Sleep Deprivation classification using BOLD fMRI Data

Alex Fung, Patrick M. Osborne, Tony M. Lee, Viswesh Krishnamurthy

Mainline Github Repo:

Powahowa/ml1030_neuroimaging

Contribute to Powahowa/ml1030_neuroimaging development by creating an account on GitHub.
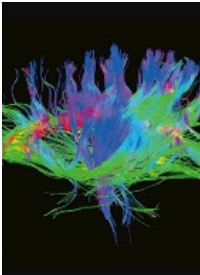
github.com

Preprocessing GitHub Repo:

edickie/ds000201_preproc

preprocessing and feature extraction for ds0000201 pulling the datalad container into envs to version control content...

github.com

Educational York University, School of Continuing Studies:

Certificate in Machine Learning

As the world becomes more automated, more companies than ever are investing in machine learning to keep ahead of the...

continue.yorku.ca

Project Sponsor:

Krembil Centre for Neuroinformatics

The team at the Krembil Centre for Neuroinfomatics is putting today's most advanced technology to work on this...

www.camh.ca

. . .

Background

Team **Data Knights** worked with the Krembil Centre for Neuroinformatics at the Centre for Addiction and Mental Health (CAMH) in Toronto, Ontario, Canada with the goal of identifying machine learning applications in the analysis BOLD fMRI neuroimaging.

The specific dataset and deliverables for this project were arrived at in consultation with David Rotenberg, Operations Director, Krembil Centre for Neuroinformatics, CAMH and Dr. Erin Dickie, Scientist and Assistant Professor, Krembil Centre for Neuroinformatics, CAMH. The goal in selecting a problem and dataset was to provide both useful academic conclusions to the CAMH scientists and to provide a challenging and educational project to the team.

Dr. Dickie studies brain connectivity in people with complex brain disorders using fMRI imaging. Considering this, we are working on a dataset that contains BOLD (blood-oxygen-level-dependent) functional Magnetic Resonance Imaging to capitalize on her area of expertise and to provide results that are useful and relevant to her studies.
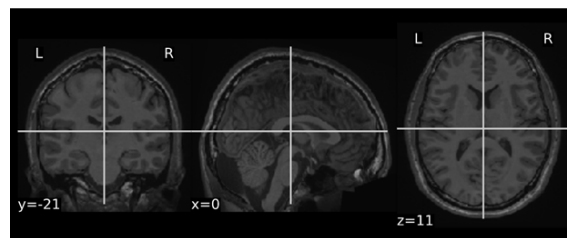
Anatomical MRI Image of a Subject

BOLD fMRI IMAGING

Neuroimaging as defined on Wikipedia, "is the use of various techniques to either directly or indirectly image the structure, function, or pharmacology of the nervous system".

BOLD fMRI stands for **B**lood **O**xygen **L**evel **D**ependent **f**unctional **M**agnetic **R**esonance **I**maging. To explain this, it is necessary to understand blood flow in the brain. In simple terms, as neurons fire they require rapid delivery of oxygen. This results in blood oxygen levels that are higher in areas of the brain that have neurological activity (relative to other areas of the brain).

The fMRI part of the acronym comes in when this data is captured. Oxygenated blood has a different magnetic signature than deoxygenated blood, so the BOLD metric can be used as a kind of contrast in the fMRI images.

Ideally, the BOLD images will help us identify physiological differences in subjects' brains. The BOLD fMRI images do have some downsides— there is a large amount of noise in the images/signal that must be processed out. Some examples are movement, cerebrospinal fluid fluctuations, and the overall "global signal". The BOLD data also only lets us compare areas of the brain relative to other areas—it does not provide an absolute measurement versus a known baseline

The Problem to Solve

Identify sleep deprivation/sleep states and whether these changes can be identified using focused feature engineering and machine learning modelling

1. Develop a binary classifier that identifies whether a subject is sleep deprived or not based on the BOLD fMRI scans.

a. Pre-process/clean the data and engineer any necessary features to support this.

b. Run both traditional and deep learning machine learning algorithms and choose the best classifier based on specific scoring metrics.

2 Engineer frequency-domain features via a Fourier transform on the time series of BOLD fMRI scans. Develop and assess a binary classifier using these features to determine whether certain specific frequencies (noted in the "BOLD signatures of sleep" paper) correlate with different sleep states.

