Detection of Infectious Disease using Wearable Sensors

– Experience from UAS's American Life in Realtime

Shiyang Sima[1], **Ritika Chaturvedi**[2], **Hossein Ghasemkhani**[3], **Alok Chaturvedi**[1]

[1] *Mitchell E. Daniels, Jr. School of Business, Purdue University*
[2] *Leonard D. Schaeffer Center for Health Policy & Economics, University of Southern California*
[3] *Michael G. Foster School of Business, University of Washington*

Motivation

Pandemic damage

- Economic disruption, Healthcare System Strain, Social Consequences

Who Might Be Infected

- Silent Spreaders, Clusters

Motivation for an Early Warning System

- Early Intervention, Resource Allocation, Public Awareness

Background

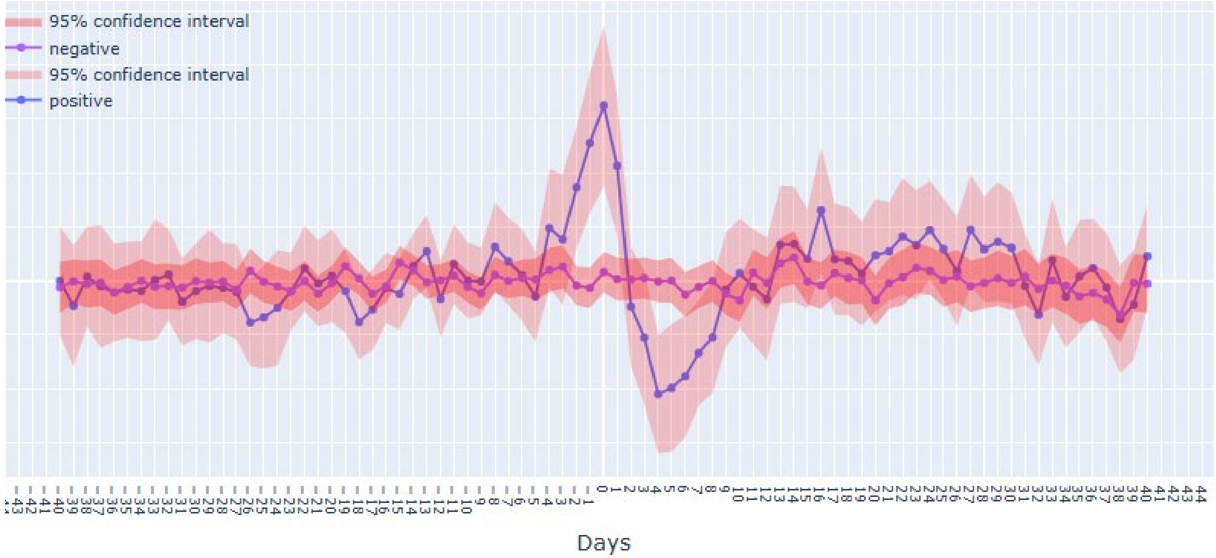
Probability Sampling: Ensures a representative cross-section of the population for accurate insights.

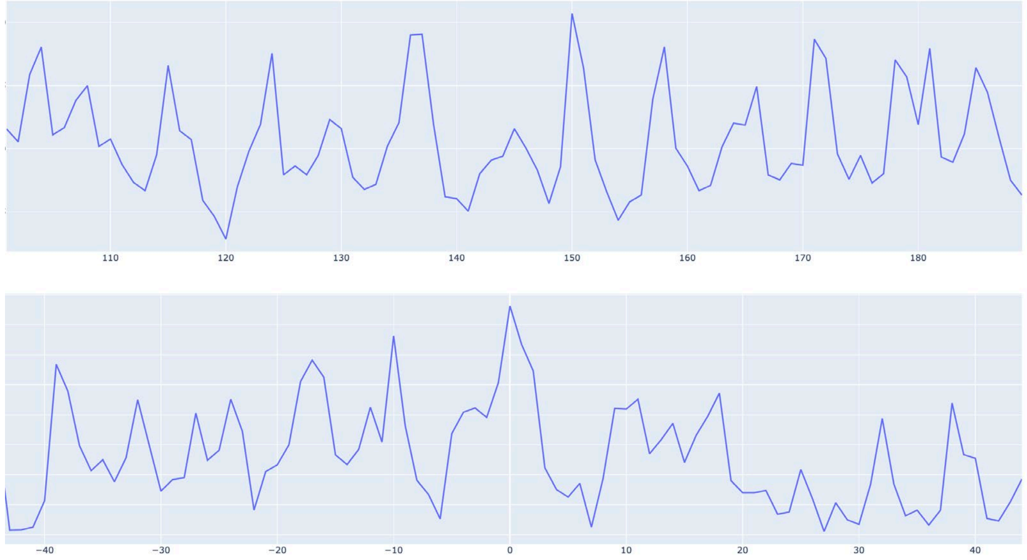
Passive Sensing: Collects data autonomously, without requiring active responses from individuals.

