

Stress Management Considering Sleeping Habits

Jesus Tafolla, B.S. Digital Systems and Robotics

Tecnológico de Monterrey, Campus Querétaro, México

A01703178@tec.mx

Abstract

Considering today's lifestyle, people just sleep forgetting the benefits sleep provides to the human body. Smart-Yoga Pillow (SaYoPillow) is proposed to help in understanding the relationship between stress and sleep and to fully materialize the idea of "Smart-Sleeping" by proposing an edge device. An edge processor with a model analyzing the physiological changes that occur during sleep along with the sleeping habits is proposed. Based on these changes during sleep, stress prediction for the following day is proposed. The secure transfer of the analyzed stress data along with the average physiological changes to the IoT cloud for storage is implemented.

Introduction

Stress is defined as a state of mental or emotional strain due to unavoidable or demanding circumstances, also known as stressors. Stress can also be defined as a specific strain on the human body caused by various stressors. Stressors cause the human body to release stress hormones. Stressors are categorized as physiological, psychological, absolute, and relative [1].

Humans develop lack of adaptation when exposed to stress due to various stressors for longer periods of time which can have major impacts on relationships, work, health and on the self by causing emotional breakdowns. Having an ability to build a self-monitoring system to tackle these stressors is particularly important. The attempt to control and monitor stress variations due to lack of sleep is the focus of this work. The better the quality of sleep, the lower the stress levels [2].

Framing the problem

The main motivation of SaYoPillow is to actualize the phrase "Smart-Sleeping" which can be termed as a sleep that is complete and which meets the ideal body requirements during sleep. The idea is to propose a smart wearable that requires no user input, a fully automated, response control system which does not compromise the user's convenience. It also aims to educate the user about the benefits and importance of excellent quality sleep and to understand the relationship between sleep and stress.

Solution

Decision trees are powerful and popular tools for classification and prediction. Decision trees represent rules, which can be understood by humans and used in knowledge system such as database. A decision tree is a hierarchical model for supervised learning whereby the local region is identified in a sequence of recursive splits in a smaller number of steps. A decision tree is composed of internal decision nodes decision node and terminal leaves [3]. Decision tree is a classifier in the form of a tree structure which consists of:

- Decision node: specifies a test on a single attribute.
- Leaf node: indicates the value of the target attribute.
- Edge: split of one attribute.
- Path: a disjunction of test to make the final decision.

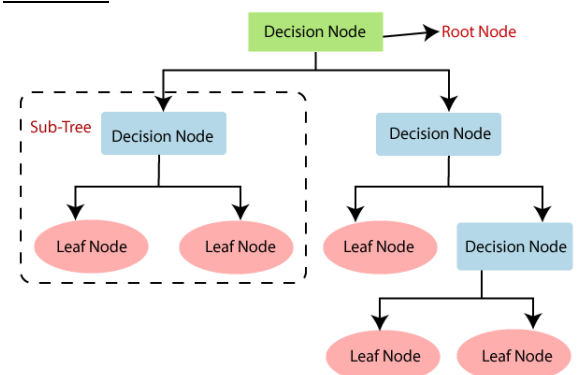


Image 1: Decision Tree diagram (1)

Explaining Data

In this dataset we will see the relationship between the parameters of:

Stress Detection Dataset
Snoring Range
Respiration Rate
Body Temperature
Limb Movement Rate
Blood Oxygen Levels
Eye Movement
Number of Hours of Sleep
Heart Rate
Stress Levels (0-4)

Table 1: Dataset Labels

The relationships between sleep and stress in research made are

1. Sleep can be stated as a state in which the nervous system remains inactive with eyes closed and relaxed muscles.
2. To identify sleep disorders, a sleep study is conducted, known as Polysomnography. This helps in monitoring the sleep stages by performing electroencephalography (EEG) [4].
3. Studies show that the quality of sleep affects the quality of day [5]. Disturbed sleep has been associated with increased instances of sickness, burnout syndrome, persistent psycho-physiological insomnia, weak immune system, PTSD, and higher risk of occupational accidents [6].