The Dataset

The selected dataset for this project is "*The Stockholm Sleepy Brain Study: Effects of Sleep Deprivation on Cognitive and Emotional Processing in Young and Old*".[i] It is available from the OpenNeuro neuroinformatics database under a creative commons CC0 license (public domain).
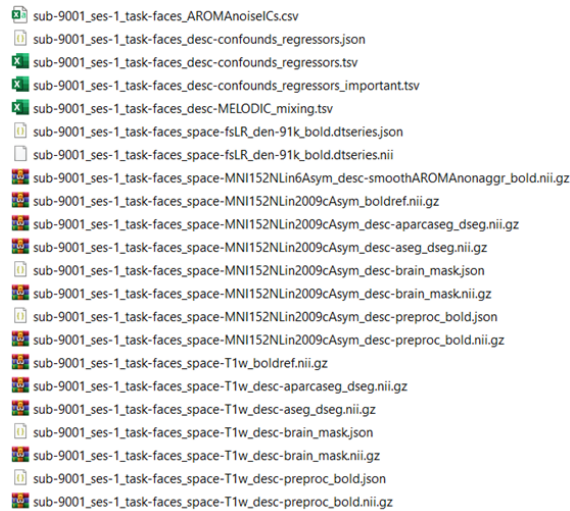
It has been chosen to align with Dr. Dickie's interest in identifying whether there are visible changes in the brain (and in the fMRI images) when a subject is sleep deprived vs not sleep deprived

Link to dataset: https://openneuro.org/datasets/ds000201/

It contains BOLD fMRI imaging data for 86 subjects, with 2 sessions each. In one of the sessions the participant was sleep-deprived (only 3 hours of sleep out their regular cycle). In the other session the participant was not sleep deprived. In each session the participant was asked to perform 5 separate tasks. BOLD fMRI scans were made of each participant during each task and session.

In total, there should be 860 BOLD fMRI scans in the dataset (86 subjects x 2 sessions x 5 tasks). Of these, 430 should represent sleep-deprived subjects and 430 should represent non-sleep deprived participants. This makes the dataset evenly balanced for our purposes.

[i] "The Stockholm Sleepy Brain Study: Effects of Sleep Deprivation on Cognitive and Emotional Processing in Young and Old"— https://openneuro.org/datasets/ds000201/versions/1.0.3

sub-9001_ses-1_task-faces_AROMAnoiseICs.csv
sub-9001_ses-1_task-faces_desc-confounds_regressors.json
sub-9001_ses-1_task-faces_desc-confounds_regressors.tsv
sub-9001_ses-1_task-faces_desc-confounds_regressors_important.tsv
sub-9001_ses-1_task-faces_desc-MELODIC_mixing.tsv
sub-9001_ses-1_task-faces_space-fsLR_den-91k_bold.dtseries.json
sub-9001_ses-1_task-faces_space-fsLR_den-91k_bold.dtseries.nii
sub-9001_ses-1_task-faces_space-MNI152NLin6Asym_desc-smoothAROMAnonaggr_bold.nii.gz
sub-9001_ses-1_task-faces_space-MNI152NLin2009cAsym_boldref.nii.gz
sub-9001_ses-1_task-faces_space-MNI152NLin2009cAsym_desc-aparcaseg_dseg.nii.gz
sub-9001_ses-1_task-faces_space-MNI152NLin2009cAsym_desc-aseg_dseg.nii.gz
sub-9001_ses-1_task-faces_space-MNI152NLin2009cAsym_desc-brain_mask.json
sub-9001_ses-1_task-faces_space-MNI152NLin2009cAsym_desc-brain_mask.nii.gz
sub-9001_ses-1_task-faces_space-MNI152NLin2009cAsym_desc-preproc_bold.json
sub-9001_ses-1_task-faces_space-MNI152NLin2009cAsym_desc-preproc_bold.nii.gz
sub-9001_ses-1_task-faces_space-T1w_boldref.nii.gz
sub-9001_ses-1_task-faces_space-T1w_desc-aparcaseg_dseg.nii.gz
sub-9001_ses-1_task-faces_space-T1w_desc-aseg_dseg.nii.gz
sub-9001_ses-1_task-faces_space-T1w_desc-brain_mask.json
sub-9001_ses-1_task-faces_space-T1w_desc-brain_mask.nii.gz
sub-9001_ses-1_task-faces_space-T1w_desc-preproc_bold.json
sub-9001_ses-1_task-faces_space-T1w_desc-preproc_bold.nii.gz

Pre-Processing

The dataset is provided in a fairly un-processed format. The MRI scans
need to be run through a complex pre-processing routine in order to
clean them up and make them usable.

Dr. Dickie assisted with this pre-processing using the fmriprep
preprocessing pipeline, running on the Niagara Cluster, the most
powerful academic supercomputing cluster in Canada.

This pre-processing involves too many discrete tasks to go into detail
on. Essentially, the pre-processing spatially normalizes BOLD images to
MNI (Montreal Neurological Institute) template, standardizing shape,
position orientation, etc. of so that an x/y/z position in one scan is
directly comparable to another.

In addition to the spatial work, confound/regressor signals were
extracted and initial mask files were generated which we will discuss
later in this article

Data Exploration/the BOLD Files

At this step, we began to use the Nilearn and Nibabel packages to do
further data exploration, cleaning and processing. We will later use
Nilearn (specifically NiftiMasker and NiftiLabelsMasker) for feature
extraction as well.

The BOLD fMRI images are stored on disk in NifTi-format images
(Neuroimaging Informatics Technology Initiative). The NifTi format
essentially stores the images as multi-dimensional arrays, including
specific header/metadata information interpreted by Nilearn and
Nibabel.

