Single-Device Multimodal Solution for Predicting Sleep Efficiency and Anticipating Next-Day's Mood

Presented By

Jamalia Sultana
SBU ID: 115834976

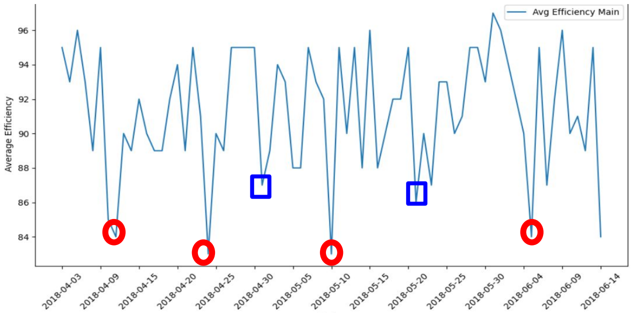
Adiba Shaira
SBU ID: 115936265

Problem Statement

$$Sleep\ Efficiency = \frac{TotalSleepTime}{TotalMinutesinBed}$$

- We aim to predict a user's sleep efficiency using their activity, location, and screen time collected over the time period.
- Detecting the lower peaks earlier so that we can make people aware.

Avg Sleeping Efficiency 91.27



Temporal Analysis of “x” subject

Motivation

- Sleep efficiency is a critical indicator of overall health, affecting cognitive performance, mood, and physical well-being.
- Early detection of inefficiencies helps adapt routines, environment, and behaviors for better outcomes.
- Leveraging activity, location, and screen time data enables predictive analytics and pattern recognition.

Related Work

Sleep Quality Prediction from wearable data [1]

- Predict sleep quality based solely on physical activity data from actigraphy
- 69% accuracy on good sleep(>90%) prediction

Sleep Trackers

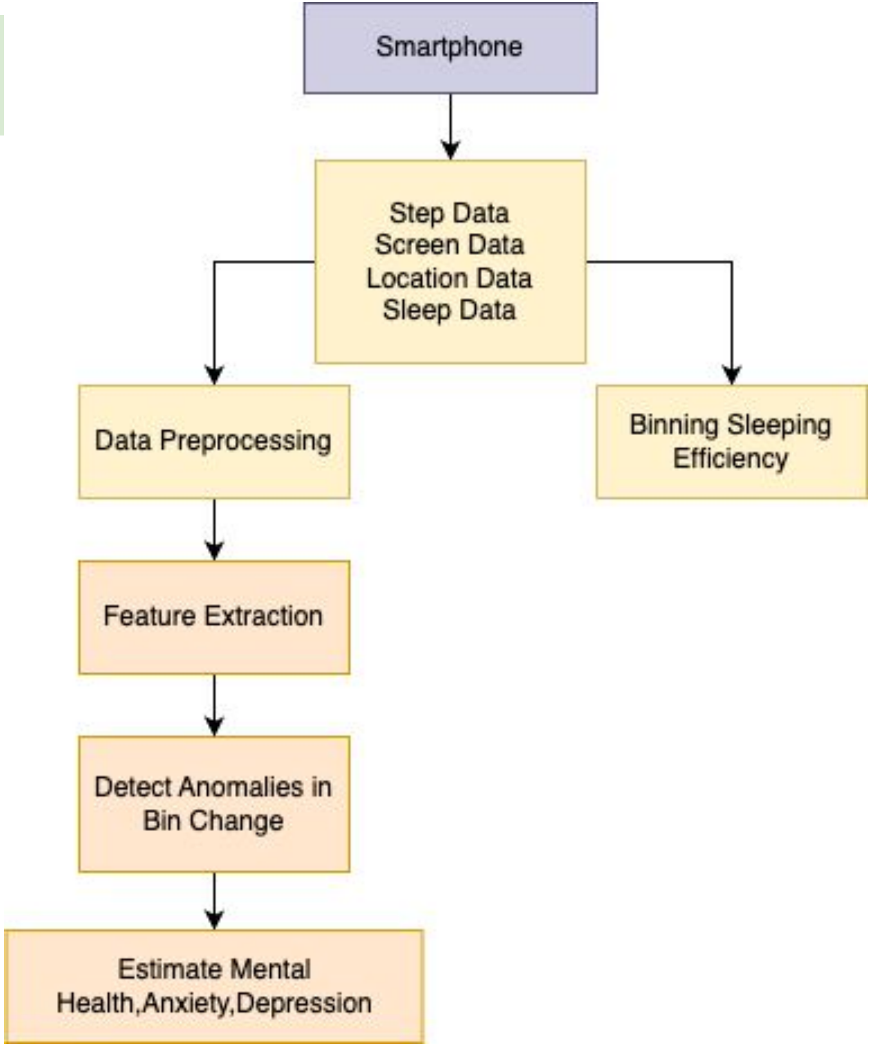
- Sleep trackers monitor movement and heart rate, analyze sleep patterns, and provide insights on sleep duration and efficiency.
- Single device solution

Bidirectional Associations of Sleep with Activities [2][3]

- Correlations between sleep quality and daily activity levels
- Observational analysis rather than predictive modeling.

