Decoding Brain Dynamics: Brain Activity Patterns Predict Nature of the Stimulus

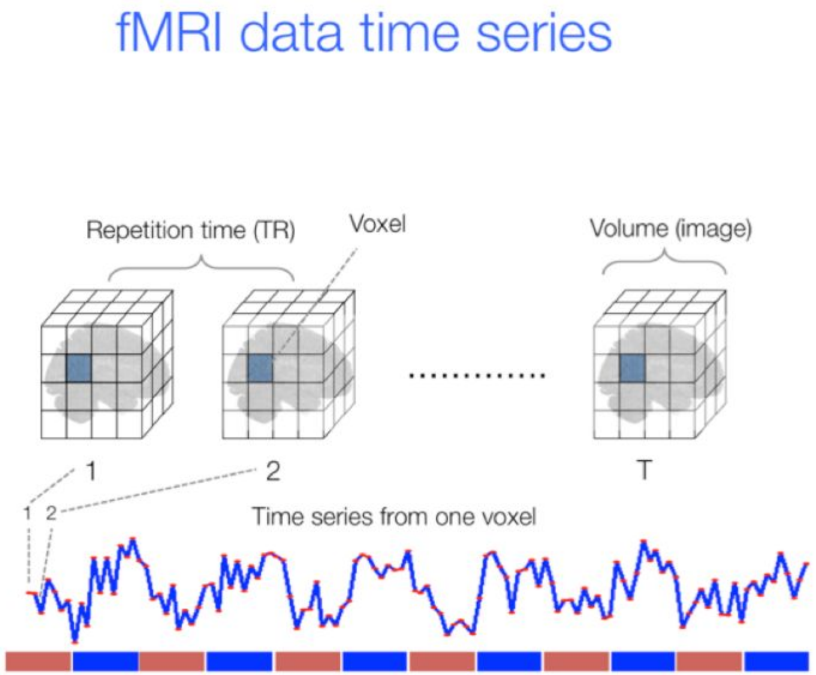
Chirag Limbachia
Capstone Project 2
December 6, 2020

Purpose/Motivation

- Different stimuli evoke different brain response, which enables an organism to respond to a particular stimulus in a particular way.
- Interesting question about the brain-stimuli relationship is: can one predict the stimuli based on the evoked brain response?
- In this project, dynamic brain activity patterns of human subjects, captured by functional magnetic resonance imaging (fMRI) scanner, in response to approaching and retreating threats is used to predict the stimulus.

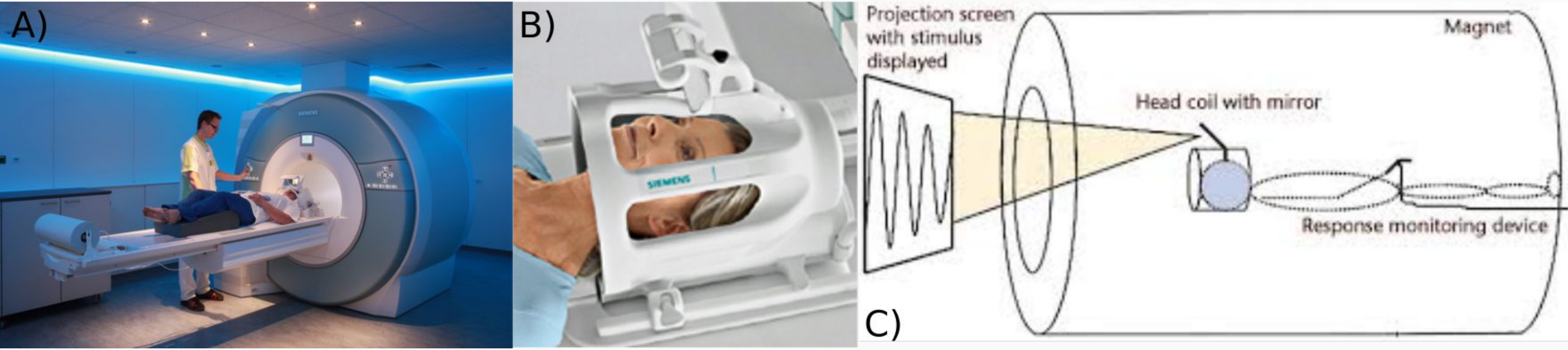
What is fMRI data?

