Analysis of Sleep Indicators and State Determinants Using Machine Learning with Lifelog Data

Jun-Ho Park
Graduate School of Management Consulting
Hanyang University
Seoul, South Korea
ict0624@gmail.com

Seung-Hyun Baek*
Division of Business Administration
Hanyang University ERICA
Ansan, South Korea
sbaek4@hanyang.ac.kr

***Abstract*—This study comprehensively analyzed the impact of physiological, behavioral, digital, environmental, and social factors on sleep using lifelog data from smart devices. The most critical finding is that digital factors, such as mobile device usage, emerged as the strongest predictors across all sleep indicators. Additionally, environmental factors showed extensive influence on sleep quality, while social factors selectively contributed to sleep duration determination. Conversely, physiological and behavioral factors, traditionally considered important, showed relatively lower predictive power, demonstrating the need for comprehensive analysis beyond simple biological signals. In conclusion, this research provides important scientific evidence for developing personalized digital sleep health management systems that consider modern lifestyle patterns.**

Keywords—Lifelog, Tree-based Classification, Sleep Quality, Sleep Indicators, Multi-sensor Data, Permutation Importance

I. Introduction

Sleep is essential for physical recovery and mental health, but sleep quality deterioration has become a significant public health issue in modern society. Previous studies have shown achievements in analyzing sleep through actigraphy or specific biosignals [1],[2], but these approaches have apparent limitations in their dependence on single-sensor data. In particular, there has been insufficient analysis of complex influences surrounding individual lives, such as environmental, digital, and social factors.

This study proposes the necessity of utilizing multi-sensor lifelog data that encompasses the overall daily life of individuals. Using decision tree-based classification models, we aim to quantitatively identify the key determinants of how physiological, behavioral, digital, environmental, and social factors affect sleep. Additionally, we will systematically analyze the contribution of each factor through time window-based feature engineering and permutation importance-based statistical validation.

This approach is expected to provide important scientific evidence for developing personalized sleep health management systems and digital therapeutic platforms.

II. Methodology

A. Datasets

This study utilized the multi-sensor lifelog dataset provided by the 2025 Human Understanding AI Paper Challenge [3]. The dataset consists of various sensor data collected through participants' Android smartphones and smartwatches, along with daily survey results, comprehensively reflecting individuals' physiological, behavioral, digital, environmental, and social activities.

The collected sensor data is categorized into smartphone-based sensors and smartwatch-based sensors, with data collection intervals ranging from several seconds to 10 minutes. All timestamps are recorded in Korea Standard Time (KST) format as YYYY-MM-DD HH:MM:SS.

Smartphone sensor data consists of 9 types as shown in Table I.

TABLE I. Smartphone-Based Sensor Data Configuration

Sensor	Feature	Description
mACStatus	m_charging	Charging status (0: not charging, 1: Charging)
mActivity	m_activity	Activity classification (0: vehicle, 1: bicycle, 2: walking, 3: still, 4: unknown, 5: tilting, 7: walking, 8: running)
mAmbience	m_ambience	Ambient noise label and probability value list
mBle	m_ble	Bluetooth device address, device class, RSSI value
mGps	m_gps	Altitude, latitude, longitude, speed information
mLight	m_light	Ambient illuminance (lx units)
mScreenStatus	m_screen_use	Screen usage status (0: not in use, 1: in use)
mUsageStats	m_usage_stats	App name and usage time (milliseconds)
mWifi	m_wifi	Base station ID (BSSID) and RSSI value

Smartwatch bio-signal and activity data are configured as shown in Table II.

TABLE II. Smartwatch-Based Sensor Data Configuration

Sensor	Feature	Description
wHr	heart_rate	Heart rate record list
wLight	w_light	Ambient illuminance (lx units)
wPedo	burned_calories	Calories burned
	distance	Distance traveled (meters)
	speed	Speed (km/h units)
	step	Step count
	step_frequency	Steps per minute frequency

Sleep-related dependent variables consist of 6 metrics. These are divided into two categories: objective sleep indicators (S1, S2, S3) and subjective sleep states (Q1, Q2, Q3). Subjective sleep states were binary classified based on the average of individual survey responses. Objective sleep indicators were classified based on compliance with National Sleep Foundation guidelines.

