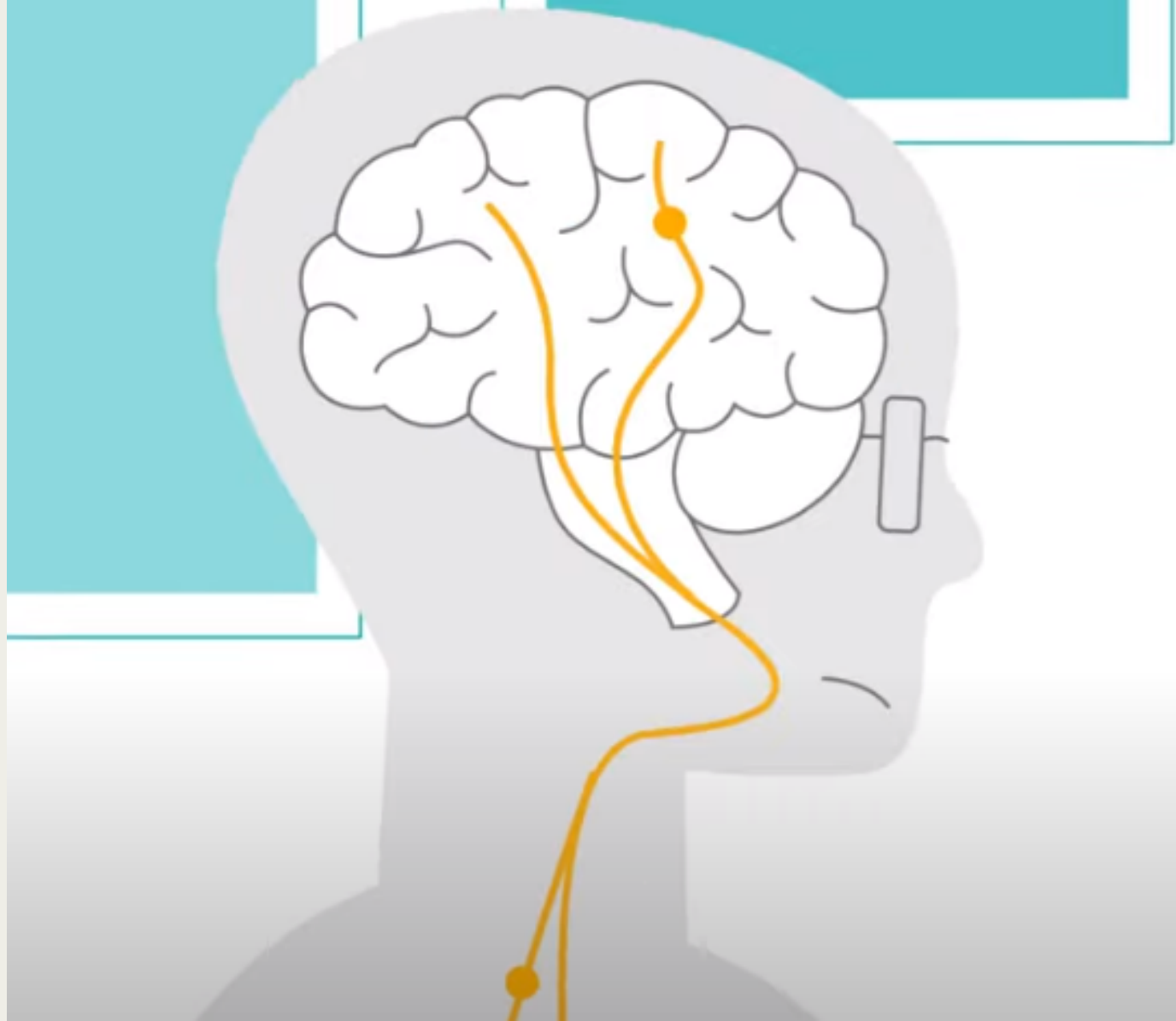# Parkinson's FoG prediction

HARSH YADAV

Goal

Our goal was to **detect freezing of gait (FOG)**, a debilitating symptom that afflicts many people with Parkinson's disease. We tested several machine learning models trained on data collected from a wearable 3D lower back sensors.

Our work will help researchers better understand **when and why FOG episodes occur**. This will improve the ability of medical professionals to optimally evaluate, monitor, and ultimately, prevent FOG events.

Context

An estimated **7 to 10 million** people around the world have Parkinson's disease, many of whom suffer from freezing of gait (FOG). **During a FOG episode, a patient's feet are "glued" to the ground, preventing them from moving forward despite their attempts.**

FOG has a profound negative impact on health-related quality of life—people who suffer from FOG are often **depressed**, have an increased **risk of falling**, are likelier to be confined to **wheelchair use**, and have **restricted independence**.

While researchers have multiple theories to explain when, why, and in whom FOG occurs, there is still no clear understanding of its causes.

- There are many methods of evaluating FOG, though most involve FOG-provoking protocols. **People with FOG are filmed while performing certain tasks** that are likely to increase its occurrence. Experts then **review the video** to score each frame, indicating when FOG occurred. While scoring in this manner is relatively **reliable and sensitive**, it is extremely **time-consuming** and requires specific expertise.
- Another method involves augmenting FOG-provoking testing with **wearable devices**. With more sensors, the detection of FOG becomes **easier**, however, compliance and **usability may be reduced**.

Therefore, a ***combination of these two methods*** may be the best approach. When combined with machine learning methods, the accuracy of detecting FOG from a lower back accelerometer is relatively high.

