



Predicting Caffeine-Related Sleep Disruption Using Supervised Machine Learning

Zahra Zare^{1,*}, Mumtaz Karatas¹

¹ Wright State University, Dayton, OH 45435, USA

*Corresponding author. E-mail addresses: zare.3@wright.edu (Z. Zare), mumtaz.karatas@wright.edu (M. Karatas)

ARTICLE INFO

Article history:

Received 27 May 2026

Received in revised form 2 July 2026

Accepted 5 July 2026

Available online 7 July 2026

Keywords:

Caffeine consumption; Sleep quality; Machine learning; Health analytics; Supervised learning; Classification models

ABSTRACT

Caffeine is one of the most widely consumed psychoactive substances, yet its effects on sleep remain complex and not fully understood. In this study, we investigate the relationship between caffeine consumption habits and their impact on sleep quality using machine learning models. For this purpose, we use a dataset that includes self-reported survey responses from 500 participants, featuring variables such as caffeine intake, beverage type, time of day, gender, and age. We frame the problem as a binary classification task to determine whether an individual's sleep was negatively affected by caffeine. After data preprocessing and feature engineering, we applied and evaluated the performance of three classification models: logistic regression, decision tree, and random forest. We assessed model performances using standard metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. Our results revealed that the random forest classifier outperformed the other two approaches.

1. Introduction

Caffeine is one of the most widely consumed stimulants worldwide, commonly found in coffee, tea, energy drinks, and soft beverages. Its effects on the human body—particularly on focus and sleep—are of great interest to both researchers and the general public. While caffeine is known to temporarily enhance alertness and cognitive performance, excessive or poorly timed consumption may lead to disrupted sleep and reduced overall well-being. Despite its widespread use, individual responses to caffeine vary significantly, influenced by factors such as age, gender, metabolism, and behavioral habits.

Recent advancements in artificial intelligence (AI) and machine learning (ML) have shown promise in modeling complex dietary and behavioral patterns, including caffeine intake, to provide more personalized health insights. X. Wang et al. [31] highlighted how AI applications can analyze multi-dimensional data to tailor dietary recommendations based on individual needs and behaviors, thereby enhancing well-being and lifestyle management. Similarly, Tsolakidis et al. [29] reviewed a range of ML algorithms used for personalized nutrition, showing how data-driven techniques—including decision trees and ensemble methods—can effectively interpret behavioral inputs for improved predictive modeling.

In this study, we explore the impact of caffeine consumption on two critical cognitive dimensions: focus level and sleep quality. Using a dataset of 500 self-reported participants—including features such as caffeine intake, beverage type, time of day, gender, and age—we aim to uncover patterns and predictive relationships through ML models. Our goal is to determine how different variables interact with caffeine to influence outcomes and whether predictive techniques can offer actionable insights.

The main contribution of our study is the development of a supervised learning framework for predicting caffeine-related sleep disruption using individual-level behavioral and demographic features. Unlike traditional descriptive analyses that primarily examine average associations between caffeine intake and sleep outcomes, in this study we evaluate whether ML models can classify individuals who are more likely to experience sleep disruption. By comparing interpretable and non-linear classification models, we provide insights into both predictive performance and the potential usefulness of data-driven approaches for personalized lifestyle recommendations.

To accomplish this, we frame a supervised learning problem focused on classification, specifically, assessing whether an individual's sleep was negatively impacted by caffeine consumption. We apply and evaluate multiple classification models, including logistic regression, decision trees, and random forests, using metrics such as accuracy, precision, recall, F1-score, and ROC-AUC to measure performance.

We believe that this work contributes to the field of health analytics, where personalized recommendations derived from behavioral data can support healthier consumption decisions and improved lifestyle awareness. Section 2 reviews relevant literature on caffeine's impact on sleep and focus, as well as prior applications of ML in health analytics. Section 3 describes the dataset, preprocessing steps, and modeling methodology. Section 4 presents the experimental results and discussion. Finally, Section 5 concludes the study and outlines directions for future research.

2. Related Work

In this section, we provide a synthesis of the related work. We first investigate existing studies that analyze the effects of caffeine on focus and sleep. Next, we draw our attention to the applications of supervised ML in behavioral health studies, with a focus on lifestyle data to predict health and cognitive outcomes.

