplay labeled data aiming to detect discriminative gameplay patterns among the two groups (ADHD, non-ADHD) and estimate a player's likelihood of having ADHD characteristics. A schema including a train-test splitting with 75-25% split ratio, k-fold cross validation (CV) with k=3, machine learning pipeline and data evaluation was designed.

Results: A total of forty-three (43) participants were recruited for this study, eighteen (18) of whom were diagnosed with ADHD while the rest twenty-five (25) were controls. Initial neuropsychological assessment confirmed that the participants in the ADHD group showed a deviation from the participants without ADHD characteristics. A preliminary analysis of collected data consisting of 30 gameplay sessions showed that the trained Machine Learning models achieve high performance (i.e. accuracy up to 0.85) in correctly predicting a user's label (ADHD or non-ADHD) from his gameplay session in the ADHD360 platform.

Conclusions: ADHD360 is characterized by notable capacity to discriminate player gameplay behavior as either ADHD or non-ADHD. Therefore, the ADHD360 platform could be a valuable complementary tool for early ADHD detection.

Trial Registration: ClinicalTrials.gov NCT04362982

Keywords: Attention Deficit Hyperactivity Disorder (ADHD); Machine learning; Web Health; Serious games; ADHD Monitoring

Introduction

Attention Deficit Hyperactivity Disorder (ADHD) is one of the most common neurodevelopmental disorders during childhood [1]. ADHD is commonly diagnosed during childhood and prolongs into adulthood to variable extent ranging from 5% to 75% [2]. ADHD is characterized by persistent symptoms of inattention and/or hyperactivity/impulsivity that interfere with or attenuate social, academic, and occupational functioning, as well as the developmental stage. The symptoms are present prior to the age of 12 years, for a period of at least 6 months in two or more settings [3].

The diagnosis procedure remains challenging as it is non-standardized, multi-parametric and highly dependent on subjective evaluation of the perceived behavior [4], [5]. Additionally, the effectiveness of treatment strategies commonly relies on systematic monitoring using pen-and-paper methods [4], [6]. Subjectivity bias, difficulty in monitoring in different settings and risk of data loss are inherent obstacles to those methods.

Addressing the limitations of the existing approaches, ADHD360 focuses on developing an integrated technology solution that includes a serious game for a probabilistic prediction indicating the presence of ADHD in players using Machine Learning models, and a mobile application for monitoring ADHD behaviors. The platform aims to (a) facilitate early detection of ADHD characteristics but also (b) serve as an adjunct intervention for ADHD management. The first goal of the project is addressed by analyzing the gameplay behavior of the users with respect to their diagnosis (i.e., ADHD, non-ADHD). Thus, discriminative gameplay patterns among the groups (ADHD, non-ADHD) are explored to estimate the user’s likelihood of having ADHD characteristics. The second goal is to investigate the effectiveness of the platform as an intervention for ADHD management.

Background

Discrepancies of ADHD prevalence have emerged over time and among studies, leading to mixed conclusions about the possible under-diagnosis or over-diagnosis of the disorder [7]. Multiple factors have been identified to affect the recognition and the diagnosis of ADHD including the parental role, school-based factors, intrinsic factors related to children as well as the role of health providers [8].

Focusing on the medical role, the access of the public to health providers [7], the limited reimbursement for specialized mental care [9] as well as differences in clinical approaches, scoring cut-off and factors related to the medical system across different countries [10] have attributed to the existing difficulties in the diagnosis process. However, the underlying inconsistencies in the definition of ADHD based on different diagnostic manuals make the recognition and the diagnosis of ADHD even more challenging [3], [11]. Moreover, ADHD diagnosis is possible to be biased due to the underlying subjectivity of the assessment procedure, the information interpretation [12] and the lack of standardization [11].

Although diagnostic challenges could be addressed through a detailed history of prenatal conditions, family status and school/academic life [7], [13], there is a need for standardized tools that could enhance the diagnosis accuracy.

As early detection of ADHD could ameliorate the disorder’s development, diminish its long-term impact [14], improve quality of life [15], overall functioning and self-esteem [16], [17], research efforts were focused on formulating programs [14], [18], [19], games [20], [21] or game-based tools [22], [23] for early diagnosis.

Addressing the existing challenges, the ADHD360 aims to develop an integrated platform having as core elements a serious game along with a mobile application for monitoring ADHD behaviors in a SMART (Specific, Measurable, Attainable, Realistic and Timely) way [24], [25]. The design of the serious game is based on both Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-V) [3] along with neuropsychological tools, easily transferred to the game design process via game mechanics implementation, on a specific ADHD behavior.

Methods

Study design

A two-phase non-randomized controlled pilot study is performed in the context of ADHD360 project. The protocol of the pilot trials was initially designed for in-person participation, but after the COVID-19 outbreak, it was adjusted to the new conditions imposed by the pandemic, providing remote participation. The procedures for each of these two ways of participating are described thoroughly below.