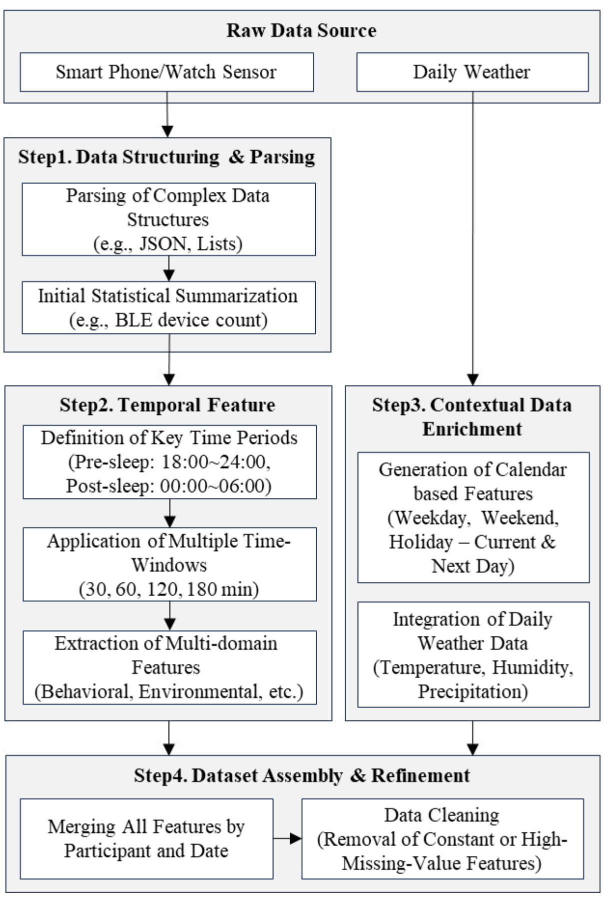
TABLE III. SLEEP-RELATED DEPENDENT VARIABLES

Metric	Category	Description	Label
Q1	Subjective	Overall sleep quality perceived immediately upon waking	0: below personal average, 1: above personal average
Q2	Subjective	Physical fatigue before sleep	0: high fatigue, 1: low fatigue
Q3	Subjective	Stress level experienced before sleep	0: high stress, 1: low stress
S1	Objective	Total Sleep Time	0: not recommended, 1: adequate, 2: recommended
S2	Objective	Sleep Efficiency	0: inadequate, 1: recommended
S3	Objective	Sleep Onset Latency	0: inadequate, 1: recommended

B. Data Pre-processing

The multi-sensor lifelog data in this study is high-dimensional raw data containing individual daily life, requiring systematic preprocessing to transform it into a form suitable for model learning to analyze relationships with sleep. This process consists of four steps as shown in Figure 1.

Fig. 1. Data Preprocessing and Variable Generation Procedure



1) *Step 1. Data structuring and parsing:*

The first step involves transforming various raw sensor data into a standardized tabular format suitable for analysis and ensuring consistency by integrating data from multiple sources. In this process, complex data structures recorded in list format, such as Bluetooth sensor data, are parsed to generate meaningful summary variables such as unique device counts and RSSI statistical values. Additionally, ambient noise data is converted from noise labels and probability values to the highest probability noise label and corresponding dB values, as shown in Table IV.

TABLE IV. AMBIENT NOISE DATA PROCESSING

Category	Data Format	Description
Before	timestamp: 2025-07-13 20:45:00 m_ambience: [['Traffic', 0.85], ['Conversation', 0.65]]	Noise labels and probability values recorded in list format
After	timestamp: 2025-07-13 20:45:00 m_ambience_cat: Traffic m_ambience_db: 75	Parse the list to generate meaningful summary variables such as unique device count at that time point and number of devices with strong RSSI signals.

