USER LOCALIZATION AND ACTIVITY RECOGNITION

EE4015

Hema Varshita(EE17BTECH11022)
Sai Manasa Pappu(EE17BTECH11036)
Vamshika K(EE17BTECH11045)

PART 1 - ACTIVITY RECOGNITION

Data Description

Dataset Generation Dataset

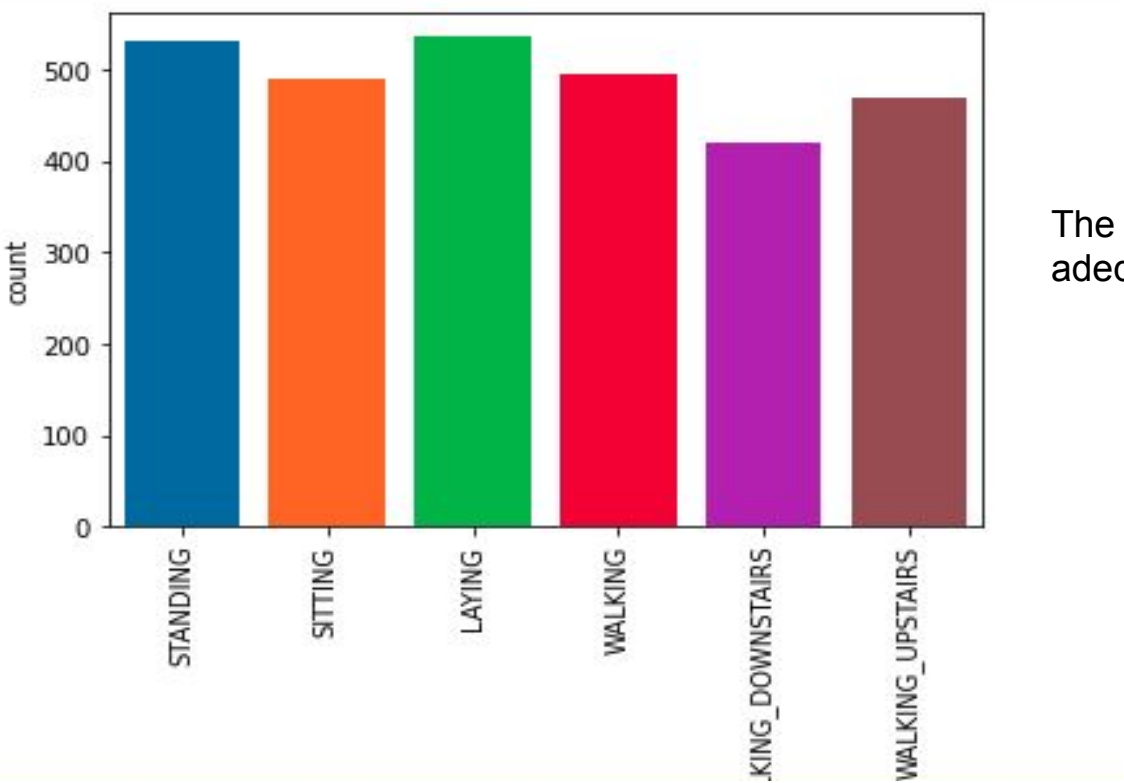
- Training data : 7352 samples, Test data: 2947 samples
- Input shape : 562 length vector mapped to 1 activity label
- Time series, multivariate data. Features come from accelerometer and gyroscope 3-axial (XYZ) raw signals.
- Time domain signals captured at 50 Hz -> filter to remove noise
- Sampled in fixed-width sliding windows of 2.56 sec and 50% overlap (128 readings/window)
- Derive acceleration values to get Jerk values.

Features extracted (for each XYZ input) in time and frequency domains

Acc
AccJerk
Gyro
GyroJerk
AccMag
AccMag
AccJerkMag
GyroMag
GyroJerkMag
mean
std

mad
max
min
energy
entropy
correlation
maxInds
meanFreq
skewness
kurtosis
bandsEnergy

