

University of Pécs, Department of Psychology

***APPLICATION OF MACHINE
LEARNING IN THE STUDY OF
SUBCLINICAL ADHD***

Consultant:

Dr. András Matuz

By:

Pegah Sadat Hemmati Hosseini

Pécs, Summer 2024

Attention Deficit Hyperactivity Disorder.....	4
History.....	4
Who is an “ADHD” person?.....	5
Diagnostic criteria for ADHD.....	6
Treatment and coping with ADHD.....	6
Subclinical ADHD.....	7
What is Subclinical ADHD?.....	7
Prevalence and Diagnostic Challenges.....	8
Consequences of Subclinical ADHD.....	8
Differences Between ADHD and Subclinical ADHD.....	9
Long-Term Considerations and Interventions.....	10
Challenges in Diagnosis.....	10
Definition of Subclinical ADHD in Our Study.....	11
Machine learning applications in psychology.....	11
What is machine learning?.....	11
Supervised machine learning.....	12
Traditional statistical analysis vs. Machine learning.....	12
Methodology.....	14
Process of data collection.....	14
Questionnaires.....	16
ASRS.....	16
DASS-21.....	16
CAGE.....	17
PIUQ.....	18
Algorithms used in machine learning.....	18
Evaluating the performance of a model.....	19
Data analysis.....	21
Results.....	22
Discussion.....	25
Limitations.....	28
References.....	29

Attention Deficit Hyperactivity Disorder

History

Records of characteristics resembling what we now recognize as ADHD dates back to almost 200 years ago. In 1902, George Still documented 43 children grappling with significant challenges in maintaining attention and adhering to moral standards in regulating their behavior. In doing so, he emphasized the importance of considering the consequences of one's actions for both oneself and others, cautioning against behaviors like “lawlessness”, spitefulness, and dishonesty within his clinical observations (Barkley et al., 2008). In 1930 this condition was known by various terminologies, such as "Minimal Brain Dysfunction" and "Hyperactive/Hyperkinetic Syndrome." Historical records from the 1930s indicate that physicians observed a paradoxical calming effect of stimulant medications, typically associated with increased activity, in children exhibiting hyperactivity and presumed to have "brain damage." However, these stimulant treatments were not employed for children with genetically rooted neurodevelopmental disorders like Down syndrome (Barkley et al., 2008; Mahone & Denckla, 2017). The term “ADHD” was coined in 1980, and its diagnostic criteria have evolved since then (Mahone & Denckla, 2017). It was initially perceived as a childhood disorder and its symptoms were believed to gradually disappear by adolescence; however, ADHD is now recognized as the most prevalent neurodevelopmental disorder. It affects approximately 5% of the school-aged population globally (Antshel et al., 2011; Cortese & Coghill, 2018; Mahone & Denckla, 2017), persisting into adulthood in 65% of cases (Bernardi et al., 2012). This shift in understanding has brought significant attention to the topic over the past 35 years, especially with advancements in neuroimaging techniques translating findings into clinical practice (Cortese & Coghill, 2018). The evolving epidemiology of ADHD highlights the importance of identifying its causes and impacts. In this paper our focus will remain on ADHD in adults.

Who is an “ADHD” person?

Core symptoms associated with ADHD are mostly known as inattentiveness, hyperactivity, and impulsivity (Anbarasan et al., 2020; Antshel et al., 2011; Barkley et al., 2008; Furman, 2005). These symptoms evolve over the lifespan of an individual diagnosed with ADHD. While inattentiveness tends to remain relatively stable; the capacity to concentrate on areas of interest shows variability, often improving but still retaining a degree of unpredictability (Ginapp et al., 2022). Hyperactivity in ADHD adults is characterized by persistent restlessness and a difficulty in relaxation, whereas impulsivity is associated with engaging in behaviors that carry significant risks, such as impulsive speech, reckless driving, and problematic gambling, etc (Bernardi et al., 2012; Ginapp et al., 2022; Kosheleff et al., 2023).

It is important to note that the presence of these symptoms alone does not constitute the primary complaint or sole criteria for diagnosis. The key factor prompting treatment is usually some type of functional impairment in various life domains (Kosheleff et al., 2023). Executive dysfunction, which can arise as a result of ADHD's core symptoms, manifest in challenges related to time management, organization, and task execution. These issues are common across all age groups but are especially problematic in adulthood, as individuals transition from a structured upbringing to greater self-dependence. This shift can significantly affect their studies, employment, and interpersonal relationships (Kosheleff et al., 2023).

Studies suggest that adults with ADHD often experience internal feelings of restlessness and exhibit external signs of disorganized lifestyles and turbulent interpersonal relationships (Ginapp et al., 2022). Additionally, these individuals are more likely to face adverse outcomes, including traumatic experiences, and are characterized by lower perceived health, diminished social support, higher stress levels, and an increased risk of accidents and significant life events. This increased vulnerability may partly explain the higher incidence of PTSD in this population (Bernardi et al., 2012). In comparison with adults without ADHD, those with the condition report a lower quality of life and suffer from anxiety, depression, and daytime sleepiness (Agarwal et al., 2012). Adults with a lifelong history of ADHD are found to have a higher incidence of psychiatric issues, poorer mental health overall, and lower life satisfaction (Weissenberger et al., 2018).

the wider symptomatology of ADHD, where co-occurring internal disorders amplify cognitive and emotional regulation difficulties (Balázs & Keresztény, 2014; Kóbor et al., 2012; Overbey et al., 2011). Large body of research has found the comorbidities of ADHD with depression and anxiety, these comorbidities often complicate the diagnosis and treatment of ADHD, as symptoms of depression and anxiety can overlap with or mask the symptoms of ADHD; making them crucial part of management of the syndrome and assessment of their broader clinical needs and helping them to improve the quality of life (Agarwal et al., 2012; Antony et al., 2022; Coutinho et al., 2021; Nemesure et al., 2021; Sevincok et al., 2020). This combination often results in a greater impairment in social, academic, and occupational functioning.

Diagnostic criteria for ADHD

In the transition from DSM-4 to DSM-5, the criteria for diagnosing ADHD underwent significant changes (Cortese & Coghill, 2018). Initially, the DSM-4 required the presence of either six symptoms of inattention or six symptoms of hyperactivity-impulsivity, with these symptoms needing to be evident for at least six months. Additionally, it was necessary for the symptoms to cause impairment in more than one setting, such as at home and school, and they had to appear before the child turned seven years old. However, recognizing that these criteria might be too stringent (Faraone et al., 2006), the DSM-5 introduced some adjustments. For individuals over 17, the threshold for symptoms was lowered to five instead of six. This change indicates that the manifestations of ADHD can be different in adults compared to children. Also, the age by which symptoms must first appear was raised from 7 to 12 years old. This shift allows for the diagnosis of ADHD in individuals who may not have shown clear symptoms until later in childhood, also accommodating the possibility of a retrospective diagnosis when symptoms become more apparent as responsibilities increase with age.

Treatment and coping with ADHD

It has been found that less than half of individuals with ADHD have ever sought treatment, typically in late adolescence or early adulthood (Bernardi et al., 2012). Furthermore,

there is a notable gap in treatment provision, with only about a quarter of those seeking treatment actually receiving it (Kessler et al., 2006; Merikangas et al., 2010).

Research also indicates that the lifetime treatment rates for ADHD are only 44.02%, and the average age of first treatment is 18.4 years, suggesting a significant undertreatment issue (Bernardi et al., 2012). Individuals living with undiagnosed ADHD often develop their own coping mechanisms, which include structured routines, maintaining to-do lists, using reminder apps, adjusting their environments to minimize distractions, or opting for an unstructured lifestyle and working from home (Ginapp et al., 2022). An interesting aspect of self-medication among individuals with ADHD is the use of cigarette smoking, which has been found to have positive effects on ADHD symptoms due to the effects of nicotine on attentiveness and alertness (Weissenberger et al., 2018).

The benefits of receiving treatment for ADHD are significant. Improvements have been observed in various aspects of life, including vitality, social, emotional, and physical functioning, which contribute to enhanced overall physical and mental health (Agarwal et al., 2012). Additionally, adults who were diagnosed reported experiencing relief, reduced self-criticism, and a better understanding of their symptoms, leading to less shame about their perceived shortcomings (Aoki et al., 2020; Ginapp et al., 2022; Toner et al., 2006).