Data Preparation

The Decision Tree algorithm manages its own data preparation internally. It does not require pretreatment of the data. Before preparing the data, we need to load the data to be processed with its corresponding labels. After this we need to split the dataset into our training data frame and test data frame, we check indexes to make a random test and train set [7].

Decision Tree is not affected by Automatic Data preparation, but in this case, we need to be sure the data we have is pure and the algorithms needs to

follow the following diagram [8].

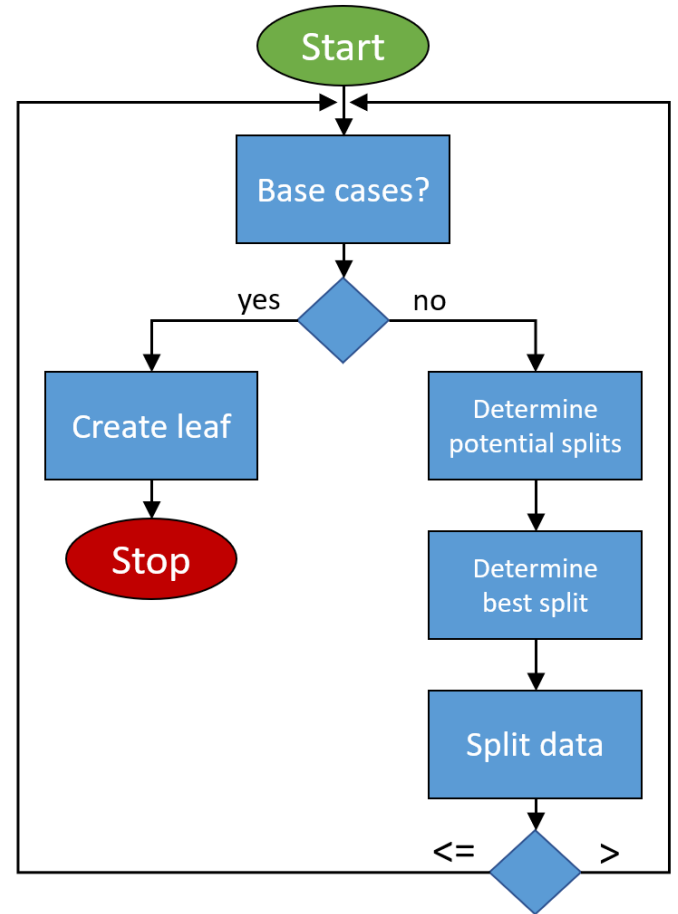


Image 2: Decision Tree diagram (2)

At the beginning of the algorithm, we need to prepare the data and determine the feature type of it in this case we want to separate the values of the sets into categorical and continuous to check which type of data we have.

Entropy and Splits

On our base case, we are checking the purity of the data and if this is correct, we proceed to create a leaf. Otherwise, we proceed with a recursive part which performs the potential splits and determine the best split for it to split it depending on the metrics of the mean square error if it is a regression task, if it's a classification task we calculate the entropy which can be calculate as follows:

$$H = - \sum p_i (\log_2(p_i))$$

Where:

$H \rightarrow$ Entropy

$p_i \rightarrow$ Event probability

The entropy calculates the impurity or uncertainty present in the data. The split of the dataset involves iterating over each row, checking if the attribute value is below or above the split value and assigning it to the left or right group [9]. By a dataset given, we must check every value on each attribute as a candidate split, evaluate the cost of the split, and find the best possible split we could make [10]. Once the split is found, we can use it as a node in our decision tree.

Build a tree recursively until we get all the leaf nodes based on two criteria: Maximum tree depth and Minimum node record.

Predictive Modelling

In the project developed we compare the variations of the trees by themselves. A code from scratch, using a framework and a technique called random forest which is several trees together to have different result and a better accuracy. According to the tests made the Tree code from scratch was a better performance and accuracy than the framework using scikit learn and the random forest code from scratch.

For the data used on the project the following trees were created using the framework to give us an idea of how a tree works.