Activity type (mActivity) data converts numeric codes to categorical variables and MET (Metabolic Equivalent of Task) values. Table V shows this conversion process and MET value mapping.

TABLE V. MET VALUE MAPPING BY ACTIVITY TYPE

Category	Data Format	Description
Before	timestamp: 2025-07-13 19:00:00 m_activity: 7	Activity type recorded as numeric code
After	timestamp: 2025-07-13 19:00:00 m_activity_cat: WALKING m_activity_met: 3.5	Convert activity type to categorical variable and MET (Metabolic Equivalent of Task) value

MET values by activity type are assigned as follows: vehicle movement 1.3, bicycle movement 8.0, stationary 1.2, unknown 3.0, walking 3.5, running 10.0.

For GPS data, individual coordinate information is converted to GPS VSD (Variability of Spatial Displacement), a location variability indicator within time windows, to quantify movement patterns as shown in equation (1).

$$VSD_t = \sigma_{alt}(t) \times \sigma_{lat}(t) \times \sigma_{lon}(t) \qquad (1)$$

Where VSD_t is the spatial variability index at time t, and $\sigma_{alt}(t)$, $\sigma_{lat}(t)$, $\sigma_{lon}(t)$ are the standard deviations of altitude, latitude, and longitude within the time window, respectively.

Illuminance data collected separately from smartphones (mLight) and smartwatches (wLight) show different measured values due to device wearing positions and exposure conditions. Since there is no guarantee that participants always wear both devices or that sensors are positioned under appropriate light exposure conditions, data loss and measurement inconsistency issues occur. To address this, a single continuous illuminance stream (uLight) was generated by integrating data from both sources based on chronological order. Table VI shows this integration process.

TABLE VI. INTEGRATED LIGHT DATA GENERATION EXAMPLE

Category	Data Format	Description
Before	mLight: 150 lx (smartphone) wLight: 0 lx (watch in sleeve)	Different illuminance values measured by each device
After	uLight: 150 lx	Continuous illuminance stream generated by combining two sources in chronological order

2) Step 2. Multi-dimensional derived variable generation based on time windows : Sleep quality and state are the result of accumulated activities, environmental exposure, and physiological states over several hours. Days are divided into evening (18:00-24:00) and dawn (00:00-06:00), and statistical features are extracted by segmenting into four-time windows: 30, 60, 120, and 180 minutes. The core of time window-based feature extraction is calculating the mean and standard deviation of each sensor data as shown in equations (2) and (3).

$$F_{S,W}^{mean} = \frac{1}{|W|}\sum_{t\in W} S(t) \tag{2}$$

$$F_{S,W}^{std} = \sqrt{\frac{1}{|W|-1}\sum_{t\in W}\left(S(t) - F_{S,W}^{mean}\right)^2} \tag{3}$$

In equations (2) and (3), $F_{S,W}^{mean}$ and $F_{S,W}^{std}$ represent the mean and standard deviation of sensor S in window W, respectively, and $|W|$ represents the number of data points in window W. Through this statistical summary, both central tendency and variability of sensor data by time interval can be captured simultaneously.

3) Step 3. Integration of external data and temporal context variables: Weather data and temporal context variables are integrated to build a more comprehensive analysis dataset. 21 types of daily weather data provided by the Korea Meteorological Administration are integrated based on lifelog recording dates, and calendar-based variables are generated to reflect the basic cycle of social activities - weekly characteristics.

4) Step 4. Final dataset construction and refinement: All derived variables are integrated into a unified analysis dataset using participant individual identifiers and lifelog recording dates as common keys. In the final refinement process, constant variables and variables with missing value rates exceeding 80% are removed. Table VII shows an example of the final dataset.

