Strong-D Activity Prediction: Background

Chentian Jiang

4/9/2019

Strong-D

Stanford's STRONG-D Study: Strength Training Regimen for Normal Weight Diabetics

- Goal: determine the best exercise regimen for normal weight participants with Type 2 Diabetes
  - Manipulated variables: strength vs aerobic vs combined exercises
  - Response variable: indirect measures of blood sugar levels (hemoglobin A1c)

Strong-D Data

Table 1: Strong-D features used for RHR estimation.

Data Set	Features
Fitbit/Fitabase: HR*	Participant ID, timestamp (seconds), HR (bpm)
Fitbit/Fitabase: Steps	Participant ID, timestamp (minutes), num. steps
Clinic	Participant ID, timestamp (date), supine HR (bpm), sitting HR (bpm), standing HR (bpm)

- Fitbit data: 423 days of measurements for 78 participants from August 1, 2017 to September 28, 2018

Goal and Motivation

Build a machine learning model for classifying aerobic vs strength using wrist-worn, PPG-based Fitbit Charge 2 data.

- Convenience: wrist-worn, consumer product
- Estimate number of minutes for performing a type of activity
  - Study compliance
  - Weekly adult lifestyle guidelines (WHO): $>$150 minutes of moderate-intensity aerobic physical activity, $>$2 days with muscle-strengthening activities
- Determine which type of activity is occuring at which timestamp
  - Extract biomarkers according to the type of activity
  - Caveat: some biomarkers are used as features
- Feature engineering based on background literature and EDA
  - I hope that you all can help me brainstorm features throughout this presentation :)

Accelerometers for Classifying Activity Type

Figure 1: Wrist and hip accelerometers.

- Wrist, hip, ankle
- Static vs dynamic acceleration
  - $1\text{g} = 9.8 \text{ m/s}^2$
  - 1-3 axes
- 20–30 Hz

Accelerometers for Classifying Activity Type

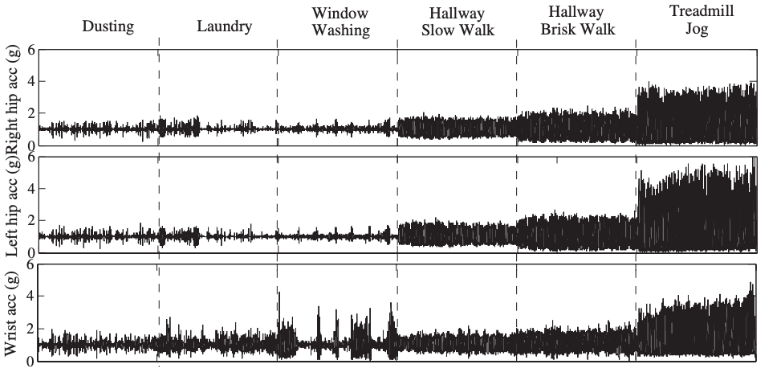


Figure 2: Example accelerometer measurements for one participant (Ellis et al., 2014).

- Activity categories: aerobic (e.g. walking, running, basketball), household (e.g. laundry), vehicle

Traditional Method for Classifying Activity Type

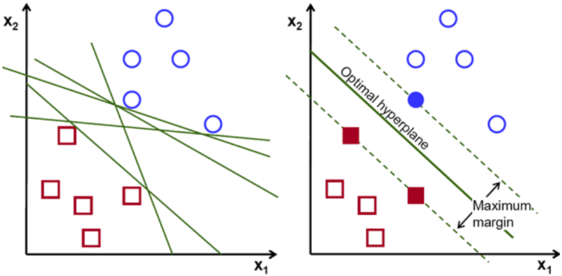
Cut-point

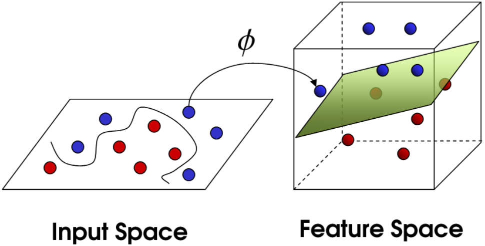
- e.g. sedentary ($<$ 100 CPM), light (100 – 1951 CPM), moderate to vigorous physical activity ($>=$ 1952 CPM) (Ellis et al., 2016)
  - CPM: counts per minute (proprietary accelerometer metric)

Common Machine Learning Models

- Support Vector Machine (SVM)
- Neural Network (NN)
- **Random Forest (RF)**
- 75 - 95% accuracy

