# RESTING STATE EEG CLASSIFICATION OF CHILDREN WITH ADHD

G. CIODARO

A. JAHANIAN NAJAFABADI

B. GODDE

# HARNESSING THE POWER OF MACHINE LEARNING TO UNDERSTAND THE HUMAN BRAIN.

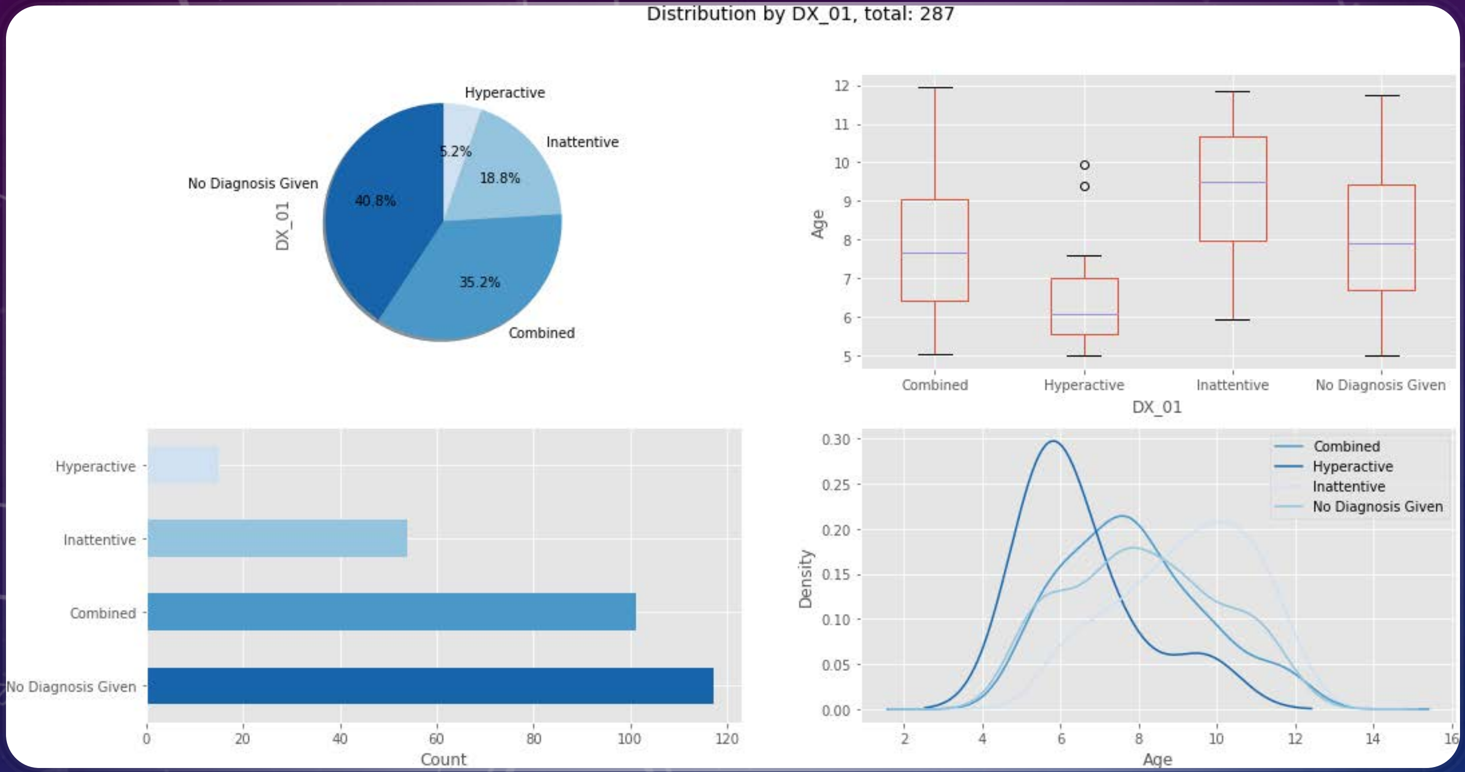
# IMPORTANCE OF  ATTENTION DEFICIT HYPERACTIVITY DISORDER

- One of the most frequent neuropsychiatric diagnoses during childhood.
- It is estimated that at least one child in every classroom could be diagnosed [1].
- Individuals with ADHD may experience difficulties with education, personal relationships, self-esteem, and quality of life [2]
- According to DSM–IV (American Psychiatric Association, 1994), there are three main clinical forms of ADHD: inattentive, hyperactive/impulsive, and combined.

