Efficient Sleep State Tracking

Leveraging Data Science for Health

Problem

- Sleep is vital for human health.
- Accurate sleep monitoring is challenging due to a lack of naturalistic data and accurate annotation (Esper et al., 2023).
- Data science can enhance sleep research and consumer health tracking.

Recommendations

1. Utilize the model's predictions to automatically annotate sleep onset and wake-up times.

2. Incorporate these annotations into sleep research to study its overall health effects.

3. Apply the annotations in a business context to help individuals understand their sleep patterns' impact on health.

Proposed Solution

- Deep learning model for predicting "onset" and "wakeup" events
- Frame as segmentation problem
- Use encoder decoder framework
- Confidence scores (0-1) for predicted events
- Inputs: data from wrist-worn accelerometer
- Labels: (asleep/awake, onset, wakeup)
- Metric: Average precision for event detection

Deliverables

Develop a predictive model for sleep onset and wakeup times.

Create a comprehensive report and presentation.

Stakeholders

Researchers studying sleep.

Companies interested in using accelerometer data for health-related sleep tracking.

Data

The dataset contains around 500 multi-day wrist-worn accelerometer recordings annotated with "onset" and "wakeup" events. Data may have gaps where the accelerometer was removed.

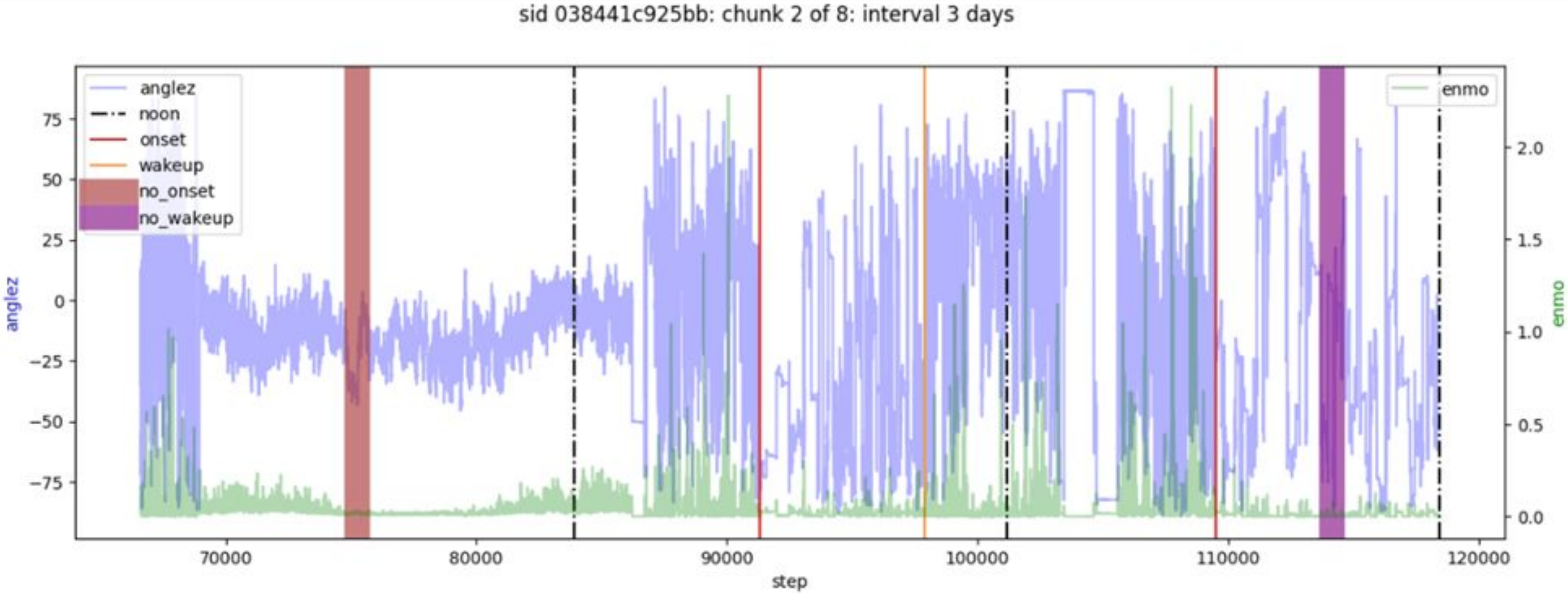
Files and Field Descriptions

- train_series.parquet: Contains continuous accelerometer data with unique series IDs.
- train_events.csv: Sleep logs with series IDs, night enumeration, event types, and timestamps.

Exploratory Data Analysis

- No missing values for enmo and anglez columns.
- Some nights lack "onset" or "wakeup" annotations, likely due to accelerometer removal.
- A few nights have only one event annotated.

During sleep, **enmo** and **anglez** values exhibit distinct patterns.

