Data Scavenger Hunt Results

In the data scavenger hunt assignment, there were nine features defined on the data: Minimum Linear Acceleration (MLA), Peak Linear Acceleration (PLA), Average Linear Acceleration (ALA), Minimum Rotational Velocity (MRV), Peak Rotational Velocity (PRV), Average Rotational Velocity (ARV), Minimum Rotational Acceleration (MRA), Peak Rotational Acceleration (PRA), and Average Rotational Acceleration (ARA). After visualizing all of the features, only three features, PLA, PRA, and ARV appeared to be helpful in determining five data instances that are "different," with PLA and PRA agreeing on the same five data instances.

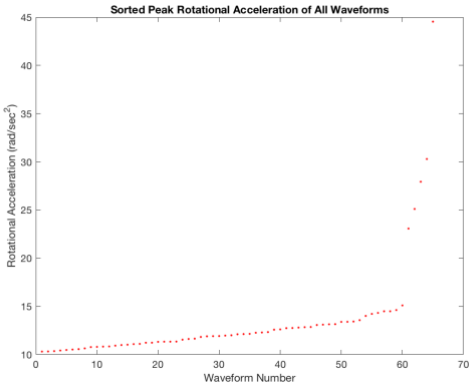

Figure 1. Sorted PRA shows five "different" data instances clearly.

I chose PLA, PRA, and ARV out of the nine features, because these features have the largest ranges, and with these features plotted it is visually clear which five data instances stand out. For example, looking at the sorted PRA feature plot in Figure 1, the last five data points are separated from the rest by a relatively large gap of at least 7.973 rad/sec^2, forming their own cluster.

While PLA, PRA, and ARV are the three features that best differentiate five data instances, only PLA and PRA agree on the same instances. With ARV, the five data instances that stand out have indexes of 8, 9, 14, 25, and 60, whereas with PLA and PRA, the indexes are 8, 9, 11, 19, and 20. This agreement makes PLA and PRA the two most helpful features in determining the five different data instances.

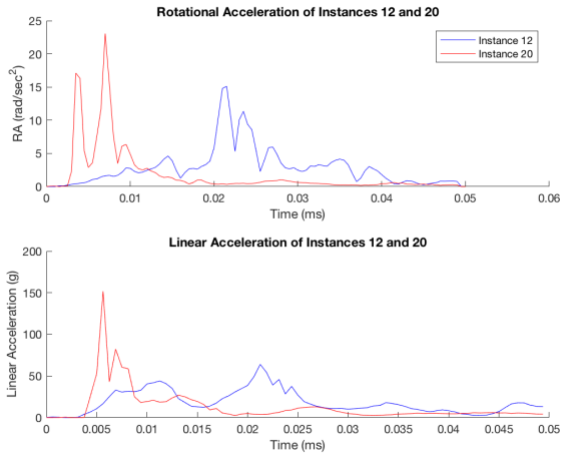

Figure 2. Comparison of PLA and PRA with a "different" and a "normal" data instance.

To compare how PLA and PRA make the instances distinct from one another, Figure 2 has plotted the overall linear acceleration and rotational acceleration of a "different" instance, 20, with a randomly chosen "normal" instance, 12. Instance 12 has a PLA value of 78.41 g and a PRA value of 15.1006 rad/sec^2, whereas instance 20 has a PLA value of 151.714 g and a PRA value of 23.0736 rad/sec^2. These differences are visible in the comparatively taller peaks in both plots of instance 20.