Evaluation of the influence of behaviors on Machine Learning prediction of sleep apnea and insomnia

Abstract — The etiology of sleep disorders has shown that their origin is multifactorial. Thus, certain genetic factors exacerbate the intensity of these troubling events. However, the main causes of these disorders originate from the daily behaviors and lifestyle choices of patients. For example, the consumption of stimulants, poor eating habits, excessive screen time, and lack of physical activity are all triggers. Yet, most predictive models of sleep disorders rely almost exclusively on patients' physiological measurements. In response to this trend, this study aims to evaluate the influence of behaviors on the prediction of sleep apnea and insomnia. For this purpose, numerous machine learning algorithms were first created and then optimized through a random search coupled with stratified k-fold cross-validation before being trained and finally compared. The Multinomial Logistic Regression, being the most performant and thus considered the most reliable model, was then selected. The calculation of Mean Absolute SHAP Values subsequently revealed that behavioral variables, with a value of 0.0055, were the third most impactful variable on this model's prediction. Finally, a t-test showed that there is no statistically significant difference between the Mean Absolute SHAP Values of behavioral variables and physiological variables. As a result, it follows that behavioral variables are equally important as physiological variables in predicting insomnia and sleep apnea.

1. Introduction

a. *Sleep disorders*

Sleep disorders are defined as all disturbances in the duration or quality of sleep. Beyond sleep, these disorders significantly affect the health, safety, and quality of life of patients, whether they are adults or children [1]. It is common to differentiate sleep disorders into 3 groups according to their impact on the patient: dyssomnias (insufficient sleep), hypersomnias (excessive sleep), and parasomnias (abnormal behavior during sleep). Specifically, sleep apnea belongs to the group of parasomnias, while insomnia belongs to the group of dyssomnias. Additionally, since 2005, the American Academy of Sleep Medicine has published a new "International Classification of Sleep Disorders," considered a reference in the field, which distinguishes sleep disorders into 8 groups following advances in the pathology of these diseases [2].

Sleep apnea is one of the most frequent causes of sleep disorders. Sleep apnea is categorized into two groups: "obstructive sleep apnea" (OSA) and "central sleep apnea" (CSA). However, OSA has a much higher prevalence than CSA. Thus, nearly 17% of women and 34% of men in the United States suffer from obstructive sleep apnea. Similar figures are reported in other countries. This disorder is all the more alarming as its prevalence increased by about 30% between 1990 and 2010. The term "obstructive" refers to the fact that the occurrence of OSA results from obstructions of the upper airways (nose, pharynx, mouth, larynx) causing repetitive episodes of temporary interruptions in breathing. Several risk factors explain this phenomenon, including obesity, alcohol, a sedentary lifestyle, sleep position, late-night eating habits, and smoking. Among the potential effects caused by sleep apnea, daytime sleepiness, heart attack, depression, and cognitive disorders are among the most cited. To avoid these complications, anyone experiencing certain symptoms such as nocturia, daytime sleepiness, or unusual fatigue is advised to undergo a diagnosis by a clinician, often composed of a questionnaire, a physical examination, laboratory polysomnography (PSG), and Home Sleep Apnea Testing (HSAT). The severity of OSA is then quantified by the Apnea–Hypopnea Index (AHI), which measures the number of apneas per hour of the patient: $AHI<5$ = Normal, $5 \leq AHI<15$ = Mild sleep apnea, $15 \leq AHI<30$ = Moderate sleep apnea, and $AHI \geq 30$ = Severe sleep apnea. To treat OSA, the most used and effective method is Positive Airway Pressure (PAP). Its operation involves applying pressure in the patient's airways to prevent their obstruction during inspiration. Various other types of methods are also used as a supplement or in case of intolerance to PAP: behavioral measures (weight loss, alcohol abstinence, prevention of the supine sleep position), oral appliances, surgical treatments (maxillomandibular advancement, uvulopalatopharyngoplasty, tracheotomy) [3][4].

Insomnia, on the other hand, is defined as insufficient sleep related to difficulties in falling asleep and/or staying asleep. Depending on its frequency, it is considered acute insomnia if it lasts only from one night to a few weeks, while if it persists beyond 3 months and occurs 3 nights per week, then it is classified as chronic insomnia. The prevalence of insomnia is estimated to be 10% of the population in Europe. Insomnia most often originates from mental disorders such as depression, stress, or anxiety. Additionally, alcohol consumption can also contribute to causing insomnia. The consequences of insomnia are multiple: fatigue, concentration or memory problems, mood disorders, irritability, daytime sleepiness, hypertension, and loss of motivation, among others. Symptoms of insomnia include difficulty falling asleep or staying asleep, waking up earlier than desired, and unwillingness to go to bed at an appropriate time. To evaluate the severity of a patient's insomnia, the use of a sleep diary for 1 or 2 weeks is common and is the only method fully approved by the scientific community, as opposed to actigraphy and PSG. Next, the treatment of insomnia is mainly carried out according to two strategies: cognitive-behavioral therapy and pharmacological therapy. Cognitive-behavioral therapy for insomnia (CBT-I), conducted in 4 to 8 sessions, involves trying to relieve the patient of their cognitive disorders through appropriate recommendations, while pharmacological therapy involves taking medications such as antidepressants, antipsychotics, melatonin, and benzodiazepines. Their side effects must still be considered: dependence and dose abuse [5][6].