2.1 *Effects of Caffeine on Focus and Sleep*

Caffeine is a widely consumed central nervous system stimulant known for its ability to enhance alertness, attention, and cognitive performance. Understanding how caffeine consumption affects cognitive performance and sleep quality has been investigated by several researchers in both health sciences and behavioral research domains. Since caffeine is one of the most commonly used stimulants, it has potential negative impacts on sleep, and this is a complex research subject. With the increasing availability of behavioral data and advancements in ML techniques, researchers can analyze and predict those effects based on individual consumption habits.

Several studies have demonstrated that moderate caffeine intake can improve reaction time, selective attention, and working memory, particularly in cases where fatigue or suboptimal alertness occur [5, 20, 21]. For example, the European Food Safety Authority (EFSA) has established a causal relationship between caffeine consumption and improved alertness and attention, with as little as 75 mg being effective [7]. Experimental evidence also reveals that caffeine enhances performance in tasks requiring sustained attention and quick responses [21, 33].

In [3], the author investigated the effects of caffeine on attention and task performance through a two-part study combining a behavioral experiment and a case study involving salivary caffeine measurement. Results indicated no significant differences in accuracy or reaction times between caffeinated and non-caffeinated conditions in an auditory attention task, though task difficulty influenced performance.

Einöther and Giesbrecht [8] reviewed the acute effects of caffeine on attention by analyzing double-blind, placebo-controlled studies. Their findings indicated that caffeine enhances performance on

both simple and complex attention tasks and impacts alerting and executive control networks, while the evidence on dose-response relationships, habitual consumption effects, and withdrawal reversal remains inconclusive. The review concludes that caffeine provides clear cognitive benefits that extend beyond reversing withdrawal symptoms. The cognitive-enhancing effects of caffeine are not uniform across all individuals or cognitive domains. While caffeine generally improves “lower” cognitive functions such as simple reaction time, its effects on “higher” cognitive processes—like complex problem-solving or memory—are less consistent [20]. Nehlig [21] also reports that caffeine may facilitate learning in passive tasks but has minimal effect on intentional learning or long-term memory. Additionally, according to [5], excessive caffeine intake can lead to adverse effects such as anxiety, jitteriness, and restlessness, which may counteract its cognitive benefits.

Caffeine’s primary mechanism of action involves antagonism of adenosine receptors, which delays the onset of sleepiness and promotes wakefulness [25, 26]. While this property underlies its stimulating effects, it also causes disruptions in sleep architecture and quality. Previous research reveals that caffeine consumption, especially in the hours leading up to bedtime, can delay sleep onset, reduce total sleep time, decrease sleep efficiency, increase the proportion of light sleep, decrease the proportion of deep sleep, and lead to more nighttime awakenings.

In [23], the authors presented an integrative review examining the relationship between caffeine consumption, sleep quality, and subsequent daytime functioning. They concluded that while caffeine improves cognitive and physical performance, these benefits may largely reflect the reversal of withdrawal effects rather than a true net gain, and that further research is needed to clarify its overall impact. In [32], the authors performed a cross-sectional study using a caffeine food frequency questionnaire (C-FFQ) and the Pittsburgh Sleep Quality Index (PSQI) to examine the relationship between caffeine intake and sleep quality in adults. Their findings revealed that higher caffeine consumption is associated with shorter sleep duration and poorer overall sleep quality.

In another study, Kerpershoek et al. [16] examined the influence of caffeine consumption timing and chronotype on sleep quality among 880 students. They found that higher total caffeine intake is linked to poorer sleep quality only for those who avoided evening caffeine. Similarly, Higbee et al. [12] conducted a comparative study among 272 undergraduate nursing students to examine the impact of caffeine and energy drink consumption on sleep quality, sleep quantity, and perceived stress. Their findings revealed that students who consumed energy drinks reported significantly poorer sleep quality, fewer sleep hours, and higher stress levels compared to those who consumed other caffeinated beverages or abstained from caffeine. In [18], the authors surveyed 262 university students to examine the relationship between caffeinated beverage intake and sleep quality, finding a significant negative correlation where higher caffeine consumption was associated with poorer sleep quality, suggesting the need to control caffeine intake to improve sleep.

