

Original Paper

Detecting Perceived Unfair Treatment Among US College Students Using Mobile Sensing: Pilot Machine Learning Study

Yiyi Ren¹, MS; Raghu Mulukutla², MS; Jennifer Mankoff³, PhD; Anind K Dey¹, PhD

¹Information School, University of Washington, Seattle, WA, United States

²College of Computing, Georgia Institute of Technology, Atlanta, GA, United States

³Paul G. Allen School of Computer Science & Engineering, University of Washington, Seattle, WA, United States

Corresponding Author:

Yiyi Ren, MS

Information School

University of Washington

1851 NE Grant Ln

Seattle, WA, 98105

United States

Phone: 1 4123308417

Email: yren98@uw.edu

Abstract

Background: Experiences of unfair treatment on college campuses are linked to adverse mental and physical health outcomes, highlighting the need for interventions. However, detecting such experiences relies mainly on self-reports. No prior research has examined the feasibility of using mobile sensing via smartphones and wearables for the passive detection of these experiences.

Objective: This pilot study explores the potential of using passive sensing to detect daily experiences of perceived unfair treatment (PUT) after they occur. It aims to develop and evaluate machine learning models against naive baselines and establish a benchmark for future research.

Methods: We analyzed data from 201 undergraduate students collected over two 10-week academic terms in 2018. PUT was self-reported at the daily level via ecological momentary assessment (EMA) surveys, with 413 of 9629 (4.3%) total responses indicating unfair treatment. We implemented two modeling approaches with distinct training schemes: (1) supervised classification models trained in a user-independent manner using data from different individuals, and (2) anomaly detection models trained in a user-dependent manner using historical data from the same individuals. Classification performance was assessed using stratified group 5-fold cross-validation for user-independent models and a chronological train-test split for user-dependent models.

Results: Of the 201 study participants, 110 reported experiencing unfair treatment at least once. On average, participants reported unfair treatment in 4.66% of their EMA responses (95% CI 3.13% to 6.19%). User-independent classification models showed mixed performance (AUC-ROC [area under the receiver operating characteristic curve]: 0.546-0.640, AUC-PR [area under the precision-recall curve]: 0.047-0.093, F_1 -score: 0.070-0.121). Tree-based models, particularly light gradient boosting machine (LightGBM) and Random Forest, outperformed all 3 baselines in AUC-ROC and AUC-PR; LightGBM also improved the F_1 -score. In comparison, user-dependent anomaly detection models performed better, with the multiday long short-term memory-AE model (50 features, 7-day window) achieving the highest recall (0.830, +73.3%, $P < .001$) and F_1 -score (0.391, +24.9%, $P < .001$) without reducing precision (0.256), and improving AUC-PR by 45.9% and AUC-ROC by 21.6% relative to naive baselines ($P < .001$). Feature importance analysis identified key behavioral patterns for population-level detection, including increased time spent off campus, elevated evening and nighttime activity, reduced indoor mobility on campus, prolonged screen use, delayed sleep onset, and shorter sleep duration.

Conclusions: Mobile sensing shows promise for detecting daily experiences of PUT in college students and identifying associated behavioral patterns. Our findings highlight opportunities for timely interventions through mobile technology to mitigate the impact of these experiences on students' mental health and well-being.

(JMIR Form Res 2025;9:e78657) doi: [10.2196/78657](https://doi.org/10.2196/78657)

KEYWORDS

passive sensing; perceived discrimination; anomaly detection; digital phenotyping; mental health; mobile health

Introduction

Unfair treatment refers to the act of denying individuals equal and just consideration based on characteristics such as race, gender, age, or disability [1]. In US college environments, perceived unfair treatment (PUT) remains a persistent issue with significant impacts on students' lives [2-5]. While the literature often uses discrimination interchangeably with unfair treatment [6,7], our study adopts the broader construct of PUT, which includes not only overt acts of discrimination but also subtle, everyday indignities known as microaggressions [8]. Drawing on past research on perceived discrimination to contextualize our work, we define PUT as an individual's subjective perception of being treated unjustly based on group characteristics. Within the university setting, this can manifest in various ways. For example, students being stereotyped by faculty, perceiving bias in academic evaluation, being unfairly blamed for dorm noise, or encountering classmates who express surprise at a minority student's success [9].

These experiences of PUT can induce acute physiological and emotional distress [10-13], contribute to increased suicidality [14-16], substance use [3,17,18], poor academic performance [19,20], and have long-lasting effects on social well-being and mental health, such as disrupted personality development [21], hindered career growth [22,23], eating disorders [24,25], and impaired social integration [26,27]. Despite their prevalence, many incidents go unreported [28-30], resulting in limited institutional awareness and response. Developing reliable methods to detect these experiences soon after they occur is crucial for enabling timely interventions [31,32] and providing social support [33,34] for at-risk students.

Currently, PUT is primarily studied and detected based on self-reports, either via standard questionnaires such as the Major Experiences of Discrimination and the Everyday Discrimination Scale [35-37], or through the Experience Sampling Method, also known as Ecological Momentary Assessment (EMA) [38-40]. While these self-reported measures provide valuable insights into individuals' experiences, they are subject to recall and nonresponse biases, inconsistent reporting, and significant participant burden [41-43], making them challenging to scale for continuous or longitudinal monitoring and detection. To the best of our knowledge, no framework or system exists that can automatically or passively detect PUT after it happens.

The health care landscape is undergoing a notable transformation, shifting towards noninvasive and accessible methods for early detection [44-46]. This shift is being largely fueled by advancements in mobile sensing technology and the growing interest in machine learning, which together offer unprecedented opportunities [47-50]. Numerous studies have highlighted the effectiveness of these technologies in addressing mental health and well-being tasks, such as depression screening and detection [51-55]. Concurrently, emerging research has begun to uncover short-term behavioral correlates of discrimination experiences [39,56,57], including changes in physical activity, sleep patterns, phone use, and social interactions—behaviors that can be objectively measured through smartphone and wearable sensors. However, most

previous studies have focused on uncovering health and behavioral associations with perceived discrimination. While these analytical approaches provide valuable insights, they do not directly address the challenge of detection. Therefore, this study aims to fill this gap by developing and evaluating models that detect PUT after it occurs based on behavioral changes. To our knowledge, this is the first work to explore passive detection using mobile sensing data. Our goal is to establish a benchmark that can inform and advance future research in this emerging field.

The combination of high-dimensional mobile sensing features and the flexibility of machine learning techniques allows for a data-driven, scalable, and personalized approach compared with traditional statistical methods [58-62]. Moreover, mobile sensing via smartphones or wearables offers a ubiquitous, continuous, and nonintrusive means of data collection, making it a powerful tool for capturing momentary experiences more effectively than traditional survey-based approaches [63-66]. Importantly, our goal is not to replace human expertise, but to augment it using technology for early detection and intervention at scale [67-69].

Passive detection of PUT presents unique challenges. Such events are often sporadic, vary greatly in form [70,71], and are perceived subjectively across individuals [72-74], leading to significant variability in experiences and reporting. This makes it difficult for longitudinal studies like ours to collect sufficient samples of day-to-day self-reported incidents, with adequate variance in the ground truth. Nonetheless, traditional instruments for measuring PUT, such as self-reports, have well-documented limitations [28], including biases [75,76] and limited ability to account for confounding factors such as physical or mental stressors. These tools also rely on repeated measures [77], which can lead to participant fatigue and reduced compliance over time. Importantly, self-reporting is inherently episodic, making it less suited for continuous, population-level screening across a campus setting. In contrast, passive sensing offers the ability to continuously and unobtrusively monitor behavioral and physiological signals over time, rendering it a promising tool to complement traditional methods in detecting and understanding these complex psychosocial phenomena.

Recent advancements in machine learning-based rare event detection (RED) have shown promising results across a range of domains, including health care and mobile sensing [78,79]. While traditional ensemble methods such as random forests and gradient boosting have been used [80-82], deep learning architectures such as autoencoders (AEs) and long short-term memory networks (LSTMs) are gaining momentum [83-88]. This is particularly relevant in multivariate time series settings [88,89], where smartphones and wearable devices continuously generate multiple streams of time-stamped data (eg, location, activity, and phone use) that capture complex behavioral patterns over time [90-93]. Despite these advancements, RED remains a challenging task [78,79], especially given the infrequent nature of the events. This often leads to reduced quantitative performance, particularly in metrics such as precision, recall, and F_1 -score, as studies across various domains frequently report only modest improvements over baseline methods [79]. For instance, Coley et al [94] reported a precision of 0.09, a recall

of 0.53, and an F_1 -score of 0.16 using a random forest model for suicide risk detection on a health care dataset with a rarity of 0.2%. Closer to our context, Pillai et al [95] proposed a multitask learning framework in the Tesseract study [92] that combined an unsupervised AE with an auxiliary sequence prediction task to detect rare life events (<2%) using mobile sensing data. This approach improved performance compared with several baselines, achieving an F_1 -score of 0.29. Such approaches have not yet been explored for detecting PUT experiences in social contexts like college environments. This gap presents an important opportunity to investigate the potential of mobile sensing-based RED methods in screening behavioral anomalies that may signal experiences of PUT.

Building on prior work [57,95], we developed and evaluated machine learning models leveraging mobile sensing data collected in 2018 as part of a multiyear study of undergraduate students [93,96]. We focused on two modeling approaches: (1) user-independent classification models, trained to identify behavioral patterns that are indicative of PUT across individuals; and (2) user-dependent anomaly detection models, personalized to detect short-term deviations in individual behavior that may signal responses to such experiences. Our objective is to assess the feasibility of using exclusively mobile sensing data, collected up until the time students wake the next morning, to detect experiences of PUT from the previous day at both the population and individual levels. We evaluated model performance against naive baselines using metrics commonly used in RED [97], including area under the receiver operating characteristic curve (AUC-ROC), area under the precision-recall curve (AUC-PR), precision, recall, and F_1 -score.

Methods

Ethical Considerations

This work was approved by the University of Washington Institutional Review Board and was assigned the ID Study00003324. All participants in the study provided their informed consent in person. All data collected in the study were kept separate from participants' personal identifiers, to provide anonymity and protect privacy. Participants could receive compensation up to US \$245/quarter in gift cards, depending on the completeness of their data collection, both passively sensed data and EMAs.

Data Collection

To be eligible for the study, participants were required to be over 18 years old, enrolled as first-year full-time undergraduates, and own an Android or iOS smartphone with an active data plan. The data collection period spanned 2 academic terms (January to June 2018), lasting approximately 20 weeks. At the beginning of the study, participants installed a smartphone app built using the AWARE framework [98] and wore a Fitbit tracker continuously throughout the study period, enabling passive data collection. This setup captured a wide range of mobile contextual information, including location, activity recognition, battery status, phone calls, screen use, and Bluetooth and Wi-Fi scans, while the Fitbit tracker provided steps and sleep data.

During the study, participants received regular EMA surveys [99] to report PUT experiences (Multimedia Appendix 1 contains survey questions). They completed EMAs twice weekly (on Sundays and Wednesdays) over an 8-week period, reporting on events from the previous day, and completed daily evening EMAs for an additional 2 weeks each academic term, reflecting on the same day's experiences. To identify which days participants experienced PUT, each EMA asked whether they had experienced unfair treatment on the reported day (today or yesterday).

Feature Extraction

In this study, we computed behavioral features from 7 smartphone data streams (activity recognition, battery, Bluetooth, call, location, screen, and Wi-Fi) and 2 wearable data streams (sleep and steps). We extended a behavioral feature extraction library [100] and followed a similar approach to prior work [90,101] by aggregating mobile sensing data into statistical summaries across various epochs of the day, including night (12 AM to 6 AM), morning (6 AM to 12 PM), afternoon (12 PM to 6 PM), and evening (6 PM to 12 AM), as well as the entire day (12 AM to 12 AM). This approach allowed us to characterize human behavior patterns within the day, providing a structured representation of daily behavior. Additionally, aligning our feature extraction process with prior work ensured methodological consistency, enabling direct comparisons and potential cross-study generalization. Notably, sleep features were computed only on a daily basis, as most student participants typically experienced a single major sleep episode each night. Some features were stream-specific, such as location variance derived solely from GPS data or the frequency of screen unlocks from screen events. Others involved multiple streams, for example, estimating indoor mobility duration by fusing location and activity recognition data. We describe the extraction of each set of behavioral features in Multimedia Appendix 2. In addition to these features, we computed the number of data samples collected for each data stream as an additional feature. This allowed us to assess data compliance and gain insights into event-based data streams such as calls, screen use, activity, and sleep, where the number of records reflects the frequency of these events.

Data Availability Analysis

Out of the 201 study participants, 110 individuals reported experiencing PUT at least once during the study period. To estimate how frequently PUT was reported overall, we calculated the proportion of positive EMA responses (indicating PUT) for each participant (ie, the number of positive responses divided by the total number of EMA responses they submitted). We then averaged these individual-level proportions across all participants, resulting in a mean reporting rate of 4.66% (95% CI 3.13% to 6.19%). The distribution of the EMA survey responses, with a total of 9629 submissions, included 413 (4.3%) positive responses and 9216 (95.7%) negative responses.

We observed a significant level of missing values in the behavioral features. Issues related to data collection, such as poor study compliance, phone battery depletion, or app crashes, directly contributed to the lack of raw sensing data. Additionally, event-based data streams, such as call logs, only record data

when specific events occur (eg, when a call is made), making it challenging to determine whether the absence of data was due to the absence of such events (eg, no calls made) or due to issues in data collection. Insufficient volume of raw sensing data per time period can also result in missing feature values. This is especially relevant for statistical features that require a reasonable number of samples for aggregated calculations. Similarly, features like Bluetooth and location often rely on a sufficient number of raw data records for effective data clustering. Last, the limited diversity of data streams per sample

can affect the computation of fused features, as they depend on nonmissing values from multiple data streams. For instance, both location and activity recognition data are required for extracting features such as study duration and indoor mobility.

In [Table 1](#), we report the availability of each data stream in the raw dataset, which serves as the foundation for feature calculation. We computed data availability as the percentage of daily samples where each data stream was available, relative to the total samples across all participants (total days \times total participants). A higher percentage indicates broader availability.

Table 1. Availability of each data stream in the raw dataset.

