WiDS - Datathon [Kaggle Competition]

* Logistic Regression
* Decision Trees
* Random Classifier
* All the datasets have to be merged (both of training dataset and test dataset)
  + One common column between all the datasets
  + participant\_Id: key column
* Feature engineering and hyperparameter tuning

Start with!

* Clean the dataset
* Baseline model: test more models
  + meta data
  + Categorical data

Binary Classification

First training: supervised learning

y : solution test

x : remaining dataset training

Pretrained model

* Must be able to fine-tune it to give better results

Identifying which functional brain network connections are associated with ADHD

Build a model that simultaneously predicts:

* Whether an individual has ADHD
* Whether the individual is female

Approach for the problem

1. Explore the dataset
   1. Functional connectivity matrices (representing brain activity relationships and labels for ADHD and biological sex
   2. Check for missing values and preprocess them
2. Visualizing brain network patterns
   1. p