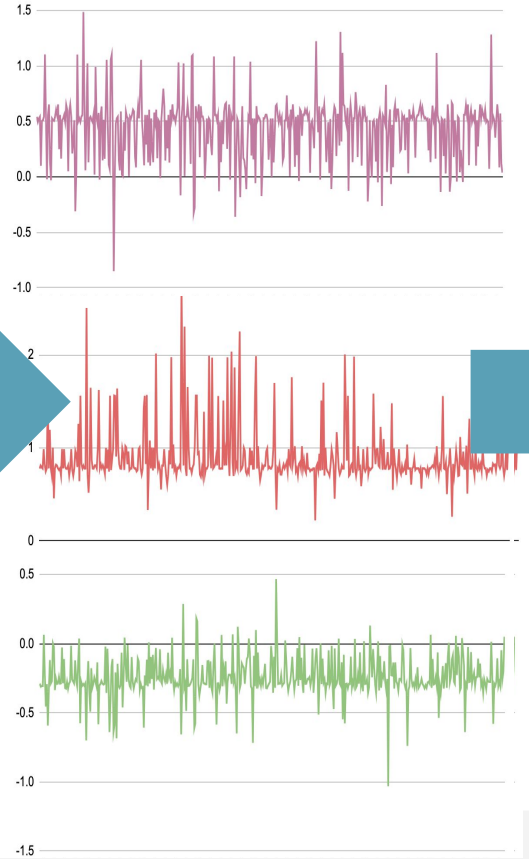
Stepping Up:
Improving Step Count ML Algorithms

Brendan Callender, Jadyn Ellis,
Martin Hsu, Kirina Sirohi

Using Machine Learning to Count Steps



8887 Steps

01
Data

Accelerometer Data (ACT24)

Device Data

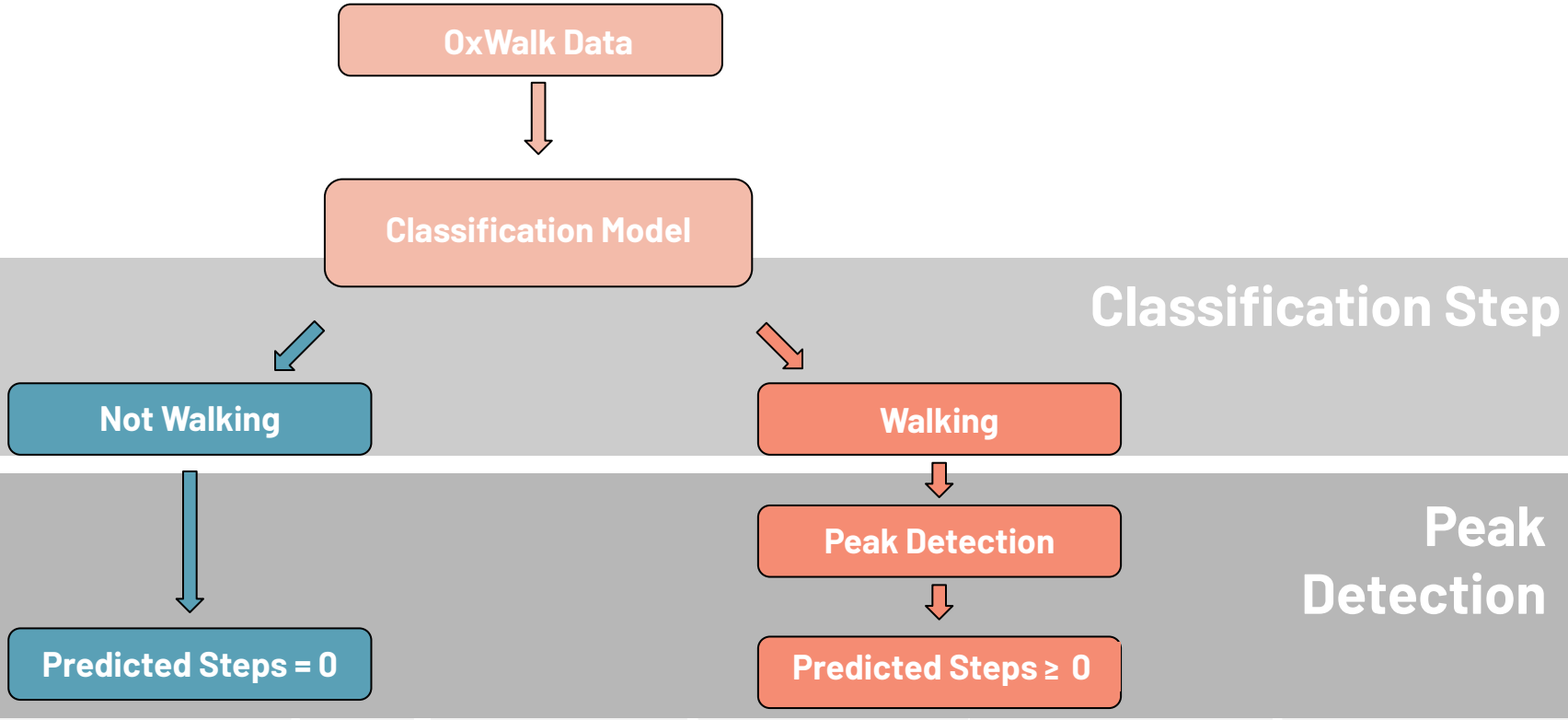
- Wrist-worn accelerometers
- 80Hz → 80 readings / second
- 3 axis → X, Y, Z

