Project in Data Intensive Systems

4DV652

Lab Lecture 3 – Weak link classification

Welf Löwe

1

Agenda

- Classification
- Lab 3 task descriptions

2

Classification

- Map predictors $X_1, \ldots X_p$ to discrete response $Y = \{c_1, \ldots, c_k\}$
- Different approaches
  - Parametric: Logistic regression, Naïve Bayesian Classifier, Linear/Quadratic Discriminant Analysis, …
  - Non-parametric: K-Nearest-Neighbors
- Assessment of classification accuracy
  - Error rate
  - Precision/recall and F-score

3

Example: Weak-Link Classification

- What joint is the weakest during a movement
- Based on the 38 movement predictors X
- What is the weakest link Y for that movement out of 14 weak link classes:
  - ForwardHead
  - LeftArmFallForward, RightArmFallForward
  - LeftShoulderElevation, RightShoulderElevation
  - ExcessiveForwardLean
  - LeftAsymmetricalWeightShift, RightAsymmetricalWeightShift
  - LeftKneeMovesInward, RightKneeMovesInward
  - LeftKneeMovesOutward, RightKneeMovesOutward
  - LeftHeelRises, RightHeelRises

4

Prepare the data

- Input: 2 Excel files
  - with 38 (33 unique) features (you worked with this one, new ID column)
  - with 14 weak link features (new)
  - Data points are related with their unique ID
- Tasks:
  - Find the weakest link for each data point, i.e., the column with the maximum score in each row; add this as a new column
    - Alternatively find the body region with the weakest link, e.g., upper body (shoulders and arms) vs lower body (hips and legs)
  - Join the two tables via their ID
  - Discard the 14 weak link scores from the joint table (you must not use them as predictors)

5

Agenda

- Classification
- Lab 3 task descriptions

6

Lab assignment 3: classify the movements

- ML
  - Classify the weakest link based on the 38 predictors of a movement using a number of classification variants (different models or model parameterizations)
  - Select an accuracy metric
  - Test and iteratively improve the accuracy of the variants
- Software development
  - Add a new endpoint for weakest link classification to the server that implements the champion variant
  - Extend the web client to use also this new endpoint
- Reporting in a third notebook:
  - The iteration(s) over ML process steps
  - Your implementation
- Deadline: 2023-02-15

7