Subclinical ADHD

What is Subclinical ADHD?

Subclinical ADHD, also known as subthreshold ADHD, is a condition where individuals exhibit ADHD-like symptoms but not to the extent required for a formal ADHD diagnosis. These symptoms often include difficulties with attention, working memory, self-regulation, and executive functions. While less severe than full ADHD syndrome, subclinical ADHD can still significantly impact quality of life, leading to lower life satisfaction and functional impairments in educational, interpersonal, and day-to-day activities (Brown & Casey, 2016; Overbey et al., 2011).

Prevalence and Diagnostic Challenges

The prevalence rate for subclinical ADHD is reported to be between 0.8% and 23.1% (Balázs & Keresztény, 2014). This wide range is due to the lack of consensus on the definition of “subclinical ADHD” within the scientific community. There are no definite diagnostic criteria for subclinical ADHD; a review study reported that most studies regarding subclinical ADHD required a minimum number of symptoms of inattention and/or hyperactivity, but not the full symptom criteria outlined in classification systems. Some studies required additional criteria, such as age of onset or functional impairment, while others adopted a more stringent approach, requiring full symptom criteria for ADHD without the presence of functional impairment. In some cases, subclinical ADHD meant that all criteria were met according to diagnostic manuals, but at least one criterion was labeled as “not certain” by clinicians. The commonality across these definitions is that individuals exhibit symptoms associated with ADHD and experience functional impairments to some degree (Balázs & Keresztény, 2014).

Consequences of Subclinical ADHD

Studies have shown that individuals who do not meet the cut-off criteria for a full ADHD diagnosis still face significant challenges. For example, a study conducted on over 900 school-aged Korean children in 2014 found that those who did not meet the diagnostic criteria for ADHD still suffered severe consequences, such as reduced attentional control, poorer behavioral and emotional regulation, and lower academic performance, albeit to a lesser extent than children who met the full DSM-4 criteria for ADHD (Hong et al., 2014). These individuals also showed lower interpersonal skills and could be wrongly perceived as having behavioral traits rather than a medical condition. This often leads to self-blame and a false self-concept.

Studies have shown children with subclinical ADHD perform lower in academic, functional, and interpersonal aspects of their lives compared to their non-ADHD peers (Balázs & Keresztény, 2014; Boot et al., 2017; Brown & Casey, 2016; Gritti et al., 2023). Such patterns could be predictors of late-onset full ADHD syndrome, substance abuse disorders, internalizing disorders, antisocial personality, and conduct disorders in later stages of life (Balázs & Keresztény, 2014; Kóbor et al., 2012; Overbey et al., 2011).

Despite reports indicating that individuals with ADHD often experience a lower quality of life (Agarwal et al., 2012), a study involving 127 Israeli adults with borderline intelligence and diagnosed with ADHD presents a more optimistic view. This research highlights that access to inclusive education, limited medical disabilities, lower scores on the Brown ADHD Symptom Scale, higher monthly income, active participation in leisure activities, and having a personal friend are all associated with higher quality of life scores (Rimmerman et al., 2007). These findings suggest that early diagnosis, coupled with efforts to provide inclusivity and accessibility, can significantly enhance the quality of life for individuals with ADHD. This study not only sheds light on potential interventions but also underscores the importance of a supportive and accommodating environment in improving life experiences for those with ADHD.

Differences Between ADHD and Subclinical ADHD

A study conducted in 2006 compared full syndrome ADHD with late-onset ADHD (patients who did not meet the age of onset criteria in retrospective clinical inspections) and individuals with subclinical ADHD. It found that subclinical ADHD patients exhibited milder symptoms of ADHD and had a different pattern of familial transmission. Compared to controls, individuals with subclinical ADHD scored one-third to half a standard deviation lower on the Wide Range Achievement Test for arithmetic and verbal IQ. They also exhibited modest psychiatric comorbidity (Faraone et al., 2006).

Comparative studies on subclinical ADHD reported that the condition is more prevalent in families of lower socioeconomic status and is associated with different biological patterns of transmission, suggesting that subclinical ADHD might result from environmental factors (Hong et al., 2014). Emphasizing early intervention and special education can improve life satisfaction and quality of life in the long run, decreasing the economic burden on the individual and society (Agarwal et al., 2012; Brown & Casey, 2016; Hong et al., 2014).

Long-Term Considerations and Interventions

Longitudinal studies on children clinically diagnosed with ADHD who no longer meet the criteria as adults report that they still struggle with subclinical ADHD. In six years, one-third and in ten years, around one-fifth of them developed subclinical ADHD, indicating the importance of reassessment and screening for subclinical ADHD (Balázs & Keresztény, 2014; Biederman et al., 2010; Hong et al., 2014).

Effective interventions for subclinical ADHD often include behavioral interventions such as behavior modification techniques, cognitive-behavioral therapy, and social skills training, which focus on modifying specific problematic behaviors and developing coping strategies. Educational support and accommodations are particularly crucial for children and adolescents, involving individualized education plans or classroom adjustments to cater to specific learning and attention challenges. Psychological counseling and psychotherapy play a significant role, helping individuals understand their behaviors and develop strategies to cope with daily challenges, including stress management and organizational skills training, alongside therapy for any co-occurring issues like anxiety or depression. Family involvement is also key, with family support and education becoming integral. Family therapy and parent training programs educate family members about the condition and teach them effective support strategies.

Challenges in Diagnosis

The challenges in diagnosing subclinical ADHD include the lack of common diagnostic criteria for the condition, with no consensus in the scientific community on a definitive definition of "subclinical ADHD". Additionally, while diagnostic manuals typically rely on a categorical classification for ADHD, many clinicians and researchers suggest a continuum model for understanding ADHD and its subclinical diagnosis (Antshel et al., 2011; Boot et al., 2017; Hong et al., 2014; Weissenberger et al., 2018). This model recognizes the varying intensities and forms of ADHD symptoms across individuals. Establishing a clear definition for subthreshold ADHD is complex, as it seems to be dimensionally distributed in the population. The exact number of symptoms required for a diagnosis of subthreshold ADHD remains unclear, and the impact of subthreshold ADHD does not necessarily depend on the symptom count.

Definition of Subclinical ADHD in Our Study

Despite the various definitions of subclinical ADHD, in our study, we use the term "subclinical ADHD" for individuals who meet the pattern of the prescreening ASRS questionnaire. This approach allows us to identify individuals who exhibit significant symptoms and functional impairments associated with ADHD, even if they do not meet the full diagnostic criteria for the disorder.

Machine learning applications in psychology

What is machine learning?

A subfield of artificial intelligence known as "machine learning" uses a variety of techniques to automatically identify patterns in data sets. Based on these newly discovered patterns, the system may subsequently predict a target within new data sets. Machine Learning (ML) utilizes complex statistical and probabilistic methods to develop systems that can autonomously learn knowledge from data. This makes it possible to more quickly and precisely spot patterns; highlight the most important variables in data and to draw more accurate conclusions from data sources. ML has yielded substantial advantages across various domains such as artificial intelligence, computer vision, speech recognition, and natural language processing. It enables researchers and developers to extract crucial information from datasets, deliver personalized experiences, and create intelligent systems (Jordan & Mitchell, 2015). The overall process typically consists of two main phases: training and testing. During the training phase, the algorithm is exposed to a portion of the dataset, learning from the examples by adjusting its parameters to minimize errors in its predictions. This phase is crucial as it forms the foundation of the algorithm's ability to make accurate predictions. After the training phase, the performance of the algorithm is evaluated using the testing phase. This involves a subset of the data that was not used during training, often referred to as the "test set." This separation is essential to objectively assess the algorithm's performance, ensuring that it can generalize its learning to new data, rather than just memorizing the training set.

Supervised machine learning

Supervised machine learning is a method of training algorithms, where the process involves feeding the algorithm a large dataset that is already labeled. This means that for each piece of data (or instance) in the training set, the desired output is known. These labels help the algorithm understand patterns and relationships within the data, enabling it to make predictions or decisions when presented with new, unseen data.

The efficiency and accuracy of a supervised machine learning algorithm depend heavily on the quality and quantity of the labeled data provided. More comprehensive and diverse datasets typically lead to more robust and reliable models.

