EEG computational experiment plan [Apr 12]

discussed Apr 12, 2023, 14:15

The first text is here. The resume was: the goal is accuracy 90 by comparing various models and pipelines.

1. Collect alternative models in code
2. Set a proper pipeline
3. Compare the models
4. Find datasets for augmentation

Set the pipeline

The pipeline from PyRiemann is convenient. However, check the following ones: MNE-Python, PyEEG, EEGTools, Brainstorm, FieldTrip, EEGLAB, NeuroKit, PyEDFlib, NiBabel
to set one which can give the fastest result.
Use additional features 1) general state of the brain, current and historical brain statistics, 2) personal statistics in static, 3) personal data - in time, like facial expression and eyeblinks convenient data representation. Fix the following lists 1) list of data preprocessing, 2) list of feature extraction, 3) list of main models, and 4) list of datasets.

Baseline and basic model

Let the baseline is the PyRiemann class of modes and pipelines. The second baseline is the HTNet/EEGnet models. See the competitions and papers with models comparison (listed in the References).

Data augmentation

Our sets contain about 700 events (objects) for all users per task. To augment the datasets and boost the accuracy, we should use similar open datasets. This strategy played well in the UW-HTNet paper when they used EEG and ECoG data in the same model. See how augmentation from ECoG works is the same target labels. Collect various types of data.
Check if **our signals are OK compared to** signals from the other collections.

Stability of models

The alternative dataset must show the same results on various models to ensure we collect our data normally.
How differs the optimal modes from task to task, from user to user?
Do we need fine-tuning before each user measurement?
Set the distance between the model parameters and structure.

Refine the data

Construct alternative pipelines to prepare datasets according to MNE standards.
Fix the data format. Fix a convenient format to exchange: FIF (MNE), EDF (European), BDF (Biosemi), EGI (Electrical Geodesics), BrainVision, EEGLAB.
Select a proper subset of the electrodes. Use metric analysis, namely:
Rerun and define the functional groups. The brain areas reflect similar electrodes and signals, corrected with some diffusion process (possibly with a strong deterministic part). Represent in cm and milliseconds (keep sample part as in MNE). Measure pairwise similarity and suggest how to plot the distance. Plot the most repeatable (though not the most informative) signal overall surface (and over patients). Advance the feature extraction (the primary was the time reaction). Compare the signals with found patterns and say, is it possible to catalog these things?
Find a model which works with the weighted importance of electrodes.

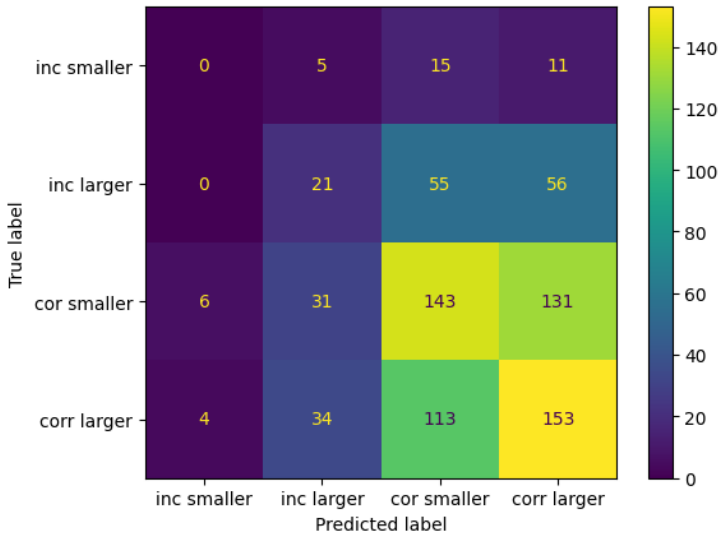
Discuss the experiment strategy

The human can get an accuracy of 90 is a very good point of view.
So accuracy goes for 1) each event, 2) each patient, and 3) human expertise. Describe the exact error measurement and error analysis protocol.
Discuss (from a medical point of view) the hierarchy of 1) block number (we have three blocks), 2) event number and event label with response label (one event matches two or three responses), 3) possible content of the event (type of stimulus) and equivalence of stimuli, 4) how the target label form.

