Human activity recognition using smartwatch gyroscope: A data mining approach

Elyas Khorasani (77264206)


Data Analytics
and Visualisation
Dr Ah-Lian Kor

01 **Introduction**

Aim & Objectives

02 **Findings**

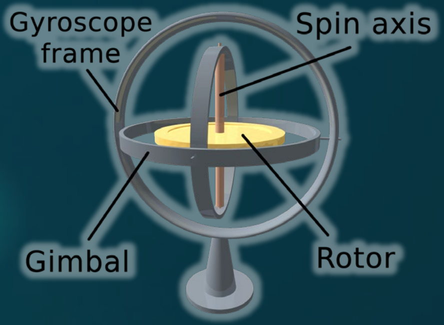
Level 1- 4

03 **Methodology**

Macro & Micro

04 **Conclusion**

Future Works



Introduction

WISDM Dataset

Jogging, Typing and Folding Clothes

AIM

The aim is to recognise human activities using the smartwatch's gyroscope biometric data of daily routines, including jogging, typing and folding clothes, by taking advantage of data mining processes.

Objectives

RO1- Level 1: Descriptive statistics analysis

RQ1: Are the biometric sensor data different in each task?

RQ2: How different are the biometric sensor data in various people?

RO2- Level 2: Inferential statistics analysis

RQ3: Is there any significant pattern for each activity compared to other activities?

RO3- Level 3: Machine Learning techniques

RQ4: Can the subject activity be predicted based on the categorised biometric data captured by the gyroscope sensor?

RO4: Level 4: Deep Learning

RQ6: Can the subject activity be predicted by the raw biometric data captured by the sensors?

Macro Methodology

Cross-industry standard process for data mining

CRISP-DM

Sample, Explore, Modify, Model, Assess

SEMMA

Knowledge Discovery in Databases

KDD

Micro Methodology

Descriptive — Imputing

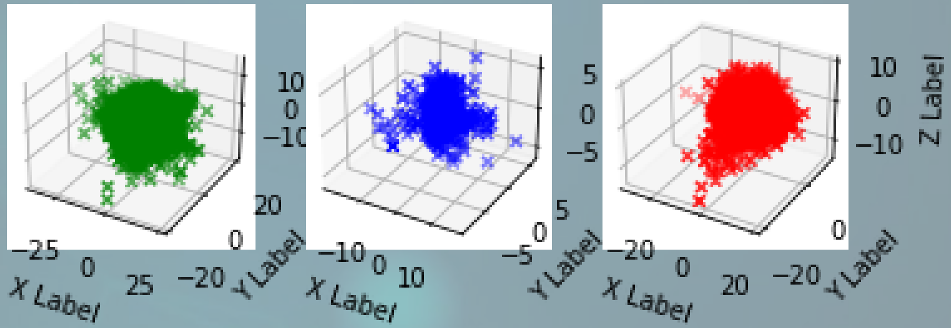
Inferential — Hypotheses

Machine Learning — Decision Tree, Random Forest

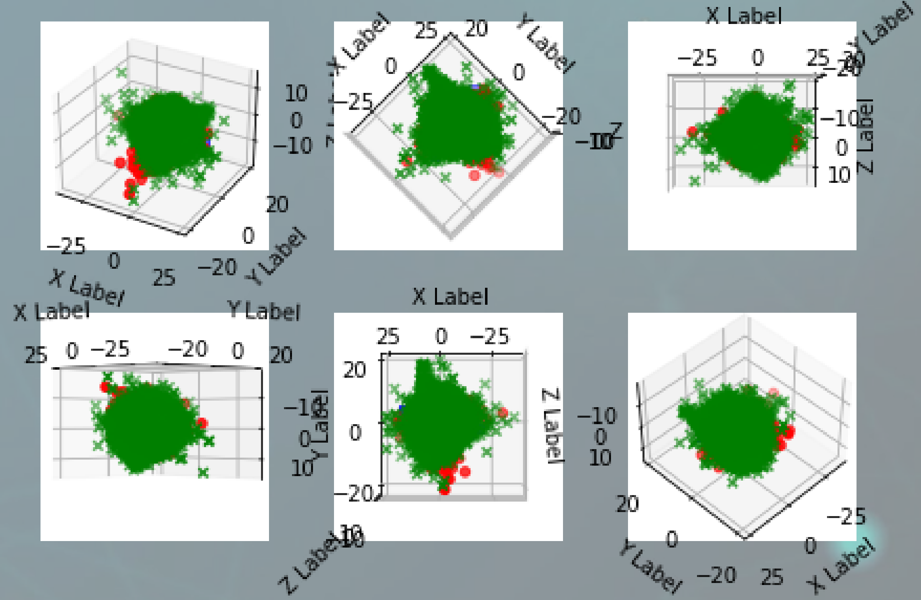
Deep Learning — SGD, RMSprop and Adam
64, 128 and 256 neurons