Applications of supervised ML algorithms in mental health and psychology are useful in several domains, such as detection and diagnosis of mental health issues, prognosis and treatment, public health and research. ML algorithms have been already used in detection of depression and anxiety (Nemesure et al., 2021; Nickson et al., 2023), alzheimer's (Diogo et al., 2022) and suicide prevention(Jung et al., 2019) and ADHD diagnosis (Cao et al., 2023; Chen et al., 2021; Rosenbusch et al., 2021)

Traditional statistical analysis vs. Machine learning

In the context of psychology, the distinction between traditional statistical methods and machine learning lies primarily in their objectives and approaches to data. Traditional statistics are geared towards understanding and modeling the underlying processes of psychological phenomena, often focusing on hypothesis testing and inference. This involves creating specific models that are interpretative and explanatory, helping in understanding the 'why' behind observed data. These methods are particularly useful when the goal is to test theories or establish causality, and they work well with structured and controlled experimental data.

Machine learning, on the other hand, excels in predictive accuracy and handling complex, high-dimensional datasets where the relationships among variables might be nonlinear or too intricate for traditional models. It is less concerned with the underlying theory and more with identifying patterns that can predict outcomes. Machine learning is particularly beneficial in scenarios where the primary goal is prediction, such as forecasting behavior or identifying at-risk individuals, and in cases where the data are too vast or complex for traditional statistical methods to handle effectively. However, machine learning models can sometimes be not that clear, offering less in the way of interpretability. In psychology, where both understanding and prediction are valuable, combining traditional statistical methods with machine learning can often yield the most comprehensive insights (Bzdok et al., 2018; Orrù et al., 2020).

The integration of machine learning in diverse fields offers a promising path for identifying key factors, particularly in developing our understanding of ADHD and its subclinical implications. This exploratory study is focused on understanding the primary predictors of subclinical ADHD with applications of machine learning; therefore, we have compiled a series of psychometric questionnaires, which, in prior research, have been associated with comorbid conditions in ADHD. In particular, ADHD symptoms are linked with internalizing disorders such as anxiety and depression. Existing literature suggests a frequent co-occurrence of ADHD with these disorders, thereby complicating their management and treatment (Antony et al., 2022; Sevincok et al., 2020).

Furthermore, individuals with ADHD exhibit a significantly elevated risk of developing substance use disorders (SUD) compared to their non-ADHD peers. This increased risk is partly due to impulsivity, a common symptom of ADHD, which may lead to heightened levels of drug usage. There is also an indication of a biologically increased susceptibility to substance dependence in ADHD individuals. Research has found that adolescents with ADHD are more likely to develop SUDs in their teenage years, and adults with a history of childhood ADHD show a higher propensity for drug and alcohol dependence, though the intricacies of this relationship warrant further examination (Harstad et al., 2014; Ivanov et al., 2022; Levy et al., 2014; Magon & Müller, 2012). Additionally, there is an observed correlation between ADHD-related symptoms and a heightened tendency towards problematic internet use (PIU). This relationship suggests that the defining characteristics of ADHD, including challenges in

attention regulation and impulsivity, may predispose individuals to problematic behaviors in internet usage (Augner et al., 2023; Overbey et al., 2011).

Methodology

Process of data collection

The recruitment criteria required all participants to be aged 18 or above. Participants received a brief overview of the study's purpose and provided their informed consent before participating. The data collection adhered to standard ethical considerations for research, ensuring anonymity throughout the process.

Training Phase:

As part of a larger protocol, Data was collected from 1,000 Hungarian participants (500 males, 500 females; mean age = 41.75 years, SD = 11.7, min 18 and max 65).

Evaluation Phase:

The evaluation phase of the process was conducted in three different ways:

1. First Phase: Evaluation was conducted on 30% of the large dataset collected during the training phase.

2. Second Phase: Data was collected from a smaller Hungarian population with the aim to test the models in a young sample. This sample included 62 participants (52 females, 10 males; mean age = 22.87 years, SD = 3.62, min 18 and max 30).

3. Third Phase: To evaluate the model across cultures, data was collected from an Iranian population. This sample consisted of 70 participants (47 females, 23 males; mean age = 30.69 years, SD = 9.96, min 19 and max 68).

Once consent was obtained, participants proceeded to answer an extensive series of demographic questions collecting basic personal and socio-economic information from participants, including gender, age, living situation, family background, marital status, education, employment details, and health conditions, types of electronic devices used, purpose of internet usage (for shopping, social media, gaming, pornographic content, etc) , time of online activities (including how many hours per day, time frames of the day being online), etc.

This initial section was followed by a selection of relevant psychometric questionnaires designed to assess various aspects of mental health and behavior. These included the Problematic Internet Use Questionnaire (PIUQ), the Depression Anxiety Stress Scale (DASS-21), the CAGE Questionnaire for Substance Use, and the Adult ADHD Self-Report Scale (ASRS). For detailed information about the internal consistency of each questionnaire, see Table 1.

Table 1. Reliabilities of the psychometric questionnaires

Subscales of questionnaires	Hungarian Train (ω)	Hungarian Test (ω)	Iranian Test (ω)
ASRS A	0.87	0.74	0.79
ASRS B	0.92	0.85	0.85
CAGE alcohol	0.81	0.64	0.54
CAGE sedatives	0.82	0.84	0.68
CAGE stimulants	0.81	0.54	0.83
CAGE smoking	0.80	0.90	0.87
DASS anxiety	0.90	0.80	0.80
DASS depression	0.93	0.88	0.90
DASS stress	0.89	0.83	0.81

PIQU control	0.86	0.81	0.80
PIQU obsession	0.92	0.83	0.70
PIQU neglect	0.86	0.81	0.80

Questionnaires

ASRS

Adult ADHD Self Report Scale, ASRS, is a tool developed for identifying adult attention-deficit/hyperactivity disorder (ADHD) in the general population.

The ASRS includes 18 questions, designed to assess how frequently an adult experiences symptoms that align with the DSM-IV criteria for ADHD. These questions are rated on a scale from 1 to 5, where 1 indicates "Never" and 5 indicates "Very Often". Among these 18 questions, six were identified as the most indicative of ADHD symptoms. These six questions form the core of the ASRS v1.1 Screener and constitute Part A of the Symptom Checklist. The remaining twelve questions are included in Part B of the Symptom Checklist.

A shorter version of the scale consists of only six questions. These six were chosen because they best align with clinical diagnoses of ADHD. The study found that the shorter, six-question version of the ASRS was more effective than the full 18-question version in several ways. It was better in terms of sensitivity (identifying those with the condition), specificity (excluding those without the condition), overall accuracy in classification, and in the kappa statistic, a measure of agreement between the ASRS and clinical diagnoses (Kessler et al., 2005, 2007).

DASS-21

The DASS-21, or the Depression Anxiety Stress Scales, is a shortened psychological assessment tool specifically designed to measure the levels of depression, anxiety, and stress in individuals. It's a shortened version of the original DASS, containing 21 questions instead of 42, divided equally, 7 questions for each subscale, to address three key emotional states.

In this scale, participants respond to questions that help assess their emotional state over the past week. Each item is rated on a scale from 0 to 3, where 0 means the statement did not apply at all, and 3 indicates it applied very much or most of the time. DASS-21 is valued for its conciseness, reliability, and ease of use. It's not just limited to clinical settings for patient evaluation but is also widely employed in research to understand how groups of people may be experiencing these common emotional states (Norton, 2007).

The “Depression” subscale of the questionnaire looks at mood states like hopelessness, a reduced sense of personal worth, and a lack of interest in activities that normally bring joy, the “Anxiety” subscale measures physical symptoms associated with anxiety, such as trembling and sweating, as well as psychological symptoms like feelings of nervousness and apprehension and the “Stress” subscale, on the other hand, evaluates more prolonged feelings of tension and strain, assessing how easily a person gets agitated or how difficult it is for them to relax.

CAGE

The CAGE questionnaire is a brief and widely used screening tool for detecting alcohol abuse and dependence. Each of the questions requires a simple 'yes' or 'no' answer, and the total score ranges from 0 to 4. The questionnaire is concise, taking about 30 seconds to administer, and is designed to be non-intimidating (Ewing, 1984).