However, the datasets used to train and test these algorithms have been relatively small and generalizability is limited to date. Furthermore, the emphasis has been on achieving high levels of accuracy, while precision, for example, has largely been ignored.
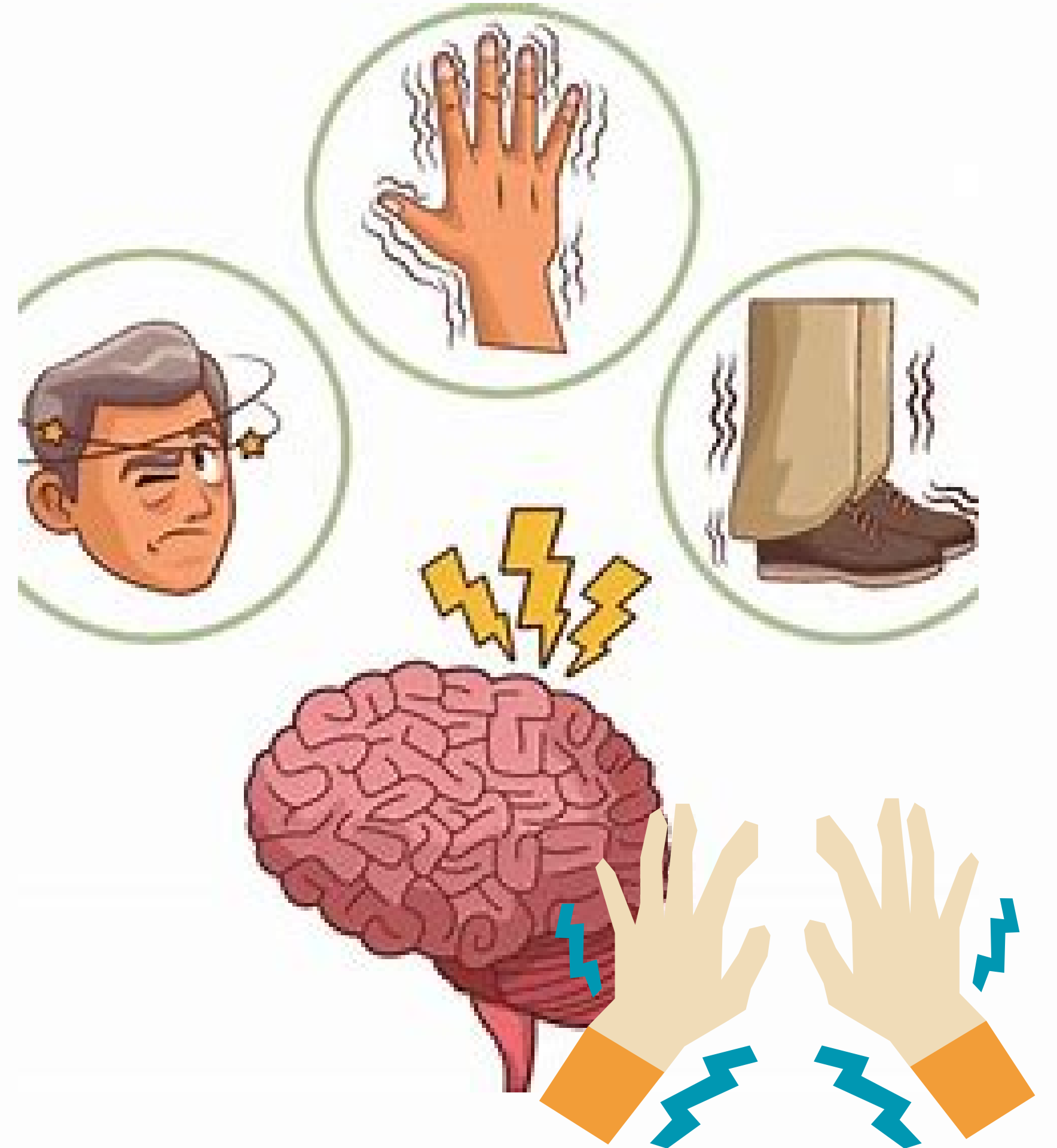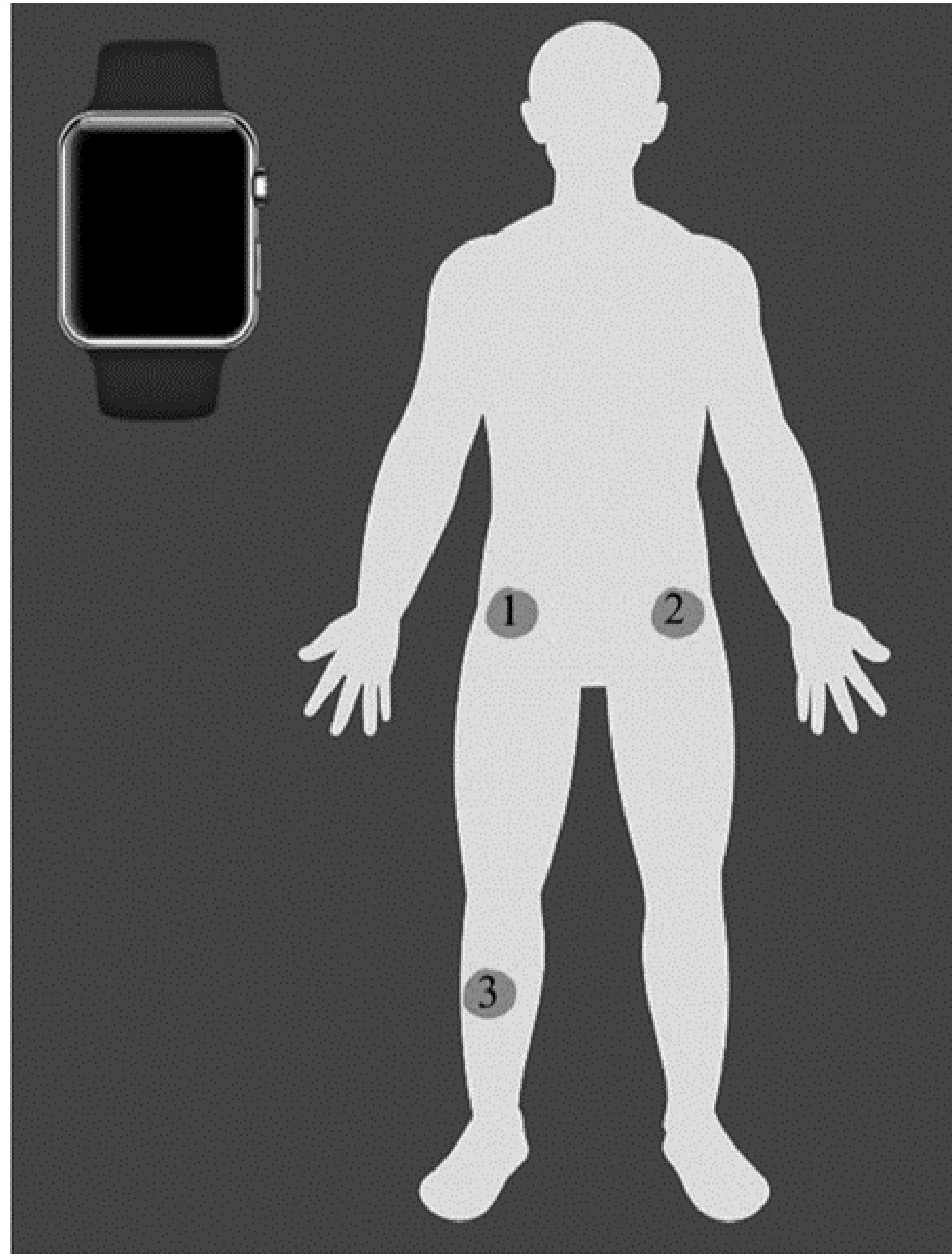
Dataset

DeFOG

Data collected in Subject's home environment🏠.

Both in "ON" and "OFF" medication state.

tDCS FOG

Data collected in the LAB🥼

Both in "ON" and "OFF" medication state.

Tasks

1. **4-meter walk test**
2. **Timed Up & Go (TUG)**
3. **Timed Up & Go (TUG) –** Dual-task (subtracting numbers while performing the TUG test)
4. **Turning task with alternating directions** (performing 4 x 360 degrees turns, each time alternating the rotation direction).
5. **Turning task –** Dual-task (same as before, but with additional number subtraction task).
6. **Hotspot Door –** A walking trial that involves opening a door, entering another room, turning, and returning to the start point.
7. **Personalized Hotspot -** walking through an area in the house that the subject describes as FoG provoking.

Trials from the tdcsfog and defog datasets were videotaped and annotated by expert reviewers documented the freezing of gait episodes. That is, the start, end and type of each episode were marked by the experts.

The Dataset comprised of **45 FoG+ PD patient** .

What we did ?