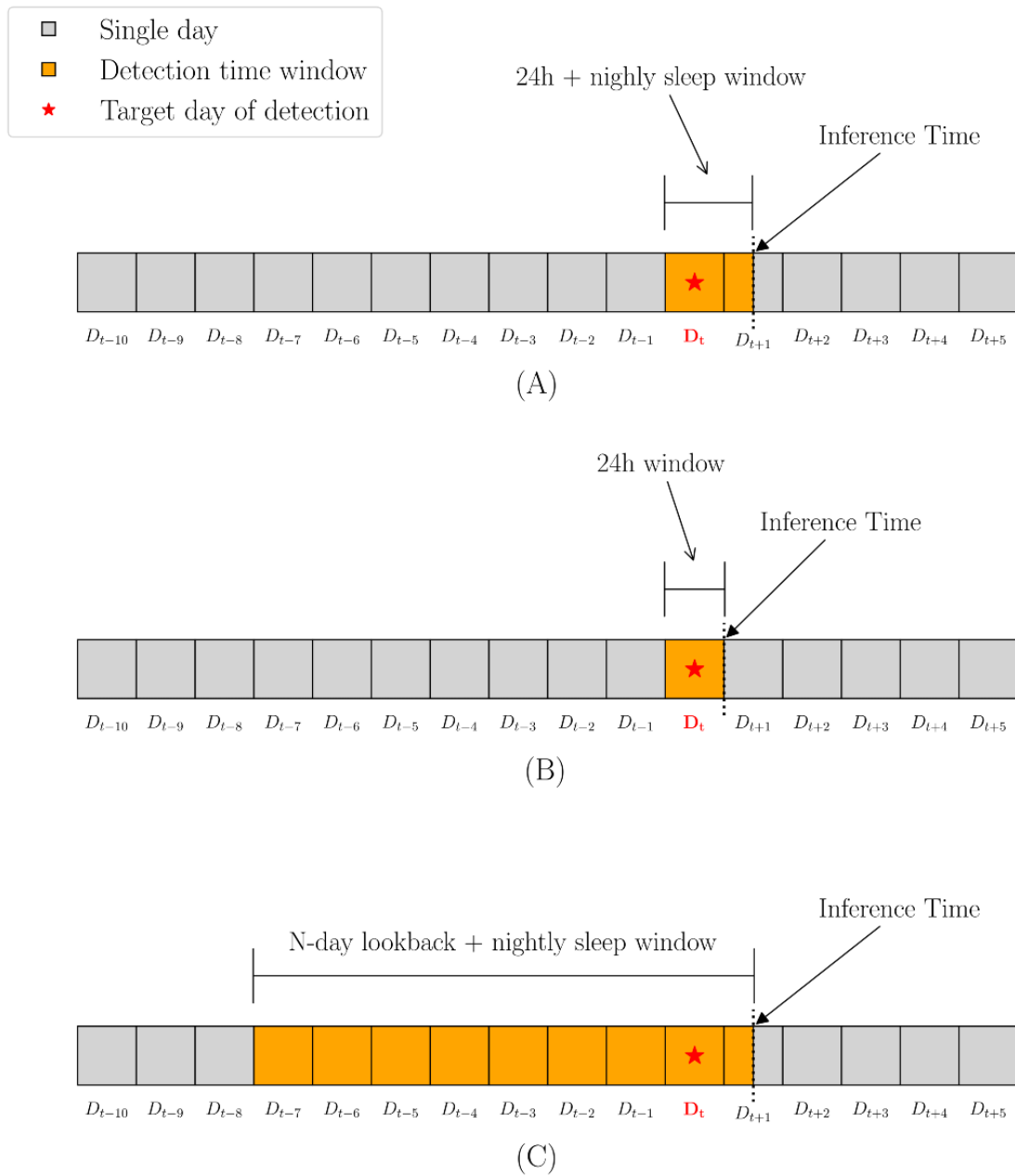
Data stream	Availability (%)
Activity	49.93
Battery	43.61
Bluetooth	48.65
Calls	35.05
Locations	51.08
Screen	51.85
Wi-Fi	52.44
Sleep	65.17
Steps	70.52

Modeling

In this study, we focus on 2 modeling approaches (user-independent and -dependent) to retrospectively detect PUT using daily inference windows ([Figure 1](#)), implemented through 3 specific model types. The user-independent classification models leverage labeled data from training participants to detect PUT in new or unseen participants. In contrast, the user-dependent anomaly detection models learn from participants' historical data to differentiate between normal and rare patterns in their own behavior. These approaches are driven by practical considerations. First, the scarcity of labeled

data makes it difficult to build a model that generalizes well to unseen users [[61,93,96,102-105](#)], though identifying broad patterns and key features remains essential in imbalanced settings [[106](#)]. Second, individuals' behavioral patterns before and after experiencing PUT can vary greatly [[21,107-110](#)] and are often moderated by various factors [[111-115](#)], highlighting the need for personalized training strategies. Third, given the rarity of target events, modeling them as anomalies is a commonly adopted approach in a wide variety of studies [[78,79](#)]. Furthermore, we believe these dual modeling approaches enhance the robustness of our models across diverse real-world deployment scenarios.

Figure 1. Illustration of detection time windows and inference times for three model types: (A) user-independent classification model, (B) intraday long short-term memory-autoencoder user-dependent model, and (C) multiday long short-term memory-autoencoder user-dependent model.



User-Independent Modeling

We selected light gradient boosting machine (LightGBM) [116] as our primary algorithm for user-independent modeling, for its ability to handle high-dimensional mobile sensing features, capture nonlinearity, mitigate overfitting through its ensemble mechanism, and natively manage missing data. LightGBM is a gradient boosting framework that builds decision trees efficiently, using a histogram-based approach to speed up training while maintaining high accuracy. Its built-in feature importance calculation enhances model interpretability and makes it an ideal choice for gaining population-wide insights in imbalanced settings. We benchmarked LightGBM with 4 classic machine learning algorithms: k-nearest neighbors (KNN), logistic regression, support vector machine (SVM), and random forest. All models were implemented using the scikit-learn

Python library [117], with binary cross-entropy loss as the objective function. For comparison, we also implemented 3 baseline classifiers. The first uses only static demographic information (Multimedia Appendix 3) without incorporating any behavioral data. The other two are naive baselines, which make random predictions without considering input features: (1) a uniform classifier that assigns labels randomly with equal probability, ignoring class distribution; and (2) a stratified classifier that assigns labels randomly while preserving the class distribution observed in the data.

Prior to training, we applied a filtering step that required each sample to contain at least 7 available data streams, resulting in a final dataset of 4720 person-day records. Of these, 167 (3.5%) were labeled positive and 4553 (96.5%) as negative. Participants contributed an average of 24 days of usable data (SD 9.6). LightGBM natively handled missing values, while for the classic

models, we applied median imputation via scikit-learn's SimpleImputer. To address class imbalance, we experimented with class weight adjustments as well as 2 widely used oversampling approaches: the synthetic minority oversampling technique and SVM-based synthetic minority oversampling technique [118,119].

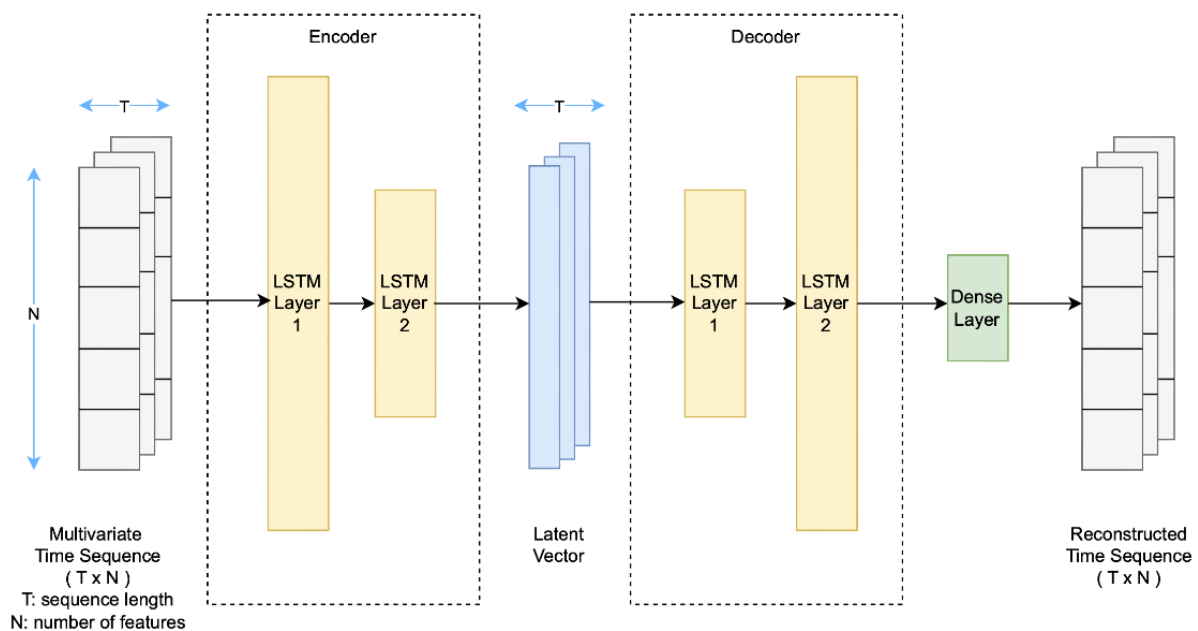
We chose the best-performing population-level model for subsequent feature analysis and selection, which was then used to build user-dependent models. As illustrated in Figure 1A, all models were trained on features extracted from a full 24-hour period on the target day, combined with sleep features from the nightly sleep window ending the following morning, after which inference was performed.

User-Dependent Modeling

For user-dependent modeling, we focused on anomaly detection with LSTM-AEs. LSTM-AEs are a type of neural network designed to learn temporal patterns in sequential data by encoding and reconstructing time-series inputs, making them effective for detecting deviations from typical behavioral patterns. Since human behavior typically follows daily routines and longer-term cycles (eg, weekly), we believe this approach helps detect anomalies in individual trajectories that may signal PUT. Our LSTM-AE architecture (Figure 2) consists of 2 stacked LSTM layers (encoder) that compress input sequences into a fixed-size representation, which is then replicated and decoded by another 2 LSTM layers (decoder). A final dense

layer reconstructs the original sequence. The model learns normal behavioral patterns by minimizing reconstruction loss (mean squared error). We implemented the model architecture using Keras and TensorFlow libraries. We explored both intraday (1-day lookback) and multiday (N-day lookback) input constructions to assess whether behavioral changes associated with PUT are better detected through short-term or longer-term temporal context. The intraday model focuses on capturing short-term daily patterns by dividing each 24-hour period into 4 fixed 6-hour epochs: night (12 AM to 6 AM), morning (6 AM to 12 PM), afternoon (12 PM to 6 PM), and evening (6 PM to 12 AM), resulting in sequences with 4 aggregated data points. In contrast, the multiday model captures longer-term trends using an N-day lookback window, where each sequence includes N days of daily features, ending with the target day. For both intraday and multiday models, we selected features based on the top behavioral features identified by the best-performing user-independent model. For the intraday model, inference occurs daily at the end of the day (Figure 1B); whereas for the multiday model, inference is performed the following morning after sleep-related features are available (Figure 1C). During inference, anomaly scores for intraday sequences were derived from the reconstruction loss of the entire sequence, whereas for multiday sequences, only the loss of the target (last) day was used. Anomalies were then defined using a threshold set at the 75th percentile of training reconstruction errors, providing a conservative and data-driven decision rule.

Figure 2. LSTM autoencoder architecture. LSTM: long short-term memory.



The models were trained exclusively on negative samples, with all positive samples reserved for testing, which is a common practice in RED [120-122]. Negative samples were split chronologically (90:10) into training and test sets. The training and test sample sizes are shown in Table 2. To prevent information leakage, multiday training samples overlapping

with positive test samples were excluded. To ensure data quality, input sequences with more than 80% missing data were removed, and for multiday models, sequences with more than 50% missingness on the target day were also excluded. Reconstruction loss was computed only for observed (non-NaN) positions to minimize bias.

Table 2. Training and test sample sizes for user-dependent modeling with long short-term memory autoencoders.

Model	Train (Neg)	Test		Per user	
		Pos	Neg	Train, mean (SD)	Test, mean (SD)
Intraday models					
50 features	6595	311	833	33 (12)	6 (3)
100 features	6558	309	827	33 (12)	6 (3)
150 features	4659	199	618	23 (8)	4 (3)
Multiday models					
50 features, 3-day window	4428	191	598	22 (9)	4 (3)
50 features, 5-day window	4341	191	598	22 (9)	4 (3)
50 features, 7-day window	4299	190	596	21 (9)	4 (3)
50 features, 9-day window	4248	189	596	21 (9)	4 (3)

Additionally, for comparison, we implemented two naive baseline models: (1) a uniform classifier that assigns labels randomly with equal probability, ignoring class distribution; and (2) a stratified classifier that assigns labels randomly while preserving the class distribution observed in the data.

Model Evaluation

For the user-independent model, we implemented nested stratified group K-fold cross-validation to address the dataset's limited size and class imbalance. The outer loop ($K=5$) split the data into training and test sets, ensuring class balance and participant-level separation. The inner loop ($K=4$) further splits the training data for feature selection and hyperparameter tuning, following the same stratified, participant-dependent structure. We evaluated performance using threshold-independent metrics (AUC-PR and AUC-ROC) and threshold-dependent metrics (precision, recall, and F_1 -score), reporting the mean and standard deviation across 25 independent performance estimates obtained from 5 repeats of 5-fold cross-validation with different random seeds. Threshold-dependent metrics are sensitive to both the chosen cutoff and outcome prevalence. To address this, we set the classification threshold to the empirical positive class probability from the training data, providing a consistent and data-driven operating point. Threshold-independent metrics complement this choice by characterizing performance across all thresholds, with AUC-PR being particularly informative under class imbalance [123,124]. Together, these metrics allow both robust model comparison and an understanding of the practical trade-offs between capturing rare events and limiting false alarms.

For the user-dependent anomaly detection model, we likewise evaluated performance using threshold-independent metrics (AUC-PR and AUC-ROC) and threshold-dependent metrics (precision, recall, and F_1 -score), reporting the mean and standard deviation across 10 randomized runs with different model initializations. During inference, the anomaly threshold was conservatively set at the 75th percentile of training

reconstruction errors, as the limited sample size precluded further fine-tuning. Threshold-independent metrics were computed from anomaly scores across test samples, while threshold-dependent metrics reflected performance at the selected threshold.