It has been validated as a tool for detecting alcohol abuse and dependence in various patient populations, including medical and surgical inpatients, ambulatory medical patients, and psychiatric inpatients. However, its performance has varied in primary care patients, and it has not been effective in specific groups like white women, prenatal women, and college students. It is also not recommended as a screening test for less severe forms of drinking (Dhalla & Kopec, 2007).

Since then CAGE questionnaire has been adapted for detecting nicotine dependence (Rustin, 2000) and substance use for elderly population (Hinkin et al., 2001) inspired by this approach, we adapted the CAGE questionnaire for Stimulants and Sedatives as well.

PIUQ

The Problematic Internet Use Questionnaire (PIUQ) is a self-report survey designed to measure problematic or potentially addictive internet use. This 18-item questionnaire assesses aspects of internet usage that may be harmful or indicative of addiction. Each item on the PIUQ is rated on a 5-point scale, ranging from “never” (1) to “always” (5), allowing respondents to self-rate the extent of their internet-related issues. The PIUQ divides these items into three 6-item subscales, each targeting a specific dimension of problematic internet use:

“Obsession” Subscale measures psychological fixation with the internet, including worries, anxiety, and depression related to not being able to access the internet, while “Neglect” Subscale focuses on the neglect of typical daily activities in favor of internet use, such as neglecting work, sleep, or spending time with friends or family and “Control” Disorder Subscale assesses the inability to control or decrease the amount of time spent online, including awareness of excessive internet use being problematic, but an inability to reduce it (Demetrovics et al., 2008).

Algorithms used in machine learning

Support Vector Machines (SVM), Decision Trees (DT), and Random Forests (RF) are three commonly used algorithms in machine learning, each with its unique characteristics and applications (Alpaydin, 2020).

Support Vector Machines (SVM):

Support Vector Machines, or SVMs, are powerful algorithms used for classification and regression tasks. SVM works by finding the optimal hyperplane that separates different classes in the data space. It aims to maximize the margin between the classes, which helps in achieving better generalization to unseen data. In simpler terms, imagine your data points as dots on a graph, and SVM draws a line (or a hyperplane in higher dimensions) to separate these dots into different groups. It does this in a way that maximizes the space between the groups, making it easier to classify new points correctly.

Decision Trees (DT):

Decision Trees are intuitive and easy-to-understand algorithms used for classification and regression tasks. They work by recursively partitioning the data into smaller subsets based on the features that best split the data into distinct classes or categories. In other words, think of a decision tree as a flowchart where each internal node represents a "decision" based on a feature, each branch represents the outcome of that decision, and each leaf node represents the final decision or classification.

Random Forests (RF):

Random Forests are an ensemble learning method that combines multiple decision trees to improve predictive performance and reduce overfitting. Each decision tree in the forest is trained on a random subset of the training data and a random subset of features. When making predictions, each tree in the forest individually predicts the target variable, and the final prediction is determined by averaging or voting across all trees.

Evaluating the performance of a model

Evaluating the performance of a machine learning model is crucial for understanding its effectiveness and reliability. Different metrics provide insights into various aspects of the model's performance. Here are some commonly used metrics and how they are interpreted:

Accuracy

Accuracy is the ratio of correctly predicted instances to the total instances in the dataset. It is a straightforward and widely used metric but can be misleading if the dataset is imbalanced (i.e., when one class is much more frequent than others).

ROC Curve and AUC

ROC (Receiver Operating Characteristic) Curve is a graphical representation of a classifier's performance across different threshold settings. It plots the True Positive Rate (Sensitivity) against the False Positive Rate (1 - Specificity).

- True Positive Rate (Sensitivity, Recall): The proportion of actual positives correctly identified by the model.

- False Positive Rate: The proportion of actual negatives that are incorrectly identified as positives by the model.

AUC (Area Under the ROC Curve) measures the entire two-dimensional area underneath the ROC curve. AUC ranges from 0 to 1, with higher values indicating better performance. An AUC of 0.5 suggests no discrimination (i.e., random guessing), while an AUC of 1.0 indicates perfect discrimination.

Sensitivity (Recall)

Sensitivity (Recall) measures the proportion of actual positives that are correctly identified by the model. It is crucial in situations where identifying positive instances is more important than identifying negative ones (e.g., in medical diagnoses).

Specificity

Specificity measures the proportion of actual negatives that are correctly identified by the model. It is important when the cost of false positives is high.

To sum up, each of the metrics presented above provides different insights into the models' performance, and the choice of which metrics to use depends on the specific context and goals of the machine learning application (Hand, 2012).

Data analysis

All programming was implemented using Python programming language and libraries such as pandas, scikit-learn (Pedregosa et al., 2011) and numpy were utilized for data manipulation, machine learning processes, and numerical computations. The focus was on implementing a methodology that involved data preprocessing, feature selection, model training, and evaluation to accurately classify subjects who met subclinical ADHD criteria. While splitting the data into a train set with 70% and a test (first phase of evaluation) with 30% of the data, we employed stratification, to maintain the proportion of the classes in both training and test sets, due to the imbalance in the subclinical ADHD and non-ADHD classes within the data. Stratification ensures that both sets are representative of the overall sample, thereby enhancing the reliability of our model evaluations.

A multi-step feature selection was employed, consisting of an initial prescreening of the features applying chi-square tests for categorical variables and point biserial correlation tests for continuous variables to identify significant features based on a p-value threshold of 0.05 ensuring that only features with a statistically significant relationship with the target variable were used for further analysis. Following the prescreening; feature selection via recursive feature elimination with five-fold cross-validation (5-CV) was performed on the training set. Three classification algorithms were used: support vector machine, random forest, and decision tree classifier. Before the training, the data were standardized using Z-transformation (except for the training of random forest and decision tree classifiers). Hyperparameters for each classifier were optimized through grid search with 5-CV. Both the internal validation and the evaluation of the classification performance on the test set were done using the area under the receiver operating characteristic curve (area under the curve [AUC]). We also calculated accuracy, sensitivity and specificity to evaluate the performance.

We introduced the new dataset to the trained model once for the smaller Hungarian population and again on the Persian population, assessing the model's performance on these new datasets. In the second phase of evaluation, the trained model was introduced to a smaller Hungarian population to assess performance on a dataset with a lower mean age. In the third phase, the model's cross-cultural applicability was evaluated using data from an Iranian population.

Results

Feature selection

The 'Depression (DASS)' feature from the DASS-21 questionnaire had the highest mean importance of 0.208 and was selected 100 times. Similarly, 'Anxiety (DASS)' had a mean importance of 0.1459 and was also chosen 100 times. 'Stress (DASS)' followed with a mean importance of 0.1523 and was selected 97 times.

From the PIUQ, 'Neglect' had a mean importance of 0.144 and was selected 100 times, 'Control' had a mean importance of 0.091 with 94 selections, and 'Obsession' had a mean importance of 0.068 with 89 selections.

The demographic feature 'Age' showed a mean importance of 0.034 and was selected 84 times. 'Years with a digital device' had a mean importance of 0.032 and was chosen 71 times. 'Alcohol (CAGE)' had a mean importance of 0.029 and was selected 77 times. Lastly, 'Depression (self report)' had a mean importance of 0.026 and was selected 78 times (figure 1).

The tables 2, 3 and 4 include the top selected features and their statistical summaries for two groups: individuals with subclinical ADHD and those without (non-subclinical ADHD).