System Overview

1. Single device (smartphone)-collected data
2. Preprocessing and feature extraction to analyze patterns
3. Detect anomalies (bad sleep) using sleep efficiency bins
4. Anomalies are linked to potential mental health indicators
5. Aims to estimate mental health status by predicting the sleeping efficiency



Data Preprocessing

Globem Dataset

INS_W1(2018)
INS_W2(2019)
INS_W3(2020)
INS_W4(2021)

- Took 50 participants from each year
- Drop the rows having more than 90% null values
- Impute the other null values with median value
- Kept the features based on day

Dataset	Avg_Sleeping_ Efficiency < 85	Avg_Sleeping_ Efficiency >= 85
INS_W1	8	42
INS_W2	7	43
INS_W3	6	44
INS_W4	4	46

Sleep Efficiency Binning and Anomaly Detection

Binning Sleep Efficiency

- Sleep efficiency is divided into **bins** based on predefined ranges: `[45-70, 70-85, 85-90, 90-100]`.
- Labels `[0, 1, 2, 3]` are assigned to these bins, with higher labels indicating better sleep efficiency.

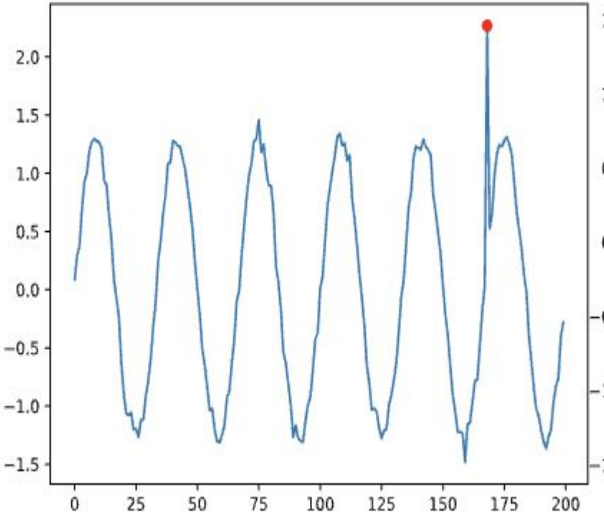
Threshold for Good Sleep

- Sleep efficiency above **85%** is considered **good sleep**.
- The bins are designed with **85% as the midpoint**, reflecting the threshold for optimal sleep quality.

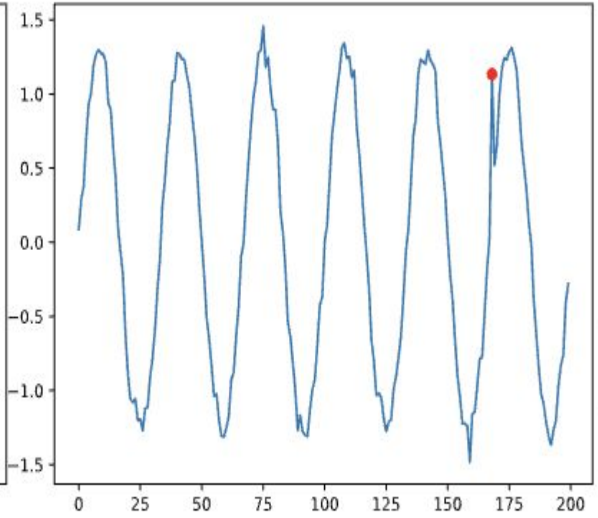
Anomaly Detection

- An **anomaly column** is generated to capture changes in sleep patterns.
- If the current bin's label is **lower** than the previous day's, it is marked as an **anomaly** (`1`), indicating a decline in sleep efficiency.
- The percentage of anomaly in our data is 0.4%

