Archives of Neuroscience
Volume 10 · Issue 2 · e134329

OPEN PEER REVIEW

Improvement of Classification Performance in High-Dimension Low-Sample-Size Modeling by Sparse Functional Connectivity States in Subjects with Attention Deficit-Hyperactivity Disorder and Healthy Controls

Zahra Zolghadr, Mahmoudreza Hadjighassem, Seyed Amirhossein Batouli, Mehdi Tehrani-Doost, Lida Shafaghi, Hamid Alavi Majd, Yadollah Mehrabi

SUBMITTED	REVISED	ACCEPTED	DOI
24 Dec 2022	21 Mar 2023	27 Mar 2023	10.5812/ans-134329

About this report

This open peer review (OPR) report shows the public comments made on this article during peer review, by the reviewers, the associate editor and the editor-in-chief, for every revision round, together with the authors' replies. Comments appear only with the consent of the people who wrote them. Reviewers who chose to sign their report are named; the others are shown as "Referee". Comments are edited for language only, and confidential remarks to the editors are not included.

More about open peer review: brieflands.com/knowledgebase/opr

Contents

1 **Original submission** Seyed-Amirabbas Ahadiat, Referee 1, Associate Editor, Editor-in-Chief 2

2 **Revision 1** Referee 1, Seyed-Amirabbas Ahadiat, Associate Editor, Editor-in-Chief 4

Authors' reply to reviewers 5

Reviewer and editor comments 10

1 Original submission

4 comments · 31 Dec 2022 – 4 Jan 2023

Seyed-Amirabbas Ahadiat Reviewer

MAJOR REVISION 31 Dec 2022

dear author

This article has an interesting topic, but it needs to be edited before it is reviewed.

There is a mistake in the topic of the article, and it needs to be edited.
What type of study is it?
What are the inclusion and exclusion criteria for this study? What are the time and place of the study?

The introduction is very long and is more like a literature review than an original article, and what is the meaning of presenting the formula in the introduction?

I would appreciate it if you could adjust the article in accordance with the journal's standard protocol and send it again.

Referee 1 Reviewer

MAJOR REVISION 3 Jan 2023

This is a study based on machine learning that categorizes the diagnostic status of attention deficit hyperactivity disorder (ADHD). The data that had been preprocessed from the ADHD-200 worldwide competition in 2011 were evaluated. The training included seven hundred sixty-eight items, while the test consisted of one hundred seventy-one data points. In order to classify ADHD, the authors evaluated the efficacy of three different models: support vector machine (SVM), distance weighted discrimination (DWD), and data maximum dispersion classifier (DMDC). According to the authors, the performance of DMDC, when applied to sparse data, significantly improved classification outcomes across the three existing models. The neural connection between subcortical regions including portions of the basal ganglia and cerebellum distinguishes between ADHD and healthy controls based on the current results.
1. The research topic as well as the primary objective of this investigation is not stated anywhere in the abstract.
2. The second paragraph of the introduction contains many statements, none of which has been cited with appropriate references.
3. Table 1 lacks the age variable, which I recommend that the authors include.
4. Please cite the following sentence with its related reference:
“The best model presented in the world competition belongs to the Johns Hopkins University team, a cumulative model with 61% overall accuracy, 57.5% balanced rating, 21% sensitivity, and 94% specificity”.
5. Most sentences in the discussion section also have not been cited with their related references.
6. Compare the findings of this research to those of the gold standard.
7. What is the medical task of this model? Prognostic or diagnostic? Please clarify in the methodology and discuss it in the discussion.
8. How was data labeling performed? How did you determine the classification of ADHD? This needs more clarification.
9. How do you interpret your findings that these data are representative of the population?
10. Were there any missing data? Please explain in detail and explain how you managed the missing data.
11. How did you determine and eliminate redundant independent variables, if applicable?
12. Are sufficient data available for the model? Assess whether sufficient data are available for a good fit of the

model.
13. Please provide the details of the training approach, including data augmentation and hyperparameters, if applicable.
14. What technique was used to split the data into partitions?
15. You merely separated the data into the train and test sets. What about validation?
16. Please explain the transferability of this model to the new clinical setting.
17. What about the replicability and generalizability of the model? How are the details of the model available for replicability?
18. Please explain the clinical utility of this model at the end of the discussion.
19. Please mention the limitations of the model.