Table 2. Training data (Subclinical ADHD (n = 103), Non-subclinical ADHD (n = 897))

Top selected features	Subclinical ADHD	Non subclinical ADHD
Numerical features (mean (SD))		
DASS Depression	10.583 (5.815)	4.434 (4.737)
DASS Anxiety	7.883 (5.275)	3.070 (3.837)
DASS Stress	10.350 (5.156)	4.929 (4.235)
PIUQ Obsession	14.612 (6.961)	9.562 (4.529)

PIUQ Neglect	16.204 (5.593)	10.775 (4.228)
PIUQ Control	15.961 (5.625)	11.164 (4.313)
Age	37.612 (11.524)	42.225 (11.605)
Age of first usage of electronic device	17.680 (9.543)	20.928 (10.499)
Categorical features (Mode (frequency of the mode in %))		
Region	Village (24%)	Town (36%)
Housing	House (37%)	House (59%)
Uses internet for surfing	Yes (65%)	Yes (66%)

Table 3. Iranian testing data (Subclinical ADHD (n = 26), Non-subclinical ADHD (n = 44))

Top selected features	Subclinical ADHD	Non subclinical ADHD
Numerical features (mean (SD))		
DASS Depression	9.962 (5.318)	4.545 (3.592)
DASS Anxiety	6.462 (3.603)	4.045 (3.263)
DASS Stress	10.692 (3.824)	6.386 (3.452)

PIUQ Obsession	12.962 (3.013)	12.159 (3.679)
PIUQ Neglect	15.923 (3.599)	13.364 (3.141)
PIUQ Control	18.077 (3.989)	15.136 (4.311)
Age	27.115 (4.999)	32.795 (11.504)
Age of first usage of electronic device	12.346 (3.633)	14.841 (5.507)

**Categorical features (Mode
(frequency of the mode in %))**

Region	City (56%)	City (47%)
Housing	Rent (38%)	Apartment (49%)
Uses internet for surfing	No (65%)	No (57%)

Table 4. Hungarian testing data (Subclinical ADHD (n = 25), Non-subclinical ADHD (n = 37))

Top selected features	Subclinical ADHD	Non subclinical ADHD
Numerical features (mean (SD))		
DASS Depression	8.520 (6.063)	4.432 (3.387)
DASS Anxiety	8.400 (4.673)	4.270 (3.330)

DASS Stress	11.920 (4.092)	6.649 (4.131)
PIUQ Obsession	10.400 (3.266)	9.405 (3.855)
PIUQ Neglect	13.040 (2.622)	11.387(2.670)
PIUQ Control	15.560 (4.104)	12.270 (3.687)
Age	22.000 (3.175)	23.459 (3.820)
Age of first usage of electronic device	11.040 (2.263)	10.243 (2.938)

**Categorical features (Mode
(frequency of the mode in %))**

Region	Capital (28%)	City (57%)
Housing	Apartment (67%)	House (52%)
Uses internet for surfing	Yes (31%)	No (40%)

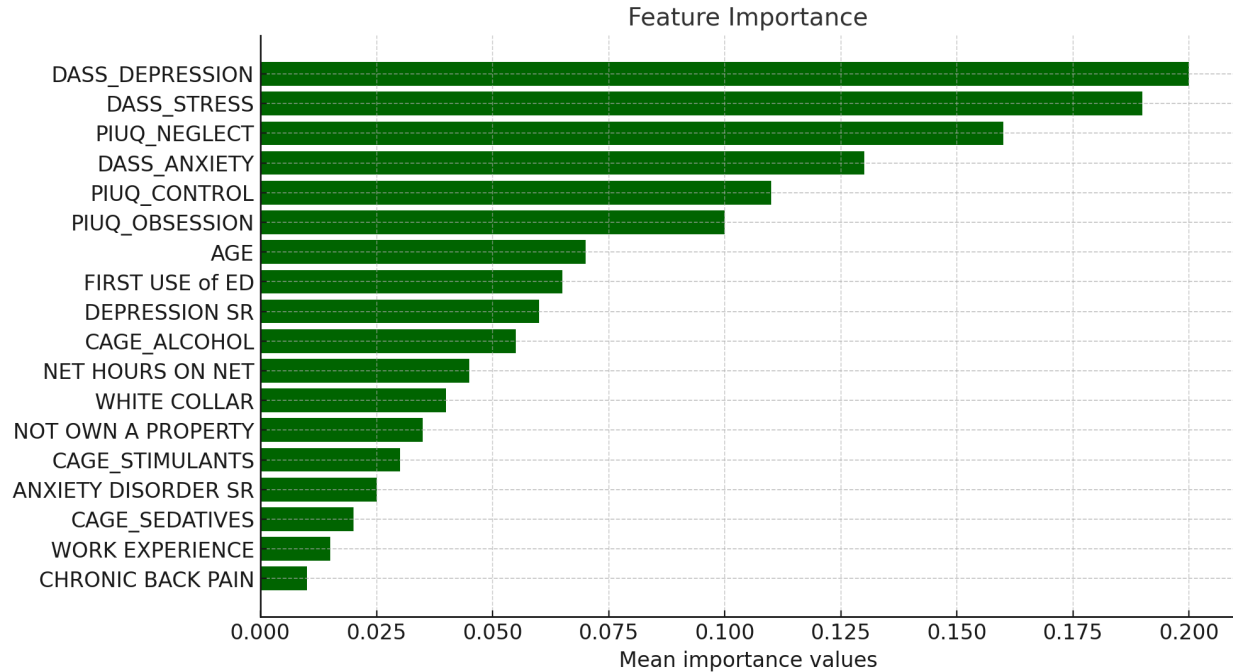


Figure 1: Importance values of the most frequently selected features for regression. (ED=Electronic Device, SR=Self Report).

The aim of the machine learning analyses was to train classification models that successfully differentiate between subclinical ADHD participants and non ADHD ones.

The analysis of the Decision Tree (DT), Random Forest (RF), and Support Vector Machine (SVM) classifiers on data from all the phases of data collection, with a focus on the ROC-AUC scores, reveals distinct levels of performance across these models.

Table 5 summarizes the model performances on the data ($n = 1000$) collected during the training phase. The DT classifier achieved an AUC score of approximately 0.761. This suggests a moderate ability to differentiate between subclinical ADHD and non-ADHD classifications, indicating a fair level of effectiveness but potentially less suitable for more complex classification tasks. In contrast, the RF classifier shows a stronger performance and an AUC of around 0.818.

Outperforming both DT and RF, the SVM classifier achieved the highest AUC score of approximately 0.831 and accuracy of 0.736. This superior performance highlights the SVM

model's robustness in handling the complexities of this data. Its higher discriminative power makes it particularly effective in medical or psychological diagnostics, where accurately distinguishing between different conditions is critical.

Table 5. Performance of each algorithm on the training phase (n = 1000)

	Metric	Random Forest	Decision Tree	Support Vector Machine
0	Accuracy	0.763333333	0.713333333	0.736666667
1	ROC	0.817604029	0.774733181	0.831634489
2	F1	0.382608696	0.358208955	0.357723577
3	Sensitivity	0.709677419	0.774193548	0.709677419
4	Specificity	0.769516729	0.706319703	0.739776952

Table 6 and 7 summarize the model performances on the younger Hungarian sample (i.e. second phase) and the Iranian sample (i.e. third phase), respectively. In the second and third phases of evaluation, our model again performed very well in classifying ADHD and non-ADHD individuals, indicating the robustness of the trained model. This performance was consistent both on new Hungarian data with AUC of 0.811 (Table 6) and accuracy of 0.741 and across different cultures with AUC of 0.821 and accuracy of 0.742 (Table 7).

Table 6. performance of each algorithm on the second phase (younger Hungarian sample, n = 62)

	Metric	Random Forest	Decision Tree	Support Vector Machine
0	Accuracy	0.661290323	0.725806452	0.741935484
1	ROC	0.774054054	0.800540541	0.811891892
2	F1	0.618181818	0.721311475	0.666666667
3	Sensitivity	0.68	0.88	0.64
4	Specificity	0.648648649	0.621621622	0.810810811

Table 7. Performance of each algorithm on the third phase (Iranian sample, n = 70)

	Metric	Random Forest	Decision Tree	Support Vector Machine
0	Accuracy	0.763333333	0.713333333	0.736666667
1	ROC	0.817604029	0.774733181	0.831634489
2	F1	0.382608696	0.358208955	0.357723577
3	Sensitivity	0.709677419	0.774193548	0.709677419
4	Specificity	0.769516729	0.706319703	0.739776952

Discussion

The primary aims of our research were to explore the predictors of subclinical ADHD and test the model on different sets of data. Our machine learning analysis and its findings align with the existing body of research on the importance of internalizing disorders, such as anxiety and depression, as well as ineffective coping mechanisms and self-medication in individuals with subclinical ADHD. This concurrence with established literature reinforces the validity of our approach and the relevance of these psychological factors in understanding subclinical ADHD. It underscores the necessity of addressing these issues in both clinical practice and future research to mitigate their impact on individuals with subclinical ADHD.