Additionally, recent years have seen a resurgence of research on co-morbid insomnia and sleep apnea (COMISA), although it was first identified in 1973 [7] and then described as "a new clinical syndrome, sleep apnea with insomnia." Indeed, a significant number of people suffer from these two disorders simultaneously. Thus, 30 to 50% of patients with OSA also suffer from insomnia, while 30 to 40% of patients with insomnia suffer from OSA [8].

b. *Gap*

To best screen and prevent these sleep disorders, numerous studies have developed machine learning prediction models in recent years. The drawback of these studies is that they are exclusively physiologically-oriented. This means that the databases chosen in these studies are constructed from clinical measurements or sleep-focused questionnaires. This approach can be justified. Indeed, the objective is notably to build models capable of substituting polysomnography (PSG), the main diagnostic tool for OSA, but whose use, complex and costly in terms of qualified personnel, is in high demand by patients. Thus, building models based on respiratory variables is appropriate [9]. However, such a procedure denies the potential influence of patients' behaviors on the prediction of sleep disorders. This study aims to determine whether behavioral variables have a significant impact on sleep disorder prediction models, particularly in comparison to physiological variables. To achieve this, several prediction models will be trained and evaluated, similar to what is done in the literature. Then, the best model, considered the most reliable, will be retained to perform an analysis of the importance of its features on its prediction.

2. Methods

a. *Dataset description*

The dataset used in this study comes from Kaggle [10]. It contains 374 patient samples organized into 13 features, including physiological features (“Blood Pressure”, “BMI Category”, “Heart Rate”), demographic features (“Gender”, “Age”), behavioral features (“Occupation”, “Daily Steps”, “Physical Activity Level”), and features that can resemble responses from sleep monitoring questionnaires (“Sleep duration”, “Quality of Sleep”, “Stress Level”). The target variable is “Sleep disorder”. It contains three classes: “None”, “Insomnia”, and “Sleep Apnea”. Finally, this dataset contains no missing values.

b. *Data processing and feature selection*

The data processing required several steps. The first was to delete the last feature "Person ID" which only served to differentiate each entry. The next step was to merge, in the "BMI Category" feature, the classes "Normal" and "Normal Weight" as they designated the same type of individuals. Then, the "Blood Pressure" feature, whose values were in the form "Systolic Pressure/Diastolic Pressure", was converted into two new features "Systolic Pressure" and "Diastolic Pressure". With these initial modifications made, several data visualizations were performed in search of outliers. As for the categorical variables, the classes of the "Occupation" feature presented uneven distributions as shown in Figure 1. The decision was then made to group these profession classes into categories based on their similarity: "Technical" = "Engineer" + “Scientist" + "Software Engineer" / "Commercial" = "Sales Representative" + "Salesperson" / "Administrative" = "Accountant" + "Manager" + "Lawyer". The class "Teacher" was left as is because it did not belong to any of these categories. "Doctor" and "Nurse" were kept separate as their large occurrences would have nullified the process. Indeed, the objective was threefold. First, to reduce the imbalance while maintaining a logical and informative data structure. Secondly, this allowed reducing the dimensionality of the problem, especially since it is the only categorical variable containing 11 different classes, whereas the other categorical variables ("Gender", "BMI Category") have only 3 distinct classes. Finally, it also helped improve the performance of the models by making the classes more homogeneous with a lower variance, which decreased from 786 to 415.