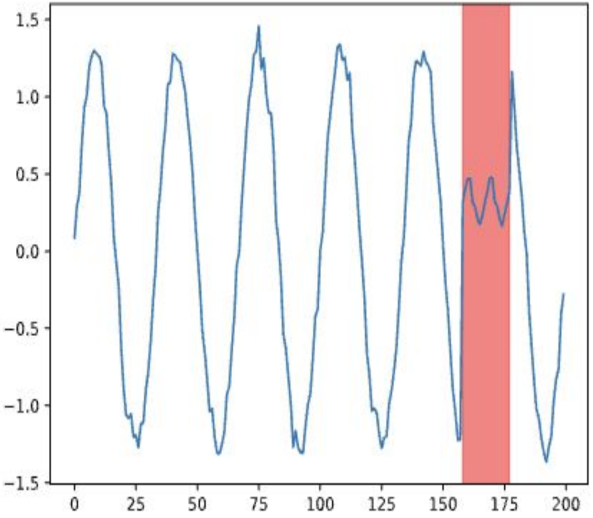
Anomaly Type in Time Series



Global Point Outliers

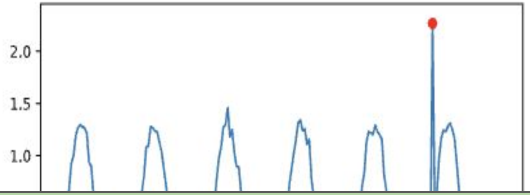


Contextual Point Outliers



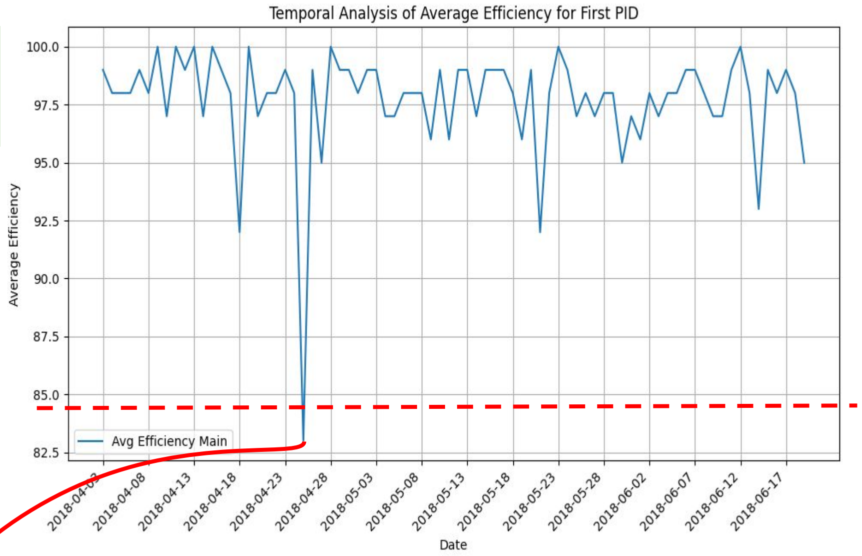
Collective/Pattern Outliers

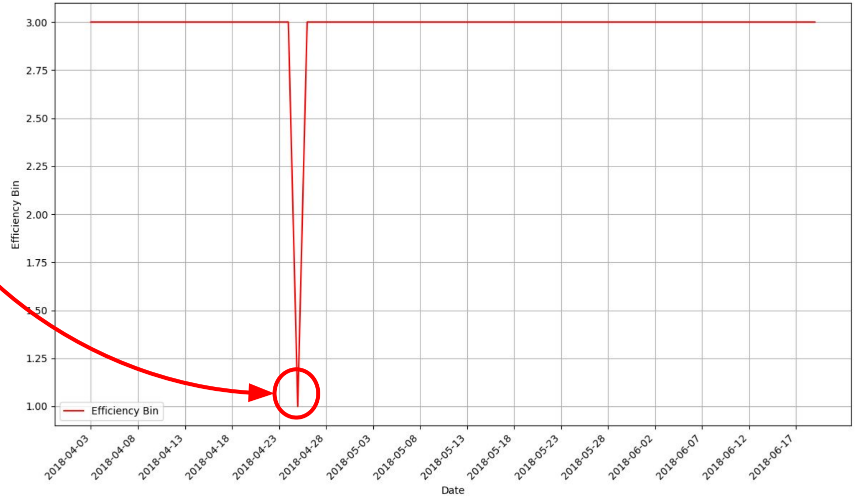
Anomaly in Sleeping Efficiency



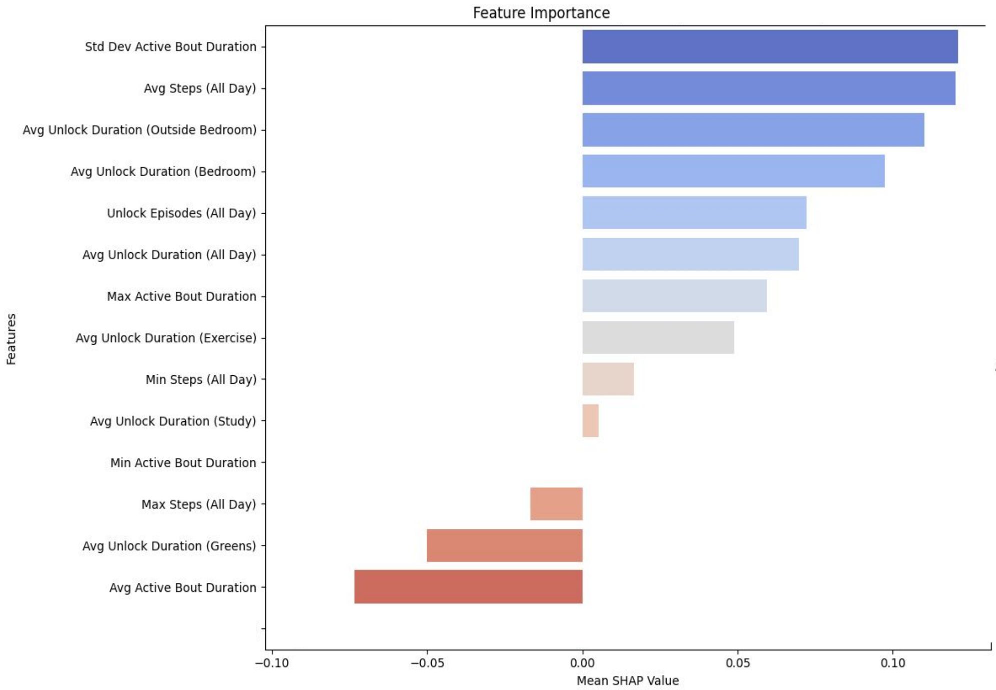
We want to predict this bad sleep before it happens.