- fMRI is a timeseries data.
- Like a movie of the brain.
- Each frame is a 3D volume that represents a timepoint.
- Unit of the 3D volume is a voxel (analogous to 2D image pixel).
- Dimensions of a voxel is generally 3mm x 3mm x 3mm.
- Rate at which fMRI data is collected is commonly referred to as repetition time (TR).
- Commonly employed TR is 0.5-2 seconds per volume. TR = 1.25 seconds for the current study.



Typical fMRI Experiment Setup

- Participant views visual stimuli on a projection screen via a mirror mounted to the scanner’s head coil.
- Head coil measures the brain signal.



Experimental Procedure/Data Collection

- fMRI data was collected from 61 participants.
- Forty-five minutes of fMRI data was collected on every participant (Six 7.5 minutes long fMRI sessions).
- Visual stimuli was presented on the projection screen.
- Two circles moved around randomly on the screen.
- When circles collided, a mild but unpleasant electric shock was delivered to the index and middle fingers of participant’s left hand.
- Shock delivery was only meant to induce fear of circle collision.
- Several “near-miss” events occured at random times during the experiment.
- Near-miss events are defined as those when the circles approach each other at least for 8 seconds, come very close (i.e., distance less than 1.5 times the circle diameter), but miss and start to retreat, again, at least for 8 seconds.
- Near-miss events were included to investigate brain’s dynamic response to approaching and retreating threats.
- Demo

Data Cleaning

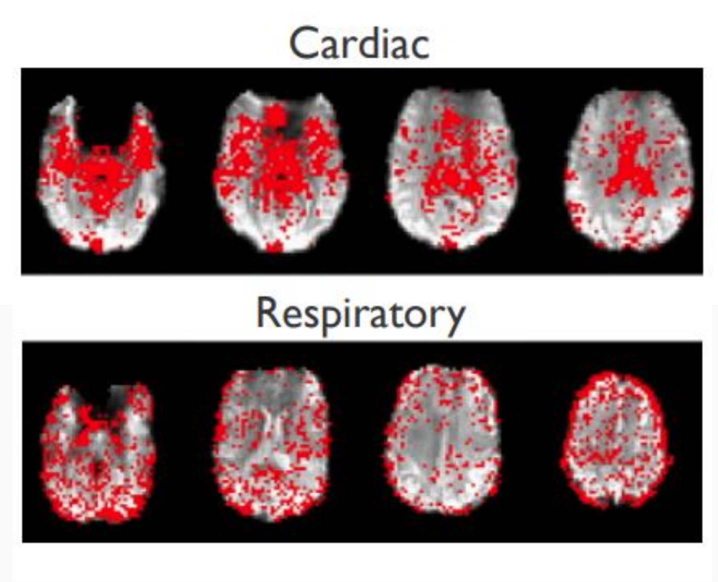
- fMRI timeseries data suffers from a lot of unwanted signals (noise) such as,
  - Physiological signals associated with respiratory and cardiac cycles.
  - Head motion.
  - Scanner noise (drifts).
- These noise were filtered out form the data using ICA implemented in fMRI data analysis software, FSL (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/)

