Continuous Inference of Psychological Stress from Sensory Measurement Collected in the Natural Environment
Plarre et al., 2011

Physiological Data Analysis for Educational Technologies

Iris Figalst, 13. Januar 2017
Agenda

- Motivation & Goal
- Procedure
- Data Collection
- Modelling Physiological Stress from Sensor Measurements
- Perceived Stress Model
- Stress from Self-Report
- Applying Stress Models to the Field Data
- Summary & Outlook
- Demo
Motivation & Goal

- Stress can lead to significant negative health consequences

- Challenges when measuring stress:
  - Physiological measurements: devices, confounders, between-person differences
  - Perceived stress measured by self-reports → unsatisfactory

→ Building a classifier that works on a large population in the natural environment
Procedure

- Collect data in a lab study
- Two models for continuous prediction of stress
  - Physiological classifier
  - Perceived stress model
- Applying stress models to field data
Data Collection

Quelle: http://www.frontfourgroup.com/assets/big-data.jpg
Participants

- 21 students (University of Minnesota)
- Meet health requirements
- Mean age 20.6, standard deviation 1.9
- 50% male, 50% female
Measures

- **Sensory measures by AutoSense wearable sensor suite**
  1. Electrical output of the heart measured by electrocardiograph (ECG)
  2. Skin conductance measured by ECG electrodes
  3. Skin temperature thermistor attached to skin mid thorax
  4. Ambient temperature sensor
  5. Three-axis accelerometer
  6. Relative lung volume measured by respiratory inductive plethysmograph (RIP)

- **Self-Report**
  - Questions describing subjective stress state
    - Cheerful?, Happy?, Angry/Frustrated?, Nervous/Stressed?, Sad?
    - Four-point scale: 0 (NO), 1 (no), 2 (yes), 3 (YES)
Lab Procedure

- Two-hour lab session
- Rest sessions before, between and after stressors
- Stressors (social, cognitive, physical):
  - Public speaking
  - Mental arithmetic
  - Cold pressor
- 14 Self-Reports, during rest sessions & before/after stressor

Quelle: Plarre et al., 2011
Field Procedure

- Same participants
- Collecting sensor data on two separate days
- Physiological samples stored on mobile phone
- Phone periodically prompts participants to complete self-report questionnaires (25 per day)
Modelling Physiological Stress from Sensor Measurements

Quelle: http://www.anatomy4beginners.com/resources/Anatomy-and-physiology-study-bones.jpg
ECG Features I

- Features derived from windows of R-R-intervals
  - Duration between successive R-peaks
  - Ratio between low- and high-frequency of heart beat
  - Heart beat frequency in 3 bands (low, medium, high)

- Preprocessing for training
  - Synchronization of physical measurements to timing of stressors
  - Computation of R-R-intervals
  - Tagging outliers
  - Ignore one minute of data immediately following each self-report
  - Base features are normalized to account for between-person differences
  - Each one-minute window labeled with ground truth

Quelle: Plarre et al., 2011
- R-R-intervals obtained from a participant
- Red dots: valid R-R-intervals
- Blue crosses: Outliers

Quelle: Plarre et al., 2011
Respiration Features I

- Respiration features from respiration cycles (inhalation/exhalation period)
  - Inhalation duration
  - Exhalation duration
  - Respiration duration
  - IE-ratio
  - Stretch
  - Minute ventilation/ minute volume
  - Breath rate

- Respiratory sinus arrhythmia (RSA)
  - Multimodal feature derived from ECG and respiration
  - Variability in R-R-intervals due to respiration
Preprocessing for training

- Respiration signal is segmented into one-minute segments
- Discard segments that are missing more than 15% of their samples
- Identify peak and valley in each respiration cycle, mark start/end
- Threshold (75th percentile) for peaks to remove spurious peaks
- Duration between two peaks at least 1.5 seconds
- Normalization of each feature to account for between-person differences

Three base Features computed from a respiration signal
Quelle: Piarre et al., 2011
Classifiers and Evaluation Metric

- Compute four statistics on features that have multiple values per minute (mean, median, quartile deviation, 80th percentile)

→ 35 features used to train the classifiers

- Three types of classifiers
  - J48 decision tree
  - J48 decision tree with adaptive boosting (AdaBoost)
  - Support vector machine (SVM)

- 10-fold cross validation to obtain performance measures of all three classifiers
Classification Results on Lab Data I

- Classification accuracy when using ECG or RIP
  → Both features highly discriminatory of stress

- Effect of normalization
  - Improves accuracy of classification
  - Decreases number of false negatives/positives

- Classification accuracy for individual stress tasks
  → Accuracies >= 95% if goal is to classify stress for specific stressors

<table>
<thead>
<tr>
<th></th>
<th>J48 Decision Tree</th>
<th>J48 with Adaboost</th>
<th>SVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not Normalized</td>
<td>82.33%</td>
<td>88.00%</td>
<td>88.17%</td>
</tr>
<tr>
<td>Normalized</td>
<td>87.67%</td>
<td>90.17%</td>
<td>89.17%</td>
</tr>
</tbody>
</table>

Accuracies of three different classifiers trained with normalized and unnormalized features

Quelle: Plarre et al., 2011
Classification accuracy when only using selected features