# SUMMARY OF THIS PROJECT.

1. Obtain EEG signals from the HBN dataset.
2. Build 46 data process pipelines (Experiments)
3. Rank every experiment using a test dataset.
4. Select the best two experiments.
5. Research how these two experiments make prediction.
6. Apply statistical tests to develop formal propositions.

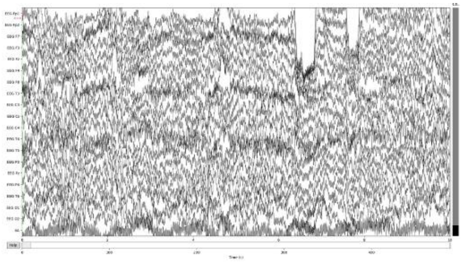
# HEALTHY BRAIN NETWORK (HBN) DATA



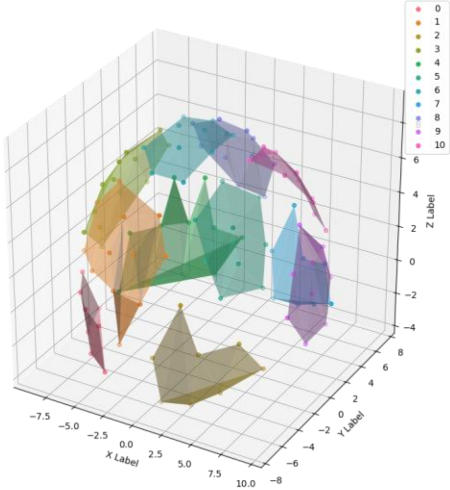

| Primary Diagnosis | Train-Set | Test-Set | Train+Test-set |
|---|---|---|---|
| ADHD-Inattentive Type | 44 | 10 | 54 |
| ADHD-Hyperactive | 13 | 2 | 15 |
| ADHD-Combined Type | 43 | 58 | 101 |
| No Diagnosis Given | 100 | 17 | 117 |
| Total Subjects | 200 | 87 | 287 |

| Spatial-Pre-processing | Feature Extraction | Statistical Learner | Feature Transforma-tion |
|---|---|---|---|
| Clustering K=11 | Relative Power bands | XGB | Polynomials |
| Clustering K=20 | Sample entropy | RF | K-Means Clustering |
| None | Image Representation | CNN Shallow, CNN Deep, resnet18 | None |