Data Visualisation



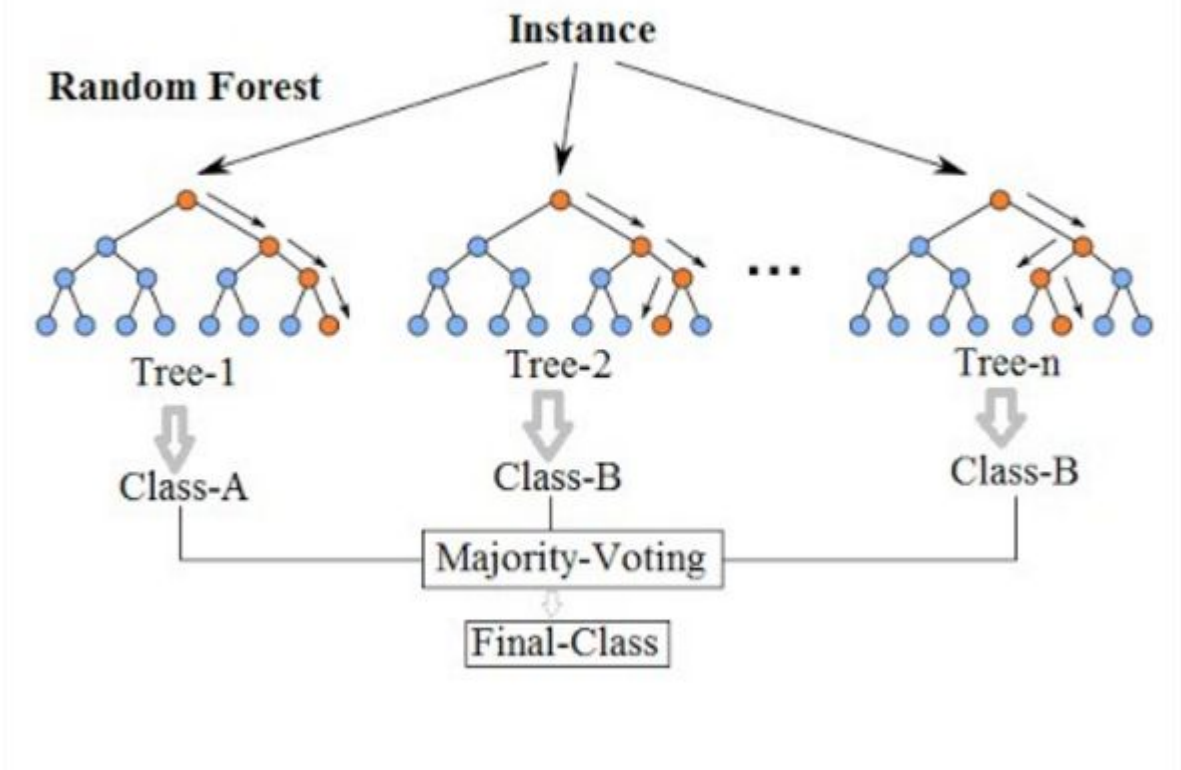
The Dataset is adequately balanced.

The problem has been solved using 3 methods using both Machine Learning and Deep Learning

1. Deep Learning - Sequential and Functional ANN
   (val_acc : 94.27%)
2. Machine Learning - K Nearest Neighbours
   (acc : 90.77%)
3. Machine Learning - Random Forest
   (acc : 92.7%)

Particularly, this dataset is used because of its reproducibility and ease of continuing with the Indoor Localization dataset.

Random Forest



Random Forest

```
no.of trees: 6  accuracy: 0.8903970139124533
no.of trees: 10  accuracy: 0.9049881235154394
no.of trees: 20  accuracy: 0.9100780454699695
no.of trees: 50  accuracy: 0.9270444519850696
no.of trees: 100  accuracy: 0.9267051238547676
no.of trees: 150  accuracy: 0.9250084832032576
```