Ground Truth

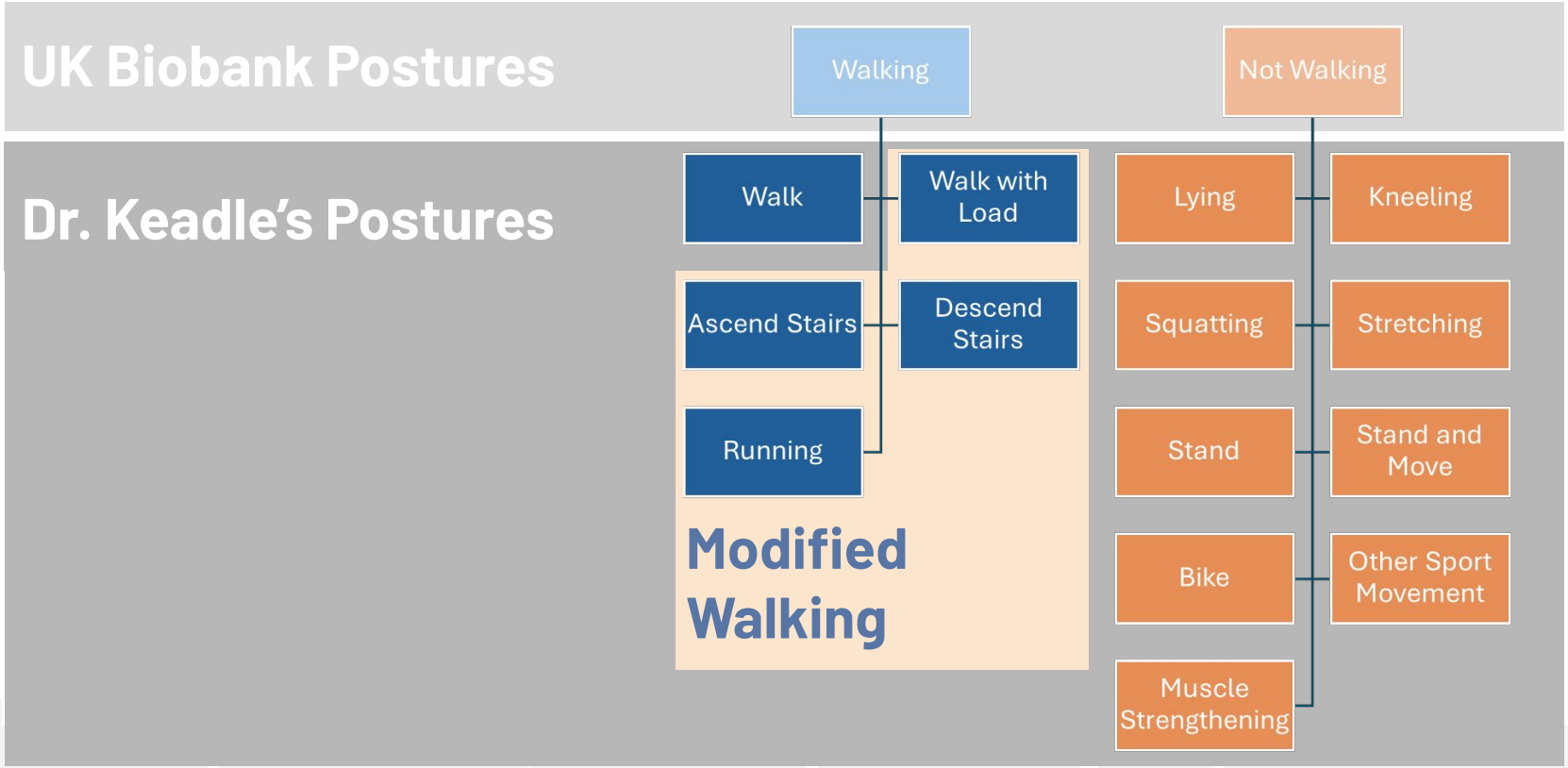
- 24 people, 2x3 hour video recordings
- Labeled by Dr. Keadle's Team
- Steps
- Activity type, posture, intensity