As you can see from the GIF below, there 3 spatial dimensions along
the x, y & z axes. As an fMRI is collected over time, this results in a 4D
data structure, with time as the 4th dimension.
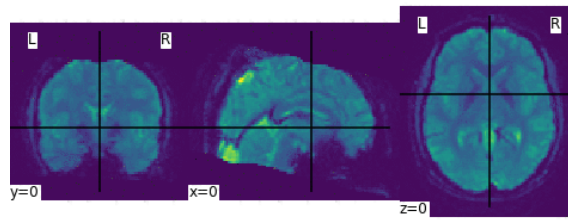
BOLD Image Visualized Over 100 Time Slices (5 minutes)

Pay close attention to the GIF above to see minute changes in the brain signal.

In examining the data, we learned that was significant variation on the scanning parameters used while capturing the fMRI images, which leads to further work in terms of standardization.

The NifTi files range from 150 MB to 450 MB per BOLD image. The full dataset (including just the NifTi images) is about 193 GB.
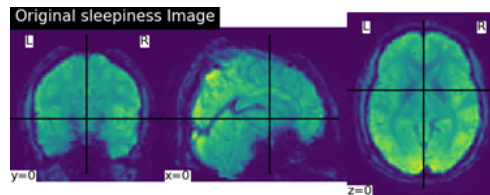
In terms of dimensions, the x, y and z dimensions were not consistent (i.e. the physical space and resolution of the scans did not match). x ranged from 87–104 voxels (3d pixels), y ranged from 103–134, z ranged from 65–81.

The scans were captured over time with one scan per time slice. Each time slice had been split up into slices with 3-second intervals in-between each slice. The number of time slices, t, ranged from 146–400.

filename	dimensions
sub-9001_ses-1_task-faces_space-MNI152NLin2009cAsym_desc-preproc_bold.nii.gz	(104, 123, 81, 165)
sub-9001_ses-1_task-hands_space-MNI152NLin2009cAsym_desc-preproc_bold.nii.gz	(104, 123, 81, 202)
sub-9001_ses-1_task-sleepiness_space-MNI152NLin2009cAsym_desc-preproc_bold.nii.gz	(87, 103, 65, 200)
sub-9001_ses-1_task-arrows_space-MNI152NLin2009cAsym_desc-preproc_bold.nii.gz	(104, 123, 81, 352)
sub-9001_ses-1_task-rest_space-MNI152NLin2009cAsym_desc-preproc_bold.nii.gz	(87, 103, 65, 193)

There were several scans that were missing BOLD fMRI data entirely or had inappropriately small dimensions, so we chose to drop these from the dataset. We were left with a final count of 808 BOLD scans.

As part of the data cleaning, we chose standard time-slice ranges to extract features on, excluding both the beginning and the end of the time series to account for anomalies that sometimes occur at these points in fMRI scans.

Data Cleaning and Preparation

Masking

In order to handle the previously mentioned differences in physical shape/resolution of the BOLD images, we needed to apply a process known as masking in neuroimaging.

As you can see from the image below, the MRI scans in the dataset are not uniform in shape. The fMRI images were captured as the subjects were performing different tasks and the scans were capture with different parameters resulting in different image output shapes. Of the 5 tasks, 3 have a significantly cut-down "viewport" and are missing data at either the upper or lower ranges of the image.

Original rest Image

Original faces Image
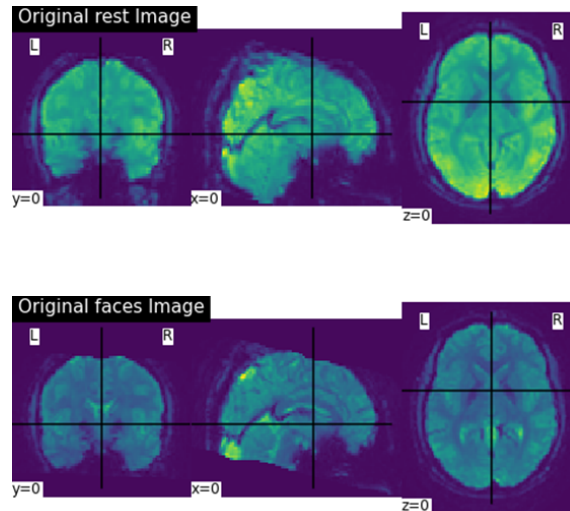
Unfortunately, the differences in the scans are not acceptable to the Machine Learning models we want to run—so we need to find a way to ensure that the images are comparable (apples to apples).

The solution to this is image resampling, masking and mask intersection. Nilearn offers functions to assist with this.

The first step is resampling. Each NifTi file has a specified shape and affine. The shape represents the physical space the image exists in and the affine is a transformation matrix that handles spatial orientation. Resampling the images using Nilearn results in a uniform resolution and transformation matrix across all images, which prepares us for the next step, masking.