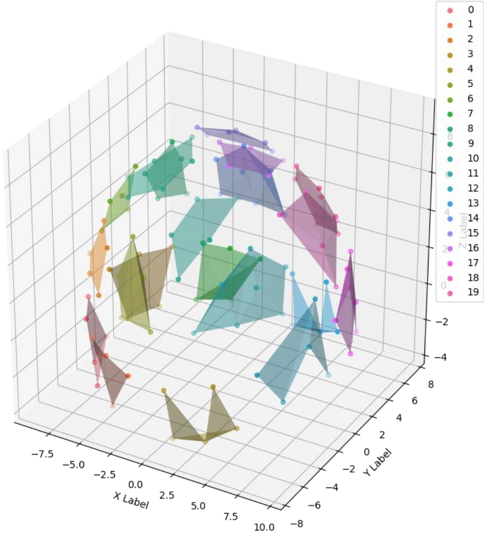
Experimental Space



Raw signal 111 channels



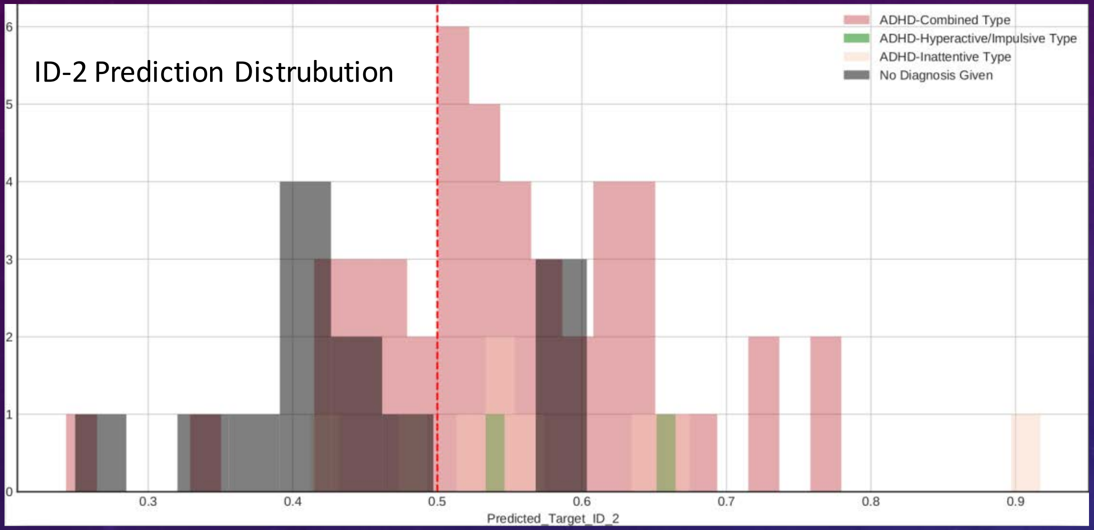

Spatial clustering 11

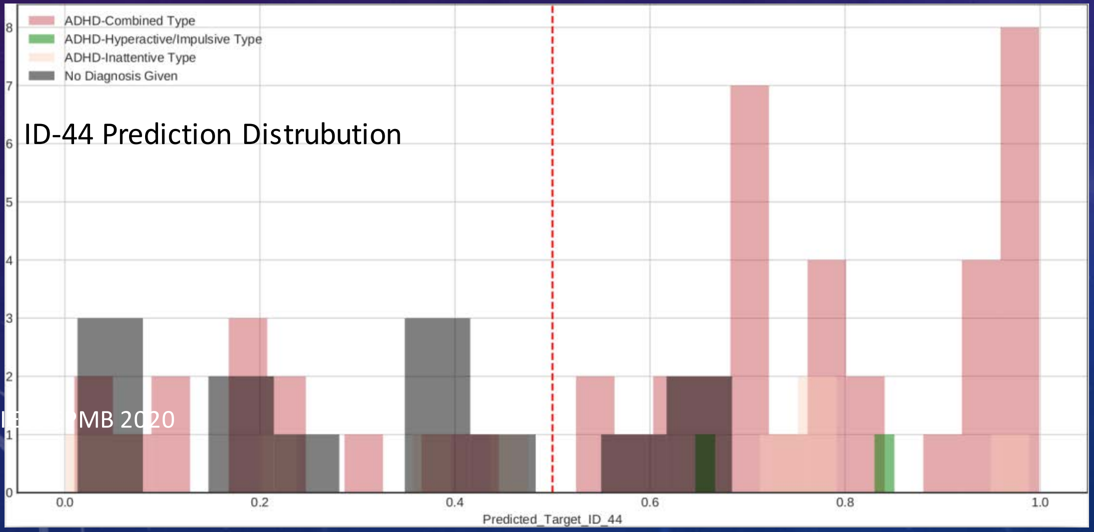


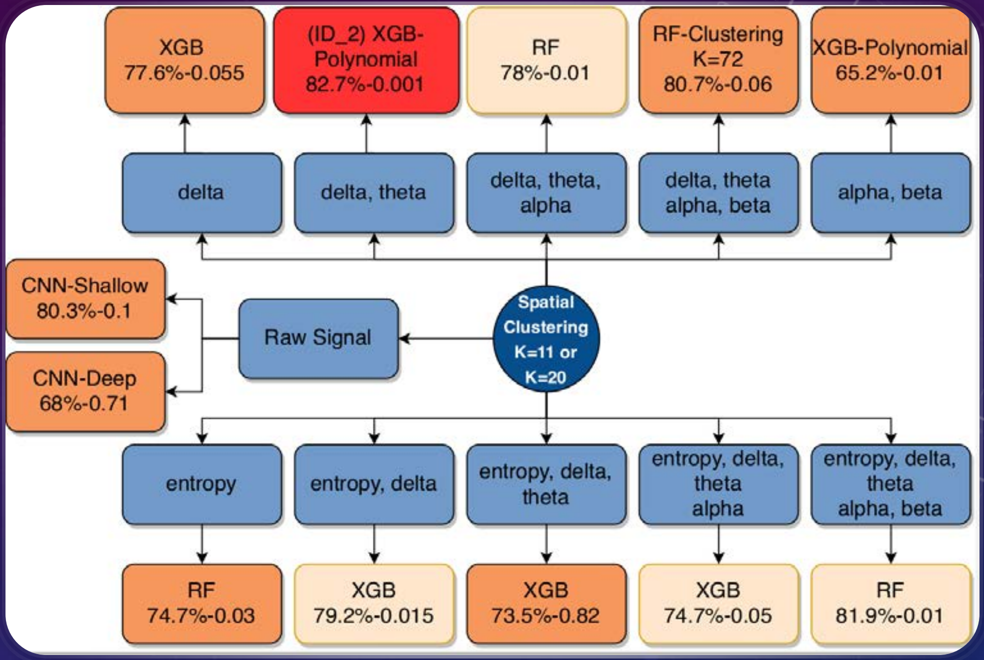

Spatial clustering 20

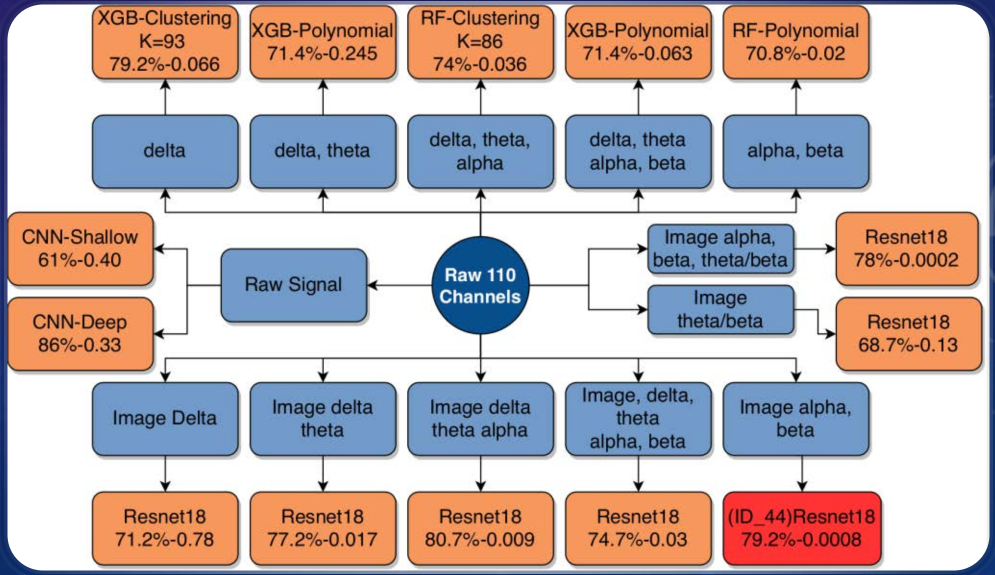
# 2.BUILD 46 DATA PROCESS PIPELINES (EXPERIMENTS)

# BEST TWO EXPERIMENTS



# GENERAL TRAINING STRATEGY

## ID-2

- Defined $D_{intrain}$ and $D_{validation}$ such as $D_{intrain} \cup D_{validation} = D_{train}$
- $D_{validation}$ contains signals from 15 subjects, 9 from ADHD and 6 from Healthy.
- use $D_{intrain}$ to fit $XGB$ on default configuration, and on tree internal transformations $Feature_{Tran} : \{Polynomials, Kmeans_{clustering}, None\}$
- Select the best performing transformation from the previous step using $D_{validation}$.
- Create 200 combinations of the set of parameters shown in table 3.2.
- Use 5-fold cross validation on random splits of the $D_{intrain}$ to select the best combination.
- Train on $D_{train}$, and return the model.

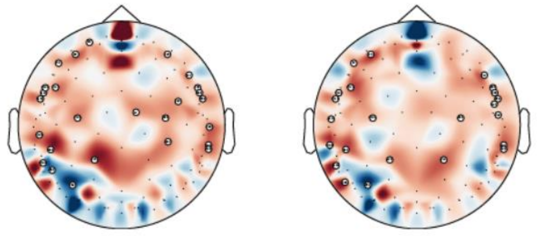
## ID-44

- Obtain an image representation of the inputs.
- Train using Cyclical learning rate technique proposed in [34]. mini-batch Stochastic gradient descent and batch.
- Stops when the negative log likehood reaches values smaller than 0.09.

# ID-44 EXPERIMENT

| | |
|---|---|
| $Spatial_{Pre}$ | None |
| $Feature_{Ext}$ | alpha and beta relative power image representation |
| $Ml_{Tech}$ | 18-Residual-CNN |

Statistically significant electrodes.

