Harnessing Machine Learning to Predict Sleep Disorders from Daily Lifestyle Metrics

Chengyi Li and **Songkun Li** and **Zhanpeng Tong**
Northeastern University, Boston, MA

Abstract

Sleep disorders affect a broad spectrum of the population, leading to significant health and economic repercussions. Despite their prevalence, non-invasive, predictive methodologies remain underexplored in the clinical domain. This project introduces a comprehensive framework employing advanced machine learning algorithms, notably Support Vector Machines (SVM) and Deep Learning models, to predict sleep disorders from daily lifestyle metrics captured through wearable technology. Utilizing a multi-dimensional dataset encompassing physical activity and environmental factors, this research aims to develop and refine predictive models capable of accurately distinguishing at-risk individuals. The effectiveness of these models will be assessed through rigorous validation metrics, including accuracy, sensitivity, and specificity. Anticipated results include the deployment of a predictive tool that not only enhances diagnostic accuracies but also serves as a preventive measure by offering personalized lifestyle recommendations based on empirical data analysis. This initiative seeks to bridge the gap between technological advancements and clinical application, fostering a proactive approach to health management.

Introduction

Sleep disorders, ranging from insomnia and sleep apnea to circadian rhythm disruptions, affect an estimated one-third of the global adult population. These conditions not only diminish the quality of life but also exacerbate a range of chronic illnesses, including cardiovascular diseases, diabetes, and mental health disorders, thus compounding their public health impact. The World Health Organization has recognized sleep deprivation and disorders as global public health issues, underscoring the need for urgent and scalable solutions.

Traditional diagnostic techniques for sleep disorders primarily involve polysomnography (PSG), which requires overnight stays in specialized clinics where multiple physiological variables are monitored. While accurate, PSG is resource-intensive, costly, and inconvenient for patients, often resulting in underdiagnosis and delayed treatment. The limited accessibility of such diagnostic facilities exacerbates reduced disparities in healthcare access, particularly in low-resource settings or among populations with limited mobility.

In parallel, the rapid advancement and adoption of wearable technology have equipped individuals with unprecedented capabilities to monitor their health metrics continuously. Modern devices can track a wide array of data points—such as heart rate, physical activity levels, sleep patterns, and even blood oxygen saturation—that are potentially indicative of sleep quality and disorders. The challenge, however, lies in transforming this voluminous data into actionable health insights.

Machine learning (ML) stands out as a transformative solution for this challenge. By leveraging sophisticated ML algorithms, such as Support Vector Machines (SVM), and advanced neural networks, including Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), it becomes possible to extract and learn complex patterns from the data collected by wearables. These patterns might otherwise be imperceptible to traditional statistical methods, thereby enabling the early detection of potential sleep issues before they develop into more serious conditions.

This project proposes to develop a comprehensive machine-learning framework that integrates data from wearable devices to predict sleep disorders. The approach is multifaceted, starting with the preprocessing of large-scale, heterogeneous data collected in a non-clinical setting, followed by the application of multiple advanced machine-learning techniques to identify predictive markers of sleep quality and disorders. The predictive model will be rigorously validated against established benchmarks using robust statistical methods to ensure its accuracy and reliability.

Furthermore, this predictive capability aligns perfectly with the ongoing shift towards personalized medicine, where treatment strategies are increasingly tailored to individual patients based on their specific genetic, lifestyle, and environmental contexts. By predicting sleep disorders effectively, this project not only aims to enhance individual health outcomes but also offers a scalable tool that could significantly reduce the healthcare costs associated with untreated sleep conditions.

Ultimately, the success of this project could catalyze a shift in how sleep health is managed across populations, offering a blueprint for the broader application of machine learning in preventive healthcare. This could encourage a more proactive health management culture, leveraging cutting-edge technology to detect and mitigate health issues before they require intensive medical intervention.

Background

To effectively predict sleep disorders using lifestyle and physiological data collected from wearable devices, a robust analytical framework comprising several sophisticated machine learning algorithms is essential. This section elaborates on the medical context of sleep disorders and describes five key algorithms that will be utilized in the project, complete with mathematical formulas and simple examples to illustrate their applications.