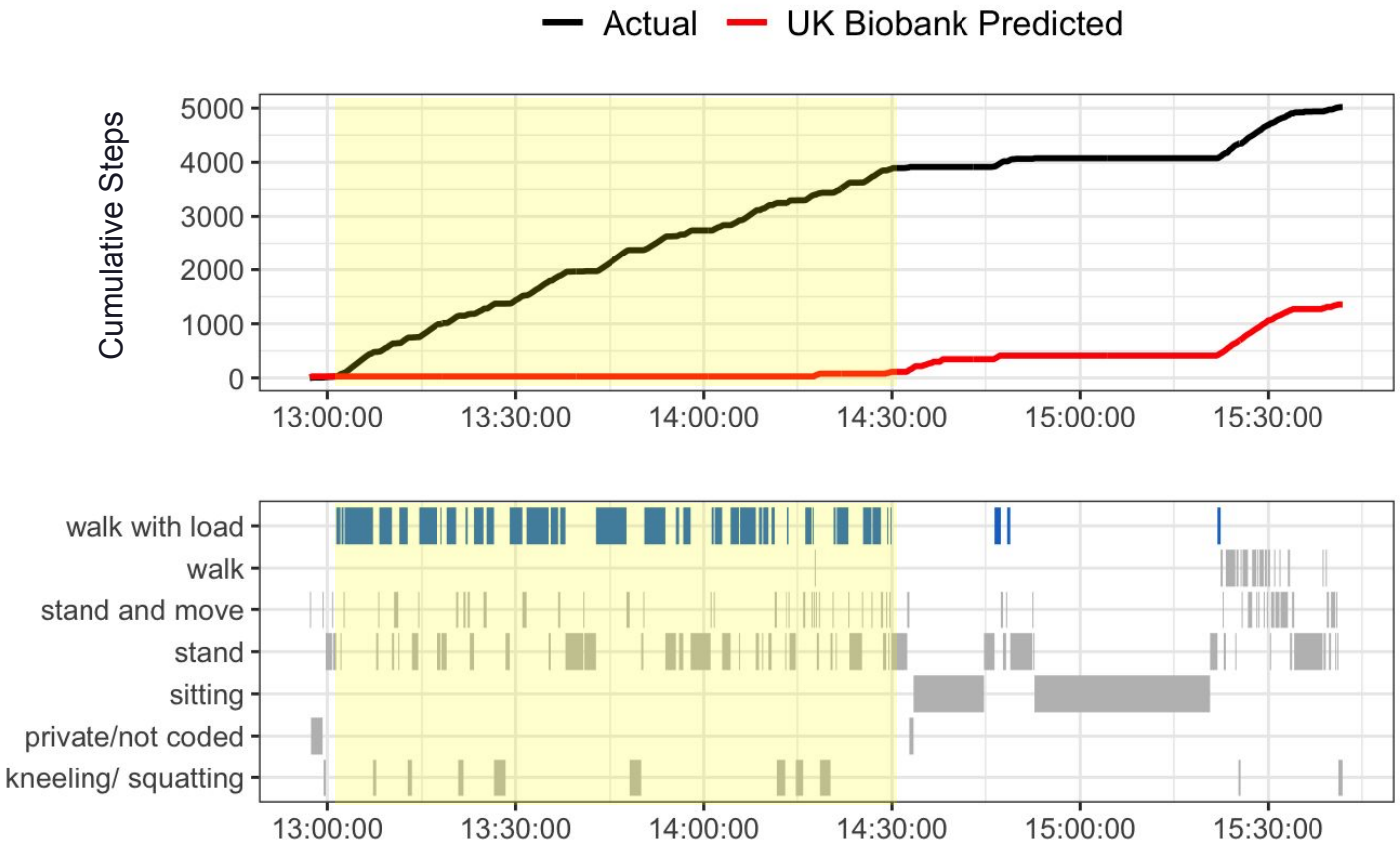
02 Evaluation of UK Biobank Algorithm

UK Biobank


OxWalk Data
Classification Model
Classification Step
Not Walking
Walking
Peak Detection
Peak Detection
Predicted Steps = 0
Predicted Steps ≥ 0

ACT24 Taxonomy is More Granular than Biobank





Biobank algorithm struggles with modified walking

Modified walking has a large impact on error

When tested on ACT24, **UK Biobank algorithm** is estimated to be **off by**

2.02

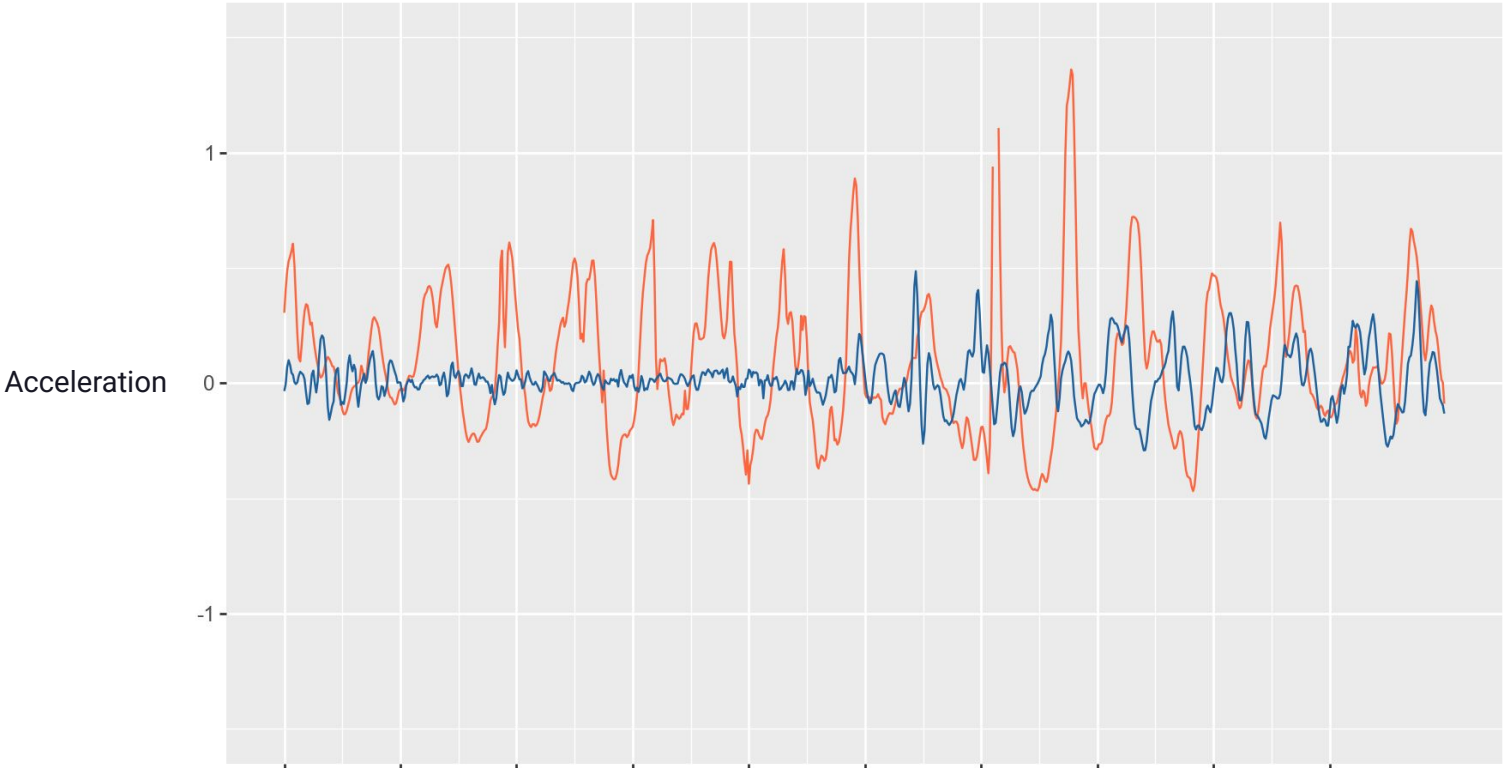
steps per minute **overall**

22.90

steps per minute in periods of **modified walking**

Different Postures Convey Distinct Signals

Walking vs Walking with Load



New Research Answers/Questions

- **Q: How can we improve step counting?**
  - A: Focus on modified walking
  - A: Classify on more types of postures/activities

- **Q: How many categories should we classify on?**
  - Q: How many categories are too few? Too many?
  - Q: What should the categories be?