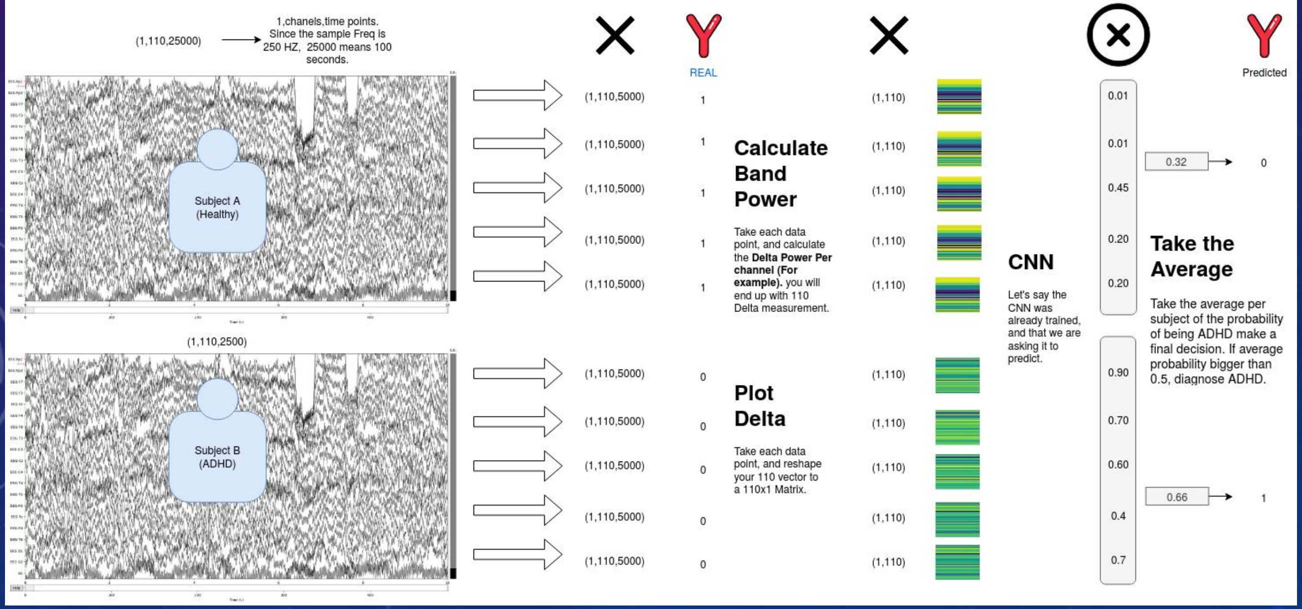


(a) Mean Difference of Alpha power band. ADHD and Healthy.

(b) Mean Difference of Alpha times Beta power band. ADHD and Healthy.

Histogram activation map comparison.

| Compared Histogram | distance ($d(H_1,H_2)$) |
|---|---|
| No Diagnosis Given vs ADHD-Hyperactive | 0.77 |
| No Diagnosis Given vs ADHD-Inattentive Type | 0.64 |
| ADHD-Combined Type vs ADHD-Inattentive Type | 0.54 |
| ADHD-Combined Type vs ADHD-Hyperactive | 0.52 |
| ADHD-Hyperactive vs ADHD-Inattentive | 0.34 |
| ADHD-Combined Type vs No Diagnosis Given | 0.23 |



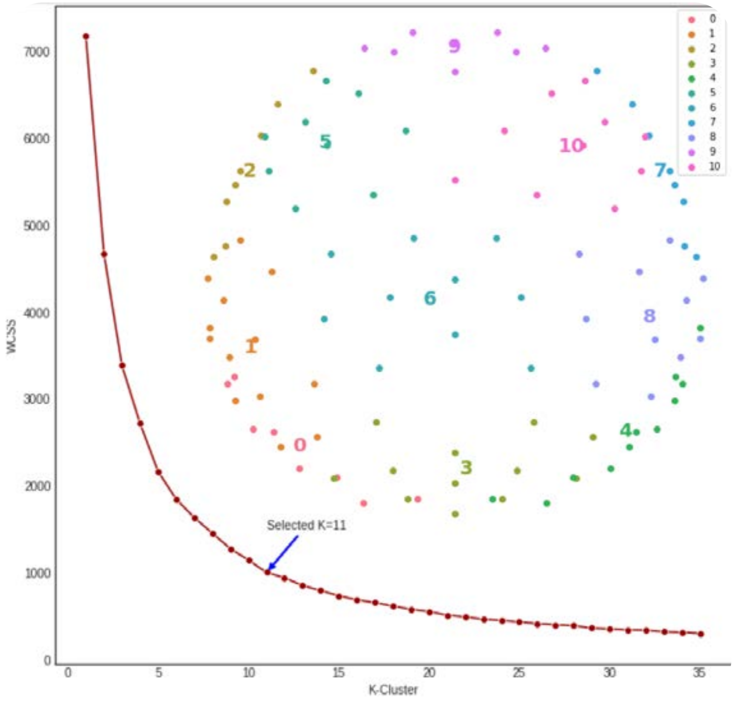

schematic representation

# ID-2 EXPERIMENT

| | |
|---|---|
| $Spatial_{Pre}$ | $Cluster_{11}$ |
| $Feature_{Ext}$ | delta, theta power bands |
| $Ml_{Tech}$ | XGB $\longrightarrow$ Polynomials |

