Stress Detection

By: Ruei-yuan Tu, Emily Xu, Lily Sheng, Hongkun Zeng, Shravani Mhatre, Yash Jain, Leena Abdelrahman, Ethar Hussein

Topic of Interest & Exigence

We are interested in this topic because as college students, we are almost always stressed due to how much work we have. Because of this, we are interested in seeing the different methods we can use to detect our stress level.

Managing stress is an essential skill that can help maintain a good balance between work life and personal life. For example, according to the Office of Disease Prevention and Health Promotion, learning how to manage stress can decrease the likelihood of getting certain diseases such as depression, while having unchecked stress can manifest in issues with one's physical health such as with sleep (Office of Disease Prevention and Health Promotion). Thus providing more exigency for the topic of stress detection in our everyday lives. If we are able to identify more ways to detect stress, we can drastically reduce the chances of these health problems. Additionally, understanding stress will lead to better efficiency in our work, greater productivity, and contentment in our education.

Data Set Details

- dataset from a smart pillow
  - Smart-Yoga Pillow (SaYoPillow)
- it is a processor with model to analyze the effects of sleep
- Usability - 8.82
- License - Data files © Original Authors

Metadata (All are quantitative)

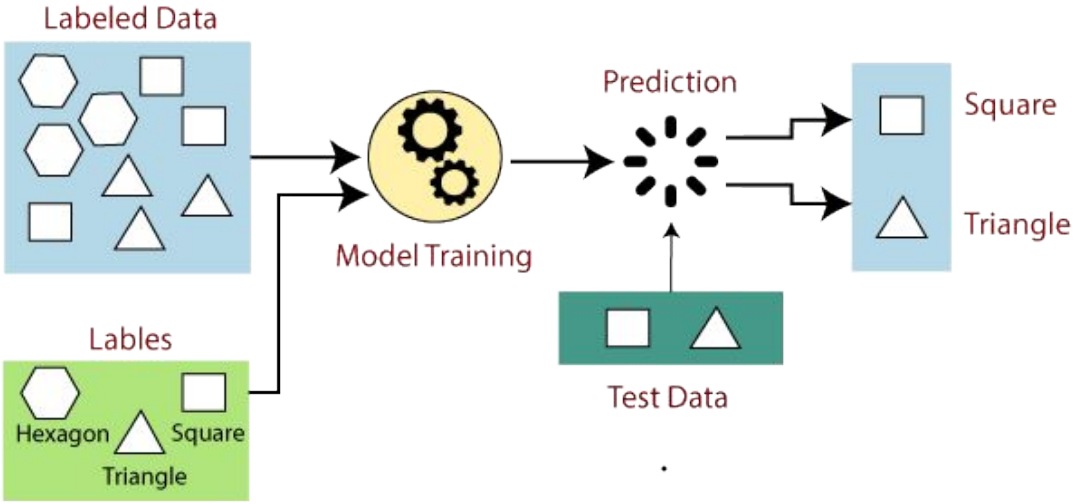
- snoring range of the user
- respiration rate
- body temperature
- limb movement rate
- blood oxygen levels
- eye movement
- number of hours of sleep
- heart rate
- stress levels
  - 0 - low/normal, 1 - medium/low, 2-medium, 3 - medium/high, 4 - high

CSV

	A	B	C	D	E	F	G	H	I
1	sr	rr	t	lm	bo	rem	sr	hr	sl
2	93.8	25.68	91.84	16.6	89.84	99.6	1.84	74.2	3
3	91.64	25.104	91.552	15.88	89.552	98.88	1.552	72.76	3
4	60	20	96	10	95	85	7	60	1
5	85.76	23.536	90.768	13.92	88.768	96.92	0.768	68.84	3
6	48.12	17.248	97.872	6.496	96.248	72.48	8.248	53.12	0
7	56.88	19.376	95.376	9.376	94.064	83.44	6.376	58.44	1

Types of Training

We used supervised learning because we label and pre-process the input data before we run the algorithm. We used a Linear Regression machine Learning model where the model finds the best fit linear line between the independent and dependent variables.