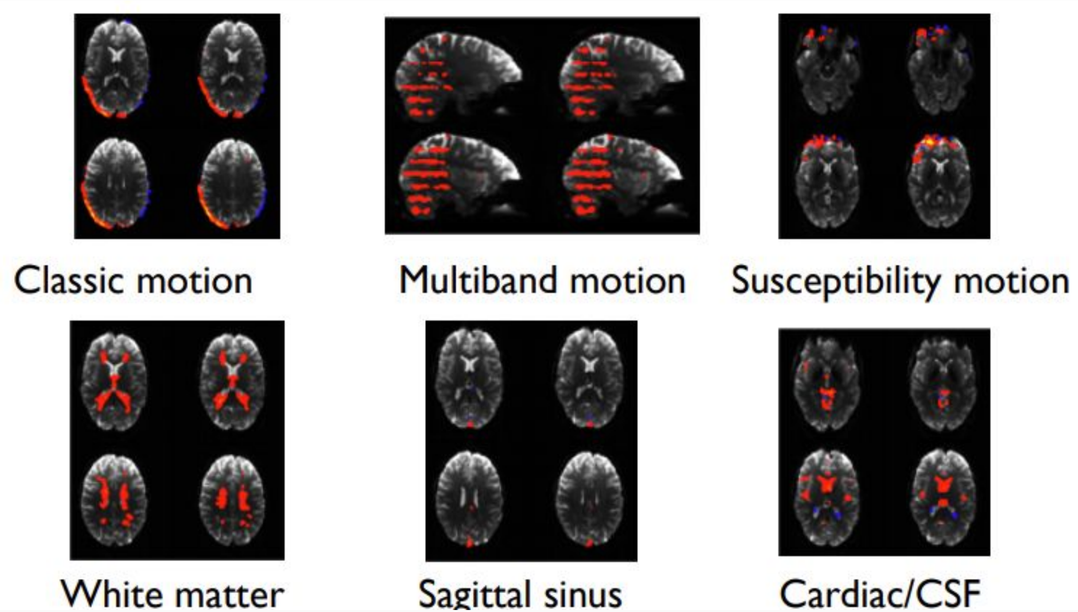
ICA Denoising

- ICA decomposes the data into independent spatio-temporal components
- Based on the spatial and the associated temporal pattern, components are classified into noise or signal
- Components identified as noise are regressed out

ICA Denoising

- These are some examples of most commonly encountered noise.



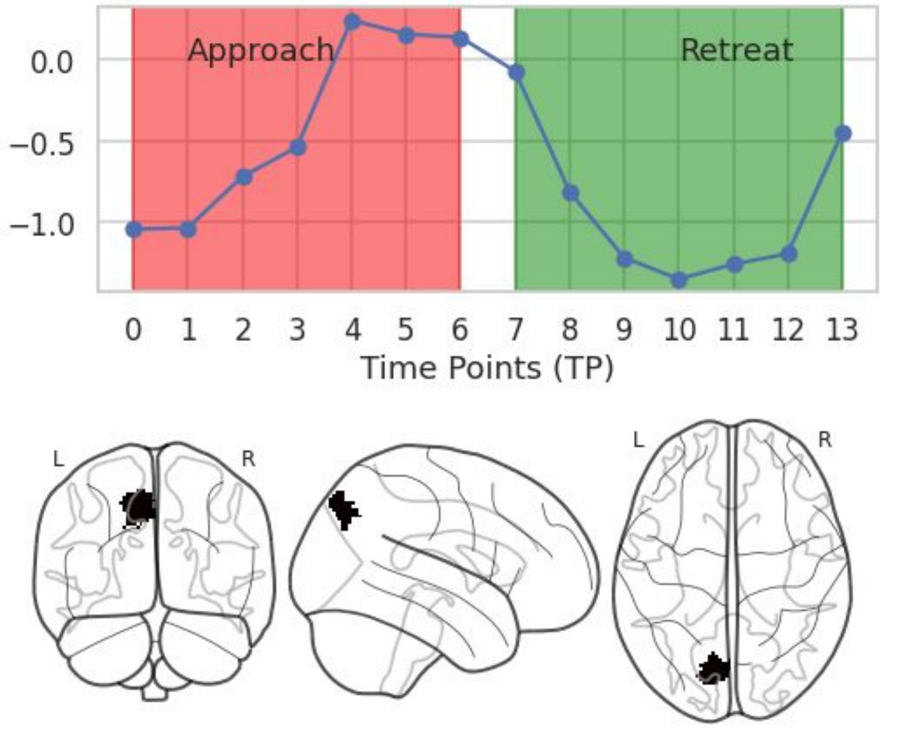


Feature Selection

- Out of the 45 minute long timeseries data, “segments” corresponding to near-miss events were extracted
- Each segment was 14 timepoints long (14 x 1.25 = 17.5 seconds): first 7 timepoints corresponded to approach, and later 7 to retreat
- Every participant had 46 segments. That is, approximately 13 out 45 minutes of data was used to train and test the predictive model