A mask is essentially a 3D image containing True/False values, indicating that the image should either show, or be removed at each voxel. The mask has the same dimensions as the original NifTi so a mask can be applied directly over-top of the bold image to remove undesired areas.

In our case, we have a default mask that was computed during pre-processing for each image which removes areas we are not interested in, skull, surrounding matter, etc. See below for an example of a mask. It is important to note that the mask file is 3d, just like the BOLD image, it is just being plotted on a slice for visualization.
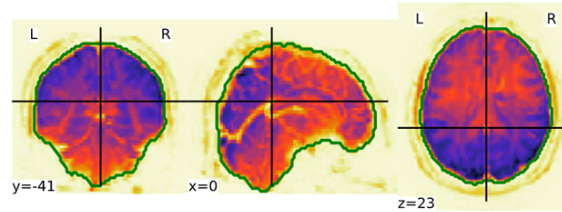
NiftiMasker
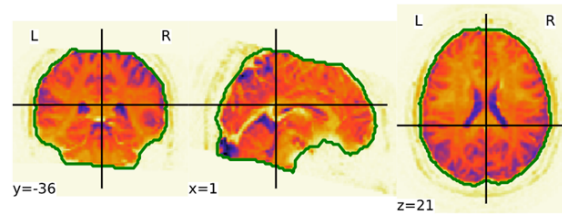
Calculated 3D mask visualized on a plane (Faces Task)

As you can see with the next mask file, the masks differ between subjects.

NiftiMasker

Calculated 3D mask visualized on a plane (Sleepiness Task)

The best way to get a uniform dataset is to calculate the intersection between the masks. Nilearn provides tools in its NiftiMasker function to do this. We have also been using NiftiMasker to visualize these mask images.

We first tried to calculate the intersection of all the masks for Subject 9001, session 1 (the first subject and session).
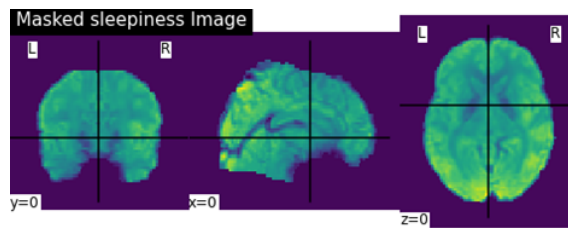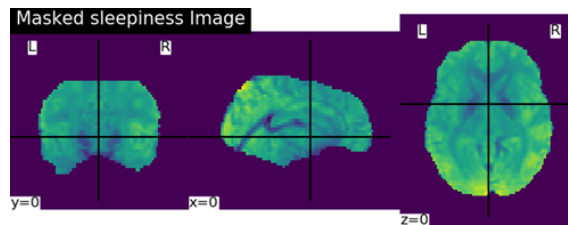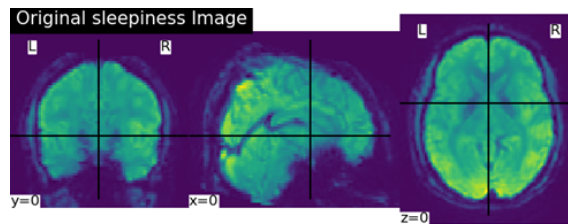
NiftiMasker
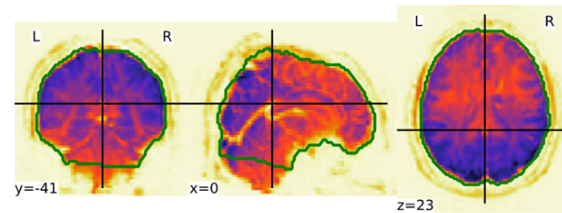
Intersected 3D mask visualized on a plane (Subject 9001—Session 1)

As you can see from above, the intersected mask is now restricted to only the common masked-areas between the given masks. See below for the end-result of this applied on the BOLD image.

Original sleepiness Image
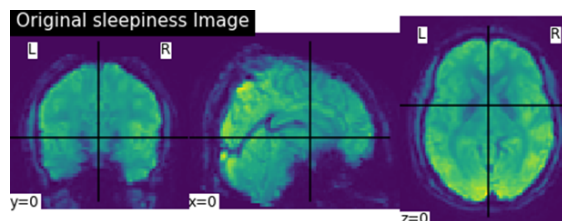
We scaled this process up and calculated an intersected mask on all 808 available BOLD files, resulting in the following.

Feature Extraction

Nifti MASKER—Feature generation was done using Nifti masker

a. Niftimasker in nilearn was used for feature extraction of some of the features.

b. Niftimasker applies data transformation on nifti images and converts them into a 2D array

Raw voxel data

The 3D nifti images are expressed as voxels (volumetric pixel). Each such voxel can be read in as a feature. Each data point would be represented as a 3D array of values. For the project, we chose 60 time slices of the signal and then the signals were either standardized and confounds removed or otherwise.