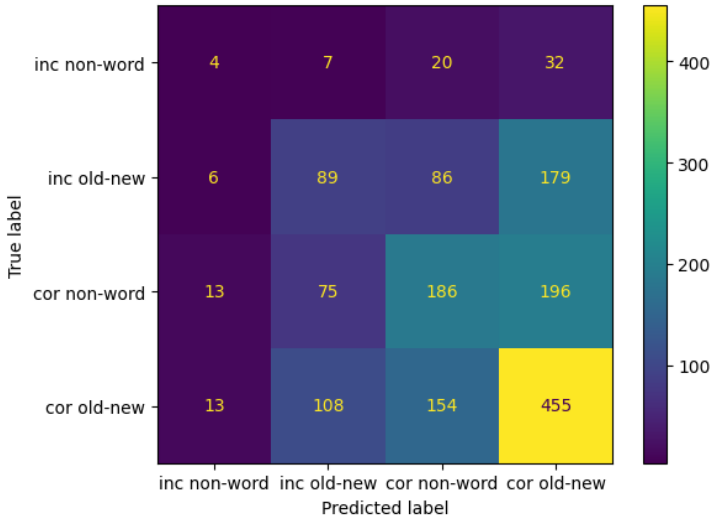
Preliminary models and results

The model pipeline is PyRiemann
XdawnCovariances(n_components), TangentSpace(metric="riemann"), LogisticRegression()

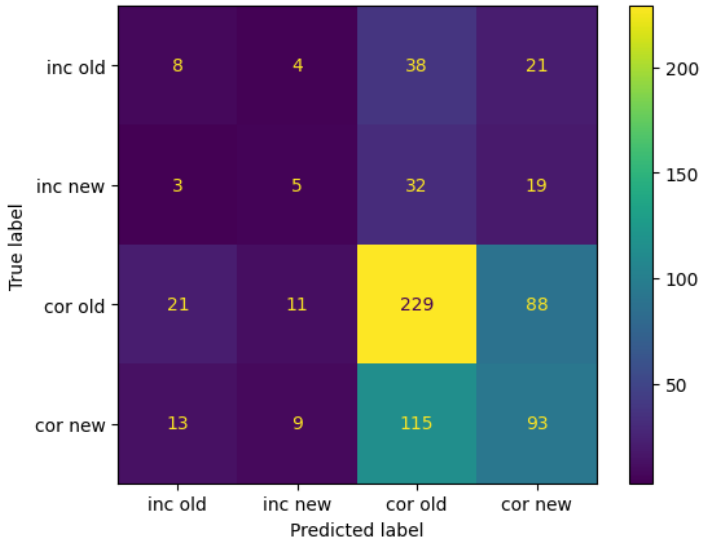
Task 1


inc smaller
inc larger
cor smaller
corr larger
True label
0
5
15
11
0
21
55
56
6
31
143
131
4
34
113
153
inc smaller
inc larger
cor smaller
corr larger
Predicted label
140
120
100
80
60
40
20
0