TABLE VII. CONCEPTUAL EXAMPLE OF FINAL ANALYSIS DATASET

subject_id	Lifelog_date	PreSleep&u_light&20h00m&120min&median	PreSleep&ambience_db&21h00m&60min&mean	…	Q1
id01	2024-06-26	150.2	68.5	…	0
id01	2024-06-27	145.8	72.1	…	0
id01	2024-06-28	162.4	65.2	…	1

Feature names are structured as time_period & sensor_type & start_time & window_size & statistical_function. Thus, PreSleep&u_light&20h00m&120min&median is the median illuminance from 20:00 for 120 minutes before sleep. See Table VIII for all engineered features.

TABLE VIII. SUMMARY OF ENGINEERED FEATURES

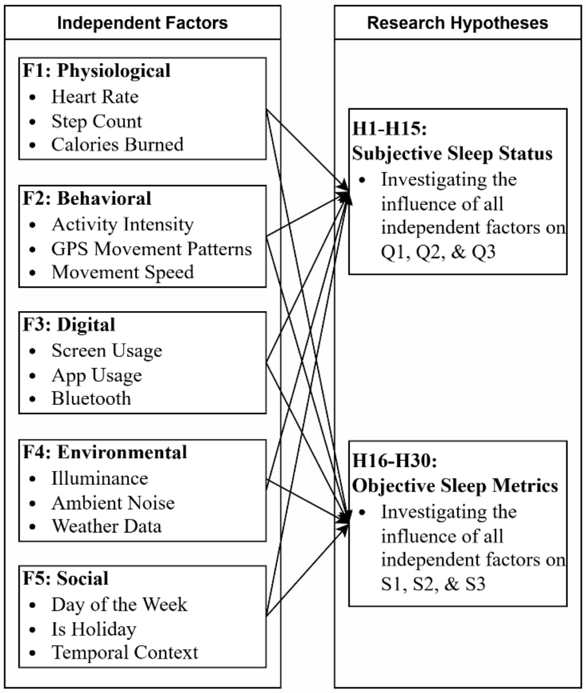
Source Data	Feature Name	Generation Method & Aggregation
mLight, wLight	Unified Illuminance	Concatenated smartphone/watch data; median over 30/60/120/180 min windows.
mAmbience	Ambient Noise (dB)	Sound labels converted to dB; mean over 30/60/120/180 min windows.
mAmbience	Ambient Noise (Type)	Mode over 30/60/120/180 min windows.
mActivity	Activity Intensity (MET)	Activity types converted to METs; mean over 30/60/120/180 min windows.
mActivity	Activity Type	Mode over 30/60/120/180 min windows.
wPedo	Step Count / Distance	Mean over 30/60/120/180 min windows.
mGps	Movement Speed	Mean over 30/60/120/180 min windows.
mGps	Mobility Variance	Product of Lat/Lon/Alt standard deviations, calculated over 30/60/120/180 min windows.
mScreenStatus	Screen Usage Ratio	Mean over 30/60/120/180 min windows (ON=1, OFF=0).
mUsageStats	App Usage (Total)	Mean of total usage duration over 30/60/120/180 min windows.
mUsageStats	App Usage (Max)	Mean of max single-app duration over 30/60/120/180 min windows.
mUsageStats	App Usage (Count)	Mean of unique apps used over 30/60/120/180 min windows.
mUsageStats	App Usage (Type)	Mode of app category over 30/60/120/180 min windows.
wHr	Heart Rate	Mean over 30/60/120/180 min windows.
mBle	BLE Device Count	Mean of unique device/class counts over 30/60/120/180 min windows.
mBle	Strong Signal BLE Count	Mean count of devices with RSSI > -70dBm over 30/60/120/180 min windows.
mWifi	WiFi AP Count	Mean of unique Access Points (APs) over 30/60/120/180 min windows.
mWifi	Strong Signal AP Count	Mean count of APs with RSSI > -67dBm over 30/60/120/180 min windows.
Meteorological Data	Temperature (6 features)	Daily values: Mean/Max/Min temp, time of max/min temp, diurnal range.
	Wind (7 features)	Daily values: Mean/Max/Gust speed, direction, time.
	Precipitation (3 features)	Daily values: Total precipitation, max hourly amount, time of max.
	Insolation (3 features)	Daily values: Sunshine duration, sunshine rate, solar radiation sum.
	Humidity (2 features)	Daily values: Mean/Min humidity.
lifelog_date	Calendar Features	Daily values: Day of the week, weekend status, holiday status.