Standardize

In extracting features using NiLearn, there is a way to standardize the input signal. When standardized, timeseries are shifted to zero mean and scaled to unit variance.

Confounds Masking

Confounds in fMRI are the signals that are non-neuronal. These can be attributed to a subject's movements and other physiological changes during an MRI. We learnt from Dr. Dickie about the 36 confound variables usually used. There are 3 translational confounds, 3 rotational confounds, 3 global signals (csf , white matter & global

signal), their squares, and the derivative of their squares. When the mask is passed over the confounds, it removes them from the dataset.

FFT

All fMRI files are 4D. In other words, there is data in 3D represented over time. Therefore, they can be expressed in the frequency domain using Fourier Transform.

a. Iterate over slices—In this approach, a 3D FFT was done over a series of time slices of the BOLD signal

b. Pass the full BOLD signal—In this approach, a 4D FFT was done over the entire BOLD signal
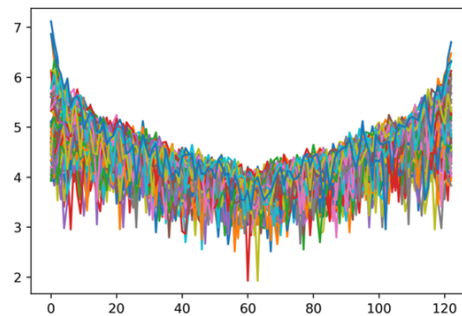
Example of 2-D FFT on the raw voxels data.

Atlas & Connectome features

A Brain Atlas, is a three-dimensional map of the brain structure. It also subdivides the brain into different Regions of Interest (ROI). As there are different map projections of the world, there too are different brain atlases to consider and use. In this project, we decided to use the Harvard Oxford Atlas provided by NiLearn. Specifically, we used the 'cort-maxprob-thr25–2mm' Harvard Oxford Atlas, which contains 48 cortical structural areas.

Using niLearn's niftiLabelsMasker with the Harvard Oxford Atlas, we can extract the time series of each of the ROIs. The time series is aggregated and normalized from the voxels defined from within the ROI.
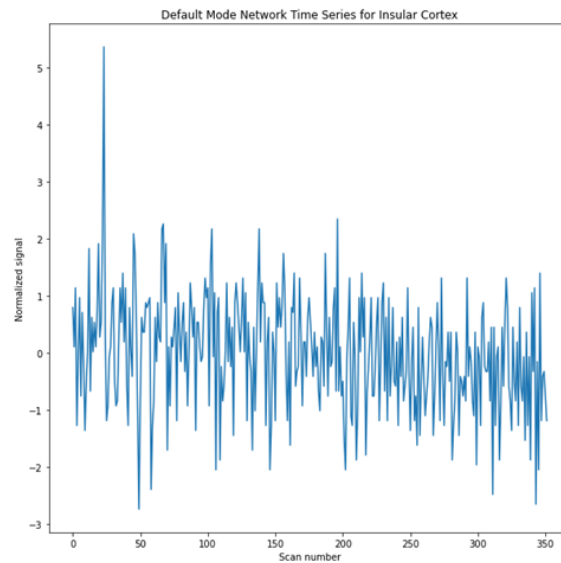
Figure: Example time series plot from the Insular Cortex region

We use the time series calculated from the ROIs to calculate pair-wise connectivity measure between all ROIs. There are three kinds of matrix correlations we could use: correlation, partial correlation, or tangential. The simplest and available connectivity is correlation, which highlights the full connectivity between ROIs. Another type of connectivity is partial correlation, which measures the correlation between pairwise ROIs after removing the effects of the remaining ROIs. Lastly, tangential correlation utilizes tangent space embedding so it can use both correlation and partial correlations.
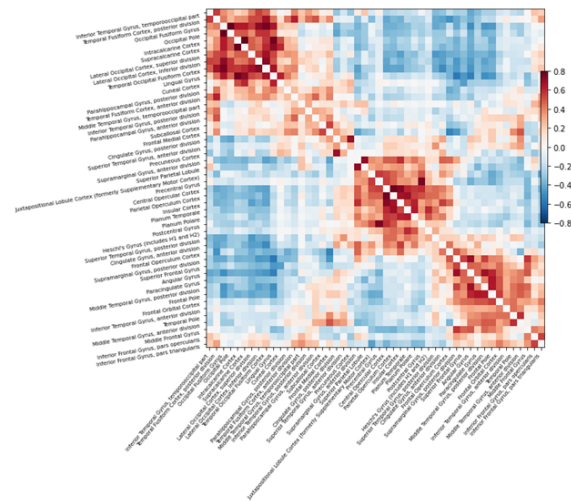

Figure: Example of a correlation matrix of ROIs with connectivity measure using correlation

After calculating connectivity measures between ROIs using correlation, partial correlation, and tangential, the connectivity measures are vectorized then are ready to be passed into ML models for binary classification.