Confusion Matrix for 50 trees

```
print(confusion_matrix(Y_test,yhat))
```

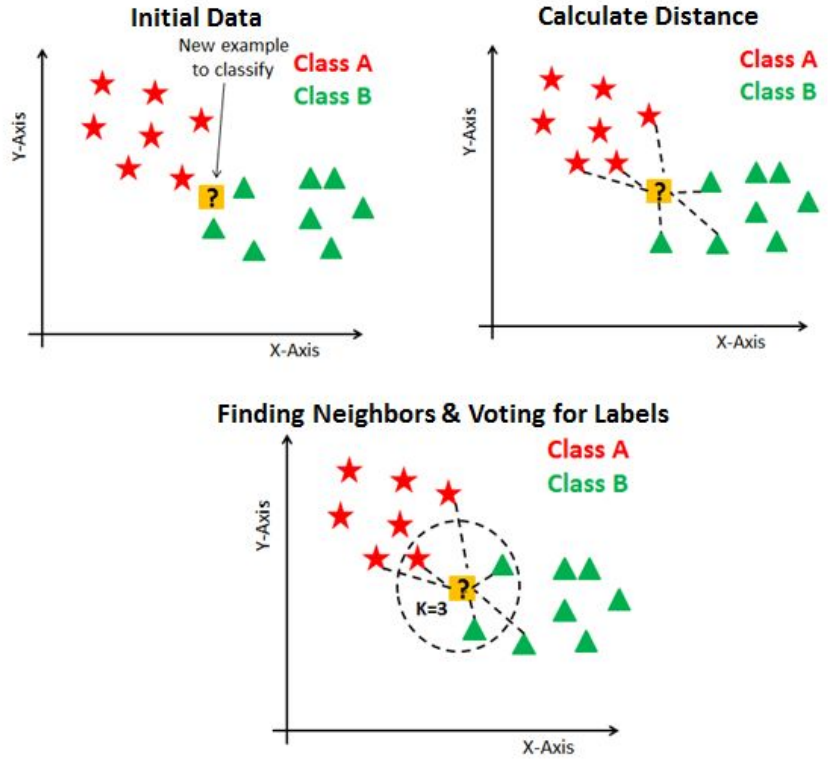
```
[[537   0   0   0   0   0]
 [  0 430  61   0   0   0]
 [  0  42 490   0   0   0]
 [  0   0   0 484   9   3]
 [  0   0   0  20 362  38]
 [  0   0   0  29   7 435]]
```

Classification Report for 50 trees

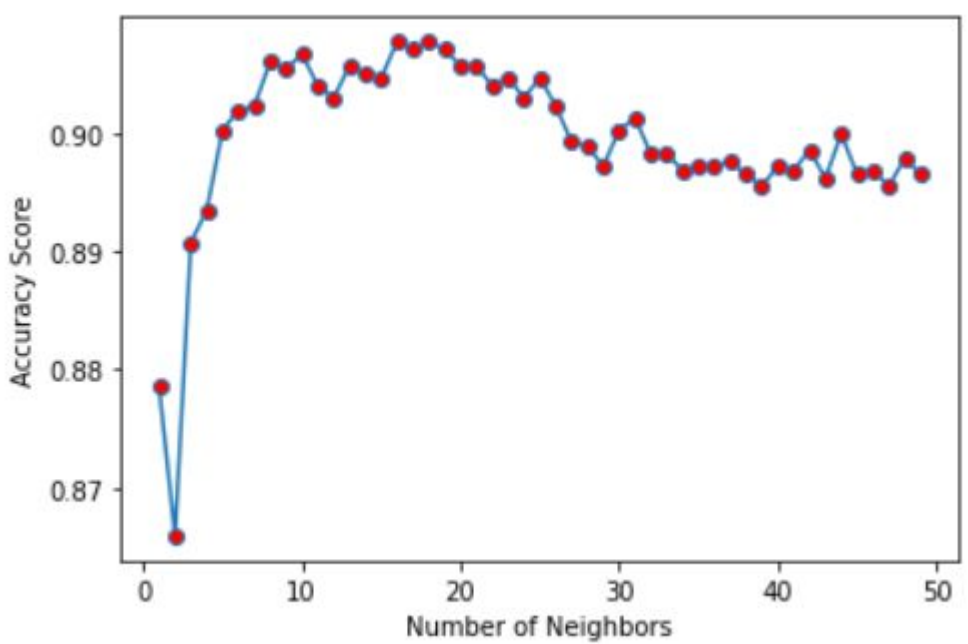
```
print(classification_report(Y_test, yhat))
```

	precision	recall	f1-score	support
0	1.00	1.00	1.00	537
1	0.91	0.88	0.89	491
2	0.89	0.92	0.90	532
3	0.91	0.98	0.94	496
4	0.96	0.86	0.91	420
5	0.91	0.92	0.92	471
accuracy			0.93	2947
macro avg	0.93	0.93	0.93	2947
weighted avg	0.93	0.93	0.93	2947

