Project in Data Intensive Systems

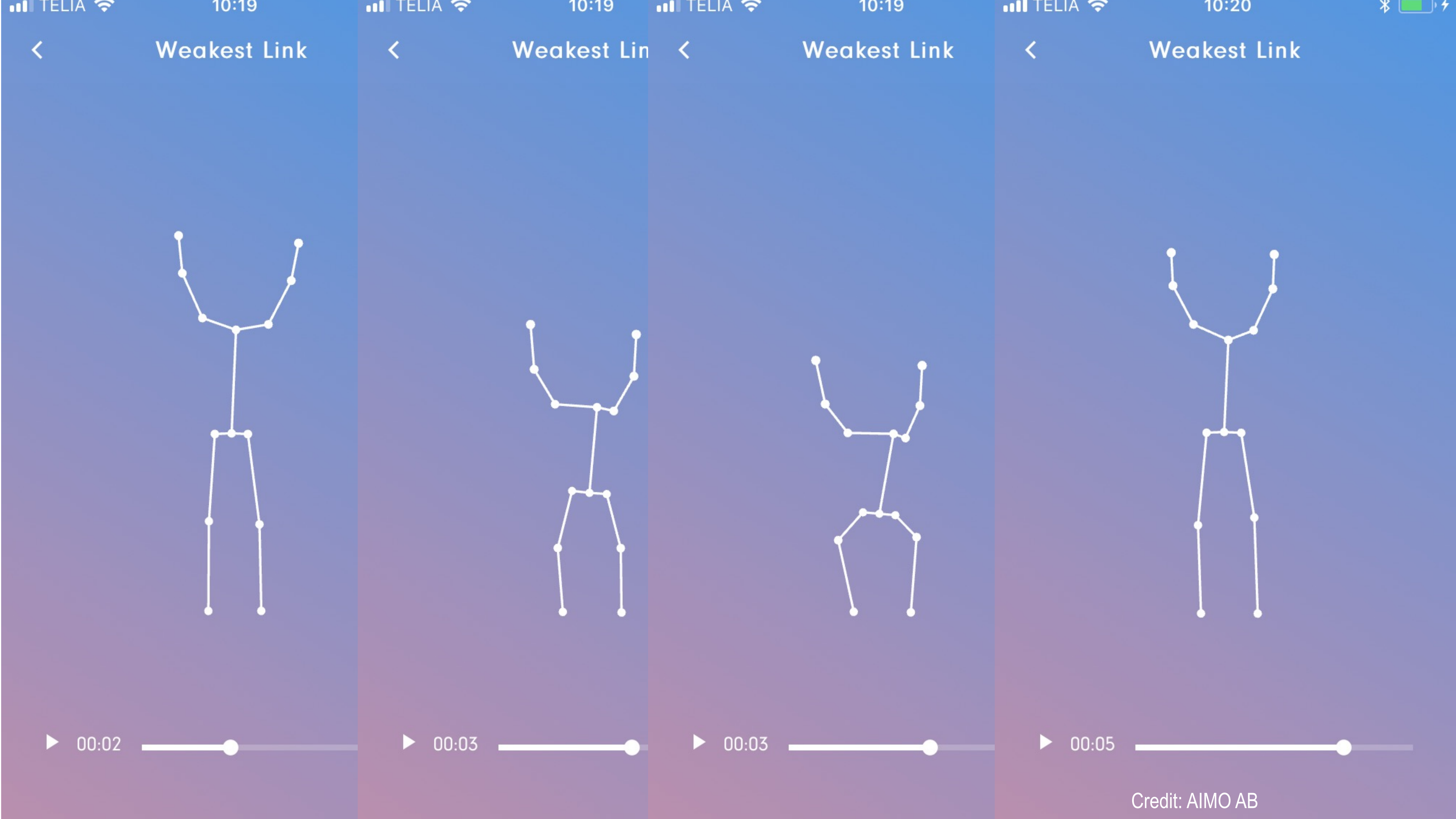
4DV652

Automated Movement Assessment

Welf Löwe

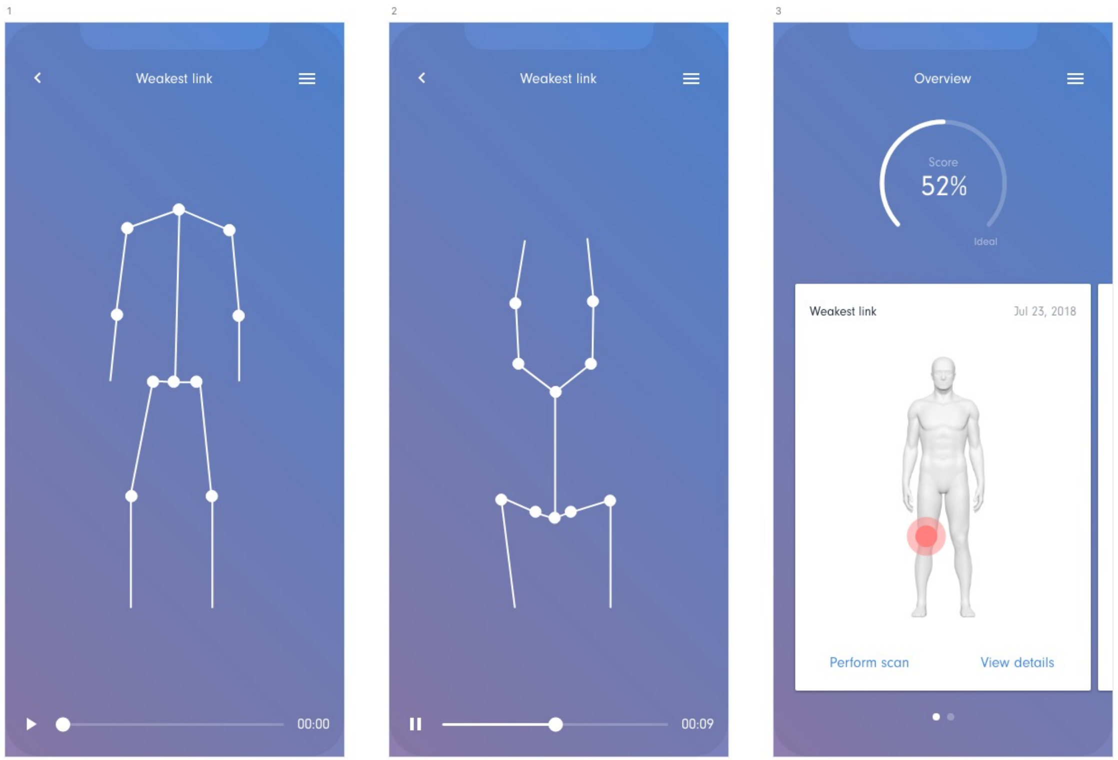
Agenda

- Data intensive system: Automated Movement Assessment
  - Context
  - Description of dataset
- Lab 2 task descriptions