Meta-analyses also indicate that caffeine can reduce total sleep time by up to 45 minutes and sleep efficiency by 7%, with even moderate doses (e.g., 200 mg) having measurable effects if consumed within 6 hours of bedtime [6, 9]. Habitual high caffeine intake (four or more caffeinated beverages daily) is associated with shorter sleep duration and altered sleep patterns in large population studies [27]. Individual sensitivity, genetic factors, and the timing of consumption all influence the degree of sleep disruption experienced [25]. In [4], the authors provided a systematic review of epidemiological studies and randomized controlled trials that were performed to understand the effect of caffeine on the quality of sleep. They highlighted that caffeine is widely used to enhance alertness and may offer health benefits, but its causal effects on health and sleep remain unclear.

2.2 Machine Learning in Behavioral Health

Supervised ML has emerged as a transformative tool in behavioral health research, particularly for predicting health and cognitive outcomes using complex lifestyle data. The increasing availability of large-scale behavioral and lifestyle datasets has enabled researchers to implement supervised ML algorithms, such as random forests, support vector machines, and logistic regression, to understand patterns and risk factors that traditional statistical models may miss. These models are trained on labeled datasets, where input features—such as demographic, behavioral, or lifestyle variables—are mapped to known outcomes, including mental health diagnoses, disease risk, or cognitive scores [13, 14].

A rich body of literature has demonstrated the utility of supervised ML in predicting the onset and progression of mental health conditions, such as depression, anxiety, and ADHD, by leveraging behavioral and lifestyle data. These approaches often outperform traditional statistical techniques in terms of predictive accuracy and the ability to handle high-dimensional data [1, 13]. Supervised ML algorithms have also been employed to forecast cognitive impairment and dementia, using longitudinal health, lifestyle, and neuropsychological data to gain insights for early intervention and care planning [11, 17]. In addition to prediction, ML models have facilitated the identification of individuals most likely to benefit from specific interventions, such as exercise programs for children with autism spectrum disorder or personalized mental health treatments [28].

Lifestyle factors, including physical activity, diet, sleep, and social behaviors, are critical predictors of a wide range of health outcomes. In a more recent study, Tsolakidis et al. [29] reviewed recent advancements in AI/ML technologies for personalized nutrition. They highlighted how AI-driven recommendation systems, combined with big data from wearable devices, sensors, and mobile applications, are increasingly used to deliver individualized diet and lifestyle recommendations. Supervised ML models trained on these lifestyle features have demonstrated high accuracy in predicting chronic disease risk, such as diabetes and cardiovascular disease, as well as all-cause mortality, using data from large-scale surveys like NHANES and BRFSS [10, 22]. Furthermore, ML approaches have been applied to predict health-related quality of life and long-term health status, integrating social determinants and behavioral data to provide a holistic view of individual well-being [2]. By continuously monitoring lifestyle patterns, ML models can also identify individuals at risk for lifestyle-related diseases before clinical symptoms emerge, supporting proactive and preventive healthcare strategies [19, 24]. More broadly, Karatas et al. [15] reviewed how ML is transforming preventive healthcare (PHC) as a whole, highlighting its growing capabilities in early disease detection, personalized risk assessment, and targeted interventions across domains such as cancer screening, cardiovascular risk prediction, and population health management. Recent studies have also highlighted the potential of ML to identify key nutritional and behavioral predictors of cognitive performance, offering new avenues for cognitive health promotion and intervention [30].

3. Methodology

3.1 Data and Features

The dataset used in this study was obtained from a publicly available caffeine intake tracker on Kaggle. It consists of 500 self-reported entries, each representing a unique individual's caffeine consumption habits along with cognitive and physiological outcomes. The dataset includes a mix of *numeric*, *categorical*, and *binary* variables. Key features include:

- **Caffeine intake (mg):** Total daily caffeine consumption.

