Second Iteration

The ActivPal accelerometer was used to track acceleration of subjects during the lab session. The ActivPal accelerometer recorded data in resample rate of 20 Hz. Each activity duration was 5 minutes equals to around 6000 rows in data. Each row contains X-axis, Y-axis, Z-axis values with datetime indexes. The X, Y, Z-axis was processed to gravitational acceleration with equation 1.

$$G = \frac{a - 127}{63} \qquad \text{[equation 1]}$$

Where G is gravitational acceleration and a is scaled acceleration from ActivPal Acceleration of X, Y or Z axis. Accurate assessment of activity duration requires analysis in small segments, but segments of larger size might carry more meaningful information on the type of activity and improve the recognition of Physical activity (PA) (Bonomi, 2009). Multiple Random Forest models were created to track the best segment size. The data going into these models were segmented from 1 to 14 seconds with steps of 0.1 seconds. Accuracy, recall and precision were considered on deciding which 3 segment sizes to uphold. In the lab sessions the lab assistant logged each activity start and end time of every respondent which was used to label the time segments with the correct activity. Within each time segment the features for the model was created. These features consist of the mean and standard deviation of the X, Y and Z axis.

The dataset was split in train, validation and test following the steps which are described in chapter 'Train, Validation and Test dataset'. The validation dataset and K-fold cross-validation were used in development of Random Forest model to determine which segment sizes gave the best result. At the same time the amount of trees was decided upon experimentally. A range of 10 to 200 random forest models were created with each having different number of trees. The best number of trees was decided upon by looking at the accuracy score. The best segment sizes were decided upon looking at metrics accuracy, precision and recall on the validation dataset and the K-fold cross-validation. K-fold cross-validation was to determine how the model reacted to unseen data in 5 folds. The data which was used was combination of training and validation dataset

The Random forest model was developed with the processed data previously described as input. The activities the respondents performed in the lab session are as follows: cycling light, cycling heavy, walking, running, jumping and walking on stairs. Jumping and walking on stairs were excluded because of missing Vyntus data, therefore missing ground truth of MET-values. The ability of the model to distinct activities from each other were analysed visually with confusion matrix on the validation dataset.

Met opmerkingen [W(1]: Overal comment: activPAL ipv ActivPal :)

Met opmerkingen [W(2]: Dit zegt mij niet zoveel. Misschien dit uitleggen

Met opmerkingen [W(3]: Beter schrijven als X, Y and Z-axis

Met opmerkingen [W(4]: 'were' omdat het meervoud is. correct me if im wrong

Met opmerkingen [W(5]: Is een afkorting, maar de hele betekenis staat nog niet in de tekst (volgens mij)

Met opmerkingen [S(6]: dit zegt dat grotere segments recognition of PA kunnen verbeteren maar in de volgende zin staat er "therefore data was cut in time segments to classify activity in a smaller time range"

Met opmerkingen [T(7]: Hele sterke introductie

Met opmerkingen [W(8]: Is een afkorting, maar de hele betekenis staat nog niet in de tekst (volgens mij)

Met opmerkingen [W(9]: Is vanaf `accuracy` een begin van een nieuwe zin?

Met opmerkingen [S(10]: is er 1 segment size gekozen of meerdere sizes?

Met opmerkingen [T(11]: Hele sterke introductie

Met opmerkingen [W(12]: Kan wat concreter, bijv: K-fold cross validation was applied on the validation dataset to determine..

Met opmerkingen [W(13]: was.. ? Volgens mij mist er iets hierna

Met opmerkingen [W(14]: Leest niet zo vloeiend dit

Met opmerkingen [T(15]: Goed onthouden dat dit vooraf wel moet zijn omschreven wat vyntus data is.

Met opmerkingen [A(16R15]: Dit moet toch in subject design of intrudoction to subjects worden beschreven?

Bonomi, A. G., Goris, A. H., Yin, B., & Westerterp, K. R. (2009). Detection of type, duration, and intensity of physical activity using an accelerometer. *Medicine & Science in Sports & Exercise*, *41*(9), 1770-1777.