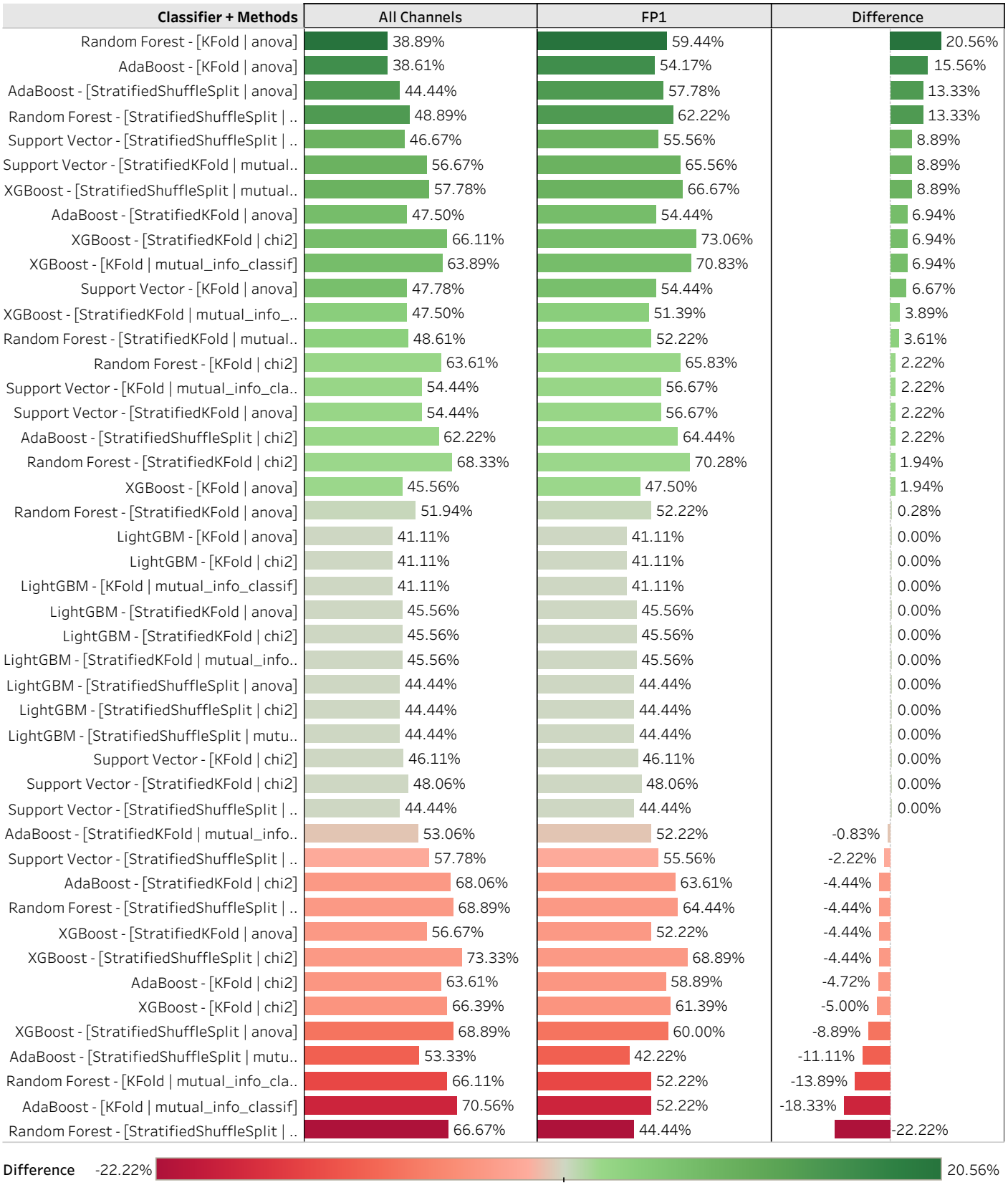
Predictions with EEG Data for Alzheimer Disease and Frontotemporal Dementia

This analysis aims to compare the predictions of cognitive diseases such as Alzheimer's and Frontotemporal Dementia using different Machine Learning models. The study seeks to evaluate how the EEG channel FP1 affects the accuracy of predictions. The results will inform future research where samples are taken using BCI devices that provide readings only from the frontotemporal channels.

Project: https://github.com/sobieddch90/mcd-udg-tfm-eeg-classification
Source Dataset: https://openneuro.org/datasets/ds004504/versions/1.0.5

Predictions with All Features

Feature Extraction from the EEG Data was performed by calculating the signal power with PSD and then calculating statistical metrics for each channel, the five frequency bands, and 60-second windows with an overlap of 20 seconds.