Pre-Processing

defog-defog

	Time	AccV	AccML	AccAP	StartHesitation	Turn	Walking	Valid	Task	file
0	0	-1.002697	0.022371	0.068304	0	0	0	False	False	be9d33541d
1	1	-1.002641	0.019173	0.066162	0	0	0	False	False	be9d33541d
2	2	-0.999820	0.019142	0.067536	0	0	0	False	False	be9d33541d
3	3	-0.998023	0.018378	0.068409	0	0	0	False	False	be9d33541d
4	4	-0.998359	0.016726	0.066448	0	0	0	False	False	be9d33541d
...	...	...	...	...	...	...	...	...	...	...
109120	109120	-0.939241	0.031564	-0.394737	0	0	0	False	False	06414383cf
109121	109121	-0.941096	0.031582	-0.392626	0	0	0	False	False	06414383cf
109122	109122	-0.940131	0.029092	-0.394385	0	0	0	False	False	06414383cf
109123	109123	-0.939872	0.028058	-0.398664	0	0	0	False	False	06414383cf
109124	109124	-0.939006	0.026628	-0.398454	0	0	0	False	False	06414383cf

13525702 rows × 10 columns

defog-metadata

	Id	Subject	Visit	Medication
0	02ab235146	e1f62e	2	on
1	02ea782681	ae2d35	2	on
2	06414383cf	8c1f5e	2	off
3	092b4c1819	2874c5	1	off
4	0a900ed8a2	0e3d49	2	on
...	...	...	...	...
132	f3a921edee	1a778d	1	off
133	f40e8c6ebe	575c60	1	off
134	f8ddbdd98d	107712	1	on
135	f9efef91fb	5d9cae	2	off
136	f9fc61ce85	040587	1	on

137 rows × 4 columns

```
In [12]:   defog_m= defog_metadata.merge(defog, how = 'inner', left_on = 'Id', right_on = 'file')
           defog_m.drop(['file','Valid','Task'], axis = 1, inplace = True)
           defog_m
```

Out[12]:

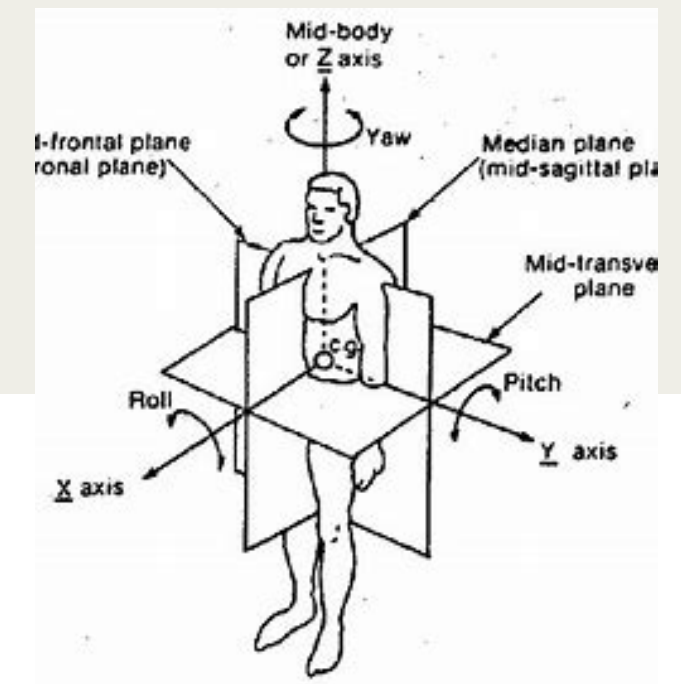|	Id	Subject	Visit	Medication	Time	AccV	AccML	AccAP	StartHesitation	Turn	Walking
0	02ea782681	ae2d35	2	on	1000	-0.970018	0.061626	-0.265625	0	0	0
1	02ea782681	ae2d35	2	on	1001	-0.984375	0.044497	-0.265625	0	0	0
2	02ea782681	ae2d35	2	on	1002	-0.984375	0.029016	-0.265625	0	0	0
3	02ea782681	ae2d35	2	on	1003	-0.984375	0.015625	-0.265625	0	0	0
4	02ea782681	ae2d35	2	on	1004	-0.984670	0.015330	-0.265625	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...
4090525	f9fc61ce85	040587	1	on	119027	-0.961216	0.142428	-0.289655	0	0	0
4090526	f9fc61ce85	040587	1	on	119028	-0.960343	0.142836	-0.290506	0	0	0
4090527	f9fc61ce85	040587	1	on	119029	-0.957958	0.145494	-0.290007	0	0	0
4090528	f9fc61ce85	040587	1	on	119030	-0.960616	0.145839	-0.291527	0	0	0
4090529	f9fc61ce85	040587	1	on	119031	-0.967076	0.144342	-0.292384	0	0	0

4090530 rows × 11 columns

Summary

data shape: (4090530, 11)

Out[14]:

	data type	#missing	%missing	#unique	min	max	first value	second value	third value
Id	object	0	0.0	91	NaN	NaN	02ea782681	02ea782681	02ea782681
Subject	object	0	0.0	38	NaN	NaN	ae2d35	ae2d35	ae2d35
Visit	int64	0	0.0	2	1.0	2.0	2	2	2
Medication	object	0	0.0	2	NaN	NaN	on	on	on
Time	int64	0	0.0	338197	1000.0	414387.0	1000	1001	1002
AccV	float64	0	0.0	3485229	-6.024701	4.458365	-0.970018	-0.984375	-0.984375
AccML	float64	0	0.0	3539384	-2.115008	4.524038	0.061626	0.044497	0.029016
AccAP	float64	0	0.0	3437121	-5.11865	4.388132	-0.265625	-0.265625	-0.265625
StartHesitation	int64	0	0.0	2	0.0	1.0	0	0	0
Turn	int64	0	0.0	2	0.0	1.0	0	0	0
Walking	int64	0	0.0	2	0.0	1.0	0	0	0

Feature Engineering

	Id	Subject	Visit	Medication	Time	AccV	AccML	AccAP	StartHesitation	Turn	Walking	event
0	02ea782681	ae2d35	2	on	1000	-0.970018	0.061626	-0.265625	0	0	0	Normal
1	02ea782681	ae2d35	2	on	1001	-0.984375	0.044497	-0.265625	0	0	0	Normal
2	02ea782681	ae2d35	2	on	1002	-0.984375	0.029016	-0.265625	0	0	0	Normal
3	02ea782681	ae2d35	2	on	1003	-0.984375	0.015625	-0.265625	0	0	0	Normal
4	02ea782681	ae2d35	2	on	1004	-0.984670	0.015330	-0.265625	0	0	0	Normal
...	...	...	...	...	...	...	...	...	...	...	...	...
4090525	f9fc61ce85	040587	1	on	119027	-0.961216	0.142428	-0.289655	0	0	0	Normal
4090526	f9fc61ce85	040587	1	on	119028	-0.960343	0.142836	-0.290506	0	0	0	Normal
4090527	f9fc61ce85	040587	1	on	119029	-0.957958	0.145494	-0.290007	0	0	0	Normal
4090528	f9fc61ce85	040587	1	on	119030	-0.960616	0.145839	-0.291527	0	0	0	Normal
4090529	f9fc61ce85	040587	1	on	119031	-0.967076	0.144342	-0.292384	0	0	0	Normal

4090530 rows × 12 columns

Train-Test Split

In [31]:
```
X
```

Out[31]:

	AccV	AccML	AccAP
0	-0.970018	0.061626	-0.265625
1	-0.984375	0.044497	-0.265625
2	-0.984375	0.029016	-0.265625
3	-0.984375	0.015625	-0.265625
4	-0.984670	0.015330	-0.265625
...	...	...	...
4090525	-0.961216	0.142428	-0.289655
4090526	-0.960343	0.142836	-0.290506
4090527	-0.957958	0.145494	-0.290007
4090528	-0.960616	0.145839	-0.291527
4090529	-0.967076	0.144342	-0.292384