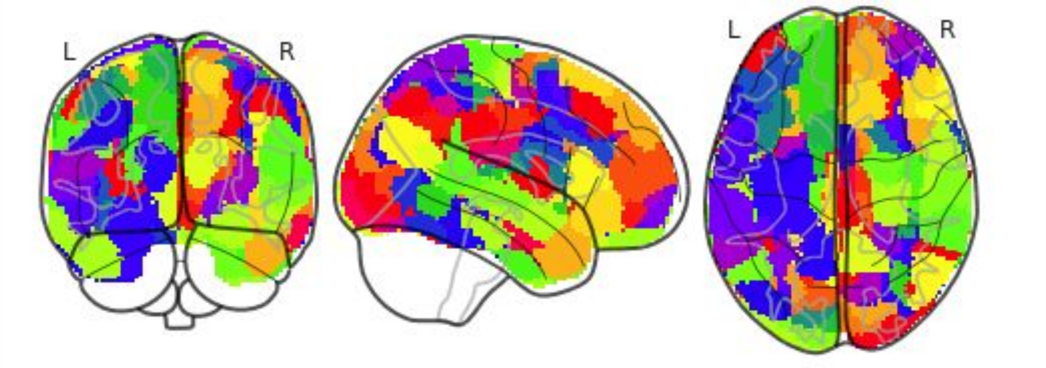
Feature Selection

- An example of a representative segment from the parietal cortex of the brain



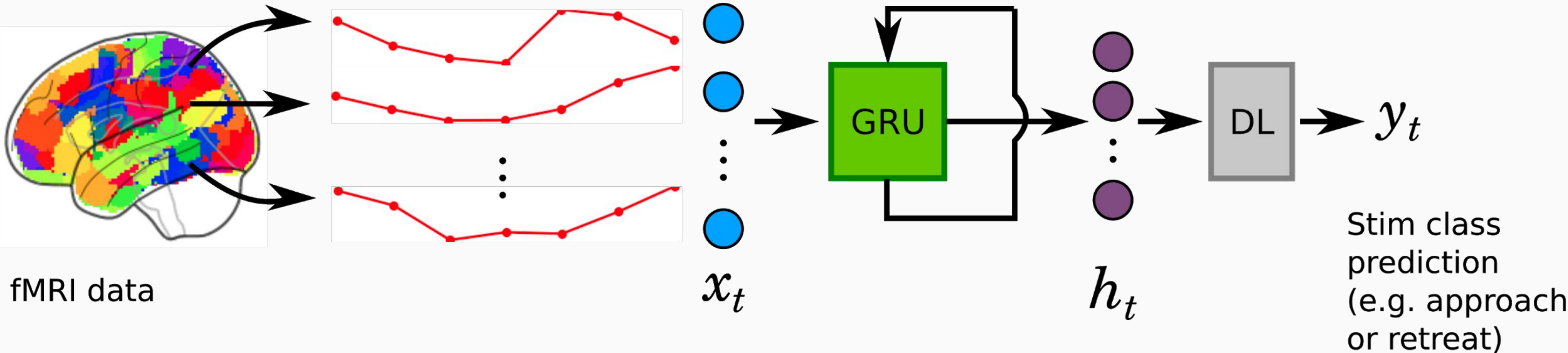
Feature Selection

- Representative Segments from a total of 316 brain regions were extracted in a similar fashion.
- Following figure shows all the 316 brain regions of interest (ROI).
- Out of 61 participants, 42 participants' data was used to train and validate the model.
- Remaining 19 participants data was used to test the model.