C. Research Hypothesis Development

To verify the impact of five factors extracted from multi-dimensional lifelog data on six sleep indicators, a total of 30 research hypotheses were established. Each hypothesis was formulated as an alternative hypothesis stating "a specific factor makes a significant contribution to predicting the corresponding sleep indicator," with the null hypothesis set as "the corresponding factor does not affect sleep indicator prediction."

Figure 2 systematically shows the relationships between 5 factors (F1-F5) and 6 sleep indicators (Q1-Q3, S1-S3), visualizing the structure of 30 research hypotheses (H1-H30).

Fig. 2. Research Hypothesis Framework



The factors used in the analysis are classified into five categories. First, there are physiological factors (F1) related to biosignals such as heart rate and step count, behavioral factors (F2) covering physical activities such as activity intensity and movement patterns, and digital factors (F3) representing digital device usage such as screen use and app usage. Additionally, environmental factors (F4) include physical environment such as illuminance, noise, and temperature, and social factors (F5) represent social context such as day of the week and holiday status.

Hypotheses related to subjective sleep states (H1-H15) verify the impact of each factor on overall sleep quality (Q1), physical fatigue (Q2), and stress level (Q3), while hypotheses related to objective sleep indicators (H16-H30) aim to identify the impact of each factor on total sleep time (S1), sleep efficiency (S2), and sleep onset latency (S3).

D. Modeling and Feature Selection

To identify factors affecting sleep indicators and states, a CatBoost model based on gradient boosting was utilized. To select only meaningful core variables for prediction among numerous variables, a two-stage variable selection procedure was applied.

In the first stage, an initial model is trained with all variables to calculate the importance score of each variable. Next, only variables with importance scores above average are selected to construct the final feature set, as expressed in equations (4) and (5).

The process of selecting only features with importance scores above average from the entire feature set is shown in equation (4).

$$\mathcal{F}_{selected} = \{f_i \in \mathcal{F} : I(f_i) \geq \bar{I}\} \quad (4)$$

Where the average importance is calculated by equation (5).

$$\bar{I} = \frac{1}{n}\sum_{i=1}^{n} I(f_i) \quad (5)$$

In equations (4) and (5), $\mathcal{F}$ is the entire feature set, $\mathcal{F}_{selected}$ is the selected feature set, $I(f_i)$ is the importance score of feature f_i, $\bar{I}$ is the average importance of all features, and n is the total number of features. This procedure contributes to preventing model overfitting and improving computational efficiency.

E. Model Training and Evaluation

To objectively and reliably evaluate model performance using the finally selected core variables, a rigorous training and validation procedure was established.

Repeated Stratified 2-Fold cross-validation was performed with 5 repetitions. This design was adopted because the total number of measurement days with provided labels was only 450 days, which is relatively limited. Preliminary experiments showed that using more folds (3-fold or higher) resulted in overfitting due to excessively reduced training data size. Therefore, to compensate for validation instability due to the small number of folds, the 2-fold cross-validation was repeated 5 times, performing a total of 10 validations to ensure reliability of model performance evaluation.