Results

Participants' Characteristics

The study focused on 201 full-time undergraduate students from the cohort enrolled in 2018 at the University of Washington. The mean age of the sample was 18.4 (SD 0.56) years. Female students comprised 64.7% (130/201 students) of the sample. Academically, students were drawn from a variety of departments, with approximately half majoring in engineering. To ensure a diverse sample, recruitment strategies included the oversampling of students from underrepresented backgrounds, specifically those with disabilities, first-generation students, and gender minorities.

User-Independent Modeling Results

Table 3 reports the performance metrics of all user-independent classification models alongside the baseline classifiers. To assess whether benchmarked models significantly outperformed the baselines, we conducted pairwise comparisons using the Wilcoxon signed-rank test (paired, 1-sided). Table 4 presents the P values from these comparisons, adjusted using the Benjamini-Hochberg false discovery rate procedure. The classification models showed mixed performance, with AUC-ROC ranging from 0.546 to 0.640, AUC-PR from 0.047 to 0.093, and F_1 -scores from 0.070 to 0.121. Compared with the baselines, some benchmarked models showed modest improvements. Both Random Forest and LightGBM achieved higher AUC-ROC and AUC-PR scores compared with KNN, Logistic Regression, SVM, and the baseline classifiers. Among all models, LightGBM achieved the highest F_1 -score. The observed variability across the cross-validation folds highlights the challenge of between-individual generalizability.

Table 3. Performance of user-independent classification models compared with baseline classifiers. Metrics are reported as mean (SD) across 25 estimates from repeated 5-fold cross-validation (5 repeats, different random seeds).

Model	AUC-ROC ^a , mean (SD)	AUC-PR ^b , mean (SD)	Precision, mean (SD)	Recall, mean (SD)	F_1 -score, mean (SD)
Baseline models					
Uniform	0.500 (0.000)	0.036 (0.018)	0.034 (0.018)	0.492 (0.121)	0.063 (0.031)
Stratified	0.501 (0.022)	0.037 (0.018)	0.046 (0.048)	0.041 (0.040)	0.039 (0.036)
Demographic	0.523 (0.126)	0.057 (0.043)	0.051 (0.044)	0.267 (0.143)	0.083 (0.065)
User-independent models					
KNN ^c	0.561 (0.062)	0.049 (0.025)	0.037 (0.020)	0.836 (0.074)	0.070 (0.036)
Logistic Regression	0.546 (0.082)	0.047 (0.026)	0.041 (0.023)	0.579 (0.145)	0.075 (0.040)
SVM ^d	0.567 (0.095)	0.065 (0.045)	0.053 (0.039)	0.268 (0.139)	0.089 (0.056)
Random Forest	0.634 (0.086)	0.093 (0.094)	0.045 (0.022)	0.709 (0.143)	0.084 (0.038)
LightGBM ^e	0.640 (0.065)	0.077 (0.043)	0.083 (0.050)	0.275 (0.118)	0.121 (0.064)

^aAUC-ROC: area under the receiver operating characteristic curve.

^bAUC-PR: area under the precision-recall curve.

^cKNN: k-nearest neighbors.

^dSVM: support vector machine.

^eLightGBM: light gradient boosting machine.

Table 4. Results of Wilcoxon signed-rank tests comparing user-independent models with baseline models. Reported *P* values were obtained from paired, one-sided Wilcoxon signed-rank tests and adjusted for multiple comparisons using the Benjamini-Hochberg false discovery rate (FDR) correction.

Baseline models	<i>P</i> values				
	AUC-ROC ^a	AUC-PR ^b	Precision	Recall	<i>F</i> ₁ -score
KNN^c					
Uniform	<.001	<.001	.03	<.001	.004
Stratified	<.001	<.001	.72	<.001	.002
Demographic	.16	.88	.97	<.001	.79
Logistic regression					
Uniform	.007	<.001	.004	.006	.003
Stratified	.007	<.001	.57	<.001	<.001
Demographic	.33	.91	.96	<.001	.65
SVM^d					
Uniform	.002	<.001	.002	>.99	.004
Stratified	.003	<.001	.16	<.001	<.001
Demographic	.16	.34	.56	.67	.54
Random forest					
Uniform	<.001	<.001	<.001	<.001	<.001
Stratified	<.001	<.001	.41	<.001	<.001
Demographic	.002	.03	.72	<.001	.23
LightGBM^e					
Uniform	<.001	<.001	<.001	>.99	<.001
Stratified	<.001	<.001	.002	<.001	<.001
Demographic	<.001	.03	.003	.47	.013

^aAUC-ROC: area under the receiver operating characteristic curve.

^bAUC-PR: area under the precision-recall curve.

^cKNN: k-nearest neighbors.

^dSVM: support vector machine.

^eLightGBM: Light Gradient Boosting Machine.

The pairwise comparisons indicated that both LightGBM and Random Forest outperformed all 3 baselines in terms of AUC-ROC and AUC-PR. LightGBM improved AUC-ROC by 22.4% (0.640 vs 0.523, $P < .001$) and Random Forest improved AUC-PR by 63.2% (0.093 vs 0.057, $P = .03$), both relative to the demographic baseline. While KNN, Logistic Regression, and Random Forest achieved significantly higher recall than the baselines, their precision values were significantly lower. LightGBM showed significant improvements in both precision and *F*₁-score compared with all 3 baselines, albeit at the cost of lower recall. Other benchmarked models showed smaller or nonsignificant differences, particularly compared with the demographic baseline. Overall, these findings suggest that

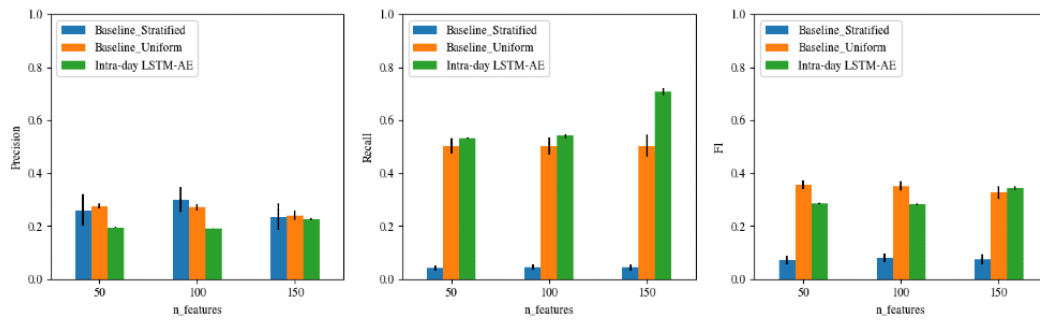
tree-based models can improve the detection of PUT events, which we explore further in the Discussion.

User-Dependent Modeling Results

Intraday Models

Figure 3 plots the performance of the intraday LSTM-AE model as the number of input features increases, compared with 2 naive baselines. Models trained with 50 and 100 features did not yield improvements over the baselines except for recall, while the model trained with 150 features showed improvements in both recall and *F*₁-score. In general, models with more input features achieved higher recall.

Figure 3. Performance of the intraday anomaly detection model using varying numbers of input features, compared against naive baseline models. LSTM-AE: long short-term memory-autoencoder.



Multiday Models

Figure 4 plots the performance of the multiday LSTM-AE model trained on input sequences with 50 features across varying window sizes (3, 5, 7, and 9 days), compared with 2 naive baselines. All models consistently outperformed both the stratified baseline and the uniform baseline with respect to recall and F_1 -score ($P < .001$ for both). The model with a 7-day window

achieved the highest recall (0.836), while the 3-day model achieved the highest precision (0.279). All multiday LSTM-AE models had comparable F_1 -scores (0.397-0.405). The 9-day window model showed a slight decline relative to the 7-day model in both recall (0.807 vs 0.836) and precision (0.397 vs 0.405), suggesting that extending the temporal window beyond 7 days yields diminishing returns.

Figure 4. Performance of the multiday anomaly detection model across different window sizes, compared against naive baseline models. LSTM-AE: long short-term memory-autoencoder.

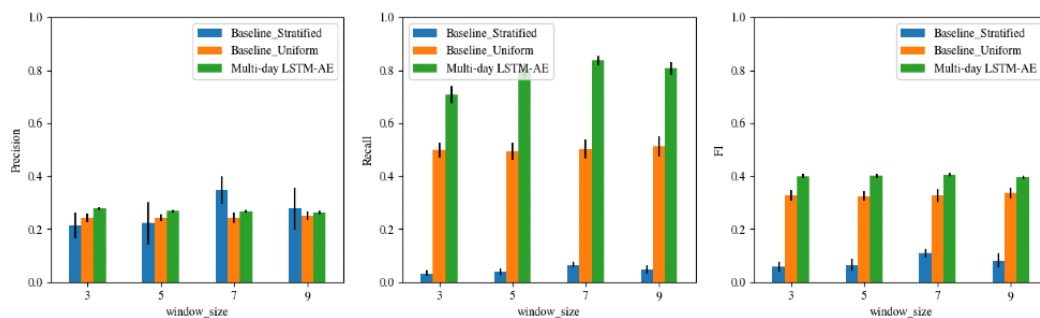


Table 5 summarizes the performance of the best-performing intraday and multiday models, alongside their respective naive baselines for reference. To assess whether benchmarked models significantly outperformed the baselines, we conducted pairwise comparisons using the Wilcoxon signed-rank test (paired, one-sided). Table 6 presents the P values from these comparisons, adjusted using the Benjamini-Hochberg false discovery rate procedure. As shown in Tables 5 and 6, the intraday model (150 features) did not show significant improvements over baselines except for recall ($P < .001$) and

F_1 -score ($P = .012$ vs uniform; $P < .001$ vs stratified). In contrast, the multiday model (50 features, 7-day window) significantly outperformed both baselines on nearly all evaluation metrics, except for precision against the stratified baseline. It achieved significantly higher recall (0.830 vs 0.479 for the uniform baseline, +73.3%, $P < .001$) and F_1 -score (0.391 vs 0.313 for the uniform baseline, +24.9%, $P < .001$) with comparable precision (0.256), while improving AUC-PR by 45.9% (0.353 vs 0.242, $P = .002$) and AUC-ROC by 21.6% (0.605 vs 0.500, $P = .002$) relative to the baselines.

Table 5. Performance of user-dependent models compared with baseline classifiers. Metrics are reported as mean (SD) across 10 randomized runs with different model initializations.

Model	AUC-ROC ^a , mean (SD)	AUC-PR ^b , mean (SD)	Precision, mean (SD)	Recall, mean (SD)	F ₁ -score, mean (SD)
Intraday (150 features)					
Baseline (“uniform”)	0.500 (0.000)	0.244 (0.000)	0.243 (0.013)	0.495 (0.030)	0.326 (0.017)
Baseline (“stratified”)	0.500 (0.000)	0.244 (0.000)	0.263 (0.069)	0.046 (0.016)	0.078 (0.027)
LSTM-AE ^c	0.447 (0.004)	0.240 (0.002)	0.227 (0.003)	0.703 (0.010)	0.343 (0.005)
Multiday (50 features, 7-day window)					
Baseline (“uniform”)	0.500 (0.000)	0.242 (0.000)	0.233 (0.015)	0.479 (0.034)	0.313 (0.020)
Baseline (“stratified”)	0.500 (0.000)	0.242 (0.000)	0.250 (0.079)	0.044 (0.014)	0.074 (0.024)
LSTM-AE	0.608 (0.008)	0.353 (0.008)	0.256 (0.005)	0.830 (0.027)	0.391 (0.007)

^aAUC-ROC: area under the receiver operating characteristic curve.

^bAUC-PR: area under the precision-recall curve.

^cLSTM-AE: long short-term memory-autoencoder.

Table 6. Results of the Wilcoxon signed-rank tests comparing user-dependent models with baseline models. Reported *P* values were obtained from paired, one-sided Wilcoxon signed-rank tests and adjusted for multiple comparisons using the Benjamini-Hochberg false discovery rate (FDR) correction.

Baseline	<i>P</i> values				
	AUC-ROC ^a	AUC-PR ^b	Precision	Recall	F ₁ -score
Intraday LSTM-AE^c (150 features)					
Uniform	>.99	>.99	>.99	<.001	.012
Stratified	>.99	>.99	>.99	<.001	<.001
Multiday LSTM-AE (50 features, 7-day window)					
Uniform	.002	.002	<.001	<.001	<.001
Stratified	.002	.002	.99	<.001	<.001

^aAUC-ROC: area under the receiver operating characteristic curve.

^bAUC-PR: area under the precision-recall curve.

^cLSTM-AE: long short-term memory-autoencoder.

Feature Importance Analysis

Analyzing feature importance in user-independent classification models is essential for identifying key behavioral features associated with reported PUT at the population level. In this study, we examined feature importance scores from the best-performing LightGBM model computed using information gain and averaged across all cross-validation folds. Higher scores indicate greater influence on the model’s decision. However, feature importance in tree-based models can be sensitive to correlated features [125] or sampling imbalance [126] and should be interpreted with caution. Future work could incorporate Shapley additive explanations [127] values to provide more robust insights into feature contributions and interactions in model decisions.

Table 7 summarizes the top 30 daily behavioral features ranked by their average importance scores. To better understand their associations with the outcome, we analyzed value distributions across positive and negative daily samples and computed standardized mean differences (SMDs) as effect sizes. Larger SMDs indicate stronger class separation. All top-ranked features

had |SMD| values below 0.5, with 17 falling in the small-to-moderate range (|SMD| between 0.2 and 0.5) and 13 in the minimal separation range (|SMD| less than 0.2). This indicates that while these features may not independently distinguish between the 2 outcomes, they likely contribute to the model’s decision through complex interactions or nonlinear associations.

Interestingly, the top 10 features are evenly distributed across a variety of sensing modalities, including 2 from steps, one fused feature derived from both activity and location, and one feature each from the remaining 6 data streams: location, activity, screen, Bluetooth, Wi-Fi, and call. This balanced distribution suggests that diverse behavioral signals, rather than any single dominant source, collectively contribute to the model’s detection ability. Across the top 30 features, the most frequently represented time epochs were the all-day window (n=9) and the evening period (n=7), indicating that both cumulative and evening-specific behavioral patterns are particularly informative for detection. As summarized in Textbox 1, these selected features touch upon key aspects of

students' daily lives, reflecting mobility, activity, phone use, social interactions, and sleep patterns.

Table 7. Top 30 daily behavioral features identified by the best-performing light gradient boosting machine model.

