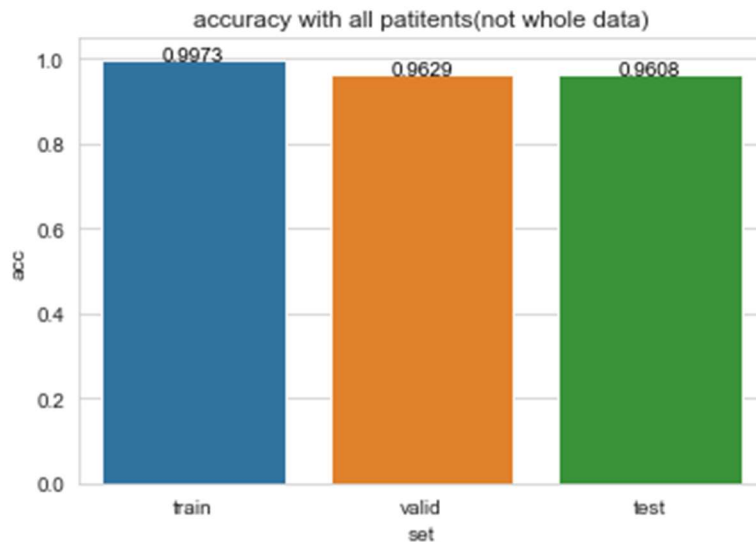# Test1

Use data include all patients but not the whole dataset. Not shuffling then 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.

No: 0,　Yes: not 0

Confusion matrix of all data(train,valid,test):

	Predict No	Predict Yes
Actual No	30354	626
Actual Yes	94	15360

Confusion matrix of valid set:

	Predict No	Predict Yes
Actual No	4162	193
Actual Yes	83	2992

Confusion matrix of test set:

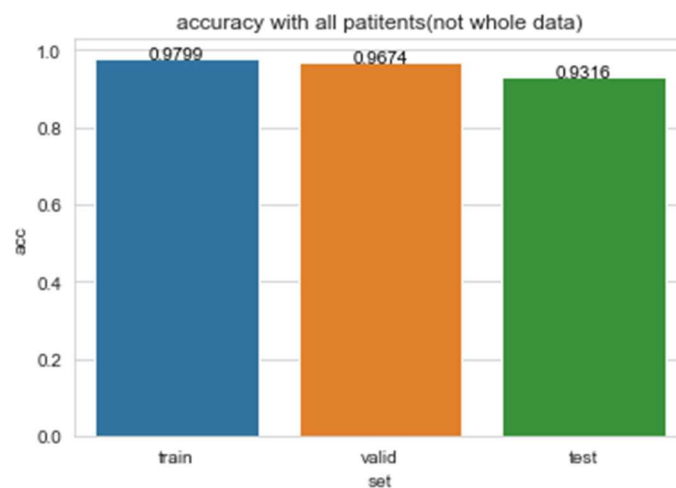|	Predict No	Predict Yes
Actual No	8923	364
Actual Yes	0	0

Test the model on some of the rest dataset which are not used. The rest data's labels are all 0.
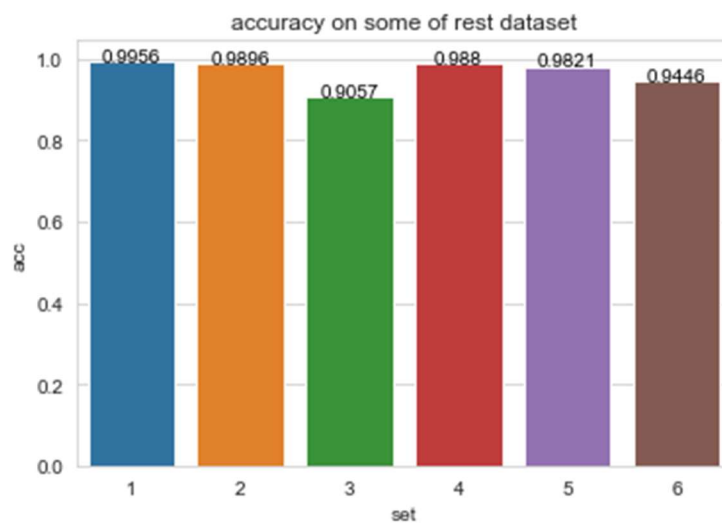
Test2

Use data include all patients but not the whole dataset. Not shuffling then split the data to 20% for test and 80%for next split. The rest data are not shuffled either and split it in 20% for validation and 80% for training.

No: 0,　Yes: not 0

Confusion matrix of all data(train,valid,test):

	Predict No	Predict Yes
Actual No	29620	1360
Actual Yes	115	15339

Confusion matrix of valid set:

	Predict No	Predict Yes
Actual No	7188	242
Actual Yes	0	0

Confusion matrix of test set:

	Predict No	Predict Yes
Actual No	8652	635
Actual Yes	0	0

Test the model on some of the rest dataset which are not used. The rest data's labels are all 0.

accuracy on some of rest dataset
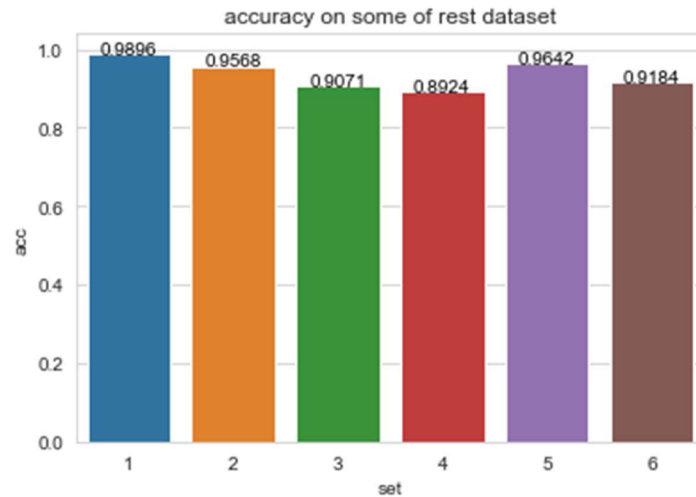
Parameter

Windowsize: 256
Step: 64
Threshold for wilson amplitude: 30
Threshold for zero crossing: 0
Threshold for slope Sign Change: 1
EMG Histogram: bins = 9, range=(-70,70)
model = xgb.XGBClassifier(max_depth=4,
                          learning_rate=0.3,
                          n_estimators=1000,
                          eval_metrics='error',
                          objective='binary:logistic',
                          seed=100,
                          sub_sabsample=0.8,
                          reg_lambda = 15,
)