Unsupervised Learning: Enables passive prediction without patient input, allowing the model to adapt dynamically in rapidly changing situations.

Transformer Model: Utilizes temporal features and attention mechanisms to analyze both short-term and long-term trends in data.

Background



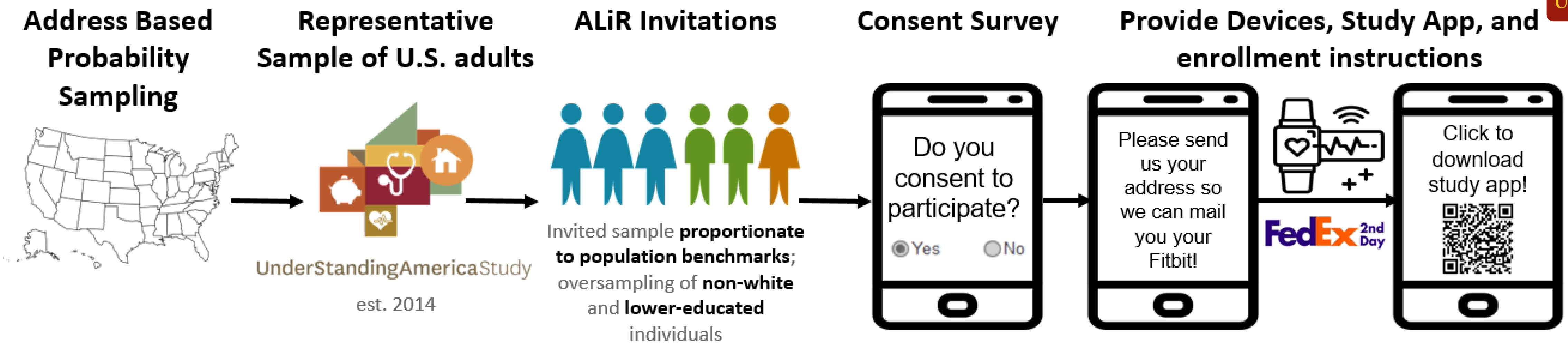


- While the signal is clear at an aggregated level and well-documented in many published studies (Radin et al., 2021), individual data shows significant variability. Our goal is to detect these signals on a per-person basis.

Research objectives

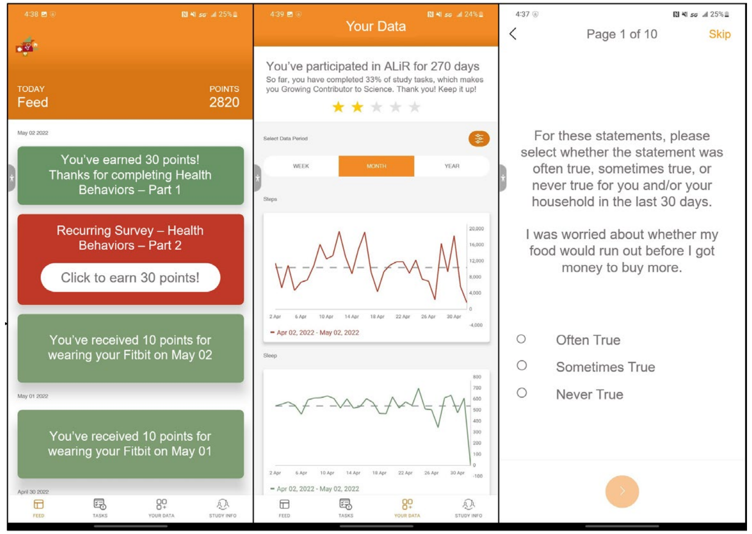
Developing an Individualized Early Warning System for Infectious Diseases

- **Data**: Using COVID-19 as an example.
- **Approach**: Build an unsupervised early warning system tailored to individuals.
- **Attention Mechanism**: Design to detect anomalies in Fitbit sensor data.
- **Verification**: Cross-check warning time points with reported diagnosis dates and symptom onset.
- **Operationalization**: Implement for real-time warning and detection.