- **Time of day:** Categorized into morning, afternoon, and evening.
- **Beverage type:** Type of caffeinated drink consumed (e.g., coffee, tea, energy drinks).
- **Gender:** Self-reported gender of the participant.
- **Age:** Participant age in years.
- **Focus level:** Self-reported score from 0 to 1.
- **Sleep quality:** Self-reported score from 0 to 1.
- **Sleep impacted:** Binary label indicating whether the person's sleep was negatively affected (0 = No, 1 = Yes).

Prior to modeling, the data was preprocessed. Categorical variables such as beverage type, gender, and time of day were one-hot encoded. The dataset had no missing values. The target variable for classification (`sleep_impacted`) showed a moderate class distribution of 297 negative (59.4%) and 203 positive (40.6%) cases, which was accounted for using evaluation metrics such as precision, recall, and ROC-AUC that are robust to mild class imbalance.

To ensure the dataset was suitable for ML algorithms, we applied preprocessing and feature engineering steps. Since the dataset did not contain any missing values, no imputation was required. Categorical variables, such as beverage type, time of day, and gender, were transformed using one-hot encoding to convert them into a machine-readable numerical format. Additionally, we conducted feature selection by examining correlations among variables and evaluating model performance using reduced feature sets. Features that did not improve predictive performance or provided limited additional information were excluded from the final models.

The target variable in this study is *sleep impacted*, a binary indicator representing whether the participant's sleep was negatively affected by caffeine consumption (0 = no, 1 = yes). This variable was used for classification tasks.

3.2 Modeling Approach

We framed this study as a binary classification task to predict whether an individual's sleep was negatively impacted by caffeine consumption. All models were implemented using Python's `scikit-learn` library. For this purpose, we employed three supervised learning algorithms: logistic regression, a linear model for binary classification; decision tree, a non-linear model based on hierarchical feature splits; and random forest, an ensemble of decision trees to improve generalization.

We split the dataset into training (70%) and testing (30%) sets. We evaluated the model performance using standard classification metrics, including accuracy, precision, recall, F1-score, and ROC-AUC.

4. Results and Discussion

Table 1 summarizes the performance of all three classification models evaluated in this study. Among these, the random forest model achieved the highest overall performance, particularly in terms of ROC-AUC, which reached 0.9566. This value indicates the discriminatory ability and demonstrates the model's strength in distinguishing between cases where sleep was impacted and cases where it was not. Logistic regression performed identically in terms of overall predictive accuracy, achieving 0.90 compared to random forest's 0.90, and both models reported the same precision

(0.9032) and recall (0.8615).. However, the slight edge in ROC-AUC for random forest suggests that it handles threshold variation better, likely due to its ability to model nonlinear relationships and interactions among features.

In contrast, the decision tree model exhibited noticeably lower performance across all metrics. Its accuracy was 0.8333, which represents a drop of approximately 6.7 percentage points compared to the other two models. Furthermore, its precision and recall both declined to 0.8000, resulting in an F1-score of 0.8000, compared to 0.8819 for both logistic regression and random forest. The most significant gap appears in the ROC-AUC metric: while logistic regression and random forest both exceeded 0.95, the decision tree model only achieved 0.8258, which is considerably lower and reflects its limited ability to capture complex relationships without ensemble techniques.

The nearly identical performance between logistic regression and random forest in terms of accuracy, precision, recall, and F1-score highlights an important trade-off between interpretability and flexibility.

Table 1
Classification performance comparison

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC
Logistic Regression	0.90	0.9032	0.8615	0.8819	0.9542
Random Forest	0.90	0.9032	0.8615	0.8819	0.9566
Decision Tree	0.8333	0.833	0.769	0.8	0.8258

Figures 1–3 illustrate the confusion matrices and Receiver Operating Characteristic (ROC) curves for each model. The confusion matrix for logistic regression in Figure 1 shows that the model correctly classified 79 out of 85 negative cases (sleep not impacted) and 56 out of 65 positive cases (sleep impacted). This results in only 15 misclassifications overall (6 false positives and 9 false negatives). The ROC curve demonstrates strong separability between classes, with an area under the curve (AUC) of 0.95, indicating that the model consistently ranks true positives higher than false positives across thresholds.