4090530 rows × 3 columns

```
y
```

```
0          0
1          0
2          0
3          0
4          0
          ..
4090525    0
4090526    0
4090527    0
4090528    0
4090529    0
Name: target, Length: 4090530, dtype: int64
```

```
y.unique()
```

```
array([0, 2, 3, 1])
```

```
y.value_counts()
```

```
0    3404683
2     586829
3      98518
1        500
Name: target, dtype: int64
```

```
# spliting dataset into training and test set
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=1004)
```

Classifiers

1. **Decision Tree**
2. **Naive Bayes**
3. **Random Forest**
4. **XG Boost**
5. **LGBM**
6. **KNN**

Decision Tree

Accuracy: 0.7644649144894834
Confusion Matrix:
[[876620   134 118662 25603]
 [124   5   20   3]
 [113695   24 58809   3743]
 [23441   2   3588   2686]]
Classification Report:

	precision	recall	f1-score	support
0	0.86	0.86	0.86	1021019
1	0.03	0.03	0.03	152
2	0.32	0.33	0.33	176271
3	0.08	0.09	0.09	29717
accuracy			0.76	1227159
macro avg	0.33	0.33	0.33	1227159
weighted avg	0.77	0.76	0.77	1227159

Naive Bayes

Accuracy: 0.8320144333374893
Confusion Matrix:
[[1021014   0   5   0]
 [152   0   0   0]
 [176271   0   0   0]
 [29717   0   0   0]]
Classification Report:

	precision	recall	f1-score	support
0	0.83	1.00	0.91	1021019
1	0.00	0.00	0.00	152
2	0.00	0.00	0.00	176271
3	0.00	0.00	0.00	29717
accuracy			0.83	1227159
macro avg	0.21	0.25	0.23	1227159
weighted avg	0.69	0.83	0.76	1227159

Random Forest

```
Accuracy: 0.8394780138515058
Confusion Matrix:
[[987819       1  32449     750]
 [    139      0     13       0]
 [134598       0  41523     150]
 [ 28305       0    581     831]]
Classification Report:
              precision    recall  f1-score   support

           0       0.86      0.97      0.91   1021019
           1       0.00      0.00      0.00       152
           2       0.56      0.24      0.33    176271
           3       0.48      0.03      0.05     29717

    accuracy                           0.84   1227159
   macro avg       0.47      0.31      0.32   1227159
weighted avg       0.81      0.84      0.81   1227159
```

XG Boost

```
Accuracy: 0.8333907831014563
Confusion Matrix:
[[1019172        0    1847        0]
 [     152      0       0        0]
 [  172740      0    3531        0]
 [   29713      0       4        0]]
Classification Report:
              precision    recall  f1-score   support

           0       0.83      1.00      0.91   1021019
           1       0.00      0.00      0.00       152
           2       0.66      0.02      0.04    176271
           3       0.00      0.00      0.00     29717

    accuracy                           0.83   1227159
   macro avg       0.37      0.25      0.24   1227159
weighted avg       0.79      0.83      0.76   1227159
```

LGBM

```
Accuracy: 0.8336955520841227
Confusion Matrix:
[[1017220        0     3799        0]
 [    152        0        0        0]
 [ 170414        0     5857        0]
 [  29701        0       16        0]]
Classification Report:
              precision    recall  f1-score   support

           0       0.84      1.00      0.91   1021019
           1       0.00      0.00      0.00       152
           2       0.61      0.03      0.06    176271
           3       0.00      0.00      0.00     29717

    accuracy                           0.83   1227159
   macro avg       0.36      0.26      0.24   1227159
weighted avg       0.78      0.83      0.77   1227159
```

KNN

```
Accuracy: 0.8303520570683994
Confusion Matrix:
[[972169        4    47315     1531]
 [   141        2        9        0]
 [130247        0    45794      230]
 [ 27661        0     1047     1009]]
Classification Report:
              precision    recall  f1-score   support

           0       0.86      0.95      0.90   1021019
           1       0.33      0.01      0.03       152
           2       0.49      0.26      0.34    176271
           3       0.36      0.03      0.06     29717

    accuracy                           0.83   1227159
   macro avg       0.51      0.31      0.33   1227159
weighted avg       0.79      0.83      0.80   1227159
```

TDCS FOG(🥼)

	Id	Subject	Visit	Test	Medication	Time	AccV	AccML	AccAP	StartHesitation	Turn	Walking
0	003f117e14	4dc2f8	3	2	on	0	-9.533939	0.566322	-1.413525	0	0	0
1	003f117e14	4dc2f8	3	2	on	1	-9.536140	0.564137	-1.440621	0	0	0
2	003f117e14	4dc2f8	3	2	on	2	-9.529345	0.561765	-1.429332	0	0	0
3	003f117e14	4dc2f8	3	2	on	3	-9.531239	0.564227	-1.415490	0	0	0
4	003f117e14	4dc2f8	3	2	on	4	-9.540825	0.561854	-1.429471	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...
7062667	ffda8fadfd	7fcee9	20	1	off	4220	-9.403467	0.089003	-3.220304	0	0	0
7062668	ffda8fadfd	7fcee9	20	1	off	4221	-9.404246	0.090531	-3.216584	0	0	0
7062669	ffda8fadfd	7fcee9	20	1	off	4222	-9.405770	0.084380	-3.224039	0	0	0
7062670	ffda8fadfd	7fcee9	20	1	off	4223	-9.403579	0.084236	-3.236686	0	0	0
7062671	ffda8fadfd	7fcee9	20	1	off	4224	-9.405036	0.082027	-3.234458	0	0	0

7062672 rows × 12 columns

Summary

data shape: (7062672, 12)

	data type	#missing	%missing	#unique	min	max	first value	second value	third value
Id	object	0	0.0	833	NaN	NaN	003f117e14	003f117e14	003f117e14
Subject	object	0	0.0	62	NaN	NaN	4dc2f8	4dc2f8	4dc2f8
Visit	int64	0	0.0	7	2.0	20.0	3	3	3
Test	int64	0	0.0	3	1.0	3.0	2	2	2
Medication	object	0	0.0	2	NaN	NaN	on	on	on
Time	int64	0	0.0	97077	0.0	97076.0	0	1	2
AccV	float64	0	0.0	7027490	-35.521119	20.906953	-9.533939	-9.53614	-9.529345
AccML	float64	0	0.0	7030366	-26.164398	27.484719	0.566322	0.564137	0.561765
AccAP	float64	0	0.0	7028071	-47.829639	30.337694	-1.413525	-1.440621	-1.429332
StartHesitation	int64	0	0.0	2	0.0	1.0	0	0	0
Turn	int64	0	0.0	2	0.0	1.0	0	0	0
Walking	int64	0	0.0	2	0.0	1.0	0	0	0

LGBM

XG Boost

LGBM:

```
Accuracy: 0.7232521672029366
Confusion Matrix:
[[924192     40  50372    192]
 [ 44500     91  15740    453]
 [238070     61  97022    305]
 [ 28600     80  12503    314]]
Classification Report:
```

	precision	recall	f1-score	support
0	0.75	0.95	0.84	974796
1	0.33	0.00	0.00	60784
2	0.55	0.29	0.38	335458
3	0.25	0.01	0.01	41497
accuracy			0.72	1412535
macro avg	0.47	0.31	0.31	1412535
weighted avg	0.67	0.72	0.67	1412535