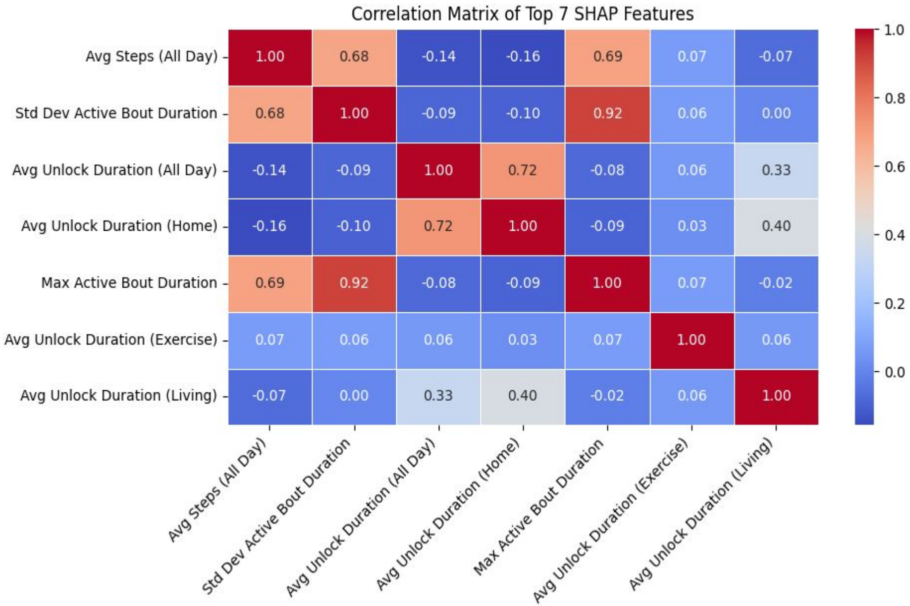
Global Point Outliers





Feature Analysis





In the barplot, features pushing the prediction higher are shown in red, those pushing the prediction lower are in blue.

Feature Analysis

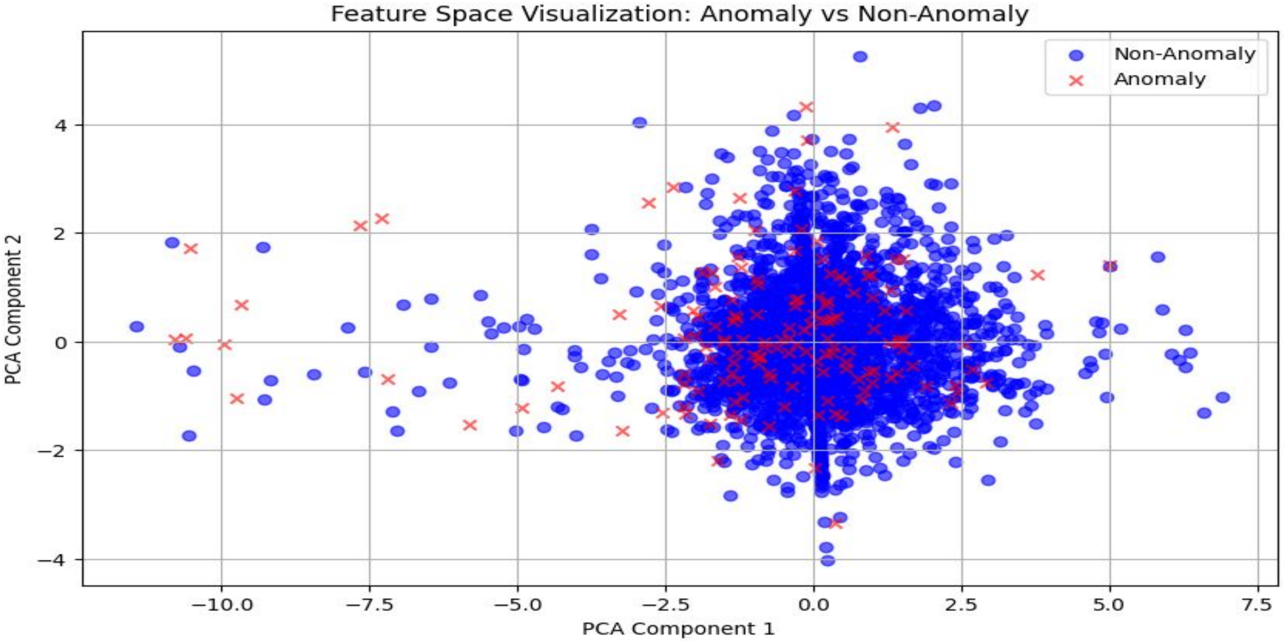
Null Hypothesis (H_0): There is **no significant difference** between the deviations of features on days when sleep efficiency score drops by more than 2 . In other words, the deviations in the selected features are the same whether or not there is a significant drop in sleep efficiency.