Feature	Data stream	Epoch	Positive, mean (SD)	Negative, mean (SD)	Importance score	Effect size (SMD ^a)
Total time spent off-campus (minutes)	Location	All day	729.5 (505.8)	535.3 (435.3)	892.1	0.44
Total step count	Steps	Evening	4140.6 (2878.8)	3106.0 (2726.1)	628.5	0.38
Activity sample count	Activity	Night	171.7 (102.4)	139.6 (96.2)	595.8	0.33
Total screen time (minutes)	Screen	Afternoon	116.7 (70.1)	90.1 (64.9)	447.9	0.41
Bluetooth sample count	Bluetooth	Morning	25.2 (76.1)	45.9 (106.0)	415.3	-0.20
Number of unique Wi-Fi access points	Wi-Fi	All day	17.8 (16.9)	17.7 (24.0)	380.1	0.01
Number of missed calls	Call	Evening	1.2 (1.9)	0.8 (1.6)	290.2	0.25
Indoor mobility duration (minutes)	Activity/location	All day	48.3 (64.0)	65.2 (72.1)	289.2	-0.24
Sleep sample count	Sleep	All day	456.6 (128.9)	469.4 (112.0)	273.5	-0.11
Average duration of sedentary bouts (minutes)	Steps	Evening	37.8 (74.8)	37.2 (69.9)	257.2	0.01
Total number of active bouts	Steps	All day	56.9 (29.4)	52.3 (21.1)	192.2	0.21
Longest stay duration at study places (minutes)	Location	All day	50.6 (98.6)	45.4 (79.6)	175.5	0.06
Average stay duration off-campus (minutes)	Location	All day	137.6 (207.9)	91.1 (160.8)	175.4	0.29
Shortest phone interaction bout (minutes)	Screen	All day	10.5 (105.4)	3.5 (46.4)	173.2	0.14
First unlock time (seconds since midnight)	Screen	Night	1244.6 (2611.8)	1436.3 (3400.3)	159.2	-0.06
Average stay duration in green spaces (minutes)	Location	All day	27.4 (57.1)	26.1 (68.1)	146.4	0.02
Total sedentary time (minutes)	Steps	Evening	289.1 (48.9)	306.0 (40.4)	146.2	-0.42
Percentage of time spent off-campus	Location	All day	0.6 (0.4)	0.4 (0.3)	141.9	0.32
Shortest phone interaction bout (minutes)	Screen	Night	14.1 (44.2)	6.9 (34.3)	140.7	0.21
Activity sample count	Activity	Morning	203.2 (94.6)	211.0 (138.1)	135.1	-0.06
Sedentary bout duration variation (minutes)	Steps	Evening	28.5 (25.6)	29.8 (24.2)	132.8	-0.05
Circadian movement	Location	Night	2.3 (0.0)	2.3 (0.0)	132.7	0.02
Duration of physical activities (minutes)	Activity	Evening	84.8 (64.3)	64.9 (53.9)	131.9	0.37
Number of unique Wi-Fi access points	Wi-Fi	Afternoon	8.8 (8.6)	9.7 (38.8)	130.6	-0.03
Total time spent off-campus (minutes)	Location	Morning	171.1 (146.6)	117.8 (125.5)	130.0	0.42
Average steps per active bout	Steps	Evening	209.7 (137.0)	169.8 (141.0)	128.8	0.28
Last active bout end time (seconds since midnight)	Steps	Morning	40,457.5 (4216.2)	41,539.9 (2925.4)	127.9	-0.36
Variation in time spent in green spaces (minutes)	Location	Morning	3.2 (11.3)	1.3 (7.6)	120.9	0.23
Start time of Nightly sleep (seconds since midnight)	Sleep	All day	34,585.9 (7774.7)	33,569.8 (8917.1)	120.5	0.11

Feature	Data stream	Epoch	Positive, mean (SD)	Negative, mean (SD)	Importance score	Effect size (SMD ^a)
Shortest phone interaction bout (minutes)	Screen	Afternoon	1.6 (18.9)	2.3 (19.8)	118.8	-0.03

^aSMD: standardized mean difference.

Textbox 1. Examples of selected sensing features and their relation to students' daily behaviors.

- Campus-map features: Features such as time spent off-campus and indoor mobility show that students who experience perceived unfair treatment (PUT) tend to spend more time away from campus and exhibit reduced movement within campus buildings.
- Physical activity features: Features such as step count and activity duration show that these students tend to be more active during the evening and night hours.
- Screen use features: Longer afternoon screen time, longer minimum interaction durations, and earlier phone unlocks at night among students who report experiences of PUT. These features provide insights into phone use patterns and digital engagement.
- Bluetooth features: Lower sample counts in the morning for students who experience PUT. These features may serve as indicators of social exposure and proximity to others.
- Call features: A higher number of missed calls during evening hours among students reporting PUT. These features may reflect phone availability and social responsiveness.
- Sleep features: Shorter sleep durations and later sleep onset times on days marked by PUT.

Discussion

Principal Findings

In this study, we reported results from both classification and anomaly detection machine learning models evaluated under user-independent and user-dependent settings for detecting PUT among college students. A novel aspect of our work is the exclusive use of passively collected mobile sensing data for training and inference, offering a nonintrusive and low-burden alternative to traditional self-reports. As shown in [Figure 1](#), our models were able to detect past events within a day of occurrence (including the subsequent nightly sleep window), indicating that mobile sensing may offer timely detection of PUT experiences and potential practical applications in future interventions. The best-performing classification model (LightGBM) significantly outperformed all 3 baseline classifiers, including the demographic baseline, in AUC-ROC, AUC-PR, precision, and F_1 -score. The top-performing anomaly detection model, which incorporated a 7-day temporal context and 50 features, significantly outperformed the baselines across most metrics, achieving notably higher recall while maintaining comparable precision. These results suggest that mobile sensing features may serve as behavioral indicators for detecting experiences of PUT. While not intended to replace traditional assessments, mobile sensing could potentially complement existing methods, especially for large-scale or continuous monitoring.

Among the user-independent classification models, ensemble tree-based classifiers such as random forest and LightGBM consistently outperformed traditional machine learning algorithms like logistic regression and SVM, achieving higher overall performance. This result is consistent with previous work from various disciplines, highlighting the robustness of ensemble methods for handling imbalanced datasets [[128-132](#)] and the advantage of nonlinear models in capturing the complex

behavioral patterns typically observed in mobile sensing data [[90,93,133,134](#)]. However, the overall limited performance across all models, along with the observed high variability across different user-independent data splits (eg, higher standard deviations in performance metrics), underscores the challenges of between-individual generalizability posed by the infrequent and subjective nature of the detection task. While user-independent models remain valuable for identifying globally informative features, future work should focus on personalized or semipersonalized modeling approaches that can better accommodate individual variability. Balancing scalability with personalization, for instance, by deploying personalized models for individuals identified as high-risk by a global model, may guide a more effective and targeted detection framework.

Our preliminary results suggest that daily behavioral signals alone, whether modeled in a user-independent setting (eg, classification models using daily features) or a user-dependent setting (eg, the intraday LSTM-AE model), were insufficient to reliably capture behavioral shifts associated with PUT. In contrast, models that incorporated multiday temporal windows demonstrated notable performance improvements. For example, all multiday anomaly detection models achieved statistically significant gains in recall and F_1 -score compared with the corresponding baselines. This suggests that a longer temporal context may help detect more gradual or subtle behavioral deviations that may not be evident within a single day. Among the multiday models, the 7-day model achieved the best overall performance, while longer windows (eg, 9-day) showed reduced effectiveness. This suggests that the 7-day window likely strikes a balance between capturing sufficient behavioral context and remaining short enough to detect localized anomalies. It may also align with natural weekly rhythms that help reveal meaningful patterns. Taken together, these findings highlight the potential value of modeling temporal dynamics across multiple days. We recommend that future work identify and apply an optimal window length that balances contextual

richness with anomaly detectability to improve sensitivity without compromising precision.

In user-dependent anomaly detection, intraday models with 150 features outperformed those with 50 or 100 features, underscoring the value of a rich feature set for capturing behavioral anomalies. Feature importance analysis further indicated that performance gains were not driven by any single type of sensor data. Instead, a diverse set of behavioral indicators spanning multiple sensing modalities, including physical activity, phone use, mobility, and sleep, all contribute meaningfully to model performance. The absence of a dominant modality or feature suggests that effective detection likely relies on multimodal inputs and potentially their interactions to capture the complex, context-dependent nature of behavioral patterns rather than on isolated signals. These findings highlight the multifaceted nature of behavioral responses to PUT, though larger studies are needed to confirm these patterns. In the following section, we further contextualize these results by relating them to prior work.

Connections to Existing Literature

Our location-based features reveal that on days with reported PUT, students spend more time off-campus, which may reflect disengagement or withdrawal from campus life. Indoor mobility within campus buildings is also lower (eg, fewer transitions between classrooms, libraries, or other study areas), potentially indicating reduced academic or social participation. Moreover, Bluetooth sample counts are lower on these days, revealing fewer nearby Bluetooth-enabled devices, which may imply decreased social exposure. These findings support prior research showing that discrimination can undermine students' sense of belonging and social well-being, often resulting in increased feelings of isolation, social withdrawal, decreased academic and campus engagement, and even truancy [27,135-139]. Regarding physical activity, our features show that evening and nighttime activity tends to be higher on days with reported PUT, with greater step counts and longer durations of physical activity. These findings align with previous studies linking higher physical activity to perceived discrimination [109,140,141], possibly reflecting altered routines or coping behaviors. Our phone-use related features show longer afternoon screen time, earlier first unlocks, and longer shortest nightly phone interactions on days with reported PUT compared with days without. These patterns align with prior findings linking perceived discrimination to problematic phone use among students [142-144]. Finally, our sleep features show that students reporting PUT tend to have a later sleep onset and shorter sleep duration on the same day. This is consistent with previous research linking perceived discrimination to reduced sleep duration and poorer sleep quality [145-150].

While closely aligning with prior research, our study extends the literature by offering a unique short-term behavioral lens on the impact of PUT, complementing the longer-term patterns typically emphasized in the field. Many of the sensed patterns we identify likely reflect students' immediate behavioral and physiological responses to PUT. These patterns are not only informative for detection but can also guide interventions in campus environments, enabling continuous monitoring and

timely, targeted support. While our models cannot prevent the initial event, they could facilitate timely interventions to mitigate its adverse effects and promote student well-being. For instance, our models could trigger personalized microinterventions delivered via push notifications [151-154], such as encouraging on-campus engagement, prompting positive social interaction, suggesting phone-use breaks, offering early sleep reminders and sleep hygiene education. These data-driven strategies can be integrated with traditional approaches, including support groups, self-care practices, and professional mental health services.

Cross-Domain Reflection: Rare Event Detection Challenges

We explored applying RED methodologies to the domain of detecting daily PUT. While we were able to connect some of our findings to existing literature on discrimination, most studies focused on retrospective self-reports, prevalence, and associated health outcomes. Relatively little attention has been given to detecting these events as they occur in daily life. To better situate our methodology and evaluate our model's performance, we conducted a comparative review of established RED approaches across diverse domains, including health care, crowd behavior, and mechanical systems. As summarized in [Multimedia Appendix 4](#), these studies [94,95,155-158] use a wide range of methods and demonstrate anomaly detection performance broadly comparable to ours.

Although our research is situated in a different domain, we find significant value in such cross-domain reflection. First, it allows us to assess whether our progress in capturing PUT experiences aligns with the advancements achieved in other novel contexts. Second, it helps validate our modeling approaches. Given the novelty of our application, we view these comparisons as reflective benchmarks rather than direct performance evaluations.

In our work, we encountered several challenges inherent to RED, including a highly skewed class distribution, difficulties in capturing correlations due to data sparsity, and similarities between rare and nonrare events. These challenges are consistent with those faced by other RED studies across various domains [79]. A further commonality is the inherent trade-off between recall and precision. High recall indicates a model's ability to capture rare instances, but it is often accompanied by lower precision, suggesting that many predicted events may not be actual occurrences. This pattern is evident not only in our results but also across other domains. For instance, Pillai et al [95], in a context most similar to ours, achieved a recall of 0.21 and a precision of 0.47 using a dataset where rare events constituted approximately 1.9% of the data. In health care, however, models often deliberately favor recall over precision to ensure that critical cases are not missed. Inspired by this perspective, our study advocates for a recall-oriented approach that prioritizes identifying as many PUT experiences as possible, thereby maximizing opportunities to support students at risk.

A distinguishing aspect of our work is the relatively small dataset compared with those used in other studies, such as Coley et al [94] in suicide risk assessment. Despite this limitation, our

models achieved recall and F_1 -scores that are comparable to, or in some cases exceed, those reported in similar domains, although direct comparisons are limited by dataset differences.

Limitations

We recognize several limitations in this pilot study. First, responses to PUT were collected at the daily level in a binary format (yes or no), without capturing the exact timing of each incident. This limited our ability to precisely align intraday sensor data patterns with specific experiences. For this reason, we cannot rule out that the intraday model's performance was affected by this limitation. Future work could benefit from more granular reporting to enable more accurate temporal analyses.

Second, our dataset was relatively small. While adequate for this exploratory pilot study, larger and more diverse datasets will be needed to validate and generalize these findings.

Third, this study focused exclusively on mobile sensing features. While this approach allowed us to isolate the predictive utility of passive sensing data, we acknowledge that the integration of additional domain knowledge, such as demographic and socioeconomic variables, or information about the type or reason for each incident, may improve model performance, interpretability, and fairness. Future research could reintroduce demographic and contextual information to account for heterogeneity across individuals and potentially enhance both predictive accuracy and fairness.

Finally, while our study EMAs were designed to directly ask about the target event and we aimed to detect PUT based on participants' behavioral responses, we acknowledge that many of the identified behavioral patterns (eg, social withdrawal, changes in phone use, sleep, and physical activity) could result from other negative experiences or mental health conditions such as depression, representing potential confounding factors. In addition, we cannot rule out the potential moderating influence of other variables on the association between PUT and the sensed features. Fully disentangling PUT-related signals from overlapping influences remains an open challenge and an important direction for future research.

Conclusions

This pilot study demonstrates the feasibility of using mobile sensing to screen for instances of PUT (4.3% or 413 of 9629 responses) among college students, providing a promising alternative to traditional self-report methods. Our machine learning models, leveraging diverse mobile sensing features and multiday temporal context, show strong potential for capturing short-term behavioral changes indicative of these infrequent experiences. We envision that future personalized and context-aware ML approaches, enhanced by larger datasets and deeper domain knowledge, will further improve detection accuracy, ultimately enabling timely interventions and support for at-risk students.