Modeling

- A variant of Long-Short Term Memory (LSTM), Gated Recurrent Units (GRU) architecture was employed to characterize the spatio-temporal pattern in the fMRI data
- The GRU architecture had three hidden layers, each with 16 GRU units, and an ouput time-distributed, dense layer (DL) with single sigmoid activation unit. The time-distributed DL returned prediction for every timepoint



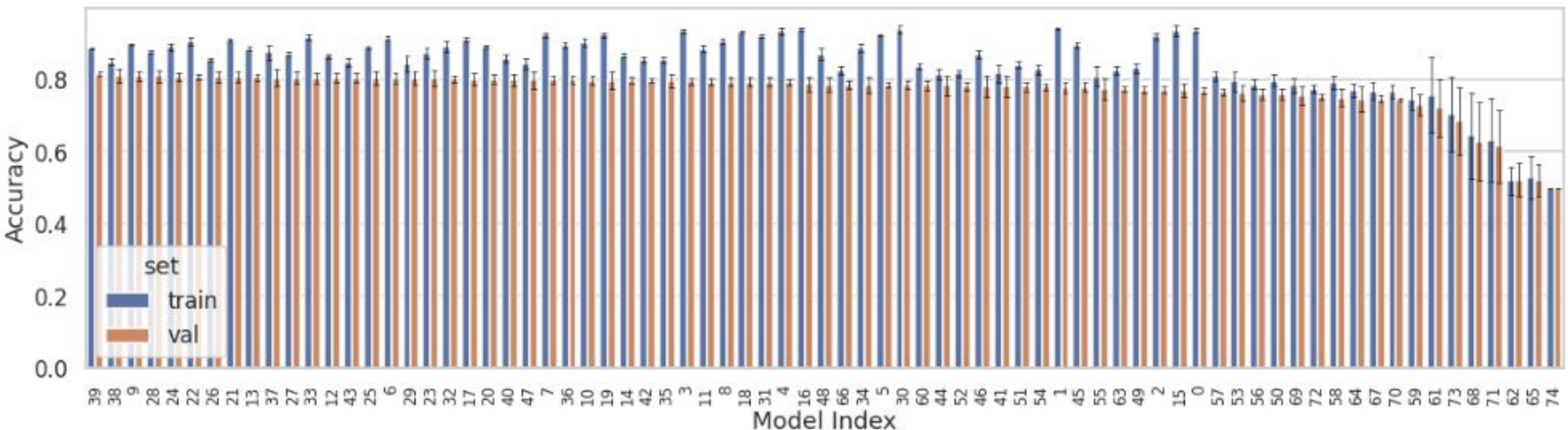
Modeling

- The model was fine tuned by finding optimal values for the following hyperparameters: 1) L2 regularization and 2) dropout rate applied to each of the hidden layers, and 3) learning_rate of the Adam optimizer.

- Optimal hyperparameters were found by doing a full grid-search over 75 combinations of hyperparameters, and cross-validating each combination with a nested 5-fold cross-validation method.

Modeling

- A total of 75 models were trained and validated.
- Following plot shows the performance of each model in terms of its training and validation accuracies.
- The models are shown in the descending order of their mean validation accuracy.
- The error bars indicate standard deviation across folds.