Associate Editor MAJOR REVISION 4 Jan 2023

Dear EIC,
The manuscript needs revision. Please ask the authors to revise it and reply to the reviewers' comments.

Editor-in-Chief MAJOR REVISION 4 Jan 2023

Dear Author,
Thank you for submitting the manuscript. It needs revision. Please reply to all comments accurately. Also refine the manuscript based on the comments.

2

Revision 1

AUTHORS' RESPONSE TO REVIEWERS

Authors' reply to reviewers

The next pages contain the authors' point-by-point reply to the comments of this revision round, as submitted with the revised manuscript. It is published as the authors wrote it.

Response to Comments & Questions

Thank you for giving us the opportunity to submit the revised draft of our manuscript titled *"Sparse Functional Connectivity States Could Improve the Classification Performance in High-dimensional Low Sample Size Modeling: Attention Deficit Hyperactivity Disorder vs Healthy Controls"*. The time and efforts that the editor and the reviewers have dedicated to provide the valuable feedback on our manuscript are highly appreciated. We are grateful to the reviewers for their insightful comments on our paper. We have been able to incorporate changes to reflect most of the suggestions provided by the reviewers. Changes were highlighted within the manuscript in two colors specific for each set of comments and questions.

Sincerely Yours

Dr. Y. M

Corresponding author

22/01/ 2023

Reviewer 1:

This article has an interesting topic, but it needs to be edited before it is reviewed

1. ***There is a mistake in the topic of the article and it needs to be edited.***

#1. Thank you for this comment. We accordingly corrected some words (low-sample and attention-deficit). We appreciate if you could inform us about the further needed modifications in the title.

2. ***What type of study is it?***

#2. The current study is a cross-sectional machine learning analysis and is based on the world-wide available fMRI database (ADHD-200). We have added an expression describing the type of the study in the Methods section of the abstract (highlighted in yellow).

3. ***What is the inclusion and exclusion criteria for this study? Time and place of the study?***

#3. Many thanks. The current analysis project is based on a multicenter (eight centers) database ***(http://fcon_1000.projects.nitrc.org/)***, and each center had its own defined I/E criteria. However, the common (necessary) items among them were as follows:

The included subjects had received a confirmed diagnosis of ADHD (combined, inattentive, hyperactive), based on the DSM-IV criteria (diagnosis of ADHD and ruling out the other disorders, specifically other neurodevelopmental disorders and anxiety/ depressive-associated states) by the same board-certified child/adolescent psychiatrist (DPD) for all participants. The whole details of the diagnostic procedures and considerations have been brought in the ***"Diagnostics"*** section of each center. There were also a battery of complementary and rating scales and questionnaires for providing additional disorder-specific information (there exists a minor variability regarding these complementary tests and assays regarding each center included). Moreover, based on your valuable comment, we added complementary information about the inclusion/exclusion criteria through the text (methods) and highlighted that in yellow.
The database has been released in 2011, it is expected that the center-specific release had been around that time. The places were 1. Bradley Hospital/Brown University, 2. Kennedy Krieger Institute, 3. NeuroIMAGE Sample, 4. New York

University Child Study Center, 5. Oregon Health & Science University, 6. Peking University, 7. University of Pittsburgh, 8. Washington University IN ST. Louis.

4. ***The volume of the introduction is very large and it is more like a literature review than an original article, and what is the meaning of presenting the formula in the introduction?***

#4. Many thanks for this comment. We accordingly omitted some sentences in the introduction part (track-changed). Regarding presenting formula in the introduction, in the analysis (specifically machine-learning modeling)-associated studies, sometimes authors present their applied formulas in the introduction for more clarification. According to your consideration, we transferred the formulas to the method part and highlighted them.

As an example of formula in the introduction part is the following item:

- *Wang, Boxiang, and Hui Zou. "Sparse distance weighted discrimination." Journal of Computational and Graphical Statistics 25.3 (2016): 826-838.*

5. **I would appreciate it if you could adjust the article in accordance with the journal's standard protocol and send it again.**

#5. Many thanks. We tried to align the text with the corresponding guidelines.