XG Boost:

```
Accuracy: 0.719957381587005
Confusion Matrix:
[[924753      2  49871    170]
 [ 45221      0  15131    432]
 [243224      0  91902    332]
 [ 29006      1  12180    310]]
Classification Report:
```

	precision	recall	f1-score	support
0	0.74	0.95	0.83	974796
1	0.00	0.00	0.00	60784
2	0.54	0.27	0.36	335458
3	0.25	0.01	0.01	41497
accuracy			0.72	1412535
macro avg	0.38	0.31	0.30	1412535
weighted avg	0.65	0.72	0.66	1412535

KNN

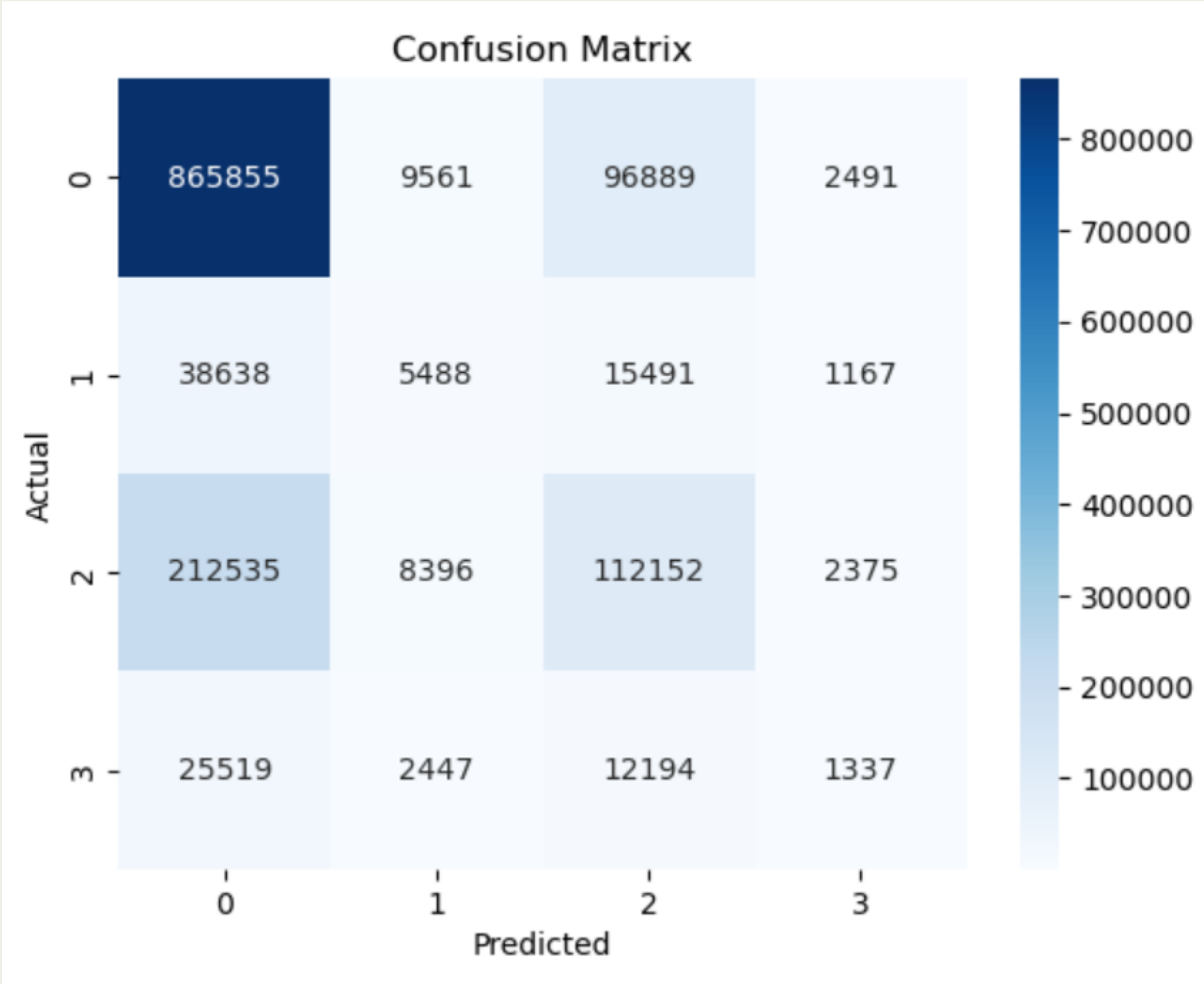```
Accuracy: 0.6972089187170584
Confusion Matrix:
[[865855   9561   96889   2491]
 [38638   5488   15491   1167]
 [212535   8396 112152   2375]
 [25519   2447   12194   1337]]
Classification Report:
              precision   recall f1-score   support

           0   0.76   0.89   0.82   974796
           1   0.21   0.09   0.13   60784
           2   0.47   0.33   0.39   335458
           3   0.18   0.03   0.05   41497