Alternative Hypothesis (H_1): There is a **significant difference** between the deviations of features on days when sleep efficiency drops by more than 2% and on normal days. This would imply that the behavior of the features is different on days with a sleep efficiency drop compared to normal days.

P-value: 2.15e-22
There is a significant difference in deviations for the top 7 features during the days sleep efficiency drops.

Feature Space Visualization

1. PCA is used for generating the feature space
2. No significant clustering for anomaly vs non-anomaly



Final Feature List

Selected Features

1. **Activity & Screen Time:**
   - `f_steps:fitbit_steps_summary_rapids_minsumsteps:allday` (Minimum daily steps)
   - `f_screen:phone_screen_rapids_firstuseafter00unlock_locmap_greens:allday` (First unlock after the day starts)
2. **Sleep Features:**
   - `f_slp:fitbit_sleep_summary_rapids_avgefficiencymain:allday` (Sleep efficiency)
   - `f_slp:fitbit_sleep_summary_rapids_firstbedtimemain:allday` (First bedtime)
   - `f_slp:fitbit_sleep_summary_rapids_lastbedtimemain:allday` (Last bedtime)

Target day Features: Activity, screen time

Derived Features: Previous Day Sleep features

Performance Metrics

1. Precision
2. Recall
3. F1-score
4. Balanced accuracy

Models

1. Unsupervised:
    a. One Class SVM
    b. Isolation Forest
2. Supervised:
    a. Random Forest
3. Ensemble:
    a. Majority Voting
    b. Soft Voting
4. Machine Learning:
    a. LSTM
    b. Auto Encoder

Benchmark Comparison

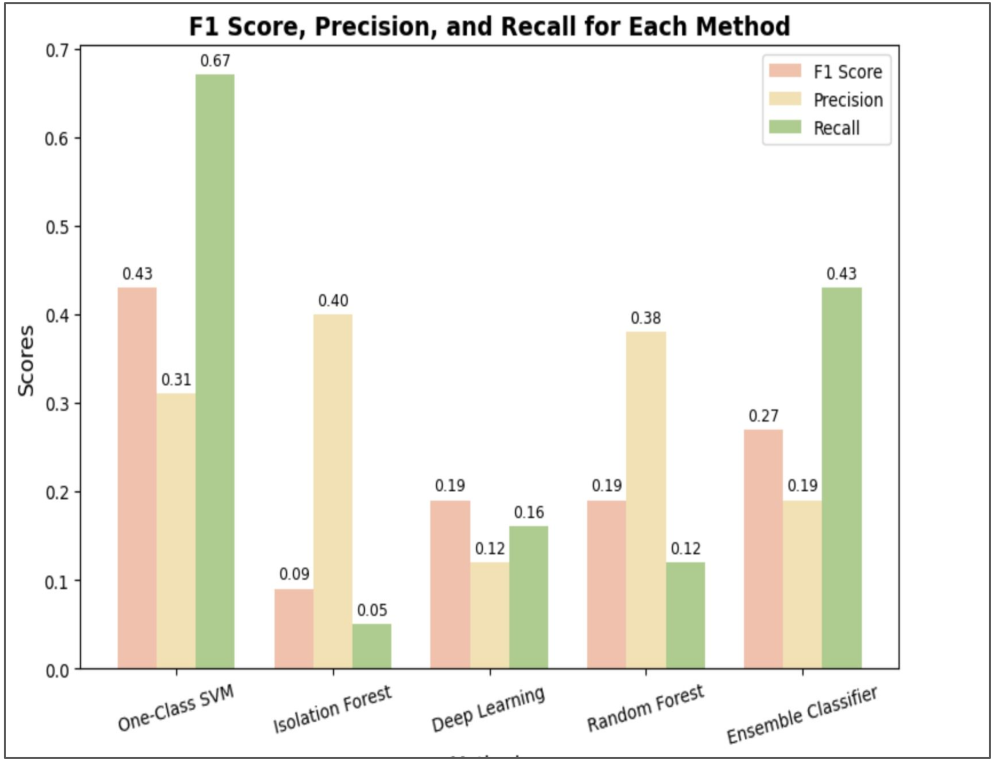
1. **OCSVM** achieved the best performance, indicating its strong ability to correctly identify anomalies and balance between precision and recall.

