BRAINHACK SCHOOL 2020

Can We Identify Gender Using fMRI?

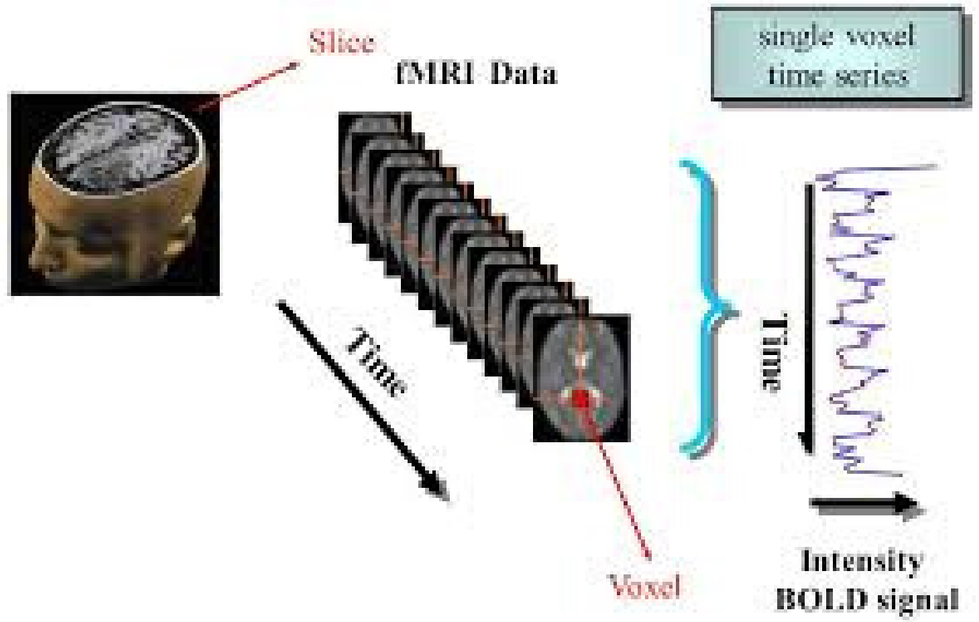
Tajwar Sultana
NED University of Engg. & Tech.
Pakistan