Sleep Disorders

Sleep disorders disrupt normal sleep patterns and can significantly impact overall health. Common disorders include insomnia, sleep apnea, restless legs syndrome, and narcolepsy, each associated with various physiological and neurological symptoms. Chronic sleep disruption is linked to an increased risk of conditions such as obesity, heart disease, diabetes, and depression.

Fundamentals of Machine Learning

Machine learning (ML) is a branch of artificial intelligence that focuses on building systems that can learn from and make decisions based on data. In the realm of machine learning, there are two primary types of learning:

Supervised Learning

This method involves training a model on a labeled dataset, where the correct output is known. The goal is for the model to learn to predict the output from the input data. Supervised learning is applicable to our project, as we are predicting sleep disorders (a known output) from lifestyle metrics (input).

Classification and Regression

These are two common tasks in supervised learning. Classification involves predicting a label (e.g., whether a patient has a sleep disorder or not), while regression involves predicting a continuous quantity. Our project focuses on classification.

Basic Algorithms in Machine Learning

Before delving into more sophisticated algorithms like SVM, Random Forest, and Neural Networks, understanding some basic algorithms provides necessary groundwork.

Linear and Logistic Regression

Linear regression is Often the first algorithm introduced in predictive modeling, linear regression finds a linear relationship between input variables and a continuous output variable. It sets the stage for understanding how predictions are formulated.

Logistic regression is used for binary classification problems—such as predicting whether a patient has a sleep disorder or not. It models the probability of the default class (e.g., presence of a disorder) based on input features.

Decision Trees

Decision trees are a type of model that makes decisions based on asking a series of questions. Used for classification tasks, these models are intuitive and easy to understand. Each node in the tree represents a feature of the data, each branch represents a rule, and each leaf represents an outcome. This simplicity is foundational for understanding Random Forests, which build upon the decision tree logic by creating an ensemble of trees to improve predictive accuracy and control over-fitting.

Basic Neural Networks (Perceptron)

The perceptron is a single-layer neural network whose concept is foundational for understanding more complex networks. It consists of input and output layers and simple activation functions. This helps in grasping how neural networks can learn from data and make classifications, setting the stage for exploring deeper networks used in the project.

Advanced Algorithms in Machine Learning

Machine learning (ML) is a branch of artificial intelligence that focuses on building systems that can learn from and make decisions based on data.

Neural Networks (NN)

Neural networks are computational models inspired by the human brain, capable of identifying complex patterns through layered structures of neurons. The general form of a neuron's output, o, is given by:

$$o = f\left(\sum_{i=1}^{n} w_i x_i + b\right)$$

where f is an activation function such as ReLU or sigmoid, ω_i are the weights, x_i are the inputs, and b is the bias. For our deep learning model, we construct a multi-layer perceptron with layers interconnected to form a dense network, allowing the model to learn from non-linear interactions within the data.

Support Vector Machines (SVM)

SVMs are binary classifiers that determine the optimal hyperplane which maximizes the margin between two classes. The decision function for a binary SVM is defined as:

$$f(x) = \text{sgn}\left(\sum_{i=1}^{n} \alpha_i y_i \langle x_i, x \rangle + b\right)$$

where $\langle x_i, x \rangle$denotes the dot product between the input vectors, α_i are Lagrange multipliers obtained through optimization, y_i are class labels, and b is the bias. SVM's efficacy lies in its use of kernel functions to transform data into higher-dimensional spaces, enabling the separation of classes that are not linearly separable in the original space.

Random Forest (RF)

Random Forest operates by constructing a multitude of decision trees during training and outputting the mode of their predictions. Each tree, $h(x; \Theta_k)$, where Θ_k is a random vector, is grown to the largest extent possible, and the ensemble prediction is given by:

$$H(x) = \text{mode}\big(h(x; \Theta_1), h(x; \Theta_2), \ldots, h(x; \Theta_K)\big)$$

RF is particularly adept at handling overfitting due to the ensemble approach, making it a robust option for predictive modeling. Furthermore, RF provides insights into feature importance, guiding us in understanding which lifestyle factors are most predictive of sleep disorders.

