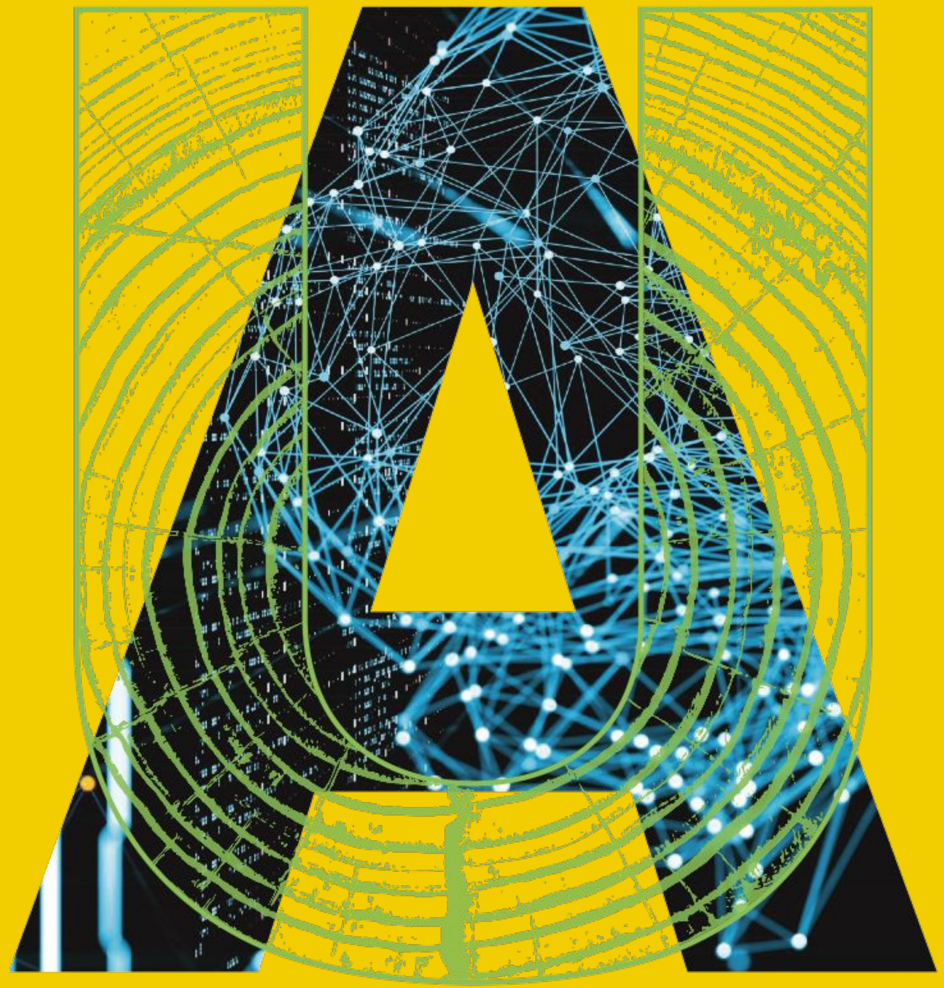
MACHINE LEARNING & THE BRAIN

Coding Workshop

Thursday 28 September 2023

Alex Murphy

UNIVERSITY OF ALBERTA



Today

Session 1: Working with EEG (Notebook link)

Session 2: Working with fMRI (Notebook link)

Session 3: Working with LLMs (Notebook link)

The notebook outputs are all contained in the uploaded version. Please feel free to "Clear All Outputs" and run along in real time with me. This will allow you to inspect outputs further and play around with settings (useful).