K-Nearest Neighbors algorithm



Number of neighbours(k) vs Accuracy attained



Confusion Matrix for k = 16

```
print(confusion_matrix(y_test,ypred))
```

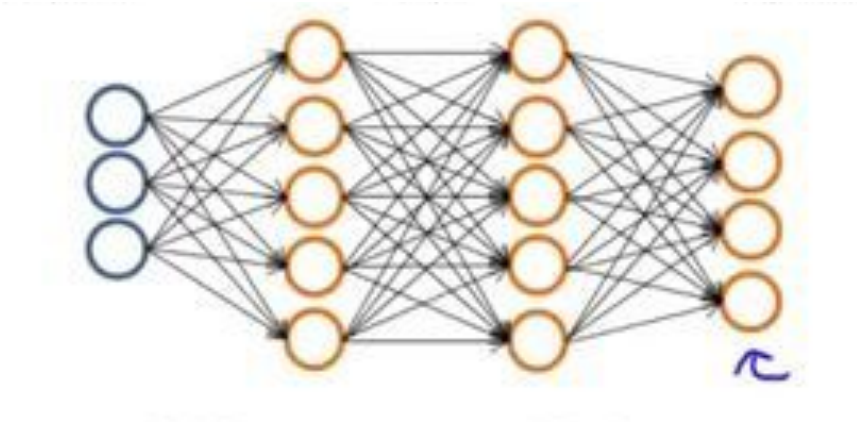
```
[[535   1   1   0   0   0]
 [  0 373 114   0   0   4]
 [  0  32 500   0   0   0]
 [  0   0   0 490   6   0]
 [  0   0   0  63 315  42]
 [  0   0   0  40   2 429]]
```

Classification Report for k = 16

```
print(classification_report(y_test,ypred))
```

```
                    precision    recall  f1-score   support

            LAYING       1.00      1.00      1.00       537
           SITTING       0.92      0.76      0.83       491
          STANDING       0.81      0.94      0.87       532
           WALKING       0.83      0.99      0.90       496
WALKING_DOWNSTAIRS       0.98      0.75      0.85       420
  WALKING_UPSTAIRS       0.90      0.91      0.91       471

          accuracy                           0.90      2947
         macro avg       0.91      0.89      0.89      2947
      weighted avg       0.90      0.90      0.90      2947
```

Using Deep Learning



Input Dimensions : (562,)
Output : 6 nodes (Number of output classes)
3 Hidden layers

SEQUENTIAL API

Layer (type)	Output Shape	Param #
dense (Dense)	(None, 128)	72064
dense_1 (Dense)	(None, 64)	8256
dense_2 (Dense)	(None, 32)	2080
dropout (Dropout)	(None, 32)	0
dense_3 (Dense)	(None, 6)	198

Total params: 82,598
Trainable params: 82,598
Non-trainable params: 0

```
[[1406    0    0    0    1    0]
 [   0 1188   98    0    0    0]
 [   0   12 1362    0    0    0]
 [   0    0    0 1223    3    0]
 [   0    0    0    0  986    0]
 [   0    0    3    2   34 1034]]
```

CONFUSION MATRIX

	precision	recall	f1-score	support
0	1.00	1.00	1.00	1407
1	0.99	0.92	0.96	1286
2	0.93	0.99	0.96	1374
3	1.00	1.00	1.00	1226
4	0.96	1.00	0.98	986
5	1.00	0.96	0.98	1073
accuracy			0.98	7352
macro avg	0.98	0.98	0.98	7352
weighted avg	0.98	0.98	0.98	7352

CLASSIFICATION REPORT

TRAINING SNAPSHOT - 20 EPOCHS

```
7352/7352 [==============================] - 1s 128us/sample - loss: 0.0808 - acc: 0.9709 - val_loss: 0.1862 -
val_acc: 0.9464
Epoch 12/20
7352/7352 [==============================] - 1s 113us/sample - loss: 0.0729 - acc: 0.9761 - val_loss: 0.1574 -
val_acc: 0.9508
Epoch 13/20
7352/7352 [==============================] - 1s 118us/sample - loss: 0.0722 - acc: 0.9743 - val_loss: 0.1587 -
val_acc: 0.9505
Epoch 14/20
7352/7352 [==============================] - 1s 132us/sample - loss: 0.0659 - acc: 0.9761 - val_loss: 0.2634 -
val_acc: 0.9301
Epoch 15/20
7352/7352 [==============================] - 1s 134us/sample - loss: 0.0680 - acc: 0.9728 - val_loss: 0.1748 -
val_acc: 0.9477
Epoch 16/20
7352/7352 [==============================] - 1s 120us/sample - loss: 0.0780 - acc: 0.9716 - val_loss: 0.3395 -
val_acc: 0.9152
Epoch 17/20
7352/7352 [==============================] - 1s 121us/sample - loss: 0.0509 - acc: 0.9820 - val_loss: 0.1475 -
val_acc: 0.9545
Epoch 18/20
7352/7352 [==============================] - 1s 125us/sample - loss: 0.0718 - acc: 0.9748 - val_loss: 0.1919 -
val_acc: 0.9464
Epoch 19/20
7352/7352 [==============================] - 1s 121us/sample - loss: 0.0495 - acc: 0.9841 - val_loss: 0.2981 -
val_acc: 0.9308
Epoch 20/20
7352/7352 [==============================] - 1s 125us/sample - loss: 0.0574 - acc: 0.9770 - val_loss: 0.2558 -
val_acc: 0.9427
```