In addition to confirming known factors, our model has successfully identified other factors not previously explored in scientific literature. These novel factors suggest potential avenues for future investigation and offer new insights into the multifaceted nature of subclinical ADHD. The newly identified factors include:

Age of First Exposure to Electronic Devices: Early exposure to electronic devices may have a significant impact on the development of subclinical ADHD, warranting further research into its long-term effects. This finding aligns with growing concerns about the influence of digital technology on mental health (Sohn et al., 2019), emphasizing the need for guidelines regarding the appropriate age and duration of exposure.

Internet Usage Hours: The amount of time spent on the internet appears to correlate with subclinical ADHD symptoms, suggesting that excessive internet use may contribute to the disorder. Problematic internet use, characterized by behaviors like those identified through the PIUQ, might be coping mechanisms or signs of underlying executive dysfunction. This phenomenon is partly attributed to the challenges faced by those with subclinical ADHD in self-regulation, which predisposes them to addictive behaviors, including excessive internet use. Online environments providing immediate feedback, rewards, and easier avenues for social interaction can compensate for real-life challenges, explaining why individuals with subclinical

ADHD are more prone to problematic internet usage compared to their healthy counterparts (Wang et al., 2017).

Housing Situation: The living conditions and the stability of one's housing situation can significantly impact the manifestation of subclinical ADHD symptoms. For instance, individuals in unstable or stressful housing environments may experience heightened anxiety and distraction, exacerbating ADHD-like symptoms. This underscores the importance of considering socio-environmental factors when assessing subclinical ADHD, as a stable and supportive living situation could potentially mitigate some of these symptoms.

Occupation: Employment status and the nature of one's job may influence subclinical ADHD symptoms. For example, high-stress occupations or jobs with low satisfaction could exacerbate symptoms like inattention and restlessness. This is particularly relevant when considering the difference between blue-collar and white-collar jobs. Blue-collar jobs, often involving manual labor and physical tasks, might have different stressors compared to white-collar jobs, which typically involve more cognitive and administrative work.

While a direct cause-effect relationship cannot be determined with our current research design, it's worth exploring whether subclinical ADHD influences various life choices and circumstances or if these circumstances exacerbate subclinical ADHD symptoms. For example, it is important to investigate if subclinical ADHD makes someone more likely to choose certain types of jobs or if the nature of these jobs worsens subclinical ADHD symptoms. Similarly, we should consider whether unstable housing conditions lead to more pronounced ADHD-like symptoms or if individuals with subclinical ADHD are more prone to finding themselves in unstable living situations. This complex interplay between individual characteristics and environmental factors highlights the need for a comprehensive approach to understanding and addressing subclinical ADHD.

Our model, initially trained on a Hungarian sample, demonstrated efficacy when applied to an Iranian sample. This cross-cultural validation underscores the robustness and generalizability of our model across diverse populations. The ability of the model to perform

well in different cultural contexts suggests that the identified factors are not limited to a specific demographic but rather are universally relevant. This cross-cultural success emphasizes its potential as a valuable tool for international research and clinical practice.

Furthermore, this study highlights the importance of establishing a unified definition for subclinical ADHD to facilitate research in both academic and clinical sectors. Our analysis reveals that individuals with subclinical ADHD face substantial challenges, evidenced by indicators such as depression, anxiety, and stress identified through our machine learning approach. These markers are significant on their own and mirror the wider symptomatology of ADHD, where co-occurring internal disorders amplify cognitive and emotional regulation difficulties (Balázs & Keresztény, 2014; Kóbor et al., 2012; Overbey et al., 2011). A large body of research has found the comorbidities of ADHD with depression and anxiety, complicating the diagnosis and treatment of ADHD as symptoms can overlap or mask each other. This overlap makes the management of these comorbidities crucial for the effective treatment of subclinical ADHD, as addressing them can significantly improve the quality of life for affected individuals (Agarwal et al., 2012; Antony et al., 2022; Coutinho et al., 2021; Nemesure et al., 2021; Sevincok et al., 2020).

The study emphasizes that recognizing and treating these co-occurring conditions is critical for managing ADHD effectively. Additionally, behaviors like problematic internet use might be coping mechanisms or signs of underlying executive dysfunction, highlighting the need for early intervention. The growing influence of digital devices, while increasing attentional issues, also opens new possibilities for assisting individuals with subclinical ADHD.

Moreover, individuals with ADHD are often characterized by impulsivity and hyperactivity, increasing the likelihood of engaging in risky behaviors, including substance abuse. The relationship between ADHD and substance use involves not only impulsivity but also significant coping strategies. Many individuals with ADHD struggle with symptoms like inattention, restlessness, and impulsivity, leading some to self-medicate with substances (Ohlmeier et al., 2008). Our study sheds light on alcohol and substance use patterns, similar to those observed in clinical ADHD as per the CAGE questionnaire responses. These patterns may

indicate ineffective coping strategies or self-medication for conditions like depression or anxiety. Such insights call for developing improved strategies to help individuals with subclinical ADHD establish healthier coping mechanisms, significantly enhancing their quality of life.

Early intervention is particularly crucial in educational settings, where it can significantly change outcomes for students with subclinical ADHD. Implementing policies for early detection and targeted support in schools could enhance academic and social success, potentially reducing the disorder's long-term effects. The process of receiving a formal ADHD diagnosis can be a critical step for affected individuals. This diagnosis often brings a significant psychological shift, reducing feelings of self-blame and contributing to a sense of relief. Understanding their condition through a medical lens allows individuals with ADHD to reframe their experiences and challenges, empowering them and serving as a catalyst for lifestyle adaptations and identity development (Ginapp et al., 2022).

Future research is recommended to focus on longitudinal studies to monitor the progression of subclinical ADHD and on developing specific diagnostic criteria. Understanding the natural progression of subclinical ADHD requires research with larger and more culturally diverse sample sizes. Equally important is examining the disorder's impact on personal and social functioning. Enhancing social skills and emotional intelligence is key to improving interpersonal relationships and social integration for those affected by subclinical ADHD, fostering a more inclusive and supportive environment for these individuals.

Limitations

Our model was trained solely using questionnaire data. Future research could enhance the robustness of our findings by incorporating additional types of data, such as biological data including EEG (Adamou et al., 2020), ECG (Khan Mamun & Elfouly, 2023), and others. Integrating diverse data sources would provide a more comprehensive understanding of SCADHD and improve the accuracy and applicability of the model across different contexts and populations.

Conclusions

Our machine learning models, particularly the Support Vector Machine (SVM), demonstrated robust performance in classifying subclinical ADHD across various samples, including younger Hungarian and Iranian populations. The SVM's superior AUC scores and accuracy rates underscore its potential for application in clinical settings, offering a reliable tool for early identification and intervention.

The study's findings align with existing literature on the impact of internalizing disorders and ineffective coping mechanisms in individuals with subclinical ADHD. Additionally, novel factors such as early exposure to electronic devices, internet usage hours, housing situation, and occupation were identified, providing new avenues for research and intervention. These insights highlight the multifaceted nature of subclinical ADHD and the importance of considering both psychological and socio-environmental factors in its assessment and treatment.

Cross-cultural validation of our model indicates its generalizability, suggesting that the identified factors are universally relevant. This underscores the potential for our model to be used in diverse populations, enhancing its applicability in international research and clinical practice.

Furthermore, the study emphasizes the importance of addressing co-occurring conditions like depression, anxiety, and problematic internet use in managing subclinical ADHD. Early intervention, especially in educational settings, is crucial for improving outcomes for individuals with subclinical ADHD. The study also highlights the potential benefits of a formal ADHD diagnosis in providing psychological relief and promoting positive lifestyle adaptations.