Credit: AIMO AB

Quality score and weak links based on movement data



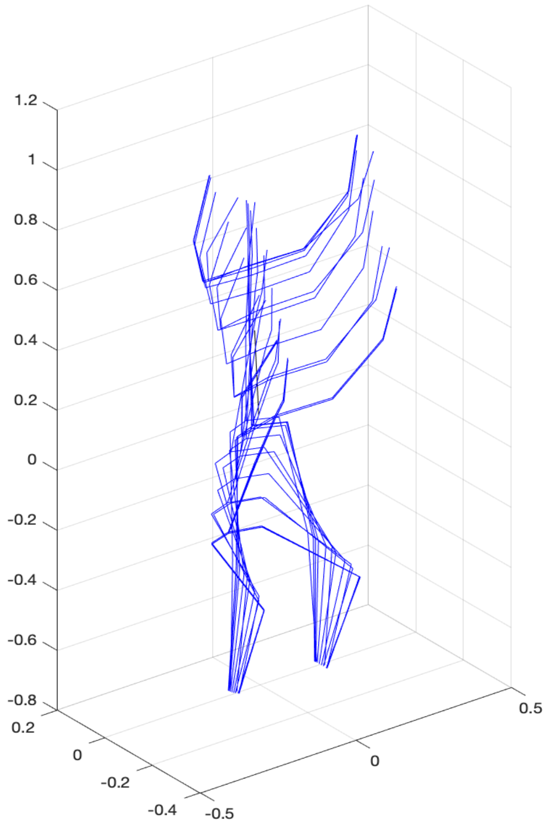
Agenda

1. Technical setup
2. Input data and direct features
3. Preprocessing
4. Compute indirect features
5. Matching with a master sequence
6. Building a statistical model
7. Scoring
8. Experiments