Predictions only Total Power and Spectral Density Features

From the previous EEG Data Feature Extraction, only Total Power and Spectral Density metrics were taken which represent the total power and variation in each channel, frequency band, and window.

Classifier + Methods	All Channels	FP1	Difference
Support Vector - [StratifiedShuffleSplit | ..	57.78%	66.67%	8.89%
XGBoost - [StratifiedKFold | anova]	45.56%	52.22%	6.67%
Support Vector - [StratifiedKFold | mutual..	42.50%	47.78%	5.28%
Support Vector - [StratifiedKFold | anova]	52.22%	56.67%	4.44%
XGBoost - [StratifiedKFold | mutual_info_..	51.67%	55.00%	3.33%
XGBoost - [StratifiedShuffleSplit | anova]	60.00%	62.22%	2.22%
Random Forest - [KFold | anova]	38.89%	41.11%	2.22%
AdaBoost - [KFold | anova]	38.89%	41.11%	2.22%
Support Vector - [KFold | anova]	45.56%	47.78%	2.22%
AdaBoost - [StratifiedKFold | anova]	43.33%	45.56%	2.22%
Random Forest - [StratifiedShuffleSplit | ..	44.44%	46.67%	2.22%
Support Vector - [StratifiedShuffleSplit | ..	44.44%	44.44%	0.00%
Support Vector - [StratifiedKFold | chi2]	48.06%	48.06%	0.00%
Support Vector - [KFold | chi2]	46.11%	46.11%	0.00%
Random Forest - [StratifiedShuffleSplit | ..	44.44%	44.44%	0.00%
Random Forest - [StratifiedKFold | chi2]	42.78%	42.78%	0.00%
Random Forest - [StratifiedKFold | anova]	45.56%	45.56%	0.00%
Random Forest - [KFold | chi2]	54.72%	54.72%	0.00%
LigthGBM - [StratifiedShuffleSplit | mutu..	44.44%	44.44%	0.00%
LigthGBM - [StratifiedShuffleSplit | chi2]	44.44%	44.44%	0.00%
LigthGBM - [StratifiedShuffleSplit | anova]	44.44%	44.44%	0.00%
LigthGBM - [StratifiedKFold | mutual_info..	45.56%	45.56%	0.00%
LigthGBM - [StratifiedKFold | chi2]	45.56%	45.56%	0.00%
LigthGBM - [StratifiedKFold | anova]	45.56%	45.56%	0.00%
LigthGBM - [KFold | mutual_info_classif]	41.11%	41.11%	0.00%
LigthGBM - [KFold | chi2]	41.11%	41.11%	0.00%
LigthGBM - [KFold | anova]	41.11%	41.11%	0.00%
AdaBoost - [StratifiedShuffleSplit | mutu..	44.44%	44.44%	0.00%
AdaBoost - [StratifiedShuffleSplit | chi2]	48.89%	48.89%	0.00%
AdaBoost - [StratifiedShuffleSplit | anova]	44.44%	44.44%	0.00%
AdaBoost - [StratifiedKFold | mutual_info..	45.56%	45.56%	0.00%
AdaBoost - [StratifiedKFold | chi2]	43.06%	43.06%	0.00%
AdaBoost - [KFold | mutual_info_classif]	41.11%	41.11%	0.00%
AdaBoost - [KFold | chi2]	41.39%	41.39%	0.00%
Support Vector - [KFold | mutual_info_cla..	47.78%	45.56%	-2.22%
Random Forest - [KFold | mutual_info_cla..	43.33%	41.11%	-2.22%
XGBoost - [KFold | mutual_info_classif]	64.17%	59.17%	-5.00%
Random Forest - [StratifiedKFold | mutual..	52.22%	45.56%	-6.67%
XGBoost - [StratifiedShuffleSplit | chi2]	53.33%	46.67%	-6.67%
Random Forest - [StratifiedShuffleSplit | ..	57.78%	48.89%	-8.89%
XGBoost - [StratifiedShuffleSplit | mutual..	68.89%	55.56%	-13.33%
XGBoost - [KFold | anova]	58.89%	45.56%	-13.33%
Support Vector - [StratifiedShuffleSplit | ..	64.44%	51.11%	-13.33%
XGBoost - [KFold | chi2]	56.67%	38.33%	-18.33%
XGBoost - [StratifiedKFold | chi2]	52.22%	31.94%	-20.28%

Difference -20.28% 8.89%

Predictions only Features from Alpha Band

From the main Feature Extraction of the EEG Data, only the corresponding metrics to the Alpha frequency band were taken, because this represents the region with the greatest alterations in patients with Alzheimer's Disease and Frontotemporal Dementia.