FUNCTIONAL API

```
Model: "model"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
input_2 (InputLayer)         [(None, 562)]             0
_________________________________________________________________
dense_4 (Dense)              (None, 128)               72064
_________________________________________________________________
dense_5 (Dense)              (None, 64)                8256
_________________________________________________________________
dense_6 (Dense)              (None, 32)                2080
_________________________________________________________________
dense_7 (Dense)              (None, 6)                 198
=================================================================
Total params: 82,598
Trainable params: 82,598
Non-trainable params: 0
```

```
[[537   0   0   0   0   0]
 [  0 438  51   0   0   2]
 [  0  14 518   0   0   0]
 [  0   0   0 482   6   8]
 [  0   0   0   6 382  32]
 [  0   0   0  10   1 460]]
```

CONFUSION MATRIX

	precision	recall	f1-score	support
0	1.00	1.00	1.00	537
1	0.97	0.89	0.93	491
2	0.91	0.97	0.94	532
3	0.97	0.97	0.97	496
4	0.98	0.91	0.94	420
5	0.92	0.98	0.95	471
accuracy			0.96	2947
macro avg	0.96	0.95	0.95	2947
weighted avg	0.96	0.96	0.96	2947

CLASSIFICATION REPORT

PART 2 - INDOOR LOCALIZATION

INDOOR LOCALIZATION

Dataset

Columns 1-520 : Received Signal Strength Indicator of 520 WiFi sources : 0 to -104 (increasing order of strength) : 100 for no signal

Longitude : Longitude of position **Latitude** : Latitude of position

Floor : Indicative values 0 to 4. Each building has different number of floors.

Building ID : Indicative values 0 to 2. 3 buildings in total

Space ID : Relative Position

User ID, Phone ID, Timestamp

Total 20 users and 25 android devices have been used.
The dataset covers three buildings with 4 or more floors and an area of almost 110.00 sq mt.

	WAP001	WAP002	WAP003	WAP004	WAP005	WAP006	WAP007	WAP008	WAP009	WAP010	...	WAP520
0	100	100	100	100	100	100	100	100	100	100	...	100
1	100	100	100	100	100	100	100	100	100	100	...	100
2	100	100	100	100	100	100	100	-97	100	100	...	100
3	100	100	100	100	100	100	100	100	100	100	...	100
4	100	100	100	100	100	100	100	100	100	100	...	100
...	...	...	...	...	...	...	...	...	...	...	...	...
19932	100	100	100	100	100	100	100	100	100	100	...	100
19933	100	100	100	100	100	100	100	100	100	100	...	100
19934	100	100	100	100	100	100	100	100	100	100	...	100
19935	100	100	100	100	100	100	100	100	100	100	...	100
19936	100	100	100	100	100	100	100	100	100	100	...	100

19937 rows × 529 columns

WAP009	WAP010	...	WAP520	LONGITUDE	LATITUDE	FLOOR	BUILDINGID	SPACEID	RELATIVEPOSITION	USERID	PHONEID	TIMESTAMP
100	100	...	100	-7541.2643	4.864921e+06	2	1	106	2	2	23	1371713733
100	100	...	100	-7536.6212	4.864934e+06	2	1	106	2	2	23	1371713691
100	100	...	100	-7519.1524	4.864950e+06	2	1	103	2	2	23	1371714095
100	100	...	100	-7524.5704	4.864934e+06	2	1	102	2	2	23	1371713807
100	100	...	100	-7632.1436	4.864982e+06	0	0	122	2	11	13	1369909710
...	...	...	...	...	...	...	...	...	...	...	...	...
100	100	...	100	-7485.4686	4.864875e+06	3	1	1	2	18	10	1371710683
100	100	...	100	-7390.6206	4.864836e+06	1	2	140	2	18	10	1371710402
100	100	...	100	-7516.8415	4.864889e+06	3	1	13	2	18	10	1371710921
100	100	...	100	-7537.3219	4.864896e+06	3	1	113	2	18	10	1371711049
100	100	...	100	-7536.1658	4.864898e+06	3	1	112	2	18	10	1371711025