| Feature (Cluster) | p-value | Criterium |
|---|---|---|
| theta (6)-theta (8) (>ADHD) | 0.01 | gain-ranked 1 |
| theta (6)-delta (10) | 0.14 | gain-ranked 2 |
| theta (6) | 0.05 | gain-ranked 3 |
| delta (0) | 0.93 | weight-ranked 1 |
| theta (5)-theta (5) (>ADHD) | 0.04 | weight-ranked 2 |
| delta (2)-delta (10) (<ADHD) | 0.03 | weight-ranked 3 |

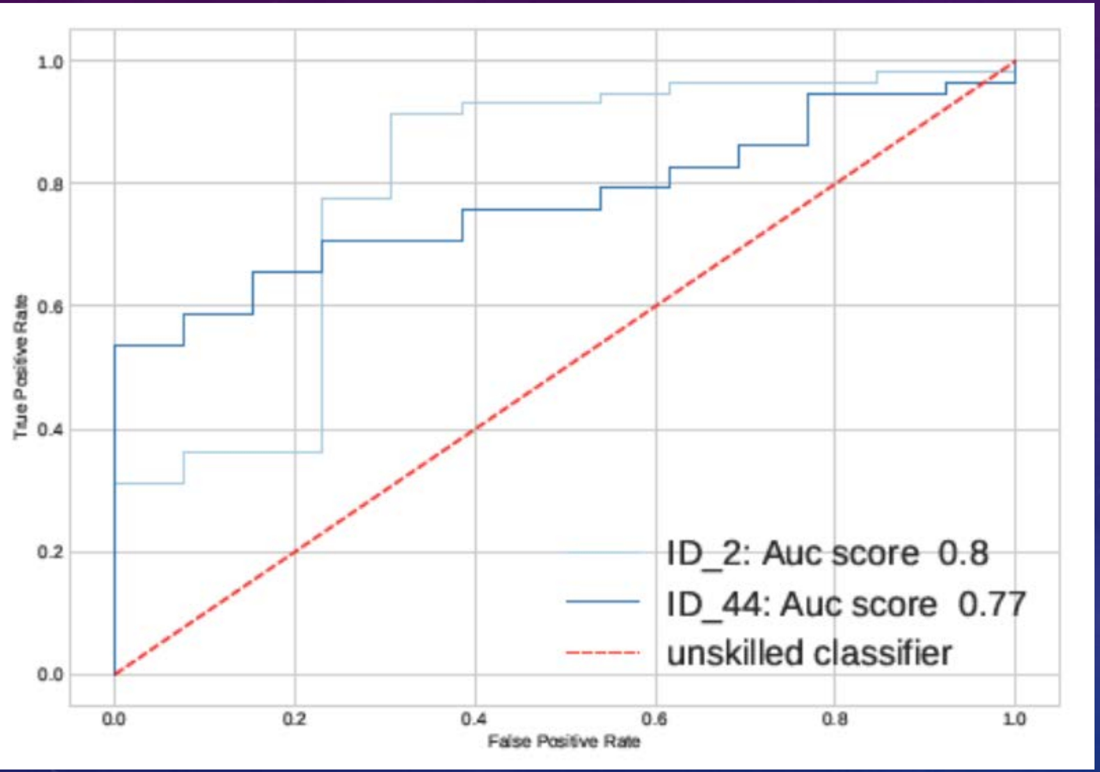
Statistically significant clustered areas.




K means clustering selection.

