May 4th, 11:00 AM - 1:00 PM

Understanding Work and Sleep Through A Machine Learning Approach

Jennifer Saucedo
Portland State University

Josh J. Prasad
Colorado State University

Tori L. Crain
Portland State University

Understanding Work and Sleep Through a Machine Learning Approach

Jennifer Saucedo[1], Josh J. Prasad[2], & Tori L. Crain[1]

1. Portland State University, 2. Colorado State University

BACKGROUND

- 1 in 3 adults do not get the recommended 7+ hours of sleep[1,2].
  - Lack of sleep linked to a variety of chronic health outcomes (e.g., diabetes, high blood pressure).
- Prior research finds that work may play a role in this unhealthy sleep trend.
  - Paid work time and commute time are strong predictors of sleep[3].
  - Studies implement the use of ordinary least squares (OLS) regression, which only allow for a minimal amount of predictor variables and variables are chosen based on theory.
- Few studies have implemented machine learning methods to examine sleep.
- Utilizing machine learning approaches will help us better understand the relative importance of work-related factor to sleep.

THE PRESENT STUDY

To address these some of these limitations, the present study will:

- Examine if the utilization of machine learning methods will better predict sleep over traditional statistical methods (e.g., OLS regression).
- Examine the individual factors that best predict sleep across all domains of life (e.g., work-related and life-style variables).

METHOD AND ANALYTICAL APPROACH

Government-funded data from the Bureau of Labor Statistics (BLS), the American Time Use Survey (ATUS) was utilized. These data provide an understanding of how households in the United States spend their time.

- Overlapping variables in the Current Population Survey (CPS), respondent, and summary files were merged.
- There were a total of 27,810 participants were included in the analyses from survey years 2018 (*n* = 9,593), 2019 (*n* = 9,435), and 2020 (*n* = 8,782).
- Before primary analyses, all variables (890) with near-zero variance were dropped to reduce dimensions of the data, resulting in a total of 295 variables.

A variety of prediction methods were utilized to evaluate and compare predictive performance[4]. Data from 2018 were used to train the model and evaluate model performance in 2019 and 2020 data.

PARTICIPANT INFORMATION

- Age: 50.56 in 2018 (*SD* = 18.11), 50.61 (*SD* = 18.09) in 2019, 51.19 (*SD* = 18.29).
- Average Sleep (in minutes): 530.07 (*SD* = 137.87) in 2018, 529.87 (*SD* = 136.32) in 2019, and 540.43 (*SD* = 134.63) in 2020.
- Race: 79.6% white in 2018, 79.5% white in 2019, 80.5% in 2020.

TABLES

Table 1. Model performance in predicting total time asleep.

Prediction Methods	2018 R^2	2019 R^2	2020 R^2
Full OLS Regression (OLS)	.135	.135	.047
Forward Stepwise Regression (LM)	.133	.055	.111
Least Angle Regression (LAR)	.260	.252	.263
Elastic Net Regression (ENET)	.259	.243	.263
Principle Component Regression (PCR)	.254	.244	.245
Partial Least Squares Regression (PLS)	.254	.244	.244
Random Forests (RF)	.316	.323	.344
Stochastic Gradient Boosted Trees (GBM)	.322	.329	.351

Table 2. Most important variables in predicting total sleep time across models.

Variables	OLS	LM	LAR	ENET	PCR	PLS	RF	GBM	Mean
Time working	100.00	100.00	100.00	2.97	100.00	54.96	94.53	100.00	81.56
Time alone at work	2.14	37.61	37.61	0.38	37.61	43.97	100.00	85.36	43.09
Time alone not at work	40.50	18.72	18.72	1.41	18.72	40.88	72.77	36.05	30.97
Time commuting to work	10.87	46.37	46.37	2.98	46.37	34.00	34.61	6.25	28.48
Household family income	16.57	13.36	13.36	25.84	13.36	100.00	24.26	12.16	27.36
Sleeplessness	42.35	13.56	13.56	5.49	13.56	46.14	35.65	24.45	24.34
Day of week	22.11	5.94	5.94	41.04	5.94	60.22	18.61	12.93	21.59
Time spent eating/drinking	39.77	2.36	2.36	3.33	2.36	48.42	37.48	18.79	19.36
Educational attainment	8.20	14.89	14.89	18.93	14.89	43.49	15.40	6.56	17.15
Socialization	53.40	1.22	1.22	1.47	1.22	24.56	18.93	25.31	15.92

Note. Importance is quantified via Relative Importance, which is on a 0 - 100 scale indicating not important to most important. To facilitate interpretation, cells are color coded according to importance as well. Red indicates not important, and blue indicates important.

RESULTS

Machine learning methods, on average, tend to outperform traditional regression methods (Table 1).

- Stochastic gradient boosted trees and random forests show improvement in predictive performance.
- Both methods incorporate regularization and handle predictors differently than traditional methods.

Time at work and commute time are important predictors of sleep (Table 2).

- Findings further indicate additional variables that have not been previously examined in prior research, including time alone not at work, time alone at work, and household family income.

DISCUSSION

Practical Implications:

- Future organizational interventions should target loneliness, financial insecurity, and long working hours. Organizations may consider reducing work hours during the day and allow flexibility to employees to support sleep and health.
- Possible implications for public policy related to working hours and income equity.

Recommendations for Future Research:

- Examine sleep duration over a longer period and consider additional dimensions of sleep other than sleep duration (e.g., sleep quality).
- Examine additional forms of sleep data that can be analyzed with machine learning methods, such as actigraphy data.
- Examine demographic variables, to examine intersectional inequity, given the importance of family income in current models.
- Explore machine learning methods as a potential informative method that can be used to robustly predict a variety of health outcomes in the workplace.

REFERENCES

1. Centers for Disease Control and Prevention. (2017, May 2). Sleep and sleep disorder: Data and statistics. Retrieved April 1, 2022, from https://www.cdc.gov/sleep/data_statistics.html
2. Consensus Conference Panel, Watson, N. F., Badr, M. S., Belenky, G., Bliwise, D. L., Buxton, O. M., Buysse, D., Dinges, D.F., Gangwisch, J., Grandner, M.A., Kushida, C., Malhotra, R.K., Martin, J.L., Patel, S.R., Quan, S.F., Tasali, E., Twery, M., Croft, J.B., Maher, E., Barrett, J.A., Thomas, S.M., Heald, J.L. (2015). Recommended amount of sleep for a healthy adult: a joint consensus statement of the American Academy of Sleep Medicine and Sleep Research Society. Journal of Clinical Sleep Medicine, 11(6), 591-592.
3. Basner, M., Spaeth, A. M., & Dinges, D. F. (2014). Sociodemographic characteristics and waking activities and their role in the timing and duration of sleep. Sleep, 37(12), 1889-1906.
4. Putka, D. J., Beatty, A. S., & Reeder, M. C. (2018). Modern prediction methods: New perspectives on a common problem. Organizational Research Methods, 21(3), 689-732.

Portland State UNIVERSITY