

Predicting ADHD Using Activity Time Series Data

Irene Foster^{1,*} 

¹ Smith College - Department of Statistical & Data Sciences Northampton, MA, USA;

* Correspondence: ifoster25@smith.edu

Version April 24, 2023 submitted to Journal Not Specified



Abstract: **TODO** [Usually, 150-200 words overview of the research paper you have conducted – includes short introduction to the research question, review to methods, review of key results, conclusion for the results in line with the research paper.]

Keywords: ADHD; behavioral activity; machine learning; time series

1. Introduction

Attention-deficit/hyperactivity disorder (ADHD) is a complex neurodevelopmental disorder that shares symptoms with many other conditions, and it is often misdiagnosed. It is estimated that 8.4% of children and 2.5% of adults have ADHD, and presentation and assessment are different in the two groups (Elmaghraby and Stephanie Garayalde [1]). There are three main types of ADHD: inattentive presentation, hyperactive/impulsive presentation, and combined presentation (Elmaghraby and Stephanie Garayalde [1]). The inattentive type is characterized by difficulty staying on task, sustaining focus, and staying organized (noa [2]). Hyperactivity is excessive movement and may present as restlessness or talking too much in adults (noa [2]). Impulsivity is when a person acts without thinking and may manifest as desire for immediate rewards or the inability to delay gratification (noa [2]). The combined type is when both symptoms of the inattentive type and the hyperactive/impulsive type are present (noa [2]). ADHD can impact individuals in many areas of their life such as academic/professional, interpersonal relationships, and daily functioning (Elmaghraby and Stephanie Garayalde [1]). In adults it can have far reaching detrimental effects and lead to poor self-worth, sensitivity towards criticism, and increased self-criticism (Elmaghraby and Stephanie Garayalde [1]). However, sometimes ADHD is not identified until a person is an adult if the symptoms were not recognized, they had mild ADHD, or they managed sufficiently well until demands of college/work (noa [3]). Due to the harmful consequences ADHD can lead to, it is important that it is diagnosed and treated.

There are many challenges to diagnosing ADHD, particularly in adults. Adult ADHD symptoms are sometimes harder to discern than ADHD symptoms in children (noa [4]). Combining this with the fact that adult ADHD symptoms are similar to those in other conditions can make diagnosis difficult (noa [4]). Stress, illness, and other mental conditions such as anxiety or mood disorders can all have symptoms that are similar to ADHD (noa [4], noa [3]). For example, emotional dysregulation present in ADHD can be diagnosed with a mood disorder or ADHD symptoms can be covered up by substance abuse (Katzman *et al.* [5]). Physicians are also usually more familiar with mood and anxiety disorders, leading to misdiagnosis and delays in ADHD treatment (Katzman *et al.* [5]). Additionally, other mental health conditions such as anxiety, mood, and substance use disorders are common in adults with ADHD (noa [3]). Studies have shown that 18.6% to 53.3% of people with ADHD have depression and almost 50% of people with ADHD have an anxiety disorder (Katzman *et al.* [5]). Some researchers suggest that in some cases stress, depression, and anxiety may be manifesting due to undiagnosed

or untreated ADHD (Katzman *et al.* [5]). These factors make ADHD difficult to recognize and treat, leading to an under-diagnosis and under-treatment of adult ADHD (Katzman *et al.* [5]). Due to the extensive effects ADHD can have, it is important that it is properly diagnosed and treated.

There is no specific procedure to diagnose ADHD and psychiatrists, neurologists, primary care doctors, clinical psychologists, or clinical social workers can all diagnose adults with ADHD (Contributors [6]). Steps to getting a diagnosis may include a physician using behavioral questionnaires to ask about the impacts ADHD has, possible symptoms present in childhood, talking to a parent or partner, and psychological tests (Contributors [6]). They may also test for learning disabilities, other mental health conditions, or physical illnesses to rule these options out (Contributors [6]).

