Breakout Session 7:

Cloud Based Neuroimaging Analysis for Identifying Traumatic Brain Injuries and Related Changes

Dr. Aparna Gullapalli
Assistant Professor, Mind Research Network

Cloud based neuroimaging analysis for identifying traumatic brain injuries and related changes

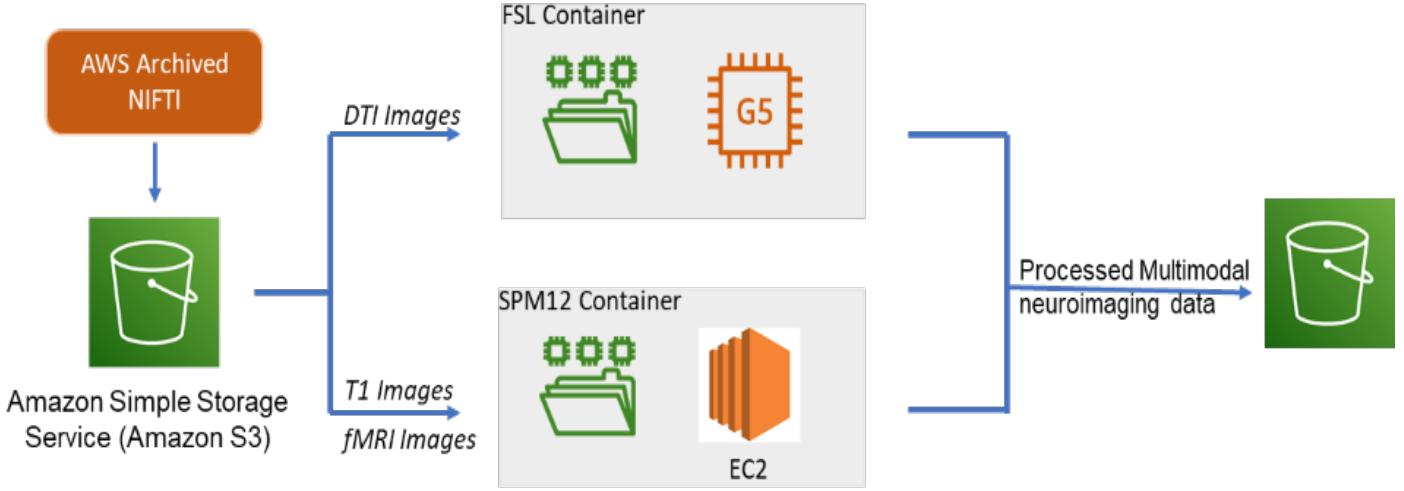
Aparna Gullapalli
Mind Research Network,
Albuquerque, NM

Motivation

- Traumatic Brain Injury (TBI) is a global health concern in the general population, and is disproportionately represented among incarcerated populations and the goal of our research is to use advanced multimodal neuroimaging data to identify reliable biomarkers in chronic mild TBI.
- Neuroimaging across all modalities can produce data on the order of 6GB – 10GB per subject.
- Current classification algorithms have computational complexity approaching $O(n^4p)$
- Running times of multiple weeks per classification run make data exploration and hyperparameter tuning an inherently long process
- Currently our raw imaging data is stored in S3 buckets on AWS as part of COINS (COllaborative Imaging and Neuroinformatics Suite) in a HIPAA compliant architecture.