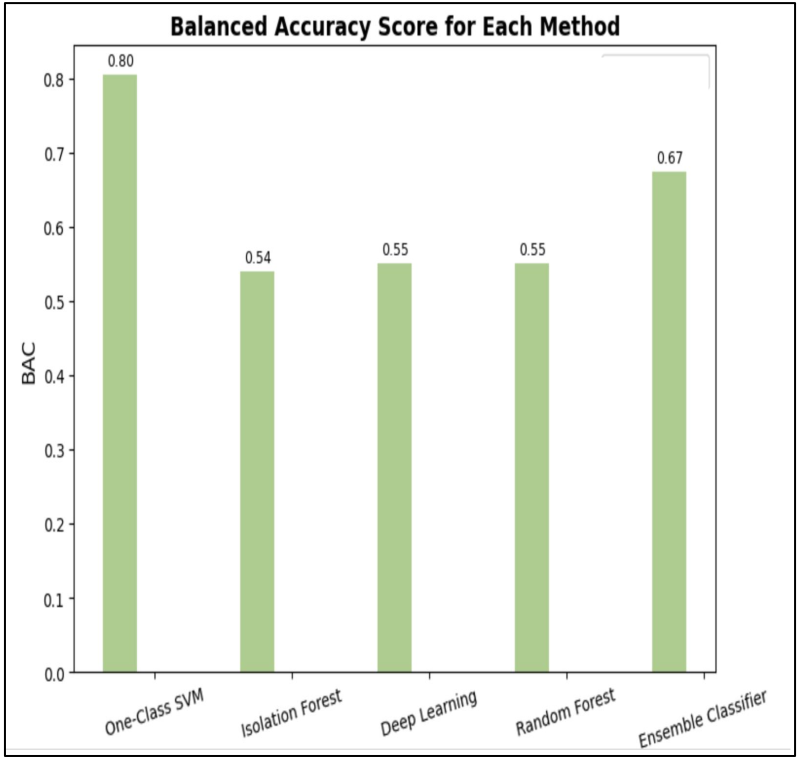
2. From the survey paper published in neurips 2024 [*] we can observe that the f1 score is relatively low for anomaly detection. In real world data, it's hard to find anomalies as there remains feature overlapping.

Dataset (Best Metrics)	Credit Card Precision	Credit Card Recall	Credit Card F1	GECCO Precision	GECCO Recall	GECCO F1	Globem Precision	Globem Recall	Globem F1
LSTM-RNN	0.004	0.11	0.007	0.343	0.275	0.305	0.12	0.16	0.14
IForest	0.098	0.569	0.168	0.439	0.353	0.391	0.13	0.66	0.21
OCSVM	0.107	0.62	0.183	0.185	0.743	0.296	**0.31**	**0.67**	**0.43**
AutoEncoder	0.103	0.598	0.176	0.424	0.34	0.377	0.11	0.15	0.13

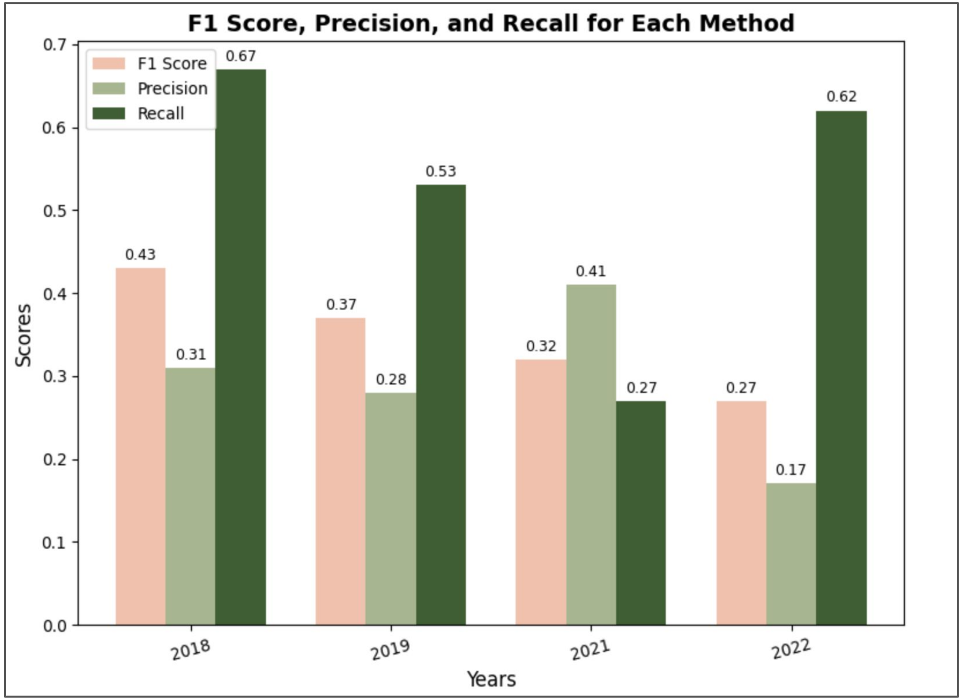
* Revisiting Time Series Outlier Detection: Definitions and Benchmarks (NeurIPS 2024)