The random forest model results in Figure 2 exhibit an identical confusion matrix to logistic regression in this experiment (79 true negatives, 56 true positives, 6 false positives, and 9 false negatives). This leads to the same accuracy (0.90) and precision-recall metrics as logistic regression. However, the ROC curve shows a marginally higher AUC of 0.96.

In contrast, the decision tree model in Figure 3 displays a noticeably lower performance. The confusion matrix reveals 75 true negatives and 50 true positives, along with 10 false positives and 15 false negatives. These higher error counts result in reduced accuracy (0.8333) and a lower F1-score (0.8000). The ROC curve for this model confirms weaker separability, with an AUC of only 0.83, below that of the other two models.

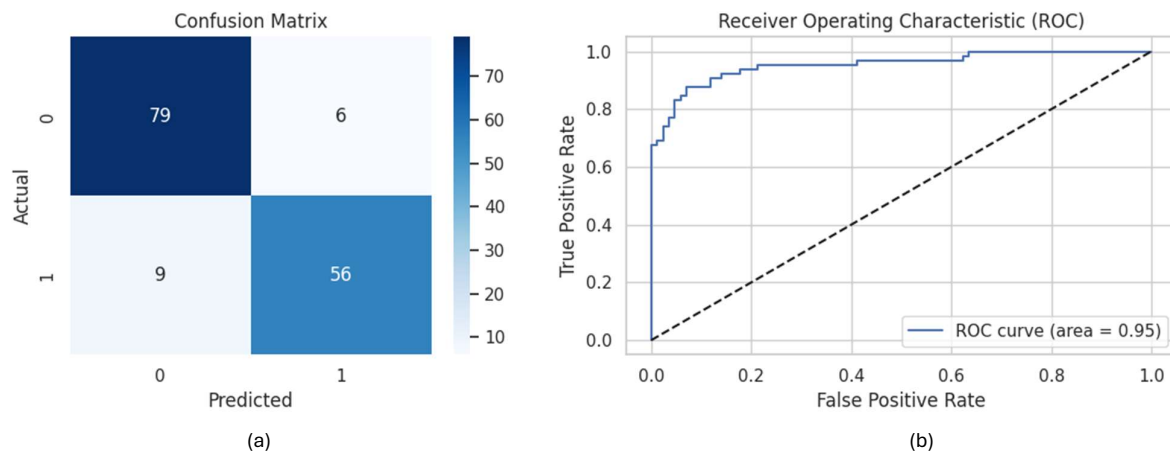


Fig. 1. Performance of logistic regression: (a) confusion matrix and (b) ROC curve