Acknowledgments

This material is based upon work supported by the National Science Foundation (grant numbers IIS1816687 and IIS7974751), and the University of Washington College of Engineering and the Paul G. Allen School of Computer Science and Engineering.

Data Availability

The datasets generated or analyzed during this study are available in the GLOBEM repository [159].

Authors' Contributions

YR and AD conceptualized the project. YR led the data curation, decisions about methodology, and the formal analysis, with support from RM and AD. AD provided the resources for the project. YR and RM led the writing of the original draft. All authors contributed to the reviewing and editing of the manuscript. AD and JM provided supervision for this project.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Ecological momentary assessment questions.

[[PDF File \(Adobe PDF File\), 51 KB-Multimedia Appendix 1](#)]

Multimedia Appendix 2

Behavioral feature extraction details.

[[PDF File \(Adobe PDF File\), 49 KB-Multimedia Appendix 2](#)]

Multimedia Appendix 3

Demographic variables used in baseline classifier.

[[PDF File \(Adobe PDF File\), 54 KB-Multimedia Appendix 3](#)]

Multimedia Appendix 4

Comparison of rare event detection studies.

[\[PDF File \(Adobe PDF File\), 60 KB-Multimedia Appendix 4\]](#)

References

1. Stuber J, Meyer I, Link B. Stigma, prejudice, discrimination and health. *Soc Sci Med*. 2008;67(3):351-357. [FREE Full text] [doi: [10.1016/j.socscimed.2008.03.023](https://doi.org/10.1016/j.socscimed.2008.03.023)] [Medline: [18440687](https://pubmed.ncbi.nlm.nih.gov/18440687/)]
2. Bravo AJ, Wedell E, Villarosa-Hurlocker MC, Looby A, Dickter CL, Schepis TS, et al. Stimulant Norms Prevalence (SNAP) Study Team. Perceived racial/ethnic discrimination among young adult college students: prevalence rates and associations with mental health. *J Am Coll Health*. 2023;71(7):2062-2073. [doi: [10.1080/07448481.2021.1954012](https://doi.org/10.1080/07448481.2021.1954012)] [Medline: [34398695](https://pubmed.ncbi.nlm.nih.gov/34398695/)]
3. Qeadan F, Azagba S, Barbeau WA, Gu LY, Mensah NA, Komaromy M, et al. Associations between discrimination and substance use among college students in the United States from 2015 to 2019. *Addict Behav*. 2022;125:107164. [doi: [10.1016/j.addbeh.2021.107164](https://doi.org/10.1016/j.addbeh.2021.107164)] [Medline: [34735979](https://pubmed.ncbi.nlm.nih.gov/34735979/)]
4. Qeadan F, Madden EF, Barbeau WA, Mensah NA, Azagba S, English K. Associations between discrimination and adverse mental health symptoms and disorder diagnoses among college students in the United States. *J Affect Disord*. 2022;310:249-257. [doi: [10.1016/j.jad.2022.05.026](https://doi.org/10.1016/j.jad.2022.05.026)] [Medline: [35561881](https://pubmed.ncbi.nlm.nih.gov/35561881/)]
5. 2023 annual report. Center for Collegiate Mental Health (CCMH). URL: https://ccmh.psu.edu/assets/docs/2023_Annual%20Report.pdf [accessed 2025-05-08]
6. Sovern J. Is discrimination unfair? *Ga State Univ Law Rev*. 2025;41(3):631. [FREE Full text]
7. Denise EJ, Hagiwara N. "Discrimination" versus "Unfair Treatment": measuring differential treatment and its association with health. *Sociol Inq*. 2019;89(4):645-676. [doi: [10.1111/soin.12277](https://doi.org/10.1111/soin.12277)]
8. Sue DW, Capodilupo CM, Torino GC, Bucceri JM, Holder AMB, Nadal KL, et al. Racial microaggressions in everyday life: implications for clinical practice. *Am Psychol*. 2007;62(4):271-286. [doi: [10.1037/0003-066x.62.4.271](https://doi.org/10.1037/0003-066x.62.4.271)]
9. Suarez-Balcazar Y, Orellana-Damacela L, Portillo N, Rowan JM, Andrews-Guillen C. Experiences of differential treatment among college students of color. *J High Educ*. 2016;74(4):428-444. [doi: [10.1080/00221546.2003.11780855](https://doi.org/10.1080/00221546.2003.11780855)]
10. Emmer C, Dorn J, Mata J. The immediate effect of discrimination on mental health: a meta-analytic review of the causal evidence. *Psychol Bull*. 2024;150(3):215-252. [doi: [10.1037/bul0000419](https://doi.org/10.1037/bul0000419)] [Medline: [38330346](https://pubmed.ncbi.nlm.nih.gov/38330346/)]
11. Thayer JF, Carnevali L, Sgifofo A, Williams DP. Angry in America: psychophysiological responses to unfair treatment. *Ann Behav Med*. 2020;54(12):924-931. [doi: [10.1093/abm/kaa094](https://doi.org/10.1093/abm/kaa094)] [Medline: [33416840](https://pubmed.ncbi.nlm.nih.gov/33416840/)]
12. Ong AD, Deshpande S, Williams DR. Biological consequences of unfair treatment: a theoretical and empirical review. *Handb Cult Biol*. 2017:279-315. [doi: [10.1002/9781119181361.ch12](https://doi.org/10.1002/9781119181361.ch12)]
13. Harrell JP, Hall S, Taliaferro J. Physiological responses to racism and discrimination: an assessment of the evidence. *Am J Public Health*. 2003;93(2):243-248. [doi: [10.2105/AJPH.93.2.243](https://doi.org/10.2105/AJPH.93.2.243)]
14. Gomez J, Miranda R, Polanco L. Acculturative stress, perceived discrimination, and vulnerability to suicide attempts among emerging adults. *J Youth Adolesc*. 2011;40(11):1465-1476. [FREE Full text] [doi: [10.1007/s10964-011-9688-9](https://doi.org/10.1007/s10964-011-9688-9)] [Medline: [21717234](https://pubmed.ncbi.nlm.nih.gov/21717234/)]
15. Busby DR, Horwitz AG, Zheng K, Eisenberg D, Harper GW, Albuher RC, et al. Suicide risk among gender and sexual minority college students: the roles of victimization, discrimination, connectedness, and identity affirmation. *J Psychiatr Res*. 2020;121:182-188. [FREE Full text] [doi: [10.1016/j.jpsychires.2019.11.013](https://doi.org/10.1016/j.jpsychires.2019.11.013)] [Medline: [31837538](https://pubmed.ncbi.nlm.nih.gov/31837538/)]
16. Mao Y, Liu L, Lu Z, Wang W. Relationships between perceived discrimination and suicidal ideation among impoverished Chinese college students: the mediating roles of social support and loneliness. *Int J Environ Res Public Health*. 2022;19(12):7290. [FREE Full text] [doi: [10.3390/ijerph19127290](https://doi.org/10.3390/ijerph19127290)] [Medline: [35742540](https://pubmed.ncbi.nlm.nih.gov/35742540/)]
17. Le TP, Iwamoto DK. A longitudinal investigation of racial discrimination, drinking to cope, and alcohol-related problems among underage Asian American college students. *Psychol Addict Behav*. 2019;33(6):520-528. [doi: [10.1037/adb0000501](https://doi.org/10.1037/adb0000501)] [Medline: [31414850](https://pubmed.ncbi.nlm.nih.gov/31414850/)]
18. Fahey MC, Morris JD, Robinson LA, Pebley K. Association between perceived discrimination and vaping among college students. *Subst Use Misuse*. 2021;56(5):738-741. [doi: [10.1080/10826084.2021.1887250](https://doi.org/10.1080/10826084.2021.1887250)] [Medline: [33682601](https://pubmed.ncbi.nlm.nih.gov/33682601/)]
19. Steele R, Rosado A, Hernandez N, Brondolo E. Discrimination, acculturative stress, and academic achievement in emerging adults. *J Vincent Soc Action*. 2022;6(1). [FREE Full text]
20. Stevens C, Liu CH, Chen JA. Racial/ethnic disparities in US college students' experience: discrimination as an impediment to academic performance. *J Am Coll Health*. 2018;66(7):665-673. [doi: [10.1080/07448481.2018.1452745](https://doi.org/10.1080/07448481.2018.1452745)] [Medline: [29565755](https://pubmed.ncbi.nlm.nih.gov/29565755/)]
21. Kim J, Song K, Sutin AR. Gender differences in the relationship between perceived discrimination and personality traits in young adulthood: evidence using sibling fixed effects. *Soc Sci Med*. 2021;286:114329. [doi: [10.1016/j.socscimed.2021.114329](https://doi.org/10.1016/j.socscimed.2021.114329)] [Medline: [34428601](https://pubmed.ncbi.nlm.nih.gov/34428601/)]
22. Schmidt CK, Miles JR, Welsh AC. Perceived discrimination and social support: the influences on career development and college adjustment of LGBT college students. *J Career Dev*. 2010;38(4):293-309. [doi: [10.1177/0894845310372615](https://doi.org/10.1177/0894845310372615)]

23. Liu X, Sun X, Hao Q. Influence of discrimination perception on career exploration of higher vocational students: chain mediating effect test. *Front Psychol.* 2022;13:968032. [FREE Full text] [doi: [10.3389/fpsyg.2022.968032](https://doi.org/10.3389/fpsyg.2022.968032)] [Medline: [35967637](https://pubmed.ncbi.nlm.nih.gov/35967637/)]
24. Kwan MY, Gordon KH, Minnich AM. An examination of the relationships between acculturative stress, perceived discrimination, and eating disorder symptoms among ethnic minority college students. *Eat Behav.* 2018;28:25-31. [doi: [10.1016/j.eatbeh.2017.12.003](https://doi.org/10.1016/j.eatbeh.2017.12.003)] [Medline: [29306178](https://pubmed.ncbi.nlm.nih.gov/29306178/)]
25. Harris CL, Haack S, Miao Z. Everyday discrimination is a stronger predictor of eating competence than food insecurity or perceived stress in college students amidst COVID-19. *Appetite.* 2022;179:106300. [FREE Full text] [doi: [10.1016/j.appet.2022.106300](https://doi.org/10.1016/j.appet.2022.106300)] [Medline: [36075312](https://pubmed.ncbi.nlm.nih.gov/36075312/)]
26. Chaw A. The correlation between perceived discrimination and social anxiety in college students who identify as LGBTQ. *Undergrad Res J.* 2023;3(2). [doi: [10.58361/2766-3590.1063](https://doi.org/10.58361/2766-3590.1063)]
27. Villegas-Gold R, Yoo HC. Coping with discrimination among Mexican American college students. *J Couns Psychol.* 2014;61(3):404-413. [doi: [10.1037/a0036591](https://doi.org/10.1037/a0036591)] [Medline: [25019544](https://pubmed.ncbi.nlm.nih.gov/25019544/)]
28. Lewis TT, Cogburn CD, Williams DR. Self-reported experiences of discrimination and health: scientific advances, ongoing controversies, and emerging issues. *Annu Rev Clin Psychol.* 2015;11(1):407-440. [doi: [10.1146/annurev-clinpsy-032814-112728](https://doi.org/10.1146/annurev-clinpsy-032814-112728)] [Medline: [25581238](https://pubmed.ncbi.nlm.nih.gov/25581238/)]
29. Shammass D. Underreporting discrimination among Arab American and Muslim American community college students. *J Mix Methods Res.* 2016;11(1):99-123. [doi: [10.1177/1558689815599467](https://doi.org/10.1177/1558689815599467)]
30. Irby-Shasanmi A, Leech TGJ. 'Because I Don't know': uncertainty and ambiguity in closed-ended reports of perceived discrimination in US health care. *Ethn Health.* 2017;22(5):458-479. [doi: [10.1080/13557858.2016.1244659](https://doi.org/10.1080/13557858.2016.1244659)] [Medline: [27741709](https://pubmed.ncbi.nlm.nih.gov/27741709/)]
31. Sue DW, Alsaidi S, Awad MN, Glaeser E, Calle CZ, Mendez N. Disarming racial microaggressions: microintervention strategies for targets, White allies, and bystanders. *Am Psychol.* 2019;74(1):128-142. [doi: [10.1037/amp0000296](https://doi.org/10.1037/amp0000296)] [Medline: [30652905](https://pubmed.ncbi.nlm.nih.gov/30652905/)]
32. Awad MN, Connors EH. Active bystandership by youth in the digital era: microintervention strategies for responding to social media-based microaggressions and cyberbullying. *Psychol Serv.* 2023;20(3):423-434. [doi: [10.1037/ser0000749](https://doi.org/10.1037/ser0000749)] [Medline: [36951730](https://pubmed.ncbi.nlm.nih.gov/36951730/)]
33. Morton SCM, Everhart R, Dautovich N, Chukmaitov A. Perceived discrimination and mental health outcomes in college students: the mediating effect of preventive health behaviors and social support. *J Am Coll Health.* 2025;73(6):2380-2389. [doi: [10.1080/07448481.2023.2286462](https://doi.org/10.1080/07448481.2023.2286462)] [Medline: [38010405](https://pubmed.ncbi.nlm.nih.gov/38010405/)]
34. Ajrouch KJ, Reisine S, Lim S, Sohn W, Ismail A. Perceived everyday discrimination and psychological distress: does social support matter? *Ethn Health.* 2010;15(4):417-434. [FREE Full text] [doi: [10.1080/13557858.2010.484050](https://doi.org/10.1080/13557858.2010.484050)] [Medline: [20582775](https://pubmed.ncbi.nlm.nih.gov/20582775/)]
35. Williams DR, Gonzalez HM, Williams S, Mohammed SA, Moomal H, Stein DJ. Perceived discrimination, race and health in South Africa. *Soc Sci Med.* 2008;67(3):441-452. [FREE Full text] [doi: [10.1016/j.socscimed.2008.03.021](https://doi.org/10.1016/j.socscimed.2008.03.021)] [Medline: [18486292](https://pubmed.ncbi.nlm.nih.gov/18486292/)]
36. Williams DR, Yan Yu, Jackson JS, Anderson NB. Racial differences in physical and mental health: socio-economic status, stress and discrimination. *J Health Psychol.* 1997;2(3):335-351. [FREE Full text] [doi: [10.1177/135910539700200305](https://doi.org/10.1177/135910539700200305)] [Medline: [22013026](https://pubmed.ncbi.nlm.nih.gov/22013026/)]
37. Williams DR. Measuring discrimination resource. *Psychology.* 1997;2(3):335-351. [doi: [10.1093/acprof:oso/9780199732166.003.0003](https://doi.org/10.1093/acprof:oso/9780199732166.003.0003)]
38. Civitillo S, Jugert P. Zooming in on everyday ethnic-racial discrimination: a review of experiencing sampling methodology studies in adolescence. *Eur J Dev Psychol Routledge.* 2024;21(4):592-611. [doi: [10.31234/osf.io/chxed](https://doi.org/10.31234/osf.io/chxed)]
39. Nam S, Jeon S, Ash G, Whittemore R, Vlahov D. Racial discrimination, sedentary time, and physical activity in African Americans: quantitative study combining ecological momentary assessment and accelerometers. *JMIR Form Res.* 2021;5(6):e25687. [FREE Full text] [doi: [10.2196/25687](https://doi.org/10.2196/25687)] [Medline: [34096870](https://pubmed.ncbi.nlm.nih.gov/34096870/)]
40. Livingston NA, Flentje A, Heck NC, Szalda-Petree A, Cochran BN. Ecological momentary assessment of daily discrimination experiences and nicotine, alcohol, and drug use among sexual and gender minority individuals. *J Consult Clin Psychol.* Dec 2017;85(12):1131-1143. [FREE Full text] [doi: [10.1037/ccp0000252](https://doi.org/10.1037/ccp0000252)] [Medline: [29189029](https://pubmed.ncbi.nlm.nih.gov/29189029/)]
41. Tarrant MA, Manfredo MJ, Bayley PB, Hess R. Effects of recall bias and nonresponse bias on self-report estimates of angling participation. *North Am J Fish Manag.* 1993;13(2):217-222. [doi: [10.1577/1548-8675\(1993\)013<0217:eorban>2.3.co;2](https://doi.org/10.1577/1548-8675(1993)013<0217:eorban>2.3.co;2)]
42. Napa SC, Prieto CK, Diener E. Diener E, editor. *Experience Sampling: Promises and Pitfalls, Strength and Weaknesses.* Netherlands. Assess Well- Collect Works Ed Diener Dordrecht Springer; 2009:157-180.
43. Van Dyke ME, Kramer MR, Kershaw KN, Vaccarino V, Crawford ND, Lewis T. Inconsistent reporting of discrimination over time using the experiences of discrimination scale: potential underestimation of lifetime burden. *Am J Epidemiol.* 2022;191(3):370-378. [FREE Full text] [doi: [10.1093/aje/kwab151](https://doi.org/10.1093/aje/kwab151)] [Medline: [34017974](https://pubmed.ncbi.nlm.nih.gov/34017974/)]
44. Abernethy A, Adams L, Barrett M, Bechtel C, Brennan P, Butte A, et al. The promise of digital health: then, now, and the future. *NAM Perspect.* 2022. [FREE Full text] [doi: [10.31478/202206e](https://doi.org/10.31478/202206e)] [Medline: [36177208](https://pubmed.ncbi.nlm.nih.gov/36177208/)]