- 13 most distinguishing features chosen by a correlation-based feature selection algorithm
- Does not improve performance
- Sometimes accuracy improves

→ smaller feature set for implementation on the mobile phone for real-time detection of stress level in the field

Best possible accuracies of classifiers

Quelle: Piarre et al., 2011
Perceived Stress Model

Quelle: http://www.islamiclife.com/2016/08/22/8558.jpg
Stress rating, initialized to the average rating of stress reported in the first self-report

\[ \hat{\pi}_k = \alpha \hat{\pi}_{k-1} + \beta x_k \]

Decay of perceived stress

Previous stress rating

Accumulation of perceived stress

\{$0,1\}$, physiological stress (output of the classification algorithm)

- \(\alpha\) and \(\beta\) personalized, calculated by least square method
- \(\rightarrow\) minimize errors between perceived stress rating and self-reported stress rating
Perceived Stress Model

- Maps physiological stress to perceived stress
- Considers value of perceived stress as hidden states in a Hidden Markov Model
  - States “stressed“/“non-stressed“
  - Output of physiological classifiers as observables
- Estimate probability of current minute being stressed as a linear function of
  - Observation from the physical classifier for the current minute
  - Probability of the previous minute being stressed
- Perceived stress score increased by accumulation factor each time a minute is marked as “stressed“
- Score decays at an individual specific exponential rate with each passing minute
- Accumulations and decay rates personalized to each subject using their self-report ratings
Evaluation of the Model on the Lab Data

- Aims to predict the self-reported rating of stress
- Computed correlation coefficient of values from $\pi(k)$ and self-reports

Probability density for the linear correlation coefficient

Quelle: Plarre et al., 2011
Stress from Self-Report

Quelle: https://shelleypsych.files.wordpress.com/2015/08/image19.jpg
Stress from Self-Report

- Self-reports provide subjective stress ratings on a scale from 0 to 3
- Machine learning approach to train self-reported stress classifier using lab data (ground truth known)
- Computed features (e.g. mean, standard deviation etc.)
  - Momentary features: features associated with a single self-report
  - Aggregate Features: features computed across all self-reports provided by an individual
- Computation of statistical features, z-scores \( \frac{\text{featureValue} - \text{meanOfFeature}}{\text{sdOfFeature}} \) and histogram bin counts over features
- J48 decision tree over features correctly classified 84% of self-reports using z-score of momentary features \( \Rightarrow \) „stressed“ if z-score > 0.6
Applying Stress Models to the Field Data

Quelle: http://uhd-wallpapers.net/images/wheat-field_421.jpeg
Screening and Cleaning of Field Data

Removal of segments of data due to noise, confounds and losses

- Outliers
- Minutes of data during a self-report
- Minutes of data that show significant motion
- Two minutes following physical activity
- Minutes of ECG data that have less than 30 valid R-R-intervals
- Minutes that have less than 66% RIP samples

Quelle: Plarre et al., 2011
Evaluation of Perceived Stress Model on Field Data

- No personalization and matching to self-reports due to data loss

→ $\alpha$ and $\beta$ for each subject from their lab data, calibration from self-reports collected in the field

- Compare average rating of stress provided by each subject to average rating produced by perceived stress model

Agreement between self-report rating of stress and perceived stress model rating

Quelle: Plarre et al., 2011
Three Measures of Stress in the Field

<table>
<thead>
<tr>
<th>Self-report classifier</th>
<th>Perceived stress model</th>
<th>Physiological classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binary ratings of stress based on subjective ratings of stress</td>
<td>Aims to predict self-reported ratings of stress</td>
<td>Binary ratings of stress based on physiological data</td>
</tr>
<tr>
<td>28.08% stressed of the time</td>
<td>26.61% stressed of the time</td>
<td>35.14% stressed of the time</td>
</tr>
</tbody>
</table>
Summary & Outlook

▪ Classifiers/models:
  ▪ Physiological classifier: classify physiological data into states (stressed/non-stressed)
  ▪ Perceived stress model: maps physiological response of stress to perceived stress
  ▪ Self-report classifier: classify self-reported ratings into states (stressed/non-stressed)

▪ Using ECG and respiration features

▪ Future work:
  ▪ Improvements in wearable sensors
  ▪ New methods to control the effect of physical activities
  ▪ Investigate additional features, e.g. body temperature, oxygen level in blood, pulse transit time,...
Demo

Quelle: https://cdn3.pcadvisor.co.uk/cmsdata/reviews/3585633/fitbit-charge-hr-heart_thumb800.jpg
Stila: Heart Rate

Aufstehen, Weg zur Uni, Big Data VL, Scala VL, Arbeit, Heimweg, Einkaufen

Pause
Stila: Computed Stress

Heart Rate Variability
Computed Stress

From Nov 8, 2016 05:43:24 To Nov 8, 2016 21:43:24

06:00 07:00 08:00 09:00 10:00 11:00 12:00 13:00 14:00 15:00 16:00 17:00 18:00 19:00 20:00 21:00

Aufstehen  Weg zur Uni  Big Data VL  Scala VL  Arbeit  Heimweg  Einkaufen

Pause
Feature: CVRR

CVRR = Coefficient of variation of R-R-intervals:

\[ CVRR = \frac{sdRR}{meanRR} \]