The connectivity measures can also be used to generate a functional connectome. A connectome represents how different ROIs interact with each other in a 3D image of a brain.
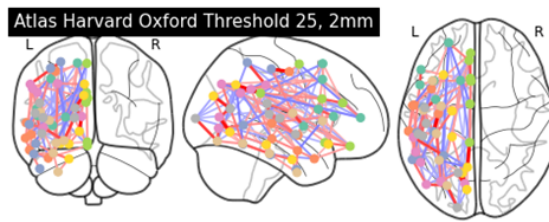
Figure: Example of a functional connectome demonstrating the various partial correlation interactions between ROIs

Role of confounds and 36 confounds

a. Confounds are extracted as part of pre-processing using *fMRIPrep*

By definition, confounds are variables that influences both the dependent variable and independent variable in an experiment, causing a spurious association. They can be considered as variables that identify noise in the signal, in this project.

Read more about confounds and what's generated here:
https://fmriprep.org/en/latest/outputs.html#confounds

Global signal, which is the spatial average of local signals from all cerebral voxels

b. Confounds that we use are the 6 motion parameters (3 translational & 3 rotational) and 3 global signals (csf, white matter & global signal). We also used their first derivatives (which brings it to 18). The squares of both the signal and their first derivatives were also included

Modelling

Traditional ML models were run on 4 combinations of the dataset based on the 'Standardize' and 'Confounds masking' explained above. Standardize = True / False, and Confounds = Masked or Unmasked.

```
cropMask =
NiftiMasker(mask_img="./finalMask/final_resamp_intersected_m
ask_v2.nii.gz", standardize=False)

cropImage =
cropMask.inverse_transform(X=cropMask.fit_transform(image.lo
ad_img(components_df['path'].iloc[i]),
confounds=confounds_selected))
```

Traditional ML models were run on the raw voxel dataset, atlas & FFT. The traditional models used were,

- Linear SVC

- Logistic Regression

- Decision Trees

- KNeighbours

- Gaussian Naïve Bayes

- Adaboost

- GBM

- Ridge Classifier

- Adaboost

- Decision Tree

- Bagging (LinearSVC)

Apart from that, neural networks were built too. We built a CNN and an LSTM, the details of which are seen below.

a. Convolutional Neural Net

- Features: Raw voxels

- 5 fold cross validation

- 4 hidden layers with convolution

- Number of Epochs per fold: 1000
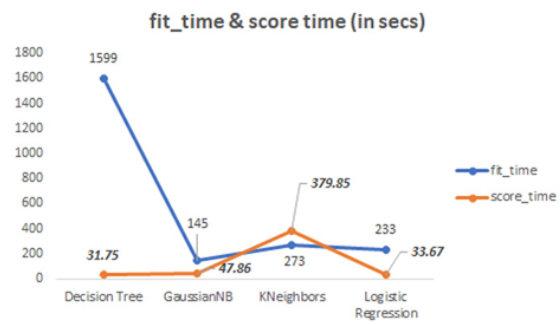
- Early stopping: 100

b. Long Short-term Recurrent Neural Net

- Features: Raw voxels

- 5 fold cross validation

- Internal Units: 100

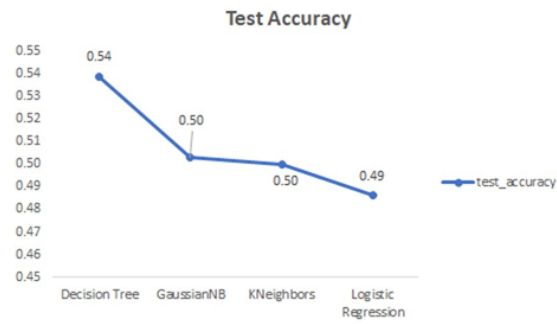- Number of Epochs per fold: 100

- Early stopping: 20

Results & Discussion

ROC, AUC & other metrics comparison by model and identifying final model

Traditional Models

Fit time is the longest with Decision Tree algorithm, as is typical. The other algorithms are reasonably speedy.
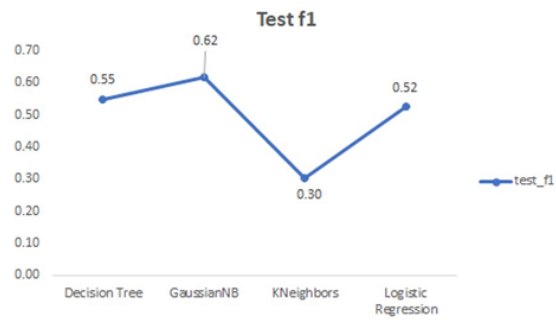
Test Accuracy

In terms of test accuracy, Decision Tree is the best here. However, an achievement of 54% accuracy is only slightly better than chance. We don't believe that this is a good classifier.
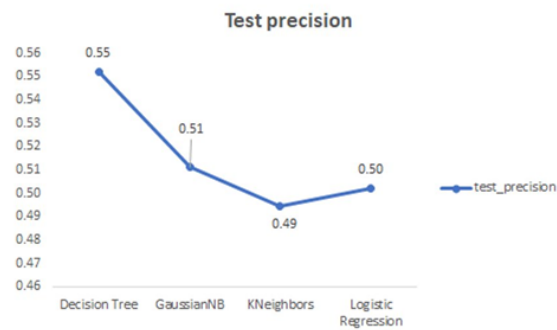