Technical Setup

- Kinect 3D camera providing a video recording of the movement
- Kinect SDK, producing a 3D skeleton avatar sequence of the movement based on the video and depth information
- Nowadays replaced by a 2D mobile camera

Input Data and Direct Features



- A skeleton **sequence**, recording of a person's movement consisting of frames
- Each **frame**: vector of features describing a posture of the movement at one point in time at 33 frames/sec
- Each **feature**: aspect of a posture, e.g., the *x*, *y*, *z*-coordinates of the left knee

Preprocessing

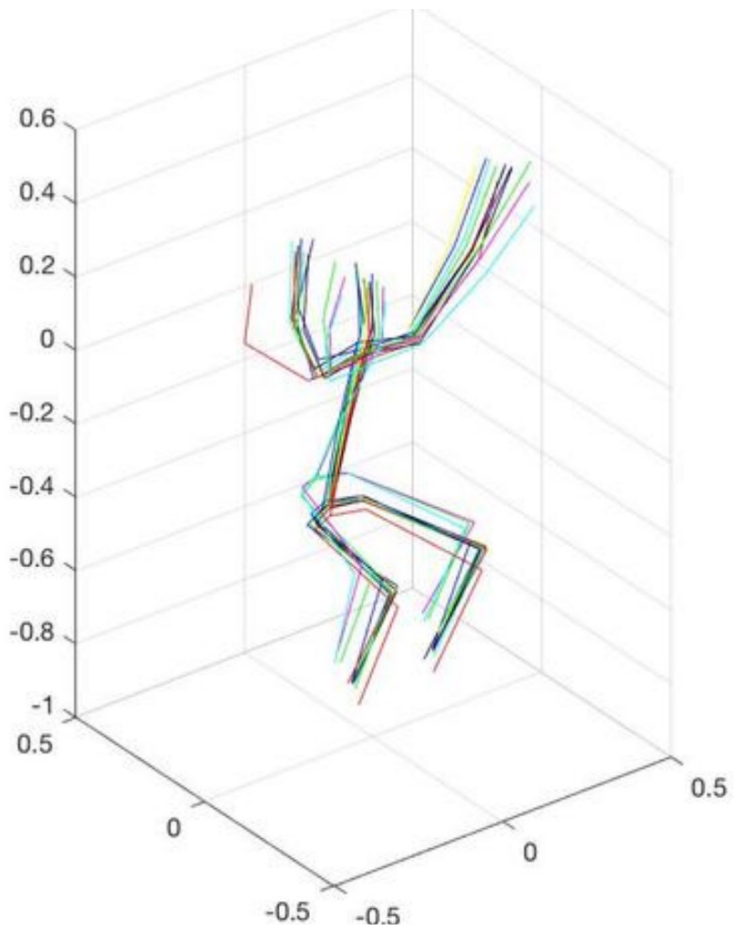
- **Smoothening**: for all direct features, a sliding window technique averages the feature values.
- **Floor plane alignment**: for each frame, the joint position vectors are rotated such that the floor plane is parallel to *x*, *z* plane.
- **Cut leading and trailing frames** of postures not belonging to the movement.
- **Scaling** transformation: (Procrustes analysis) moves each joint of the first frame of the sequence to the corresponding joint position of the first frame of a master sequence and applies it to all other frames of the sequence.
- **Moving** and **Rotating** transformation: moves the Spine Base joint of the first frame of the sequence to the origin of the coordinate system and rotates it so that it “faces” the camera and applies it to all other frames of the sequence.
- **Interpolate or error**: if a joint was not visible for **less** than *k* sequential frames, then its position is interpolated. If a joint was not visible for *k* or **more** sequential frames, then the joint is not tracked in the sequence (error).