INSPIRATION

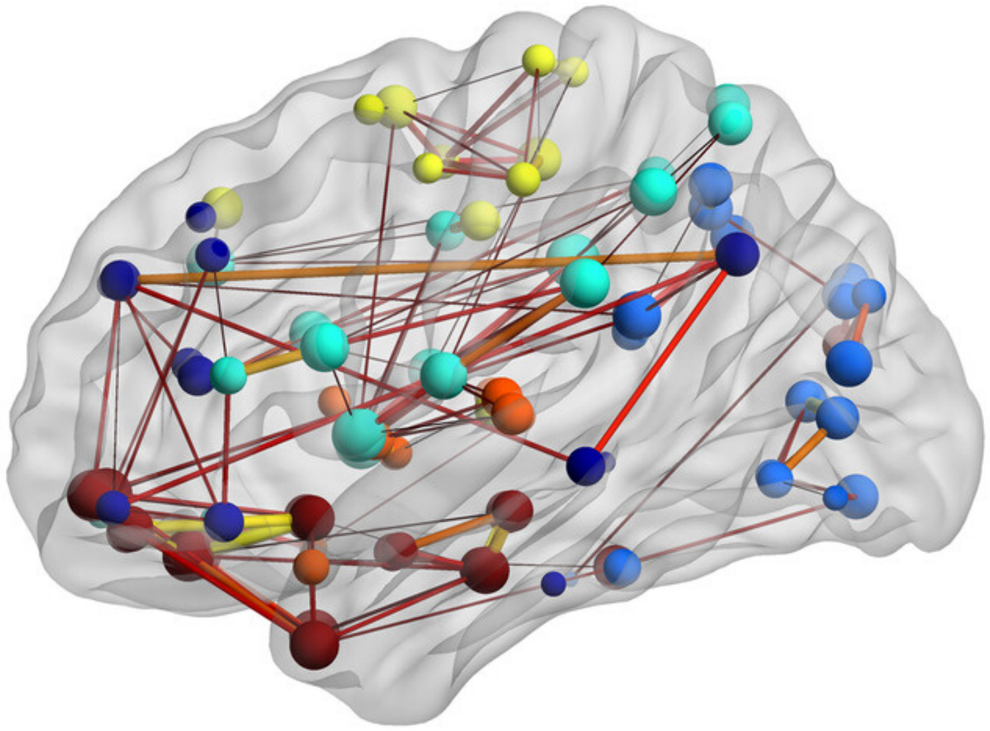
PREVIOUS RESEARCH => GENDER DIFFERENCE RESULTS IN FC DIFFERENCE

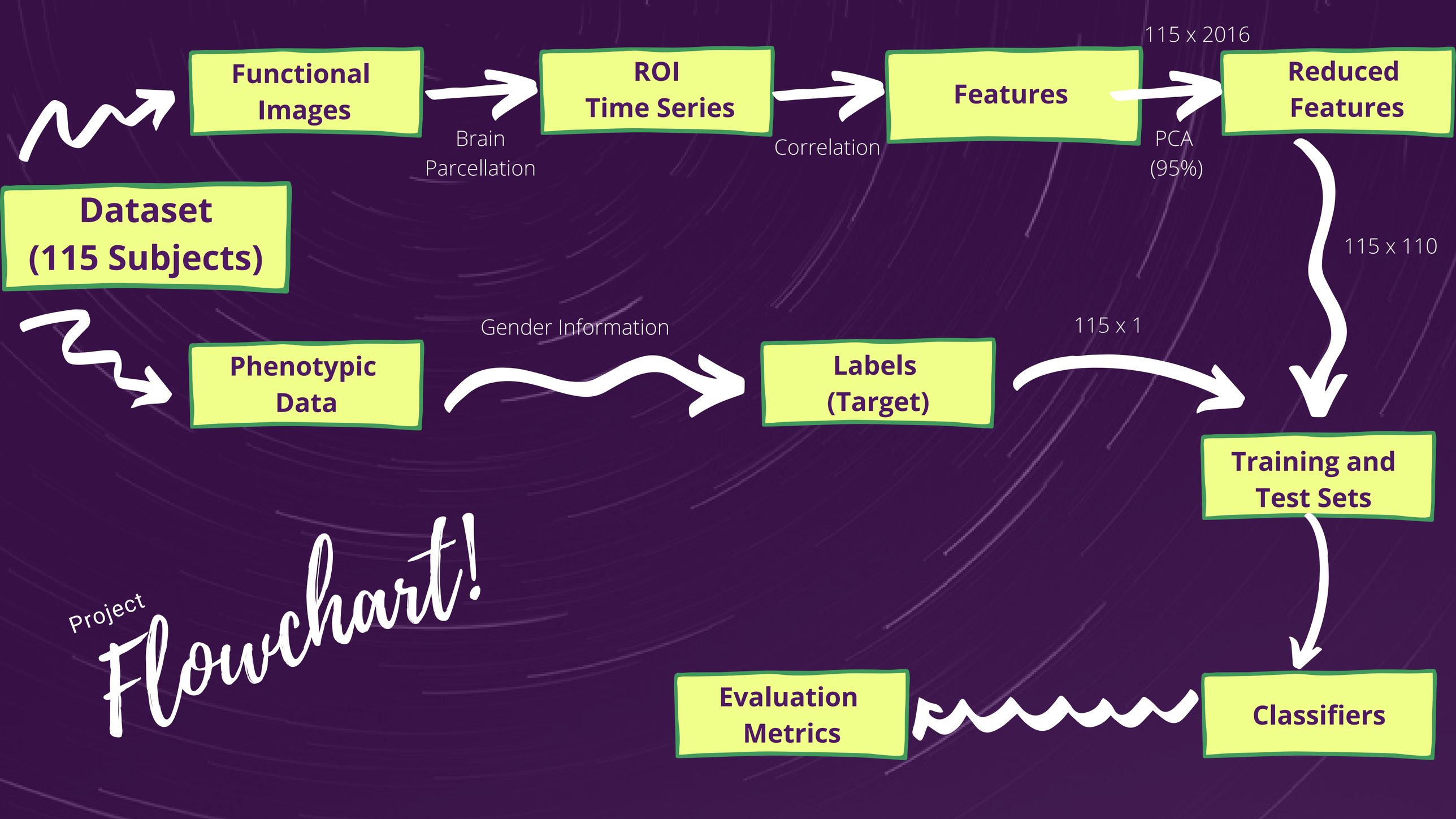
- Functional Connectivity Predicts Gender: Evidence for Gender Differences in Resting State Connectivity
- Sex Classification by Resting State Brain Connectivity

BOLD contrast and fMRI Time Series



Functional Connecitivity




Functional Images
Brain Parcellation
ROI Time Series
Correlation
Features
115 x 2016
PCA (95%)
Reduced Features
115 x 110
Dataset (115 Subjects)
Phenotypic Data
Gender Information
Labels (Target)
115 x 1
Training and Test Sets
Classifiers
Evaluation Metrics
Project
Flowchart!

Development Dataset

FUNCTIONAL IMAGES

- Used to determine Connectivity Matrices
- Features are built using Connectivity matrices

PHENOTYPIC DATA

- Used to determine Labels for supervised learning
- Gender Information as Target Data