    accuracy   0.70   1412535
   macro avg   0.41   0.34   0.35   1412535
weighted avg   0.65   0.70   0.66   1412535
```
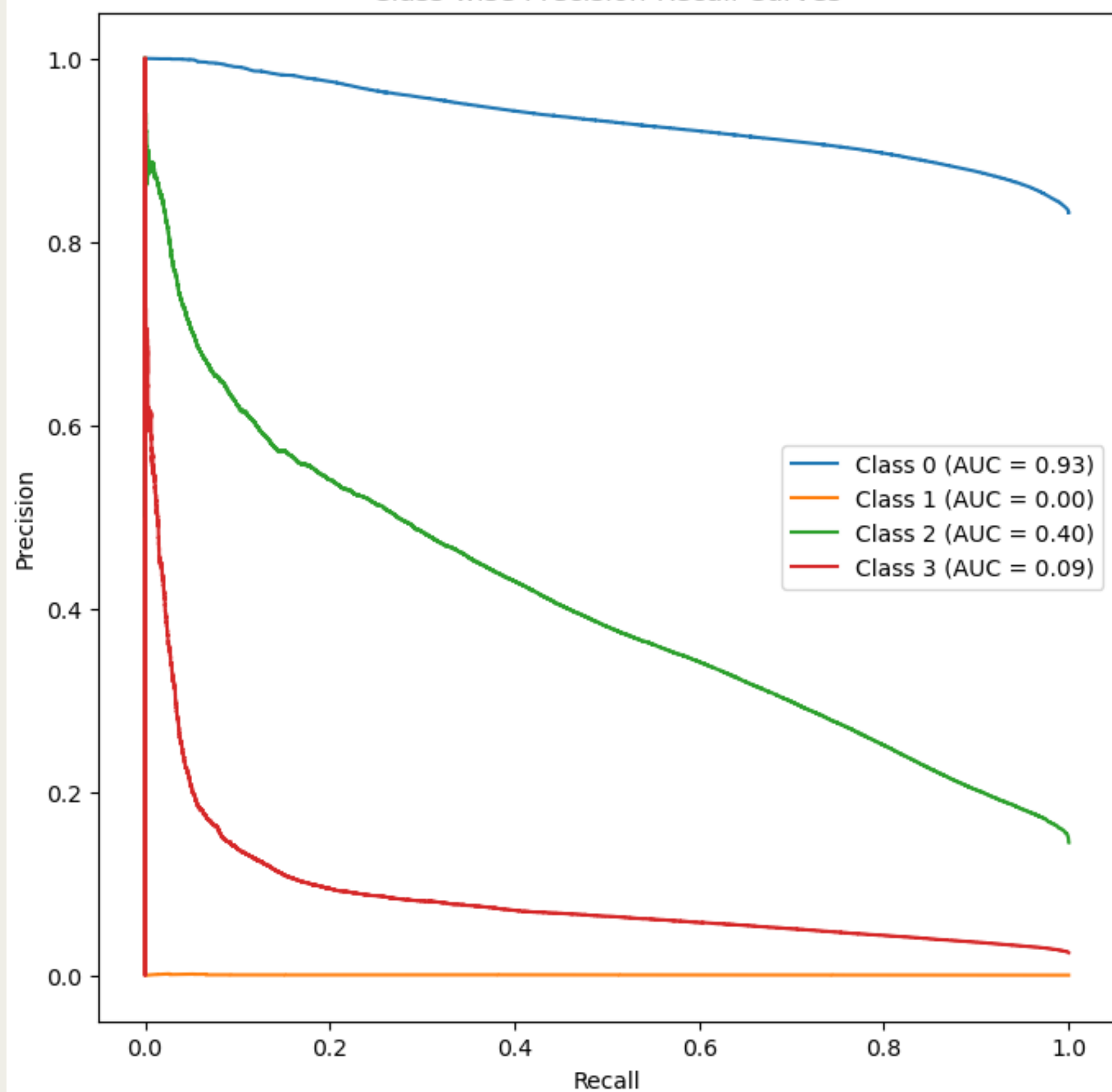


Confusion Matrix

| Actual \ Predicted | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 0 | 865855 | 9561 | 96889 | 2491 |
| 1 | 38638 | 5488 | 15491 | 1167 |
| 2 | 212535 | 8396 | 112152 | 2375 |
| 3 | 25519 | 2447 | 12194 | 1337 |

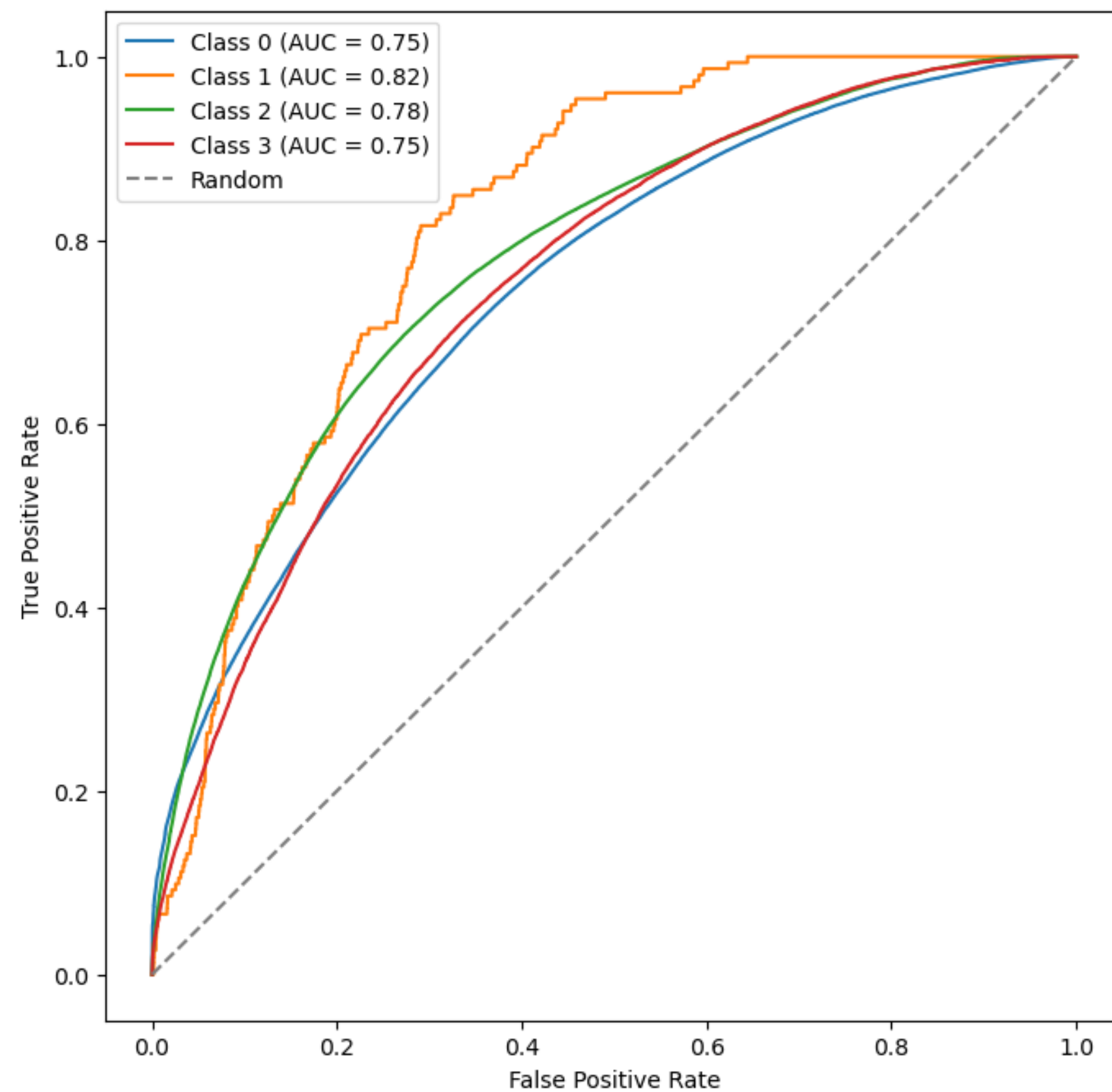Class-wise Precision-Recall Curves

Class-wise ROC Curves

# Previous Results

```
0        3404683

2         586829

3          98518

1            500
```

Decision Tree

```
Accuracy: 0.7644649144894834
Confusion Matrix:
[[876620    134 118662  25603]
 [   124      5     20      3]
 [113695     24  58809   3743]
 [ 23441      2   3588   2686]]
Classification Report:
              precision    recall  f1-score   support

           0       0.86      0.86      0.86   1021019
           1       0.03      0.03      0.03       152
           2       0.32      0.33      0.33    176271
           3       0.08      0.09      0.09     29717

    accuracy                           0.76   1227159
   macro avg       0.33      0.33      0.33   1227159
weighted avg       0.77      0.76      0.77   1227159
```

KNN

```
Accuracy: 0.8303520570683994
Confusion Matrix:
[[972169      4  47315   1531]
 [   141      2      9      0]
 [130247      0  45794    230]
 [ 27661      0   1047   1009]]
Classification Report:
              precision    recall  f1-score   support

           0       0.86      0.95      0.90   1021019
           1       0.33      0.01      0.03       152
           2       0.49      0.26      0.34    176271
           3       0.36      0.03      0.06     29717

    accuracy                           0.83   1227159
   macro avg       0.51      0.31      0.33   1227159
weighted avg       0.79      0.83      0.80   1227159
```

UpSampling (Smoting)

```
Class distribution before SMOTE:
0    2383356
2     410744
3      68918
1        353
Name: target, dtype: int64

Class distribution after SMOTE:
0    2383356
2    2383356
3    2383356
1    2383356
Name: target, dtype: int64
```

KNN

NEW-DT

```
Accuracy: 0.8303520570683994
Confusion Matrix:
[[972169       4  47315    1531]
 [    141       2      9       0]
 [130247       0  45794     230]
 [ 27661       0   1047    1009]]
Classification Report:
              precision    recall  f1-score   support

           0       0.86      0.95      0.90   1021019
           1       0.33      0.01      0.03       152
           2       0.49      0.26      0.34    176271
           3       0.36      0.03      0.06     29717

    accuracy                           0.83   1227159
   macro avg       0.51      0.31      0.33   1227159
weighted avg       0.79      0.83      0.80   1227159
```

Smoted-DT