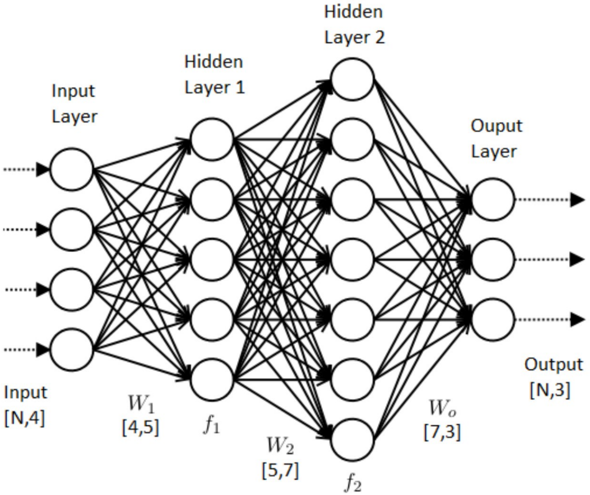
Support Vector Machine





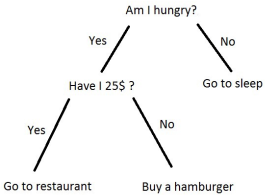
- Feature augmentation

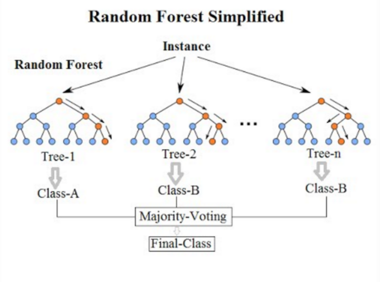
Neural Network



- Can represent complex relationships between features
  - Linear function ($W\mathbf{x}$)
  - Nonlinear (activation) function, e.g. ReLU: $f(a) = max(0, a)$

Random Forest





- Generalization
- Feature importance
- Successful with minor hyperparameter tuning

Features

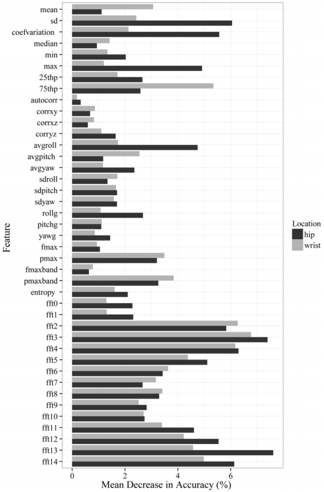


Figure 3: Example feature importance for a random forest classifier (Ellis et al., 2016).

Features

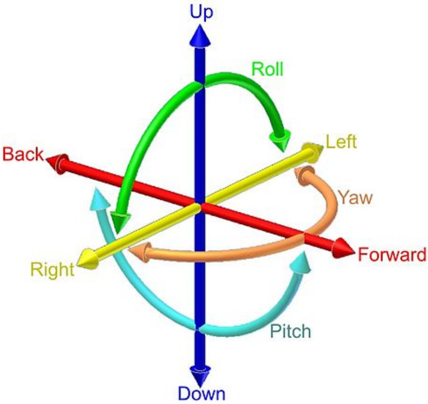


Figure 4: Pitch, yaw and roll.

HR Feature

- "Maximum HR was estimated for each participant by maxHR = 220 - age." (Ellis et al, 2014, p. 2195)
- Classification performance did *not* improve

Evaluation

	Laundry	Window	Dust	Dish	Sweep	Stairs	Walk	Jog
Laundry	29	8	4	45	7	0	3	0
Window	4	43	9	24	15	0	1	0
Dusting	4	11	23	12	45	0	1	0
Dishes	14	7	7	99	17	0	0	0
Sweeping	5	15	38	21	61	1	3	0
Stairs	0	2	0	4	0	110	22	6
Walk	3	1	1	12	0	21	435	7
Jog	0	0	0	1	0	1	12	196

Figure 5: Example confusion matrix, where the true labels are on the rows and the predicted labels are on the columns (Ellis et al., 2016).

- Leave-one-subject-out (LOSO) validation
- Precision: TP/(TP+FP)
- Recall: TP/(TP+FN)
- F-score: $\frac{2*\text{precision}*\text{recall}}{(\text{precision}+\text{recall})}$

Participants and Controlled vs Free-Living

"the high predictive accuracy of laboratory-calibrated models has not been reproducible in free-living settings." (Farrahi et al., 2019)

- Ellis et al. (2014): lab environment
- Staudenmayer et al. (2015): "as similar to free-living conditions (e.g., gardening and raking were performed outside and basketball was performed in a gym on a court) as possible." (p. 397)
- Ellis et al. (2016): free-living training measurements + labels via wearable camera

Gaps

- Classify strength/resistance vs aerobic
- Fitbit data
- Feature engineering informed by background literature and EDA
    - Personalized HR features
- Hybrid controlled and free-living conditions for training data collection

Feature Engineering Brainstorm

Build a machine learning model for classifying aerobic vs strength using wrist-worn, PPG-based Fitbit Charge 2 data.

Potential Features for a Random Forest Model

- rolling window summary functions for number of steps
- rolling window summary functions for HR
  - Rate of increase
- $HR - RHR_{est}$
  - RHR estimation model or clinical RHR measurements (supine, sitting, standing)