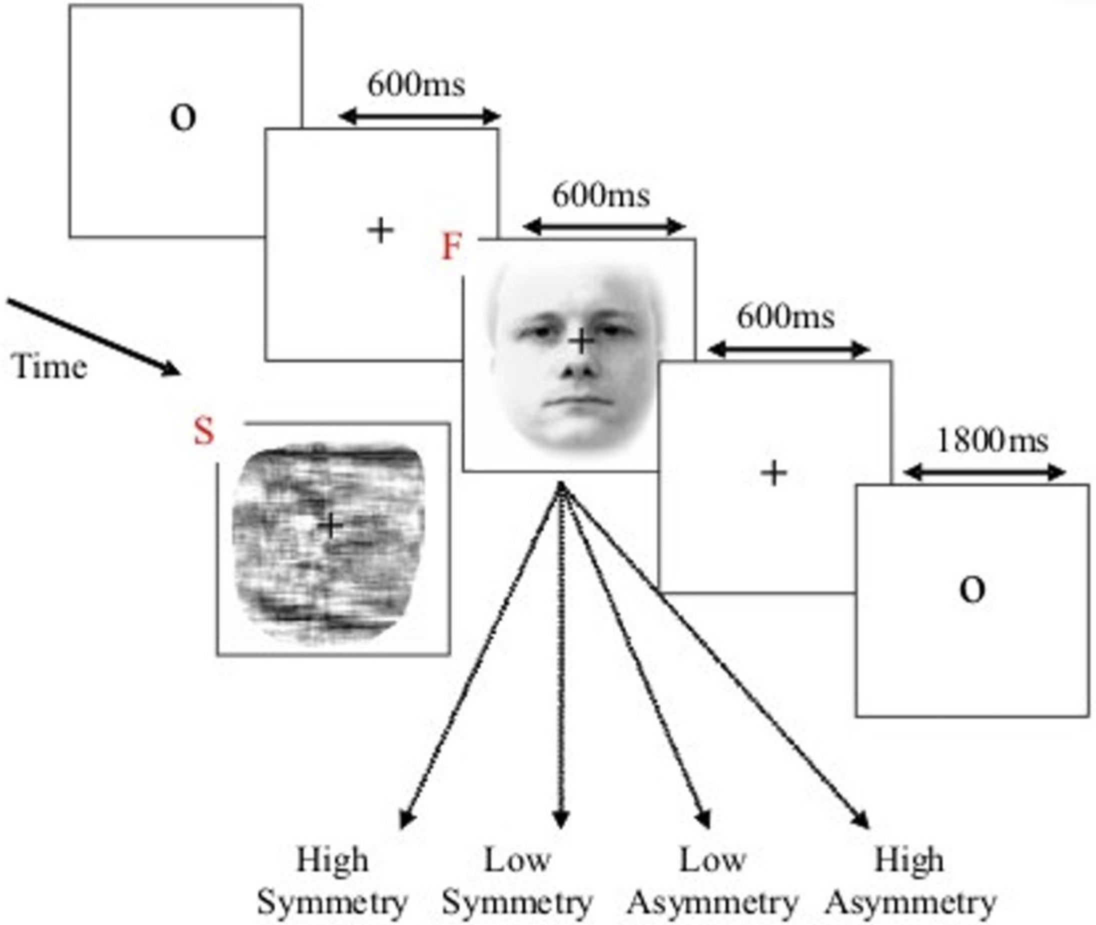
If you have cleared all outputs and want to restore everything, just come back to the original links and refresh. The cells used to preprocess data from the Google Collab file system **should only be run once**.

Today's Dataset

Today, we will be working with a multimodal dataset. A person was recorded with EEG, fMRI and MEG while performing a vision task to determine the various mechanisms involved when viewing **intact faces** and **scrambled faces.**

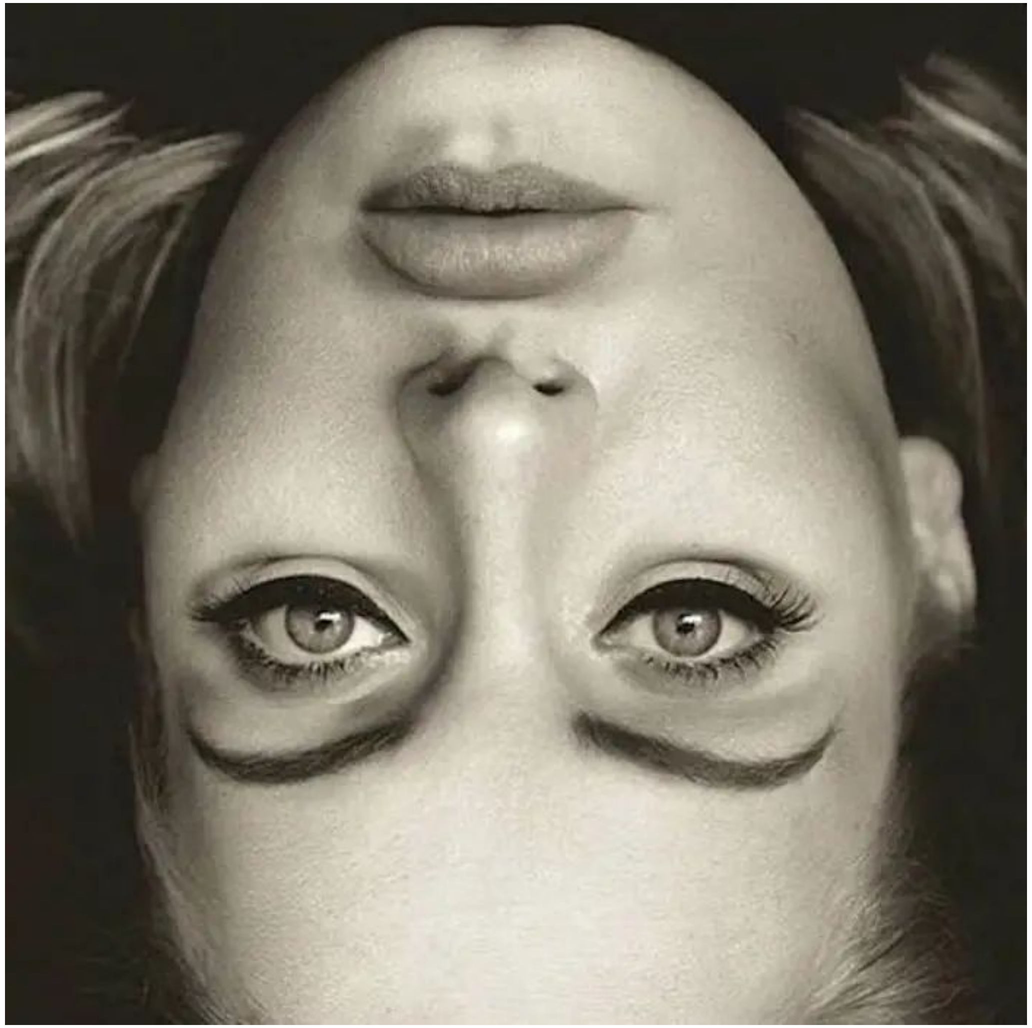
Can we train a classifier to tell whether or not a person is looking at a normal or a scrambled face, purely based on recorded brain activity?