American Life in Realtime (ALiR)

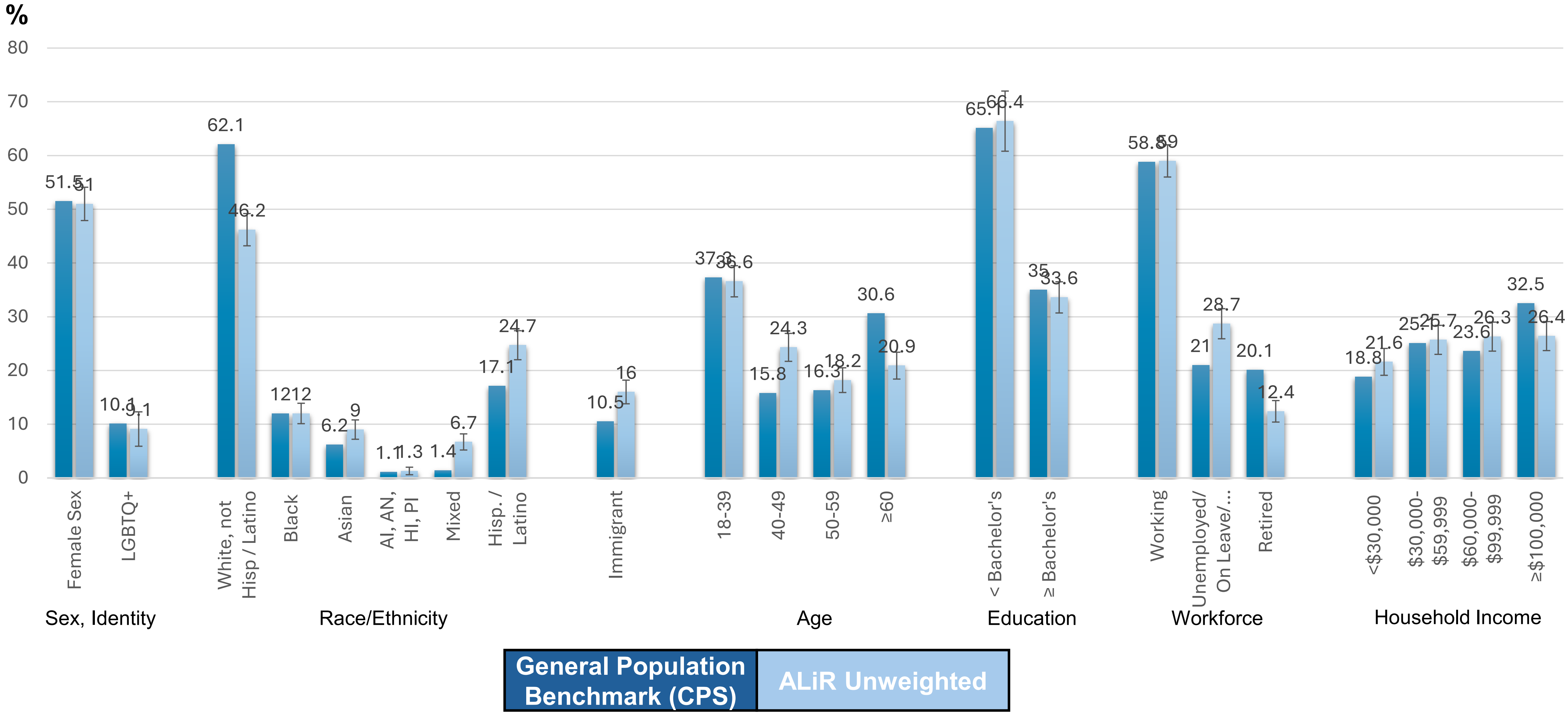
- Probability sample
- Sensor data from Fitbit
- Monthly surveys