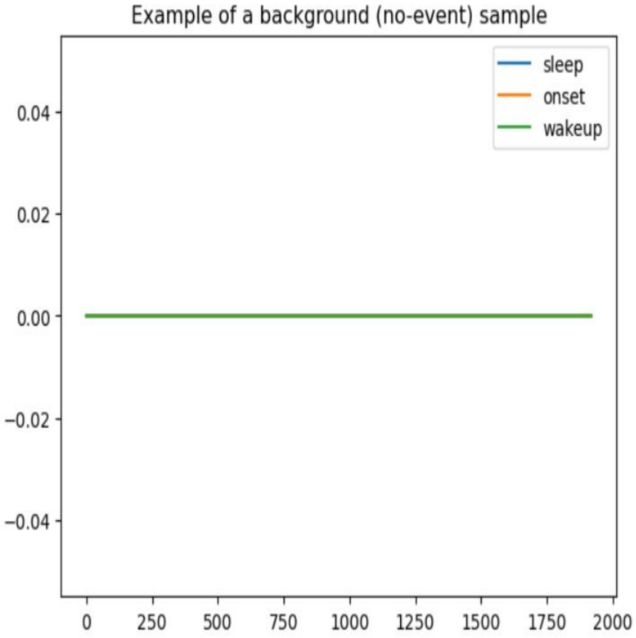


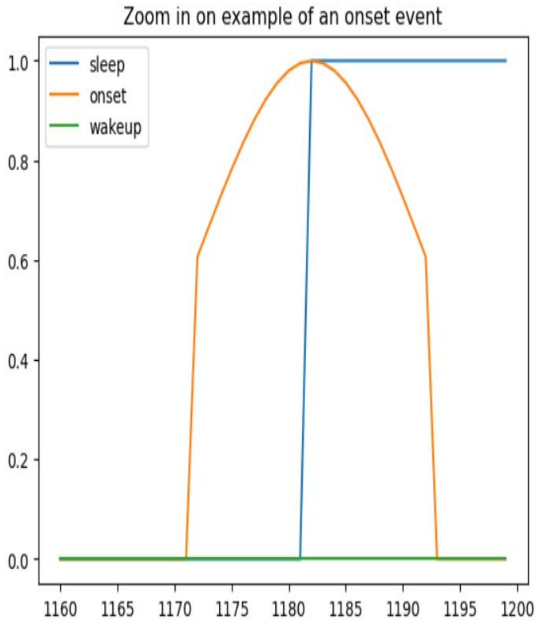
Preprocessing

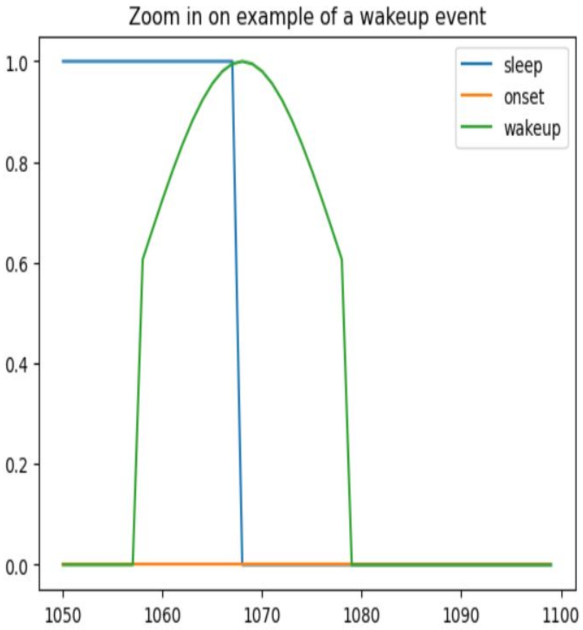
Features: shape, sine, cosine, differences, rolling medians.

Labels: shape, values, Gaussian labels.

Labels


Example of a background (no-event) sample
sleep
onset
wakeup
0.04
0.02
0.00
−0.02
−0.04
0
250
500
750
1000
1250
1500
1750
2000


Zoom in on example of an onset event
sleep
onset
wakeup
1.0
0.8
0.6
0.4
0.2
0.0
1160
1165
1170
1175
1180
1185
1190
1195
1200


Zoom in on example of a wakeup event
sleep
onset
wakeup
1.0
0.8
0.6
0.4
0.2
0.0
1050
1060
1070
1080
1090
1100

Final Model

- Segmentation model with encoder and decoder
- Encoder: LSTM feature extractor
- Decoder: UNet decoder
- Post processing: Select peak predictions for onset and wakeup events.

Validation

- 20% validation set (valid_set 1).
- Kaggle public leaderboard (valid_set 2).
- Kaggle private leaderboard (final test set).

Notable Experiments

Model	Brief Description	Valid_1 Score
Baseline	Config defaults	0.74
v1_ds3	Downsample_rate 2 to 3	0.7546
v2_ds3	Added rolling median features	0.7565
v2_ds3_fe_LSTMFeatureExtractor	Chosen Model	0.7598

Manual Post-processing

- Lower threshold (threshold == 0.005, valid_1 score: 0.765).
- Filter_by_min_max_th (Remove predictions if min(max of onset and wakeup score) < 0.03, valid_1 score: 0.767).
- Filter_(onset of wakeup)_threshold (Keep max prediction if above 0.82, valid_1 score: 0.768).
- Filter_max_score_by_night (Eliminate predictions if max is not above 0.03, valid_1 score: 0.769).
- Inflate_max_wakeup (Inflate max wakeup score by a multiplier of 3.2, valid_1 score: 0.771).
- Inflate_max_onset (Inflate max onset score by a multiplier of 13.4, valid_1 score: 0.774).

Final test score: .758

Sample Final Predictions