We'll look at EEG & fMRI to find out.

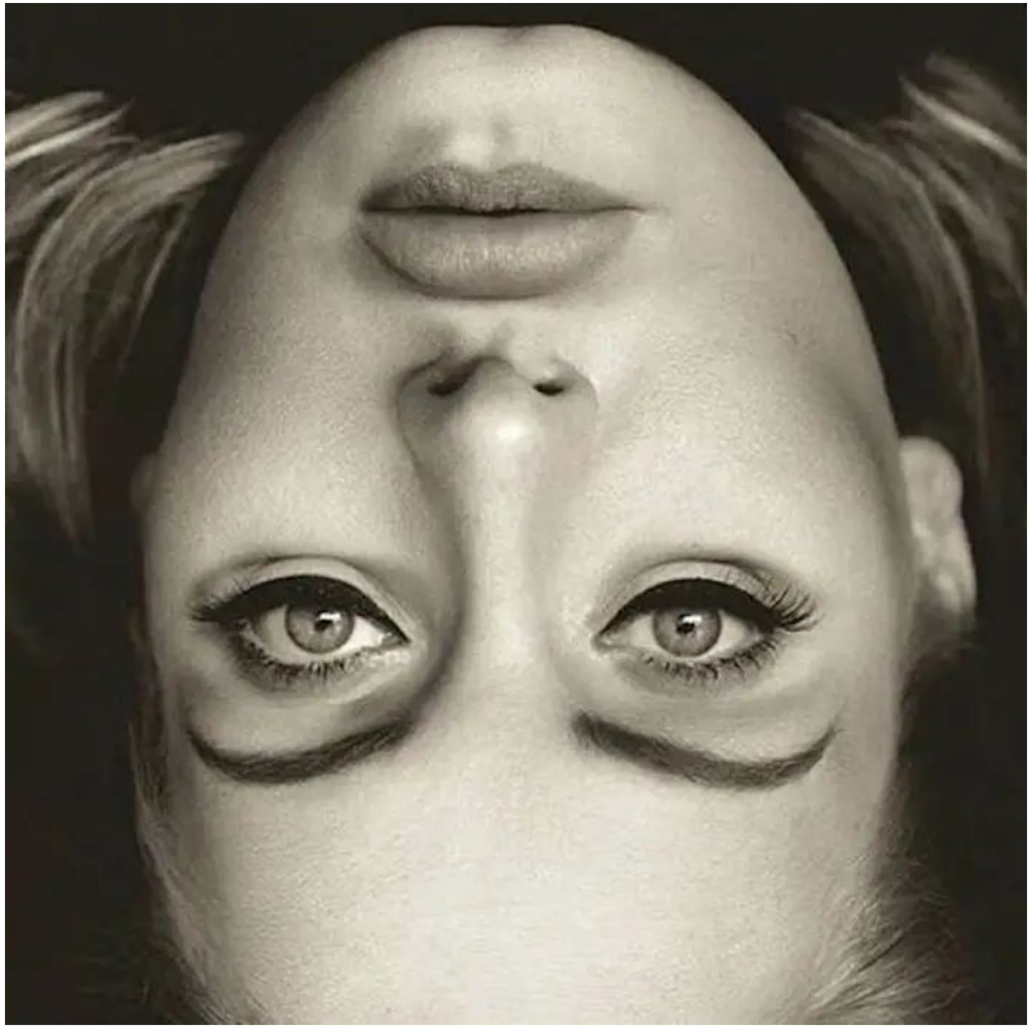


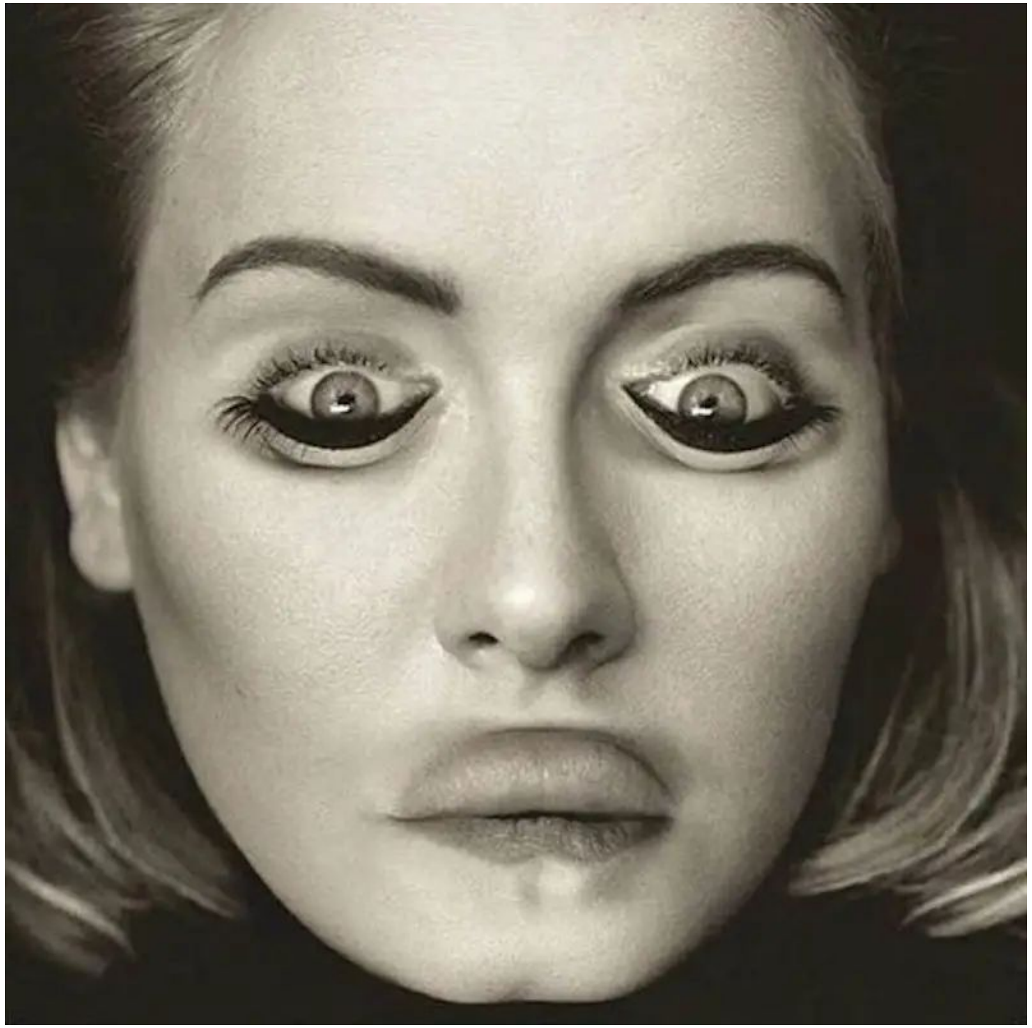
https://www.fil.ion.ucl.ac.uk/spm/data/mmfaces/

Sidenote on Face Illusions



Sidenote on Face Illusions





The Thatcher Illusion

The Thatcher illusion: Are faces special?

Mind gamers: why was a picture of Margaret Thatcher so important for understanding the human visual system?

The Thatcher Illusion

📷 The Thatcher illusion shown the correct way up. Composite: Peter Thompson

How it works

The Thatcher illusion was an important demonstration, because it was one of the first to highlight some of the underlying mechanisms by which our brains process information about faces. **By and large, faces are made up of the same, consistent features – two eyes, a nose, a mouth, some ears, and so on.** One way in which our brains could process faces is to analyse them as a collection of these separate, individual features. If that were the case though, we might expect to be easier to pick out any discrepancies in an upside-down face. **The fact that we don't suggests that we process and recognise faces in a more holistic manner – in other words, we don't just process individual features, but also the positions and relationships between those features.**

https://www.theguardian.com/science/head-quarters/2016/sep/19/the-thatcher-illusion-are-faces-special

The Brain's *Face-Response* Signal

The Thatcher illusion tells us that there is something special about face processing in the brain. What do we know from cognitive neuroscience?