45. Naik N, Hameed BMZ, Sooriyaperakasam N, Vinayahalingam S, Patil V, Smriti K, et al. Transforming healthcare through a digital revolution: a review of digital healthcare technologies and solutions. *Front Digit Health*. 2022;4:919985. [FREE Full text] [doi: [10.3389/fdgh.2022.919985](https://doi.org/10.3389/fdgh.2022.919985)] [Medline: [35990014](https://pubmed.ncbi.nlm.nih.gov/35990014/)]
46. The promise of digital healthcare technologies. *Frontiers*. URL: <https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2023.1196596/full> [accessed 2025-05-10]
47. Sameh A, Rostami M, Oussalah M, Korpelainen R, Farrahi V. Digital phenotypes and digital biomarkers for health and diseases: a systematic review of machine learning approaches utilizing passive non-invasive signals collected via wearable devices and smartphones. *Artif Intell Rev*. 2024;58(2):66. [doi: [10.1007/s10462-024-11009-5](https://doi.org/10.1007/s10462-024-11009-5)]
48. Bhatt P, Liu J, Gong Y, Wang J, Guo Y. Emerging artificial intelligence-empowered mHealth: scoping review. *JMIR Mhealth Uhealth*. 2022;10(6):e35053. [FREE Full text] [doi: [10.2196/35053](https://doi.org/10.2196/35053)] [Medline: [35679107](https://pubmed.ncbi.nlm.nih.gov/35679107/)]
49. Boukhechba M, Baglione AN, Barnes LE. Leveraging mobile sensing and machine learning for personalized mental health care. *Ergon Des*. 2020;28(4):18-23. [doi: [10.1177/1064804620920494](https://doi.org/10.1177/1064804620920494)]
50. Mohr DC, Zhang M, Schueller SM. Personal sensing: understanding mental health using ubiquitous sensors and machine learning. *Annu Rev Clin Psychol*. 2017;13:23-47. [FREE Full text] [doi: [10.1146/annurev-clinpsy-032816-044949](https://doi.org/10.1146/annurev-clinpsy-032816-044949)] [Medline: [28375728](https://pubmed.ncbi.nlm.nih.gov/28375728/)]
51. Khoo LS, Lim MK, Chong CY, McNaney R. Machine learning for multimodal mental health detection: a systematic review of passive sensing approaches. *Sensors (Basel)*. 2024;24(2):348. [FREE Full text] [doi: [10.3390/s24020348](https://doi.org/10.3390/s24020348)] [Medline: [38257440](https://pubmed.ncbi.nlm.nih.gov/38257440/)]
52. Wang R, Wang W, Dasilva A, Huckins JF, Kelley WM, Heatherton TF, et al. Tracking depression dynamics in college students using mobile phone and wearable sensing. *Proc ACM Interact Mob Wearable Ubiquitous Technol*. 2018;2(1):1-26. [doi: [10.1145/3191775](https://doi.org/10.1145/3191775)] [Medline: [39449996](https://pubmed.ncbi.nlm.nih.gov/39449996/)]
53. Opoku AK, Visuri A, Ferreira DST. Towards early detection of depression through smartphone sensing. Association for Computing Machinery; 2019. Presented at: Adjun Proc 2019 ACM Int Jt Conf Pervasive Ubiquitous Comput Proc 2019 ACM Int Symp Wearable Comput; 2019 May 10:1158-1161; New York, NY, USA. [doi: [10.1145/3341162.3347075](https://doi.org/10.1145/3341162.3347075)]
54. Levine LM, Gwak M, Kärkkäinen K, Fazeli S, Zadeh B, Peris T, et al. Alam MM, Hämäläinen M, Mucchi L, Niazi IK, Le Moulec Y, editors. *Anxiety Detection Leveraging Mobile Passive Sensing*. Cham. Springer International Publishing; 2020:212-225.
55. Acikmese Y, Alptekin SE. Prediction of stress levels with LSTM and passive mobile sensors. *Procedia Comput Sci*. 2019;159:658-667. [doi: [10.1016/j.procs.2019.09.221](https://doi.org/10.1016/j.procs.2019.09.221)]
56. Tapia AL, Wallace ML, Hasler BP, Holmes J, Pedersen SL. Effect of daily discrimination on naturalistic sleep health features in young adults. *Health Psychol*. 2024;43(4):298-309. [doi: [10.1037/hea0001359](https://doi.org/10.1037/hea0001359)] [Medline: [38190204](https://pubmed.ncbi.nlm.nih.gov/38190204/)]
57. Sefidgar YS, Seo W, Kuehn KS, Althoff T, Browning A, Riskin E, et al. Passively-sensed behavioral correlates of discrimination events in college students. *Proc ACM Hum Comput Interact*. 2019;3(CSCW):1-29. [FREE Full text] [doi: [10.1145/3359216](https://doi.org/10.1145/3359216)] [Medline: [34322658](https://pubmed.ncbi.nlm.nih.gov/34322658/)]
58. Rajula HSR, Verlato G, Manchia M, Antonucci N, Fanos V. Comparison of conventional statistical methods with machine learning in medicine: diagnosis, drug development, and treatment. *Medicina (Kaunas)*. 2020;56(9):455. [FREE Full text] [doi: [10.3390/medicina56090455](https://doi.org/10.3390/medicina56090455)] [Medline: [32911665](https://pubmed.ncbi.nlm.nih.gov/32911665/)]
59. Bennett M, Kleczyk EJ, Hayes K, Mehta R, Bennett M, Kleczyk EJ, et al. Evaluating similarities and differences between machine learning and traditional statistical modeling in healthcare analytics. In: *Artificial Intelligence Annual Volume 2022*. London, UK. IntechOpen; 2022.
60. Mann J, Lyons M, O'Rourke J, Davies S. Machine learning or traditional statistical methods for predictive modelling in perioperative medicine: a narrative review. *J Clin Anesth*. 2025;102:111782. [doi: [10.1016/j.jclinane.2025.111782](https://doi.org/10.1016/j.jclinane.2025.111782)] [Medline: [39977974](https://pubmed.ncbi.nlm.nih.gov/39977974/)]
61. Adler DA, Wang F, Mohr DC, Choudhury T. Machine learning for passive mental health symptom prediction: generalization across different longitudinal mobile sensing studies. *PLoS One*. 2022;17(4):e0266516. [FREE Full text] [doi: [10.1371/journal.pone.0266516](https://doi.org/10.1371/journal.pone.0266516)] [Medline: [35476787](https://pubmed.ncbi.nlm.nih.gov/35476787/)]
62. Cho H, She J, Marchi DD, El-Zaatari H, Barnes EL, Kahkoska AR, et al. Machine learning and health science research: tutorial. *J Med Internet Res*. 2024;26:e50890. [FREE Full text] [doi: [10.2196/50890](https://doi.org/10.2196/50890)] [Medline: [38289657](https://pubmed.ncbi.nlm.nih.gov/38289657/)]
63. Ságvári B, Gulyás A, Koltai J. Attitudes towards participation in a passive data collection experiment. *Sensors (Basel)*. 2021;21(18):6085. [FREE Full text] [doi: [10.3390/s21186085](https://doi.org/10.3390/s21186085)] [Medline: [34577291](https://pubmed.ncbi.nlm.nih.gov/34577291/)]
64. Cornet VP, Holden RJ. Systematic review of smartphone-based passive sensing for health and wellbeing. *J Biomed Inform*. 2018;77:120-132. [FREE Full text] [doi: [10.1016/j.jbi.2017.12.008](https://doi.org/10.1016/j.jbi.2017.12.008)] [Medline: [29248628](https://pubmed.ncbi.nlm.nih.gov/29248628/)]
65. Zhang H, Ibrahim A, Parsia B, Poliakoff E, Harper S. Passive social sensing with smartphones: a systematic review. *Computing*. 2022;105(1):29-51. [doi: [10.1007/s00607-022-01112-2](https://doi.org/10.1007/s00607-022-01112-2)]
66. Maharjan SM, Poudyal A, van Heerden A, Byanjankar P, Thapa A, Islam C, et al. Passive sensing on mobile devices to improve mental health services with adolescent and young mothers in low-resource settings: the role of families in feasibility and acceptability. *BMC Med Inform Decis Mak*. 2021;21(1):117. [FREE Full text] [doi: [10.1186/s12911-021-01473-2](https://doi.org/10.1186/s12911-021-01473-2)] [Medline: [33827552](https://pubmed.ncbi.nlm.nih.gov/33827552/)]