In-person participation

The first part of the clinical trials begins with a first meeting between the involved members of the research team and the parents of each participant at the Laboratory of Medical Physics and Digital Innovation, School of Medicine of Aristotle University of Thessaloniki (iMedPhys). The aim of this first meeting is to thoroughly inform the parents and the children regarding the nature and the scope of the project, as well as the exact procedures that take place according to the research protocol. In this context, parents can express any possible question they might have and sign the informed consent form. Afterwards, the participants have a brief neuropsychological assessment by an experienced psychologist from the research team. This neuropsychological assessment includes six (6) subtests of the Wechsler Intelligence Scale for Children, Fifth Edition (WISC-V), an intelligence test that measures a child's intellectual ability and five (5) cognitive domains that impact performance [26]. In particular, the six (6) subtests of WISC-V deployed here are the following: 1. Similarities, 2. Vocabulary, 3. Block Design, 4. Figure Weights, 5. Digit Span and 6. Coding. The main aim of this assessment is to obtain a more integrated view of the participant's general intellectual ability, in several domains such as verbal comprehension, visual-spatial perception, working memory, processing speed and fluid reasoning. The average time to complete the subtests is about 48-50 minutes. If any of the participants has been assessed by this scale during the past two years, he/she is not reassessed by our research team and the children's parents are kindly asked to provide the already existing scores at the aforementioned six (6) subtests, if possible. Furthermore, the ADHD RATING SCALE-IV [27] is a 4-point Likert brief questionnaire completed by the parents regarding the presence and frequency of ADHD symptoms. Two subscales, Inattention and Hyperactivity-Impulsivity are integrated into the ADHD RATING SCALE-IV. The total raw score of the scale is calculated by summing the scores in the two subscales.
Subsequently, the participants interact with the serious game "PIZZA ON TIME" that has been specifically developed for the ADHD360 project, for about 30-45 minutes, in the iMedPhys. During this interaction with the Serious Game (SG), a form is completed regarding the conditions, the duration and the behaviors that will possibly be occurred. Participants interact again with the SG for 30-45 minutes, in the iMedPhys, continuing from the level at which they stopped at the previous visit, in order to complete all the levels of the first setting (the urban setting).
After the aforementioned procedures, there is a telephone conversation between the participants' parents and the psychologist who conducted the neuropsychological evaluation to discuss if there are any behaviors characteristic of ADHD that affect the daily functioning of the participants. Once the psychologist has reached a consensus with the parents that there are certain behaviors that could be observed, the behaviors are introduced into the participant's account in the mADHD360 app. Instructions regarding the installation and the use of the mobile application are provided by the research team through email. The mADHD360 app can be used on smartphone or tablet. An observation plan is scheduled including at least two to four (2-4) times per week of behavioral monitoring regarding the frequency or the duration of these behaviors occurred during the time they

usually spend with the participant. Each round of behavior monitoring has a duration of twelve (12) minutes. The duration of this phase is determined to be two (2) weeks. During the behavior monitoring phase, the participants do not interact with the SG.

Once the behavior monitoring is complete, the research team invites the participants to continue playing the game, i.e. the second (jungle) and third (space) setting on the premises of iMedPhys. The mean duration of this part is ten (10) weeks with a frequency of at least two (2) times a week, depending on the participants' availability. In this second part of the pilot testing phase, the goal is to exploit the serious game as an intervention to improve certain ADHD symptoms.

Afterwards, a second round of behavioral monitoring is performed following the same procedures as the previous one. Furthermore, a neuropsychological assessment is conducted following the same procedures that took place during the first evaluation in the beginning of their participation.

Remote participation

All the procedures are identical to those followed in the in-person participation, apart from the interactions with the SG which are done remotely, from their home using their personal computers.

Patient Population

The eligibility criteria are reviewed by a member of the research team before the assignment to the study. The criteria should be met before enrollment into the study.

Inclusion criteria

The eligible participants should satisfy the following inclusion criteria:

- Participants should be between seven (7) to sixteen (16) years old.
- Participants of the ADHD group should be diagnosed by an approved body of the Ministry of Health.
- Participants should be willing to follow the study protocol and procedures.
- The ADHD symptoms should not be attributed to organic disease.
- Parents voluntarily provided written consent for their child's participation in the study.

Exclusion criteria

Participants presenting any of the following criteria will be excluded from the study:

- Participants suffering from other comorbid conditions (other than ADHD).
- Parents who are not willing to provide written consent for their child's participation in the study.

All participants are comprehensively informed regarding the procedures of the study by means of a document named "Participant Information & Consent Form". Moreover, the participants are informed about the purpose of the study as well as the following characteristics of their participation:

- Their participation is voluntary.
- They can ask questions about the study procedures before participating in the study.
- They will be aware of the underlying risk or burden of their participation.
- They will know who will benefit from conducting the research.
- They will know how data collection and data protection will be performed during the

project’s lifetime as well as whether data will be destroyed or reused (in case there is a possibility to reuse them).

- They will be fully informed, and they will agree with the further use of their data.
- They can leave the study at any time or/and withdraw their data from the study.
- They will be aware of the possible commercial exploitation of the research.

Each informed consent is signed by the researcher who informs the participant and the principal investigation of the study. The informed consent explicitly states that the study was approved by the Committee for Bioethics and Ethics of Medical School (Aristotle University of Thessaloniki) and all personal information are anonymized using a unique user identification code.

Ethics approval

The study protocol was approved by the Ethics and Bioethics Committee of the School of Medicine at the Aristotle University of Thessaloniki (6.225/29.7.2020).

Recruitment of participants

We initially considered recruiting at least twenty (20) participants (10 ADHD, 10 non-ADHD) in our pilot trials. We estimated the sample size for the definite study by performing power analysis using GPower software (version 3.1). We conducted a paired t-test analysis using 0.80 power, a significance level of 0.50 and an effect size equal to 0.20. The total sample size was estimated to be one hundred ninety-nine (199) participants. As Billingham et al. [28] proposed, the sample size in pilot trials does not need calculation but it still needs justification. Thus, the small number of participants that is planned to be recruited seems to be sufficient according to the suggestions of Stallard [29] and similar to what was proposed by Julious [30].
The ADHD group is mainly composed of patients from the Community Center for Mental Health of Children and Adolescents of General Hospital G. Papanikolaou which supports the implementation of the pilot trials. Furthermore, dissemination activities such as social media posts, mass media interviews, information events open to the public, project presentations at conferences and workshops are considered essential to recruit participants in pilot trials.

The ADHD360 platform design

The ADHD360 platform consists of two main components, a serious game and a mobile application for behavior monitoring.