How We Cleaned Our Data

- we checked the CSV for any null values but was not able to find any null values
- computer generated values
- we did not have to replace or delete any values
- updated column names from abbreviations

```
dataset = pd.read_csv('SaYoPillow.csv')


#Renaming each individual column

dataset = dataset.rename(columns={'sr': 'snoring_range',
                        'rr': 'respiration_rate',
                        't': 'temp',
                        'lm': 'limb_movement_rate',
                        'bo': 'blood_oxy',
                        'rem': 'eye_movement',
                        'sr.1': 'num_hours_sleep',
                        'hr': 'heart_rate',
                        'sl': 'stress_level'})
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 630 entries, 0 to 629
Data columns (total 9 columns):
 #   Column              Non-Null Count  Dtype
---  ------              --------------  -----
 0   snoring_range       630 non-null    float64
 1   respiration_rate    630 non-null    float64
 2   temp                630 non-null    float64
 3   limb_movement_rate  630 non-null    float64
 4   blood_oxy           630 non-null    float64
 5   eye_movement        630 non-null    float64
 6   num_hours_sleep     630 non-null    float64
 7   heart_rate          630 non-null    float64
 8   stress_level        630 non-null    int64
dtypes: float64(8), int64(1)
memory usage: 44.4 KB
```

Algorithm

Our dataset had many different columns that are factors that contribute to the last column stress level. This code generates separate graphs to visualize linear regressions for each column vs Stress Level.

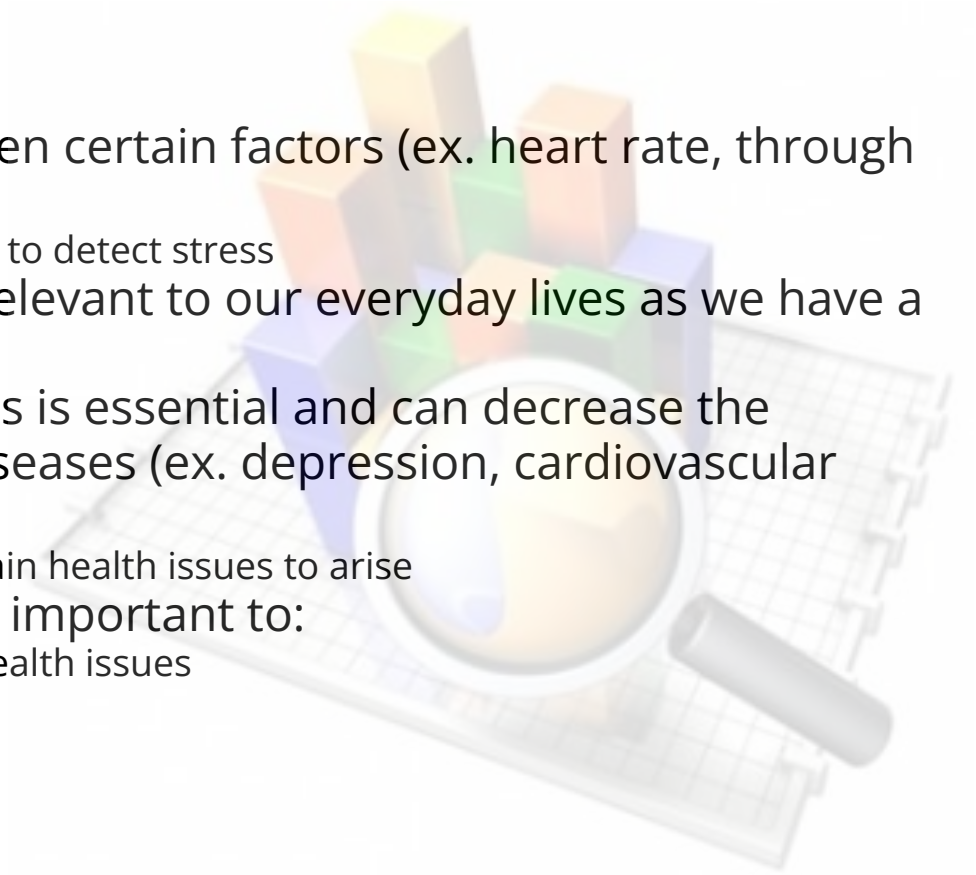
It divides the data into training and testing sets using the train_test_split()