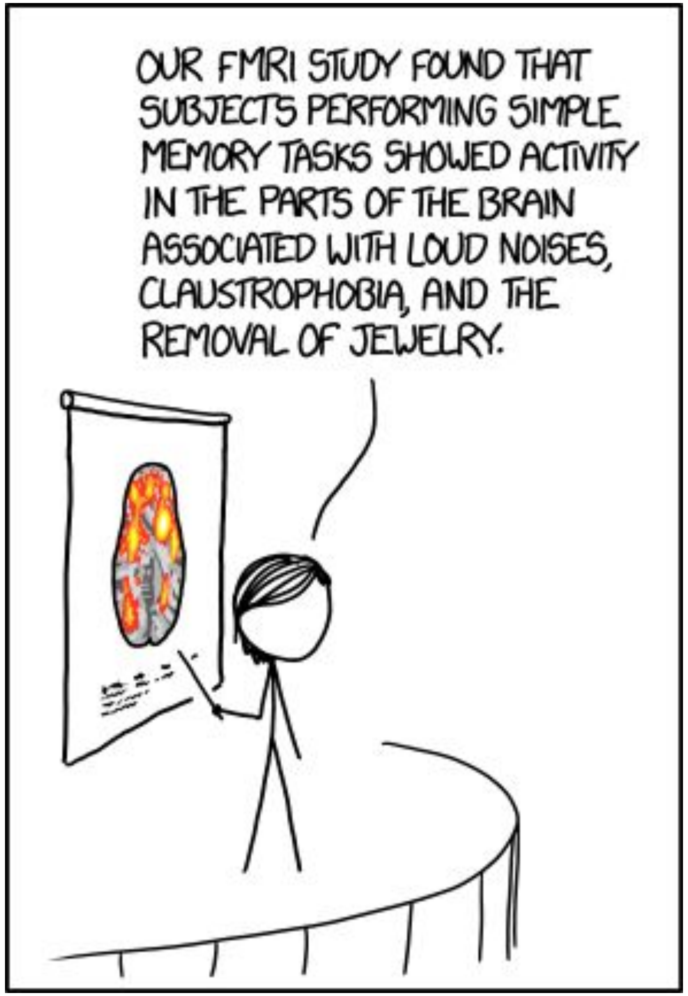
>> Enter the N170

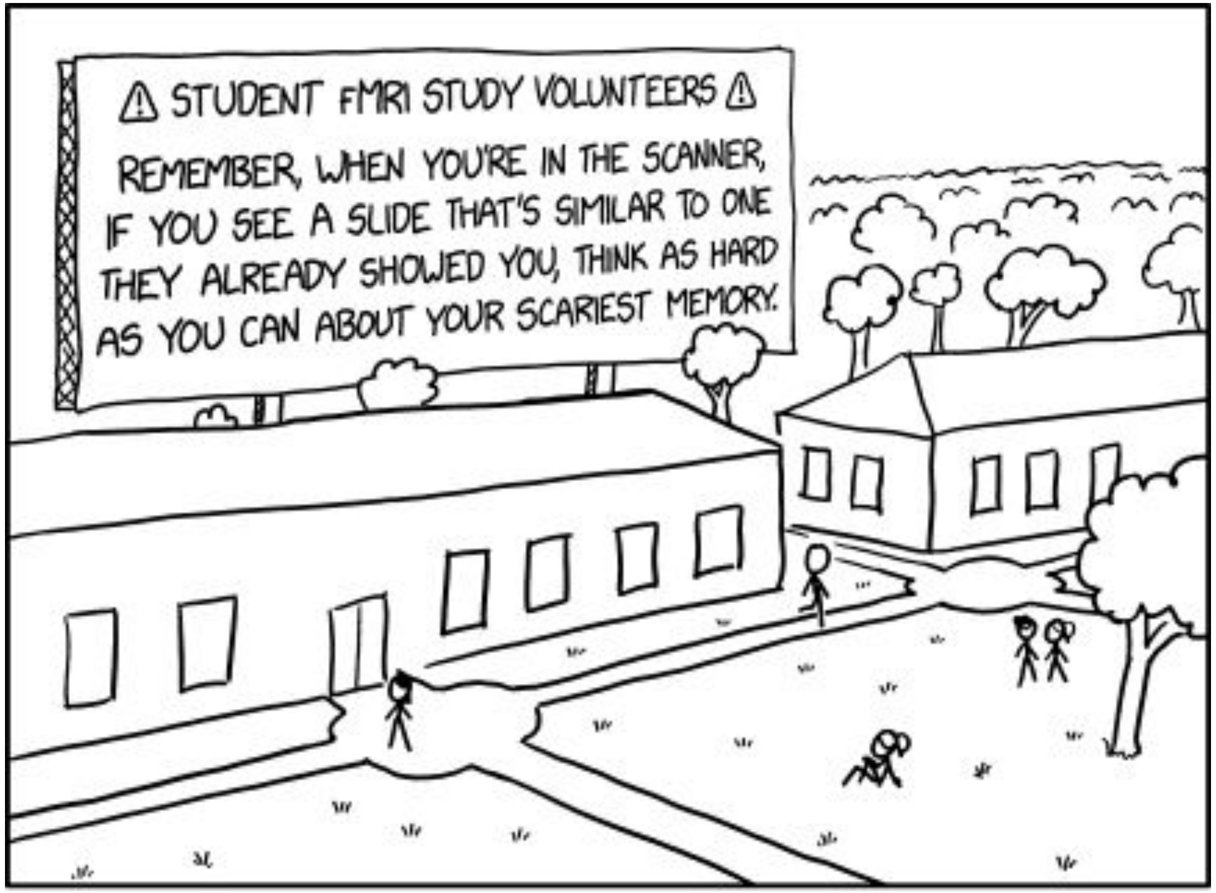
- The N170 effect is a negative peak in EEG ~170 ms after seeing a face.
- Also detectable in MEG (where it's called M170)
- This effect can be seen in EEG analysis via taking two well-matched conditions:
  - viewing faces
  - viewing scrambled versions of the same faces
- By averaging the result in each condition and subtracting one from the other, we can see the net effect that one (experimental) condition has that another (control) condition does not
- If there is only activity around zero after subtraction, brain response is deemed to be equal for both conditions (might not be true in reality, but is in the recorded signal)