A serious game for ADHD

Game design covers a wide range of activities in designing games, including story, aesthetics, mechanics and technology [31]. These variables have to be properly considered for a successful game design. Hunicke et. al propose the Mechanics-Dynamics-Aesthetics (MDA) framework and define game mechanics as the “mechanisms that describe the specific elements of the game, at the level of data presentation and algorithms (...), also the mechanics relate to the behaviors and control mechanisms provided to the player in the game context” [32].
The MDA framework standardizes the basic components of a game into distinct parts such as rules, systems, and fun. This distinction helps us in the design of the mechanisms since the basic components of the game can be translated into game-design elements such as mechanics, dynamics and aesthetics that are visible to both the end-users and the game designer but from different perspectives. There are some games that have been developed specifically to improve the symptoms of ADHD, while commercial game titles have also been used for research purposes to measure the

performance of users with or without ADHD diagnosis [33].

The two games that focus on improving ADHD symptoms are Plan it Commander [34] and Antonyms [35]. Two main gaps emerge in the existing research: i) The interconnection of game mechanics with the diagnostic criteria according to DSM-V and ii) the correlation of user performance with the standard results of ADHD tests through machine learning algorithms.

Bridging this gap, we focused on developing a serious game called “PIZZA ON TIME” (Figure 1), which is a runner game. The player tries to avoid obstacles and collect coins to deliver the pizza, and each time the player hits an obstacle a pizza slice is lost. When all the pizza slices are lost, the player needs to start over at the same level. There are three different in-game world levels (city, jungle and space) with a total of one hundred twenty (120) predefined sub-levels. Additionally, four (4) mini-games were integrated to the main runner game to enrich the existing mechanics, retaining the concept of pizza delivery.

Mobile application for behavior monitoring

The ADHD360 monitoring application (mADHD360) developed in the project provides teachers, parents, and health professionals with features to easily monitor specific targeted behaviors related to ADHD. While in the past years a significant amount of related ICT solutions tried to replace traditional pen-and-paper observation charts [24], not all of them succeeded [6]. The mADHD360 app differs from other digital monitoring applications, which are mainly based on pre-existing traditional forms of assessment. Instead, the app addresses the need to include in a unique, easy-to-use tool the range of observation features needed to conduct a complete Functional Behavioral Assessment. Finally, it allows teachers, parents, and clinicians to use a social-centric technology to create a network of people, develop and monitor a behavioral intervention plan shared among all persons involved in the care of the child with ADHD.

The process of creating a new subject and monitoring its behavior can be described as follows:

First, the associate user (usually one of the parents) creates an account and enters the basic data of the subject to be monitored in the mADHD360 app. There is a strict policy regarding the sensitive personal data of the children, which does not allow real names to be used, or any other significant detail that can reveal the identity of the child in the real world.

The next step is to create a network of people, usually people that spend time with the child (such as a relative or a class teacher) and a health professional. The app supports a directory of ADHD experts, which can be added to the child's observation network. The invitation of other people is a very simple process, and it can be performed via email.

A limited number of ADHD-related behaviors is associated with the child. The associations come from a predefined vocabulary [36] but the user can also enter custom behaviors. The network of people can gather data to unveil the function of a child‘s behavior and plan, with the help of a health professional, an intervention to reduce or eliminate the undesirable behavior. Data are gathered in sessions, and for each behavior, the application can monitor the frequency or/and the duration. Usually, there are two (2) rounds of monitoring, one round before the intervention and the other round during or even after the intervention plan.

Finally, the efficacy of each treatment can be evaluated by health professionals through a visual analysis of the data gathered by the network members during the assessment periods.

The mADHD360 app is based on the WHAAM application [36], and although the basic principles are retained, the latter app is enhanced with more features, such as asynchronous chat capabilities, and brings advanced usability and user experience.

From a technology point of view, the mADHD360 app, is designed to run on mobile devices, smartphones, and tablets. It is delivered as a progressive web app accompanied for free with the main game.

Analysis

Statistical analysis

Statistical analysis is performed using Statistical Package for the Social Sciences (SPSS, Chicago, IL, USA) (version 26.0) and the statistical significance level is set at 0.05. Continuous variables are explored for normality by means of the Shapiro-Wilk test to calculate the appropriate descriptive statistics. Continuous variables that are approximately normally distributed are reported as mean±standard deviation (SD) while those that are not normally distributed as median and interquartile range [Q1, Q3] where Q1 and Q3 are the first and the third distribution quartile respectively. Neuropsychological data expressed as raw scores were treated as continuous variables. Categorical data are described as frequencies and percentages.

Machine Learning methodology

State-of-the-art machine learning (ML) algorithms are used to analyze gameplay labeled data. Specifically, we attempt to classify game-generated data based on the user's known label (ADHD, non-ADHD). All in-game events of a gameplay session are recorded to form a timeseries that one can use to reconstruct the whole gameplay session. These events correspond to player actions and environment parameters, such as performing a jump/move left/move right action, obstacle collision and coins collected. Informative features extraction is performed by splitting players' timeseries depending on the game level. Different feature extraction techniques are assumed for each level depending on its type. In the main runner game, several aggregates are calculated by capturing different events during the gameplay i.e., movement of a player or the ability to avoid obstacles. In the mini-games, feature extraction is based on the characteristic game mechanics and events of the respective mini-game. The final feature vector of each player is produced by concatenating the individual level feature vectors.

A schema that includes train-test splitting with 75-25% split ratio, *k*-fold cross validation (CV) with *k=3*, machine learning pipeline and data evaluation was designed for experimental purposes. The train-test splitting is used for model specification (e.g., feature selection and model tuning), model selection and training of the final model. Modelling decisions are made on the training set. Thus, the test set is used only to evaluate and report the statistics of algorithms' performance. The purpose of the 3-fold CV is to finetune the hyperparameters of the supervised learning algorithms (classifiers) that we investigate.