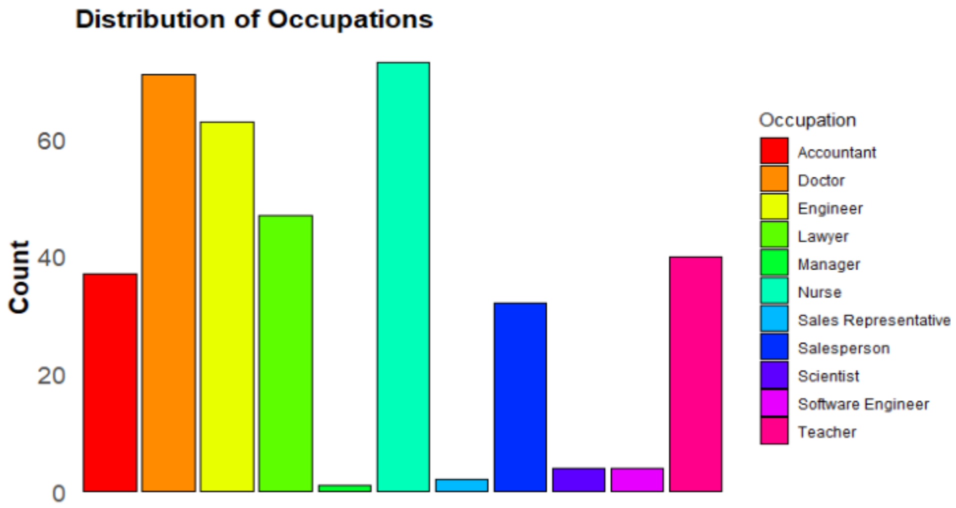

Figure 1 - Distribution of classes in “Occupation” before grouping

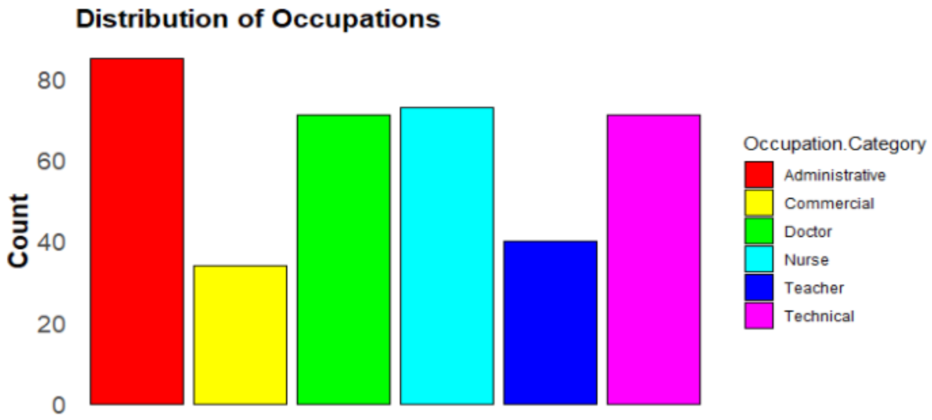

Figure 2 - Distribution of classes in “Occupation” after grouping

I then applied one-hot encoding to the 3 categorical variables (“Gender”, “BMI Category”, “Occupation”) because the classes of these features are not ordered. Regarding the numerical features, the first step was to standardize the data to improve the performance of certain distance-based classification models such as KNN or SVM [11]. This standardization was done using mean-var scaling. Then, in a second step, the correlation coefficients were calculated using the correlation matrix, with the aim of removing highly correlated variables to further reduce the dimensionality of the problem and avoid redundancy that can affect prediction models and cause overfitting. Several variables showed high correlation $(|correlation| > 0.8)$ as illustrated in Figure 3. In particular, the triplet of features “Quality of Sleep”, “Sleep Duration”, and “Stress Level” was highly correlated. “Quality of Sleep” and “Stress Level” were removed because they are more subjective compared to “Sleep Duration”. “Systolic Pressure” and “Diastolic Pressure” also had a very high correlation (0.97). To solve this problem, instead of keeping only one of the two, I decided to remove both and create a new variable from them: “Mean arterial pressure”. Indeed, mean arterial pressure (MAP) is a clinical measure often used by health professionals that can be approximated [12] by the formula: $MAP = DBP + \frac{1}{3}(SBP - DBP)$ with DBP (Diastolic Blood Pressure) and SBP (Systolic Blood Pressure).

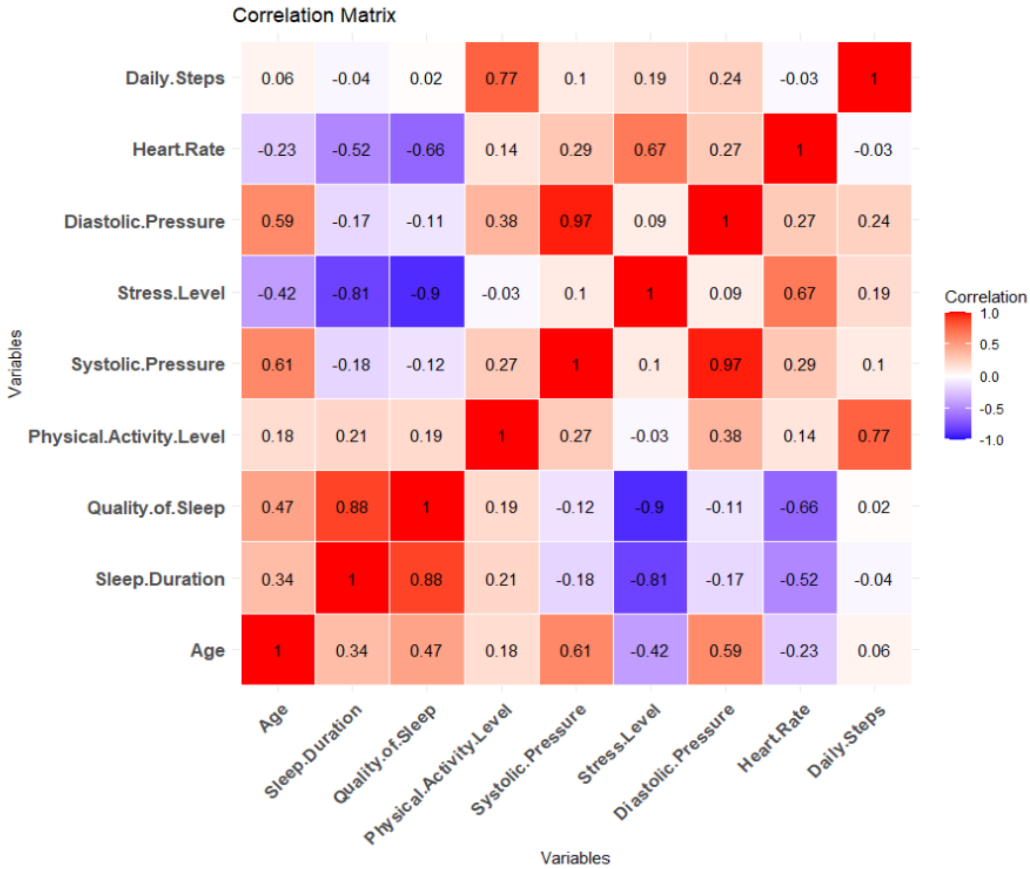

Figure 3 - Correlation matrix of numerical features

c. *ML models*

The target variable "Sleep Disorder" is composed of three classes "None", "Insomnia", and "Sleep Apnea". Therefore, I am dealing with a multiclass problem. To obtain the most reliable classifier possible, 7 different models were built and compared:

- Multinomial Logistic Regression

Multinomial logistic regression is an extension of ordinary logistic regression that allows dealing with classification problems with more than two categories. Thus, multinomial logistic regression estimates the probabilities of belonging to each of the k possible classes simultaneously.

- Naive Bayes

Naive Bayes is a type of simple probabilistic Bayesian classification based on Bayes' theorem with strong independence assumptions. A naive Bayesian classifier assumes that the presence of a particular feature in a class is independent of the presence of other features.

- KNN

The K-Nearest Neighbors (KNN) method aims to classify target points based on their distances to the points in the training sample. For classification, the algorithm assigns an unknown sample to the most frequent class among its k nearest neighbors. Distances between points are calculated using metrics like Euclidean distance.