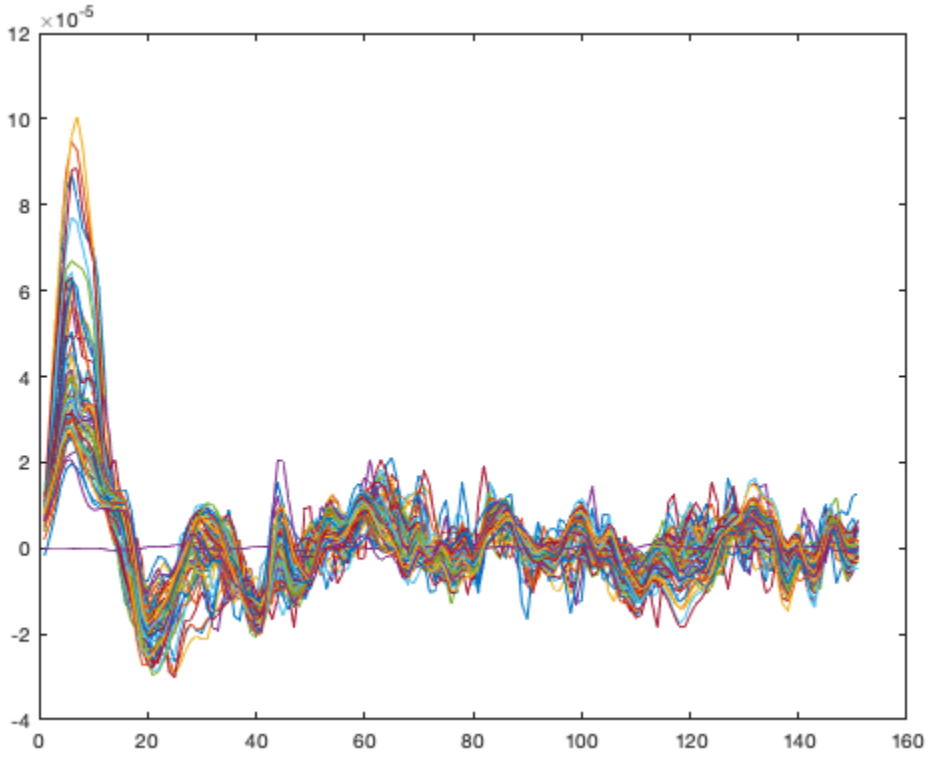
Task 2


inc non-word
inc old-new
cor non-word
cor old-new
True label
4
7
20
32
6
89
86
179
13
75
186
196
13
108
154
455
inc non-word
inc old-new
cor non-word
cor old-new
Predicted label
400
300
200
100

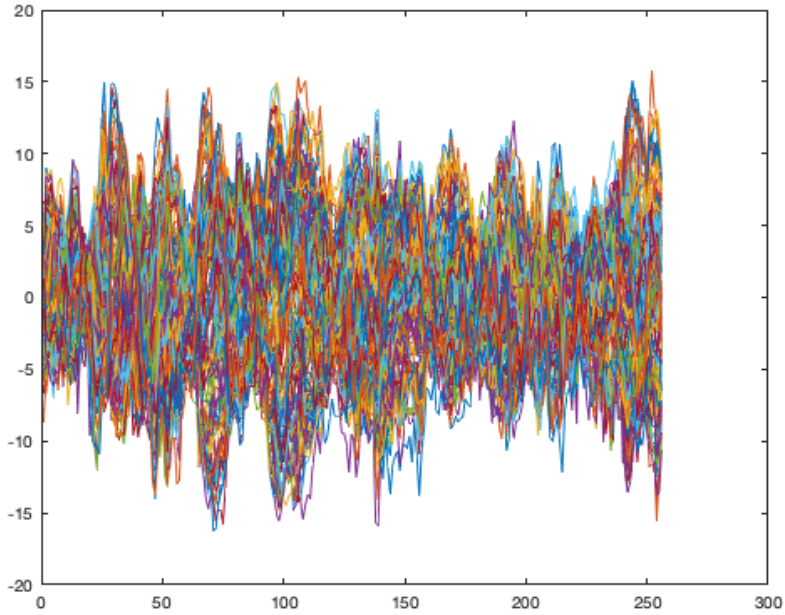
Task 3

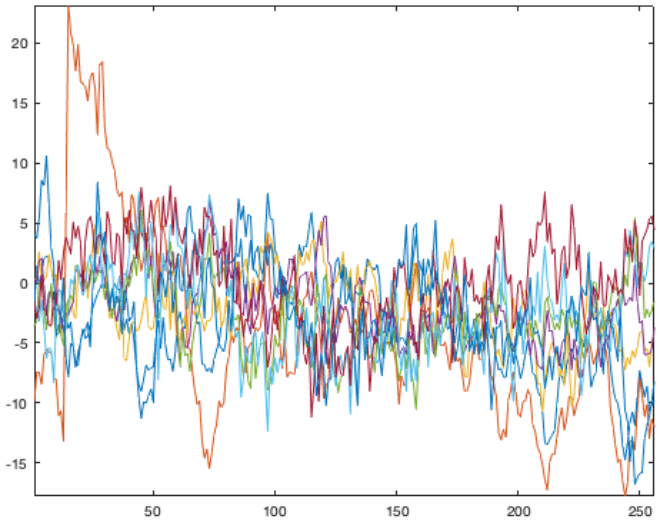


Prepared data (not ours)