For each classifier, a machine learning pipeline was designed to include the following procedures: (a) removal of features with zero variance, (b) univariate feature selection according to a statistical criterion (Analysis of Variance F-value between label and feature), (c) Principal Components Analysis (PCA) for dimensionality reduction and (d) classifier training. Therefore, the hyper-parameters of all steps are tuned (e.g., the number of components in PCA) in the 3-fold CV process. The F1-score metric is used as the optimization criterion in 3-fold CV, and the best configuration of each method's pipeline is then trained on the entire training set (fully trained pipeline).

Afterwards, the algorithm with the best-performing pipeline in 3-fold CV is selected as the proposed method. The general performance of our system on unlabeled data is estimated by evaluating the predictive performance of the proposed method that consists of a fully trained pipeline on the test set. The performance of ML system is also evaluated based on methods other than the proposed one. The learning algorithms that were applied are the following: k-Nearest Neighbors [37] (kNN), Logistic Regression [38] (LogReg), Support Vector Machine [39] (SVM), Random Forest [40] (RF), Ridge Classifier [41] (RC), Passive Aggressive Classifier [42] (PAC), Stochastic Gradient Descent [43] (SGD) and Naive Bayes [44] (NB). For each classifier, we tuned its most important parameters by using grid-search cross-validation on a sufficient range of parameter values. The objective of the optimization is to maximize the micro F1-Score across all validation folds. In addition, we report the performance metrics of Accuracy, Precision, Recall and F1 scores.

Results

Demographics

In total, forty-three (43) participants were recruited with a mean age of 11.82±2.81 years. Twenty-eight (28) out of forty-three (43) participants (65%) were male while fifteen (15) out of forty-three (43) subjects (35%) were female. Eighteen (18) out of forty-three (43) participants (42%) were diagnosed with ADHD while twenty-five (25) out of forty-three (43) (58%) participants were considered as controls (non-ADHD). About 70% of participants (30 subjects) completed the first part of pilot phase remotely.

Neuropsychological data

Neuropsychological data were available from thirty (30) participants. These subjects were evaluated using WISC-V by the psychologist of the ADHD360 team or by an approved national body within the last two (2) years. Ten (10) out of thirty (30) participants (33%) were diagnosed with ADHD while twenty (20) out of thirty (30) (67%) were considered as healthy controls (non-ADHD). The 70% of participants were male (21 subjects) while the remaining nine (9) participants (30%) were female. In the ADHD group, all participants were male while in the non-ADHD group eleven (11) subjects (55%) were male and nine (9) subjects (45%) were female. The mean age of participants screened was 11.40±2.85 years. In the ADHD group, the mean age of participants was 12.90±2.64 whereas in the non-ADHD group was 10.65±2.70.

Analyzing data from the first neuropsychological assessment, the mean score in Figure Weights subtest in the ADHD group was 14.00±5.79 while the median score of this subtest in the non-ADHD group was 19.00, [13.00, 24.00]. The mean scores in the subtests Block Design, Similarities, Digit Span, Coding and Vocabulary are displayed in the following Table 1.

Table 1: Mean scores of ADHD and non-ADHD participants in the subtests Block Design, Similarities, Digit Span, Coding and Vocabulary.

WISC subtest	**ADHD**	**non-ADHD**
Block Design	19.60±4.30	22.85±8.02
Similarities	22.00±7.45	23.50±9.05
Digit Span	18.30±4.03	19.75±6.26
Coding	27.80±9.69	32.20±13.46
Vocabulary	20.20±5.55	23.05±8.17

The mean score of the participants in the Inattention subscale of the ADHD RATING SCALE-IV was 12.70±7.04 while in the Hyperactivity/Impulsivity subscale was 10.97± 6.65. In the full group of participants, the mean total score was 23.67±12.56. Grouping the participants based on their diagnosis, the mean scores in the ADHD RATING SCALE-IV are described in Table 2.

Table 2: Mean scores of ADHD and non-ADHD participants in the ADHD RATING SCALE-IV.

ADHD RATING SCALE-IV	**ADHD**	**non-ADHD**
Inattention subscale	18.50±5.84	9.80±5.74
Hyperactivity-Impulsivity subscale	15.30±6.58	8.80±5.67
Total score	33.80±9.94	18.60±10.62

Gameplay data

Following the methodology described above, we present the results of a preliminary analysis of gameplay data collected during pilot trials. Gameplay data were collected by all recruited participants. However, technical difficulties (i.e., unstable internet connection and game installation issues) led to insufficient or no data generation in thirteen (13) participants (30%). Therefore, gameplay data collected by thirty (30) subjects (70%) were used in the following preliminary analysis. In more detail, we describe the cross-validation results for a training set consisting of data collected by twenty-two (22) participants (73%) as well as the selection of the best model. Subsequently, we evaluate the predictive performance on a test set consisting of data gathered by eight (8) participants (27%) using point estimates and confidence intervals.
A graphical illustration of the cross-validation F1 results for each learning method is displayed in Figure 2. Furthermore, detailed results for F1 and the other performance measures for the cross-validation procedure are given in Table 3.