- Random Forest

A random forest consists of a set of independent decision trees. Each tree has a partial view of the problem due to a double random sampling: random sampling with replacement on the observations (tree bagging) and random sampling on the variables (feature sampling). In the end, all these independent decision trees are assembled. The prediction made by the random forest is then the vote of all the trees.

- Support Vector Machine

The principle of SVMs is to reduce a classification problem to a hyperplane in which the data is separated into several classes whose boundary is as far as possible from the data points. In a multiclass problem, SVM can use two different approaches: "one-vs-one" or "one-vs-all". Here, I used the "one-vs-one" approach to build our models. That is, an SVM was built for each pair of classes, so in my case 3 SVMs (Sleep Apnea vs Insomnia, Insomnia vs None, None vs Sleep Apnea).

- Gradient Boosting

Gradient Boosting is a boosting algorithm whose objective is to train a forest of trees sequentially so that each new tree compensates for the weaknesses of the forest. The algorithm seeks to minimize the residuals (the difference between a model's prediction and the target value) of the previous trees.

- MLP

The multilayer perceptron is a type of neural network organized in multiple layers. A multilayer perceptron has at least three layers: an input layer, at least one hidden layer, and an output layer. Each layer consists of a number, potentially different, of neurons. An MLP uses these layers to map sets of input data onto appropriate outputs.

d. Model selection

Before comparing these classifiers with each other, the primary objective was to build the most performant model possible for each of these 7 algorithms. For this, a combination of the k-fold cross-validation technique and the random search method was conducted. First, the initial dataset was divided into two parts: a training set containing 70% of the samples and a test set containing the remaining 30%. This split was stratified. Indeed, the distribution of the classes of the target variable "Sleep disorder" was imbalanced (see Figure 4). Then, to build the best possible model, a random search was implemented to test several random combinations of the hyperparameters of each model. The advantages of random search over grid search include significant time savings and the ability to test an unlimited number of hyperparameters. In total, 150 random combinations of hyperparameters per model were performed. These 150 different models were then evaluated by dividing the training set using the k-fold cross-validation method, which is appropriate here due to the small dataset size. The number of folds was set to 5, and care was taken to ensure that the k-fold cross-validation was also stratified so that each fold had a proportional distribution of classes. Otherwise, the risk is that the model might appear performant simply because it has learned well to predict the majority class while failing to generalize for the minority classes. Once the optimal combination of hyperparameters was found based on accuracy, the "optimum" model of each of the 7 types of classifiers was retrained on the entire training set and finally evaluated on the test set.

	Sleep Disorder	Count	Percentage
1	Insomnia	77	20.59
2	None	219	58.56
3	Sleep Apnea	78	20.86

Table 1 - Distribution of classes in the target variable “Sleep disorder”

3. Methods

a. *Chosen metrics*

To compare each of the 7 classifiers, I chose to use accuracy, defined as the fraction of samples that are correctly classified, and average AUC (Area Under the Curve), defined, in the case of a multiclass problem, as the average of the AUC of each of the one vs one or one vs rest scenarios. Here, the selected strategy is one vs rest. Thus, for each class (Insomnia, Sleep Apnea, and None), the ROC curve is plotted where this class is considered the positive class and the other two as the negative class. From there, the AUC of each of the three cases is calculated, and the final result for a model is the average of the three AUCs.

b. *Comparison of model performances*

Figure 5 shows the performance results of the “optimum” models of each classifier type. Each of these models performs well on the test set with a minimum accuracy of 0.838 (KNN and Naive Bayes) and a minimum average AUC of 0.817 (Gradient Boosting). The Multinomial Logistic Regression stands out particularly with the best results in both accuracy (0.865) and average AUC (0.9). Therefore, the Multinomial Logistic Regression was selected as the best model.

	Model	Accuracy	Average AUC
1	Logistic Regression	0.8649	0.8998
2	Naive Bayes	0.8378	0.8284
3	KNN	0.8378	0.855
4	Random Forest	0.8559	0.8386
5	SVM	0.8559	0.8522
6	Gradient Boosting	0.8468	0.8166
7	MLP	0.8559	0.878

Table 2 - Results of the 7 "optimum" classifiers of each model type

c. *Verification of the uniformity of the best model's performance*

The risk of working with the average AUC and accuracy is that, ultimately, behind these figures, the model may not treat each class equally in terms of precision, specificity, and sensitivity. Thus, the model's results could be biased by one class outperforming the other two. It is therefore necessary to verify if the results of the multinomial logistic regression accurately reflect the overall performance of the model across the 3 classes. I first verified the uniformity of the average AUC across the 3 classes. For this, we can plot the three ROC curves corresponding to the 3 one-vs-rest scenarios of each class on the same graph (see Figure 6). We can first observe that each of these curves is close to the top-left corner of the graph, indicating good overall performance. Moreover, their overlap means that the model is balanced because it discriminates uniformly between the classes. Thus, the average AUC is a good measure of the discrimination of the multinomial logistic regression model. Its high value is reliable and representative. As for accuracy, displaying its confusion matrix gives a direct view of the uniformity of the error distribution (see Figure 7). As can be seen, the success rates of each class are almost similar. Thus, the model has a generally uniform precision for each class.