Type of Activity	Descriptive Analysis								
		count	mean	std	min	25%	50%	75%	max
Jogging	0	187834.0	-0.251163	2.384040	-34.858818	-1.495725	-0.191051	1.033047	24.217474
	1	187834.0	-0.017893	2.758686	-18.896862	-1.853837	0.079589	1.886868	19.104145
	2	187834.0	0.025555	2.632422	-17.444334	-1.542882	0.015315	1.602832	14.072606
		count	mean	std	min	25%	50%	75%	max
Typing	0	187175.0	0.000827	0.376592	-18.084496	-0.096836	0.000000	0.095923	15.814603
	1	187175.0	-0.000038	0.210979	-8.617611	-0.051872	0.000899	0.052285	5.622283
	2	187175.0	0.000090	0.212114	-5.717322	-0.042782	0.000583	0.047937	5.517232
		count	mean	std	min	25%	50%	75%	max
Folding Clothes	0	193373.0	-0.017581	2.032569	-29.129051	-0.863939	-0.004474	0.815522	19.209053
	1	193373.0	-0.010080	1.161938	-20.301973	-0.534951	0.003196	0.522558	11.141437
	2	193373.0	0.001540	1.163480	-13.460214	-0.484236	0.013848	0.559211	9.442167

0: x-axis, 1: y-axis, 2: z-axis



3D scatter plot of the features (Green: Jogging, Blue: Typing and Red: Folding clothes)



3D scatter plot of the features in one graph from six different angles (x: Jogging, star: Typing and dot: Folding clothes)

Oneway ANOVA

Null Hypotheses are as follows for each axis (H0): mean of Jogging activity data = mean of Typing activity data = mean of Folding clothes activity data

Alternative Hypotheses for each axis are as follows (Ha): At least one means for the activities is not equal to the other

F Value (x) = 1116.34

P Value (x) = 0

The null hypothesis can be rejected

F Value (y) = 5.02549

P Value (y) = 0.00656867

The null hypothesis can be rejected

F Value (z) = 13.9398

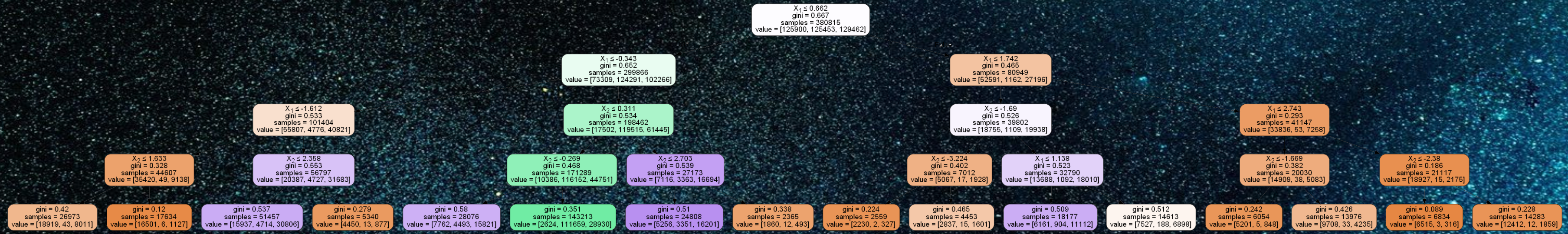
P Value (z) = 8.83457e-07

The null hypothesis can be rejected

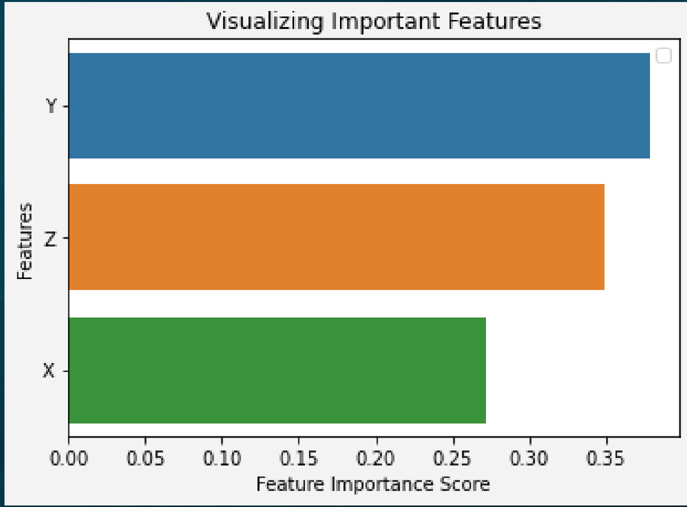
Max depth	Decision tree Accuracy				
	Attempt #1	Attempt #2	Attempt #3	Attempt #4	Attempt #5
1	46.479%	46.479%	46.479%	46.479%	46.479%
2	59.891%	59.891%	59.891%	59.891%	59.891%
3	67.296%	.296%67	67.296%	67.296%	67.296%
4	71.729%	.729%71	71.729%	71.729%	71.729%
5	72.970%	72.970%	72.970%	72.970%	72.970%
6	73.801%	.801%73	73.801%	73.801%	73.801%
7	74.438%	74.438%	74.438%	74.438%	74.438%
8	74.632%	74.634%	74.632%	74.634%	74.633%
9	74.754%	74.754%	74.755%	74.755%	74.756%
10	74.858%	74.858%	74.862%	74.856%	74.860%