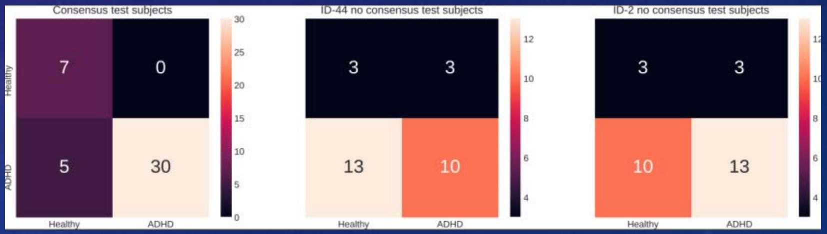
| Cluster | Electrodes |
|---|---|
| 0 | E82, E83, E89, E90, E95 |
| 1 | E96, E100, E101, E108 |
| 2 | E1, E2, E8, E114, E115, E116, E121, E122 |
| 3 | E60, E61, E62, E66, E67, E71, E72, E75, E76, E77, E78, E84 |
| 4 | E45, E50, E51, E57, E58, E59, E64, E65, E69, E70, E74 |
| 5 | E3,E4,E5,E111,E112,E117,E118,E123,E124 |
| 6 | E7, E31, E54, E55, E79, E80, E87, E105, E106 |
| 7 | E25, E26, E32, E33, E34, E38, E39, E44 |
| 8 | E30, E35, E36, E37, E40, E41, E42, E46, E47, E52, E53 |
| 9 | E9, E10, E11, E14, E15, E16, E18, E21, E22 |
| 10 | E6, E12, E13, E19, E20, E23, E24, E27, E28, E29 |

# COMPARISON ID-44, ID-2




| Experiment | Class-1 | Class-2 | p-value |
|---|---|---|---|
| ID_44 | ADHD-Combined Type (63) | No Diagnosis Given (32) | 0.0005 |
| ID_2 | ADHD-Combined Type (55) | No Diagnosis Given (43) | 0.002 |
| ID_2 | ADHD-Inattentive Type (58) | No Diagnosis Given(43) | 0.019 |

Rejected null Hypotheses. Dependency between experiment prediction and ADHD sub-types. We also tested, sexdependency and secondary diagnosis-dependency. For none of them we could reject the null hypothesis.




Confusion matrices on test subjects. from left to right: consensus predictions, ID-44 no consensus and ID-2 no consensus. In general, 93% of the test subjects were correctly classified by either ID-44 or ID-2.

# SUMMARY ID-2

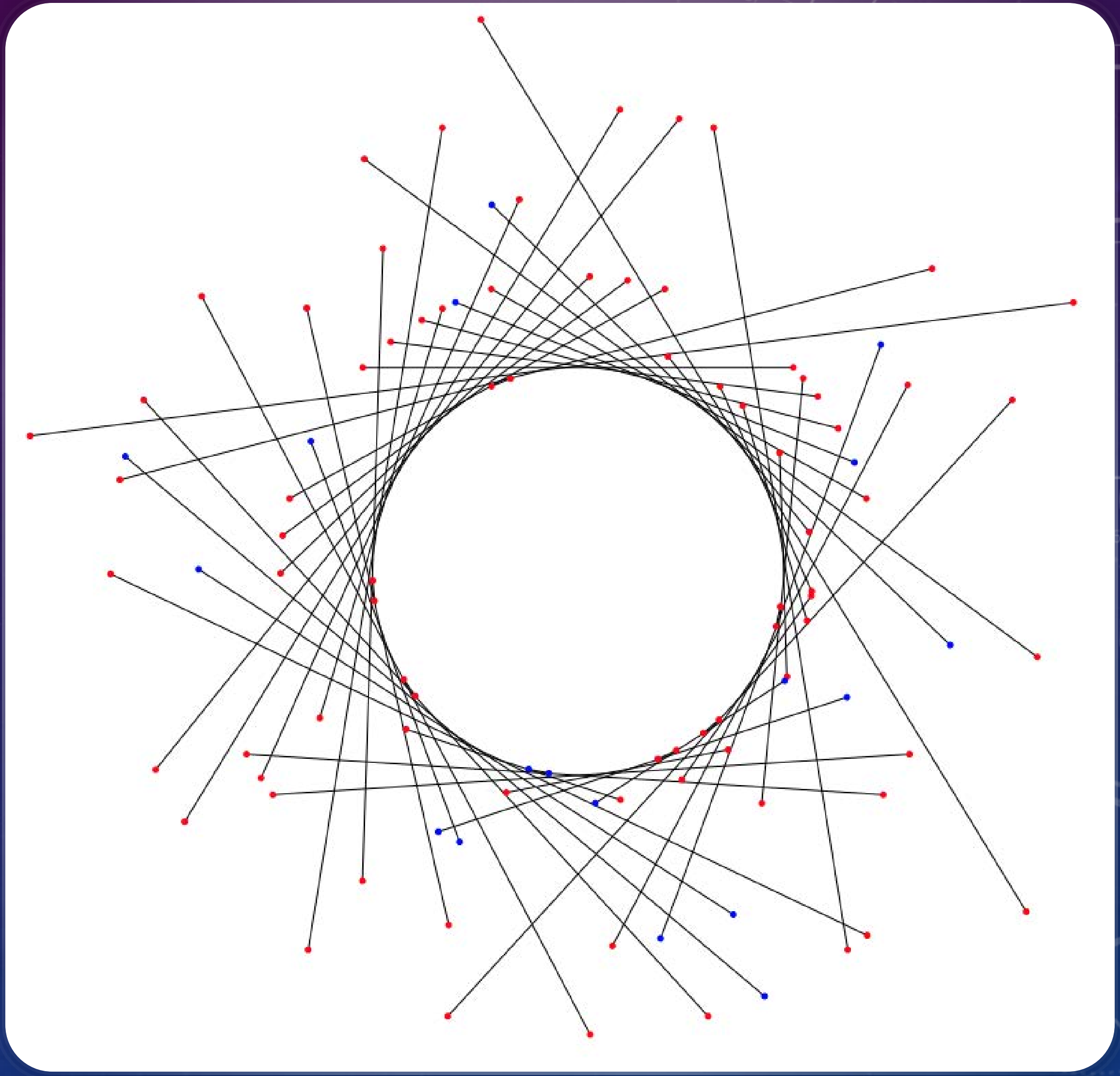
- P.1 . The activation maps used by ID-44(M L tech : Resnet18) are different for predicting ADHD-Combined Type and No diagnosis giving. Histogram comparison.
- P.2. ID-44 is suitable for detecting ADHD-Combined Type and Hyperactive. A subject predicted with probability 0.63± 0.08 Is likely to be ADHD-Combined Type. A subject predicted with probability 0.74 ± 0.198 is likely to be ADHD-Hyperactive Type. Subjects with a predicted probability of 0.32 ± 0.12 are likely to be Healthy.
- P.3. ADHD subjects exhibit symmetrical over activation on general frontal region, with localized over activation on temporal region. Additionally, alpha times beta symmetrical over activation of frontal and parietal area, with localized temporal region.