67. Singh S, Melendez K, Sezginis N. Examining the effect of discrimination and stigma on utilization of mental health services among college students. *J Am Coll Health*. 2023;71(8):2398-2405. [doi: [10.1080/07448481.2021.1970561](https://doi.org/10.1080/07448481.2021.1970561)] [Medline: [34519631](https://pubmed.ncbi.nlm.nih.gov/34519631/)]
68. Madrid-Cagigal A, Kealy C, Potts C, Mulvenna MD, Byrne M, Barry MM, et al. Digital mental health interventions for university students with mental health difficulties: a systematic review and meta-analysis. *Early Interv Psychiatry*. 2025;19(3):e70017. [doi: [10.1111/eip.70017](https://doi.org/10.1111/eip.70017)] [Medline: [40033658](https://pubmed.ncbi.nlm.nih.gov/40033658/)]
69. Topooco N, Fowler LA, Fitzsimmons-Craft EE, DePietro B, Vázquez MM, Firebaugh M, et al. Digital interventions to address mental health needs in colleges: perspectives of student stakeholders. *Internet Interv*. 2022;28:100528. [FREE Full text] [doi: [10.1016/j.invent.2022.100528](https://doi.org/10.1016/j.invent.2022.100528)] [Medline: [35378846](https://pubmed.ncbi.nlm.nih.gov/35378846/)]
70. Mikula G, Petri B, Tanzer N. What people regard as unjust: types and structures of everyday experiences of injustice. *Eur J Soc Psychol*. 2006;20(2):133-149. [doi: [10.1002/ejsp.2420200205](https://doi.org/10.1002/ejsp.2420200205)]
71. Vargas SM, Huey SJ, Miranda J. A critical review of current evidence on multiple types of discrimination and mental health. *Am J Orthopsychiatry*. 2020;90(3):374-390. [doi: [10.1037/ort0000441](https://doi.org/10.1037/ort0000441)] [Medline: [31999138](https://pubmed.ncbi.nlm.nih.gov/31999138/)]
72. Gong F, Xu J, Takeuchi DT. Racial and ethnic differences in perceptions of everyday discrimination. *Sociol Race Ethn*. 2016;3(4):506-521. [doi: [10.1177/2332649216681587](https://doi.org/10.1177/2332649216681587)]
73. Gonzalez D, McDaniel M, Kenney GM, Skopec L. Perceptions of unfair treatment or judgment due to race or ethnicity in five settings. Robert Wood Johnson Foundation. URL: <https://www.rwjf.org/en/insights/our-research/2021/07/perceptions-of-discriminatory-experiences-in-health-care-and-other-settings.html> [accessed 2021-08-05]
74. Yang P, Henderson S. Race, gender, class, and perceived everyday discrimination. *J Ethn Cult Stud*. Jul 30, 2024;11(3):51-66. [FREE Full text] [doi: [10.29333/ejecs/1801](https://doi.org/10.29333/ejecs/1801)]
75. Kaiser CR, Major B. A social psychological perspective on perceiving and reporting discrimination. *Law Soc Inq*. 2006;31(4):801-830. [doi: [10.1111/j.1747-4469.2006.00036.x](https://doi.org/10.1111/j.1747-4469.2006.00036.x)]
76. Foster MD, Dion KL. Dispositional hardiness and women's well-being relating to gender discrimination: the role of minimization. *Psychol Women Q*. 2003;27(3):197-208. [doi: [10.1111/1471-6402.00099](https://doi.org/10.1111/1471-6402.00099)]
77. Gaston SA, Jackson CL. Invited commentary: the need for repeated measures and other methodological considerations when investigating discrimination as a contributor to health. *Am J Epidemiol*. 2022;191(3):379-383. [FREE Full text] [doi: [10.1093/aje/kwab223](https://doi.org/10.1093/aje/kwab223)] [Medline: [34431505](https://pubmed.ncbi.nlm.nih.gov/34431505/)]
78. Abubakar YI, Othmani A, Siarry P, Sabri AQM. A systematic review of rare events detection across modalities using machine learning and deep learning. *IEEE Access*. 2024;12:47091-47109. [doi: [10.1109/access.2024.3382140](https://doi.org/10.1109/access.2024.3382140)]
79. Shyalika C, Wickramarachchi R, Sheth AP. A comprehensive survey on rare event prediction. *ACM Comput Surv*. 2024;57(3):1-39. [doi: [10.1145/3699955](https://doi.org/10.1145/3699955)]
80. Bai Y, Huang Z, Lam H, Zhao D. Rare-event simulation for neural network and random forest predictors. *ACM Trans Model Comput Simul*. 2022;32(3):1-33. [doi: [10.1145/3519385](https://doi.org/10.1145/3519385)]
81. Islam MDK, Hridi P, Hossain MDS, Narman HS. Network anomaly detection using LightGBM: a gradient boosting classifier. 2020. Presented at: 2020 30th Int Telecommun Netw Appl Conf ITNAC; 2020 Nov 25:1-7; Melbourne, Australia. [doi: [10.1109/itnac50341.2020.9315049](https://doi.org/10.1109/itnac50341.2020.9315049)]
82. Blagus R, Lusa L. Gradient boosting for high-dimensional prediction of rare events. *Comput Stat Data Anal*. 2017;113:19-37. [doi: [10.1016/j.csda.2016.07.016](https://doi.org/10.1016/j.csda.2016.07.016)]
83. Provotar O, Linder Y, Veres M. Unsupervised anomaly detection in time series using LSTM-based autoencoders. 2019. Presented at: 2019 IEEE International Conference on Advanced Trends in Information Theory; 2019 Dec 18:513-517; Kyiv, Ukraine. [doi: [10.1109/atit49449.2019.9030505](https://doi.org/10.1109/atit49449.2019.9030505)]
84. Said EM, Le-Khac NA, Dev S, Jurcut AD. Network anomaly detection using LSTM based autoencoder. Association for Computing Machinery; 2020. Presented at: Proceedings of the 16th ACM Symposium on QoS and Security for Wireless and Mobile Networks; 2020 Nov 06:37-45; New York, NY, USA. [doi: [10.1145/3416013.3426457](https://doi.org/10.1145/3416013.3426457)]
85. Wei Y, Jang-Jaccard J, Xu W, Sabrina F, Camtepe S, Boulic M. LSTM-autoencoder-based anomaly detection for indoor air quality time-series data. *IEEE Sensors J*. 2023;23(4):3787-3800. [doi: [10.1109/jsen.2022.3230361](https://doi.org/10.1109/jsen.2022.3230361)]
86. Wang Z, Kasongo Dahouda M, Hwang H, Joe I. Explanatory LSTM-AE-based anomaly detection for time series data in marine transportation. *IEEE Access*. 2025;13:23195-23208. [doi: [10.1109/access.2025.3535695](https://doi.org/10.1109/access.2025.3535695)]
87. Malhotra P, Ramakrishnan A, Anand G, Vig L, Agarwal P, Shroff G. LSTM-based encoder-decoder for multi-sensor anomaly detection. *arXiv*. 2016. [doi: [10.48550/arXiv.1607.00148](https://doi.org/10.48550/arXiv.1607.00148)]
88. Pang G, Shen C, Cao L, Hengel AVD. Deep learning for anomaly detection. *ACM Comput Surv*. 2021;54(2):1-38. [doi: [10.1145/3439950](https://doi.org/10.1145/3439950)]
89. Li G, Jung JJ. Deep learning for anomaly detection in multivariate time series: approaches, applications, and challenges. *Inf Fusion*. 2023;91:93-102. [doi: [10.1016/j.inffus.2022.10.008](https://doi.org/10.1016/j.inffus.2022.10.008)]
90. Wang R, Chen F, Chen Z, Li T, Harari G, Tignor S, et al. StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. Association for Computing Machinery; 2014. Presented at: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing; 2014 Jun 10:3-14; New York, NY, USA. [doi: [10.1145/2632048.2632054](https://doi.org/10.1145/2632048.2632054)]

91. Yao S, Hu S, Zhao Y, Zhang A, Abdelzaher T. DeepSense: A Unified Deep Learning Framework for Time-Series Mobile Sensing Data Processing. International World Wide Web Conferences Steering Committee; 2017. Presented at: Proceedings of the 26th International Conference on World Wide Web; 2017 Apr 03:351-360; Geneva. [doi: [10.1145/3038912.3052577](https://doi.org/10.1145/3038912.3052577)]
92. Mattingly SM, Gregg JM, Audia P, Bayraktaroglu AE, Campbell A, Chawla N, et al. The tesserae project: large-scale, longitudinal, in situ, multimodal sensing of information workers. Association for Computing Machinery; 2019. Presented at: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems; 2019 May 03:1-8; New York, NY, USA. [doi: [10.1145/3290607.3299041](https://doi.org/10.1145/3290607.3299041)]
93. Xu X, Liu X, Zhang H, Wang W, Nepal S, Sefidgar Y, et al. GLOBEM: Cross-dataset generalization of longitudinal human behavior modeling. Association for Computing Machinery; 2023. Presented at: Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies; December 2022:1-34; New York, NY, United States. [doi: [10.1145/3569485](https://doi.org/10.1145/3569485)]
94. Coley RY, Liao Q, Simon N, Shortreed SM. Empirical evaluation of internal validation methods for prediction in large-scale clinical data with rare-event outcomes: a case study in suicide risk prediction. BMC Med Res Methodol. 2023;23(1):33. [doi: [10.1186/s12874-023-01844-5](https://doi.org/10.1186/s12874-023-01844-5)] [Medline: [36721082](https://pubmed.ncbi.nlm.nih.gov/36721082/)]
95. Pillai A, Nepal S, Campbell A. Rare life event detection via mobile sensing using multi-task learning. arXiv. 2023. [doi: [10.48550/arXiv.2305.20056](https://doi.org/10.48550/arXiv.2305.20056)]
96. Xu X, Zhang H, Sefidgar Y, Ren Y, Liu X, Seo W, et al. GLOBEM dataset: multi-year datasets for longitudinal human behavior modeling generalization. arXiv. 2023. [doi: [10.1145/3686138.3686147](https://doi.org/10.1145/3686138.3686147)]
97. Adhikari S, Normand SL, Bloom J, Shahian D, Rose S. Revisiting performance metrics for prediction with rare outcomes. Stat Methods Med Res. 2021;30(10):2352-2366. [FREE Full text] [doi: [10.1177/09622802211038754](https://doi.org/10.1177/09622802211038754)] [Medline: [34468239](https://pubmed.ncbi.nlm.nih.gov/34468239/)]
98. Ferreira D, Kostakos V, Dey AK. AWARE: mobile context instrumentation framework. Front ICT. 2015;2:2. [doi: [10.3389/fict.2015.00006](https://doi.org/10.3389/fict.2015.00006)]
99. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol. 2008;4:1-32. [doi: [10.1146/annurev.clinpsy.3.022806.091415](https://doi.org/10.1146/annurev.clinpsy.3.022806.091415)] [Medline: [18509902](https://pubmed.ncbi.nlm.nih.gov/18509902/)]
100. Doryab A, Chikarsel P, Liu X, Dey AK. Extraction of behavioral features from smartphone and wearable data. arXiv. 2019;21(2):8-10. [doi: [10.5260/chara.21.2.8](https://doi.org/10.5260/chara.21.2.8)]
101. Wang R, Harari G, Hao P, Zhou X, Campbell A. SmartGPA: how smartphones can assess and predict academic performance of college students. Association for Computing Machinery; 2015. Presented at: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing; 2015 Sep 11:295-306; New York, NY, USA. [doi: [10.1145/2750858.2804251](https://doi.org/10.1145/2750858.2804251)]
102. Saeb S, Lonini L, Jayaraman A, Mohr DC, Kording KP. The need to approximate the use-case in clinical machine learning. Gigascience. 2017;6(5):1-9. [FREE Full text] [doi: [10.1093/gigascience/gix019](https://doi.org/10.1093/gigascience/gix019)] [Medline: [28327985](https://pubmed.ncbi.nlm.nih.gov/28327985/)]
103. Gu X, Deligianni F, Han J, Liu X, Chen W, Yang GZ, et al. Beyond supervised learning for pervasive healthcare. IEEE Rev Biomed Eng. 2024;17:42-62. [doi: [10.1109/rbme.2023.3296938](https://doi.org/10.1109/rbme.2023.3296938)]
104. Meegahapola L, Droz W, Kun P, de Götzen A, Nutakki C, Diwakar S, et al. Generalization and personalization of mobile sensing-based mood inference models. Association for Computing Machinery; 2023. Presented at: Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies; December 2022:1-32; New York, NY, United States. [doi: [10.1145/3569483](https://doi.org/10.1145/3569483)]
105. Woll S, Birkenmaier D, Biri G, Nissen R, Lutz L, Schroth M, et al. Applying AI in the context of the association between device-based assessment of physical activity and mental health: systematic review. JMIR Mhealth Uhealth. 2025;13:e59660. [doi: [10.2196/59660](https://doi.org/10.2196/59660)] [Medline: [40053765](https://pubmed.ncbi.nlm.nih.gov/40053765/)]
106. Japkowicz N, Stephen S. The class imbalance problem: a systematic study. Intell Data Anal. 2002;6(5):429-449. [doi: [10.3233/ida-2002-6504](https://doi.org/10.3233/ida-2002-6504)]
107. Kim Y, Jung J, Na J. Socioeconomic status differences in psychological responses to unfair treatments: behavioral evidence of a vicious cycle. PLoS One. 2022;17(6):e0268286. [FREE Full text] [doi: [10.1371/journal.pone.0268286](https://doi.org/10.1371/journal.pone.0268286)] [Medline: [35687607](https://pubmed.ncbi.nlm.nih.gov/35687607/)]
108. Barclay LJ, Kiefer T. In the aftermath of unfair events: understanding the differential effects of anxiety and anger. J Manag. 2017;45(5):1802-1829. [doi: [10.1177/0149206317739107](https://doi.org/10.1177/0149206317739107)]
109. Pascoe EA, Smart Richman L. Perceived discrimination and health: a meta-analytic review. Psychol Bull. 2009;135(4):531-554. [FREE Full text] [doi: [10.1037/a0016059](https://doi.org/10.1037/a0016059)] [Medline: [19586161](https://pubmed.ncbi.nlm.nih.gov/19586161/)]
110. Williams DR, Mohammed SA. Discrimination and racial disparities in health: evidence and needed research. J Behav Med. 2009;32(1):20-47. [FREE Full text] [doi: [10.1007/s10865-008-9185-0](https://doi.org/10.1007/s10865-008-9185-0)] [Medline: [19030981](https://pubmed.ncbi.nlm.nih.gov/19030981/)]
111. Sloan MM. Unfair treatment in the workplace and worker well-being. Work Occup. 2011;39(1):3-34. [doi: [10.1177/0730888411406555](https://doi.org/10.1177/0730888411406555)]
112. Cariello AN, Perrin PB, Williams CD, Espinoza GA, Paredes AM, Moreno OA. Moderating influence of social support on the relations between discrimination and health via depression in Latinx immigrants. J Lat Psychol. 2022;10(2):98-111. [FREE Full text] [doi: [10.1037/lat0000200](https://doi.org/10.1037/lat0000200)] [Medline: [35434535](https://pubmed.ncbi.nlm.nih.gov/35434535/)]
113. Moore J, Tilki M, Clarke L, Waters E. The moderating effect of functional social support on the association between unfair treatment and self-rated health: a study of the resilience of a community-based sample of Irish migrants in London. Ir J Sociol. 2018;26(3):267-288. [doi: [10.1177/0791603518811775](https://doi.org/10.1177/0791603518811775)]