Decision Tree

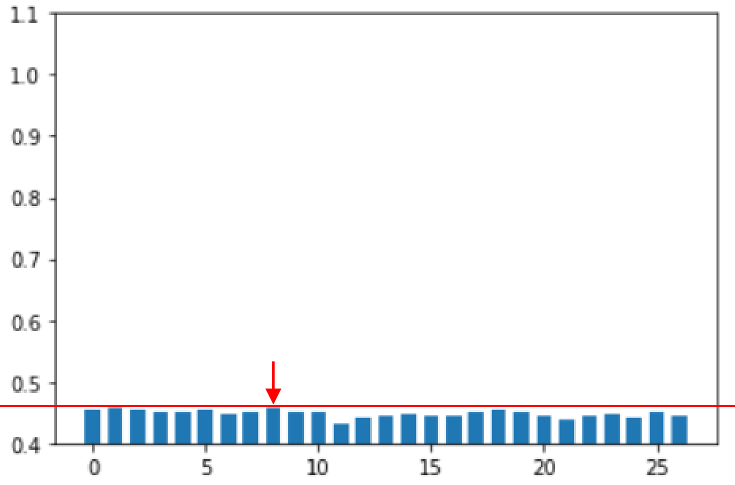
(Max depth = 4)



Random Forest Accuracy				
Attempt #1	**Attempt #2**	**Attempt #3**	**Attempt #4**	**Attempt #5**
74.489%	74.474%	74.520%	74.472%	74.509%
Attempt #6	Attempt #7	Attempt #8	Attempt #9	Attempt #10
74.545%	74.470%	74.491%	74.539%	74.475#



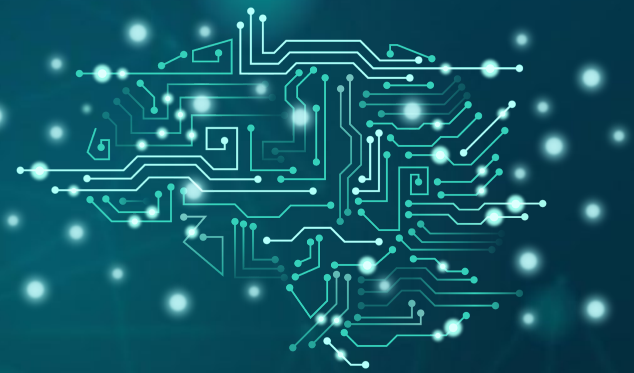
Deep Learning



Execution Time 21746.080644845963 seconds:

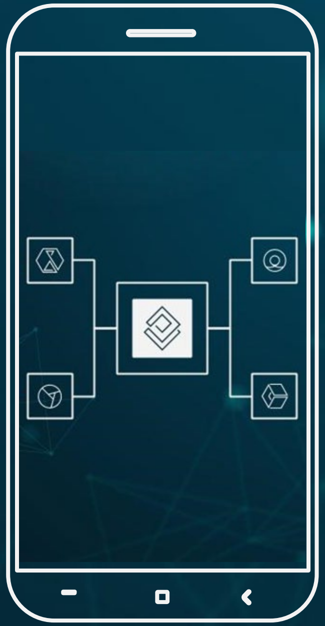
```
Model: "sequential_30"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
dense_90 (Dense)             (None, 14)                56
_________________________________________________________________
activation_87 (Activation)   (None, 14)                0
_________________________________________________________________
dense_91 (Dense)             (None, 128)               1920
_________________________________________________________________
activation_88 (Activation)   (None, 128)               0
_________________________________________________________________
dense_92 (Dense)             (None, 3)                 387
_________________________________________________________________
activation_89 (Activation)   (None, 3)                 0
=================================================================
Total params: 2,363
Trainable params: 2,363
Non-trainable params: 0
```

Conclusion



Future Work

- Accelerometer data
- Smartphone data
- Testing with real-world data



THANKS!