Analysing EM waves from electronic devices

—

Ulrich Mbou Sob – DSI twist 2 challenge

Goal of the challenge

- Retrieve and clean a huge dataset
- Set up a database
- Apply a regression model that they have learned in module 2
- Build an interactive dashboard to show your data and your model

My personal objectives

- Understand databases and get familiar with SQL

- Building dashboards

Dataset & Database

Geo-Magnetic field and WLAN dataset for indoor localisation from wristband and smartphone Data Set.

- Locations files
- Timestamp files
- Sensor data for smartphone and smartwatch
- Wifi data from a smartphone

UCI Machine Learning Repository: Geo-Magnetic field and WLAN dataset for indoor localisation from wristband and smartphone Data Set

A view of the data

	timestamp	AccelerationX	AccelerationY	AccelerationZ	MagneticFieldX	MagneticFieldY	MagneticFieldZ
count	4.200000e+01	42.000000	42.000000	42.000000	42.000000	42.000000	42.000000
mean	1.423128e+12	-0.447952	5.894310	7.932048	-14.838095	-1.009524	-33.709524
std	1.227355e+03	0.167519	0.135728	0.295671	0.342838	0.315308	0.257388
min	1.423128e+12	-0.944000	5.388000	6.885000	-15.400000	-1.200000	-34.100000
25%	1.423128e+12						
50%	1.423128e+12						
75%	1.423128e+12						
max	1.423128e+12						

MagneticFieldY	MagneticFieldZ	Z-AxisAgle(Azimuth)	X-AxisAngle(Pitch)	Y-AxisAngle(Roll)	GyroX	GyroY	GyroZ
42.000000	42.000000	42.00000	42.000000	42.000000	42.0	42.0	42.0
-1.009524	-33.709524	40.55650	-36.345857	-2.425524	0.0	0.0	0.0
0.315308	0.257388	1.74828	1.149993	0.915897	0.0	0.0	0.0
-1.200000	-34.100000	37.30400	-39.150000	-4.660000	0.0	0.0	0.0
-1.200000	-33.700000	39.51100	-36.892000	-2.990000	0.0	0.0	0.0
-1.200000	-33.700000	40.24400	-36.398000	-2.310500	0.0	0.0	0.0
-0.800000	-33.700000	41.55275	-35.635500	-1.861750	0.0	0.0	0.0
-0.200000	-33.300000	45.35000	-32.730000	-0.023000	0.0	0.0	0.0

Machine learning model

Using xgboost to perform a regression and predict the Z-axis angle

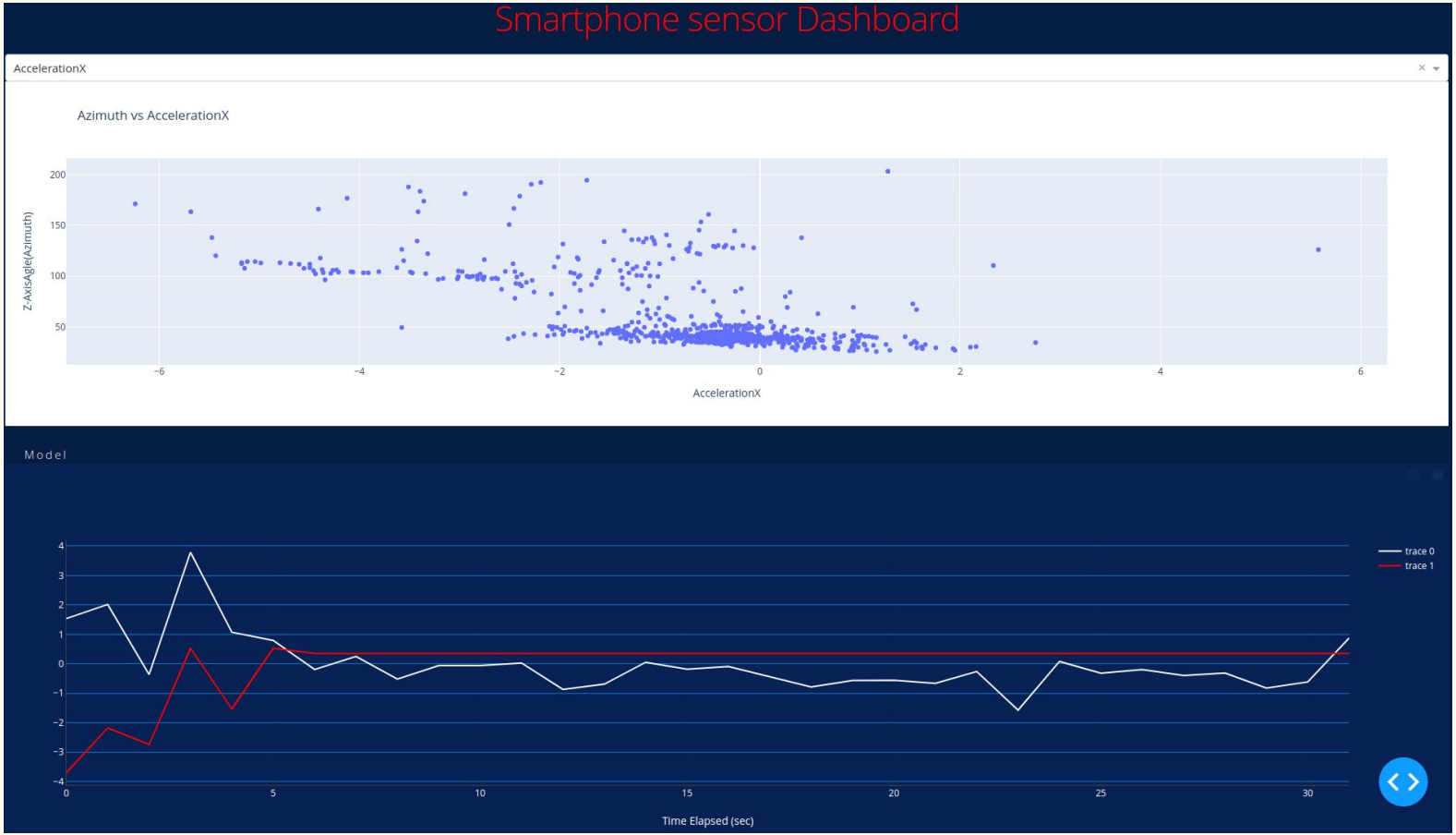
Did not spend much time fine tuning the model

Dashboard

Build a simple dashboard with two components

1) A drop down option that renders a scatter plot of the target variable vs the chosen option.
2) A continuous live plot of the target and predicted variable. This is achieved by continuously querying the timestamps for the database.

Demo


Smartphone sensor Dashboard
AccelerationX
Azimuth vs AccelerationX
Z-AxisAgle(Azimuth)
200
150
100
50
−6
−4
−2
0
2
4
6
AccelerationX
Model
4
3
2
1
0
−1
−2
−3
−4
0
5
10
15
20
25
30
Time Elapsed (sec)
trace 0
trace 1

Conclusions

This was a very interesting challenge.

I achieved my personal goals of getting familiar with databases and dashboards.

Struggled with the finding the datasets (especially how to use the amazon and google bigquery platforms)

Didn’t dive much into the machine learning model (but this was a little bit of the choice on my side give the time constraints).