DATA SIZE

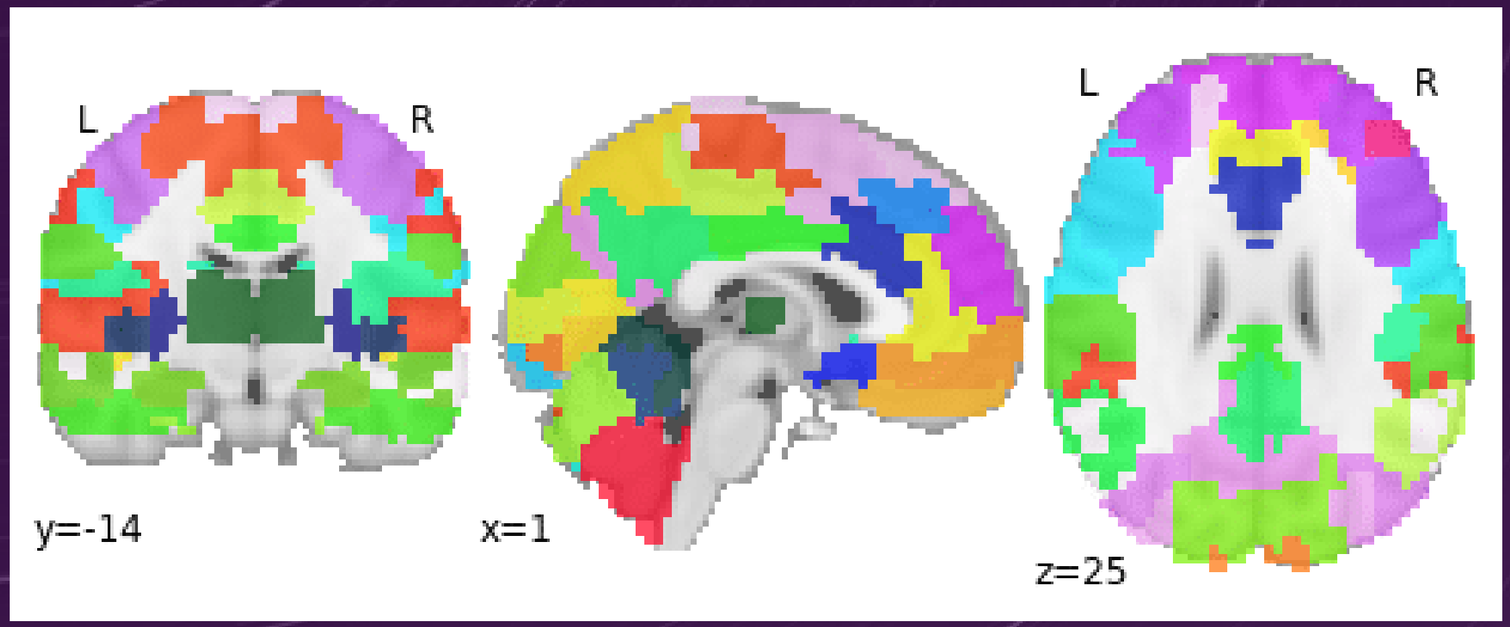
- 155 subjects
- 84 Female
- 71 Male

EXTRACTING

Region Of Interest

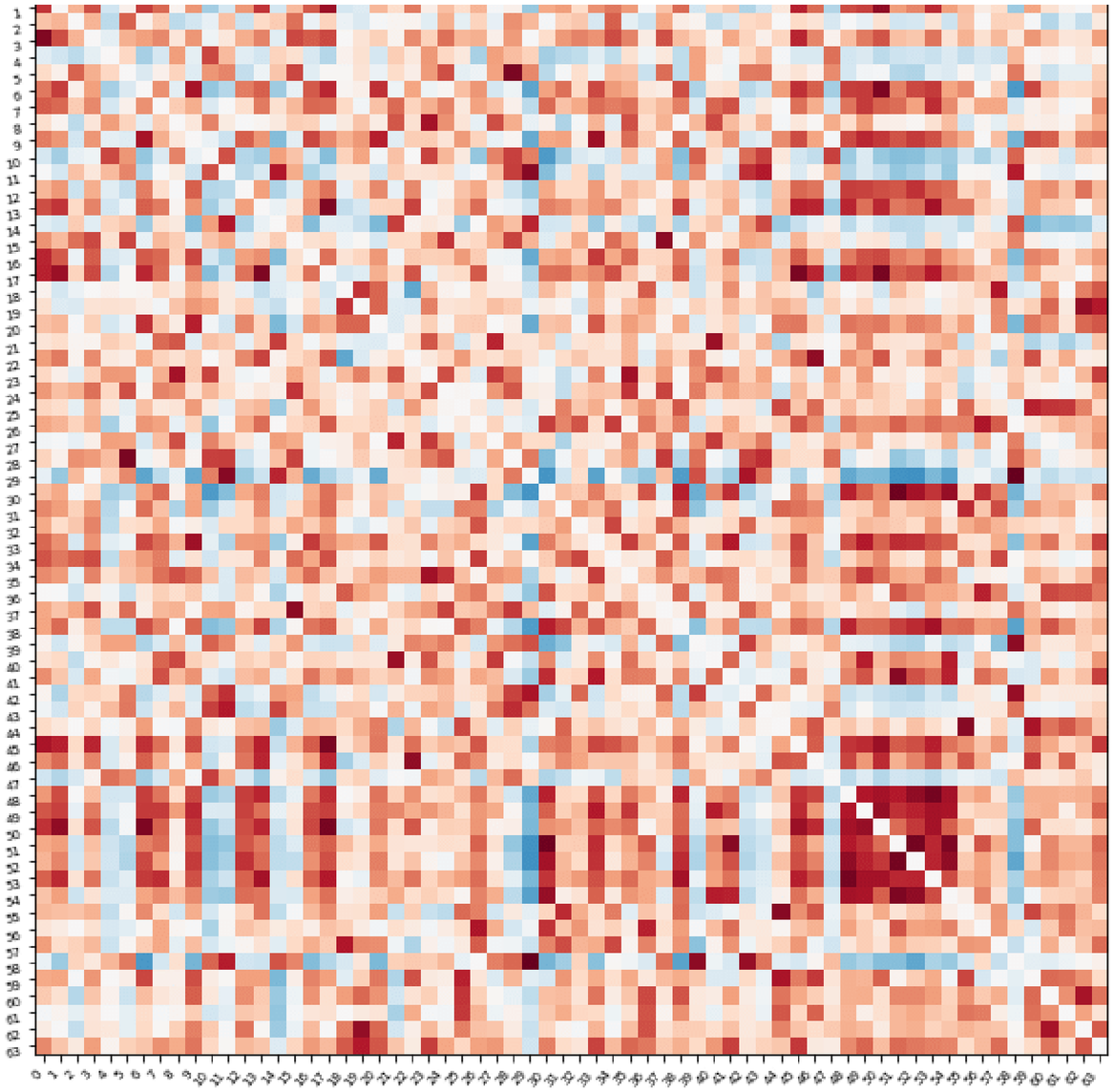
TIME SERIES

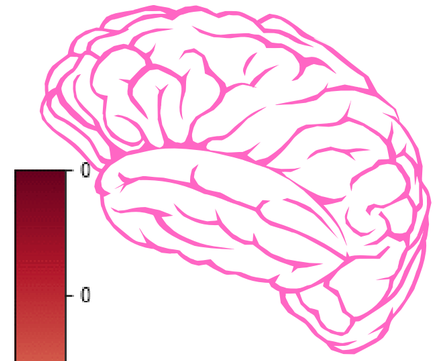
- datasets.fetch_atlas_basc_multiscale_2015
- Resolution = 64