Fig. 3. Repeated Stratified 2-Fold Cross-Validation Procedure

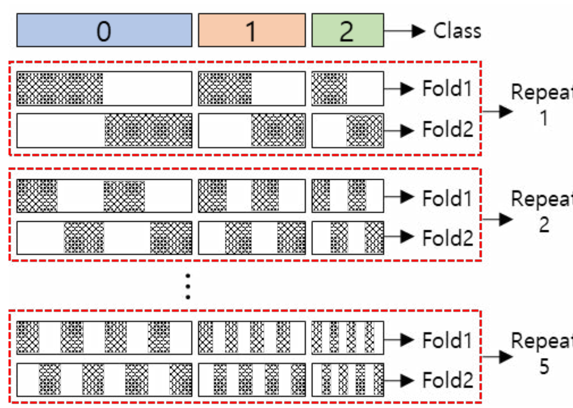


Figure 3 visualizes the repeated stratified 2-fold cross-validation process applied in class imbalance situations. Data is split while maintaining the ratio of each class (0, 1, 2), and this process ensures robust model evaluation through multiple validation rounds.

To address the class imbalance problem inherent in sleep learning data, Random OverSampling technique was applied, which randomly duplicates minority class data. Additionally, automated hyperparameter optimization based on Random Search was conducted to maximize model performance. The search process was performed to maximize Macro F1 score, which fairly evaluates performance between classes in data imbalance situations.

First, precision and recall, which are basic performance indicators of the model, are defined as in equation (6).

$$Precision = \frac{TP}{TP+FP}, \quad \text{Recall} = \frac{TP}{TP+FN} \quad (6)$$

Where TP (True Positive) represents correctly predicting actual positives as positive, FP (False Positive) represents incorrectly predicting actual negatives as positive, and FN (False Negative) represents incorrectly predicting actual positives as negative.

The F1 score, which combines precision and recall through harmonic mean, is calculated as in equation (7).

$$F_1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{7}$$

The F1 score in equation (7) is an indicator of balance between precision and recall, reaching its maximum when both indicators are high.

To evaluate overall model performance in multi-class classification problems, Macro F1 score was used as in equation (8).

$$F_{1\_macro} = \frac{1}{|C|} \sum_{c \in C} F_{1_c} \tag{8}$$

In equation (8), C is the entire class set, $|C|$ is the total number of classes, and $F1_c$ is the individual F1 score for class c. Macro F1 score evaluates the balanced performance of the entire model by equally considering the performance of each class.

Hypothesis testing was performed using a statistical methodology combining Permutation Importance technique and One-sample t-test. This method is a model-agnostic approach that quantitatively evaluates the influence of specific factors by measuring changes in model prediction performance after randomly shuffling the values of all variables belonging to a specific factor.

The importance when variables belonging to factor F are permuted is calculated as in equation (9).

$$PI_F = Score_{baseline} - Score_{permuted\_F} \tag{9}$$

In equation (9), PI_F is the permutation importance of factor F, $Score_{baseline}$ is the model accuracy on original data, and $Score_{permuted_F}$ is the model accuracy after permuting variables of factor F. The larger this difference value, the more important contribution the factor makes to model performance.

The specific verification procedure is as follows. First, the entire data is split into training set (70%) and validation set (30%), then baseline prediction accuracy is calculated for each factor variable group. Next, permutation is applied by simultaneously randomly shuffling the values of all variables belonging to a specific factor, and permutation importance is obtained by measuring the reduced new prediction accuracy. To ensure statistical reliability, this process is repeated 30 times to obtain the distribution of permutation importance, and finally, One-sample t-test is used to statistically verify whether the average importance of each factor is significantly greater than 0.

One-sample t-test for statistical significance verification is performed as in equation (10).

$$t = \frac{\overline{PI_F} - 0}{s_{PI_F}/\sqrt{n}} \tag{10}$$

In equation (10), t is the t-statistic, $\overline{PI_F}$ is the sample mean of permutation importance, s_{PI_F} is the sample standard deviation of permutation importance, and n is the number of repetitions (30 times). This test allows determining whether the influence of the factor is statistically significant or due to chance.

Statistical significance was determined at the $p < 0.05$ level, and 95% confidence intervals and effect sizes were presented together to evaluate the stability and practical meaning of effects.

