Pison – Data Scientist Technical Interview Challenge
Henry Valk
November 6th 2020

I solved this challenge by using dead-reckoning techniques and feature-based clustering to classify wrist motions embedded in whole-body movements. While the classification of unlabeled time-series data with an unknown number of classes proved difficult, I greatly enjoyed the challenge, and I hope you find my approach both creative and statistically robust. You can find the jupyter notebooks I used to conduct my analyses in this GitHub repository.

Exploratory Data Analysis

After observing summary statistics of the dataset, I decided to remove the gravitation component from the accelerometer data. I accomplished this in two steps. First, I negated the x-axis and y-axis accelerometer data so that the orientation of those measures matched that of the quaternion data. Second, I removed the acceleration due to gravity from each sample from each of the three axes accounting for the orientation of the device as represented by the quaternion data.

From there, I created a 'trial' column based on unique combinations of body movement label and repetition in order to more easily segment my data later in the process.

Finally, I visualized the ENG, quaternion, accelerometer, and gyroscope data in order to develop an intuition about the underlying structure of the dataset.

Gesture Reconstruction

To further develop an intuition of the data, I reconstructed both the orientation and position of the device across time. By adapting the view.orientation() function from the scipy-kinematics module, I was able to create an accurate animation of device orientation. The animation reveals three distinct movements, one in each of the three body movement repetitions. Next, I plotted position in 3d space for each of the 15 trials. To prevent error propagation created by simply double integrating the acceleration, I applied a Kalman filter that accounts for estimated variance in the accelerometer signal. The position reconstruction again revealed three distinct movements corresponding to repetition number.

Feature Engineering

I segmented the high pass channel 0, high pass channel 1, accelerometer, and gyroscope data with a 500-sample sliding window with 50% overlap between windows. Then, I calculated the area under the curve, variance, and absolute value of the mean for each window within each feature. Additionally, I calculated the number of zero-crossing for the accelerometer and gyroscope data. Finally, I standardized the features and projected them onto a composite feature space using Principal Component Analysis.

Cluster Analysis

I used the first four principal components, which cumulatively account for just over 63% of the variance in the engineered feature space, in the cluster analysis. Agglomerative clustering revealed three distinct clusters in the PCA feature space, with a silhouette score of 0.42. Furthermore, the agglomerative clusters, or presumptive unique gestures, are aligned in time according to the repetition number, as I found in the gesture reconstruction. To further validate the discretization into three clusters, I applied the Elbow Method for K selection in a subsequent K-means cluster analysis. The Elbow Method indeed reveals diminishing returns in distortion loss after increasing the number of clusters past three.

Conclusions and Future Directions

Intuitions gained through data reconstruction and quantitative estimates of underlying clusters in the principal component feature space revealed three distinct gestures embedded in the whole-body movements. The first gesture, Unknown, is present in all five conditions during the first repetition. The second gesture, which I'm calling Punch-and-Twist, combines an upward thrust with an internal rotation and is present in all five conditions during the second repetition. The third gesture, which I'm calling Open-and-Close, features a sweeping motion across the transverse plane of the body and is present in all five conditions during the third repetition.

While I am encouraged by the results of my classification analysis, there are limitations that I would like to address in future work. For example, my decision to engineer features in sliding time windows allowed for potential information loss. There could be information not captured by my engineered features, and the constraint of segmentation doesn't allow for information to be captured on multiple time-scales. Given the opportunity to work on this challenge further, I would explore the use of autoencoders for gesture classification. The autoencoder pipeline found here not only allows for entire timeseries as inputs, but also preserves temporal continuity in an unspecified number of unique patterns. Additionally, assessing generalizability in unsupervised learning tasks is an open problem. In future work, I will implement the generalizability criterion described here.

I hope you find my approach and solution compelling. I look forward to receiving your feedback and working on similar problems in the future.