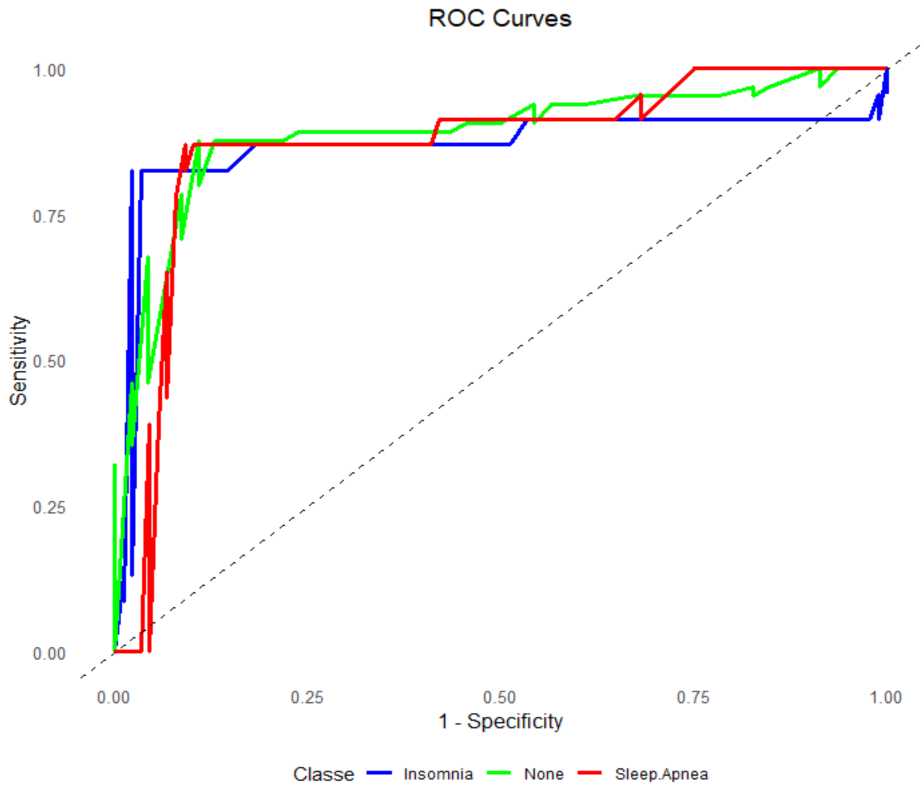

Figure 4 - ROC Curves for each one-vs-rest scenario of each class

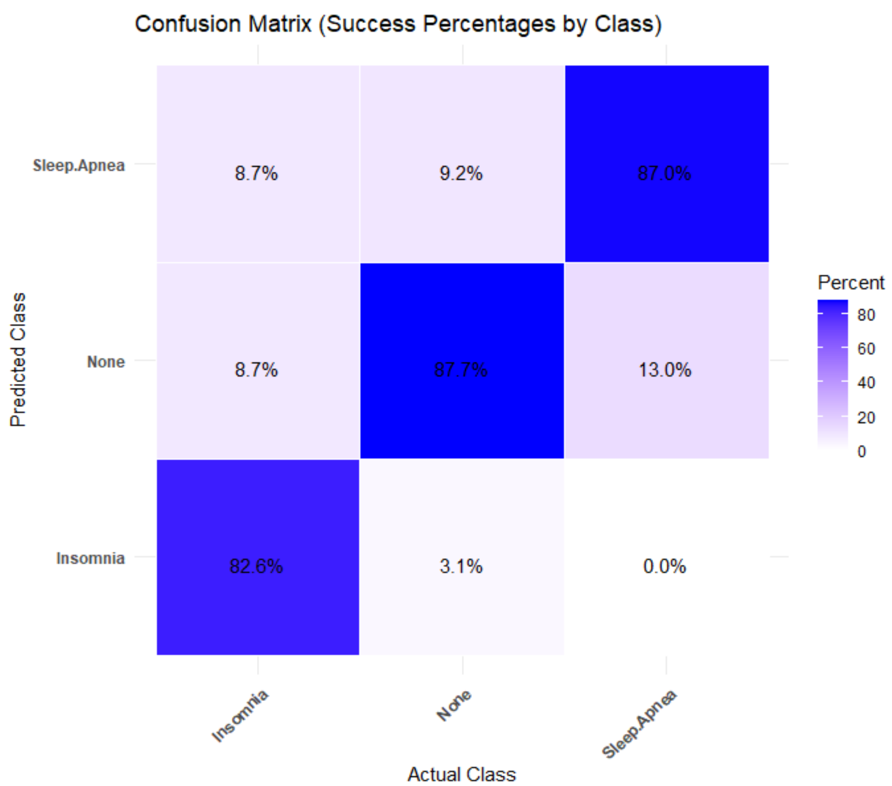

Figure 5 - Confusion Matrix (success rate by class)

d. Evaluation of the influence of behaviors on the prediction model

After obtaining the most reliable model possible, I was able to evaluate the influence of behavioral features on the prediction of insomnia and sleep apnea. To estimate this impact, I decided to calculate the Mean Absolute SHAP Values. This was justified because I wanted to know the average contribution of each feature across all predictions. For this calculation, I grouped the features related to behavior, namely: “Occupation”, “Daily Steps” and “Physical Activity Level” and I also grouped the features related to the patient's physiology: “Blood Pressure”, “BMI Category” and “Heart Rate.” Thus, in addition to understanding the influence of behavioral features, I wanted to compare these two categories of features to see whether or not physiological features indeed have much more influence than behavioral features. Therefore, for these two groups of features, we calculate the average of the Mean Absolute SHAP Values of the features that compose these two categories. Finally, we obtain the graph in Figure 8. The values are displayed in Figure 9. Thus, the behavioral features have a Mean Absolute SHAP Value of 0.0055, which corresponds to the third most influential category of features on this model's predictions. The physiological features have a Mean Absolute SHAP Value of 0.0069, the highest value among all features.

