September 29, 2020

Activity Recognition in Healthy Older People
Project Phase I

Jordan Lanius
CS 4300
University of Missouri-St. Louis

Abstract

As she has grown older, my grandmother has experienced dizziness and infrequent fainting spells. She has a button that she wears around her neck that she can use to call for help in case of a fall, but it is risky to rely on her ability to press it during an emergency. I am interested in the potential for an artificial intelligence to monitor a person's activity and be able to report any concerning behavior.

1 Dataset

The dataset was obtained by the UCI Machine Learning Repository and is titled "Activity recognition with healthy older people using a batteryless wearable sensor Data Set"[1]. It was provided by Roberto Luis Shinmoto Torres, Damith Ranasinghe, and Renuka Visvanathan from the University of Adelaide. For this project, I am only using the S1 setting that uses four RFID reader antennas. There are sixty participants that took part in this trial, with varying amounts of data collected from each.

1.1 Dataset Description

The data was originally separated into different files for each participant. During the cleaning process, where I merged it from multiple files to a single *.csv file, I included columns for the participant number and their gender. I also adjusted the activity column, condensing it from four different activities into two categories: 'In Bed', expressed as 0, and 'Out of Bed', expressed as 1.

In total, the condensed dataset has 52482 rows and 11 columns. For this project, I will be using 8 of the 11 columns for the input and use the 'Activity' column for the output. The input attributes that will be used are:

- Gender of the Participant
- Acceleration reading in G for frontal axis
- Acceleration reading in G for vertical axis
- Acceleration reading in G for lateral axis
- ID of antenna reading sensor
- Received signal strength indicator (RSSI)
- Phase
- Frequency

1.2 Input Data Visualization

Histograms plot the distribution of each input feature. Figure 1 shows these plots for each input column of the dataset.

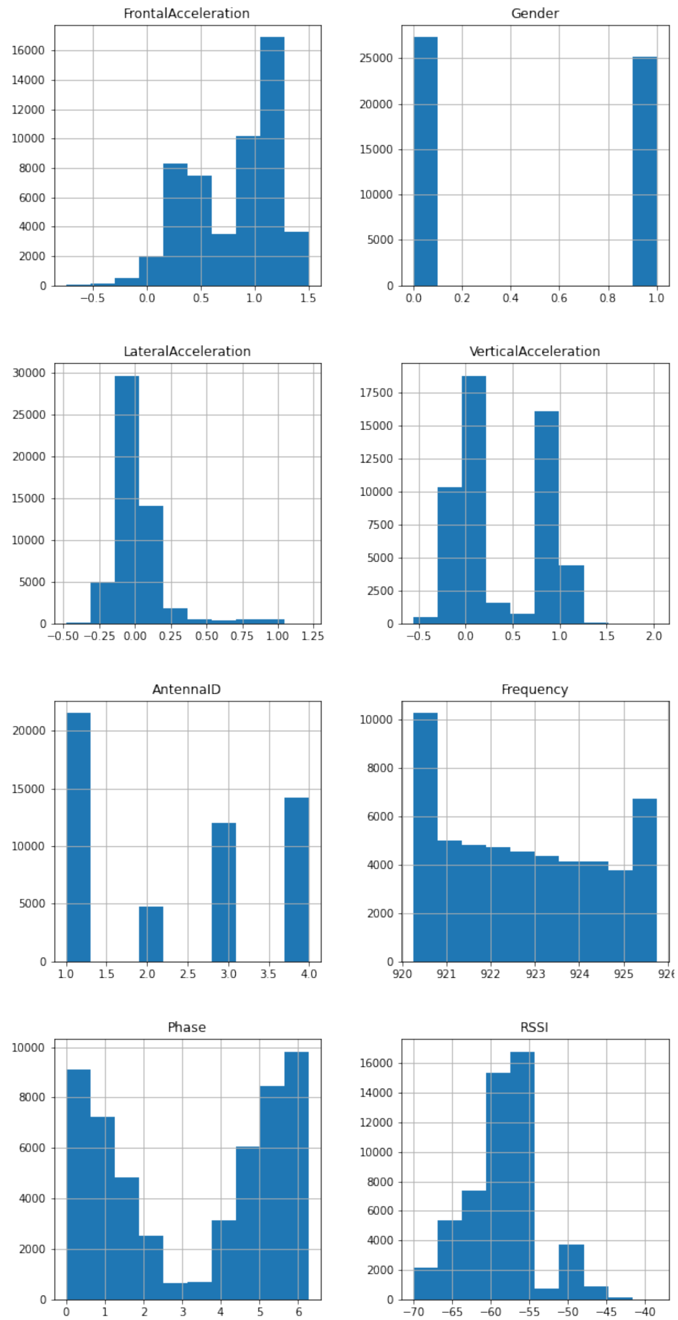

Figure 1: Input Data Distribution Histograms

Table 1 contains the minimum and maximum values of each attribute as well as the calculated mean and standard deviation values that are used for normalization.

Table 1: Input Feature Statistics

	Min	Max	Mean	Std
Gender	0.0	1.0	0.479993	0.499604
Frontal Acceleration	-0.74808	1.5032	0.805042	0.39636
Vertical Acceleration	-0.55349	2.0302	0.377804	0.468899
Lateral Acceleration	-0.48121	1.2178	0.00771	0.180674
Antenna ID	1.0	4.0	2.360752	1.261542
RSSI	-70.0	-38.5	-58.430814	4.61122
Phase	0.0	6.2817	3.275907	2.240341
Frequency	920.25	925.75	922.762261	1.693769

1.3 Output Data Visualization

Figure 2 shows the distribution of the binary output. There is some imbalance between the two. About 87.9% (46145/52482) of the sensor observations are classified as 'In Bed' while only 12.1% (6337/52482) are classified as 'Out of Bed'. While significant, this imbalance isn't extreme and there are still plenty of observations of both classes.

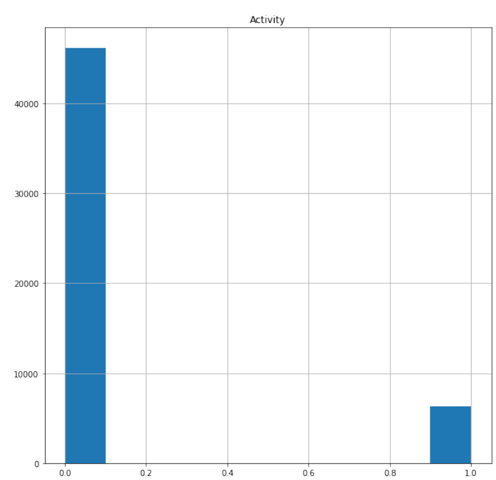

Figure 2: Output Data Distribution Histogram

2 Data Processing

2.1 Data Normalization

Neural networks work best when data has been distributed uniformly. The best way to achieve this is to apply a normalization technique to the dataset. For this project I have used Standardization which utilizes Equation 1.

$$X = (X - mean)/std \quad (1)$$

3 Conclusion

This phase of the project focused finding a good dataset to study. The dataset provided by Shinmoto Torres, Ranasinghe, and Visvanathan has a large amount of instances and plenty of unique attributes. The data it contains will provide a strong foothold to develop a neural network that can classify the activities performed by a person wearing a sensor.

References

[1] Shinmoto Torres, R. L., Ranasinghe, D. C., Shi, Q., Sample, A. P. (2013, April). Sensor enabled wearable RFID technology for mitigating the risk of falls near beds. In 2013 IEEE International Conference on RFID (pp. 191-198). IEEE.