References

- Adamou, M., Fullen, T., & Jones, S. L. (2020). EEG for Diagnosis of Adult ADHD: A Systematic Review With Narrative Analysis. *Frontiers in Psychiatry, 11*.
<https://doi.org/10.3389/fpsyt.2020.00871>
- Agarwal, R., Goldenberg, M., Perry, R., & Ishak, W. W. (2012). The Quality of Life of Adults with Attention Deficit Hyperactivity Disorder. *Innovations in Clinical Neuroscience, 9*(5–6), 10–21. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3398685/>
- Alpaydin, E. (2020). *Introduction to machine learning*. MIT Press. (Fourth). MIT Press.
- Anbarasan, D., Kitchin, M., & Adler, L. A. (2020). Screening for Adult ADHD. *Current Psychiatry Reports, 22*(12), 72. <https://doi.org/10.1007/s11920-020-01194-9>
- Antony, E. M.-A., Pihlajamäki, M., Speyer, L. G., & Murray, A. L. (2022). Does emotion dysregulation mediate the association between ADHD symptoms and internalizing problems? A longitudinal within-person analysis in a large population-representative study. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 63*(12), 1583–1590. <https://doi.org/10.1111/jcpp.13624>
- Antshel, K. M., Hargrave, T. M., Simonescu, M., Kaul, P., Hendricks, K., & Faraone, S. V. (2011). Advances in understanding and treating ADHD. *BMC Medicine, 9*(1), 72.
<https://doi.org/10.1186/1741-7015-9-72>
- Aoki, Y., Tsuboi, T., Furuno, T., Watanabe, K., & Kayama, M. (2020). The experiences of receiving a diagnosis of attention deficit hyperactivity disorder during adulthood in Japan: A qualitative study. *BMC Psychiatry, 20*(1), 373.
<https://doi.org/10.1186/s12888-020-02774-y>

- Augner, C., Vlasak, T., & Barth, A. (2023). The relationship between problematic internet use and attention deficit, hyperactivity and impulsivity: A meta-analysis. *Journal of Psychiatric Research*, 168, 1–12. <https://doi.org/10.1016/j.jpsychires.2023.10.032>
- Balázs, J., & Keresztény, Á. (2014). Subthreshold attention deficit hyperactivity in children and adolescents: A systematic review. *European Child & Adolescent Psychiatry*, 23(6), 393–408. <https://doi.org/10.1007/s00787-013-0514-7>
- Barkley, R. A., Murphy, K. R., & Fischer, M. (2008). *ADHD in adults: What the science says*. Guilford Press.
- Bernardi, S., Faraone, S. V., Cortese, S., Kerridge, B. T., Pallanti, S., Wang, S., & Blanco, C. (2012). The lifetime impact of attention deficit hyperactivity disorder: Results from the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). *Psychological Medicine*, 42(4), 875–887. <https://doi.org/10.1017/S003329171100153X>
- Biederman, J., Petty, C. R., Evans, M., Small, J., & Faraone, S. V. (2010). How persistent is ADHD? A controlled 10-year follow-up study of boys with ADHD. *Psychiatry Research*, 177(3), 299–304. <https://doi.org/10.1016/j.psychres.2009.12.010>
- Boot, N., Nevicka, B., & Baas, M. (2017). Subclinical symptoms of attention-deficit/hyperactivity disorder (ADHD) are associated with specific creative processes. *Personality and Individual Differences*, 114, 73–81. <https://doi.org/10.1016/j.paid.2017.03.050>
- Brown, A., & Casey, B. (2016). Subclinical ADHD-Symptoms Are Associated with Executive-Functioning and Externalizing Problems in College Students without ADHD-Diagnoses. *Journal of Educational and Developmental Psychology*, 6, 204. <https://doi.org/10.5539/jedp.v6n1p204>

- Bzdok, D., Altman, N., & Krzywinski, M. (2018). Statistics versus machine learning. *Nature Methods*, 15(4), 233–234. <https://doi.org/10.1038/nmeth.4642>
- Cao, M., Martin, E., & Li, X. (2023). Machine learning in attention-deficit/hyperactivity disorder: New approaches toward understanding the neural mechanisms. *Translational Psychiatry*, 13, 236. <https://doi.org/10.1038/s41398-023-02536-w>
- Chen, T., Antoniou, G., Adamou, M., Tachmazidis, I., & Su, P. (2021). Automatic Diagnosis of Attention Deficit Hyperactivity Disorder Using Machine Learning. *Applied Artificial Intelligence*, 35(9), 657–669. <https://doi.org/10.1080/08839514.2021.1933761>
- Cortese, S., & Coghill, D. (2018). Twenty years of research on attention-deficit/hyperactivity disorder (ADHD): Looking back, looking forward. *Evidence Based Mental Health*, 21(4), 173–176. <https://doi.org/10.1136/ebmental-2018-300050>
- Coutinho, D., Farias, A. C., Felden, E. P. G., & Cordeiro, M. L. (2021). ADHD Comorbid With Major Depression on Parents and Teachers Perceptions. *Journal of Attention Disorders*, 25(4), 508–518. <https://doi.org/10.1177/1087054718815574>
- Demetrovics, Z., Szeredi, B., & Rózsa, S. (2008). The three-factor model of Internet addiction: The development of the Problematic Internet Use Questionnaire. *Behavior Research Methods*, 40(2), 563–574. <https://doi.org/10.3758/BRM.40.2.563>
- Dhalla, S., & Kopec, J. A. (2007). The CAGE Questionnaire for Alcohol Misuse: A Review of Reliability and Validity Studies. *Clinical and Investigative Medicine*, 33–41. <https://doi.org/10.25011/cim.v30i1.447>
- Diogo, V. S., Ferreira, H. A., Prata, D., & Alzheimer's Disease Neuroimaging Initiative. (2022). Early diagnosis of Alzheimer's disease using machine learning: A multi-diagnostic, generalizable approach. *Alzheimer's Research & Therapy*, 14(1), 107.

<https://doi.org/10.1186/s13195-022-01047-y>

- Ewing, J. A. (1984). Detecting Alcoholism: The CAGE Questionnaire. *JAMA*, 252(14), 1905–1907. <https://doi.org/10.1001/jama.1984.03350140051025>
- Faraone, S. V., Biederman, J., Doyle, A., Murray, K., Petty, C., Adamson, J. J., & Seidman, L. (2006). Neuropsychological studies of late onset and subthreshold diagnoses of adult attention-deficit/hyperactivity disorder. *Biological Psychiatry*, 60(10), 1081–1087. <https://doi.org/10.1016/j.biopsych.2006.03.060>
- Furman, L. (2005). What Is Attention-Deficit Hyperactivity Disorder (ADHD)? *Journal of Child Neurology*, 20(12), 994–1002. <https://doi.org/10.1177/08830738050200121301>
- Ginapp, C. M., Macdonald-Gagnon, G., Angarita, G. A., Bold, K. W., & Potenza, M. N. (2022). The lived experiences of adults with attention-deficit/hyperactivity disorder: A rapid review of qualitative evidence. *Frontiers in Psychiatry*, 13, 949321. <https://doi.org/10.3389/fpsyt.2022.949321>
- Gritti, E. S., Matta, M., & Lang, M. (2023). Emotion dysregulation and ADHD subclinical manifestations in late adolescents: A study with a focus on inattention. *BPA Applied Psychology Bulletin*, 81(297), 20–36. <https://doi.org/10.26387/bpa.2023.00008>
- Hand, D. J. (2012). Assessing the Performance of Classification Methods. *International Statistical Review*, 80(3), 400–414. <https://doi.org/10.1111/j.1751-5823.2012.00183.x>
- Harstad, E., Levy, S., & Committee on Substance Abuse. (2014). Attention-deficit/hyperactivity disorder and substance abuse. *Pediatrics*, 134(1), e293-301. <https://doi.org/10.1542/peds.2014-0992>
- Hinkin, C. H., Castellon, S. A., Dickson-Fuhrman, E., Daum, G., Jaffe, J., & Jarvik, L. (2001). Screening for Drug and Alcohol Abuse Among Older Adults Using a Modified Version

- of the CAGE. *The American Journal on Addictions*, 10(4), 319–326.
<https://doi.org/10.1111/j.1521-0391.2001.tb00521.x>
- Hong, S.-B., Dwyer, D., Kim, J.-W., Park, E.-J., Shin, M.-S., Kim, B.-N., Yoo, H.-J., Cho, I.-H., Bhang, S.-Y., Hong, Y.-C., Pantelis, C., & Cho, S.-C. (2014). Subthreshold attention-deficit/hyperactivity disorder is associated with functional impairments across domains: A comprehensive analysis in a large-scale community study. *European Child & Adolescent Psychiatry*, 23(8), 627–636. <https://doi.org/10.1007/s00787-013-0501-z>
- Ivanov, I., Bjork, J. M., Blair, J., & Newcorn, J. H. (2022). Sensitization-based risk for substance abuse in vulnerable individuals with ADHD: Review and re-examination of evidence. *Neuroscience & Biobehavioral Reviews*, 135, 104575.
<https://doi.org/10.1016/j.neubiorev.2022.104575>
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255–260. <https://doi.org/10.1126/science.aaa8415>
- Jung, J. S., Park, S. J., Kim, E. Y., Na, K.-S., Kim, Y. J., & Kim, K. G. (2019). Prediction models for high risk of suicide in Korean adolescents using machine learning techniques. *PLOS ONE*, 14(6), e0217639. <https://doi.org/10.1371/journal.pone.0217639>
- Kessler, R. C., Adler, L. A., Gruber, M. J., Sarawate, C. A., Spencer, T., & Van Brunt, D. L. (2007). Validity of the World Health Organization Adult ADHD Self-Report Scale (ASRS) Screener in a representative sample of health plan members. *International Journal of Methods in Psychiatric Research*, 16(2), 52–65.
<https://doi.org/10.1002/mpr.208>
- Kessler, R. C., Adler, L., Ames, M., Demler, O., Faraone, S., Hiripi, E., Howes, M. J., Jin, R., Secnik, K., Spencer, T., Ustun, T. B., & Walters, E. E. (2005). The World Health