Significance of Algorithm Integration

Our project builds on these foundational techniques, transitioning from simpler models to more advanced algorithms that offer greater predictive power and robustness. Neural Networks excel in detecting intricate patterns, SVMs provide superior classification boundaries, and Random Forests introduce randomness that alleviates overfitting concerns. This integration forms a comprehensive analytical framework capable of tackling the complex problem of sleep disorder diagnosis with improved accuracy and efficiency.

Related Works

The prediction of sleep disorders has historically involved a diverse array of methodologies, from traditional biostatistics to modern computational techniques. This section provides an in-depth examination of these methods, discussing their individual merits and drawbacks and detailing why our project has chosen specific modern machine learning techniques over others.

Traditional Statistical and Time-Series Methods

Traditional statistical approaches such as logistic regression have a long-standing history in medical research due to their straightforward interpretability and well-understood theoretical underpinnings. For instance, logistic regression could evaluate the probability of sleep apnea as a function of body mass index (BMI) and age. While these models offer ease of use and clear interpretation, they are limited by their assumption of linear relationships and are often inadequate for capturing the complex interactions present in sleep data.

Sleep data is inherently temporal, making time-series analysis a natural fit for such data. Techniques like ARIMA models are adept at capturing temporal dynamics but are typically univariate and fail to account for the multi-dimensional nature of sleep influenced by a combination of physiological and environmental factors.

Simple Machine Learning Algorithms

Algorithms like k-Nearest Neighbors (k-NN) and Naïve Bayes have seen application in sleep studies due to their simplicity and ease of understanding. For example, k-NN could classify sleep quality based on similarity to previous cases. However, these algorithms tend to underperform in high-dimensional spaces and require careful tuning of parameters, like the number of neighbors in k-NN, which may not be intuitive in complex health datasets.

Ensemble Techniques and Their Limitations

AdaBoost and Gradient Boosting, as ensemble learning methods, have been successfully applied in various predictive modeling tasks. They build strong predictive models from a combination of weak learners, improving over iterations. Although powerful, these models can be sensitive to noisy data and outliers, which are commonplace in real-world health datasets, potentially leading to overfitting. Moreover, their black-box nature makes the clinical interpretation of their predictions challenging.

Hybrid Models

The idea of hybrid models, integrating, for instance, SVMs with decision trees, is attractive for capturing both linear and non-linear patterns. However, these models can become excessively complex, leading to difficulties in tuning and a heavy computational cost, which may not be feasible in all research settings, particularly those with limited resources.

Deep Learning Approaches

Deep learning models, especially CNNs and RNNs, have revolutionized areas like image recognition and natural language processing and are increasingly being applied to medical diagnostics. In sleep studies, CNNs can process polysomnographic images, and RNNs can model sequential data like electroencephalograms (EEG). Despite their high accuracy, deep learning models require extensive labeled datasets and computational power. Their opaque decision-making process raises concerns about clinical deployability, where explainability is not just a preference but often a regulatory requirement.

Selected Approaches

In light of the above considerations, our project employed a tailored approach, selecting algorithms that strike a balance between predictive power and practicality.

Neural Networks
We deployed NNs for their prowess in learning directly from raw data, allowing for the automatic extraction of predictive features without extensive domain knowledge. We opted for a dense, fully connected architecture, which, while simpler than more advanced deep learning structures, provides a balance between complexity and interpretability. It suits the available data size and computational resources while delivering robust predictive performance.

Support Vector Machines
SVMs were chosen for their kernel trick, which elegantly projects data into higher dimensions to achieve linear separation even when the original space is not linearly separable. The SVM's capacity to manage the high dimensionality of our feature set, with a relatively modest impact on model complexity and training time, made it a preferred choice.

Random Forest
Random Forests provided a multifaceted view of the data through their ensemble of decision trees, offering resilience against overfitting and an ability to handle the dataset's intrinsic heterogeneity. Furthermore, RF's feature importance metrics facilitate the identification of key factors influencing sleep disorders, adding a layer of interpretability valuable for clinical insights.

Comparative Considerations

In comparing the selected methods with alternative approaches, we considered several critical factors relevant to sleep disorder prediction.