03
Developing Models

Process

1. **Define** different levels of classification granularity
2. **Modify** algorithm to handle more categories and use ACT24 data instead of UK Biobank data
3. **Cross-validate** on ACT24 data
4. **Train** on ACT24, **test** on UK Biobank data
5. **Calculate** error metrics, **compare** across granularity levels and to original algorithm

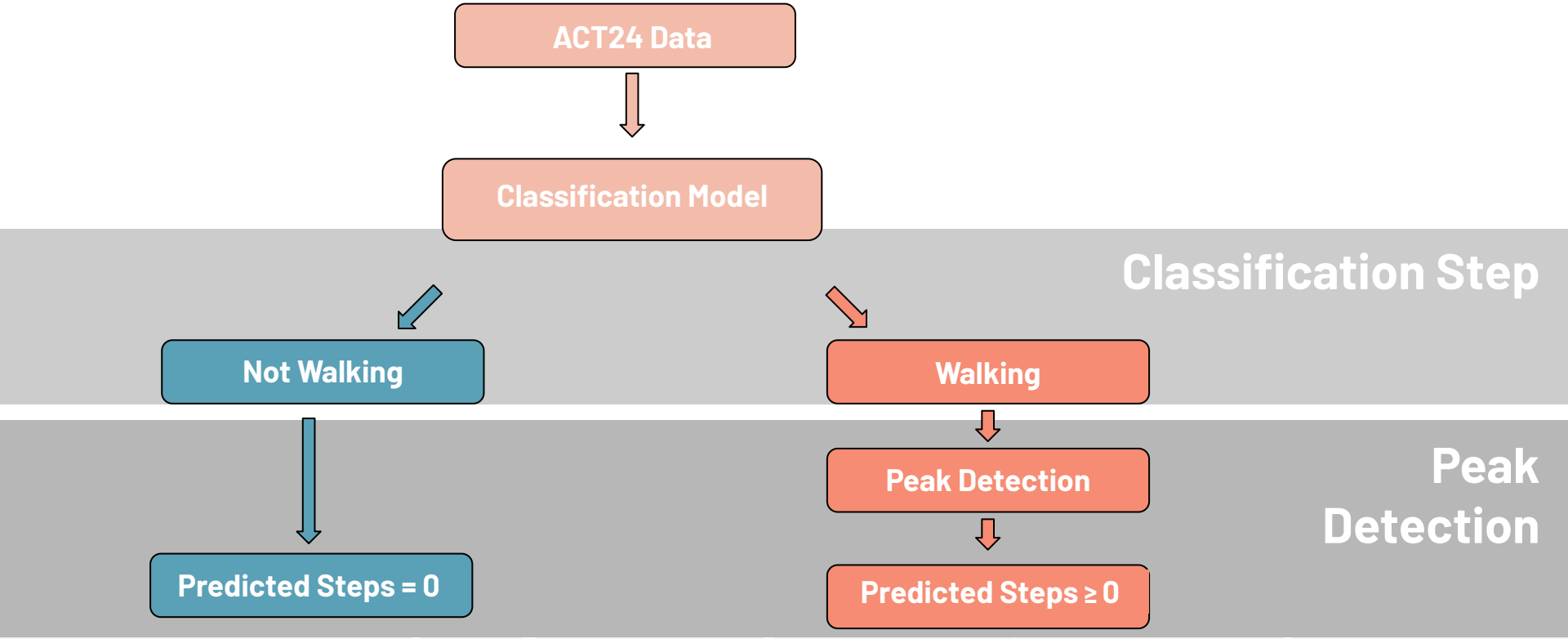
Three Classification Levels

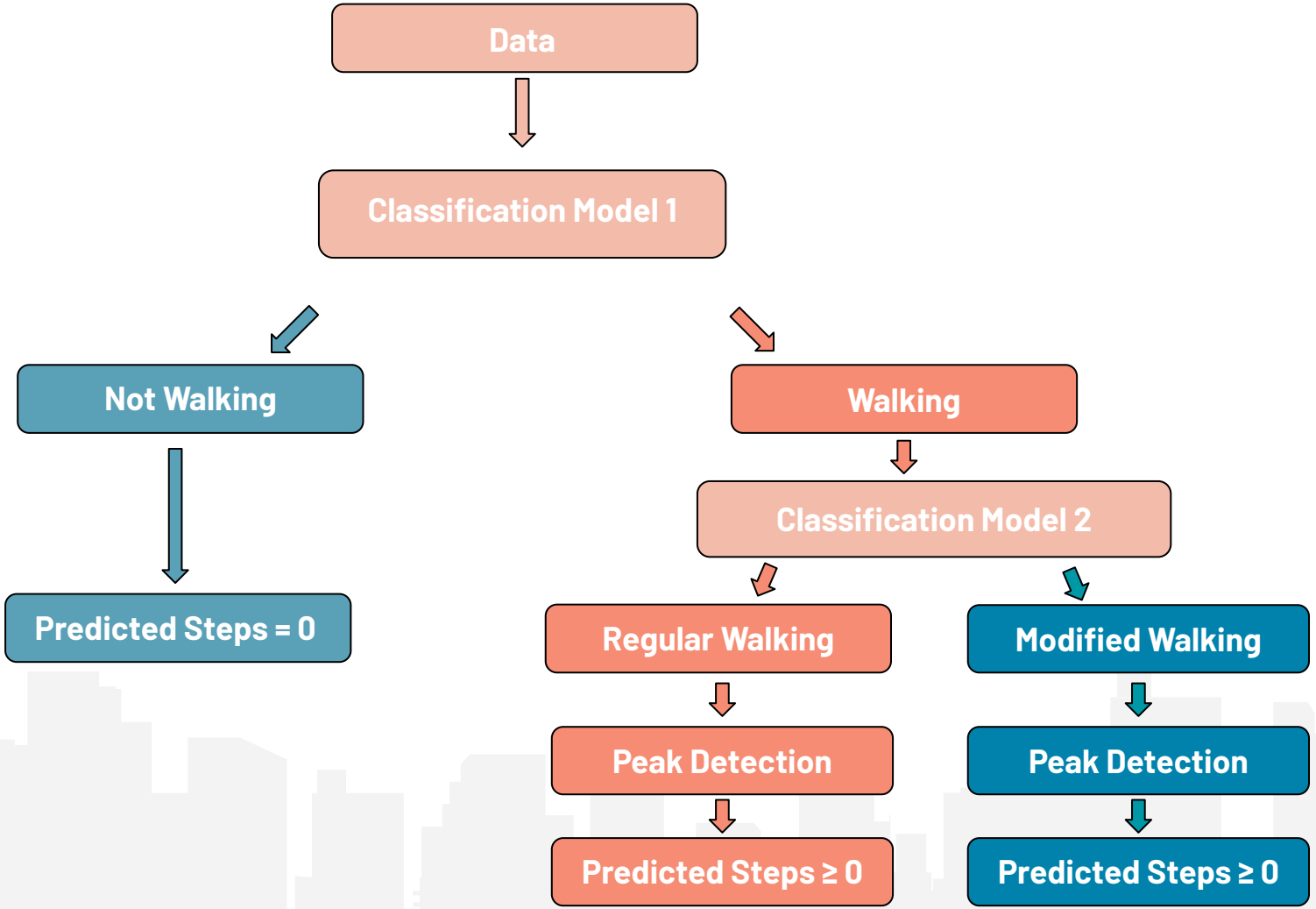
Walk/Not Walk