USING DEEP LEARNING

1. The problem is a case of Multi-Label Classification.
2. Solution : Identify Building ID, Floor ID, Longitude and Latitude as separate labels or together
3. Can be identified individually/ together as a unique position.

Data Preprocessing -

Scaling RSSI values from -**104 to 0 TO 0 to 105**

Normalising Scaled values for columns 1-520

5]:

AP008	WAP009	WAP010	...	LONGITUDE	LATITUDE	FLOOR	BUILDINGID	SPACEID	RELATIVEPOSITION	USERID	PHONEID	TIMESTAMP	NEWID
.00000	0.0	0.0	...	-7541.2643	4.864921e+06	2	1	106	2	2	23	1371713733	1151
.00000	0.0	0.0	...	-7536.6212	4.864934e+06	2	1	106	2	2	23	1371713691	1126
.07619	0.0	0.0	...	-7519.1524	4.864950e+06	2	1	103	2	2	23	1371714095	1036
.00000	0.0	0.0	...	-7524.5704	4.864934e+06	2	1	102	2	2	23	1371713807	1067
.00000	0.0	0.0	...	-7632.1436	4.864982e+06	0	0	122	2	11	13	1369909710	1445
...	...	...	...	...	...	...	...	...	...	...	...	...	...
.00000	0.0	0.0	...	-7485.4686	4.864875e+06	3	1	1	2	18	10	1371710683	915
.00000	0.0	0.0	...	-7390.6206	4.864836e+06	1	2	140	2	18	10	1371710402	550
.00000	0.0	0.0	...	-7516.8415	4.864889e+06	3	1	13	2	18	10	1371710921	1027
.00000	0.0	0.0	...	-7537.3219	4.864896e+06	3	1	113	2	18	10	1371711049	1134
.00000	0.0	0.0	...	-7536.1658	4.864898e+06	3	1	112	2	18	10	1371711025	1118

1996 unique locations were obtained

RESULTS

Hyperparameters :

Hidden layers - 3

Epochs - 150

Dropout - 0.2

Batch size - 6

Output Neurons - 1996 (No of NEWIDs)

Results :

Training accuracy - 95.8%

Validation accuracy - 79.5%

TRAINING SET - 80% of train data set

VALIDATION SET - 20% of train data set

TEST SET - Test data set

TRAINING SNAPSHOT - 150 EPOCHS


```
Epoch 136/150
15949/15949 [==============================] - 3s 214us/step - loss: 0.1852 - categorical_accuracy: 0.9530
Epoch 137/150
15949/15949 [==============================] - 4s 221us/step - loss: 0.1862 - categorical_accuracy: 0.9503
Epoch 138/150
15949/15949 [==============================] - 4s 243us/step - loss: 0.1841 - categorical_accuracy: 0.9513
Epoch 139/150
15949/15949 [==============================] - 4s 221us/step - loss: 0.1823 - categorical_accuracy: 0.9528
Epoch 140/150
15949/15949 [==============================] - 4s 222us/step - loss: 0.1819 - categorical_accuracy: 0.9525
Epoch 141/150
15949/15949 [==============================] - 4s 243us/step - loss: 0.1796 - categorical_accuracy: 0.9522
Epoch 142/150
15949/15949 [==============================] - 4s 253us/step - loss: 0.1813 - categorical_accuracy: 0.9523
Epoch 143/150
15949/15949 [==============================] - 4s 253us/step - loss: 0.1784 - categorical_accuracy: 0.9525
Epoch 144/150
15949/15949 [==============================] - 4s 244us/step - loss: 0.1776 - categorical_accuracy: 0.9539
Epoch 145/150
15949/15949 [==============================] - 4s 223us/step - loss: 0.1773 - categorical_accuracy: 0.9529
Epoch 146/150
15949/15949 [==============================] - 3s 217us/step - loss: 0.1773 - categorical_accuracy: 0.9527
Epoch 147/150
15949/15949 [==============================] - 3s 217us/step - loss: 0.1743 - categorical_accuracy: 0.9544
Epoch 148/150
15949/15949 [==============================] - 3s 214us/step - loss: 0.1751 - categorical_accuracy: 0.9532
Epoch 149/150
15949/15949 [==============================] - 3s 218us/step - loss: 0.1738 - categorical_accuracy: 0.9539
Epoch 150/150
15949/15949 [==============================] - 4s 223us/step - loss: 0.1722 - categorical_accuracy: 0.9535
15949/15949 [==============================] - 3s 203us/step
3988/3988 [==============================] - 2s 592us/step
```