Compute Indirect Features

- Direct features: the x, y, z-coordinates of 25 joints
- Ignore unstable joints (finger and toe tips) leaving 17 joints
- Add indirect features
  - Angles between the limbs adjacent to the remaining joints
  - Selected angles between limbs projected to planes and the x, y, z-axes,

    For instance, to check if the body is tilting to the left or right, calculate angle between the spine projected to the x, y-plane and the y axis

Matching – Problem and Idea



- Statistical model: expected feature values in the different movement phases?
- Different movement speed: we cannot synchronize on time or frame number

- Therefore: select a master sequence $\mathcal{M}$
- **Matching**: Find a minimum distance mapping, map of each frame in $\mathcal{M}$ to one or more frames in the user sequence $\mathcal{N}$
  - *Distance of mapping*: average of distances of all mapped frames
  - *distance*[n, m] *of two frames*: Euclidian distance of feature vector $\mathcal{N}(n)$ to feature vector $\mathcal{M}(m)$ in user frame n to master frame m
- Dynamic programming problem

Building a Statistical Model

- Split the master sequence $\mathcal{M}$ into j phases, each with the same number of frames
- Assume k features
- Match each user sequence $\mathcal{N}$ with $\mathcal{M}$
  - For each phase m of the master sequence $\mathcal{M}$
    - For each feature, add the feature value of frames in $\mathcal{N}$ matched to a frame in m to a separate sample of this feature
    - This leads to k x j sample sets
- For each of the k x j sample sets, interpolate the probability density functions PDF of that sample distribution

- 15 phases with ca. 10 frames, ca. 100 features, ca. 2000 sequences

Single feature scores

- Match each user sequence $\mathcal{N}$ with the master sequence $\mathcal{M}$
  - For each phase m of the master sequence $\mathcal{M}$
    - For each feature, compute the z-scores of the feature value in the corresponding (interpolated) PDF of that feature and phase
    - Take the average of the three largest z-scores (most deviating feature values) for each feature
- Result is a z-score vector for each of the k features $[z_1 \ \ldots \ z_k]$

- Input training data in Lab 2.

Single overall score

- Aggregates single values per feature to a single overall score value per sequence
- Scoring function comes in variants $\mathrm{S}_1 \ldots \mathrm{S}_4$
  - $\mathbf{S_1}(z_1 \ldots z_k) = 1 - \Pr(Z_1 \leq z_1, \ldots, Z_k \leq z_k)$, probability of a worse user sequence
  - $\mathbf{S_2}(z_1 \ldots z_k) = [z_1 \ldots z_k][w_1 \ldots w_k]^{\mathrm{T}}$, weighted sum with weights set by experts
  - $\mathbf{S_3}(z_1 \ldots z_k) = [z_1 \ldots z_k][w_1 \ldots w_k]^{\mathrm{T}}$, weighted sum with weights $w_1 \ldots w_k$ computed in linear regression against scores set by experts (response)
  - $\mathbf{S_4}(z_1 \ldots z_k) = [z_1 \ldots z_k][w_1 \ldots w_k]^{\mathrm{T}}$, weights as S_3 subject to left-right symmetry constraints such as $w_i = w_j$
- You will add your own variants in Lab 2

Ground truth: expert scores (response)

- Expert overall scores are values between 0% (bad exercise) and 100% (excellent exercise)
  - Weighted average of expert weak link scores weak links
- Expert weak link scores come from a digitalization of a deep squat assessment method (not data driven)
  - Validated on <100 sequences with real expert scores
- How good is a data driven scoring variant, e.g., a variant $S_1 \ldots S_4$ or the variant that you will suggest in Lab 2, compared to the expert overall score?
  - Increasing correlation of expert overall score and scoring variants $S_1 \ldots S_4$

Conclusion of Application Context

- High correlation of S_4-scores with expert scores
- Can be improved using other ML approaches (future Labs)
- Adding new movements requires expert scoring (60-500 expert scores, the equivalent of 1-7 person days) to get a high correlation with high probability

- Supported by **AIMO AB** https://aimo-fit.com/english

- **Danny Dressler, Pavlo Liapota**, Welf Löwe, "Data driven human movement assessment," in 11th Int. Conf. Innovation in Knowledge Based and Intelligent Engineering Systems, (KES), Invited Session on Digital Health, Distance Learning and Decision Support for eHealth, Springer Intelligent Decision Technologies, 2019.
- **Johan Hagelbäck, Pavlo Liapota, Alisa Lincke**, Welf Löwe, "The Performance of Some Machine Learning Approaches in Human Movement Assessment", 11th Int. Conf. e-Health (EH), 2019