Condensed Postures

All Postures

ACT24 Walk/Not Walk Classification




Data
Classification Model 1
Not Walking
Predicted Steps = 0
Walking
Classification Model 2
Regular Walking
Peak Detection
Predicted Steps ≥ 0
Modified Walking
Peak Detection
Predicted Steps ≥ 0

ACT24 Walk/Not Walk Evaluation

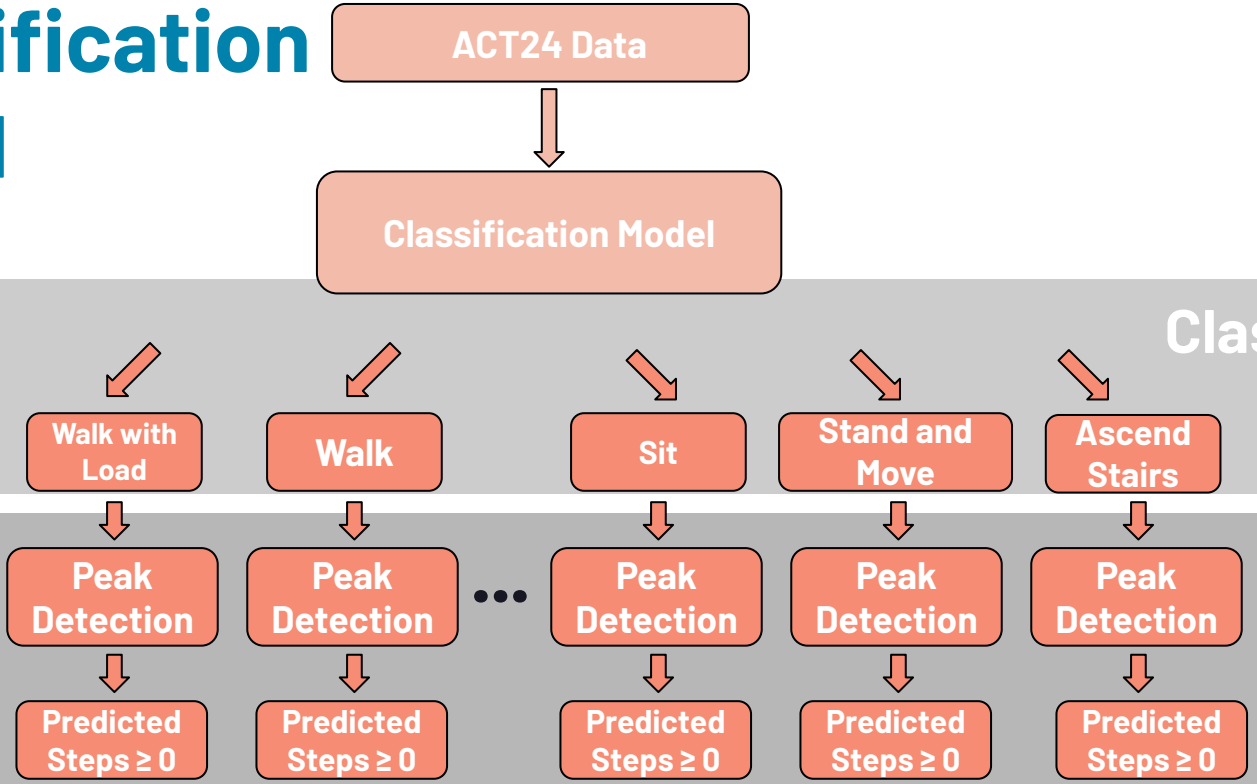
Pros

- Exposed to modified walking w/ ACT24 training data
- Binary classification problem (low classification error)
- Less sensitive to changes in dataset (low variance)