Test precision

Test precision, or the ability of the model to identify *only* the relevant data points, is not good across the board.
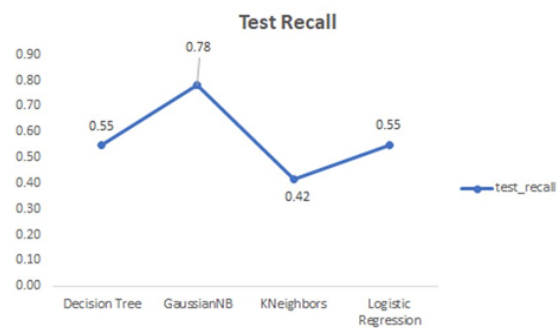
Test Recall

Test recall, or the ability of the model to find *all* the relevant cases within a dataset, shows Gaussian Naïve Bayes is strongest; everything else is bad.

Test f1

The F1 score is a function of precision and recall. Gaussian Naïve Bayes did the best here, however, keep in mind the accuracy isn't good, and the ROC AUC below for GNB doesn't paint a good picture.
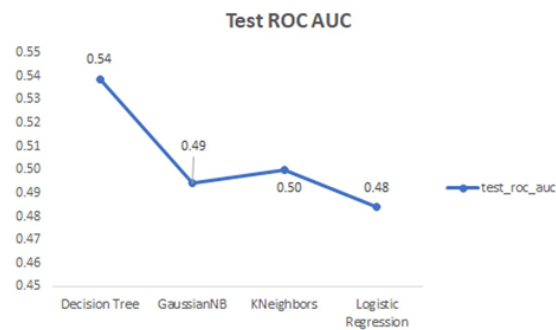
Test ROC AUC

A test ROC AUC that approaches 1.0 is a good indicator of a good classifier. In this case, nothing scores over 0.54 which means the classifiers are bad at predicting.

Connectome Features

CONNECTOME, Standardize = True, Confounds Masked
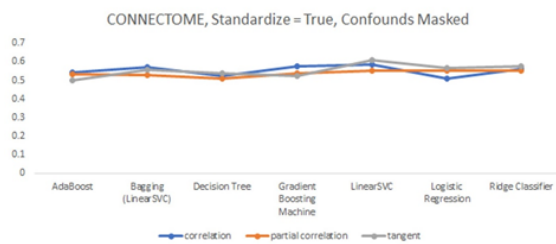
Accuracy using more algorithms is shown above for connectivity features. The highest achieved was 58% using Linear SVC on tangent connectivity.

CNN

- Average Train 5 Folds Accuracy: 1.0

- Average Val 5 Folds Accuracy: 0.5333256721496582

LSTM

- Average Train 5 Folds Accuracy: 0.5715012550354004

- Average Val 5 Folds Accuracy: 0.4703243374824524

Future Work

Despite the lack of promising results generated by the ML models on the generated features, we believe that further exploration in the creation of other features, as well as refining our current features, can improve the ML models' ability to classify sleep deprivation amongst fMRI and their BOLD signals.

Firstly, we believe that exploring methods that may reduce noise from the signals we extracted from ROI. When manually exploring the connectivity matrices of certain NIFTI images, there was visibly still a lot of noise affecting the pair-wise correlations between ROI. An example of the connectivity matrix is given in the figure below; as one can observe, the majority of the correlation matrix results were between 0.8 and 0.4 correlation score, thus indicating that the images were not cleaned further enough, despite using masks and confounds previously to clean up the signal. Therefore, we believe there is still some progress to be made in further cleaning up the signal-to-noise ratio when extracting signals from ROIs.
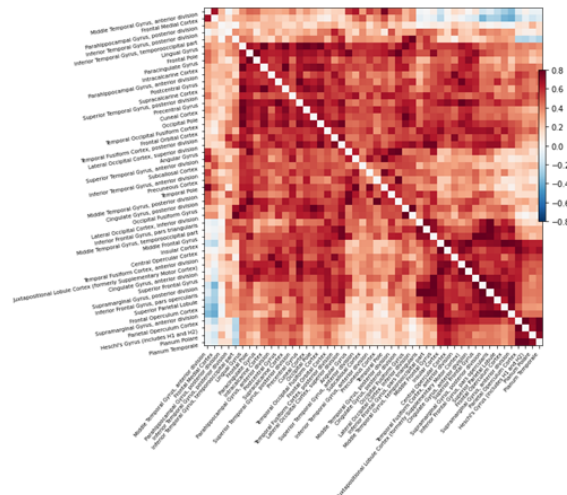

Figure: Sample Correlation Matrix showing the lack of variance in correlation score amongst pair-wise correlations between ROIs