Then the code loops through each independent variable column in the dataset and creates a linear regression using the training data. Then the code plots the regression line using the testing data.

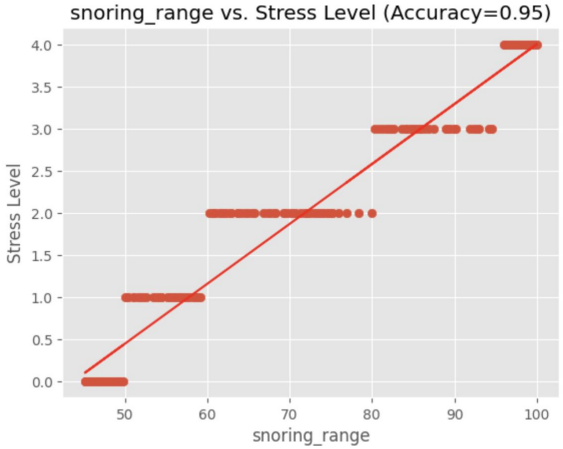
By splitting our code into training and testing data the code allows us the ability to evaluate the model's performance on data it is unfamiliar with. Therefore we can see how well the model can predict outcomes on new data

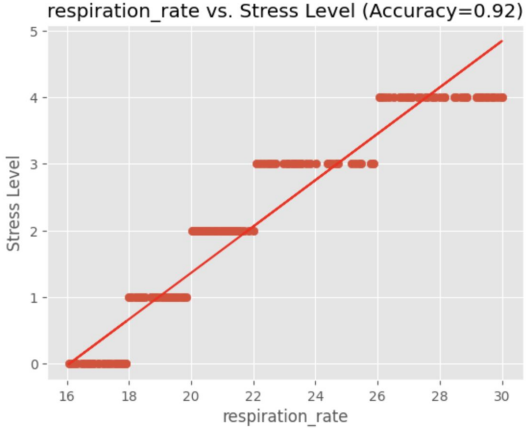
Purpose of Study

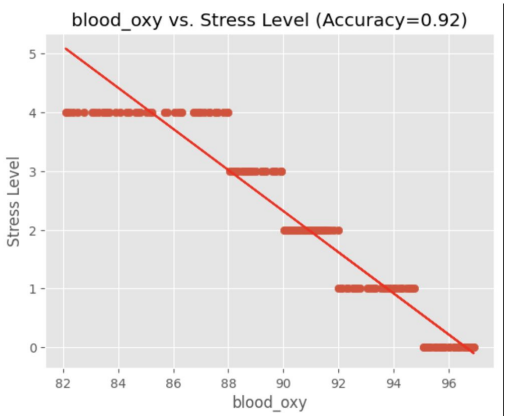
- To find the relationship between certain factors (ex. heart rate, through sleep) vs. the degree of stress
  - i.e. different methods we can use to detect stress
- As college students, stress is relevant to our everyday lives as we have a lot of work to do
- Learning how to manage stress is essential and can decrease the likelihood of getting certain diseases (ex. depression, cardiovascular diseases, etc.)
  - Unchecked stress can cause certain health issues to arise
- Therefore, identifying stress is important to:
  - reduce the probability of these health issues
  - Better efficiency in work
  - Greater productivity