Data

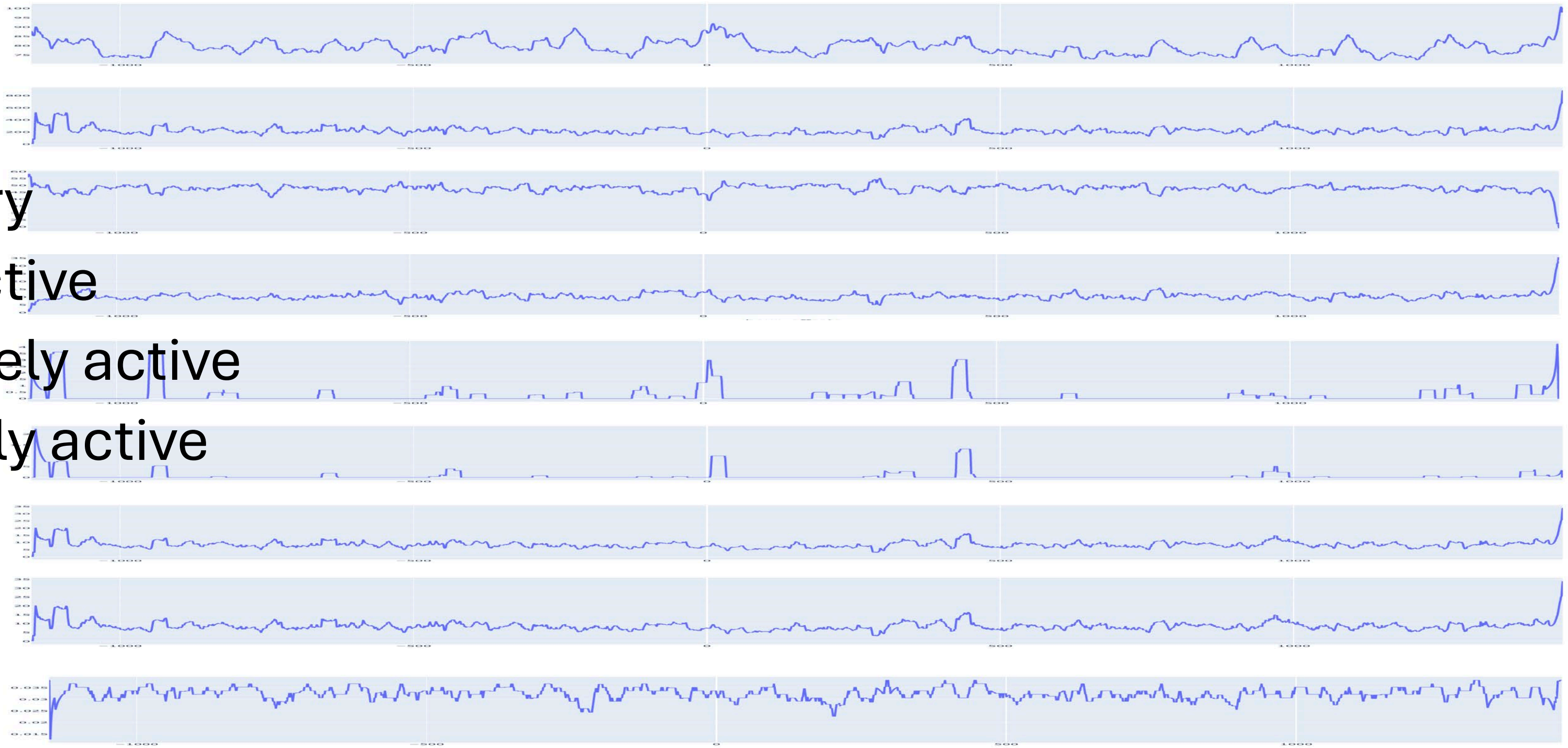
NIH funded projects: UAS (U01AG077280) and ALiR (R01LM013237)