Cons

- Walking serves as a catchall category (high peak detection error)

All Posture Classification Model



All Posture Classification Evaluation

Pros	Cons
• Walking no longer a catch-all category • Fined tuned peak detection for each category (lower peak detection error)	• Large number of classification categories • Highly sensitive to changes in the dataset (high variance, overfitting, high classification error) • More computationally expensive

Condensed Postures Mapping

Walk → **Walk**

Stand, Sit, Stretch, Kneel/Squat, Lying → **No Movement**

Stand and Move → **Stand and Move**

Ascend Stairs → **Ascend Stairs**

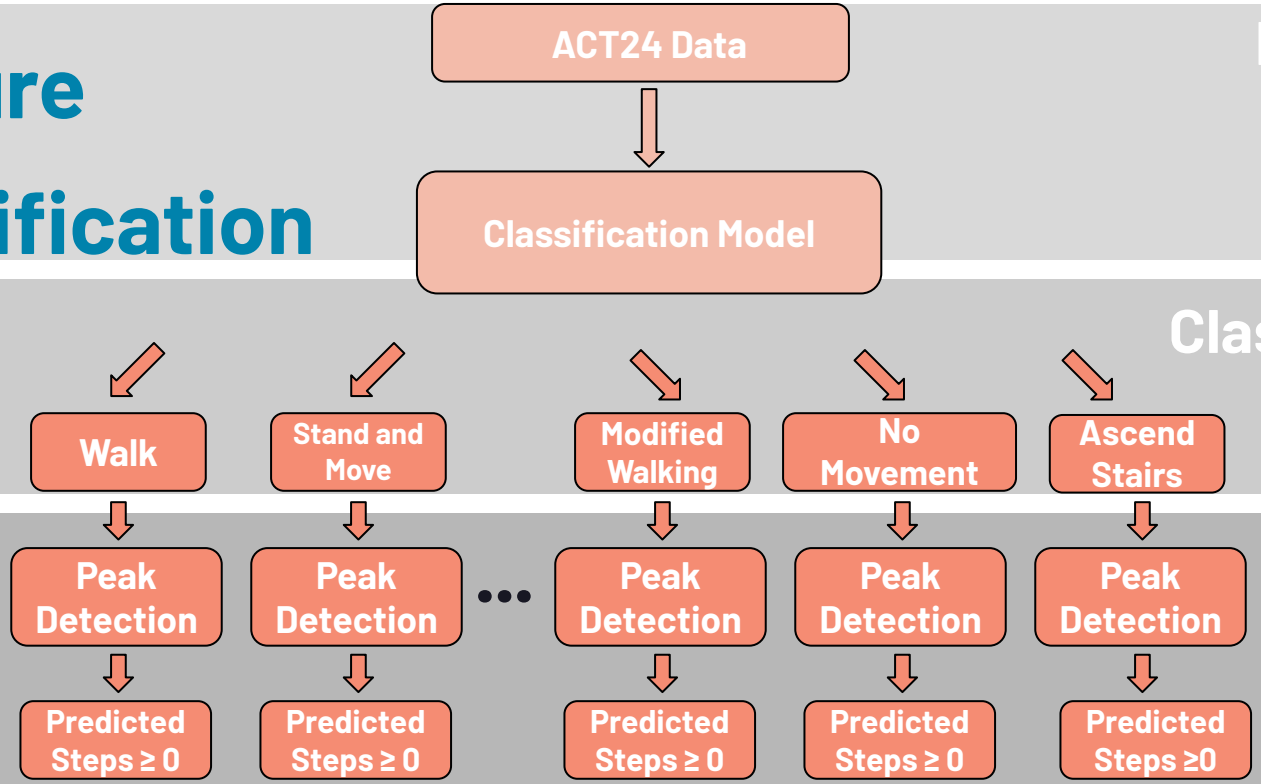
Descend Stairs, Walk with Load → **Modified Walking**

Bike → **Bike**

Muscle Strengthening → **Muscle Strengthening**

Other Sport Movement → **Other Sport Movement**

Condensed Posture Classification



Condensed Posture Classification Evaluation

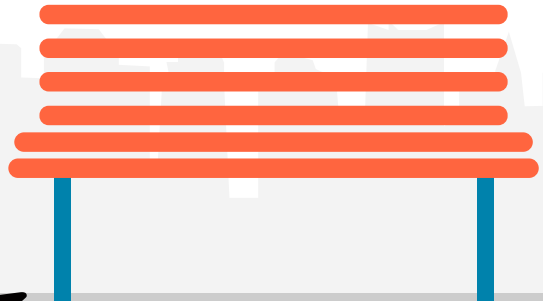
Pros

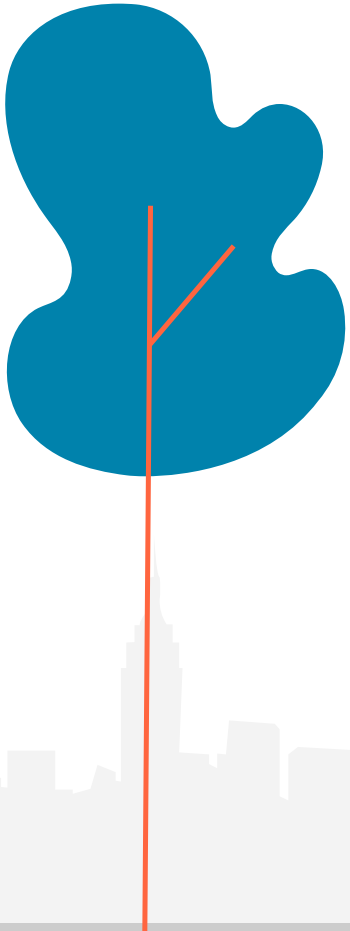
- Exposed to modified walking in training on ACT24
- Less classification categories
- Less computationally expensive than all postures