As presented in the introduction, these physiological features are favored by health professionals and researchers in the development of predictive models for sleep disorders. To determine if the difference in values between the Mean Absolute SHAP Values of the “Comportment features” and “Physiological features” is significant, I conducted a t-test. For this, the SHAP Values data must be normally distributed, and there must be homogeneity of variances between the two groups (“Comportment features” and “Physiological features”). To verify the first criterion, I plotted the Q-Q plot curves of the SHAP Values data for these two groups. Since the points on these two graphs approximately follow a straight line, the data can be considered normally distributed. To confirm this conclusion, I conducted a Shapiro-Wilk test. I obtained a p-value of 0.06674 for the “Physiological features” and a p-value of 0.4475 for the “Comportment features.” Since both p-values are greater than 0.05, the null hypothesis of normality cannot be rejected. Thus, the SHAP Values data can be considered normally distributed. For the homogeneity of variances, I conducted a Levene's test which gave a p-value of 0.8331, greater than 0.05. Thus, this hypothesis cannot be rejected. Therefore, I was able to perform a t-test to compare the Mean Absolute SHAP Values. With this test, I obtained a p-value of 0.3469, above the significance threshold of 0.05. Thus, the null hypothesis cannot be rejected. There is therefore no statistically significant difference between the Mean Absolute SHAP Values of the "Comportment features" and the "Physiological features". In conclusion, although the "Physiological features" have a higher Mean Absolute SHAP Value (0.0069 versus 0.0055), this difference is not sufficient to be statistically significant. Ultimately, this suggests that these two types of features have comparable importance in this sleep disorder prediction model and that each should be considered.

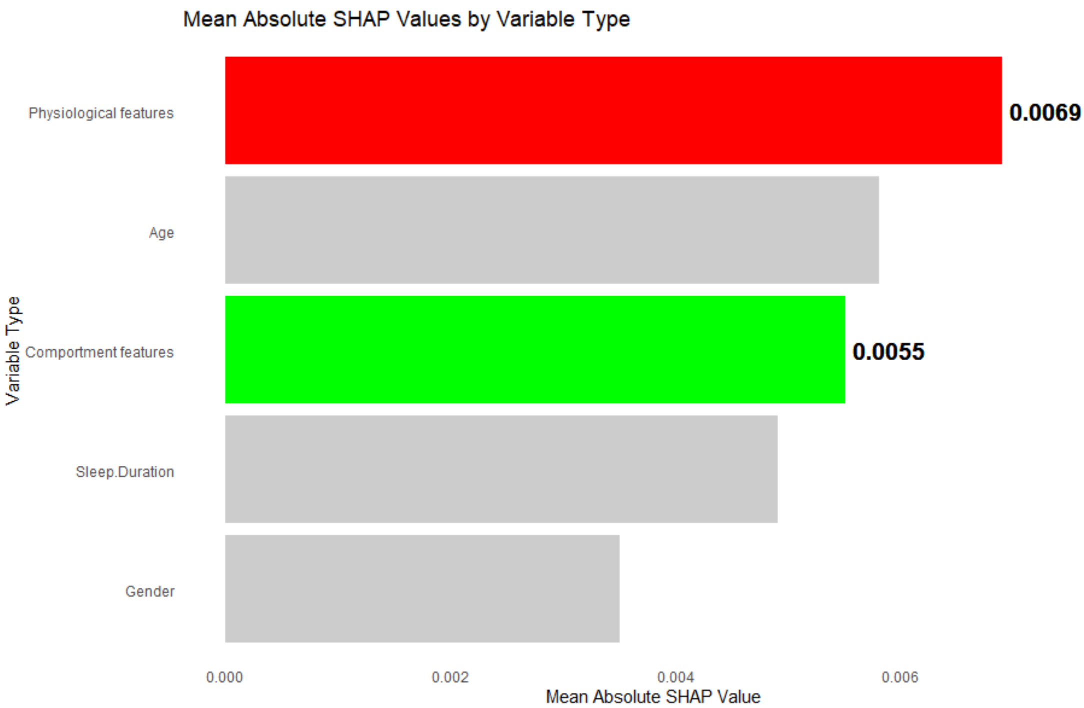

Figure 6 - Mean Absolute SHAP Values by features (Graph)

variable_type	mean_abs_shap
Age	**0.0058**
Comportment features	**0.0055**
Gender	**0.0035**
Physiological features	**0.0069**
Sleep.Duration	**0.0049**

Table 3 - Mean Absolute SHAP Values by features (Table)