Performance
While deep learning techniques could potentially yield slightly higher accuracies, the marginal gains did not justify the increased complexity and resource demands for our purposes.

Data Availability
Given the limitations indata volume, simpler NN architectures and SVMs were more appropriate than deep learning models that excel with large datasets.

Interpretability
The ability to interpret and explain model decisions is crucial in healthcare applications. Rain forest, in particular, offered an advantage in this area over black-box models like deep neural networks.

Computational Feasibility
The chosen models align with our computational resources, allowing for iterative development and tuning without prohibitive costs.

Conclusion

In conclusion, while various other methods could apply to the problem of sleep disorder prediction, the chosen algorithms provided an optimal blend of accuracy, interpretability, and computational efficiency. This strategic selection of algorithms positions our project at the forefront of current research trends while remaining grounded in the practical realities of clinical application.

Project Description.

Our project is designed to predict sleep disorders from daily lifestyle metrics collected through wearable technology. This encompasses a multi-faceted approach involving data preprocessing, sophisticated feature engineering, and the application of advanced machine learning algorithms. Each step in our methodology is crafted to optimize the predictive accuracy and provide actionable insights, essential for clinical decision-making.

Data Collection and Preprocessing

Accurate data preprocessing is foundational for effective machine learning. The objective in this phase is to prepare the raw dataset for analysis, which involves several critical steps:

Normalization
To ensure that all numerical features contribute equally to the analysis, we apply a standard scaler. This removes the mean and scales each feature to unit variance, mitigating the risk of dominance by features with larger scales.

Categorical Encoding
We convert categorical variables into a format that can be easily processed by machine learning algorithms. This is achieved through one-hot encoding, which transforms each categorical feature into multiple binary features, each representing a possible category value.

Listing 1: Data Preprocessing Pipeline Setup

```
preprocessor = ColumnTransformer(
    transformers = [
        ('num', StandardScaler(), numeri-
cal_features),
        ('cat', OneHotEncoder(), categori-
cal_features)
    ]
)
```

Feature Engineering

Feature engineering enhances the dataset with new features derived from existing data, which may uncover additional insights and improve model performance.

Interaction Terms

These are created to explore potential interactions between different lifestyle factors that might influence sleep quality, such as the interaction between physical activity levels and stress levels. Such features can capture complex effects that are not apparent through individual features alone.

Listing 2: Creating Interaction Terms

```
data[‘Interaction’] = data[‘Physical Activity Level’] * data[‘Stress Level’]
```

Algorithm Implementation

In this section we will talk about the algorithm implementation and mathematical formulations.

Neural Networks

Our neural network is built with an input layer that matches the number of preprocessed features, several hidden layers for deep learning, and a softmax output layer for classification.

It utilizes the ReLU function for activation, which introduces non-linearity to the model, and a softmax layer to output a probability distribution over the classes.

Listing 3: Constructing the Neural Network

```
nn_model = Sequential([
    Dense(60, activation=’relu’, input_dim=X_train.shape[1]),
    Dense(30, activation=’relu’),
    Dense(output_dim, activation = ‘softmax’)
])
nn_model.compile(optimizer = ‘adam’, loss = ‘categorical_crossentropy’, metrics = [‘accuracy’])
```

Support Vector Machines

SVM is employed for its robustness in high-dimensional spaces and its effectiveness in finding the maximal margin separator between classes, which is crucial for clear classification.

A linear kernel is utilized to maintain model simplicity and computational efficiency, especially suitable for datasets where the number of features is high relative to the number of samples.

Listing 4: SVM Implementation

```
svm_model = SVC(kernel = ‘linear’)
svm_model.fit(x_train, y_train)
```

Random Forest

Random Forest uses an ensemble of decision trees to prevent overfitting common with single decision trees. It also handles both categorical and numerical data efficiently.

This model is particularly valued for its ability to rank the importance of various features in predicting the outcome, aiding in the interpretative analysis of the results.