Probability sampling



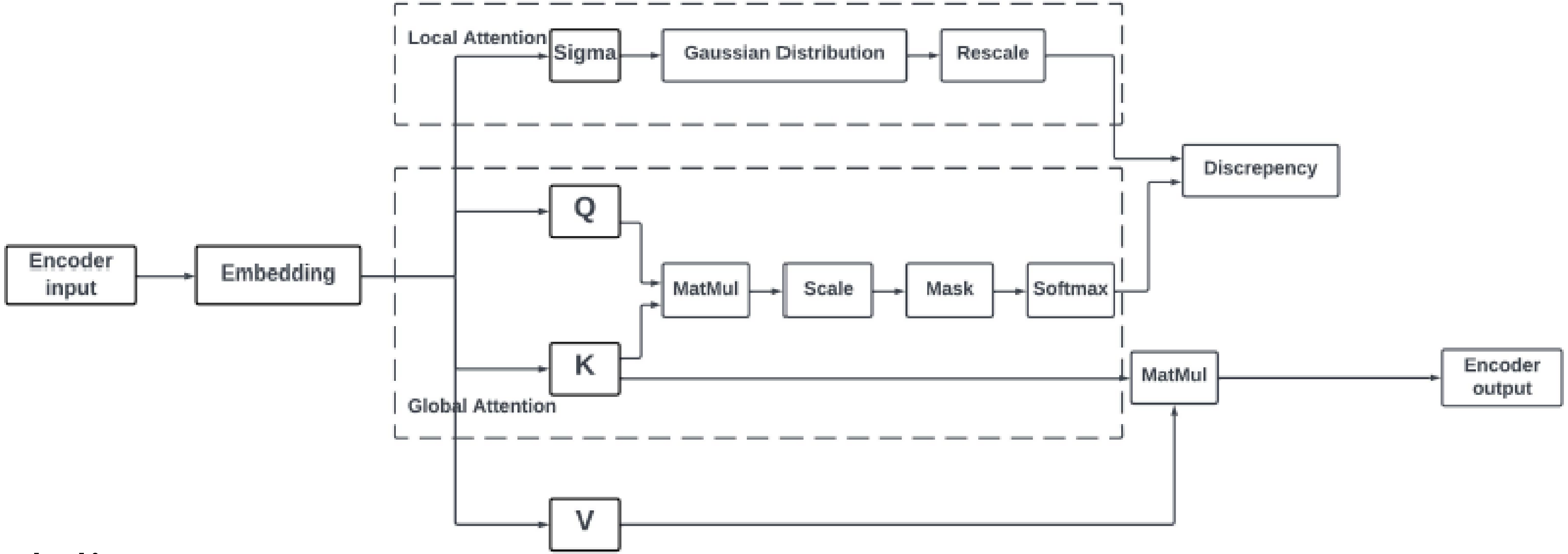
Data

- Fitbit features at hourly level
  - Heart rate
  - Step count
  - Minutes in sedentary
  - Minutes in lightly active
  - Minutes in moderately active
  - Minutes in vigorously active
  - Walking distance
  - Walking speed
  - Walking step length



Method

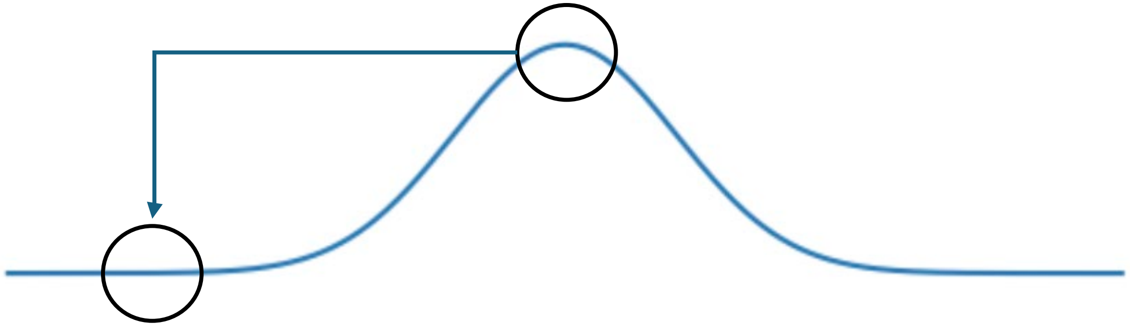
- **Hourly-level Fitbit data**: Apply a 24-hour rolling average window to fill in missing values when users remove devices for charging.
- **During covid:**
  - **Training Period (-90, -31)**: This period is used to establish the baseline because it is very rare to observe two COVID-19 cases within three months.
  - **Testing Period (-30, +30)**
- **After Recovery** :
  - **Training Period (+90, +149)**: individual has developed new health baseline in post-recovery days
  - **Testing Period (+150, +210)**



Global-local discrepancy (Hodge and Austin, 2004)