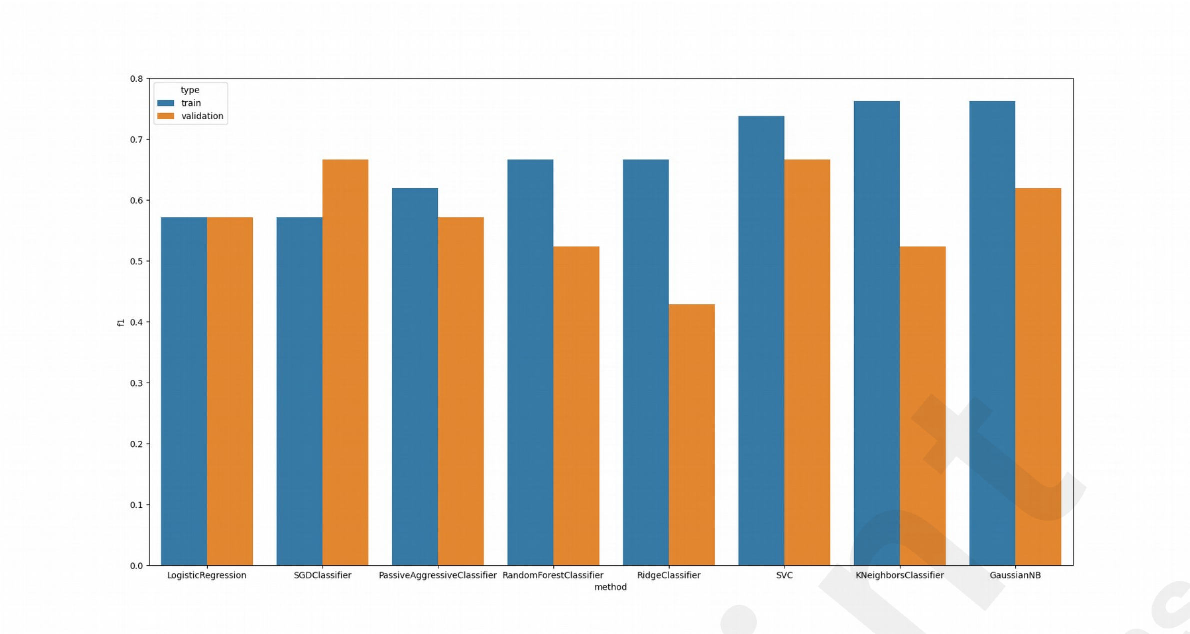


Table 3: Cross-validation results.

Method	Accuracy		Precision		Recall		F1	
	Train	Val.	Train	Val.	Train	Val.	Train	Val.
NB	0.7619	0.6190	0.7011	0.5556	0.8333	0.7778	0.7619	0.6190
k-NN	0.7619	0.5238	0.7083	0.4500	0.7778	0.5556	0.7619	0.5238
LogReg	0.5714	0.5714	0.0000	0.0000	0.0000	0.0000	0.5714	0.5714
PAC	0.6190	0.5714	0.5337	0.5000	0.8333	0.8889	0.6190	0.5714
RF	0.6667	0.5238	0.5935	0.4444	0.7222	0.5556	0.6667	0.5238
RC	0.6667	0.4286	0.6074	0.3833	0.6667	0.5556	0.6667	0.4286
SGD	0.5714	0.6667	0.4932	0.5667	0.6667	1.0000	0.5714	0.6667
SVC	0.7381	0.6667	0.7238	0.5833	0.6667	0.6667	0.7381	0.6667

Most classifiers achieve an F1-score above 0.5. We observe significant overfitting for Random Forest, kNN and Naive Bayes, with a 0.1 difference between train and validation scores. Test performance scores are given in Table 4.

Table 4: Test results evaluation.

Method	Accuracy	Precision	Recall	F1

	CI lower	CI upper	Test	CI lower	CI upper	Test	CI lower	CI upper	Test	CI lower	CI upper	Test
NB	0.14	0.85	0.42	0.14	0.85	0.42	1.00	1.00	1.00	0.25	0.92	0.60
k-NN	0.14	0.71	0.42	0.00	1.00	0.33	0.00	1.00	0.33	0.00	0.67	0.33
LogReg	0.14	0.85	0.42	0.14	0.85	0.42	1.00	1.00	1.00	0.25	0.92	0.60
PAC	0.13	0.71	0.28	0.13	0.75	0.33	0.23	1.00	0.66	0.23	0.83	0.44
RF	0.28	0.85	0.57	0.15	1.00	0.50	0.23	1.00	0.66	0.23	0.90	0.57
RC	0.27	0.85	0.57	0.00	1.00	0.50	0.00	1.00	0.66	0.00	0.88	0.57
SGD	0.14	0.71	0.42	0.00	1.00	0.33	0.00	0.76	0.33	0.00	0.66	0.33
SVC	0.71	1.00	0.85	0.50	1.00	0.75	1.00	1.00	1.00	0.66	1.00	0.85

The previously mentioned models seem to be substandard solutions in this classification problem, all dominated by LogReg, SGD, SVM and PAC. The latter models feature a trade-off between validation performance and goodness of fit. Among these methods, we choose to proceed with SVM for the following reasons: (1) it achieves the highest validation F1-score (=0.66), (2) while SGD achieves the same validation score as SVM, the latter also achieves a better fit with approximately +0.7 difference between train and validation, which for the former is about -0.9.

While the results are indicative of the potential that machine learning has in the domain of ADHD prediction based on gameplay data, we presume that experimenting with a larger dataset could give more accurate and concrete conclusions.

Discussion

Limitations of existing screening procedures for ADHD diagnosis are summarized in high dependence on pen-and-paper practices, subjectivity bias, difficulty in monitoring in different settings and risk of data loss. The ADHD360 platform, consisting of the serious game, "PIZZA ON TIME" and the mADHD360 app, attempts to be a holistic technological solution for early detection of ADHD characteristics and a training tool against ADHD-related symptoms. It facilitates real-time data collection in different settings and from different individuals involved in the user's care. Moreover, the ADHD360 platform enables quantitative data analysis assessing the behavior and the game performance of the user while providing a more integrated view of the user's behavioral characteristics. Investigating different state-of-the-art ML methods, we showed that our platform is characterized by a notable capacity to discriminate players based on their in-game patterns in those who have ADHD characteristics from those who do not. Thus, it is expected to serve as a complementary screening tool affecting positively healthcare professionals, educators, and people of ADHD.