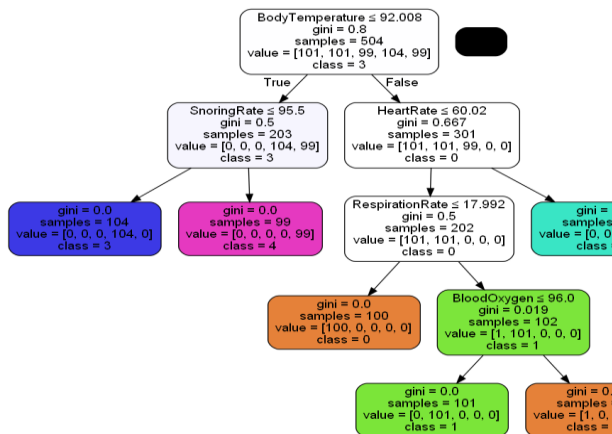


Image 3: Decision Tree example (1)

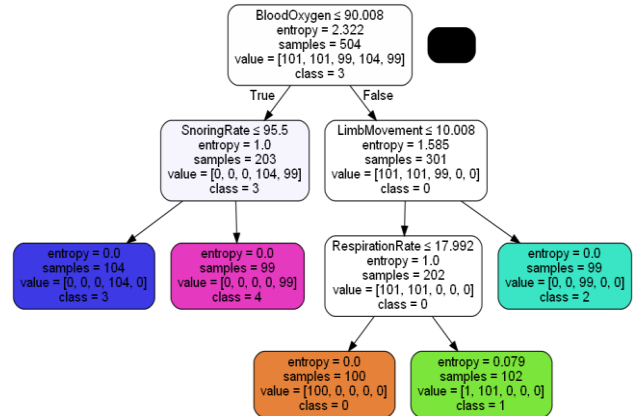


Image 4: Decision Tree example (2)

In this case the tree is interpreted in each node with to detect a pattern and how each data is related with each other. In this case we analyze each node with a true or false depending on some of the features such as blood oxygen, snoring rate, limb movement or any other one to get the data in the correct way.

With these trees performed we can make predictions and start making queries with the model.

Results

In this project the results obtained were good according to the data we have. The main use of trees is the classification of different data depending on different tasks such as classification by itself or regression task. The task performed depends on the data we have and if the data is continuous or categorical. Another point obtained is that in the random forest we split the data in several datasets to have several classifications of it and different results to have a better result than with one tree. It just depends on the data and how is processed internally.

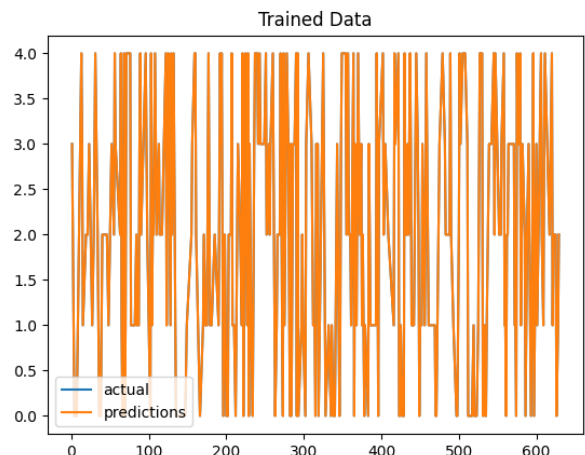


Image 5: Training the dataset.

Accuracy: 0.9984761984761985

	SnoringRate	RespirationRate	BodyTemperature	...	label	Prediction	Correct
142	78.40	21.840	93.840	...	2	2.0	True
287	89.00	24.400	91.200	...	3	3.0	True
72	57.68	19.536	95.536	...	1	1.0	True
356	45.00	16.000	96.000	...	0	0.0	True
553	49.04	17.616	98.424	...	0	0.0	True
..
565	54.72	18.944	94.944	...	1	1.0	True
74	48.68	17.472	98.208	...	0	0.0	True
570	99.84	29.840	89.800	...	4	4.0	True
296	69.12	20.912	92.912	...	2	2.0	True
158	96.32	26.320	85.400	...	4	4.0	True

[315 rows x 11 columns]

Image 6: Results obtained from the Tree program without FW.

Accuracy of the model is 99.20634920634922

	precision	recall	f1-score	support
0	1.00	1.00	1.00	25
1	1.00	1.00	1.00	25
2	1.00	0.96	0.98	27
3	0.96	1.00	0.98	22
4	1.00	1.00	1.00	27
accuracy			0.99	126
macro avg	0.99	0.99	0.99	126
weighted avg	0.99	0.99	0.99	126

Image 7: Results obtained from the Tree FW.