A large part of the ADHD diagnosis process includes objective information: the patient's perspective of themselves in behavioral questionnaires, the perspective of parents and significant others, and the opinion/view of the clinician. Rating scales, which are often used in diagnosis, are systematic but they are not objective (Gualtieri and Johnson [7]). Raters are prone to let their view of the subject and the outcome they want skew their results and different raters often differ in their view of the same subject (Gualtieri and Johnson [7]). Patients are also evaluated by a clinician during diagnosis, which can be a primary care physician in the United States (Contributors [6]). Recent research has indicated in the United States children who are older for their grade level are less likely to be diagnosed with ADHD (Dalsgaard *et al.* [8]). However in Denmark only specialists diagnose ADHD and these results were not replicated (Dalsgaard *et al.* [8]). This supports the hypothesis that non-specialists diagnosing ADHD could be the reason for the lower rate of ADHD diagnosis in children who are older for their grade (Dalsgaard *et al.* [8]). Clinicians may also be subject to bias that affects their diagnoses of ADHD and there is evidence for racial and gender disparities in diagnosis (Stanborough [9]). Boys are more likely to be diagnosed with ADHD in childhood than girls, though this may be due to different presentation of symptoms or differences in building compensation skills (Stanborough [9]). Black, Hispanic, and Asian children and adults received ADHD diagnoses less often than their non-Hispanic white counterparts (Stanborough [9]). In addition to rating scales and clinical evaluations, computerized tests can also be used to help diagnose ADHD. A computer test alone is not enough for a diagnosis, but can supplement other diagnostic tools (Gualtieri and Johnson [7]). Continuous performance tests are used in ADHD diagnoses and test vigilance or sustained attention (Gualtieri and Johnson [7]). However, there is limited correlation between CPT results and rating scales (Gualtieri and Johnson [7]). Additionally, the most common CPTs have about a 85% success rate in indicating ADHD in children that have been diagnosed with ADHD and a false positive rate of 30% in controls (Gualtieri and Johnson [7]). All together these make for a subjective, potentially inaccurate diagnostic method.

Given the difficulties in diagnosing ADHD and the subjective methods used, it's important that researchers investigate more accurate and objective measures of diagnosis. **TODO: add previous research on other predictive algorithms**

I will be continuing the research on predicting ADHD using machine learning methods. **TODO: add why im using activity data**

2. Methods

2.1. Dataset

TODO: update, add more description

I am planning to use a dataset called HYPERAKTIV, which has 103 subjects. It contains subjects' sex, age, prescribed medications, which psychiatric disorders are diagnosed (including ADHD/ADD), and if substance abuse exists. For every subject it contains time series of motor activity and heart rate over a 24-hour period and the results of a neuropsychological computer test. Additionally, a paper was written on it that contains robust metadata and the researchers did some beginning exploratory

work on the data using machine learning algorithms to predict ADHD. I intend to use the time series of motor activity and age, sex, psychiatric disorders and prescriptions medications to predict ADHD.

I am planning to use a dataset called HYPERAKTIV, which has 103 subjects. It contains subjects' sex, age, prescribed medications, which psychiatric disorders are diagnosed (including ADHD/ADD), and if substance abuse exists. For every subject it contains time series of motor activity and heart rate over a 24-hour period and the results of a neuropsychological computer test. Additionally, a paper was written on it that contains robust metadata and the researchers did some beginning exploratory work on the data using machine learning algorithms to predict ADHD. I intend to use the time series of motor activity and age, sex, psychiatric disorders and prescriptions medications to predict ADHD.

[Contrast this to what would have been your ideal dataset. How did you arrive at choosing this dataset. Provide details on the dataset, size, gender split and count, age, how it was collected, ethical concerns, etc. Also provide details on data collection and access. Usually 100-200 words.]

2.2. Data Cleaning

In the Hyperaktiv data set the time series data for each subject is stored in a separate CSV file. To analyze the data, I combined all of the activity CSVs into one file, keeping track of which subject each data point came from. *TODO: add more detail if needed* There were 108 subjects, yet only some had activity data. I only used subjects with activity data, which reduced the data set to 83.

TODO: comment on if I make all activity csvs same length, cleaning time series (seasonality, get rid of noise, etc.)

TODO: Add about heart rate if you get there

2.3. Variables of Interest

TODO, also include summary stats of variables

[Describe the variables that you are using in your datasets. Describe the rationale for using these variables. Varies in length.]

2.4. Other Important Features

TODO: Researchers did some preliminary ML on dataset - discuss

[Also comment on other things which are relevant to your research paper. Varies in length.]

2.5. Data Analyses

TODO: this is info from draft, update with actual procedure taken

First, I will calculate summary statistics of the variables I'm using introductory data visualizations. It looks like to be compared, time series have to be the same length. Since the data I am using is not all the same length, if I am unable to find a work-around I will have to preprocess the data by finding the mean of the lengths and then truncating the series that are too long and padding the ones that are too short.