```
Accuracy: 0.6681171714504803
Confusion Matrix:
[[734217    2431  190130   94549]
 [    70      30      35      12]
 [ 82274     530   78538   14743]
 [ 16625      76    5798    7101]]
Classification Report:
              precision    recall  f1-score   support

           0       0.88      0.72      0.79   1021327
           1       0.01      0.20      0.02       147
           2       0.29      0.45      0.35    176085
           3       0.06      0.24      0.10     29600

    accuracy                           0.67   1227159
   macro avg       0.31      0.40      0.31   1227159
weighted avg       0.78      0.67      0.71   1227159
```

DownSampling (Sampling)

```python
y.unique()
```

```
array([0, 2, 3, 1])
```

```python
y.value_counts()
```

```
0    3404683
2     586829
3      98518
1        500
Name: target, dtype: int64
```

Out[17]:

	event
StartHesitation	500
Turn	500
Walking	500
Normal	500

Smoted-DT

```
Accuracy: 0.6681171714504803
Confusion Matrix:
[[734217    2431 190130  94549]
 [    70      30     35     12]
 [ 82274     530  78538  14743]
 [ 16625      76   5798   7101]]
Classification Report:
              precision    recall  f1-score   support

           0       0.88      0.72      0.79   1021327
           1       0.01      0.20      0.02       147
           2       0.29      0.45      0.35    176085
           3       0.06      0.24      0.10     29600

    accuracy                           0.67   1227159
   macro avg       0.31      0.40      0.31   1227159
weighted avg       0.78      0.67      0.71   1227159
```

Sampled-RF

```
Accuracy: 0.5733333333333334
Confusion Matrix:
[[ 59  17  40  35]
 [  6 119  12   9]
 [ 22  21  83  20]
 [ 34  20  20  83]]
Classification Report:
              precision    recall  f1-score   support

           0       0.49      0.39      0.43       151
           1       0.67      0.82      0.74       146
           2       0.54      0.57      0.55       146
           3       0.56      0.53      0.55       157

    accuracy                           0.57       600
   macro avg       0.57      0.58      0.57       600
weighted avg       0.56      0.57      0.57       600
```

Downsample+UpSample

```
Class distribution before SMOTE:
0     2383356
2      410744
3       68918
1         353
Name: target, dtype: int64

Class distribution after SMOTE and subsampling:
2     410744
3     410744
1     410744
0     410744
Name: target, dtype: int64
```

Sampled-RF

```
Accuracy: 0.5733333333333334
Confusion Matrix:
[[ 59  17  40  35]
 [  6 119  12   9]
 [ 22  21  83  20]
 [ 34  20  20  83]]
Classification Report:
              precision    recall  f1-score   support

           0       0.49      0.39      0.43       151
           1       0.67      0.82      0.74       146
           2       0.54      0.57      0.55       146
           3       0.56      0.53      0.55       157

    accuracy                           0.57       600
   macro avg       0.57      0.58      0.57       600
weighted avg       0.56      0.57      0.57       600
```

Sampled+ Smoted-RF

```
Accuracy: 0.5922150267406261
Confusion Matrix:
[[595868    6989 254131 164339]
 [    37      50      39      21]
 [ 39099    1434 116338   19214]
 [  9152     248    5714   14486]]
Classification Report:
              precision    recall  f1-score   support

           0       0.93      0.58      0.72   1021327
           1       0.01      0.34      0.01       147
           2       0.31      0.66      0.42    176085
           3       0.07      0.49      0.13     29600

    accuracy                           0.59   1227159
   macro avg       0.33      0.52      0.32   1227159
weighted avg       0.82      0.59      0.66   1227159
```

6features

AccV	AccML	AccAP
-0.970018	0.061626	-0.265625
-0.984375	0.044497	-0.265625
-0.984375	0.029016	-0.265625
-0.984375	0.015625	-0.265625
-0.984670	0.015330	-0.265625
...	...	...
-0.961216	0.142428	-0.289655
-0.960343	0.142836	-0.290506
-0.957958	0.145494	-0.290007
-0.960616	0.145839	-0.291527
-0.967076	0.144342	-0.292384

Visit	Medication	Time	AccV	AccML	AccAP
2	1	1000	-0.970018	0.061626	-0.265625
2	1	1001	-0.984375	0.044497	-0.265625
2	1	1002	-0.984375	0.029016	-0.265625
2	1	1003	-0.984375	0.015625	-0.265625
2	1	1004	-0.984670	0.015330	-0.265625
...	...	...	...	...	...
1	1	119027	-0.961216	0.142428	-0.289655
1	1	119028	-0.960343	0.142836	-0.290506
1	1	119029	-0.957958	0.145494	-0.290007
1	1	119030	-0.960616	0.145839	-0.291527
1	1	119031	-0.967076	0.144342	-0.292384

6features+Downsample(500)

Sampled+ Smoted-RF

```
Accuracy: 0.5922150267406261
Confusion Matrix:
[[595868    6989 254131 164339]
 [    37      50     39      21]
 [ 39099    1434 116338  19214]
 [  9152     248   5714  14486]]
Classification Report:
              precision    recall  f1-score   support

           0       0.93      0.58      0.72   1021327
           1       0.01      0.34      0.01       147
           2       0.31      0.66      0.42    176085
           3       0.07      0.49      0.13     29600

    accuracy                           0.59   1227159
   macro avg       0.33      0.52      0.32   1227159
weighted avg       0.82      0.59      0.66   1227159
```

Sampled-lgbm

```
Accuracy: 0.77
Confusion Matrix:
[[ 90    2   38   21]
 [  0  146    0    0]
 [ 30    3  106    7]
 [ 16    2   19  120]]
Classification Report:
              precision    recall  f1-score   support

           0       0.66      0.60      0.63       151
           1       0.95      1.00      0.98       146
           2       0.65      0.73      0.69       146
           3       0.81      0.76      0.79       157

    accuracy                           0.77       600
   macro avg       0.77      0.77      0.77       600
weighted avg       0.77      0.77      0.77       600
```

6features+Downsample+UpSample(1,50,000)

Sampled-lgbm

```
Accuracy: 0.77
Confusion Matrix:
[[ 90    2   38   21]
 [  0  146    0    0]
 [ 30    3  106    7]
 [ 16    2   19  120]]
Classification Report:
```

	precision	recall	f1-score	support
0	0.66	0.60	0.63	151
1	0.95	1.00	0.98	146
2	0.65	0.73	0.69	146
3	0.81	0.76	0.79	157
accuracy			0.77	600
macro avg	0.77	0.77	0.77	600
weighted avg	0.77	0.77	0.77	600

Sampled+ Smoted-DT

```
Accuracy: 0.9062703623878502
Confusion Matrix:
[[39627     3  4423  1162]
 [    6   141     1     0]
 [ 4191     8 40195   430]
 [  704     0   292 28523]]
Classification Report:
```

	precision	recall	f1-score	support
0	0.89	0.88	0.88	45215
1	0.93	0.95	0.94	148
2	0.89	0.90	0.90	44824
3	0.95	0.97	0.96	29519
accuracy			0.91	119706
macro avg	0.91	0.92	0.92	119706
weighted avg	0.91	0.91	0.91	119706

```
Class distribution after SMOTE:
0      2383356
2      2383356
3      2383356
1      2383356
Name: event, dtype: int64