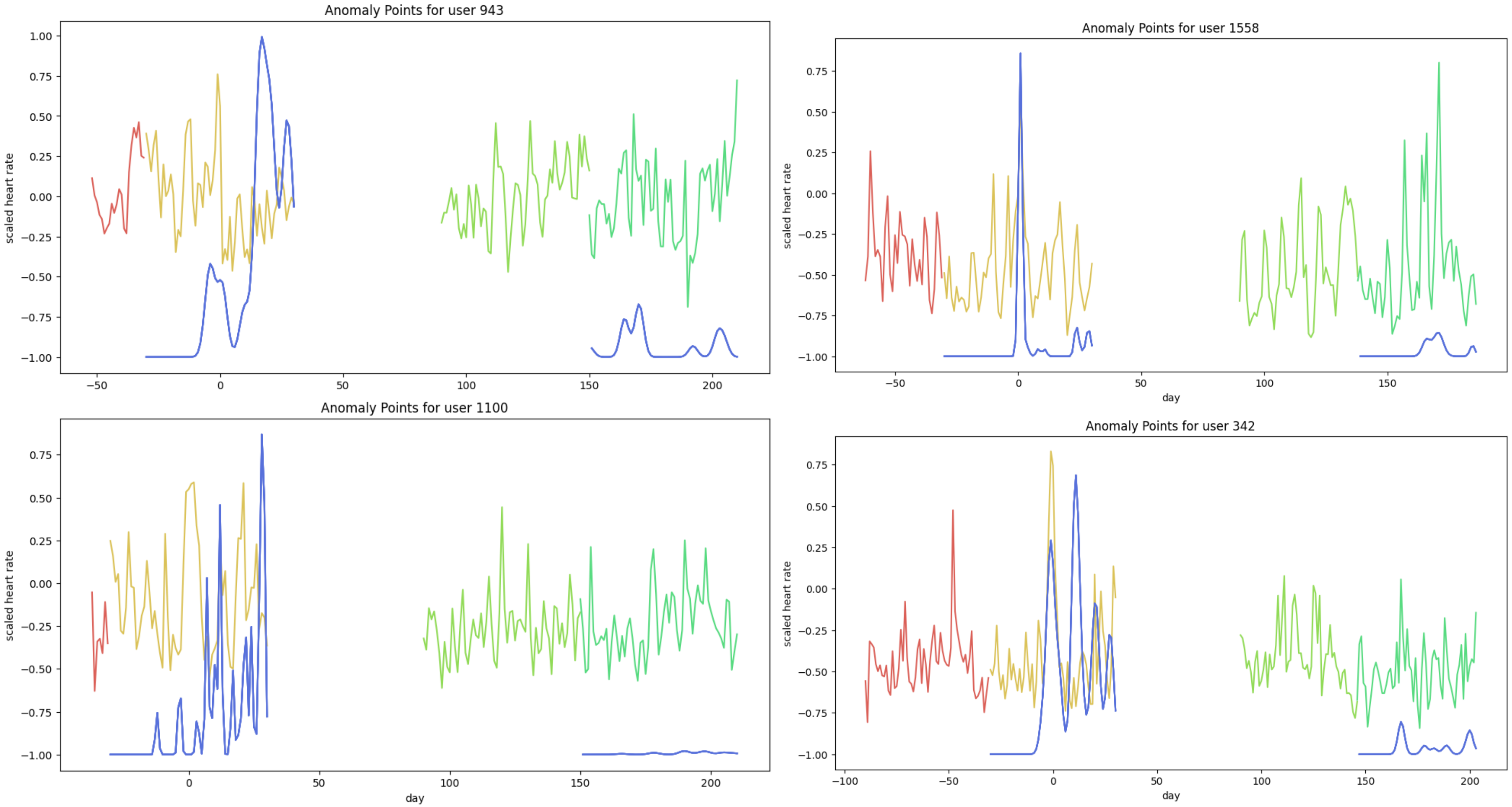
Global attention: full self attention

Local attention: Gaussian distribution



Model structure

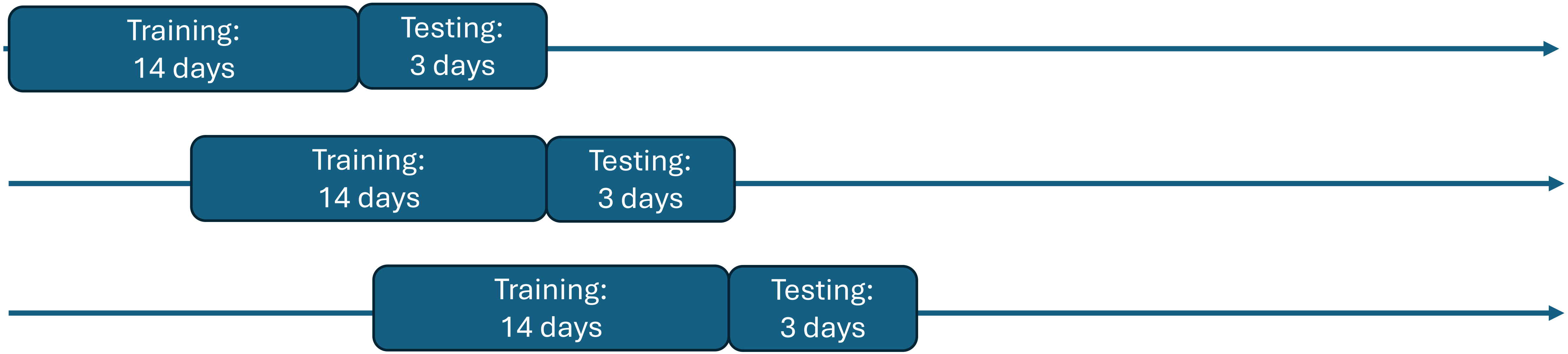
Results

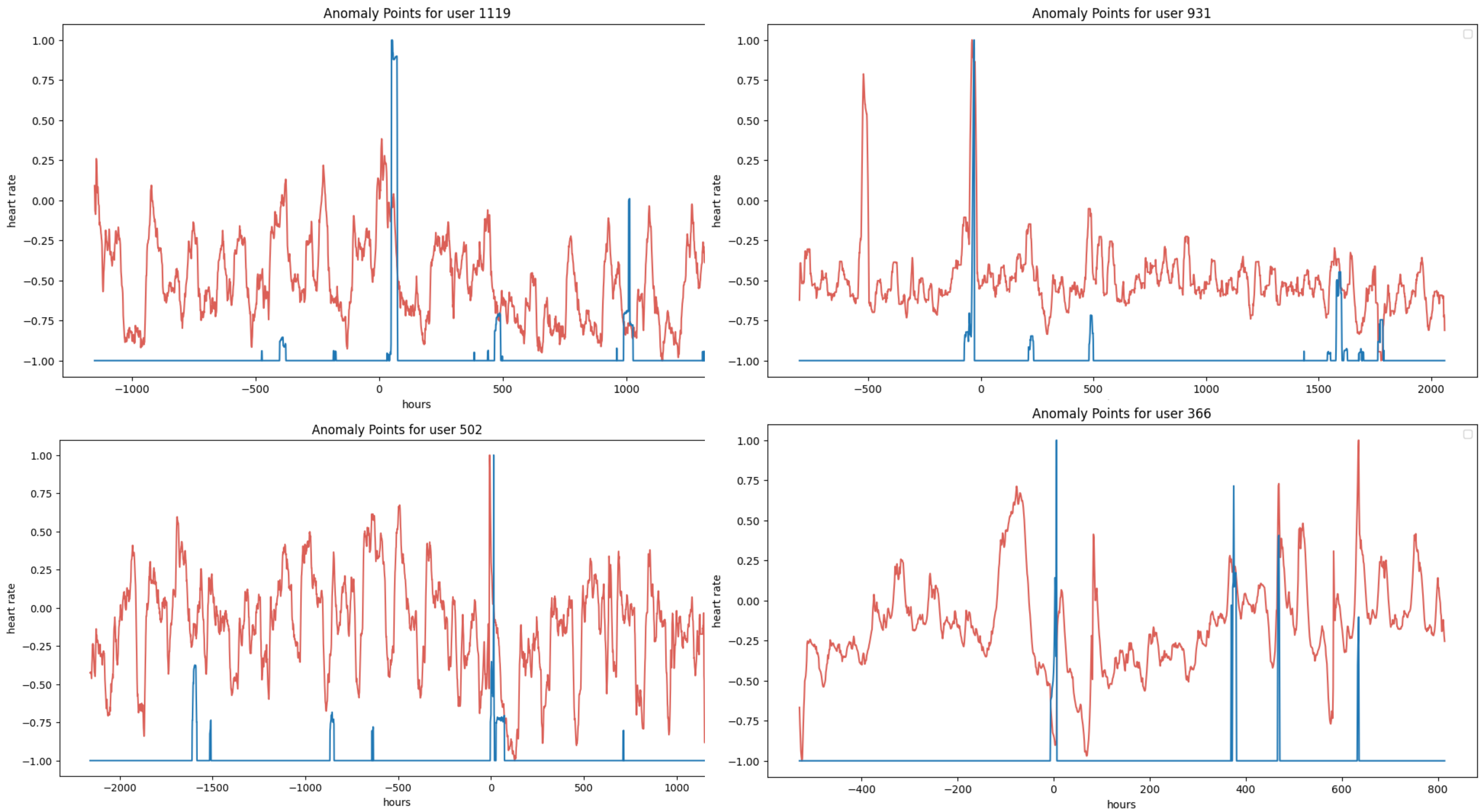


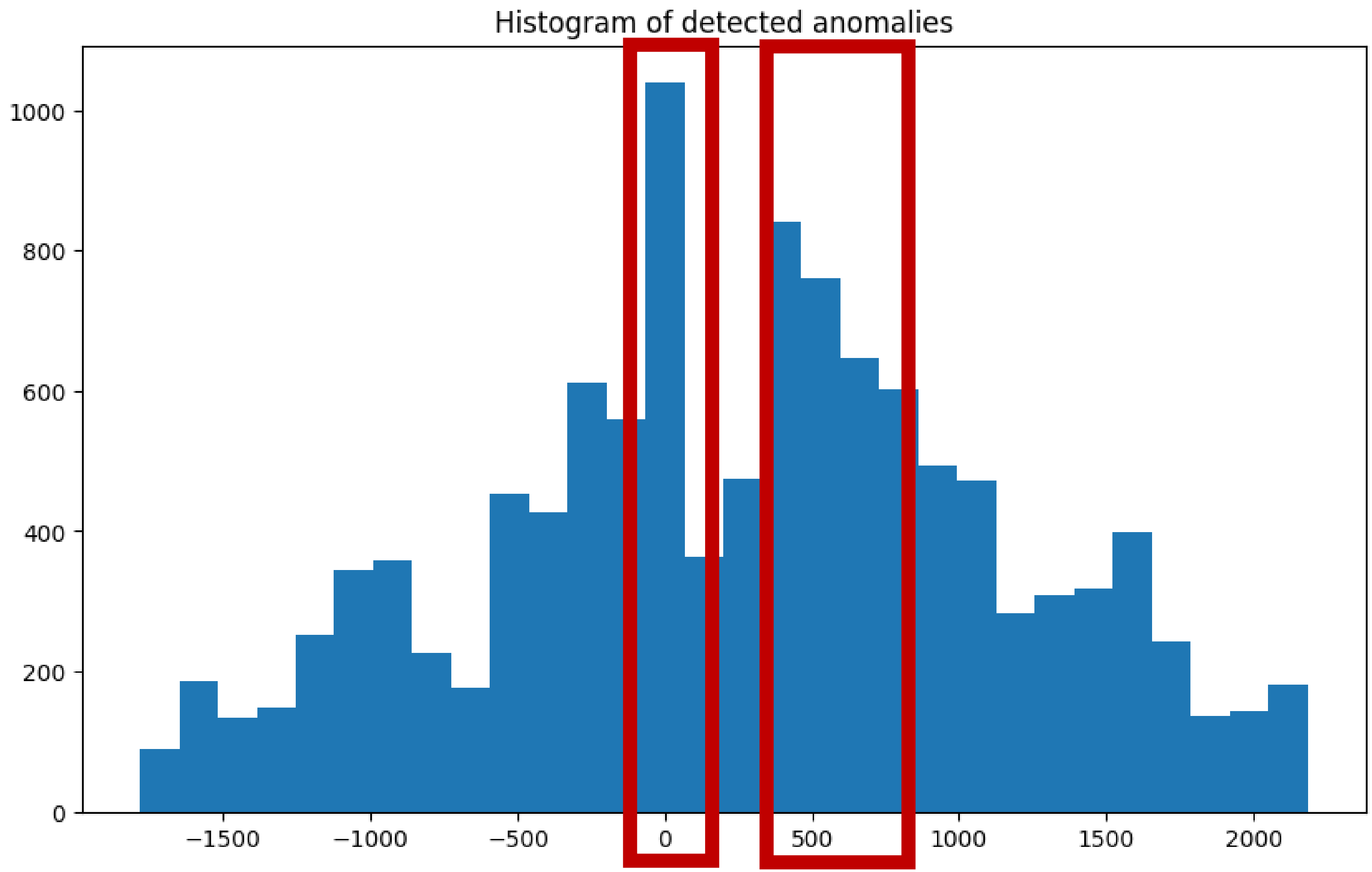
- Detected symptom onset occurred days earlier than the diagnosis date, emphasizing the potential for early warning
- significant differences observed between positive and negative periods, highlighting the system's ability to accurately differentiate and identify critical patterns.

Simulating Online detection

- Although we don't have permission from Fitbit to stream their data in real-time, we simulate an online streaming environment by using a sliding train/test window.




Anomaly Points for user 1119
Anomaly Points for user 931
Anomaly Points for user 502
Anomaly Points for user 366
heart rate
hours
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
−1000
−500
0
500
1000
1500
2000
−2000
−1500
−400
−200
200
400
600
800



Histogram of detected anomalies

Discussion and Future Directions

- A transformer-based unsupervised model can be utilized as an early warning system, demonstrated through the COVID-19 example.
- Pattern recognition – In addition to individual sensor data, the model incorporates known disease patterns derived from aggregated sample data as inputs.
- Self-improving – The system leverages explainable AI and unsupervised metrics to enhance its performance autonomously.

References:

- **Iglewicz, B., & Hoaglin, D. C.** (1993). *A Survey of Outlier Detection Methodologies*. Journal of Applied Statistics, 20(1), 1-23.
- **Radin, J. et al.** (2021). *Assessment of Prolonged Physiological and Behavioral Changes Associated with COVID-19 Infection*. JAMA Network Open, 4(7), e2113615. https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2781687
- **Radin, J. et al.** (2021). *Wearable Sensor Data and Self-Reported Symptoms for COVID-19 Detection*. Medscape. https://www.medscape.org/viewarticle/955844