Random Forest Regressor

```
Preprocessing: Values to start from 0 - subtract minimum value from all

Before:  min:-7632.143599998206 ; max: -7300.818990092725

After :  min: 0.0               ; max:  331.32460990548134
```

Results

```
==================== LONGITUDE ===============

MAE:  1.641792703790993

==================== LATITUDE ===============

MAE:  1.240997795678485
```

Random Forest Classifier

```
==================== FLOOR ===============
              precision    recall  f1-score   support

       -99.0       1.00      1.00      1.00      5002
       -98.0       1.00      1.00      1.00      4416
       -97.0       0.99      1.00      1.00      5048
       -96.0       1.00      1.00      1.00      1102
       105.0       1.00      0.99      1.00      4369

    accuracy                           1.00     19937
   macro avg       1.00      1.00      1.00     19937
weighted avg       1.00      1.00      1.00     19937


[[5001    0    1    0    0]
 [   0 4415    1    0    0]
 [   0    0 5048    0    0]
 [   0    0    0 1102    0]
 [   0    0   37    0 4332]]
Accuracy: 0.9980438380899834

==================== BUILDINGID ===============
              precision    recall  f1-score   support

       -99.0       1.00      0.99      1.00      5196
       -98.0       1.00      1.00      1.00      9492
       105.0       1.00      1.00      1.00      5249

    accuracy                           1.00     19937
   macro avg       1.00      1.00      1.00     19937
weighted avg       1.00      1.00      1.00     19937


[[5159   37    0]
 [   0 9492    0]
 [   0    1 5248]]
Accuracy: 0.9980939960876761
```

Random Forest Classifier

```
==================== SPACEID ===============
     precision    recall  f1-score   support

       154.0       1.00      1.00      1.00        10

    accuracy                           0.97     19937
   macro avg       0.97      0.96      0.96     19937
weighted avg       0.98      0.97      0.98     19937


[[65  0  0 ...  0  0  0]
 [ 0 80  0 ...  0  0  0]
 [ 1  0 82 ...  0  0  0]
 ...
 [ 0  0  0 ... 20  0  0]
 [ 0  0  0 ...  0 20  0]
 [ 0  0  0 ...  0  0 10]]
Accuracy: 0.9695540954005116
```

```
==================== RELATIVEPOSITION ===============
              precision    recall  f1-score   support

       -99.0       1.00      0.99      0.99      3329
       -98.0       1.00      1.00      1.00     16608

    accuracy                           1.00     19937
   macro avg       1.00      0.99      1.00     19937
weighted avg       1.00      1.00      1.00     19937


[[ 3283    46]
 [    2 16606]]
Accuracy: 0.9975924161107489
```

K Nearest Neighbors Regression

```
==================== LONGITUDE ===============
Score  0.9963894264614553
==================== LATITUDE ===============
Score  0.9939766852066294
```

K-Nearest Neighbors

```
==================== FLOOR ===============
                precision    recall  f1-score   support

         -99.0       1.00      1.00      1.00      5002
         -98.0       1.00      1.00      1.00      4416
         -97.0       1.00      0.99      1.00      5048
         -96.0       1.00      1.00      1.00      1102
         105.0       0.99      1.00      1.00      4369

      accuracy                           1.00     19937
     macro avg       1.00      1.00      1.00     19937
  weighted avg       1.00      1.00      1.00     19937


[[5001    0    0    0    1]
 [   0 4415    0    0    1]
 [   0    0 5011    0   37]
 [   0    0    0 1102    0]
 [   0    0    0    0 4369]]
Accuracy: 0.9980438380899834
```