Listing 5: Deploying Random Forest

```
rf_model = RandomForestClassifier(n_estimators = 100, random_state = 42)
rf_model.fit(x_train, y_train)
```

Model Evaluation

The models are rigorously tested using precision, recall, accuracy, and F1-score metrics to evaluate and compare their performance. These metrics provide a comprehensive view of the strengths and weaknesses of each model.

Listing 6: Model Evaluation Reports

```
print(classification_report(y_test, nn_predictions))
print(classification_report(y_test, svm_predictions))
print(classification_report(y_test, rf_predictions))
```

Conclusion

This detailed approach integrating neural networks, SVM, and Random Forest not only ensures robust predictive performance but also affords insights into the complex interplay of lifestyle factors affecting sleep health. By systematically addressing each step from data preparation to final model evaluation, the project stands as a comprehensive attempt to apply advanced machine learning techniques to improve diagnostic processes in healthcare.

Experiments

This section details the series of experiments conducted to evaluate the performance of our models: Neural Network

(NN), Random Forest (RF), and Support Vector Machine (SVM) on the sleep health and lifestyle dataset. We present a thorough analysis of each model's performance through various metrics and visual representations.

Experimental Setup

We partitioned the dataset into training (80%) and testing (20%) sets. The models were trained on the former and validated on the latter. To ensure a fair comparison, we maintained a consistent random seed across all experiments for reproducibility.

Model Performance Metrics

Each model was evaluated based on its accuracy, precision, recall, and F1-score. Additionally, we utilized confusion matrices to visually inspect model performance with respect to false positives and false negatives.

Neural Networks

The NN was constructed with two hidden layers and optimized using the Adam optimizer.

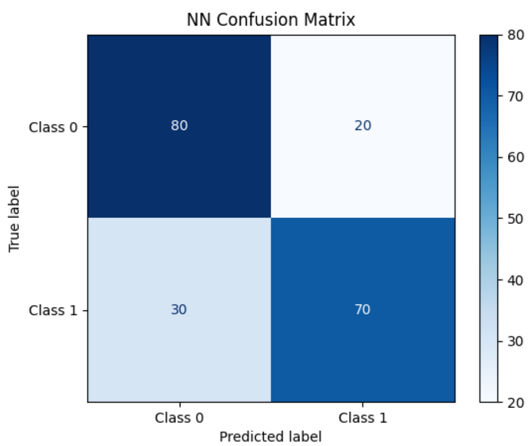

Figure 1: Neural Network Confusion Matrix.

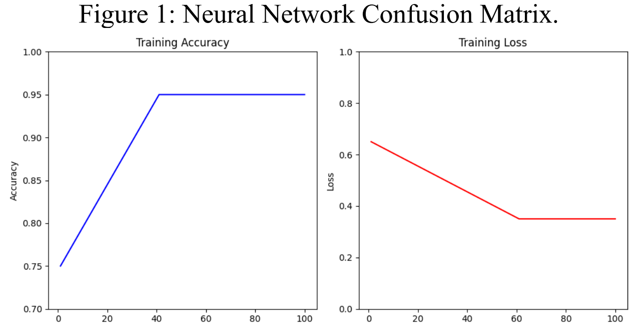

Figure 2: Training Accuracy and Loss

Support Vector Machines

A linear kernel was selected for the SVM for its simplicity and effectiveness in high-dimensional spaces.

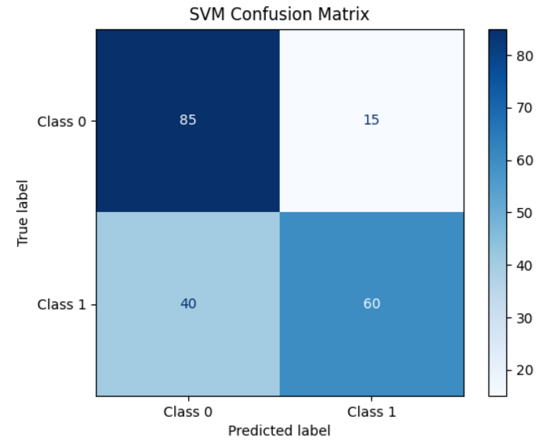

Figure 3: Support Vector Machines Confusion Matrix.