Cons

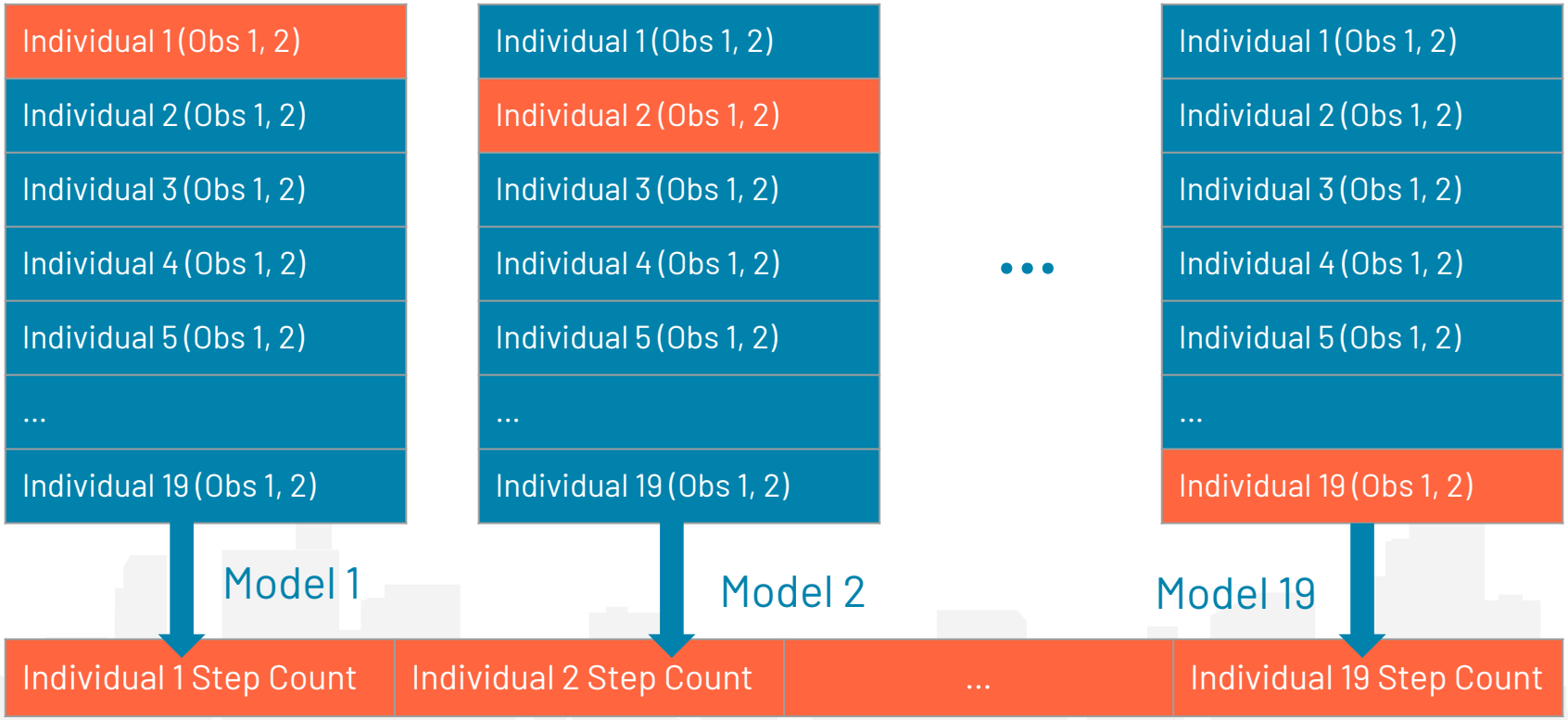
- Might still question overfitting to postures
- Could overlook certain unique posture signal tendencies