# SUMMARY ID-44

- P.4: ID-2 is suitable for detecting ADHD-Combined Type and ADHD-Inattentive. A subject predicted with probability 0.55±0.02 is likely to be ADHD-Combined Type. A subject predicted with probability 0.59 ± 0.09 is likely to be ADHD-Inattentive Type. A subject predicted with probability 0.43 ± 0.05 is likely to be Healthy. 95% confidence level.
- P.5: ADHD subjects exhibit similarity under activation of frontal lobe in delta power band. asymmetrical over activation of theta in frontal, central and temporal regions. See significance tested features from subject of XGB most importance features.

# CONCLUSION.

- We proposed a diagnosis pipeline for ADHD using machine learning. Using experiment ID-44 in combination with ID-2, a profile of the subject can be constructed for further treatment and evaluation. Similarity prediction network suggest that some subjects (closely-spaced nodes) exhibit ADHD traits on both spatial relative. alpha-beta interaction, and clustered region delta-theta interaction. On the other hand, further apart nodes offer the opportunity to treat the patient on a pure spatial alpha and beta, directed treatment. We suggest the latter given that the p-value of predictions of ID-44 (0.0008) is smaller than that of ID-2 (0.001). Consensus among the ID-44 and ID-2 experiments was obtained on 57% of the test subjects.
- In both experiments, the prediction distribution did not show enough statistical evidence to suggest a dependency on age or sex. In the case of ADHD subtypes and secondary diagnosis, we believe there are not enough samples to conclude the dependency.



Similarity network of predictions True Class Healthy (Blue) and ADHD (Red). A pair of connected nodes represent a single subject, the distance is proportional to the 10-binned prediction difference (mean:2.30, standard deviation:1.44). It only includes test-subjects predicted in the same class by ID-44 and ID-2. There are a total of 49 subjects and the mean 10-binned prediction difference on subject classified in different classes by ID-44 and ID-2 is 4.37 with standard deviation of 1.83.

# BIBLIOGRAPHY


- [1] Russell A Barkley. Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment (3e éd.). New York: Guilford, 2006
- [2] Kayleah M Groeneveld, Anna M Mennenga, Robert C Heidelberg, Rachel E Martin, Rachel K Tittle, Kyle D Meeuwsen, Linda A Walker, and Elyse K White. Z-score neurofeedback and heart rate variability training for adults and children with symptoms of attention-deficit/hyperactivity disorder: A retrospective study. Applied psychophysiology and biofeedback, 44(4):291–308, 201