However, the health emergency imposed by COVID-19 affected the implementation of pilot trials by forcing an adaptation of the protocol to remote conditions. Moreover, the lockdown and social distancing may have contributed to the low sample size and the high number of dropouts during the study. Additionally, the long duration of the experimental protocol and the need for commitment to study procedures may have negatively affected users' adherence to pilot trials. Moreover, ADHD has been characterized by "a dislike of mental effort" [45]. Taking into account the limitations listed above, we recruited about twice as many participants as originally considered. However, the drop-out rate was around 30% (13 participants) as thirty (30) subjects (70%) completed the first part of pilot phase remotely.

Even though the preliminary analysis of the collected data has shown promising results, the platform can be further improved by training the machine learning models in a larger dataset that can be developed by recruiting more participants. This would allow the learning models to improve further their accuracy in correctly distinguishing ADHD from non-ADHD gameplay behaviors. The monitoring application will soon get a dedicated enrollment service for health experts: parents will be able to see a listing of related professionals, near their area, and contact a person to enjoy a child's network. Moreover, we will add asynchronous chatting capabilities and strict monitoring schedule functionality, with device and calendar notifications. The "PIZZA ON TIME" game is planned to be released to the public for both Android and iOS devices, as well as for PC (Windows and Mac).

Acknowledgements

ADHD360 is co-financed by the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship, and Innovation, under the call RESEARCH–CREATE–INNOVATE [T1EΔK-01680].

Data availability

The data sets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Abbreviations

ADHD: Attention Deficit Hyperactivity Disorder
CV: cross validation
DSM-V: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
iMedPhys: Medical Physics and Digital Innovation, School of Medicine of Aristotle University of Thessaloniki
kNN: k-Nearest Neighbors
LogReg: Logistic Regression
mADHD360: ADHD360 monitoring application
MDA: Mechanics-Dynamics-Aesthetics
ML: Machine Learning
NB: Naive Bayes
PAC: Passive Aggressive Classifier
PCA: Principal Components Analysis
RC: Ridge Classifier
RF: Random Forest
SD: Standard Deviation
SG: Serious Game
SGD: Stochastic Gradient Descent
SVM: Support Vector Machine
WISC-V: Wechsler Intelligence Scale for Children, Fifth Edition

References

1. T. E. Wilens and T. J. Spencer, “Understanding attention-deficit/hyperactivity disorder from childhood to adulthood.,” Postgrad. Med.;2010, 122(5):97–109, doi: 10.3810/pgm.2010.09.2206.
2. M. H. Sibley et al., “Defining ADHD symptom persistence in adulthood: optimizing sensitivity and specificity.,” J. Child Psychol. Psychiatry.;2017, 58(6):655–662, doi: 10.1111/jcpp.12620.
3. American Psychiatric Association, Diagnostic and Statistical Manual of Mental Disorders, 5th ed. Washington, VA: American Psychiatric Publishing, 2013.
4. C. T. Gualtieri and L. G. Johnson, “ADHD: Is Objective Diagnosis Possible?,” Psychiatry (Edgmont);2005, 2(11):44–53, PMID: 21120096.
5. A. R. Adesman, “The Diagnosis and Management of Attention-Deficit/Hyperactivity Disorder in Pediatric Patients.,” Prim. Care Companion J. Clin. Psychiatry; 2001, 3(2): 66–77, doi: 10.4088/pcc.v03n0204.
6. N. Pandria, D. Spachos, and P. D. Bamidis, “The future of mobile health ADHD applications: Projecting WHAAM application on future mobile health,” in Proceedings of 2015 International Conference on Interactive Mobile Communication Technologies and Learning, IMCL 2015, 2015, doi: 10.1109/IMCTL.2015.7359603.
7. A. M. Hamed, A. J. Kauer, and H. E. Stevens, “Why the Diagnosis of Attention Deficit Hyperactivity Disorder Matters.,” Front. psychiatry;2015, 6 (168), doi: 10.3389/fpsyt.2015.00168.