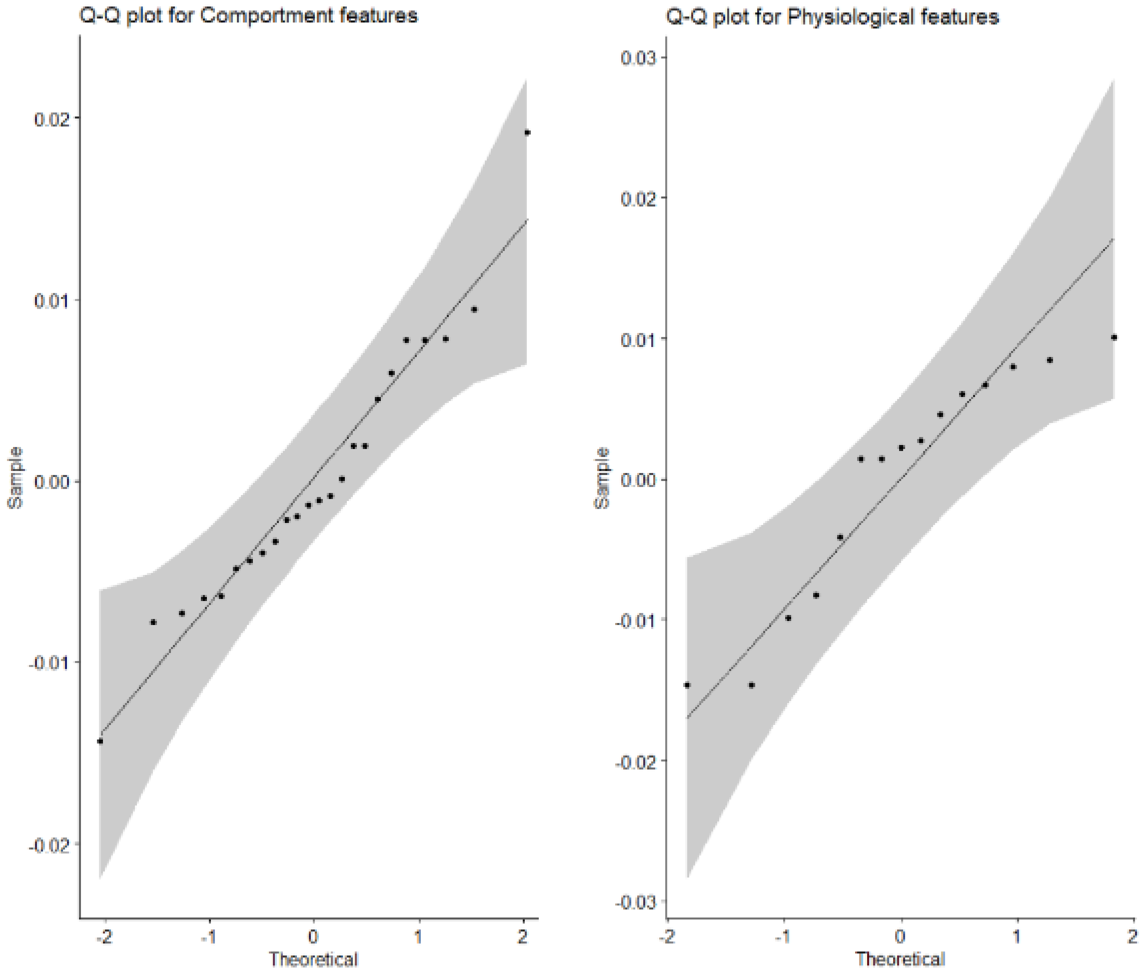

Figure 7 et 8 - Q-Q plot for Comportment (left) and Physiological features (right)

4. Discussion

a. Related works

Predicting the presence of a sleep disorder is a major concern for several scientific studies. In particular, sleep apnea and insomnia have been the subject of numerous and diverse implementations of machine learning prediction models.

Thus, Li S, Zhu P, Cai G, Li J, Huang T, and Tang W [13] used 165 responses from the Spiegel Sleep Questionnaire (SSQ) to evaluate the severity of insomnia in these patients. Random Forest Classifier (RFC), Support Vector Classifier (SVC), and K-Nearest Neighbors (KNN) models were then trained to predict the severity score of insomnia, and the results showed that the RFC was the most performant.

Regarding sleep apnea, Mencar C, Gallo C, Mantero M, and others [14] differently used a dataset of 313 patients containing 32 features, most related to physiological measurements and symptoms. After selecting 19 features by Principal Component Analysis (PCA), they split their project into two parts: on one side, comparing classification models capable of predicting the severity class of OSA, and on the other side, comparing regression models capable of directly predicting the Apnea Hypopnea Index (AHI). Evaluated using the k-fold cross-validation method, the SVM achieved the most balanced results in terms of accuracy, precision, and recall for the classification problem. For the regression problem, the SVM was again the most performant model with the lowest Root Mean Squared Error (RMSE) regardless of the number of input features.

An example of a study that combines these two sleep disorders is the one conducted by Ha S, Choi SJ, Lee S, Wijaya RH, Kim JH, Joo EY, and Kim JK [15]. Their objective was to predict the risk of a person having OSA, insomnia, or COMISA. With a total of 9 features, including 5 from a sleep-related questionnaire and 4 demographic features, an XGBoost classification model was trained. Its performances are excellent, regardless of the predicted class, with almost all AUC values above 0.9. Comparison with a Linear Regression, a Random Forest and an SVM showed that XGBoost was the best of all.

What all these studies have in common is the almost exclusive use of physiological features and the absence of behavioral features to build the prediction models. Yet, daily lifestyle choices such as lack of physical activity or irregular meal times play a role in the development of sleep disorders [16]. This is why, in this study, I evaluated the influence of behavioral features on the prediction of sleep disorders using machine learning tools. After extensive research, it appears that no similar work exists in the literature.

b. *Limitations*

Regarding the limitations of my study, a notable flaw is the small size of my dataset. Thus, the results of my study may be questioned in terms of representativeness for a larger population of patients. One can also question the decision not to include the variable "Stress Level" in the "Comportment features." Indeed, stress can be defined as a response of the organism to external life events. However, studies from the last 20 years have shown that genetic factors, such as the monoamine oxidase A genes, serotonin transporter, CRF receptor-1, ADRA2B [17], can also contribute to a higher sensitivity to stress. This duality between external and internal factors in the genesis of stress led me not to include it in the behavioral features.

5. Conclusion

In this study, based on a dataset composed of physiological features, behavioral features, and demographic features, I developed several prediction models for sleep disorders, namely insomnia and apnea. To build the most performant models of each type, a strategy of 150-trial random search coupled with stratified k-fold cross-validation was carried out to optimize the hyperparameters of these algorithms. Following this, a comparison between these classifiers in terms of accuracy and average AUC on the test set allowed for the selection of the best and most reliable one: the Multinomial Logistic Regression. Finally, to address the issue raised by this study, an analysis of the Mean Absolute SHAP Value of behavioral features indicated that it is not significantly different from the Mean Absolute SHAP Value of physiological features, and consequently, their importance in constructing prediction models for sleep disorders is identical.