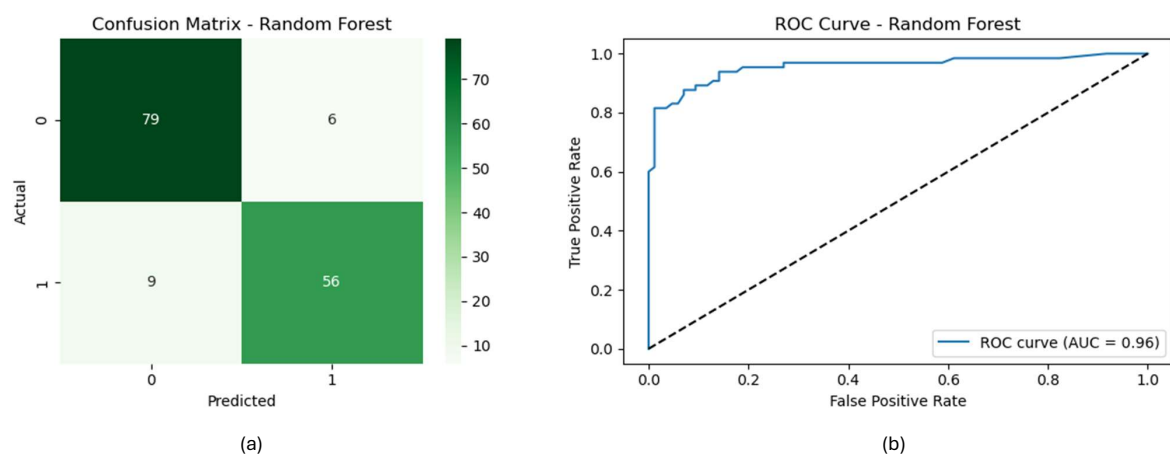


Fig. 2. Performance of random forest: (a) confusion matrix and (b) ROC curve

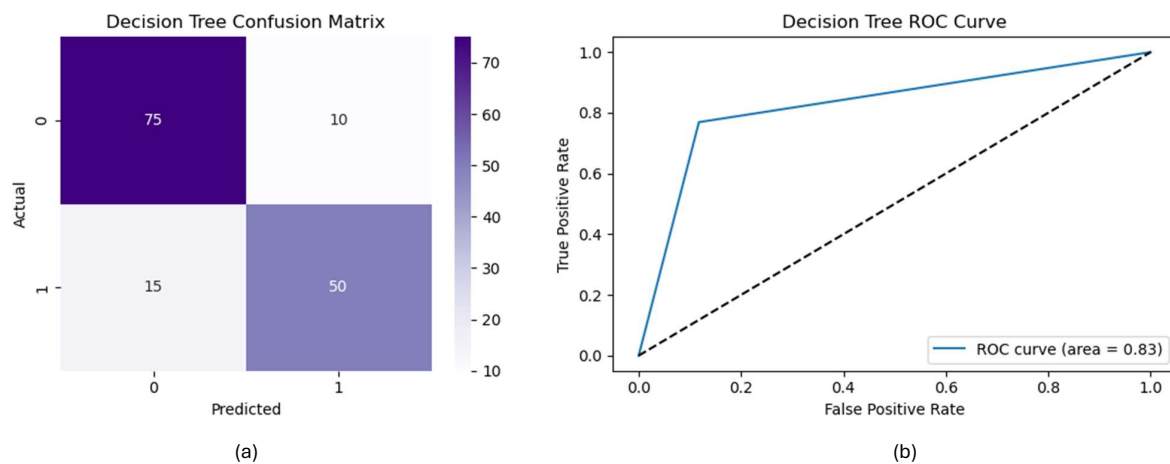


Fig. 3. Performance of decision tree: (a) confusion matrix and (b) ROC curve

The ability to predict sleep disruption with an accuracy of 90% has practical implications for personalized health recommendations. For example, individuals could receive alerts to limit caffeine intake at specific times of day, reducing the risk of impaired sleep quality. Although this study uses self-reported data, the findings provide a foundation for developing real-time predictive tools integrated with wearable devices or mobile health applications.

Although our numerical results are promising, they should be interpreted with caution due to the

nature of the dataset. The data are self-reported and relatively limited in size, which may introduce reporting bias and restrict the generalizability of the findings. In addition, important factors such as caffeine tolerance, sleep schedule, stress level, screen exposure, physical activity, medication use, and individual metabolic differences were not included in the dataset. Therefore, the models should not be interpreted as clinical decision tools, but rather as an exploratory demonstration of how supervised ML can be used to identify patterns between caffeine-related behaviors and sleep outcomes.

5. Conclusions

In this study we explored the use of supervised ML models to predict whether caffeine consumption negatively impacts sleep quality. Using a self-reported dataset of 500 individuals, we developed and compared logistic regression, decision tree, and random forest classifiers. The results showed that both logistic regression and random forest achieved strong predictive performance, while the decision tree model performed comparatively lower. The random forest model achieved the highest ROC-AUC value, suggesting that ensemble-based methods may better capture non-linear relationships among caffeine intake, timing, beverage type, and demographic factors.

Our results indicate that random forest provided the best trade-off between accuracy and generalization, making it the most suitable model among those tested. Logistic regression, while slightly behind in some settings, remains a competitive and highly interpretable option, which is important for health-related applications where transparency is critical. The decision tree, although easy to interpret, was outperformed by the other two approaches and showed a tendency to overfit; single-tree models therefore appear less reliable for this classification problem without further optimization.