Session 1: Working with EEG

colab

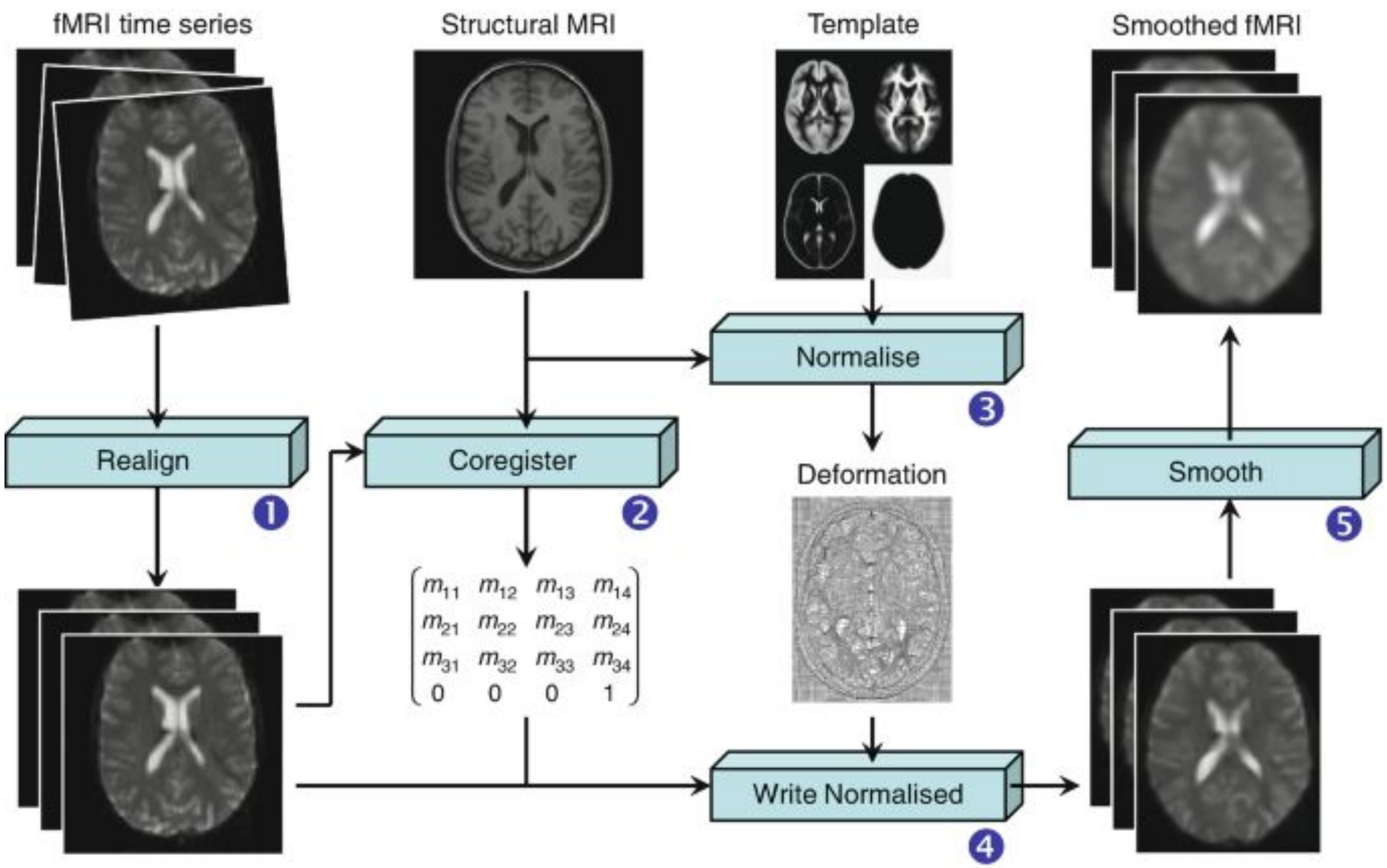
Session 2: Working with fMRI





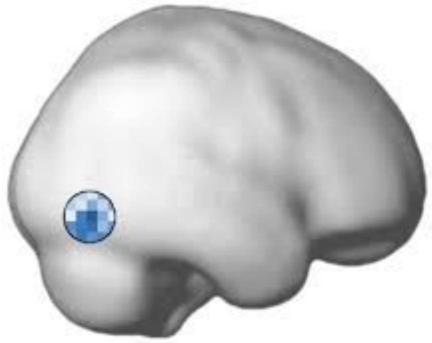
A RIVAL NEUROSCIENCE DEPARTMENT KEEPS TRYING TO SABOTAGE OUR EXPERIMENTS.