Shapes after SMOTE:
X_train_resampled: (9533424, 6)
y_train_resampled: (9533424,)
X_test: (1227159, 6)
y_test: (1227159,)
```

After handling train test split mistake

```
y_test.value_counts()
```

```
0      1021327
2       176085
3        29600
1          147
Name: event, dtype: int64
```

6features+UpSample(23,83,356)

Sampled+ Smoted-DT

```
Accuracy: 0.9062703623878502
Confusion Matrix:
[[39627     3  4423  1162]
 [    6   141     1     0]
 [ 4191     8 40195   430]
 [  704     0   292 28523]]
Classification Report:
              precision    recall  f1-score   support

           0       0.89      0.88      0.88     45215
           1       0.93      0.95      0.94       148
           2       0.89      0.90      0.90     44824
           3       0.95      0.97      0.96     29519

    accuracy                           0.91    119706
   macro avg       0.91      0.92      0.92    119706
weighted avg       0.91      0.91      0.91    119706
```

Smoted-DT

```
Accuracy: 0.9515001723493044
Confusion Matrix:
[[974575      9  42921   3822]
 [     4    143      0      0]
 [ 10937      4 164705    439]
 [  1021      0    360  28219]]
Classification Report:
              precision    recall  f1-score   support

           0       0.99      0.95      0.97   1021327
           1       0.92      0.97      0.94       147
           2       0.79      0.94      0.86    176085
           3       0.87      0.95      0.91     29600

    accuracy                           0.95   1227159
   macro avg       0.89      0.95      0.92   1227159
weighted avg       0.96      0.95      0.95   1227159
```

Conclusion

Sampled-RF

3-features

```
Accuracy: 0.7644649144894834
Confusion Matrix:
[[876620     134 118662  25603]
 [   124       5     20      3]
 [113695      24  58809   3743]
 [ 23441       2   3588   2686]]
Classification Report:
              precision    recall  f1-score   support

           0       0.86      0.86      0.86   1021019
           1       0.03      0.03      0.03       152
           2       0.32      0.33      0.33    176271
           3       0.08      0.09      0.09     29717

    accuracy                           0.76   1227159
   macro avg       0.33      0.33      0.33   1227159
weighted avg       0.77      0.76      0.77   1227159
```

```
Accuracy: 0.5733333333333334
Confusion Matrix:
[[ 59  17  40  35]
 [  6 119  12   9]
 [ 22  21  83  20]
 [ 34  20  20  83]]
Classification Report:
              precision    recall  f1-score   support

           0       0.49      0.39      0.43       151
           1       0.67      0.82      0.74       146
           2       0.54      0.57      0.55       146
           3       0.56      0.53      0.55       157

    accuracy                           0.57       600
   macro avg       0.57      0.58      0.57       600
weighted avg       0.56      0.57      0.57       600
```

Accuracy: 0.7644649144894834
Confusion Matrix:
[[876620   134 118662 25603]
 [124   5   20   3]
 [113695   24 58809   3743]
 [23441   2   3588   2686]]
Classification Report:

	precision	recall	f1-score	support
0	0.86	0.86	0.86	1021019
1	0.03	0.03	0.03	152
2	0.32	0.33	0.33	176271
3	0.08	0.09	0.09	29717
accuracy			0.76	1227159
macro avg	0.33	0.33	0.33	1227159
weighted avg	0.77	0.76	0.77	1227159

Smoted-DT

Accuracy: 0.9515001723493044
Confusion Matrix:
[[974575   9 42921   3822]
 [4   143   0   0]
 [10937   4 164705   439]
 [1021   0   360 28219]]
Classification Report:

	precision	recall	f1-score	support
0	0.99	0.95	0.97	1021327
1	0.92	0.97	0.94	147
2	0.79	0.94	0.86	176085
3	0.87	0.95	0.91	29600
accuracy			0.95	1227159
macro avg	0.89	0.95	0.92	1227159
weighted avg	0.96	0.95	0.95	1227159

6-features

Same is with tdcsFOG

	event
Normal	4871262
Turn	1678782
StartHesitation	304790
Walking	207838

3-features

Accuracy: 0.6972089187170584
Confusion Matrix:
[[865855   9561   96889   2491]
 [38638   5488   15491   1167]
 [212535   8396 112152   2375]
 [25519   2447   12194   1337]]
Classification Report:

	precision	recall	f1-score	support
0	0.76	0.89	0.82	974796
1	0.21	0.09	0.13	60784
2	0.47	0.33	0.39	335458
3	0.18	0.03	0.05	41497
accuracy			0.70	1412535
macro avg	0.41	0.34	0.35	1412535
weighted avg	0.65	0.70	0.66	1412535

Sampled-lgbm

Accuracy: 0.463833139874061
Confusion Matrix:
[[24425 6427 7997 2331]
 [3581 19358 8500 9968]
 [8901 9119 16755 6878]
 [5348 12356 7743 16584]]
Classification Report:

	precision	recall	f1-score	support
0	0.58	0.59	0.59	41180
1	0.41	0.47	0.44	41407
2	0.41	0.40	0.41	41653
3	0.46	0.39	0.43	42031
accuracy			0.46	166271
macro avg	0.47	0.46	0.46	166271
weighted avg	0.46	0.46	0.46	166271

6-features

Left panel:

```
Accuracy: 0.6972089187170584
Confusion Matrix:
[[865855    9561   96889    2491]
 [ 38638    5488   15491    1167]
 [212535    8396  112152    2375]
 [ 25519    2447   12194    1337]]
Classification Report:
              precision    recall  f1-score   support

           0       0.76      0.89      0.82    974796
           1       0.21      0.09      0.13     60784
           2       0.47      0.33      0.39    335458
           3       0.18      0.03      0.05     41497

    accuracy                           0.70   1412535
   macro avg       0.41      0.34      0.35   1412535
weighted avg       0.65      0.70      0.66   1412535
```

Right panel — **Sampled+ Smoted-RF**

```
Accuracy: 0.8603291666666667
Confusion Matrix:
[[48800  3900  6346  1195]
 [ 2139 56170   931   362]
 [ 9506  2978 44917  2741]
 [ 1581   641  1201 56592]]
Classification Report:
              precision    recall  f1-score   support

           0       0.79      0.81      0.80     60241
           1       0.88      0.94      0.91     59602
           2       0.84      0.75      0.79     60142
           3       0.93      0.94      0.94     60015

    accuracy                           0.86    240000
   macro avg       0.86      0.86      0.86    240000
weighted avg       0.86      0.86      0.86    240000
```

References

- https://pubmed.ncbi.nlm.nih.gov/32580330/
- https://www.sciencedirect.com/science/article/abs/pii/S0952197622002433
- https://pubmed.ncbi.nlm.nih.gov/37634658/
- https://www.kaggle.com/code/prashant111/lightgbm-classifier-in-python
- https://www.kaggle.com/code/pppppooijh/lgbm-for-both-defog-tdcsfog/edit
- @misc{tlvmc-parkinsons-freezing-gait-prediction, author = {Addison Howard, amit salomon, eran gazit, HCL-Jevster, Jeff Hausdorff, Leslie Kirsch, Maggie, Pieter Ginis, Ryan Holbrook, Yasir F Karim}, title = {Parkinson's Freezing of Gait Prediction}, publisher = {Kaggle}, year = {2023}, url = {https://kaggle.com/competitions/tlvmc-parkinsons-freezing-gait-prediction}}
- https://www.youtube.com/watch?v=j86omOwx0Hk

Thank you!