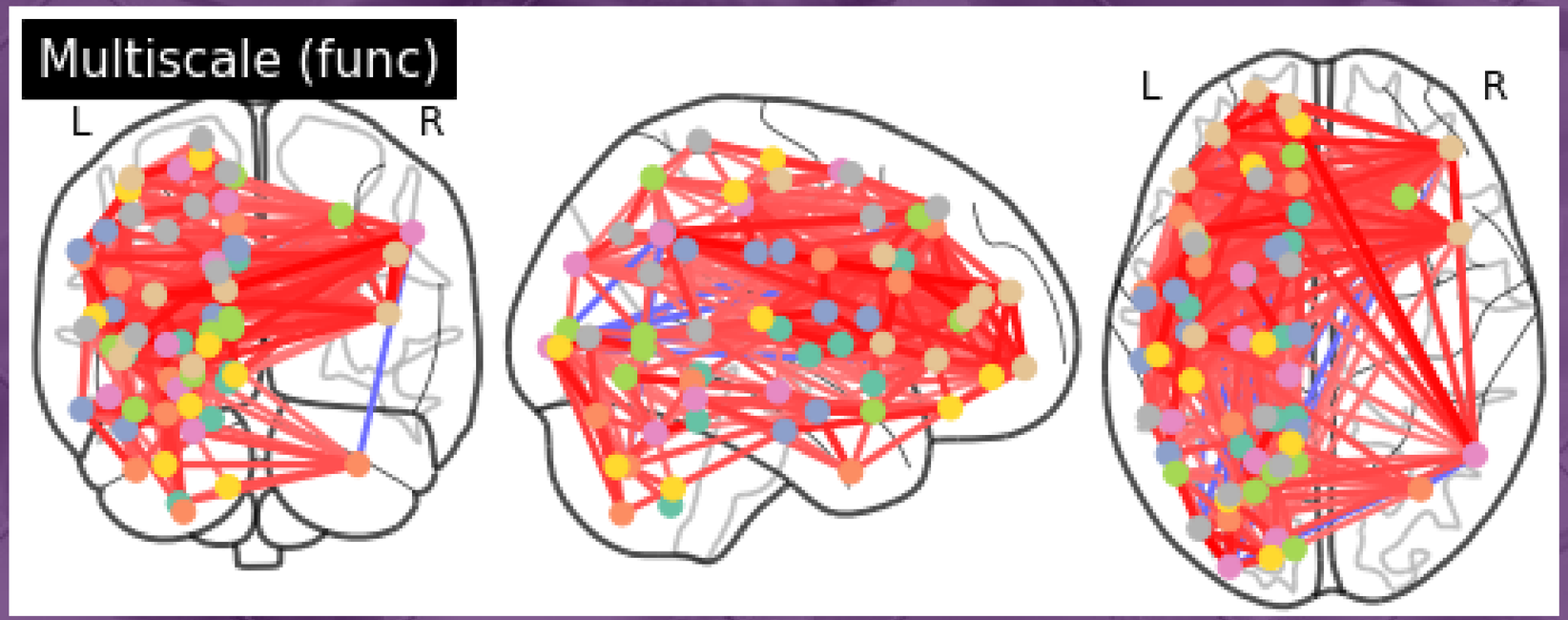
Connectivity Matrix

SINGLE SUBJECT



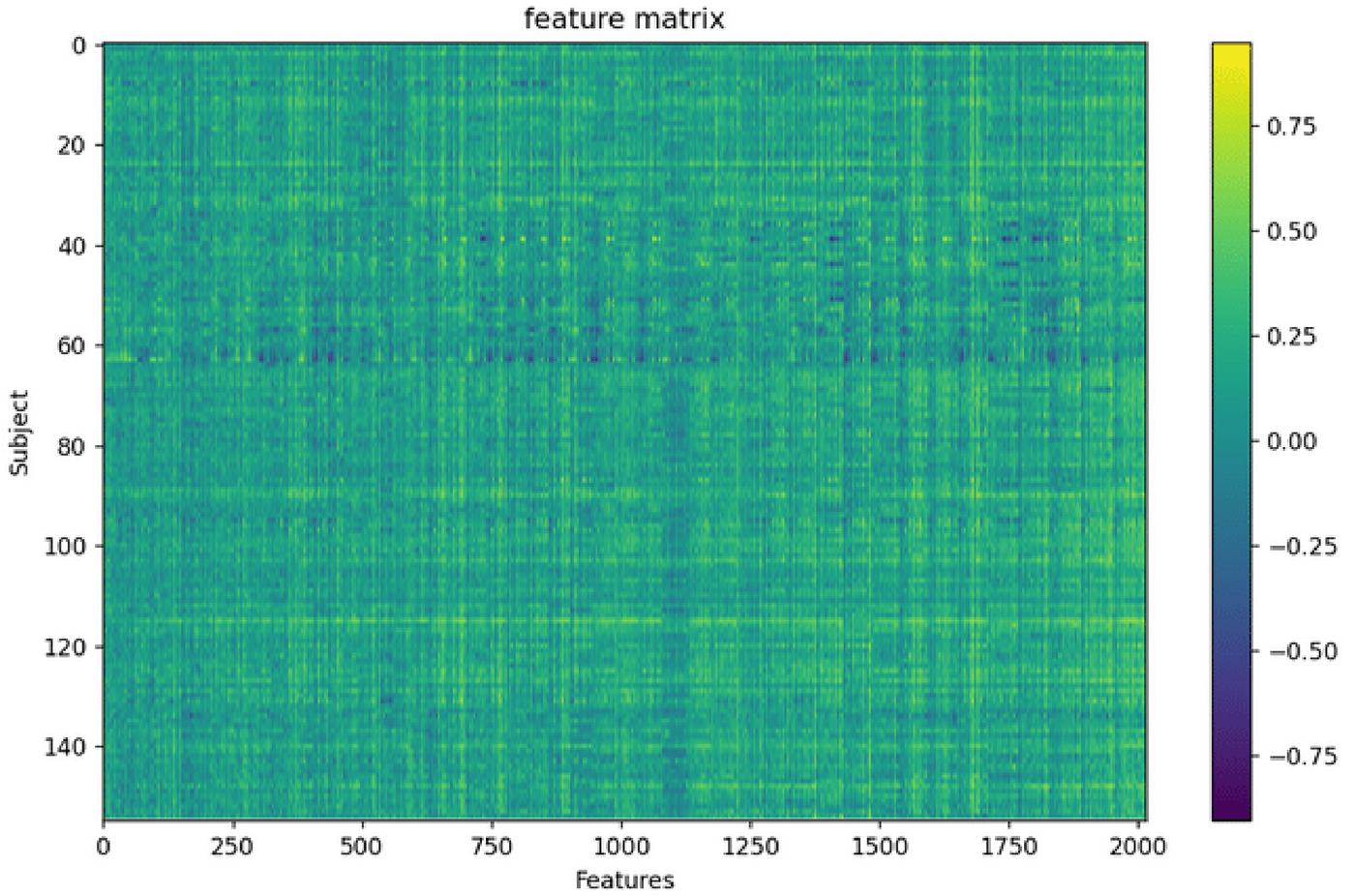


Connectome



Features

155 x
2016



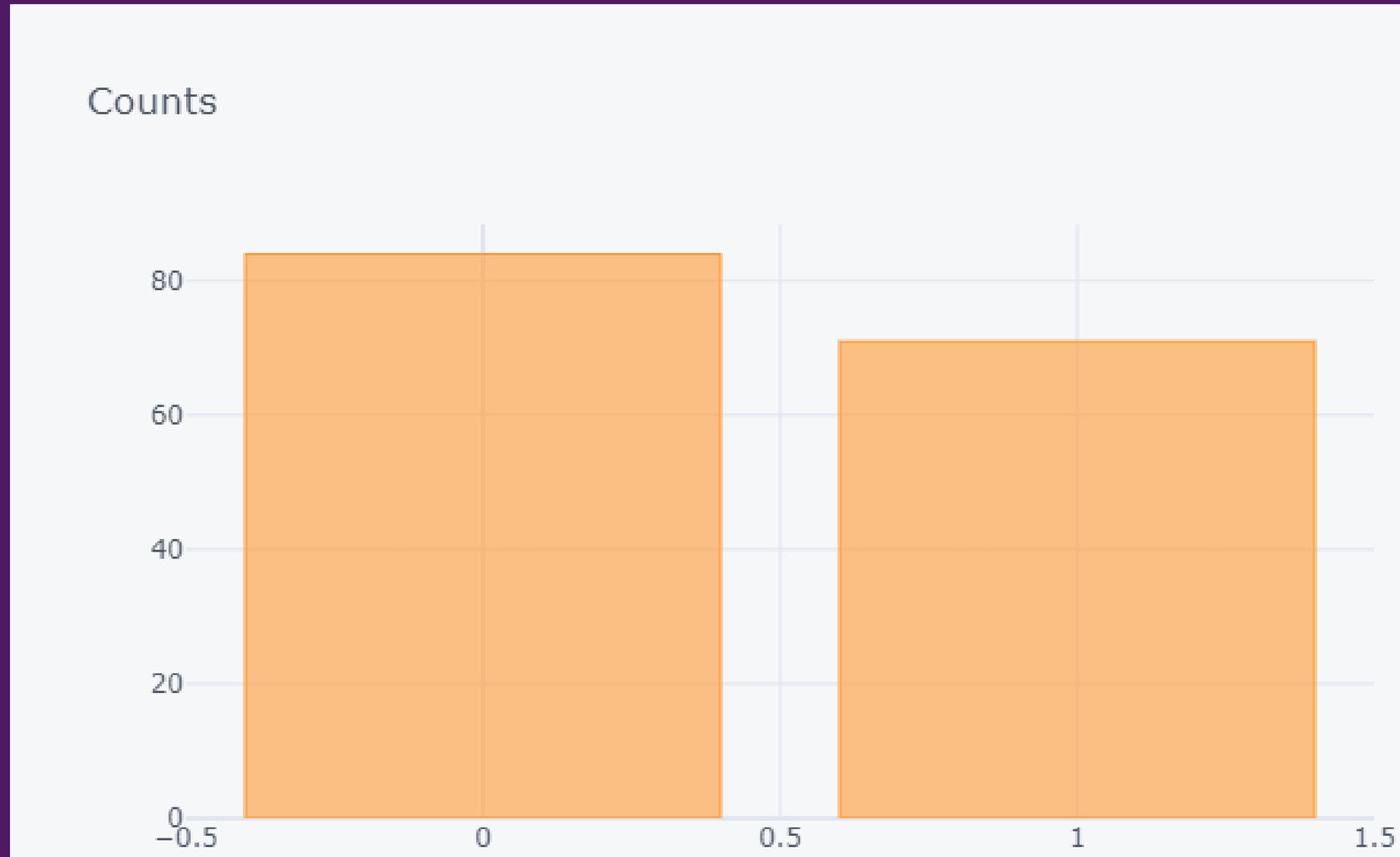
LABEL ENCODING

Label Encoder

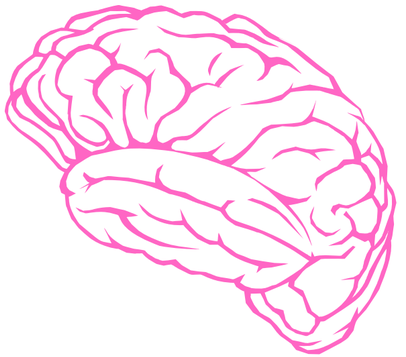
Female → 0

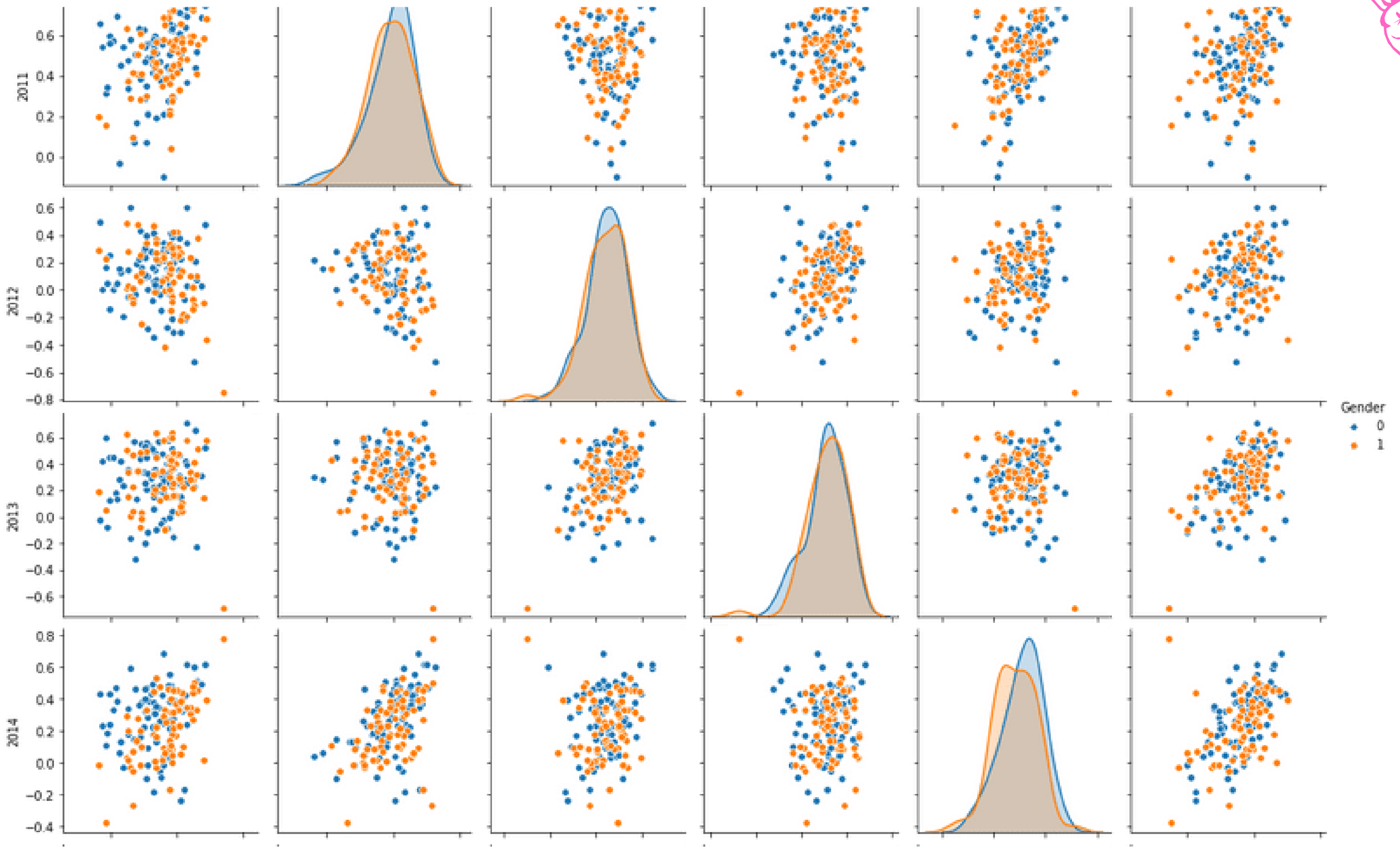
Male → 1