8. M. Zwaanswijk, P. F. M. Verhaak, J. M. Bensing, J. van der Ende, and F. C. Verhulst, "Help seeking for emotional and behavioural problems in children and adolescents: a review of recent literature.," Eur. Child Adolesc. Psychiatry;2003, 12(4): 153–61, doi: 10.1007/s00787-003-0322-6.
9. Subcommittee on Attention-Deficit/Hyperactivity Disorder et al., "ADHD: clinical practice guideline for the diagnosis, evaluation, and treatment of attention-deficit/hyperactivity disorder in children and adolescents.," Pediatrics;2011,128(5):1007–22, doi: 10.1542/peds.2011-2654.
10. P. J. Santosh et al., "Refining the diagnoses of inattention and overactivity syndromes: A reanalysis of the Multimodal Treatment study of attention deficit hyperactivity disorder (ADHD) based on ICD-10 criteria for hyperkinetic disorder," Clin. Neurosci. Res.;2005, 5(5–6):307–314, doi: 10.1016/j.cnr.2005.09.010.
11. I. Berger, "Diagnosis of attention deficit hyperactivity disorder: much ado about something.," Isr. Med. Assoc. J.;2011, 13(9):571–4, PMID: 21991721.
12. B. Kiely and A. Adesman, "What we do not know about ADHD… yet.," Curr. Opin. Pediatr.;2015, 27(3): 395–404, doi: 10.1097/MOP.0000000000000229.
13. J. Zhu, K. P. Lee, T. J. Spencer, J. Biederman, and P. G. Bhide, "Transgenerational transmission of hyperactivity in a mouse model of ADHD.," J. Neurosci.;2014, 34(8):2768–73, doi: 10.1523/JNEUROSCI.4402-13.2014.
14. E. J. S. Sonuga-Barke, J. Koerting, E. Smith, D. C. McCann, and M. Thompson, "Early detection and intervention for attention-deficit/hyperactivity disorder.," Expert Rev. Neurother.;2011, 11(4):557–63, doi: 10.1586/ern.11.39.
15. R. Agarwal, M. Goldenberg, R. Perry, and W. W. IsHak, "The quality of life of adults with attention deficit hyperactivity disorder: a systematic review.," Innov. Clin. Neurosci.;2012, 9(5–6): 10–21, PMID: 22808445.
16. V. Harpin, L. Mazzone, J. P. Raynaud, J. Kahle, and P. Hodgkins, "Long-Term Outcomes of ADHD: A Systematic Review of Self-Esteem and Social Function.," J. Atten. Disord.;2016, 20(4): 295–305, doi: 10.1177/1087054713486516.
17. M. Shaw et al., "A systematic review and analysis of long-term outcomes in attention deficit hyperactivity disorder: effects of treatment and non-treatment.," BMC Med.;2012, 10(1):99, doi: 10.1186/1741-7015-10-99.
18. Armayani, M. Yati, C. Yusnayanti, and Firman, "Minimizing attention of deficit hyperactivity disorder in children ages 7–10 years through early detection in state 1st SD 1 Poasia Kendari 2017," Enfermería Clínica;2020, 30:81–83, doi: 10.1016/j.enfcli.2019.11.026.
19. M. Vukojević, A. Dizdarević, and D. Novaković, "Early detection and recognition of children with ADHD (attention deficit hiperactivity disorder) symptoms.," Coll. Antropol.;2012, 36(4):1183–8, PMID: 23390809.
20. A. Mendez et al., "Temporal perception and delay aversion: A videogame screening tool for the early detection of ADHD," Rev. Argent. Cienc. Comport.;2015, 7:90–101, doi: 10.30882/1852.4206.V7.N3.10110.
21. A. E. Alchalabi, S. Shirmohammadi, A. N. Eddin, and M. Elsharnouby, "FOCUS: Detecting ADHD patients by an EEG-based serious game," IEEE Trans. Instrum. Meas.;2018, 67(7):1512–1520, doi: 10.1109/TIM.2018.2838158.
22. M. F. Pascual, B. G. Zapirain, and A. M. Zorrilla, "Diagnosis of the attention deficit disorder using 'D2' and 'Symbols Search' tests through a game-based tool," in 2012 17th International Conference on Computer Games (CGAMES); 2012, pp. 116–119, doi: 10.1109/CGames.2012.6314561.
23. A. Serrano-Barroso et al., "Detecting Attention Levels in ADHD Children with a Video

Game and the Measurement of Brain Activity with a Single-Channel BCI Headset.," Sensors (Basel);2021, 21(9):3221, doi: 10.3390/s21093221.
24. D. Spachos, A. Chifari, G. Chiazzese, G. Merlo, G. Doherty, and P. Bamidis, "WHAAM: A mobile application for ubiquitous monitoring of ADHD behaviors," in 2014 International Conference on Interactive Mobile Communication Technologies and Learning (IMCL2014);2014, pp. 305–309, doi: 10.1109/IMCTL.2014.7011153.
25. G. Merlo et al., "The WHAAM Application: a Tool to Support the Evidence-Based Practice in the Functional Behaviour Assessment.," J. Innov. Heal. informatics; 2018, 25(2):63–70, doi: 10.14236/jhi.v25i2.919.
26. Raiford, S. E. "The Wechsler Intelligence Scale for Children—Fifth Edition Integrated";2018, In D. P. Flanagan & E. M. McDonough (Eds.), Contemporary intellectual assessment: Theories, tests, and issues. The Guilford Press, ISBN: 9781462535781.
27. G. J. DuPaul, T. J. Power, A. D. Anastopoulos, and R. Reid, ADHD Rating Scale—IV: Checklists, norms, and clinical interpretation. New York, NY, US: Guilford Press, 1998.
28. S. A. Billingham, A. L. Whitehead, and S. A. Julious, "An audit of sample sizes for pilot and feasibility trials being undertaken in the United Kingdom registered in the United Kingdom Clinical Research Network database," BMC Med. Res. Methodol.;2013, 13(1):104, doi: 10.1186/1471-2288-13-104.
29. N. Stallard, "Optimal sample sizes for phase II clinical trials and pilot studies," Stat. Med.;2012, 31(11–12):1031–1042, doi: 10.1002/sim.4357.
30. S. A. Julious, "Sample size of 12 per group rule of thumb for a pilot study," Pharm. Stat.;2005, 4(4):287–291, doi: 10.1002/pst.185.
31. K. Oksanen, T. Lainema, and R. Hämäläinen, "Learning from social collaboration: A paradigm shift in evaluating game-based learning," in Handbook of research on serious games for educational applications, IGI Global; 2017, pp. 41–65, doi: 10.4018/978-1-5225-0513-6.ch003.
32. R. Hunicke, M. LeBlanc, and R. Zubek, "MDA: A formal approach to game design and game research," in Proceedings of the AAAI Workshop on Challenges in Game AI; 2004, 4(1):1722.
33. R. Shaw, A. Grayson, and V. Lewis, "Inhibition, ADHD, and computer games: the inhibitory performance of children with ADHD on computerized tasks and games.," J. Atten. Disord.;2005, 8(4):160–8, doi: 10.1177/1087054705278771.
34. Bul KCMM, Franken IHAA, Van der Oord S, Kato PM, Danckaerts M, Vreeke LJ, et al. Development and User Satisfaction of "Plan-It Commander," a Serious Game for Children with ADHD. Games Health J;2015, 4(6):502–12, doi: 10.1089/g4h.2015.0021.
35. V. Colombo, D. Baldassini, S. Mottura, M. Sacco, M. Crepaldi, and A. Antonietti, "Antonyms: A serious game for enhancing inhibition mechanisms in children with Attention Deficit/Hyperactivity Disorder (ADHD),"International Conference on Virtual Rehabilitation (ICVR); 2017, doi: 10.1109/ICVR.2017.8007457.
36. N. Pandria, D. Spachos, and P. D. Bamidis, "Evaluation of Neurofeedback on ADHD Using Mobile Health Technologies," in IFMBE Proceedings;2016, 57: 1292–1296.
37. N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," Am. Stat.;1992, 46(3):175–185, doi: 10.2307/2685209.
38. P. McCullagh and J. A. Nelder, Generalized linear models. Routledge; 2019, ISBN 9780412317606.
39. V. Vapnik, The nature of statistical learning theory. Springer science & business media; 1999, ISBN: 978-1-4757-3264-1.
40. L. Breiman, "Random forests," Mach. Learn.;2001, 45(1):5–32, doi: 10.1023/A:1010933404324.