Our data





Expected our data (this is not ours)

References (datasets and tools)

Competitions and competition **winners**

https://github.com/pyRiemann/pyRiemann

~~https://www.researchgate.net/figure/EEG-ERP-Portal-User-Interface-Preview_fig7_261194196~~

BCI Competition IV (2008) https://www.bbci.de/competition/iv/#datasets
2020 International BCI Competition https://osf.io/pq7vb/
2020 International brain-computer interface competition: A review
Classifying Numbers from EEG Data - Which Neural Network Architecture Performs Best?
https://paperswithcode.com/search?q_meta=&q_type=&q=EEG See more!!!
ERP Detection. No markup; one has to detect ERP in a plain signal.

EEG signals and models

A Brief Introduction to the Use of Event-Related Potentials (ERPs) in Studies of Perception and Attention, 2013
Why Blinks and Eye Movements Corrupt Your Data - multi-modeling
A similar cross-participant alignment issue occurs with EEG cortical dipoles following blind source separation [35].
Reconstructing ERP Signals Using Generative Adversarial Networks for Mobile Brain-Machine Interface
A Review of Classification Algorithms for EEG-based Brain-Computer Interfaces: A 10-year Update
Computation of the electroencephalogram (EEG) from network models of point neurons
Youtube
1 EEG feature extraction and Machine Learning classification in PYTHON
Meta-Research: Lessons from a catalog of 6674 brain recordings
https://pyriemann.readthedocs.io/en/latest/auto_examples/index.html#classification-of-erp
https://fr.mathworks.com/help/signal/ref/edffileanalyzer-app.html