04 Model Testing





Leave-One-Out Cross Validation

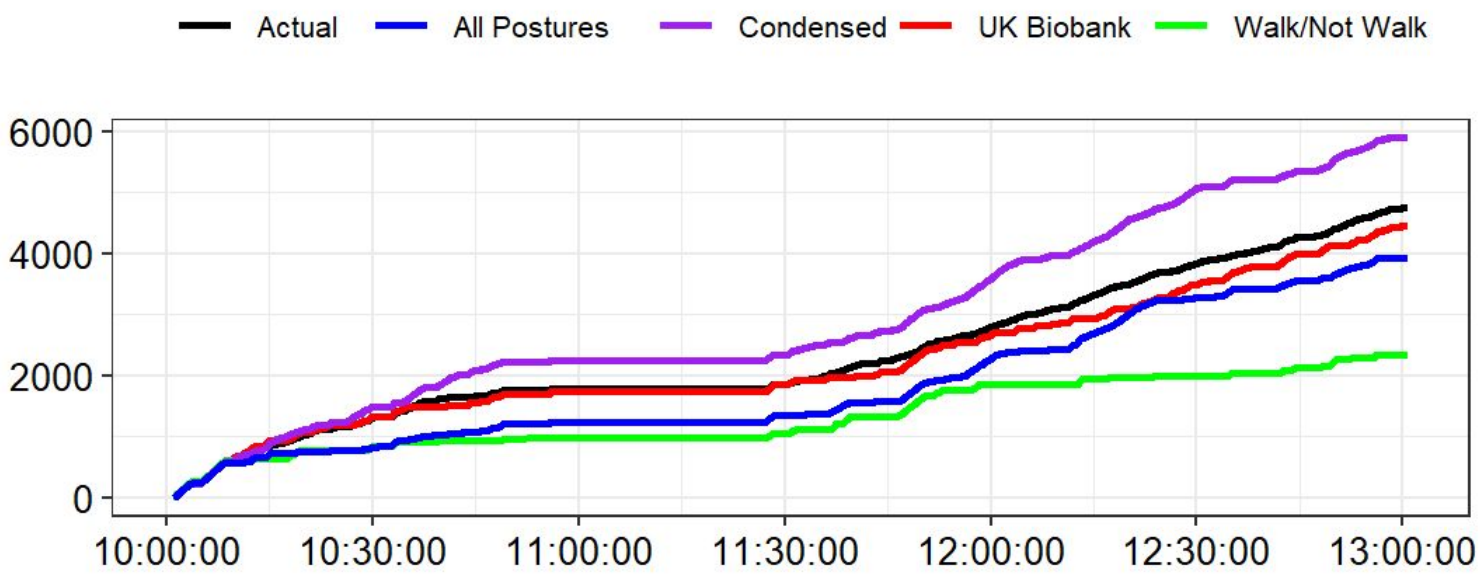


"External" Predicted Step Counts for Entire ACT24 Dataset

Cross-Validation Pros/Cons

- **Test** on data with ground truth that **includes postures**

- **Computationally expensive** (Fit 57 Neural Networks)
  - 9 hour runtime with GPU acceleration!
- Postures **not equally represented** in each individual
  - “Running” only has one 10-second epoch in one individual
  - Model changes heavily depending on which individual excluded


Actual
All Postures
Condensed
UK Biobank
Walk/Not Walk
6000
4000
2000
0
10:00:00
10:30:00
11:00:00
11:30:00
12:00:00
12:30:00
13:00:00

On average, the models are estimated to be off by

steps per minute overall

On average, the models are estimated to be off by

steps per minute in periods of modified walking

External Data Testing

- **Test** on **true external** UK Biobank accelerometer data (**OxWalk**)
- **No posture ground truth**

On average, the models are estimated to be off by...

Walk/Not Walk

3.810

All Postures

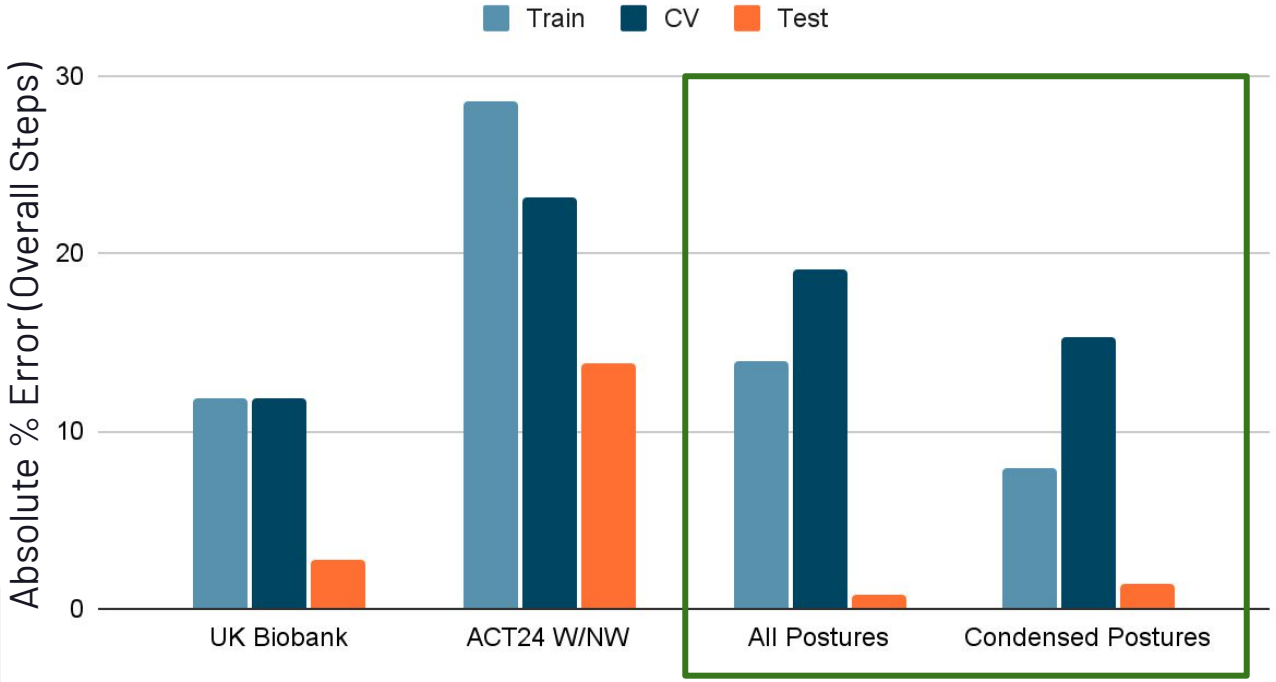
0.231

Condensed Postures

0.400

steps per minute overall

05 Results



In general, classifying on more postures results in noticeable improvement

Ranking Models (Test Data)

Model	1st	2nd	3rd
Condensed	18	12	5
All Postures	14	17	5
Walk / Not Walk	3	6	29

Results

- Classifying on **more** categories **improves** step counting outcomes
  - **Unclear** if there are specific improvements for **modified walking**

- We determined **preference** for **condensed** postures (not too few, not too many)
  - Metrics show **mixed** preference between all postures and condensed
  - Condensed is less computationally expensive
  - Less classification error

Final Thoughts

- ACT24 training data does not **represent** postures/activities equally
  - Having **more data** that represents more postures may **improve/change** model step counting algorithm outcomes

- **More combinations** of condensed postures can be explored

- **Other model specifications** have not been explored
  - Using **less** algorithm levels (e.g. one pass)
  - Using **more** algorithm levels (e.g. walk/not walk -> regular/modified walk -> peak detection)

Thanks!