Description of movement assessment dataset

- Deep squat data assessed for ca 2000 persons

- 1 expert score (AIMO Score, large score is good)
- 38 movement features (deviations, large is bad)
- 2 time-features (deviations in time and frames, large is bad)
- 1 estimated score (indirect feature, large score is good, ignore)

- Features are correlated, some are even identical (to be checked and removed)

Functional Movement Scan (FMS) features

Symmetric FSM features (angles) at positions [4 6][5 7][8 11][9 12][10 13]

1 SpineBase, -Mid, -Shoulder

2 SpineMid, -Shoulder, Neck

3 SpineShoulder, Neck, Head

4 SpineShoulder, ShoulderLeft, ElbowLeft

5 ShoulderLeft, ElbowLeft, WristLeft

6 SpineShoulder, ShoulderRight, ElbowRight

7 ShoulderRight, ElbowRight, WristRight

8 SpineMid, -Base, HipLeft

9 SpineBase, HipLeft, KneeLeft

10 HipLeft, KneeLeft, AnkleLeft

11 SpineMid, -Base, HipRight

12 SpineBase, HipRight, KneeRight

13 HipRight, KneeRight, AnkleRight

National Academy of Sports Medicine (NASM)

Symmetric NASM features (angles) at position: [14 15][17 18][21 22][24 25][26 27][28 29][31 32]

14 (= 5 FMS) ShoulderLeft, ElbowLeft, WristLeft

15 (=7 FMS) ShoulderRight, ElbowRight, WristRight

16 (=1 FMS) SpineBase, -Mid, -Shoulder

17 (=10 FMS) HipLeft, KneeLeft, AnkleLeft

18 (=13 FMS) HipRight, KneeRight, AnkleRight

19 ShoulderLeft, ShoulderRight, x-axis

20 Neck, Head, z

21 ShoulderLeft, WristLeft, z

22 ShoulderRight, WristRight, z

23 SpineBase, -Shoulder, z

24 KneeLeft, AnkleLeft, z

25 KneeRight, AnkleRight, z

…

26 HipLeft, KneeLeft, y

27 HipRight, KneeRight, y

28 HipLeft, AnkleLeft, z

29 HipRight, AnkleRight, z

30 SpineBase, SpineShoulder, z

31 ShoulderLeft, WristLeft, z

32 ShoulderRight, WristRight, z

33 WristLeft, WristRight, x

34 ShoulderLeft, ShoulderRight, x

35 HipLeft, HipRight, x

36 KneeLeft, KneeRight, x

37 AnkleLeft, AnkleRight, x

38 Arms symmetric

Features

- Movement features (angle deviations)
  - 13 Functional Movement Scan (FMS) features at positions 1—13 with weights of importance [1 1 1 1 1 1 1 2 2 2 2 2 2]
    - E.g., features 8+9/11+12: left/right asymmetrical weight shift; 10/13: left/right knee moves inwards or outwards
  - 5+20 National Academy of Sports Medicine (NASM) features at positions 14—38 with weights of importance [1 1 1 2 2] + [1 1 1 1 2 4 4 2 2 2 2 2 1 1 1 2 2 2 2 2]
    - E.g., features 24/25: left/right heels up with the highest weight of 4
- Time features
  - 2 features expressing the relative deviation of movement phase length and speed with weights of importance [1 1]
- OBS: Large feature score indicates a high deviation from the expectation

Left-Right Feature Symmetry

- E.g. an asymmetrical weight shift left should be treated just as bad as an asymmetrical weight shift right
- Symmetric FSM features at positions
  - [4 6][5 7][8 11][9 12][10 13]
- Symmetric NASM features at positions
  - [14 15][17 18][21 22][24 25][26 27][28 29][31 32]

Lab assignment 2: Movement Assessment

- ML
  - Train a linear regression model that maps features to expert scores based on the dataset
  - Try different variants using, e.g., feature selection, combined features, removal of outlier and leverage data points, symmetry constraints, weights
  - Assess the accuracy of the models and select a champion variant
- Software development
  - Deploy a service with the champion variant that receives a record of feature values and returns an expert score
  - Deploy a remote test Web client that interacts with that service
  - Define a set to automatically and continuously test and deploy new variants of the regression model
- Reporting
  - Report in a second notebook: iterations over ML process steps, deployed client server system (usage, architecture, design, and implementation) and DevOps/MLops process
- Deadline: 2024-02-07

Process of 7 Steps in ML Projects