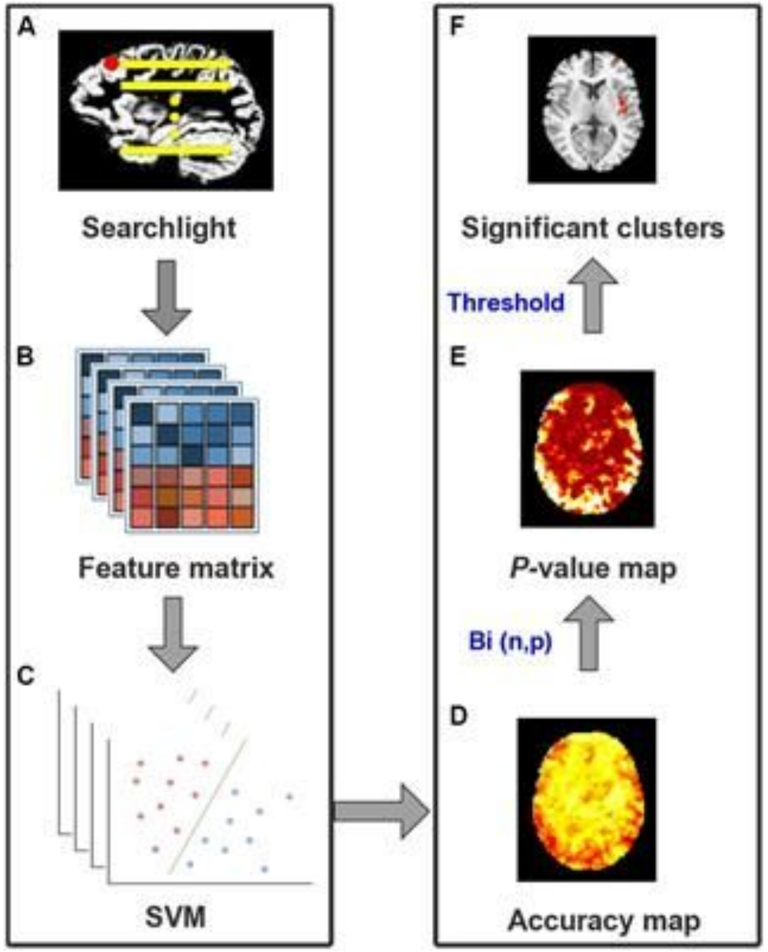
Session 2: fMRI Preprocessing



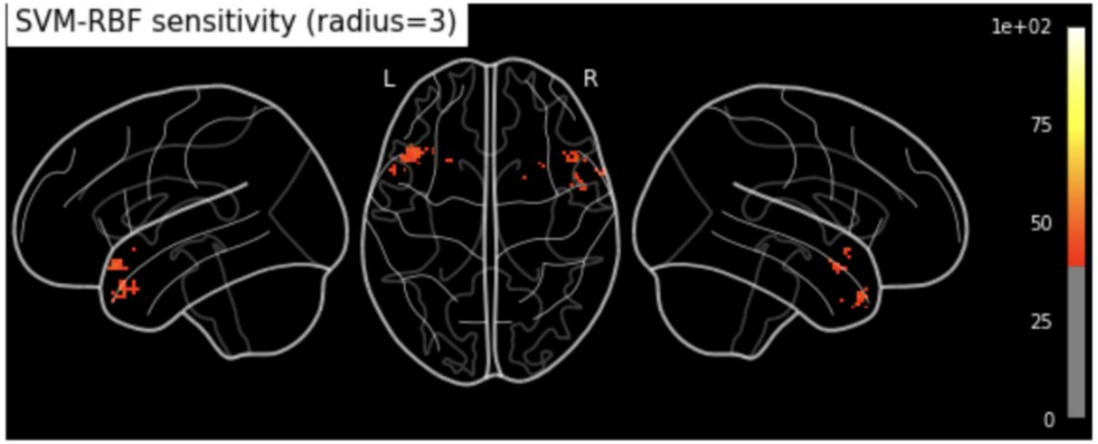
https://www.researchgate.net/figure/Flow-chart-of-preprocessing-steps-of-the-fMRI-analysis_fig5_299561305

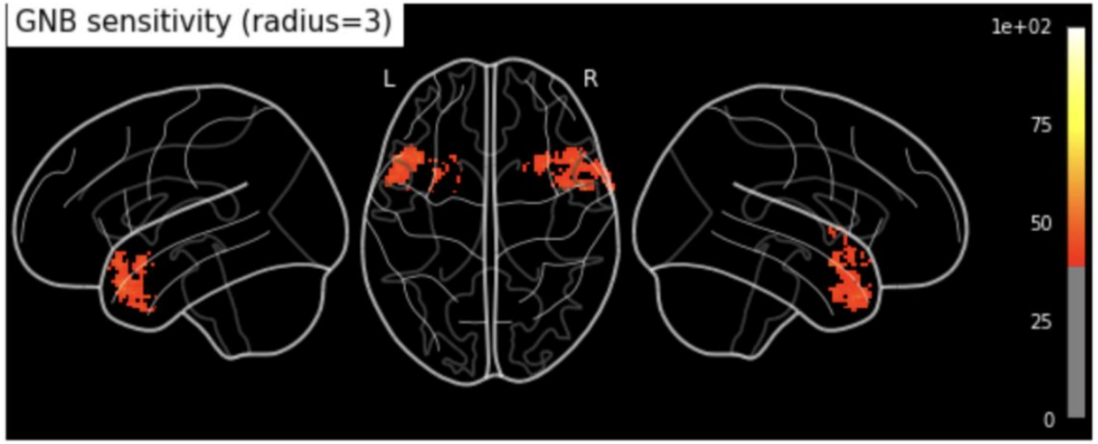
Session 2: “*Searchlights*”





Session 2: “Searchlights”





Session 2: Working with fMRI

Session 3: LLMs

Today we will review basic loading and concepts needed to extract language embeddings from large language models, which will be potential feature representations to use in your projects that touch upon language in some way.