Additional work can be considered on this subject. A more in-depth study on the individual impact of various lifestyle choices, whether related to physical activity, diet, or the patient's occupation, would allow designers of predictive models for sleep disorders to have a comprehensive view of the specific nature of behaviors that these models should or should not include. In another area, this study focuses exclusively on the prediction of insomnia and apnea, both considered the most frequent sleep disorders. Therefore, an additional study including other sleep disorders could be interesting.

6. References

[1] Karna B, Sankari A, Tatikonda G. Sleep Disorder. [Updated 2023 Jun 11]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2024 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK560720/

[2] Thorpy MJ. Classification of sleep disorders. Neurotherapeutics. 2012 Oct;9(4):687-701. doi: 10.1007/s13311-012-0145-6. PMID: 22976557; PMCID: PMC3480567.

[3] Gottlieb DJ, Punjabi NM. Diagnosis and Management of Obstructive Sleep Apnea: A Review. *JAMA.* 2020;323(14):1389–1400. doi:10.1001/jama.2020.3514

[4] Salamah, F. S. B., Alfayez, H. M. A., & Melibary, R. T. (2022). Diagnosis and treatment of obstructive sleep apnea. *International Journal Of Community Medicine And Public Health*, *9*(2), 987–994. https://doi.org/10.18203/2394-6040.ijcmph20220016

[5] Riemann D, Espie CA, Altena E, Arnardottir ES, Baglioni C, Bassetti CLA, Bastien C, Berzina N, Bjorvatn B, Dikeos D, Dolenc Groselj L, Ellis JG, Garcia-Borreguero D, Geoffroy PA, Gjerstad M, Gonçalves M, Hertenstein E, Hoedlmoser K, Hion T, Holzinger B, Janku K, Jansson-Fröjmark M, Järnefelt H, Jernelöv S, Jennum PJ, Khachatryan S, Krone L, Kyle SD, Lancee J, Leger D, Lupusor A, Marques DR, Nissen C, Palagini L, Paunio T, Perogamvros L, Pevernagie D, Schabus M, Shochat T, Szentkiralyi A, Van Someren E, van Straten A, Wichniak A, Verbraecken J, Spiegelhalder K. The European Insomnia Guideline: An update on the diagnosis and treatment of insomnia 2023. J Sleep Res. 2023 Dec;32(6):e14035. doi: 10.1111/jsr.14035. PMID: 38016484.

[6] Krystal AD, Prather AA, Ashbrook LH. The assessment and management of insomnia: an update. World Psychiatry. 2019 Oct;18(3):337-352. doi: 10.1002/wps.20674. PMID: 31496087; PMCID: PMC6732697.

[7] Ragnoli, B.; Pochetti, P.; Raie, A.; Malerba, M. Comorbid Insomnia and Obstructive Sleep Apnea (COMISA): Current Concepts of Patient Management. *Int. J. Environ. Res. Public Health* 2021, *18*, 9248. https://doi.org/10.3390/ijerph18179248

[8] Sweetman, A. Co-morbid Insomnia and Sleep Apnoea (COMISA): Latest Research from an Emerging Field. *Curr Sleep Medicine Rep* **9**, 180–189 (2023). https://doi.org/10.1007/s40675-023-00262-9

[9] Bedoya, O.; Rodríguez, S.; Muñoz, J.P.; Agudelo, J. Application of Machine Learning Techniques for the Diagnosis of Obstructive Sleep Apnea/Hypopnea Syndrome. *Life* 2024, *14*, 587. https://doi.org/10.3390/life14050587

[10] https://www.kaggle.com/datasets/uom190346a/sleep-health-and-lifestyle-dataset/data

[11] Notes on Machine Learning © 2024 by Claudio Cusano

[12] Kundu, Ramendra & Biswas, Subir & Das, Mithun. (2017). Mean Arterial Pressure Classification: A Better Tool for Statistical Interpretation of Blood Pressure Related Risk Covariates. Cardiology and Angiology: An International Journal. 6. 1-7. 10.9734/CA/2017/30255.

[13] Li S, Zhu P, Cai G, Li J, Huang T, Tang W. Application of machine learning models in predicting insomnia severity: an integrative approach with constitution of traditional Chinese medicine. Front Med (Lausanne). 2023 Oct 19;10:1292761. doi: 10.3389/fmed.2023.1292761. PMID: 37928471; PMCID: PMC10625410.

[14] Mencar C, Gallo C, Mantero M, Tarsia P, Carpagnano GE, Foschino Barbaro MP, Lacedonia D. Application of machine learning to predict obstructive sleep apnea syndrome severity. Health Informatics J. 2020 Mar;26(1):298-317. doi: 10.1177/1460458218824725. Epub 2019 Jan 30. PMID: 30696334.

[15] Ha S, Choi SJ, Lee S, Wijaya RH, Kim JH, Joo EY, Kim JK. Predicting the Risk of Sleep Disorders Using a Machine Learning-Based Simple Questionnaire: Development and Validation Study. J Med Internet Res. 2023 Sep 21;25:e46520. doi: 10.2196/46520. PMID: 37733411; PMCID: PMC10557018.

[16] Review on Sleep Disorder Prediction Using Machine Learning Ajay Santhosh , George. M. Karakattu , Muhammed Ansal T.A , Muhammed Roshan and Shiju Shaikh Manakkulam ISSN 2582-7421 VOLUME 5, Issue 2, February 2024

[17] Fink, George. (2010). Stress: Definition and history. Stress Sci. 3-9. 10.1016/B978-008045046-9.00076-0.