Classifier + Methods	All Channels	FP1	Difference
Support Vector - [StratifiedKFold | mutual..	40.83%	59.44%	18.61%
Random Forest - [StratifiedShuffleSplit | ..	40.00%	55.56%	15.56%
Random Forest - [StratifiedShuffleSplit | ..	42.22%	57.78%	15.56%
AdaBoost - [StratifiedShuffleSplit | mutu..	40.00%	53.33%	13.33%
XGBoost - [StratifiedShuffleSplit | mutual..	31.11%	44.44%	13.33%
Random Forest - [KFold | anova]	43.33%	54.72%	11.39%
XGBoost - [StratifiedShuffleSplit | chi2]	42.22%	53.33%	11.11%
AdaBoost - [KFold | anova]	36.67%	45.83%	9.17%
AdaBoost - [StratifiedShuffleSplit | anova]	42.22%	51.11%	8.89%
Random Forest - [StratifiedShuffleSplit | ..	33.33%	42.22%	8.89%
Support Vector - [KFold | mutual_info_cla..	45.28%	50.28%	5.00%
Support Vector - [StratifiedKFold | anova]	45.28%	48.06%	2.78%
XGBoost - [KFold | anova]	54.72%	56.67%	1.94%
AdaBoost - [StratifiedKFold | chi2]	45.56%	46.94%	1.39%
Support Vector - [StratifiedShuffleSplit | ..	48.89%	48.89%	0.00%
Support Vector - [StratifiedShuffleSplit | ..	44.44%	44.44%	0.00%
Support Vector - [StratifiedShuffleSplit | ..	44.44%	44.44%	0.00%
Support Vector - [StratifiedKFold | chi2]	48.06%	48.06%	0.00%
Support Vector - [KFold | chi2]	46.11%	46.11%	0.00%
LigthGBM - [StratifiedShuffleSplit | mutu..	44.44%	44.44%	0.00%
LigthGBM - [StratifiedShuffleSplit | chi2]	44.44%	44.44%	0.00%
LigthGBM - [StratifiedShuffleSplit | anova]	44.44%	44.44%	0.00%
LigthGBM - [StratifiedKFold | mutual_info..	45.56%	45.56%	0.00%
LigthGBM - [StratifiedKFold | chi2]	45.56%	45.56%	0.00%
LigthGBM - [StratifiedKFold | anova]	45.56%	45.56%	0.00%
LigthGBM - [KFold | mutual_info_classif]	41.11%	41.11%	0.00%
LigthGBM - [KFold | chi2]	41.11%	41.11%	0.00%
LigthGBM - [KFold | anova]	41.11%	41.11%	0.00%
XGBoost - [KFold | mutual_info_classif]	52.50%	51.67%	-0.83%
AdaBoost - [KFold | mutual_info_classif]	40.28%	38.33%	-1.94%
AdaBoost - [StratifiedShuffleSplit | chi2]	48.89%	46.67%	-2.22%
AdaBoost - [StratifiedKFold | anova]	52.22%	47.50%	-4.72%
XGBoost - [StratifiedKFold | chi2]	50.00%	45.28%	-4.72%
Random Forest - [StratifiedKFold | anova]	52.22%	45.83%	-6.39%
Random Forest - [StratifiedKFold | chi2]	54.44%	48.06%	-6.39%
Random Forest - [StratifiedKFold | mutual..	50.00%	43.06%	-6.94%
AdaBoost - [KFold | chi2]	52.78%	45.83%	-6.94%
Support Vector - [KFold | anova]	54.72%	46.11%	-8.61%
XGBoost - [KFold | chi2]	56.94%	47.78%	-9.17%
Random Forest - [KFold | mutual_info_cla..	46.94%	36.11%	-10.83%
Random Forest - [KFold | chi2]	56.67%	45.83%	-10.83%
XGBoost - [StratifiedKFold | mutual_info_..	49.72%	38.33%	-11.39%
AdaBoost - [StratifiedKFold | mutual_info..	56.67%	43.06%	-13.61%
XGBoost - [StratifiedShuffleSplit | anova]	64.44%	48.89%	-15.56%
XGBoost - [StratifiedKFold | anova]	68.06%	47.50%	-20.56%

Difference -20.56% 18.61%

Predictions without Frequency Bands

Similar to the first feature extraction from the EEG data, statistical metrics were calculated from the power of the signal obtained with PSD, at this time without separating the frequency bands.