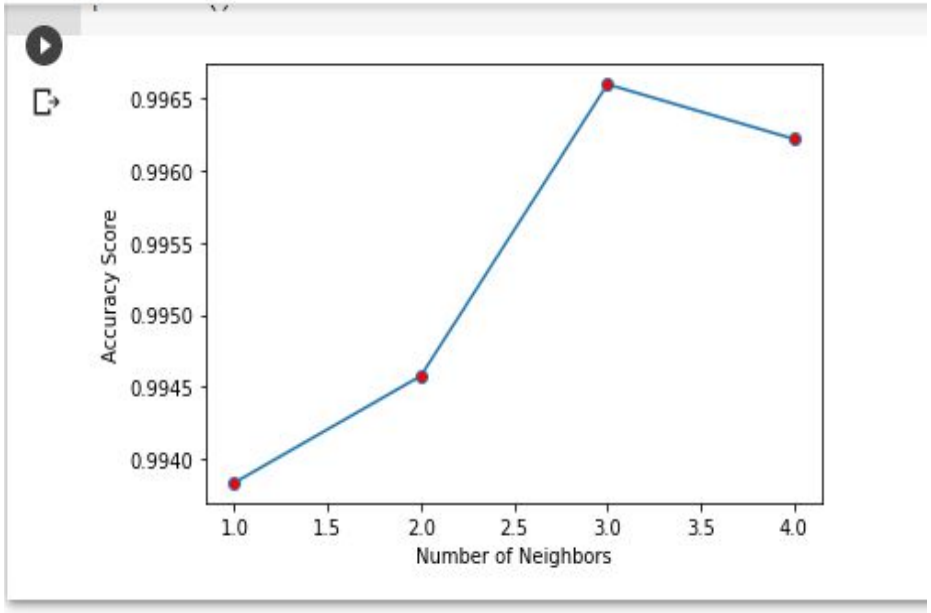
```
==================== BUILDINGID ===============
                precision    recall  f1-score   support

         -99.0       1.00      0.99      1.00      5196
         -98.0       1.00      1.00      1.00      9492
         105.0       1.00      1.00      1.00      5249

      accuracy                           1.00     19937
     macro avg       1.00      1.00      1.00     19937
  weighted avg       1.00      1.00      1.00     19937


[[5159   37    0]
 [   0 9492    0]
 [   0    1 5248]]
Accuracy: 0.9980939960876761
```

```
==================== SPACEID ==============
              precision    recall  f1-score   support

       -99.0       0.98      1.00      0.99        65
       -98.0       1.00      1.00      1.00        80
       153.0       1.00      1.00      1.00        20
       154.0       1.00      1.00      1.00        10

    accuracy                           0.97     19937
   macro avg       0.97      0.95      0.95     19937
weighted avg       0.98      0.97      0.97     19937


[[65  0  0 ...  0  0  0]
 [ 0 80  0 ...  0  0  0]
 [ 1  0 80 ...  0  0  0]
 ...
 [ 0  0  0 ... 20  0  0]
 [ 0  0  0 ...  0 20  0]
 [ 0  0  0 ...  0  0 10]]
Accuracy: 0.9681998294628078
```

```
/usr/local/lib/python3.6/dist-packages/sklearn/metrics/_classification.py:1272
  _warn_prf(average, modifier, msg_start, len(result))
==================== RELATIVEPOSITION ==============
              precision    recall  f1-score   support

       -99.0       0.99      0.99      0.99      3329
       -98.0       1.00      1.00      1.00     16608

    accuracy                           1.00     19937
   macro avg       0.99      0.99      0.99     19937
weighted avg       1.00      1.00      1.00     19937


[[ 3294    35]
 [   41 16567]]
Accuracy: 0.9961879921753524
```

Number of neighbours(k) vs Accuracy attained

For Regression:



COMPARISON

1. ACTIVITY RECOGNITION			
Existing	**Deep Learning**	**KNN**	**RF**
89.3	94.27	90.77	92.7

A Hardware Friendly Multi Class SVM has been implemented already. It gave a highest accuracy of 89.3%

2. USER LOCALIZATION			
Existing	**Deep Learning**	**KNN**	**RF**
NA	79.5	99.8,99.8,96.8,99.6	99.8,99.8,96.9,99.75

No pre-existing state of the art models have been published. KNN and RF algorithms individually predict each of Building ID, Floor ID, Space ID and Relative Position while Deep Learning merges 4 columns into a single label

THANK YOU