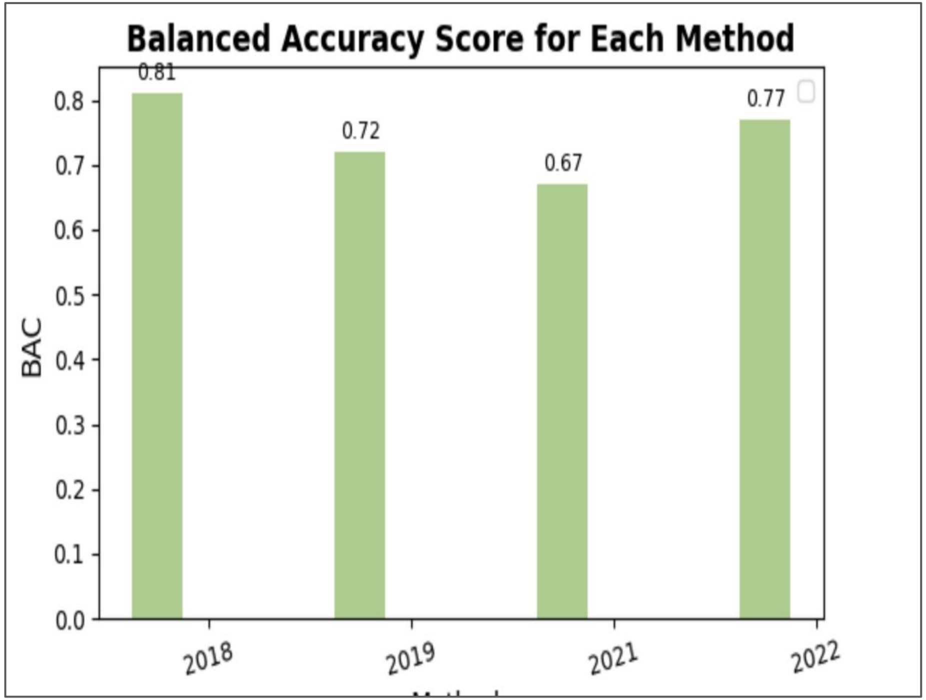
Performance across different models





Performance across different year



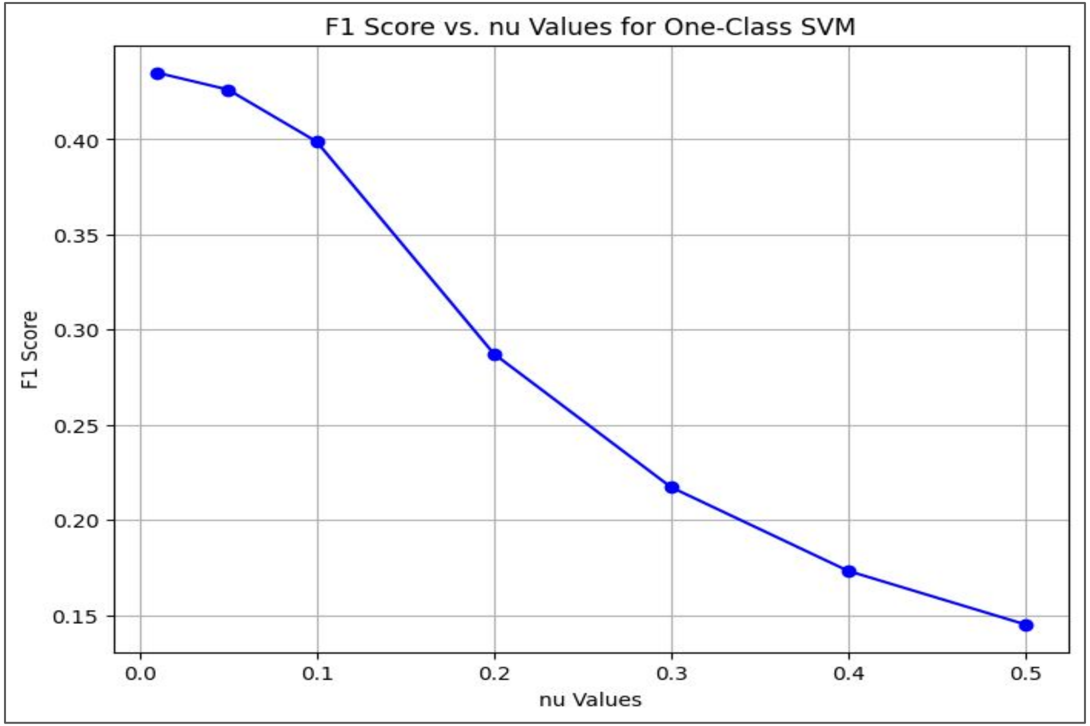


Feature Ablation Study

Days	F1 Score (Anomaly)
Previous Day	0.67
7 Consecutive Days	0.10
6 Consecutive Days	0.12
5 Consecutive Days	0.14
4 Consecutive Days	0.14
3 Consecutive Days	0.15
2 Consecutive Days	0.15
7 Average Days	0.13
6 Average Days	0.13
5 Average Days	0.13
4 Average Days	0.15
3 Average Days	0.13