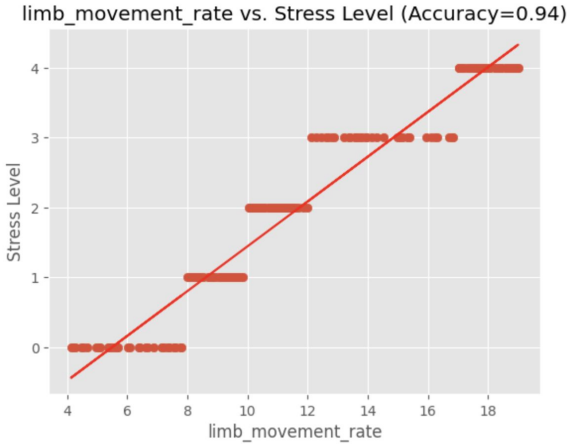


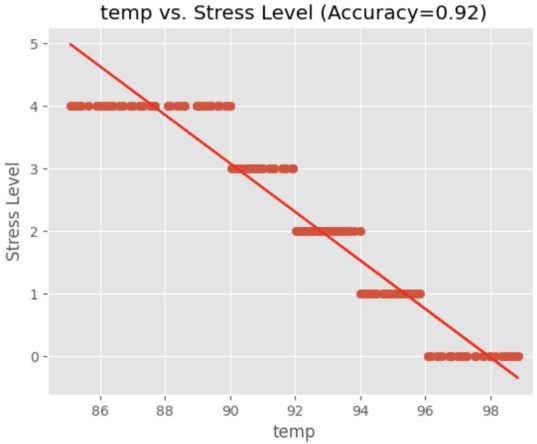
Graphs Cont.