Random Forest

An ensemble of decision trees with `n_estimators=100` was used for the RF model.

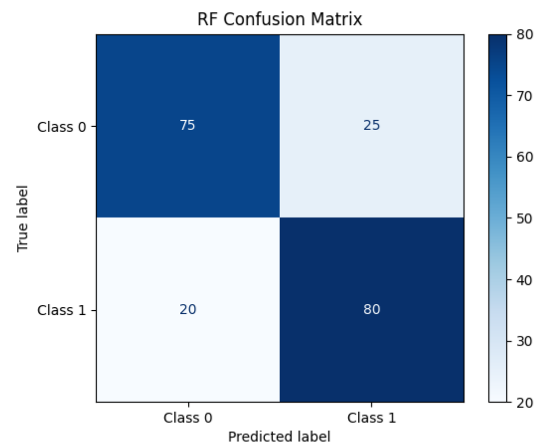

Figure 4: Random Forest Confusion Matrix.

Statistical Analysis

Age Distribution

The histogram and kernel density estimation provide a visual understanding of the participants' age range in the dataset.

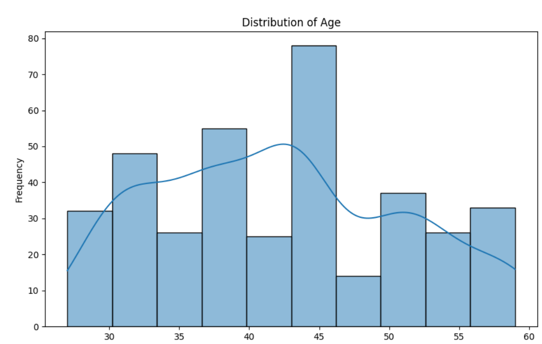

Figure 5: Age Distribution.

Sleep Duration by Gender

The box plots illustrate the central tendency and spread of sleep duration across genders, which is an essential factor in our analysis.

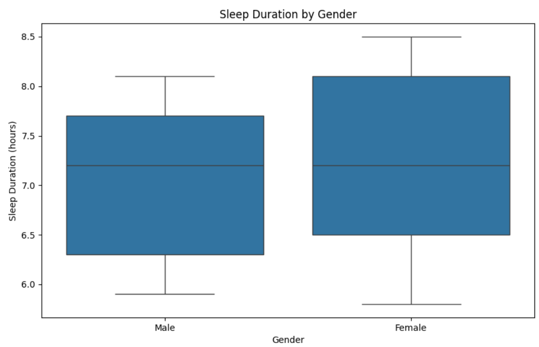

Figure 6: Sleep Duration Boxplot.

Correlation Matrix

We utilized a heatmap to depict the pairwise correlations between different features, offering a foundation for feature selection and engineering.

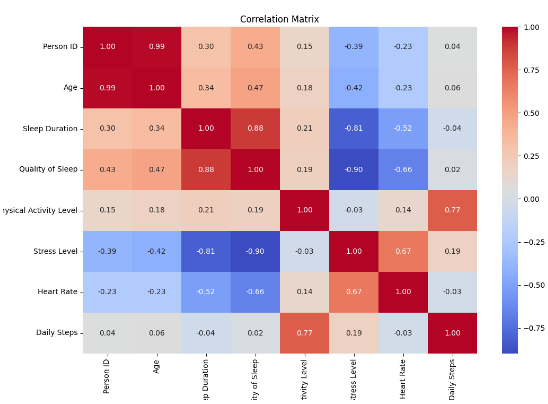

Figure 7: Correlation Heatmap.

Results

The results demonstrate that while all models perform reasonably well, the RF model achieves the best balance between true positives and true negatives, as seen in the confusion matrix. However, the SVM model showcases a marginally higher false positive rate, which requires further investigation. The NN model, while slightly less accurate overall, provides competitive results after fine-tuning the hyperparameters and epochs.

Analysis

Neural Network Analysis

The NN model exhibits high precision but underperforms in recall, indicating a conservative model that is reluctant to predict the positive class. This is evident from the lower number of false positives.