HUGGING FACE

🤗 Transformers

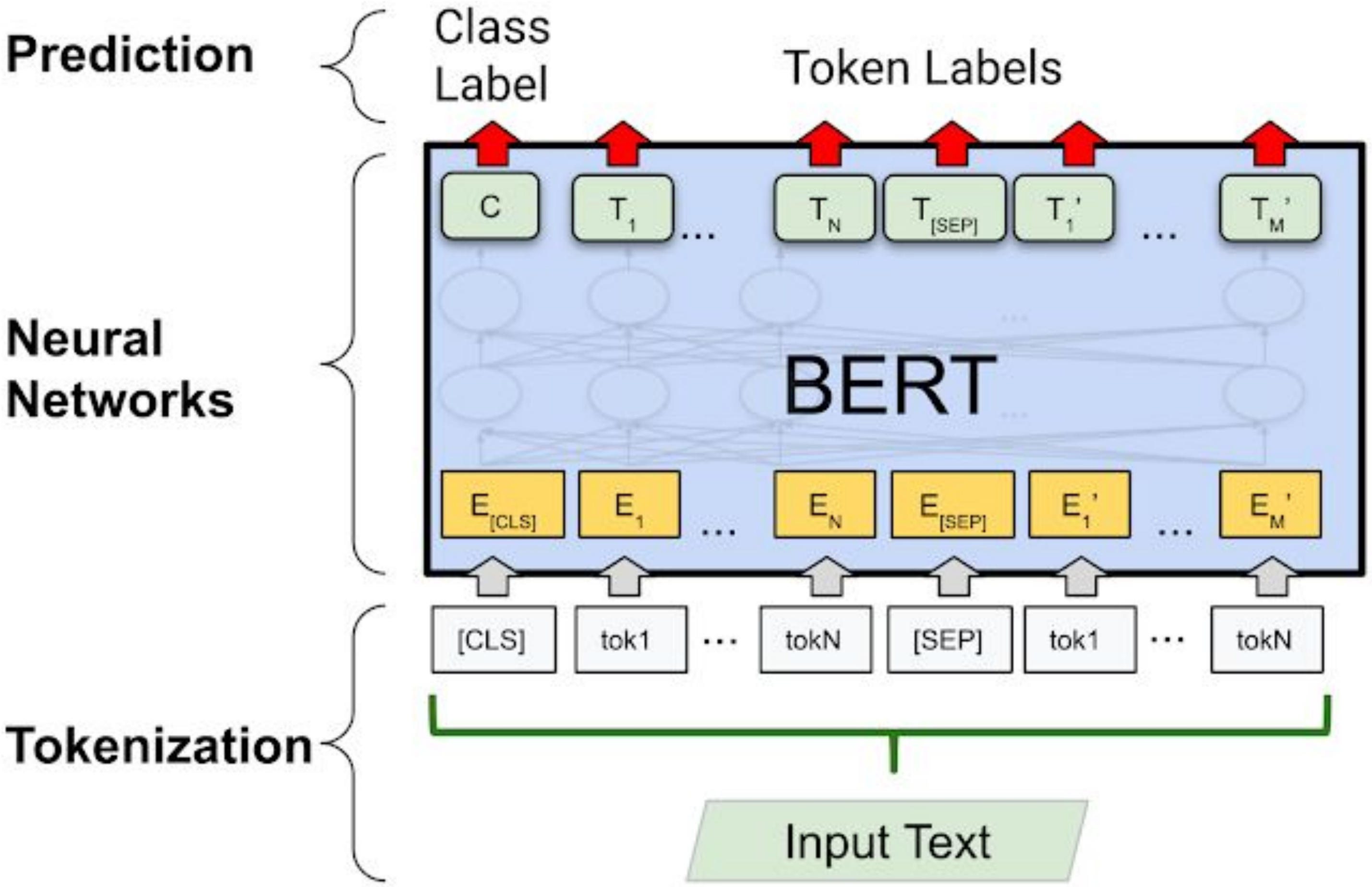
State-of-the-art Machine Learning for PyTorch, TensorFlow, and JAX.

🤗 Transformers provides APIs and tools to easily download and train state-of-the-art pretrained models. Using pretrained models can reduce your compute costs, carbon footprint, and save you the time and resources required to train a model from scratch. These models support common tasks in different modalities, such as:

📝 **Natural Language Processing**: text classification, named entity recognition, question answering, language modeling, summarization, translation, multiple choice, and text generation.
🖼️ **Computer Vision**: image classification, object detection, and segmentation.
🗣️ **Audio**: automatic speech recognition and audio classification.
🐙 **Multimodal**: table question answering, optical character recognition, information extraction from scanned documents, video classification, and visual question answering.

🤗 Transformers support framework interoperability between PyTorch, TensorFlow, and JAX. This provides the flexibility to use a different framework at each stage of a model's life; train a model in three lines of code in one framework, and load it for inference in another. Models can also be exported to a format like ONNX and TorchScript for deployment in production environments.

Session 3: LLMs - Tokenisation



https://blog.research.google/2021/12/a-fast-wordpiece-tokenization-system.html

Session 3: Working with LLMs

Final Points

- Ask us if you get stuck during your data preprocessing
- Look in the literature with keywords to explore method ideas
- Always apply the ***Defensive Coding*** principle
  - *Test and check every single assumption you implicitly make about your data*
- Machine learning with brain data is a very cool topic - don't forget to enjoy the process (it will certainly have its challenges)

Happy Coding!

UNIVERSITY OF ALBERTA