Several opportunities exist to enhance this study. First, hyperparameter tuning using methods such as grid search or randomized search could improve model performance, particularly for random forest and decision tree classifiers. Applying cross-validation rather than a single train-test split would also yield more stable and generalizable performance estimates. Another important direction is analyzing feature importance to identify which variables contribute most to predicting sleep disruption. Additionally, expanding the feature set to include behavioral and lifestyle variables such as stress level, screen time before sleep, and physical activity could capture additional variance and improve prediction accuracy. Finally, future research should explore advanced ML techniques, including gradient boosting methods such as XGBoost or LightGBM, and validate models on larger and more diverse datasets to ensure robustness and generalizability.

Acknowledgement

The authors did not receive any specific grant from funding agencies in the public, commercial, or non-profit sectors.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] Abdul Rahman, H., Kwicklis, M., Ottom, M., Amornsriwatanakul, A., H. Abdul-Mumin, K., Rosenberg, M., & Dinov, I. D. (2023). Machine learning-based prediction of mental well-being using health behavior data from university students. *Bioengineering*, 10(5). <https://doi.org/10.3390/bioengineering10050575>

- [2] Abegaz, T. M., Ahmed, M., Ali, A. A., & Bhagavathula, A. S. (2025). Predicting health-related quality of life using social determinants of health: A machine learning approach with the all of us cohort. *Bioengineering*, 12(2), 166.
- [3] Berger, C. R. (2021). *Caffeine modulation of attention and focus in task performance* [Master's Thesis]. City University of New York (CUNY).
- [4] Clark, I., & Landolt, H. P. (2017). Coffee, caffeine, and sleep: A systematic review of epidemiological studies and randomized controlled trials. *Sleep medicine reviews*, 31, 70–78.
- [5] Creyos Team. (2023). Does coffee help with focus? [Accessed: 2025-07-12]. <https://creyos.com/blog/does-coffee-help-with-focus>
- [6] Drake, C., Roehrs, T., Shambroom, J., & Roth, T. (2013). Caffeine effects on sleep taken 0, 3, or 6 hours before going to bed. *Journal of Clinical Sleep Medicine*, 9(11), 1195–1200.
- [7] EFSA. (2014, April). Scientific opinion on the substantiation of a health claim related to caffeine and increased alertness [At least 75 mg per serving causes improved attention]. <https://www.efsa.europa.eu/en/efsajournal/pub/3574>
- [8] Einöther, S. J., & Giesbrecht, T. (2013). Caffeine as an attention enhancer: Reviewing existing assumptions. *Psychopharmacology*, 225(2), 251–274.
- [9] Gardiner, C., Weakley, J., Burke, L. M., Roach, G. D., Sargent, C., Maniar, N., Townshend, A., & Halson, S. L. (2023). The effect of caffeine on subsequent sleep: A systematic review and meta-analysis. *Sleep medicine reviews*, 69, 101764.
- [10] Guo, X., Ma, M., Zhao, L., Wu, J., Lin, Y., Fei, F., Tarimo, C. S., Wang, S., Zhang, J., Cheng, X., et al. (2025). The association of lifestyle with cardiovascular and all-cause mortality based on machine learning: A prospective study from the nhanes. *BMC Public Health*, 25(1), 319.
- [11] Harvey, J., Reijnders, R. A., Cavill, R., Duits, A., Köhler, S., Eijssen, L., Rutten, B. P., Shireby, G., Torkamani, A., Creese, B., et al. (2022). Machine learning-based prediction of cognitive outcomes in de novo parkinson's disease. *npj Parkinson's Disease*, 8(1), 150.
- [12] Higbee, M. R., Gipson, C. S., & El-Saidi, M. (2022). Caffeine consumption habits, sleep quality, sleep quantity, and perceived stress of undergraduate nursing students. *Nurse Educator*, 47(2), 120–124.
- [13] Iyortsuun, N. K., Kim, S.-H., Jhon, M., Yang, H.-J., & Pant, S. (2023). A review of machine learning and deep learning approaches on mental health diagnosis. *Healthcare*, 11(3), 285.
- [14] Jiang, T., Gradus, J. L., & Rosellini, A. J. (2020). Supervised machine learning: A brief primer. *Behavior therapy*, 51(5), 675–687.
- [15] Karatas, M., Zare, Z., & Zheng, Y.-J. (2025). Transforming preventive healthcare with machine learning technologies. *Journal of Operations Intelligence*, 3(1), 113–129.
- [16] Kerpershoek, M. L., Antypa, N., & Van den Berg, J. F. (2018). Evening use of caffeine moderates the relationship between caffeine consumption and subjective sleep quality in students. *Journal of sleep research*, 27(5), e12670.
- [17] Kong, A. H., Lai, M. M., Finnigan, S., Ware, R. S., Boyd, R. N., & Colditz, P. B. (2018). Background eeg features and prediction of cognitive outcomes in very preterm infants: A systematic review. *Early human development*, 127, 74–84.
- [18] Lee, B., Mi, K., Kim, B., Kim, B., Kim, J., Lee, I., In, E., & Jung, S. (2014). Caffeine contained beverage intake and sleep quality of university students. *Journal of the Korean Society of School Health*, 27(1), 31–38.
- [19] Manikanta, G., & Vijaya Lakshmi, G. (2023). Lifestyle based disease prediction using random forest algorithm. *Journal of Emerging Technologies and Innovative Research*, 10(6), 424–427.
- [20] McLellan, T. M., Caldwell, J. A., & Lieberman, H. R. (2016). A review of caffeine's effects on cognitive, physical and occupational performance. *Neuroscience & Biobehavioral Reviews*, 71, 294–312.