- Organization adult ADHD self-report scale (ASRS): A short screening scale for use in the general population. *Psychological Medicine*, 35(2), 245–256.
<https://doi.org/10.1017/S0033291704002892>
- Kessler, R. C., Adler, L., Barkley, R., Biederman, J., Conners, C. K., Demler, O., Faraone, S. V., Greenhill, L. L., Howes, M. J., Secnik, K., Spencer, T., Ustun, T. B., Walters, E. E., & Zaslavsky, A. M. (2006). The prevalence and correlates of adult ADHD in the United States: Results from the National Comorbidity Survey Replication. *The American Journal of Psychiatry*, 163(4), 716–723. <https://doi.org/10.1176/ajp.2006.163.4.716>
- Khan Mamun, M. M. R., & Elfouly, T. (2023). AI-Enabled Electrocardiogram Analysis for Disease Diagnosis. *Applied System Innovation*, 6(5), Article 5.
<https://doi.org/10.3390/asi6050095>
- Kóbor, A., Takács, Á., Urbán, R., & Csépe, V. (2012). The latent classes of subclinical ADHD symptoms: Convergences of multiple informant reports. *Research in Developmental Disabilities*, 33(5), 1677–1689. <https://doi.org/10.1016/j.ridd.2012.04.008>
- Kosheleff, A. R., Mason, O., Jain, R., Koch, J., & Rubin, J. (2023). Functional Impairments Associated With ADHD in Adulthood and the Impact of Pharmacological Treatment. *Journal of Attention Disorders*, 27(7), 669–697.
<https://doi.org/10.1177/10870547231158572>
- Levy, S., Katusic, S. K., Colligan, R. C., Weaver, A. L., Killian, J. M., Voigt, R. G., & Barbaresi, W. J. (2014). Childhood ADHD and risk for substance dependence in adulthood: A longitudinal, population-based study. *PloS One*, 9(8), e105640.
<https://doi.org/10.1371/journal.pone.0105640>
- Magon, R., & Müller, U. (2012). ADHD with comorbid substance use disorder: Review of

- treatment. *Advances in Psychiatric Treatment*, 18(6), 436–446.
<https://doi.org/10.1192/apt.bp.111.009340>
- Mahone, E. M., & Denckla, M. B. (2017). Attention-Deficit/Hyperactivity Disorder: A Historical Neuropsychological Perspective. *Journal of the International Neuropsychological Society : JINS*, 23(9–10), 916–929. <https://doi.org/10.1017/S1355617717000807>
- Merikangas, K. R., He, J.-P., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., Benjet, C., Georgiades, K., & Swendsen, J. (2010). Lifetime prevalence of mental disorders in U.S. adolescents: Results from the National Comorbidity Survey Replication--Adolescent Supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry*, 49(10), 980–989. <https://doi.org/10.1016/j.jaac.2010.05.017>
- Nemesure, M. D., Heinz, M. V., Huang, R., & Jacobson, N. C. (2021). Predictive modeling of depression and anxiety using electronic health records and a novel machine learning approach with artificial intelligence. *Scientific Reports*, 11(1), 1980.
<https://doi.org/10.1038/s41598-021-81368-4>
- Nickson, D., Meyer, C., Walasek, L., & Toro, C. (2023). Prediction and diagnosis of depression using machine learning with electronic health records data: A systematic review. *BMC Medical Informatics and Decision Making*, 23, 271.
<https://doi.org/10.1186/s12911-023-02341-x>
- Norton, P. J. (2007). Depression Anxiety and Stress Scales (DASS-21): Psychometric analysis across four racial groups. *Anxiety, Stress, & Coping*, 20(3), 253–265.
<https://doi.org/10.1080/10615800701309279>
- Ohlmeier, M. D., Peters, K., Te Wildt, B. T., Zedler, M., Ziegenbein, M., Wiese, B., Emrich, H. M., & Schneider, U. (2008). Comorbidity of alcohol and substance dependence with

- attention-deficit/hyperactivity disorder (ADHD). *Alcohol and Alcoholism (Oxford, Oxfordshire)*, 43(3), 300–304. <https://doi.org/10.1093/alcalc/agn014>
- Orrù, G., Monaro, M., Conversano, C., Gemignani, A., & Sartori, G. (2020). Machine Learning in Psychometrics and Psychological Research. *Frontiers in Psychology*, 10. <https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02970>
- Overbey, G. A., Snell, W. E., & Callis, K. E. (2011). Subclinical ADHD, Stress, and Coping in Romantic Relationships of University Students. *Journal of Attention Disorders*, 15(1), 67–78. <https://doi.org/10.1177/1087054709347257>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., & Cournapeau, D. (2011). Scikit-learn: Machine Learning in Python. *MACHINE LEARNING IN PYTHON*.
- Rimmerman, A., Yurkevich, O., Birger, M., Azaiza, F., & Elyashar, S. (2007). Quality of life of Israeli adults with borderline intelligence quotient and attention-deficit/hyperactivity disorder. *International Journal of Rehabilitation Research. Internationale Zeitschrift Fur Rehabilitationsforschung. Revue Internationale De Recherches De Readaptation*, 30(1), 55–60. <https://doi.org/10.1097/MRR.0b013e328013d8a0>
- Rosenbusch, H., Soldner, F., Evans, A. M., & Zeelenberg, M. (2021). Supervised machine learning methods in psychology: A practical introduction with annotated R code. *Social and Personality Psychology Compass*, 15(2), e12579. <https://doi.org/10.1111/spc3.12579>
- Rustin, T. A. (2000). Assessing nicotine dependence. *American Family Physician*, 62(3), 579–584, 591–592.
- Sevincok, D., Ozbay, H. C., Ozbek, M. M., Tunagur, M. T., & Aksu, H. (2020). ADHD symptoms in relation to internalizing and externalizing symptoms in children: The

- mediating role of sluggish cognitive tempo. *Nordic Journal of Psychiatry*, 74(4), 265–272. <https://doi.org/10.1080/08039488.2019.1697746>
- Sohn, S. Y., Rees, P., Wildridge, B., Kalk, N. J., & Carter, B. (2019). Prevalence of problematic smartphone usage and associated mental health outcomes amongst children and young people: A systematic review, meta-analysis and GRADE of the evidence. *BMC Psychiatry*, 19(1), 356. <https://doi.org/10.1186/s12888-019-2350-x>
- Toner, M., Donoghue, T., & Houghton, S. (2006). Living in Chaos and Striving for Control: How adults with Attention Deficit Hyperactivity Disorder deal with their disorder. *International Journal of Disability*, 53, 247–261. <https://doi.org/10.1080/10349120600716190>
- Wang, B.-Q., Yao, N.-Q., Zhou, X., Liu, J., & Lv, Z.-T. (2017). The association between attention deficit/hyperactivity disorder and internet addiction: A systematic review and meta-analysis. *BMC Psychiatry*, 17(1), 260. <https://doi.org/10.1186/s12888-017-1408-x>
- Weissenberger, S., Ptacek, R., Vnukova, M., Raboch, J., Klicperova-Baker, M., Domkarova, L., & Goetz, M. (2018). ADHD and lifestyle habits in Czech adults, a national sample. *Neuropsychiatric Disease and Treatment*, 14, 293–299. <https://doi.org/10.2147/NDT.S1>