III. Experiments

A. Experimental Setup

Multi-sensor lifelog data was utilized to systematically analyze the impact of physiological, behavioral, digital, environmental, and social factors on objective sleep indicators and subjective sleep states. The collected raw data was transformed into multi-dimensional derived variables through time window-based feature engineering. After removing constant variables and variables with missing rates exceeding 80%, optimal feature sets were selected for each target through CatBoost-based variable selection. Each variable was classified according to its generation principle and meaning as follows:

- **Physiological Factors:** Variables related to heart rate (wHr), and step count & distance (wPedo).
- **Behavioral Factors:** Variables related to activity intensity (mActivity) and GPS movement patterns (mGps).
- **Digital Factors:** Variables related to mobile app usage (mUsageStats), screen status (mScreenStatus), Bluetooth (ble), and Wi-Fi (wifi).
- **Environmental Factors:** Variables related to integrated illuminance (uLight), ambient noise (mAmbience), and weather data (temperature, wind, precipitation, insolation, humidity).
- **Social Factors:** Variables related to the day of the week (weekday), weekend status (is_weekend), and public holiday status (is_holiday).

B. Model Implementation and Performance

Prediction models for each sleep indicator were implemented based on the CatBoost gradient boosting algorithm. Random oversampling technique was applied to address class imbalance problems during model training, and model performance stability was ensured by performing repeated stratified 2-fold cross-validation 5 times. Hyperparameters were optimized through Random Search to maximize Macro F1 score for each prediction model.

Model performance evaluation confirmed good predictive power for most sleep indicators. Particularly, the sleep onset latency (S3) prediction model showed the highest performance with an F1 score of 0.645, while subjective sleep indicators (Q1, Q2, Q3) and sleep efficiency (S2) models recorded stable performance between 0.616 and 0.631. In contrast, the total sleep time (S1) model, which is classified into 3 classes, showed relatively low performance with an F1 score of 0.465, which is attributed to the complexity of classification due to ambiguity between sleep time categories and individual differences.

C. Results and Analysis

Statistical verification results for 30 research hypotheses showed that null hypotheses were rejected in 20 hypotheses, confirming that corresponding factors have significant effects on sleep indicator prediction. Factor-wise analysis revealed that digital factors showed the most consistent and strong influence across all 6 sleep indicators, environmental factors showed significant effects in 5 indicators, and physiological factors showed significant effects in 4 indicators. Table IX presents hypothesis testing results for subjective sleep states and objective sleep indicators.

TABLE IX. HYPOTHESIS TESTING RESULTS

Hypothesis	Target	Factor	Significance	Decision
H1	Q1	Physiological	***	Rejected
H2	Q1	Behavioral	*	Rejected
H3	Q1	Digital	***	Rejected
H4	Q1	Environmental	***	Rejected
H5	Q1	Social	n.s.	Not rejected
H6	Q2	Physiological	***	Rejected
H7	Q2	Behavioral	n.s.	Not rejected
H8	Q2	Digital	***	Rejected
H9	Q2	Environmental	***	Rejected
H10	Q2	Social	n.s.	Not rejected
H11	Q3	Physiological	n.s.	Not rejected
H12	Q3	Behavioral	n.s.	Not rejected
H13	Q3	Digital	***	Rejected
H14	Q3	Environmental	***	Rejected
H15	Q3	Social	n.s.	Not rejected
H16	S1	Physiological	***	Rejected
H17	S1	Behavioral	***	Rejected
H18	S1	Digital	***	Rejected
H19	S1	Environmental	n.s.	Not rejected
H20	S1	Social	***	Rejected
H21	S2	Physiological	n.s.	Not rejected
H22	S2	Behavioral	***	Rejected
H23	S2	Digital	***	Rejected
H24	S2	Environmental	***	Rejected
H25	S2	Social	n.s.	Not rejected
H26	S3	Physiological	***	Rejected
H27	S3	Behavioral	n.s.	Not rejected
H28	S3	Digital	***	Rejected
H29	S3	Environmental	***	Rejected
H30	S3	Social	***	Rejected