- [21] Nehlig, A. (2010). Is caffeine a cognitive enhancer? *Journal of Alzheimer's Disease*, 20(Suppl 1), S85–S94.
- [22] Nguyen, B., & Zhang, Y. (2025). A comparative study of diabetes prediction based on lifestyle factors using machine learning. *arXiv preprint arXiv:2503.04137*.
- [23] O'callaghan, F., Muurlink, O., & Reid, N. (2018). Effects of caffeine on sleep quality and daytime functioning. *Risk management and healthcare policy*, 263–271.
- [24] Patil, M., Lobo, V. B., Puranik, P., Pawaskar, A., Pai, A., & Mishra, R. (2018). A proposed model for lifestyle disease prediction using support vector machine. *2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 1–6.
- [25] Sleep Foundation. (2025). Caffeine and sleep [Accessed: 2025-07-12]. <https://www.sleepfoundation.org/nutrition/caffeine-and-sleep>
- [26] Sleep.me. (2025). Caffeine and sleep [Accessed: 2025-07-12]. <https://sleep.me/post/caffeine-and-sleep>
- [27] Stucky, B., Henckel, L., Maathuis, M. H., Haba-Rubio, J., Marques-Vidal, P., Siclari, F., Heinzer, R., & Landolt, H.-P. (2024). Community-based causal evidence that high habitual caffeine consumption alters distinct polysomnography-derived sleep variables. *medRxiv*, 2024–12.
- [28] Sun, Z., Yuan, Y., Dong, X., Liu, Z., Cai, K., Cheng, W., Wu, J., Qiao, Z., & Chen, A. (2023). Supervised machine learning: A new method to predict the outcomes following exercise intervention in children with autism spectrum disorder. *International journal of clinical and health psychology*, 23(4), 100409.
- [29] Tsolakidis, D., Gymnopoulos, L. P., & Dimitropoulos, K. (2024). Artificial intelligence and machine learning technologies for personalized nutrition: A review. *Informatics*, 11(3), 62.
- [30] Verma, S., Holthaus, T. A., Martell, S., Holscher, H. D., Zhu, R., & Khan, N. A. (2025). Predicting cognitive outcome through nutrition and health markers using supervised machine learning. *The Journal of Nutrition*.
- [31] Wang, X., Sun, Z., Xue, H., & An, R. (2025). Artificial intelligence applications to personalized dietary recommendations: A systematic review. *Healthcare*, 13(12), 1417.
- [32] Watson, E. J., Coates, A. M., Kohler, M., & Banks, S. (2016). Caffeine consumption and sleep quality in australian adults. *Nutrients*, 8(8), 479.
- [33] Wu, S.-H., Chen, Y.-C., Chen, C.-H., Liu, H.-S., Liu, Z.-X., & Chiu, C.-H. (2024). Caffeine supplementation improves the cognitive abilities and shooting performance of elite e-sports players: A crossover trial. *Scientific Reports*, 14(1), 2074.