Continuous and metric approach

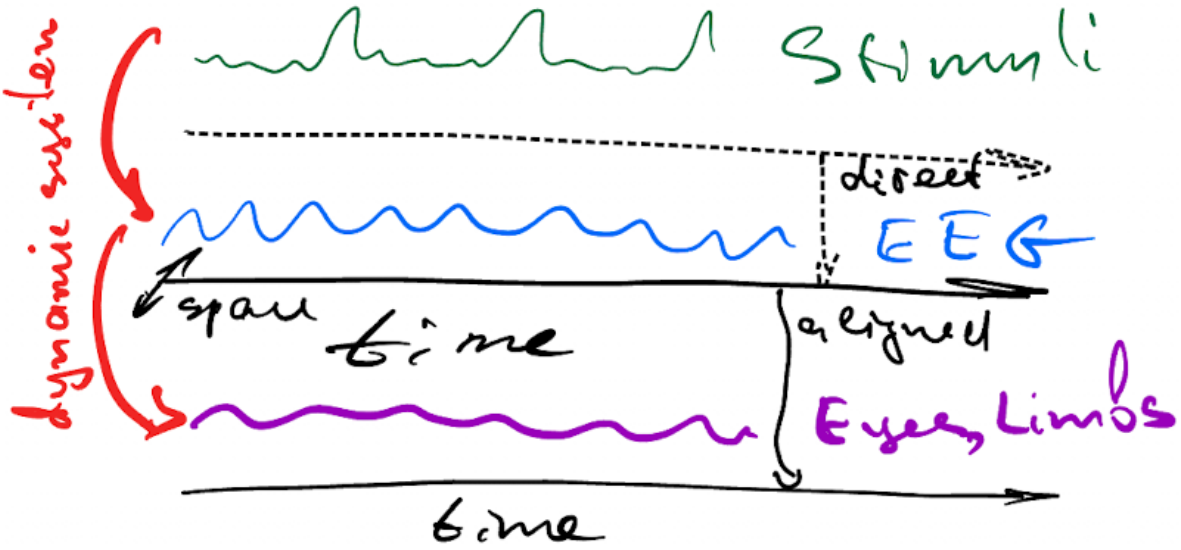
Basic materials

How do we connect N-ODE with DDD (S4 Maxim), with a basis on ECGnet (with PLS/CCA statement)?
—Main—
S4 Maxim et al.
https://github.com/Intelligent-Systems-Phystech/Khristolyubov-BS-Thesis
DDD, D4 UToronto
https://github.com/MrRezaeiUofT/Deep_Direct_Discriminative_Decoder-D4-
The paper https://arxiv.org/pdf/2205.10947.pdf
ECGnet UW
https://github.com/BruntonUWBio/HTNet_generalized decoding
(based on https://github.com/vlawhern/arl-eegmodels)
—Appx—
(Impact PyRiemann to ECGNet)
https://alexandre.barachant.org/papers/conferences/EMBC-2017-Cecotti/
https://scholar.google.com/citations?user=d7ZVePwAAAAJ&hl=fr
Natalia Varenik and Sviatoslav Panchenko
https://github.com/Intelligent-Systems-Phystech/Varenik-MS-Thesis
—Other—

~~A machine learning approach for predicting suicidal thoughts and behaviors among college students~~
~~nature.com/scientificreports~~
~~https://doi.org/10.1038/s41598-021-90728-z~~
~~—Hardware—~~
Commercial devices to record multiview simultaneously
https://www.deliberate.ai/
~~—Intheon's Neuro pipe—~~
https://www.neuropype.io/docs/index.html
https://www.neuropype.io/docs/select_examples/index.html#data-acquisition
https://intheon.io/
~~—Feature engineering—~~
~~Automated Feature Extraction on AsMap for Emotion Classification using EEG~~
~~https://arxiv.org/abs/2201.12055~~
~~—Data sets—~~
SEED Dataset A dataset collection for various purposes using EEG signals
https://bcmi.sjtu.edu.cn/~seed/seed.html
DEAP Dataset A Dataset for Emotion Analysis using EEG, Physiological, and Video Signals
https://www.eecs.qmul.ac.uk/mmv/datasets/deap/
—OpenNeuro EEG—
https://openneuro.org/search/modality/eeg?query=%7B%22modality_selected%22%3A%22EEG%22%7D

Basic models

— Part 0 —
https://github.com/srush/annotated-s4/edit/main/s4.ipynb
~~— Part 1 —~~
https://github.com/vadim-vic/HTNet_generalized_decoding
https://github.com/BruntonUWBio/HTNet_generalized_decoding
https://github.com/bahleg/arl-eegmodels/blob/master/examples/ERP.py
https://github.com/vlawhern/arl-eegmodels/tree/master/examples
https://stackoverflow.com/questions/48905127/importing-py-files-in-google-colab
https://medium.com/analytics-vidhya/importing-your-own-python-module-or-python-file-into-colab-3e365f0a35ec
https://github.com/Alina-Samokhina/MasterThesis/blob/master/experiments/Experiments.ipynb
http://strijov.com/papers/Samokhina2022P300.pdf
https://github.com/Alina-Samokhina/MasterThesis/blob/master/experiments/Experiments.ipynb
https://gin.g-node.org/v-goncharenko/neiry-demons
— Part 2 —
https://atcold.github.io/pytorch-Deep-Learning/en/week08/08-3/
https://medium.com/dataseries/variational-autoencoder-with-pytorch-2d359cbf027b
https://avandekleut.github.io/vae/
https://github.com/Atcold/pytorch-Deep-Learning
https://github.com/Atcold/pytorch-Deep-Learning/blob/master/11-VAE.ipynb



See also miscellaneous

Simon Ladouce, Lukáš Vařeka: Supporting datasets for "Predicting Navigational Decisions Through Visually Evoked P300 Event-Related Potentials Using A Real-World Brain-Computer Interface."
Single-trial P300 classification algorithm based on centralized multi-person data fusion CNN

Datasets collect here

Thinking out loud, an open-access EEG-based BCI dataset for inner speech recognition Nicolás Nieto et al https://www.nature.com/articles/s41597-022-01147-2