Prediction of Apnea Hypopnea Index (AHI) for Obstructive Sleep Apnea Diagnosis based on Using Sleep Tracking Wearable Devices

Uyen Nguyen[1], Trung Le[1], Tony Ng[2]

[1]University of South Florida, [2]Bentley University

Introduction

Obstructive Sleep Apnea (OSA):

- One of the most common sleep breathing disorders
- Significant impact on public health and economy due to high prevalence & comorbidity consequences

Apnea Hypopnea Index (AHI):

- Measured by number of apnea & hypopnea events per hour over sleep time
- A principal metric to diagnose OSA severity

Polysomnography (PSG):

- Current "gold standard" for OSA diagnosis, still contains many constraints
- Barriers in having access to PSG: expensive costs, shortage of sleep specialist workforce & laboratories.
- PSG Intrusive setup: many electrodes - patients have low incentives to undergo PSG

=> Dire need for non-intrusive OSA monitoring with sleep tracking devices.

Challenges of OSA diagnosis using sleep tracking devices:

- Portable monitors use machine learning algorithms for automatic sleep stage classification.
- Trained & tested on public datasets, not validated from real lab datasets
- Inability to distinguish between sleep and wake & lack of sleep staging channels record

=> Accuracy remains constrained & measurement error is unavoidable

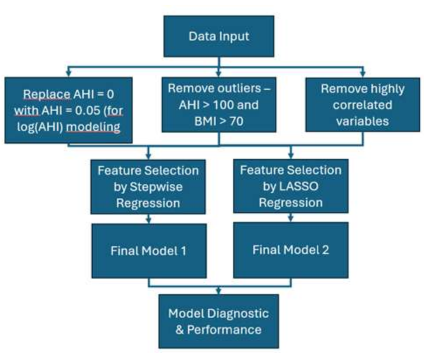
OBJECTIVES:

Hypnogram, a graph representing sleep stages over time, contains a wealth of information. Not many research done on how hypnograms are interpreted & which aspects are important.

In this paper, we identify variables collected from sleep wearable devices, that can also be collected from hypnogram, to predict the value of Apnea Hypopnea Index (AHI) of a patient.

Based on variables, we build a stepwise regression and LASSO models to evaluate the risk of a patient having OSA.

Methodology



Results & Discussion

Feature Selection

Step 1. Randomly divide the data set with 993 observations into ten approximately equal parts

Step 2. Use all 9 datasets beside the first data set to train regression model based on stepwise variable selection or LASSO regression approach

Step 3. Apply the trained model to predict AHI values and obtain the mean square errors (MSE) for this model

Step 4. Repeat the process for the rest of the datasets and obtain 9 more models and MSEs

Step 5. Determine the final model by selecting the variables that appear in more than six of the models (60% of the time)

Step 6. Compute the overall mean square error across 10 model as a value for comparing different models

Stepwise Regression
log(ahi) ~ gender + bmi + SPT + TST + onset + WK + REMper + REMsleep + REMlightoff

LASSO Regression
log(ahi) ~ gender + bmi + SPT + sleepeff + onset + N1 + REMper + REMlightoff

Variables	Stepwise Regression	LASSO Regression
Dependent variable: log(ahi)		
(Intercept)	-2.1869	-3.3117
Gender	0.8106	0.7784
bmi	0.1043	0.1041
SPT (Sleep period time): Elapsed time from sleep onset to last epoch of sleep	0.0059	0.0023
TST (Total sleep time): Time spent on sleeping during sleep period time	-0.0049	
sleepeff (Sleep efficiency): Ratio of total sleep time to sleep period time		-0.0049
Onset (Sleep onset): The duration of time when the subject undergoes a transition from wakefulness into sleep	-0.0015	-0.0016
WK (Percentage of Wake stage): Ratio of total Wake time over total sleep time	-0.0184	
N1 (Percentage of N1 stage): Ratio of total N1 time over total sleep time		0.0156
REMper (Percentage of REM stage): Ratio of total REM sleep time over total sleep time	-0.0281	-0.0243
REMsleep (REM latency from sleep onset): The duration from sleep onset to the first REM sleep	-0.0044	
REMlightoff (REM latency from light off): The duration from light off to the first REM sleep	0.0043	0.0001

Stepwise Regression

	Predicted			
Observed	Normal	Mild	Moderate	Severe
Normal	349	117	7	0
Mild	124	142	19	1
Moderate	47	88	12	4
Severe	14	48	10	7
Accuracy: 51.57%				

	Predicted	
Observed	Normal	OSA
Normal	349	124
OSA	185	331
Accuracy: 68.76%		
Sensitivity: 73.78%		
Specificity: 64.15%		

Stepwise Regression:

- For each cross validation, use stepAIC package in R to select the best model (lowest AIC)
- The combination of features in the final model are gender, bmi, SPT, TST, onset, WK, REMper, REMsleep, REMlightoff
- The average MSE across 10 models is 12.347
- The final model achieves an accuracy of 68.76%, sensitivity of 73.78% and specificity of 64.15%.

LASSO Regression

	Predicted			
Observed	Normal	Mild	Moderate	Severe
Normal	357	110	6	0
Mild	126	141	18	1
Moderate	48	90	9	4
Severe	14	45	13	7
Accuracy: 51.97%				

	Predicted	
Observed	Normal	OSA
Normal	357	116
OSA	188	328
Accuracy: 69.26%		
Sensitivity: 75.48%		
Specificity: 63.57%		

LASSO Regression:

- 10-fold cross validation to find the optimal alpha value of 0.07 for the final model
- Features selected in the final model are gender, bmi, SPT, sleepeff, onset, N1, REMper, REMlightoff
- The average MSE across 10 models is 12.360
- The final model achieves an accuracy of 69.26%, sensitivity of 75.48% and specificity of 63.57%.

Conclusion

The results from LASSO model are consistent with Stepwise Regression Results, with a similar set of features selected. Our proposed solutions would enhance the accuracy of OSA detection using sleep tracker devices while addressing the limitations of PSG. Our research also demonstrates the potential of utilizing features extracted from hypnogram in sleep monitors to detect OSA.