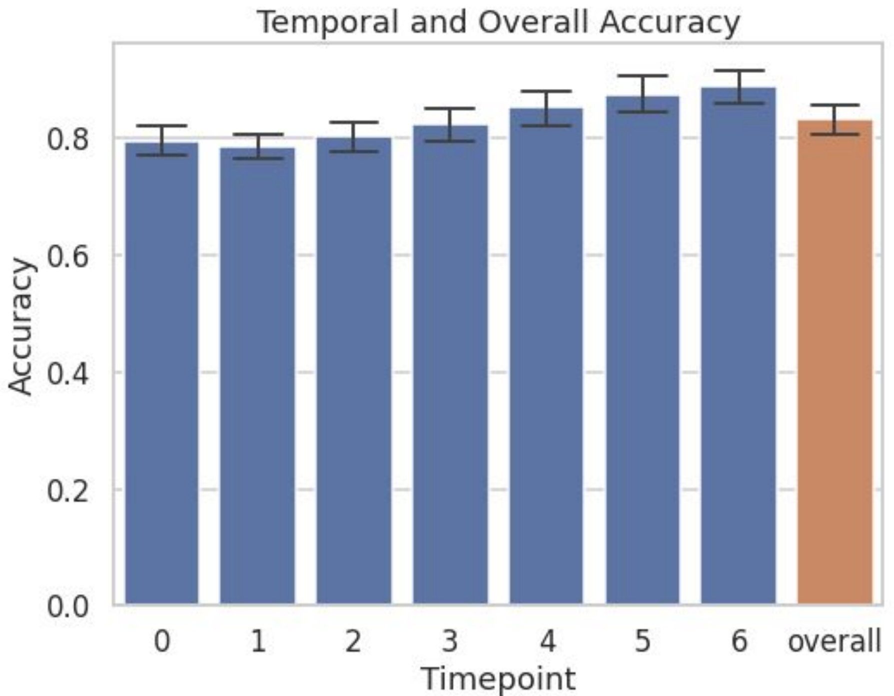
Model Performance

- Best performing model yielded mean training and validation accuracies of 0.89 and 0.82, respectively. Its hyperparameters were: L2 = 0.003, dropout = 0.3, and learning_rate = 0.001.

- The trained model was tested on the near-miss segments of the 19 held-out participants.

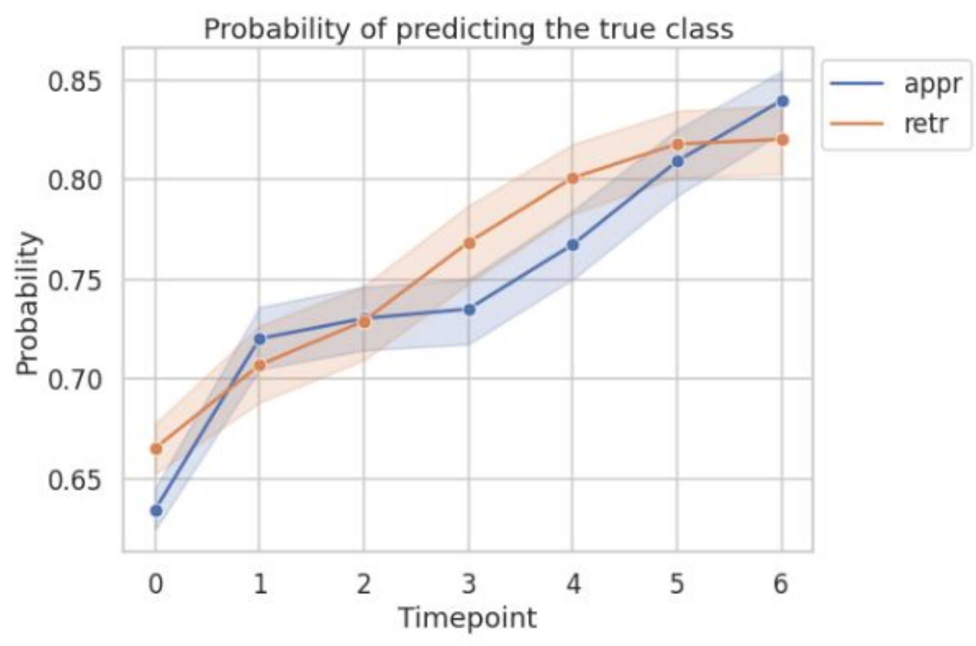
Model Performance

- Figure shows temporal and overall accuracies on the held-out participants.
- The model performs reasonably well from the 1st timepoint (TP) itself, with a mean accuracy of 0.79.
- The mean accuracy steadily increases to 0.89 by the 7th TP.
- “Overall” accuracy is the mean accuracy across TP, which is 0.83.