Gender Count



Pair Plot





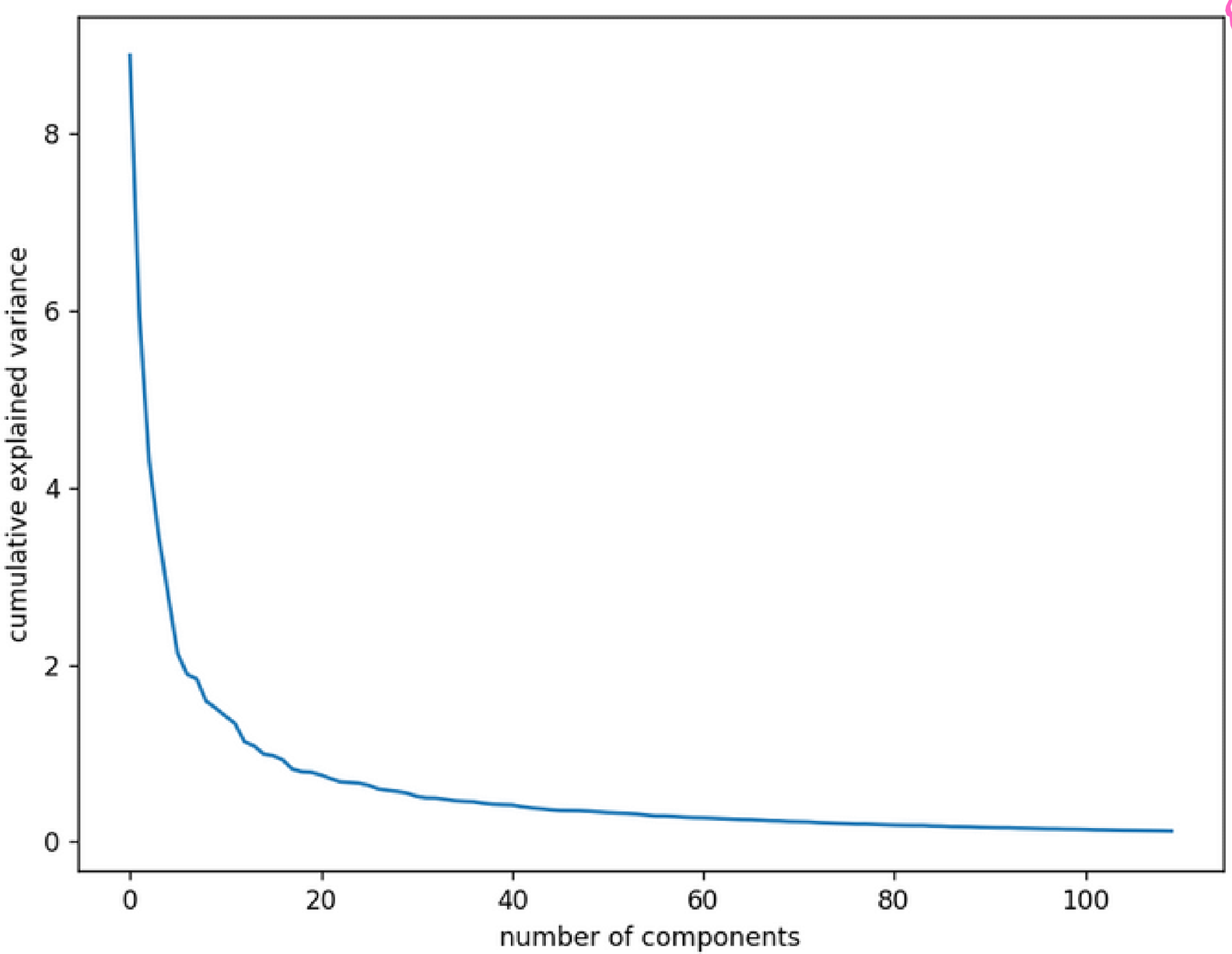
Dimensionality Reduction

155 x 2016

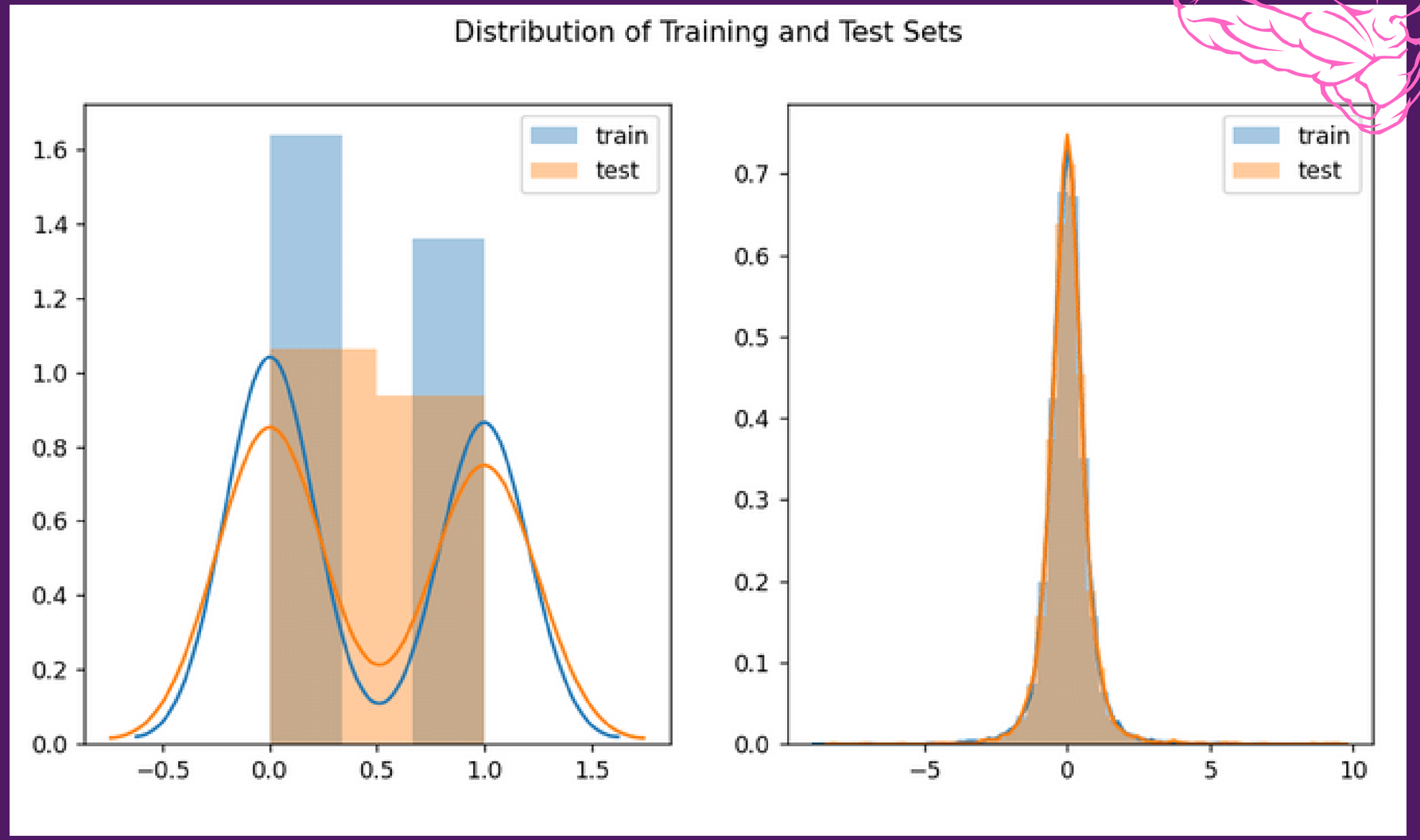
Principal Component Analysis

155 x 110

Scree Plot


cumulative explained variance
0
2
4
6
8
0
20
40
60
80
100
number of components

Training and Test Dataset



CLASSIFICATION

CLASSIFIERS

- Logistic Regression
- K-Nearest Neighbors
- Multilayer Perceptron
- Gaussian Naive Bayes
- Decision Tree
- Random Forest

CROSS VALIDATION

10 FOLD

PERFORMANCE METRICS

- Accuracy
- ROC-AUC Curve
- F1 Score
- Confusion Matrix

Evaluation Metrics

Classification Model	Accuracy	AUC	F1 Score	True Positive	False Positive	True Negative	False Negative
Logistic Regression	0.58	0.61	0.51	25	24	38	21
K-Nearest Neighbors	0.51	0.49	0.47	23	26	32	27
Multilayer Perceptron	0.6	0.59	0.47	20	29	45	14
Gaussian Naïve Bayes	0.47	0.48	0.32	16	33	35	24
Decision Tree	0.55	0.54	0.52	25	24	30	29
Random Forest	0.55	0.52	0.35	18	31	43	16

FUTURE GOALS

- Improving classification accuracy by:
- Selecting those networks whose connectivity differ between male and female
- Using advanced ML algorithms like PLS
- Choosing different atlases

- Explore Nilearn => RegionExtractor, Masker, plotting, atlases, seed based FC

- Containerization using Docker

- More of Interactive plotting

Thank you Brainhack!