Use data include all patients but not the whole dataset. Not shuffling and split the data to

20% for test and 80% for next split. Then shuffle the rest and split it in 20% for validation and

80% for training. Features extracted from 8 channels. During training is early stopping used. If valid set is not improved in 30 rounds, training will be stopped.

## Test different window size

Here test 4 window sizes: 64, 128, 256, 512

The stride is 1/4 of the window size.

accuracy

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | W64 | W128 | W256 | W512 |
| train | 0.983459 | 0.989837 | 0.985994 | 0.999424 |
| valid | 0.921379 | 0.93305 | 0.943616 | 0.951997 |
| test | 0.895815 | 0.915267 | 0.939756 | 0.931797 |
| Rest data | 0.898510 | 0.933318 | 0.954702 | 0.954827 |

In generally the accuracy get higher with increasing of the window size, but it will not effect to much after window size 256. Window size 256 has the highest accuracy of test set and the accuracy on the rest data is also closed to the highest accuracy.

Test different stride: 1/8, 1/4, 1/2 of the window size 256

accuracy

|  |  |  |  |
| --- | --- | --- | --- |
|  | S32 | S64 | S128 |
| train | 0.999272 | 0.985994 | 0.972823 |
| valid | 0.973888 | 0.943616 | 0.915667 |
| test | 0.935507 | 0.939756 | 0.931085 |

## compare features from left and right limb

Train the model on data from left or right limb.

Window size: 256 stride: 64

accuracy

|  |  |  |
| --- | --- | --- |
|  | Right | Left |
| train | 0.986858 | 0.987540 |
| valid | 0.920335 | 0.912332 |
| test | 0.862776 | 0.900466 |
| Rest data | 0.894972 | 0.910650 |

The feature from left limb has better performance on test set and the rest data.

## Compare Feature from each 2 channels

Train the model on data from each 2 channels.

Window size: 256 stride: 64

top 3 accuracy on test set

|  |  |  |  |
| --- | --- | --- | --- |
|  | LEFT\_RF\_RIGHT\_BF | LEFT\_BF\_RIGHT\_BF | LEFT\_BF\_RIGHT\_RF |
| test | 0.911816 | 0.865978 | 0.859721 |

Drop one signal from 8 out

|  |  |  |  |
| --- | --- | --- | --- |
|  | train | valid | test |
| drop\_LEFT\_TA | 0.994088 | 0.947981 | 0.92826 |
| drop\_LEFT\_TS | 0.976671 | 0.936886 | 0.930297 |
| drop\_LEFT\_BF | 0.999454 | 0.948527 | 0.92142 |
| drop\_LEFT\_RF | 0.979854 | 0.934522 | 0.9367 |
| drop\_RIGHT\_TA | 0.99286 | 0.94307 | 0.945431 |
| drop\_RIGHT\_TS | 0.989177 | 0.94016 | 0.941211 |
| drop\_RIGHT\_BF | 0.997499 | 0.947435 | 0.912835 |
| drop\_RIGHT\_RF | 0.990086 | 0.941433 | 0.945867 |

After dropping RIGHT\_BF, LEFT\_BF, LEFT\_TA the accuracy on test set get lowest. When testing on data from each 2 channels, the top 3 accuracy on test set are include RIGHT\_BF, LEFT\_BF.

## Parameter

model = xgb.XGBClassifier(max\_depth=4,

learning\_rate=0.3,

n\_estimators=1000,

silent=True,

eval\_metrics='error',

objective='binary:logistic',

seed=100,

sub\_sabsample=0.8,

reg\_lambda = 15,

)