Model Performance

- This figure shows probability of predicting the true class as a function of time. The probability of predicting the true class increases with time.
  - Approach: “appr”
  - Retreat: “retr”

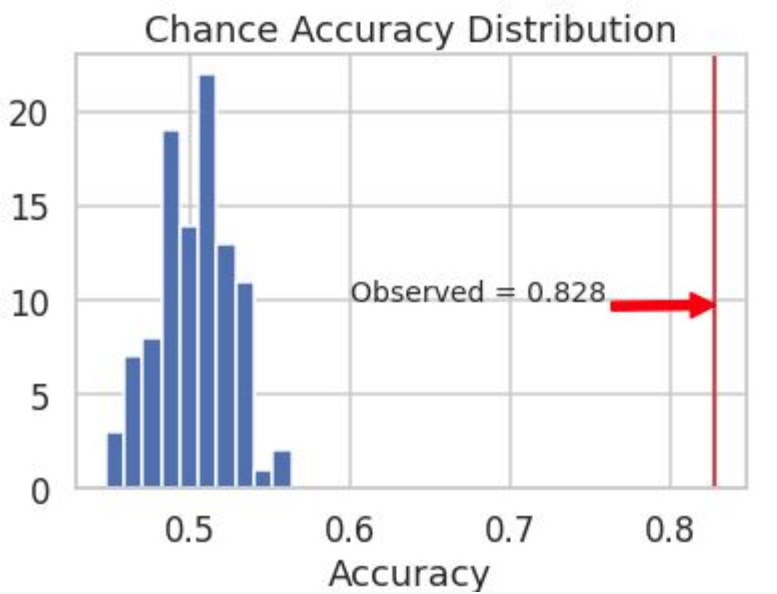


Model Evaluation

- Observed vs. Chance Accuracy
    - To assess significance of the model performance, the observed test accuracy was compared against the accuracy that would be observed if the model was to guess one of the two classes at random.
    - Model with the best hyperparameter settings was trained on the training set a hundred times, each time with randomly shuffled labels.
    - At every iteration, the model was tested on the test set with “non-shuffled” (i.e., true) labels.
    - This process was meant to simulate a chance accuracy distribution.
    - The mean of the chance accuracy distribution formed the baseline performance measure against the observed performance of the model when trained on true labels.
    - The observed test accuracy was significantly greater than the average chance accuracy ($p < 0.009$). See the figure below.