Reviewer 2:

This is a research based on machine learning that categorizes the diagnostic status of attention deficit hyperactivity disorder (ADHD). The data that had been preprocessed from the ADHD-200 worldwide competition in 2011 were evaluated. The training included seven hundred sixty-eight items, while the test consisted of one hundred seventy-one data points. In order to classify ADHD, the authors evaluated the efficacy of three different models: support vector machine (SVM), distance weighted discrimination (DWD), and data maximum dispersion classifier (DMDC). According to the authors, the performance of DMDC, when applied to sparse data, significantly improved classification outcomes across the three existing models. The neural connection between subcortical regions including portions of the basal ganglia and cerebellum distinguishes between ADHD and healthy controls based on the current results.

1. The research topic as well as the primary objective of this investigation are not stated anywhere in the abstract.

#1. Many thanks for this brilliant comment. We accordingly added this suggestion to the end of the abstract background and highlighted in turquois.

2. The second paragraph of the introduction contains many statements, none of which has been cited with appropriate references.

#2. Thanks. We added the references and tried to add the most relevant ones.

3. Table 1 lacks the age variable, which I recommend the authors to include.

#3 Many thanks for considering this. We added this to the table.

4.Please cite the following sentence with its related reference: "The best model presented in the world competition belongs to the Johns Hopkins University team, a cumulative model with 61% overall accuracy, 57.5% balanced rating, 21% sensitivity, and 94% specificity".

#4. We added the associated reference.

5. Most sentences in the discussion section also have not been cited with their related references.

#5. Many thanks for this comment. We accordingly added the associated references and citations.

6. Compare the findings of this research to those of the gold standard.

#6. The gold standard of the diagnosis of ADHD is the standard clinical interview that some studies stated the superiority of this approach (clinical psychiatric interview) compared to the partially bothersome, relatively expensive, and not sufficiently accessible Neuromapping and more specifically neuroimaging techniques. Additionally, applying machine learning-based algorithm require sufficient infrastructures and facilities. However, there are also serious concerns and limitations regarding the conventional interview-based attitudes. Through the interview process, the psychiatrist could not obtain the specific information about the brain structural and functional themes; this is important since a large body of research stated the neuroendophenotypic differences at the level of neurocognitive processing which might affect the prognosis and treatment responses. This clarification could also be efficacious in designing the cognitive-focused and emotion-centered and also neurostimulation treatments. Since there are growing concerns about the longitudinal use of pharmacotherapeutic agents, non-pharmacological and subsequent need for knowing the neuronal basis would be trend. Moreover, gaining deep understanding about the pathophysiological aspects is beneficial in designing more specific and selective pharmacological treatments. For instance, the specific role of cerebellar parts in the neurodevelopmental disorders and also their feature-like contribution in large analyses is an achievement of the recent progress in these fields. Trading off these dimensions, it seems to be a more applicable context for future use of the integrated neuroimaging-machine learning methods.

7. What is the medical task of this model? Prognostic or diagnostic? Please clarify in the methodology and discuss in the discussion.

#7. Many thanks for this consideration and note. The applications Neuroimaging-based technologies and corresponding computational approaches in analyzing and interpretation of results in medical and clinical settings are relatively new frameworks and despite widespread utilization of novel neuroimaging parameters (specifically functional neuroimaging theme) in research and pre-clinical stages and also remarkable potentials in clinical practices, we have still in the infancy level and it requires more studies and more intensive interpretations. About the approaches and obtained findings in the current study, it clarifies a very narrow spot within the black box of the pathophysiology of neurodevelopmental disorders. Either prognostic and diagnostic applications need more prolonged and more controlled studies, that based on our knowledge, there is a significant lack of these well-designed and well controlled studies to draw a confident conclusion. However, prognostic considerations (prognosis of diagnosis and treatment) seem to be more feasible regarding applying the integrated machine learning and neuroimaging. Accordingly, we have added two parts (for method and discussion) to the text (with relevant references) and highlighted them.

8. How was data labeling performed? How did you determine the classification of ADHD? This needs more clarification.

#8. Thanks. The available dataset of ADHD-200 has provided the neuroimaging data with corresponding label of "ADHD" or "Healthy". These already existing labels are based on the specific criteria of diagnosing primarily according to the DSM-IV criteria and related confirmatory and severity scaling complementary tests and assessments. Thus, we performed the classification procedures with the labeled data and categories.