114. Xu YE, Chopik WJ. Identifying moderators in the link between workplace discrimination and health/well-being. *Front Psychol*. 2020;11:458. [FREE Full text] [doi: [10.3389/fpsyg.2020.00458](https://doi.org/10.3389/fpsyg.2020.00458)] [Medline: [32256433](https://pubmed.ncbi.nlm.nih.gov/32256433/)]
115. Meier LL, Semmer NK, Hupfeld J. The impact of unfair treatment on depressive mood: the moderating role of self-esteem level and self-esteem instability. *Pers Soc Psychol Bull*. 2009;35(5):643-655. [doi: [10.1177/0146167208331337](https://doi.org/10.1177/0146167208331337)] [Medline: [19244135](https://pubmed.ncbi.nlm.nih.gov/19244135/)]
116. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. LightGBM: a highly efficient gradient boosting decision tree. 2017. Presented at: 31st Conference on Neural Information Processing Systems; 04 December 2017; Long Beach, CA, USA. URL: https://proceedings.neurips.cc/paper_files/paper/2017/hash/6449f44a102fde848669bdd9eb6b76fa-Abstract.html
117. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in python. *J Mach Learn Res*. 2011;12(85):2825-2830. [FREE Full text] [doi: [10.5555/1953048.2078195](https://doi.org/10.5555/1953048.2078195)]
118. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. *J Artif Intell Res*. 2002;16:321-357. [doi: [10.1613/jair.953](https://doi.org/10.1613/jair.953)]
119. Nguyen HM, Cooper EW, Kamei K. Borderline over-sampling for imbalanced data classification. *Int J Knowl Eng Soft Data Paradigm*. 2011;3(1):4. [doi: [10.1504/ijkesdp.2011.039875](https://doi.org/10.1504/ijkesdp.2011.039875)]
120. Chandola V, Banerjee A, Kumar V. Anomaly detection: a survey. *ACM Comput Surv*. 2009;41(3):1-58. [doi: [10.1145/1541880.1541882](https://doi.org/10.1145/1541880.1541882)]
121. Pimentel MAF, Clifton DA, Clifton L, Tarassenko L. A review of novelty detection. *Signal Process*. 2014;99:215-249. [doi: [10.1016/j.sigpro.2013.12.026](https://doi.org/10.1016/j.sigpro.2013.12.026)]
122. Schölkopf B, Platt JC, Shawe-Taylor J, Smola AJ, Williamson RC. Estimating the support of a high-dimensional distribution. *Neural Comput*. 2001;13(7):1443-1471. [doi: [10.1162/089976601750264965](https://doi.org/10.1162/089976601750264965)] [Medline: [11440593](https://pubmed.ncbi.nlm.nih.gov/11440593/)]
123. Sofaer HR, Hoeting JA, Jarnevich CS. The area under the precision - recall curve as a performance metric for rare binary events. *Methods Ecol Evol*. 2019;10(4):565-577. [doi: [10.1111/2041-210x.13140](https://doi.org/10.1111/2041-210x.13140)]
124. Cook J, Ramadas V. When to consult precision-recall curves. *Stata J*. Mar 24, 2020;20(1):131-148. [FREE Full text] [doi: [10.1177/1536867X20909693](https://doi.org/10.1177/1536867X20909693)]
125. Salih AM. Explainable artificial intelligence and multicollinearity?: a mini review of current approaches. *arXiv*. 2024. [doi: [10.48550/arXiv.2406.11524](https://doi.org/10.48550/arXiv.2406.11524)]
126. Dube L, Verster T. Interpretability of the random forest model under class imbalance. *Data Sci Finance Econ*. 2024;4(3):446-468. [doi: [10.3934/dsfe.2024019](https://doi.org/10.3934/dsfe.2024019)]
127. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Curran Associates Inc; 2017. Presented at: Proceedings for the 31st International Conference on Neural Information Processing Systems; 2017 Dec 04:4768-4777; NY, USA.
128. Kaur H, Pannu HS, Malhi AK. A systematic review on imbalanced data challenges in machine learning. *ACM Comput Surv*. 2019;52(4):1-36. [doi: [10.1145/3343440](https://doi.org/10.1145/3343440)]
129. Haixiang G, Yijing L, Shang J, Mingyun G, Yuanyue H, Bing G. Learning from class-imbalanced data: review of methods and applications. *Expert Syst Appl*. 2017;73:220-239. [doi: [10.1016/j.eswa.2016.12.035](https://doi.org/10.1016/j.eswa.2016.12.035)]
130. Tanha J, Abdi Y, Samadi N, Razzaghi N, Asadpour M. Boosting methods for multi-class imbalanced data classification: an experimental review. *J Big Data*. 2020;7(1):70. [doi: [10.1186/s40537-020-00349-y](https://doi.org/10.1186/s40537-020-00349-y)]
131. Liu L, Wu X, Li S, Li Y, Tan S, Bai Y. Solving the class imbalance problem using ensemble algorithm: application of screening for aortic dissection. *BMC Med Inform Decis Mak*. 2022;22(1):82. [doi: [10.1186/s12911-022-01821-w](https://doi.org/10.1186/s12911-022-01821-w)] [Medline: [35346181](https://pubmed.ncbi.nlm.nih.gov/35346181/)]
132. Khan AA, Chaudhari O, Chandra R. A review of ensemble learning and data augmentation models for class imbalanced problems: combination, implementation and evaluation. *Expert Syst Appl*. 2024;244:122778. [doi: [10.1016/j.eswa.2023.122778](https://doi.org/10.1016/j.eswa.2023.122778)]
133. Doryab A, Villalba DK, Chikersal P, Dutcher JM, Tumminia M, Liu X, et al. Identifying behavioral phenotypes of loneliness and social isolation with passive sensing: statistical analysis, data mining and machine learning of smartphone and fitbit data. *JMIR Mhealth Uhealth*. 2019;7(7):e13209. [FREE Full text] [doi: [10.2196/13209](https://doi.org/10.2196/13209)] [Medline: [31342903](https://pubmed.ncbi.nlm.nih.gov/31342903/)]
134. Saeb S, Zhang M, Karr CJ, Schueller SM, Corden ME, Kording KP, et al. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. *J Med Internet Res*. 2015;17(7):e175. [FREE Full text] [doi: [10.2196/jmir.4273](https://doi.org/10.2196/jmir.4273)] [Medline: [26180009](https://pubmed.ncbi.nlm.nih.gov/26180009/)]
135. Verkuyten M, Thijs J, Gharaei N. Discrimination and academic (dis)engagement of ethnic-racial minority students: a social identity threat perspective. *Soc Psychol Educ*. 2019;22(2):267-290. [doi: [10.1007/s11218-018-09476-0](https://doi.org/10.1007/s11218-018-09476-0)]
136. Smith HJ, Jaurique A, Ryan D. The mistreatment of others: discrimination can undermine university identification, student health, and engagement. *Soc Just Res*. 2016;29(4):355-374. [doi: [10.1007/s11211-016-0274-x](https://doi.org/10.1007/s11211-016-0274-x)]
137. Nakhaie R. Discrimination, psychological isolation, and flight from school. *J Int Migr Integr*. 2021;23(3):1515-1541. [doi: [10.1007/s12134-021-00901-6](https://doi.org/10.1007/s12134-021-00901-6)]
138. Gyan C, Baskh B, Song W, Yeboah AS. "Withdrawal Syndrome": the effects of acts of microaggression in the classroom on racialized students. *Appl Res Qual Life*. 2024;19(6):3169-3187. [doi: [10.1007/s11482-024-10373-2](https://doi.org/10.1007/s11482-024-10373-2)]

139. Diehl D, Houseworth J, Grier-Reed T. Examining the variable relationship between race and considerations of campus withdrawal. *Coll Stud J*. 2020;53(4):417-429. [[FREE Full text](#)]
140. Martins Neto C, Confortin SC, Lima ABS, Mouzinho LSN, Oliveira BLCAD. Association between perceived discrimination and physical activity among adolescents. *Ciênc Saúde Coletiva*. 2022;27(10):4003-4013. [doi: [10.1590/1413-812320222710.05992022en](https://doi.org/10.1590/1413-812320222710.05992022en)]
141. Sims M, Diez-Roux AV, Gebreab SY, Brenner A, Dubbert P, Wyatt S, et al. Perceived discrimination is associated with health behaviours among African-Americans in the Jackson heart study. *J Epidemiol Community Health*. 2016;70(2):187-194. [doi: [10.1136/jech-2015-206390](https://doi.org/10.1136/jech-2015-206390)] [Medline: [26417003](https://pubmed.ncbi.nlm.nih.gov/26417003/)]
142. Zhu J, Xie R, Chen Y, Zhang W. Relationship between parental rejection and problematic mobile phone use in Chinese university students: mediating roles of perceived discrimination and school engagement. *Front Psychol*. 2019;10:428. [[FREE Full text](#)] [doi: [10.3389/fpsyg.2019.00428](https://doi.org/10.3389/fpsyg.2019.00428)] [Medline: [30890982](https://pubmed.ncbi.nlm.nih.gov/30890982/)]
143. Guan W, Wang S, Liu C. Influence of perceived discrimination on problematic smartphone use among Chinese deaf and hard-of-hearing students: serial mediating effects of sense of security and social avoidance. *Addict Behav*. 2023;136:107470. [doi: [10.1016/j.addbeh.2022.107470](https://doi.org/10.1016/j.addbeh.2022.107470)] [Medline: [36055057](https://pubmed.ncbi.nlm.nih.gov/36055057/)]
144. Li W, Xu T, Diao L, Wu Q. The impact of perceived discrimination on mobile phone addiction among Chinese higher vocational college students: a chain mediating role of negative emotions and learning burnout. *Psychol Res Behav Manag*. 2024;17:401-411. [[FREE Full text](#)] [doi: [10.2147/PRBM.S440958](https://doi.org/10.2147/PRBM.S440958)] [Medline: [38343428](https://pubmed.ncbi.nlm.nih.gov/38343428/)]
145. Slopen N, Lewis TT, Williams DR. Discrimination and sleep: a systematic review. *Sleep Med*. 2016;18:88-95. [[FREE Full text](#)] [doi: [10.1016/j.sleep.2015.01.012](https://doi.org/10.1016/j.sleep.2015.01.012)] [Medline: [25770043](https://pubmed.ncbi.nlm.nih.gov/25770043/)]
146. Park K, Kim J. Longitudinal association between perceived discrimination and sleep problems among young adults in the United States: tests of moderation by race/ethnicity and educational attainment. *Soc Sci Med*. 2023;321:115773. [doi: [10.1016/j.socscimed.2023.115773](https://doi.org/10.1016/j.socscimed.2023.115773)] [Medline: [36801753](https://pubmed.ncbi.nlm.nih.gov/36801753/)]
147. Majeno A, Tsai KM, Huynh VW, McCreath H, Fuligni AJ. Discrimination and sleep difficulties during adolescence: the mediating roles of loneliness and perceived stress. *J Youth Adolesc*. 2018;47(1):135-147. [[FREE Full text](#)] [doi: [10.1007/s10964-017-0755-8](https://doi.org/10.1007/s10964-017-0755-8)] [Medline: [29164378](https://pubmed.ncbi.nlm.nih.gov/29164378/)]
148. Johnson DA, Lewis TT, Guo N, Jackson CL, Sims M, Wilson JG, et al. Associations between everyday discrimination and sleep quality and duration among African-Americans over time in the Jackson heart study. *Sleep*. 2021;44(12):zsab162. [[FREE Full text](#)] [doi: [10.1093/sleep/zsab162](https://doi.org/10.1093/sleep/zsab162)] [Medline: [34197610](https://pubmed.ncbi.nlm.nih.gov/34197610/)]
149. Gordon AM, Prather AA, Dover T, Espino-Pérez K, Small P, Major B. Anticipated and experienced ethnic/racial discrimination and sleep: a longitudinal study. *Pers Soc Psychol Bull*. 2020;46(12):1724-1735. [doi: [10.1177/0146167220928859](https://doi.org/10.1177/0146167220928859)] [Medline: [32571161](https://pubmed.ncbi.nlm.nih.gov/32571161/)]
150. Fuller-Rowell TE, Nichols OI, Burrow AL, Ong AD, Chae DH, El-Sheikh M. Day-to-day fluctuations in experiences of discrimination: associations with sleep and the moderating role of internalized racism among African American college students. *Cultur Divers Ethnic Minor Psychol*. 2021;27(1):107-117. [doi: [10.1037/cdp0000342](https://doi.org/10.1037/cdp0000342)] [Medline: [32309971](https://pubmed.ncbi.nlm.nih.gov/32309971/)]
151. Nahum-Shani I, Smith SN, Spring BJ, Collins LM, Witkiewitz K, Tewari A, et al. Just-in-time adaptive interventions (JITAs) in mobile health: key components and design principles for ongoing health behavior support. *Ann Behav Med*. 2018;52(6):446-462. [[FREE Full text](#)] [doi: [10.1007/s12160-016-9830-8](https://doi.org/10.1007/s12160-016-9830-8)] [Medline: [27663578](https://pubmed.ncbi.nlm.nih.gov/27663578/)]
152. Hsu TC, Whelan P, Gandrup J, Armitage CJ, Cordingley L, McBeth J. Personalized interventions for behaviour change: a scoping review of just-in-time adaptive interventions. *Br J Health Psychol*. 2025;30(1):e12766. [doi: [10.1111/bjhp.12766](https://doi.org/10.1111/bjhp.12766)] [Medline: [39542743](https://pubmed.ncbi.nlm.nih.gov/39542743/)]
153. Carpenter SM, Menictas M, Nahum-Shani I, Wetter DW, Murphy SA. Developments in mobile health just-in-time adaptive interventions for addiction science. *Curr Addict Rep*. 2020;7(3):280-290. [[FREE Full text](#)] [doi: [10.1007/s40429-020-00322-y](https://doi.org/10.1007/s40429-020-00322-y)] [Medline: [33747711](https://pubmed.ncbi.nlm.nih.gov/33747711/)]
154. Improving engagement and efficacy of mHealth micro-interventions for stress coping: an in-the-wild study. arxiv. URL: <https://arxiv.org/abs/2407.11612> [accessed 2025-05-10]
155. Dong Y, Pinelli F, Gkoufas Y, Nabi Z, Calabrese F, Chawla N. Inferring unusual crowd events from mobile phone call detail records. In: *Machine Learning and Knowledge Discovery in Databases*. Springer, Cham. Springer International Publishing; 2015:474-492.
156. Cheon SP, Kim S, Lee SY, Lee CB. Bayesian networks based rare event prediction with sensor data. *Knowl-Based Syst*. 2009;22(5):336-343. [doi: [10.1016/j.knosys.2009.02.004](https://doi.org/10.1016/j.knosys.2009.02.004)]
157. Dangut MD, Jennions IK, King S, Skaf Z. Application of deep reinforcement learning for extremely rare failure prediction in aircraft maintenance. *Mech Syst Signal Process*. 2022;171:108873. [doi: [10.1016/j.ymsp.2022.108873](https://doi.org/10.1016/j.ymsp.2022.108873)]
158. Wong ZSY. Statistical classification of drug incidents due to look-alike sound-alike mix-ups. *Health Inf J*. 2016;22(2):276-292. [[FREE Full text](#)] [doi: [10.1177/1460458214555040](https://doi.org/10.1177/1460458214555040)] [Medline: [25391848](https://pubmed.ncbi.nlm.nih.gov/25391848/)]
159. Generalization of longitudinal behavior modeling. GLOBEM. URL: <https://the-globem.github.io> [accessed 2025-08-15]

Abbreviations

AE: autoencoder

AUC-PR: area under the precision-recall curve
AUC-ROC: area under the receiver operating characteristic curve
EMA: ecological momentary assessment
KNN: k-nearest neighbors
LightGBM: light gradient boosting machine
LSTM: long short-term memory
PUT: perceived unfair treatment
RED: rare event detection
SMD: standardized mean difference
SVM: support vector machine

Edited by A Mavragani; submitted 11.Jun.2025; peer-reviewed by GE Fronk; comments to author 10.Jul.2025; revised version received 03.Sep.2025; accepted 21.Sep.2025; published 31.Oct.2025

Please cite as:

Ren Y, Mulukutla R, Mankoff J, Dey AK

Detecting Perceived Unfair Treatment Among US College Students Using Mobile Sensing: Pilot Machine Learning Study

JMIR Form Res 2025;9:e78657

URL: <https://formative.jmir.org/2025/1/e78657>

doi: [10.2196/78657](https://doi.org/10.2196/78657)

PMID:

©Yiyi Ren, Raghu Mulukutla, Jennifer Mankoff, Anind K Dey. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 31.Oct.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.