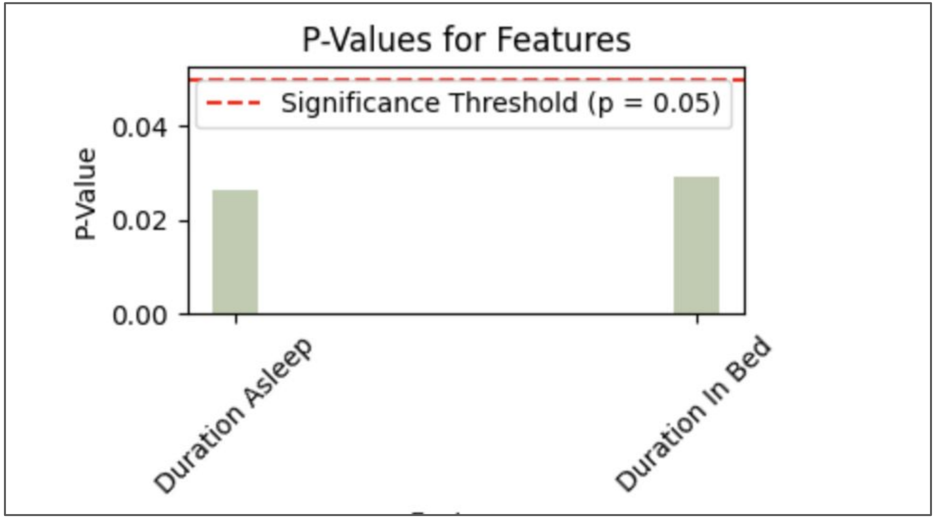
Hyperparameter Tuning in OCSVM

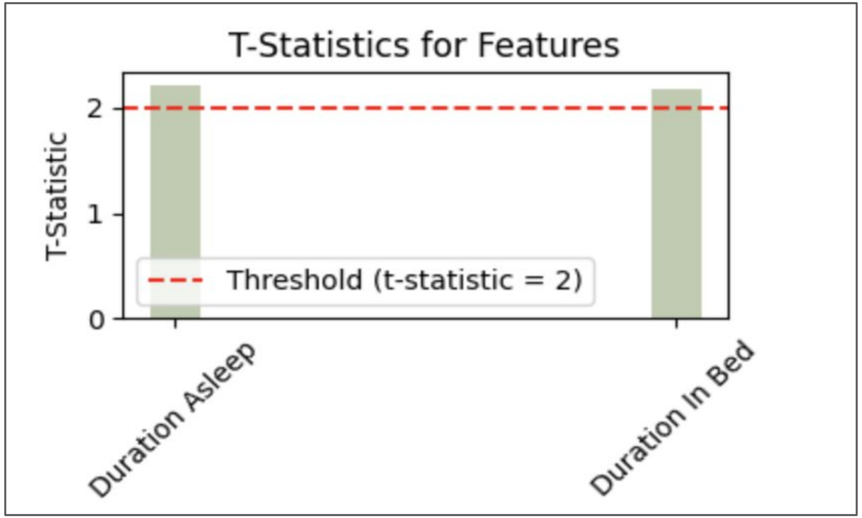
1. **nu** parameter (contamination ratio) controls the trade-off between anomalies and non-anomalies in dataset.
2. Higher **nu** values result in more false positives
3. Optimal performance is observed at lower **nu** values (~0.05 - 0.1)



Sleeping Efficiency Bin change and Mood

1. Measured by PHQ > 2 indicates poor mental health and sleep metrics
2. **Duration Asleep p-value = 0.0265**, **Duration in Bed p-value = 0.0293** confirming statistical significance.
3. **Duration Asleep t = 2.22**, **Duration in Bed t = 2.18** exceeded the critical value of 2





Future Work

- Improve precision by implementing filtering techniques to reduce false positives.
- Develop an online system for real-time anomaly detection and evaluate it with user studies.
- Enhance the robustness of the personalized recommendation system.

Conclusion

- This is the first work, to our knowledge, to predict sudden drops in sleeping efficiency, referred to as anomalies.
- Utilizes a smartphone-based system for real-time data collection and analysis.
- Achieves comparable results with state-of-the-art methods in anomaly detection.
- Provides insights into the potential impact of sleep efficiency drops on next-day mood and mental health.

Work Contribution

Jamalia Jisha : Dataset Processing, Problem Formulation, Model Development, Feature Analysis

Adiba Shaira : Related Work Study, Feature Extraction, Model Development, Feature Analysis