0.9886621315192744

```
{'HeartRate <= 59.96': [{'HeartRate <= 54.88': [0.0, 1.0]},
{'HeartRate <= 64.88': [2.0,
{'HeartRate <= 75.0': [3.0,
4.0]}]}],
{'HeartRate <= 60.0': [{'HeartRate <= 54.88': [0.0, 1.0]},
{'HeartRate <= 64.96': [2.0,
{'HeartRate <= 75.0': [3.0,
4.0]}]}]}],
{'HeartRate <= 60.0': [{'HeartRate <= 54.8': [0.0, 1.0]},
{'HeartRate <= 64.8': [2.0,
{'HeartRate <= 75.0': [3.0,
4.0]}]}]}],
{'HeartRate <= 60.0': [{'HeartRate <= 54.88': [0.0, 1.0]},
{'HeartRate <= 64.96': [2.0,
{'HeartRate <= 74.84': [3.0,
4.0]}]}]}]}
```

Image 8: Trees structures from the Random Forest

Accuracy of the model is: 0.9682539682539683

Predictions

	SnoringRate	RespirationRate	BodyTemperature	...	label	Prediction	Correct
289	64.320	20.432	92.432	...	2	2.0	True
91	86.600	23.760	90.880	...	3	3.0	True
590	94.520	25.872	91.936	...	3	4.0	False
238	83.120	22.832	90.416	...	3	3.0	True
305	62.560	20.256	92.256	...	2	2.0	True
..
574	46.160	16.464	96.696	...	0	0.0	True
437	99.456	29.456	89.320	...	4	4.0	True
224	88.120	22.032	98.016	...	3	3.0	True
389	55.040	19.008	95.008	...	1	1.0	True
276	70.080	21.008	93.008	...	2	2.0	True

Image 9: Random Forest Predictions

Future work

Future exploration on the data preparation step can be included by exploring the change in the depth and by obtaining more data for further predictions and different classification models to have the best accuracy and technique for the problem

Conclusion

The simplicity of the Random Forest and the Trees as well as the diversity, robustness and reliability trump the alternative algorithms available. Classifications in data mining assigns data samples to target classes. Also, decision trees are prone to overfittgwhile random forest is ensemble of decision trees, so it helps predicting data more accurate. A random forest is typically made up with several trees.

References

- [1] N. Schneiderman, G. Ironson, and S. Siegel, "Stress and Health: Psychological, Behavioral, and Biological Determinants." Annu Rev Clin Psychol, vol. 1, p. 607628, 2005.
- [2] K. S. Han, L. Kim, and I. Shim, "Stress and Sleep Disorder," Exp. Neu., vol. 21, no. 4, pp. 141-150, 2012.
- [3] <https://www.javatpoint.com/machine-learning-decision-tree-classification-algorithm>
- [4] S. van Liempt, J. Arends, P. J. M. Cluitmans, H. G. M. Westenberg, R. S. Kahn, and E. Vermetten, "Sympathetic Activity and Hypothalamo Pituitary-Adrenal Axis Activity during Sleep in Post-Traumatic Stress Disorder: A study Assessing Polysomnography with Simultaneous Blood Sampling," Psychoneuroendocrinology, vol. 38, no. 1, pp. 155- 165, 2013.
- [5] E.-J. Kim and J. E. Dimsdale, "The Effect of Psychosocial Stress on Sleep: A Review of Polysomnographic Evidence," Beh. sleep med., 2007.
- [6] T. Akerstedt, "Psychosocial Stress and Impaired Sleep," Scand J Work Environ. Hea., vol. 6, no. 32, pp. 493-501, 2006.
- [7]<https://www.keboola.com/blog/decision-trees-machine-learning>
- [8]https://www.youtube.com/watch?v=LN8Kjn3mf5I&ab_channel=EducationalResearchTechniques
- [9]<https://machinelearningmastery.com/information-gain-and-mutual-information/>
- [10]<https://www.analyticsvidhya.com/blog/2020/06/4-ways-split-decision-tree/>