My research question involves investigating if ADHD can be predicted using motor activity. I would

Before modeling of performing significance tests on the data, I will check if the data is stationary by using a Phillips-Perron Unit Root Test or by plotting the sample autocorrelation function. If it isn't stationary I will remove the trends or seasonality from the data (trends: difference the data or perform least squares trend removal, seasonality: seasonal differencing, seasonal means, moving averages). To compare the two groups and see if there is a significant difference between them, I am going to try fitting a model to each group of the time series, use the model to simulate new data, and then run statistics on the simulated data. I am going to continue looking to see if there is a more direct method for comparing two groups of time series data.

In addition to looking at the data as time series data, I could also try picking out features of the motor activity data and see if there is a significant difference between ADHD vs non-ADHD groups

and if prediction is possible. I could pick out features such as averages, standard deviations, or number of peaks in the motor activity time series data. This would simplify the modeling, significance testing, and prediction algorithms steps since this data would be easier to use than time series data. It would also allow me to include confounding variables such as sex, age, and other psychiatric disorders because I have not yet found a way to include confounding variables in time series data. If I have time I could also include heart rate of people as a predictor variable in addition to motor activity to predict if people have ADHD or not. There has been some evidence of difference in heart rates between the two groups, so it could be beneficial to look at this (Hicks et al., 2021)

[What analyses are you doing why. Varies in length.]

3. Results

[Summary of results analyses 1]

[Add visualizations from analyses. Varies in length.]

[Summary of results analyses 2]

[Add visualizations from analyses. Varies in length.]

4. Discussion

- Talk about what the results mean
- Depending on results - could have positive impact on diagnosing ADHD
- Diagnosis in data is still coming from people - so still has subjectivity
- Couldn't include confounding variables??
- Most practitioners don't have the technical skill to do it themselves

[Explain what your results mean in the context of the literature cited in the introduction. Minimum of 750 words.]

5. Future Directions

- Larger sample size
- Combine with heart rate data
- Could try on children
- Make accessible to practitioners if it works
- If I found another dataset, I could also explore if lifestyle factors (maybe things such as amount of exercise, level of satisfaction, income level, etc.) can be used to predict ADHD.

TODO: change citation style

[How could someone continue the work? Around 150-200 words.]

6. Personal Reflection

[What was your original research question? Why did you have to deviate? What ideal dataset would be recommend collecting? How was this process of developing a research paper? 250 words]

7. Code availability

All analysis code for this article is available at: <https://github.com/i-m-foster/sds300np-ireneFoster>

8. Acknowledgements

[Mention who you would like to thank. Any grants or people. Varies in length.]

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Elmaghraby, R.; Stephanie Garayalde, S. What is ADHD? <https://www.psychiatry.org:443/patients-families/adhd/what-is-adhd>, 2022.

2. Attention-Deficit/Hyperactivity Disorder. <https://www.nimh.nih.gov/health/topics/attention-deficit-hyperactivity-disorder-adhd>, 2022.
3. Attention-Deficit/Hyperactivity Disorder in Adults: What You Need to Know. <https://www.nimh.nih.gov/health/publications/adhd-what-you-need-to-know>, 2021.
4. Adult attention-deficit/hyperactivity disorder (ADHD) - Symptoms and causes. <https://www.mayoclinic.org/diseases-conditions/adult-adhd/symptoms-causes/syc-20350878>, 2023.
5. Katzman, M.A.; Bilkey, T.S.; Chokka, P.R.; Fallu, A.; Klassen, L.J. Adult ADHD and comorbid disorders: clinical implications of a dimensional approach. *BMC Psychiatry* **2017**, *17*, 302.
6. Contributors, W.E. Diagnosing ADHD in Adults: How Adults Are Tested for ADHD. <https://www.webmd.com/add-adhd/diagnosing-adhd-adults>, 2023.
7. Gualtieri, C.T.; Johnson, L.G. ADHD: Is Objective Diagnosis Possible? *Psychiatry (Edgmont)* **2005**, *2*, 44–53.
8. Dalsgaard, S.; Humlum, M.K.; Nielsen, H.S.; Simonsen, M. Relative standards in ADHD diagnoses: The role of specialist behavior. *Economics Letters* **2012**, *117*, 663–665. doi:<https://doi.org/10.1016/j.econlet.2012.08.008>.
9. Stanborough, R.J. What to Know About Adult ADHD Diagnosis. <https://www.healthline.com/health/adult-adhd-diagnosis>, 2022.

© 2023 by the authors. Submitted to *Journal Not Specified* for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).