Secondly, a more detailed and in-depth analysis of which ROIs are better correlated to sleep deprivation would be useful in feature engineering, as we could selectively choose which ROIs to generate features from, thus reducing the number of useless features being input into the model. For example, academic literature has shown that sleep-deprived subjects have reduced functional connectivity in ROIs of the Default Mode Network, and less anti-correlation between aforementioned ROIs of the Default Mode Network, and the ROIs of the task-positive network; in layman's terms, this means amongst sleep-deprived subjects, their ROIs in the Default Mode Network were less correlated with one another, and the ROIs in the Default Mode Network were more correlated with the ROIs used for the specific task (certain tasks trigger different amplitudes of signals in different ROIs). In addition, academic journals have also pointed out that sleep deprivation also has an effect on the variability of the global signal. Doing these kinds of investigations would allow us to further identify which ROIs are not of interest to the subject of sleep deprivation, as well as identify what other features we could generate to improve model performance.

Lastly, a deep dive into exploring FFT would be useful for future works. Although we were unable to get a model running on FFT features due to their size, we could reduce the dimensionality of FFT features by simply aggregating and standardizing the FFT by simply extracting them from ROIs rather than on a per voxel basis. Certain frequency wavelengths have been found to be correlated to different stages of sleep, although it is not known if they are present in sleep-deprived subjects during micro-sleep intervals, should they exist.

References:

- https://www.open.edu/openlearn/body-mind/health/health-sciences/how-fmri-works

- https://brainder.org/2012/09/23/the-nifti-file-format/

- https://www.coursera.org/lecture/neurohacking/the-nifti-format-pKtlS

- Gorgolewski, K.J., Auer, T., Calhoun, V.D., Craddock, R.C., Das, S., Duff, E.P., Flandin, G., Ghosh, S.S., Glatard, T., Halchenko, Y.O., Handwerker, D.A., Hanke, M., Keator, D., Li, X., Michael, Z., Maumet, C., Nichols, B.N., Nichols, T.E., Pellman, J., Poline, J.-B., Rokem, A., Schaefer, G., Sochat, V., Triplett, W., Turner, J.A., Varoquaux, G., Poldrack, R.A., 2016. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Sci Data 3, 160044

- "Erin Dickie (She/Her) Scientist, Assistant Professor, Krembil Centre for Neuroinfomatics, Centre for Addiction and Mental Health, University of Toronto" https://ohbm.github.io/osr2020/speakers/erin_dickie.html

- "The Stockholm Sleepy Brain Study: Effects of Sleep Deprivation on Cognitive and Emotional Processing in Young and Old"— https://openneuro.org/datasets/ds000201/versions/1.0.3

- "BOLD signatures of sleep"— https://www.biorxiv.org/content/10.1101/531186v1.full

- https://radiopaedia.org/articles/bold-imaging

- https://www.nature.com/articles/s41597-019-0052-3

- https://computeontario.ca/niagara-supercomputer/

- https://www.analyticsvidhya.com/blog/2019/09/step-by-step-deep-learning-tutorial-video-classification-python/

- https://nilearn.github.io/modules/generated/nilearn.masking.intersect_masks.html

- https://nilearn.github.io/modules/generated/nilearn.input_data.NiftiMasker.html#nilearn.input_data.NiftiMasker

- https://nipy.org/nibabel/images_and_memory.html#using-uncache

- https://nilearn.github.io/manipulating_images/masker_objects.html

- https://fmriprep.org/en/latest/outputs.html

- http://www.leixulab.net/paper/2011HandbookfMRI.pdf

- https://nilearn.github.io/auto_examples/plot_decoding_tutorial.html

- https://nilearn.github.io/auto_examples/03_connectivity/plot_signal_extraction.html

- https://nilearn.github.io/auto_examples/03_connectivity/plot_group_level_connectivity.html

- https://fmriprep.org/en/latest/outputs.html#confounds

- https://neurostars.org/

- https://github.com/edickie/ciftify/blob/master/ciftify/bin/ciftify_falff.py

- https://pubmed.ncbi.nlm.nih.gov/29278773/

- https://neurostars.org/t/confounds-from-fmriprep-which-one-would-you-use-for-glm/326/19

- https://www.imaios.com/en/e-Courses/e-MRI/The-Physics-behind-it-all/K-space

- https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases

- https://www.nature.com/articles/s41598-017-09744-7#:~:text=Sleep%20deprivation%20reduces%20default%20mode,1745%E2%80%931751%20(2012).&text=Decreased%20Thalamocortical%20Functional%20Connectivity%20after,8%2C%20e78830%20(2013)

- https://pubmed.ncbi.nlm.nih.gov/31072560/

- https://www.sciencedirect.com/topics/neuroscience/default-mode-network

- [1] "The Stockholm Sleepy Brain Study: Effects of Sleep Deprivation on Cognitive and Emotional Processing in Young and Old"—https://openneuro.org/datasets/ds000201/versions/1.0.3