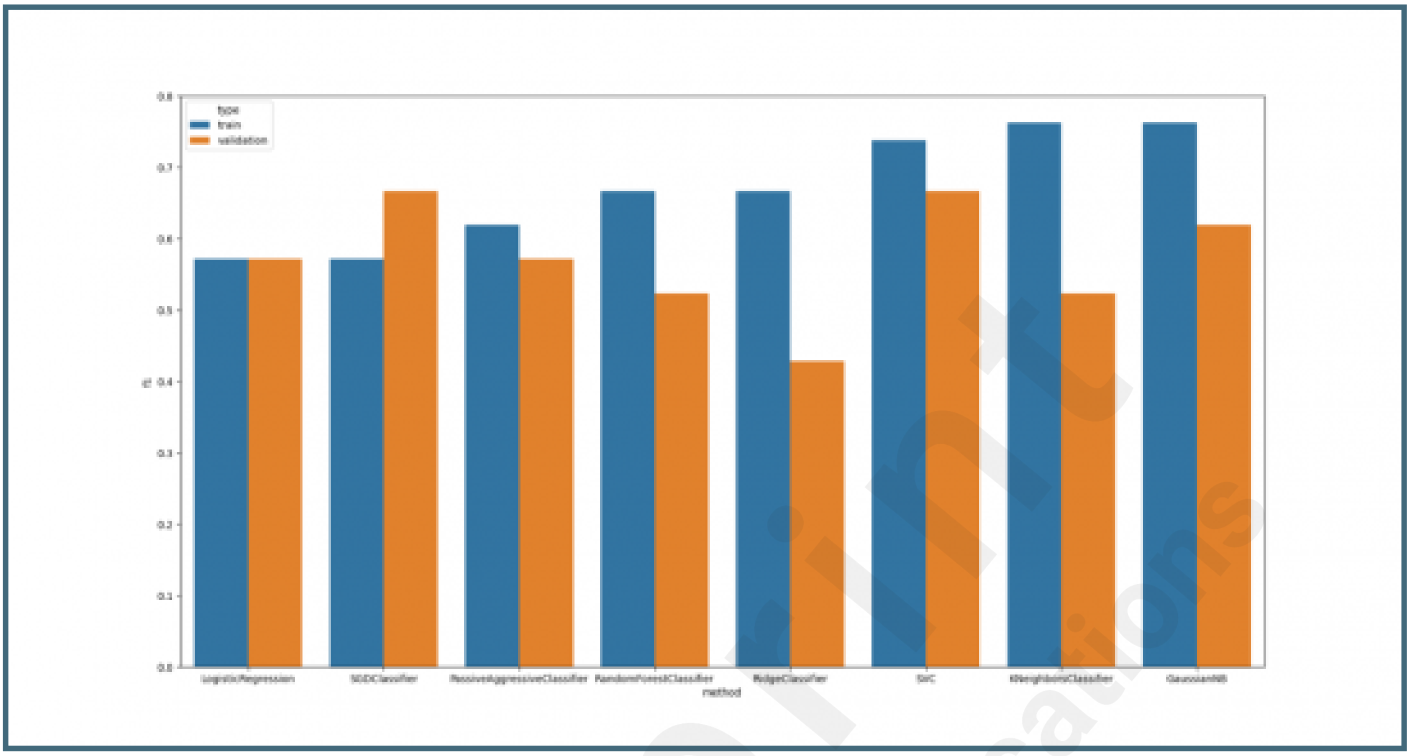
41. H. Gan, Q. She, Y. Ma, W. Wu, and M. Meng, “Generalization improvement for regularized least squares classification,” Neural Comput. Appl.;2019, 31(S2):1045–1051, doi: 10.1007/s00521-017-3090-9.
42. K. Crammer, O. Dekel, J. Keshet, S. Shalev-Shwartz, and Y. Singer, “Online passive aggressive algorithms,” 2006, doi: 10.5555/1248547.1248566.
43. N. Ketkar, “Stochastic gradient descent,” in Deep learning with Python, Springer; 2017, pp. 113–132, doi: 10.1007/978-1-4842-2766-4.
44. I. Rish and others, “An empirical study of the naive Bayes classifier,” in IJCAI 2001 workshop on empirical methods in artificial intelligence; 2001, 3(22):41–46.
45. S. H. Mitchell and D. Sevigny-Resetco, “Effort-Related Decision-Making in ADHD.,” J. psychiatry brain Sci.;2020, doi: 10.20900/jpbs.20200027.

Supplementary Files

Figures

Screenshot of the “PIZZA ON TIME” game.

Cross-Validation F1-scores.



Multimedia Appendixes

Peer review report after evaluating the ADHD360 project proposal.
URL: http://asset.jmir.pub/assets/e030bb6fd2605563c0165352220b0175.docx