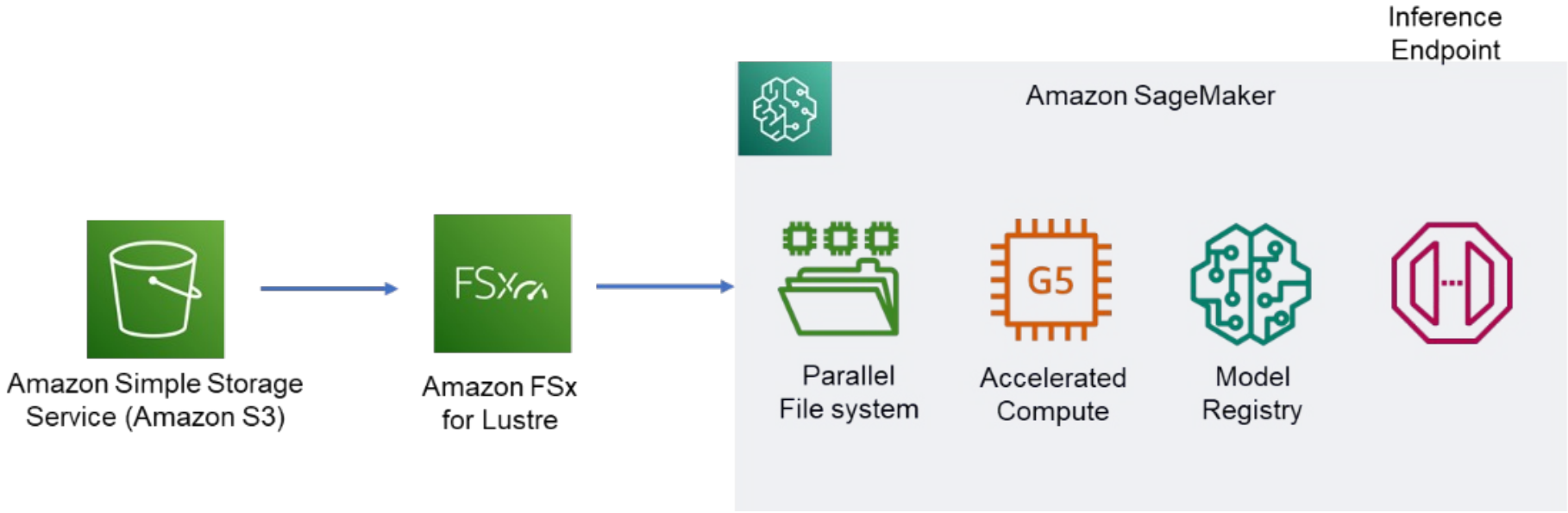
Design

- Containerize our pre-processing pipelines for neuroimaging data
- Specifically, use the G5 suite of processors for processing diffusion tensor imaging data and EC2 for structural T1 MRI images
- Validation will be performed by comparing containerized outputs to local in-house processed data.
- The main aim of this step is to enable the generalization of the pre-processing pipeline of raw NIFTI images in the cloud to a data analysis format within the AWS ecosystem.



Design

- We will explore the data science and machine learning features of AWS through Amazon Sagemaker, a HIPAA eligible, fully managed data science platform with features enabling every stage of AI/ML development life cycle
- Each neuroimaging modality can consist of several million voxels which require careful feature selection and union approaches.
- AWS designed training and inference accelerator, namely AWS Trainium and Inferentia, deliver high performance model training and inference with the lowest cost allowing us to use “off the shelf ML” algorithms for accelerated prototyping and testing of data analysis approaches.



Design

- AWS SageMaker supports “Bring-Your-Own” algorithms which will let us easily containerize our existing computationally intensive prediction algorithm onto the AWS platform.
- This allows us to measure the cost to time benefit ratio of using the advanced memory caching, parallel file system and high performance computing capabilities cloud ecosystems such as AWS.
- One of the specific aims of our parent grant was to test the generalizability of our prediction algorithms and biomarkers to data in Federal Interagency Traumatic Brain Injury Research (FITBIR) Informatics System.
- The cloud based pre-processing containers and TBI biomarker detection pipeline will be implemented on neuroimaging data from FITBIR for generalizability.

Expected Outcomes

- Containerizing pre-processing pipelines will allow us to develop easily deployable cloud based software to process NIFTI images that people with minimal knowledge of SPM12 and FSL6 can use. We hope to get a measure of the cost benefit of doing this memory intensive task on the cloud as opposed to locally on a dedicated computer.
- We expect that we will be able to replicate our existing ML pipeline on AWS Sagemaker and cut down the processing time quantifiably.
- We hypothesize that we will be able to take advantage of several inbuilt ML approaches within AWS Sagemaker and identify the best data analytic approach for TBI induced neural deficits and biomarkers.
- Finally, an end-to-end cloud based approach using data acquired from FITBIR will serve as a good test for the reproducibility and reliability of our approach.