9. How do you interpret your findings that these data are representative of the population?

#9. Thank you for this consideration. There are two points: although these data are acquired from different centers of neuroimaging and neuropsychiatry, but since the selection criteria (the global standard diagnosis process alongside the valid and reliable psychological tools) and also the imaging acquisition protocols have been identical, the data characteristics could be considered as sufficiently comparable to both "between centers" and then to the "other people with the same diagnosis based on the currently accepted guidelines. Another point is that, despite the centers are mainly located in USA-related universities, there are also other centers from other countries; this might increase the representativeness.

10. Was there any missing data? Please explain in detail and explain how did you manage the missing data?

#10. Many thanks for this question. As you considered, we have had some missing values (less than 5 percent) in personal characteristics data (age, sex, handedness). We imputed these missing data with CART (classification and regression tree) algorithm. These methods consider any variable with associated missing values, as the target variable and other variables as the predictor ones. Then the method runs a classification or regression tree and predicts a value for missing ones based on the predictor variables. We added these associated parts to the method section of the text and highlighted in turquoise.

11. How did you determine and eliminate redundant independent variables, if applicable?

#11. Since the current data are high-dimension, we cannot apply the usual methods, with providing the sparsity for the input matrix, this elimination has occurred.

12. Are sufficient data available for the model? Assess whether sufficient data are available for a good fit of the model.

#12. Albeit not. The main focus of this research is that the models that we have used for analyses are specifically developed for high dimensional but low sample size data.

13. Please provide the details of training approach, including data augmentation and hyperparameters, if applicable.

#13. Thank you for this comment. We did not run the augmentation process, since we wanted to provide an interpretation based on the original functional connectivity. Regarding the hyperparameters, we mentioned in the Method part that from a collection of parameters, using the cross-validation method, we selected the best one.

14. What technique was used to split the data into partitions?

#14. We have used readily available dataset with predefined partitioning and labels, we also applied the determined categories.

15. You merely separated the data into the train and test sets. What about validation?

#15. Many thanks. Since the current data are high-dimensional low-sample size, we did not consider a separate validation set and performed a cross-validation procedure throughout the train data.

16. Please explain the transferability of this model to the new clinical setting.

#16. Model parameters that have been estimated in the current context (the "estimation context") could be used to create parameters for use in the application context, which might be a new clinical setting. There are probabilities of challenges.

These challenges might be known as temporal transferability and spatial transferability. Within a specific duration, variations in demographic parameters, prevalence and significance or clinical manifestation pattern, are probable. There could be also differences between imaging protocols. Considering the applicability and performance stability of the model, perhaps the present algorithm would require minor modification according to the specified clinical population. These modifications could be performed through techniques such as transfer learning which is expressed as a linear combination of the direct estimation results and the previously estimated parameters.

17. What about the replicability and generalizability of the model? How the details of the model are available for replicability?

#17. Many thanks for this note. Since the current model is fitted to a multicenter dataset with minor intrinsic variabilities in each setting, it could provide an acceptable degree of generalizability and replicability and it can be used and evaluated in every setting of high-dimensionality and specifically when the research group have a vast number of ROIs derived from a limited sample size. The details of the model will be provided for every researcher that request the protocol and modeling features.

18. Please explain about the clinical utility of this model at the end of the discussion.

#18. Thank you. We added this part.

19. Please mention the limitations of the model.

#19. Compared to the deep-learning methods, it may show lower performance.

2 Revision 1

4 comments · 23 Mar 2023 – 27 Mar 2023

Referee 1 Reviewer

ACCEPT 23 Mar 2023

Dear authors,
Thank you very much for the revisions.
Best regards
Reviewer

Seyed-Amirabbas Ahadiat Reviewer

ACCEPT 25 Mar 2023

Thank you for revising.

Associate Editor

ACCEPT 26 Mar 2023

Dear EIC,
The manuscript has been revised successfully, and all questions have been answered by the authors.

Editor-in-Chief

ACCEPT 27 Mar 2023

Dear Author,
Thank you for your explanation. I find your manuscript acceptable.