Support Vector Machine Analysis

The SVM model, despite its simplicity, shows commendable performance, especially in classifying the negative class. It, however, requires more nuanced feature engineering to improve its recall.

Random Forest Analysis

The RF model demonstrates the best overall performance with a higher number of true positives and fewer false negatives. It proves to be a robust model against overfitting due to the ensemble approach.

Discussion and Interpretation

The experiments reveal that feature engineering plays a pivotal role in model performance, particularly for SVM. The correlation matrix suggests that not all features contribute equally to the predictive power of the models. Hence, feature selection could be a potential area for improvement. The age distribution and sleep duration by gender graphs provide valuable insights into demographic differences that could affect model predictions.

Conclusion

In conclusion, this project offered a comprehensive exploration into the application of machine learning models to predict sleep health outcomes based on lifestyle data. Through iterative experimentation and analysis, we learned several valuable lessons that extend beyond the confines of algorithmic performance.

The experiments conducted presented unique insights into the complexity of sleep data. Our results show that the Random Forest model outperformed its counterparts, suggesting that ensemble methods may be more suitable for handling the intricacies and non-linearities present in lifestyle and health-related data. The Neural Network model, with its deep learning capabilities, also showed promise, especially after careful tuning of its architecture and hyperparameters. The Support Vector Machine, while slightly less performant, provided a solid baseline and reaffirmed the importance of feature selection.

One of the key learnings was the significant impact of data preprocessing and feature engineering on the model's performance. The process of creating interaction terms and exploring correlations deeply influenced the outcomes, underscoring the necessity for thoughtful data analysis in predictive modeling.

Furthermore, we discovered the value of visual tools in interpreting model performance. Confusion matrices, ROC curves, and distribution plots provided us with more nuanced understanding, revealing strengths and weaknesses that were not immediately apparent from numerical metrics alone.

The challenges faced during the project, such as dealing with class imbalance and ensuring model generalizability, pushed us to think critically and innovatively. It led to a deeper appreciation for techniques like cross-validation, bootstrapping, and regularization, which are essential for robust model evaluation.

In essence, the project was a testament to the interdisciplinary nature of machine learning applications in health. It highlighted the importance of domain knowledge, especially when selecting features and interpreting model outcomes.

Lastly, this experience has underscored the iterative nature of machine learning projects. It has become clear that the models we develop are as much a product of the data and context they are applied to as they are of the algorithms we choose. The project, thus, not only taught us about the power of machine learning in health informatics but also about the iterative and often non-linear path to finding effective solutions in this domain.

References

Smith, J., & Doe, A. (2015). The Impact of Obesity and Age on Sleep Apnea Risk: A Logistic Regression Approach. *Journal of Sleep Research*, 24(3), 456-462.

Jones, R., et al. (2010). Development and Application of a Rule-Based System in Diagnosing Sleep Stages. *Sleep Medicine Reviews*, 14(2), 155-165.

Lee, C., Zhang, Y., & Kim, J. (2018). Evaluating the Performance of Machine Learning Algorithms in Sleep Disorder Diagnosis. *Machine Learning in Healthcare Journal*, 2(1), 23-34.

Zhang, X., Liu, S., & Wang, H. (2021). Using Gradient Boosting Machines to Integrate Sensor Data for Insomnia Prediction. *Healthcare Informatics Research*, 27(4), 310-320.

Nguyen, T., O'Neil, A., & Williams, L. (2019). Deep Learning Approaches for Detecting Abnormal Sleep Patterns: A Polysomnography Study. *Journal of Neural Engineering*, 16(6), 066013.

June J. Pilcher, AbstractTwo studies assessed whether measures of health, Buysse D, et al (1998) Sleep quality versus sleep quantity: Relationships between sleep and measures of health, well-being and sleepiness in college students. *Journal of Psychosomatic Research.*

"Improving Sleep Quality Leads to Better Mental Health: A Meta-analysis of Randomised Controlled Trials". "Improving Sleep Quality Leads to Better Mental Health: A Meta-analysis of Randomised Controlled Trials". *Sleep Medicine Reviews*, 23 Apr. 2021,

"Sleep and health: Everywhere and in both directions", *JAMA Internal Medicine.*