***$p<0.001$, **$p<0.01$, *$p<0.05$, n.s. = not significant

Digital factors showed consistently the highest predictive power for all sleep indicators, particularly recording the highest importance of 0.129 for total sleep time (S1) ($t=24.56$, $p<0.001$). Environmental factors showed significant effects in 5 indicators, with high importance of 0.034 for sleep onset latency (S3) ($t=14.02$, $p<0.001$). Social factors showed significant effects only for total sleep time (S1) and sleep onset latency (S3), while physiological factors were statistically significant in 4 indicators but showed relatively low importance. Behavioral factors showed significant effects in only 3 indicators, recording importance of 0.070 for total sleep time (S1) ($t=14.70$, $p<0.001$).

D. Discussion

The experimental results provide important insights that sleep quality and state are more significantly influenced by modern digital lifestyles and environmental contexts rather than traditionally emphasized physiological indicators. Particularly, the consistent high influence of digital factors suggests that smartphone usage patterns have emerged as key determinants of sleep health.

The extensive influence of environmental factors reconfirms the importance of sleep environment optimization, while the selective influence of social factors shows that individual sleep patterns interact complexly with social timetables. Conversely, the relatively low predictive power of physiological factors suggests the need for comprehensive lifestyle analysis beyond simple biosignal monitoring.

These findings contrast with existing wearable-based sleep analysis that focused primarily on biosignals, demonstrating the superiority of comprehensive approaches integrating individual digital behavior patterns and environmental contexts. The experimental approach and statistical verification results provide practical guidelines for developing personalized sleep health management systems, particularly highlighting the importance of digital behavior monitoring and environmental sensing functions.

IV. CONCLUSION

This study empirically identified that digital lifestyles and physical environments, rather than traditional physiological indicators, are the key factors determining modern sleep health through integrated analysis of multi-sensor lifelog data. Digital factors represented by smartphone usage patterns showed overwhelming influence in predicting all sleep indicators, while environmental factors such as illuminance and noise also had significant effects. Conversely, physiological factors such as heart rate and activity level showed relatively limited predictive power.

These findings suggest the need for a paradigm shift in sleep management beyond simple bio-signal monitoring to comprehensively consider individual lifestyle habits and surrounding environments. This study presents a methodology for systematically analyzing multi-sensor data and provides guidelines for personalized sleep management system design and wearable device improvements.

In conclusion, this research laid the foundation for developing practical solutions to solve modern sleep problems through comprehensive lifestyle analysis.

REFERENCES

[1] C. Lang, N. Kalak, S. Brand, E. Holsboer-Trachsler, U. Pühse, and M. Gerber, "The relationship between physical activity and sleep from mid adolescence to early adulthood: A systematic review of methodological approaches and meta-analysis," Sleep Medicine Reviews, vol. 28, pp. 32-45, 2016.

[2] A. Sathyanarayana, S. Joty, L. Fernandez-Luque, F. Ofli, J. Srivastava, A. Elmagarmid, T. Arora, and S. Taheri, "Sleep quality prediction from wearable data using deep learning," JMIR mHealth and uHealth, vol. 4, no. 4, e6562, 2016.

[3] ETRI, “ETRI Lifelog Dataset 2024,” 2025. [Online]. Available: https://nanum.etri.re.kr/share/human/ETRILifelogDataset2024

[4] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, "CatBoost: Unbiased boosting with categorical features," in Advances in Neural Information Processing Systems, vol. 31, 2018.

[5] M. Hirshkowitz, K. Whiton, S. M. Albert, C. Alessi, O. Bruni, L. DonCarlos, N. Hazen, J. Herman, E. S. Katz, L. Kheirandish-Gozal, D. N. Neubauer, A. E. O'Donnell, M. Ohayon, J. Peever, R. Rawding, R. C. Sachdeva, B. Setters, M. V. Vitiello, J. C. Ware, and P. J. Adams Hillard, "National Sleep Foundation's sleep time duration recommendations: Methodology and results summary," Sleep Health, vol. 1, no. 1, pp. 40-43, 2015.