Model Evaluation

- Observed vs. Chance Accuracy

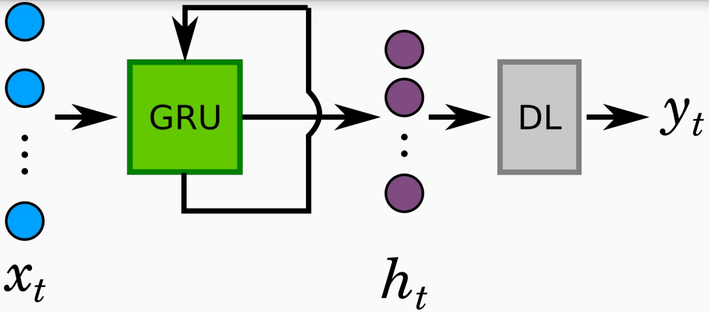


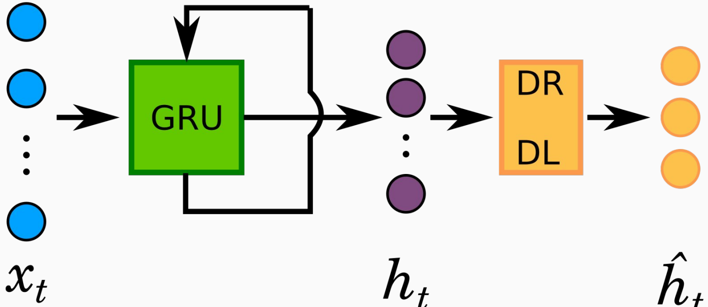
Model Evaluation

- Comparison with Random Forest
  - GRU belongs to the family of recurrent neural networks, hence it learns class separability from the sequential aspect of the data.
  - Current data is an fMRI time-series data, hence GRU was a reasonable model choice.
  - Interesting question that can be asked is, how well would a model that does not take into account the sequential aspect of the data perform? Would it perform as well as the GRU model? Or would it perform poorly?
  - To make this comparison, a Random Forest classifier was trained on the current data.
  - The Random Forest classifier was also fine tuned using the nested cross validation method, and best hyperparameters (n_estimators = 1500, max_feature = 'sqrt') were obtained.
  - The test accuracy of Random Forest classifier was only 0.58 which is significantly low compared to that of the GRU model.

Temporal Trajectories

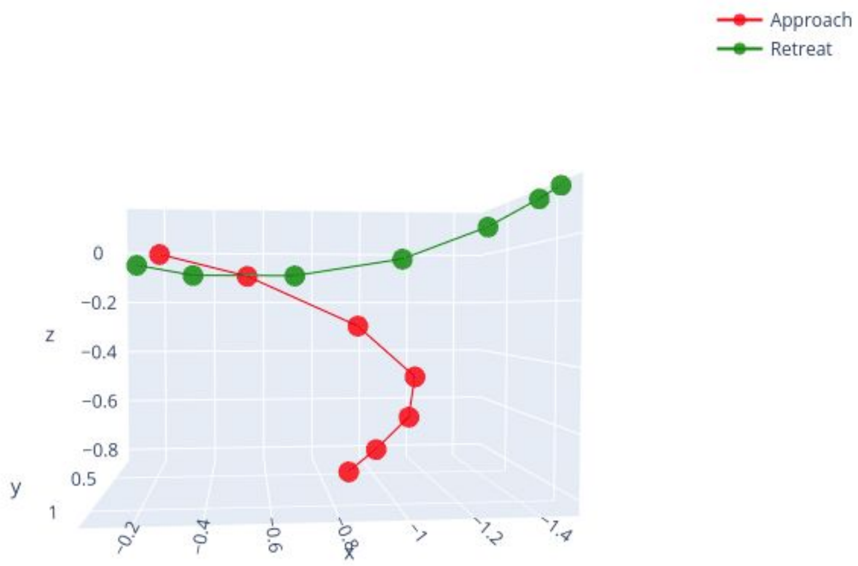
- GRU outputs hidden states that are typically high dimensional. Hidden state (ht) capture spatio-temporal variance that is most useful in maintaining class separability.
- To visualize dynamics, ht was linearly projected onto a lower (3D) dimensional space.
- This was done by replacing the output layer with a *Dimensionality Reduction Dense Layer (DRDL)* with three linear units. In essence, this is a supervised non-linear dimensionality reduction step.





Temporal Trajectories

- The 3-dimensional representations of the hidden states for both stimulus class are plotted along the three axes of the coordinate system.
- The plot represents the temporal trajectories of the two classes.
- At the first timepoint the two classes are closest to each other.
- Distance between them increases with every timepoint, indicating class separability improves with time.
- Interactive plot.



Conclusion

- The GRU model was successful in characterizing the spatio-temporal patterns of the brain during approaching and retreating threats.
- Check out the project's github and interactive jupyter-book.

Demonstration

- The video clip inth link demonstrates model performance on one of the held out participant's fMRI data.
- The video shows the visual paradigm presented to the participant during his/her fMRI scan, along with model predictions at the top-right of the screen.
- Prediction is either "Approach" or "Retreat".
- If the prediction is correct, the color of the text remains green; and if it is incorrect, it turns red.
- Note that when the circles touch, the screen turns white and a red wheel appears around the circles to indicate delivery of the physical shock.
- Also note that the speed of the video has been increased by 4x for quick demonstration purposes.