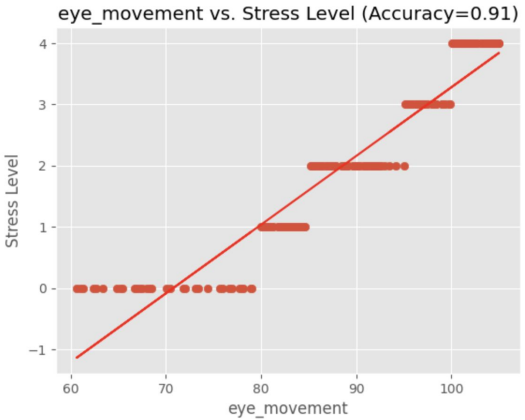

snoring_range vs. Stress Level (Accuracy=0.95)
Stress Level
snoring_range

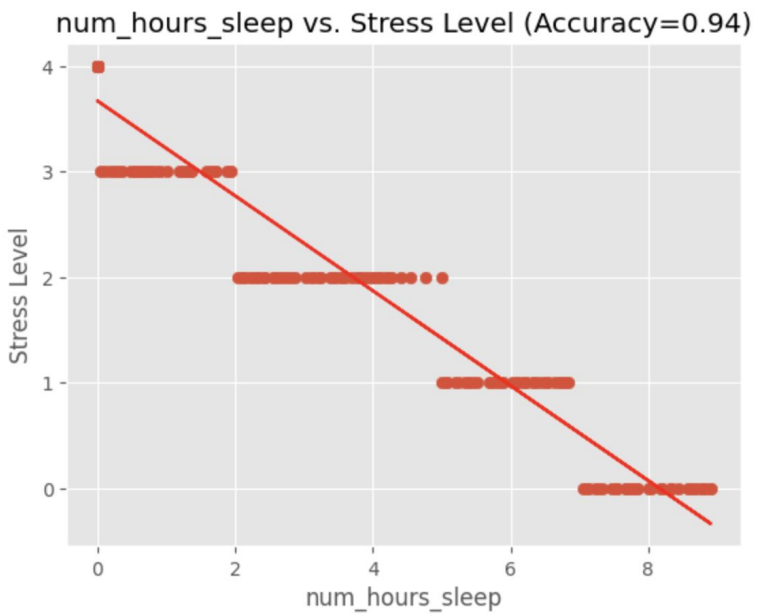

respiration_rate vs. Stress Level (Accuracy=0.92)
Stress Level
respiration_rate

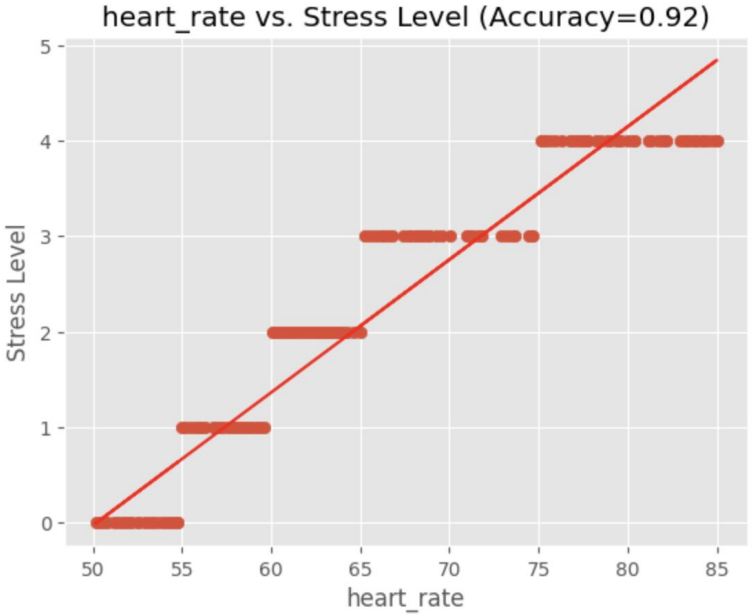

blood_oxy vs. Stress Level (Accuracy=0.92)
Stress Level
blood_oxy


limb_movement_rate vs. Stress Level (Accuracy=0.94)
Stress Level
limb_movement_rate

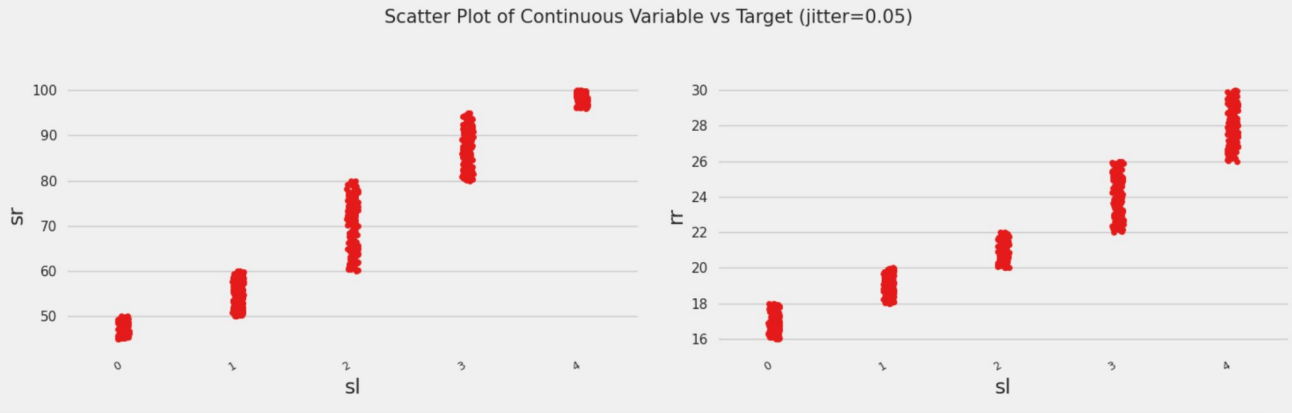

temp vs. Stress Level (Accuracy=0.92)
Stress Level
temp


eye_movement vs. Stress Level (Accuracy=0.91)
Stress Level
eye_movement


num_hours_sleep vs. Stress Level (Accuracy=0.94)
Stress Level
num_hours_sleep
0
1
2
3
4
0
2
4
6
8


heart_rate vs. Stress Level (Accuracy=0.92)
Stress Level
heart_rate
0
1
2
3
4
5
50
55
60
65
70
75
80
85

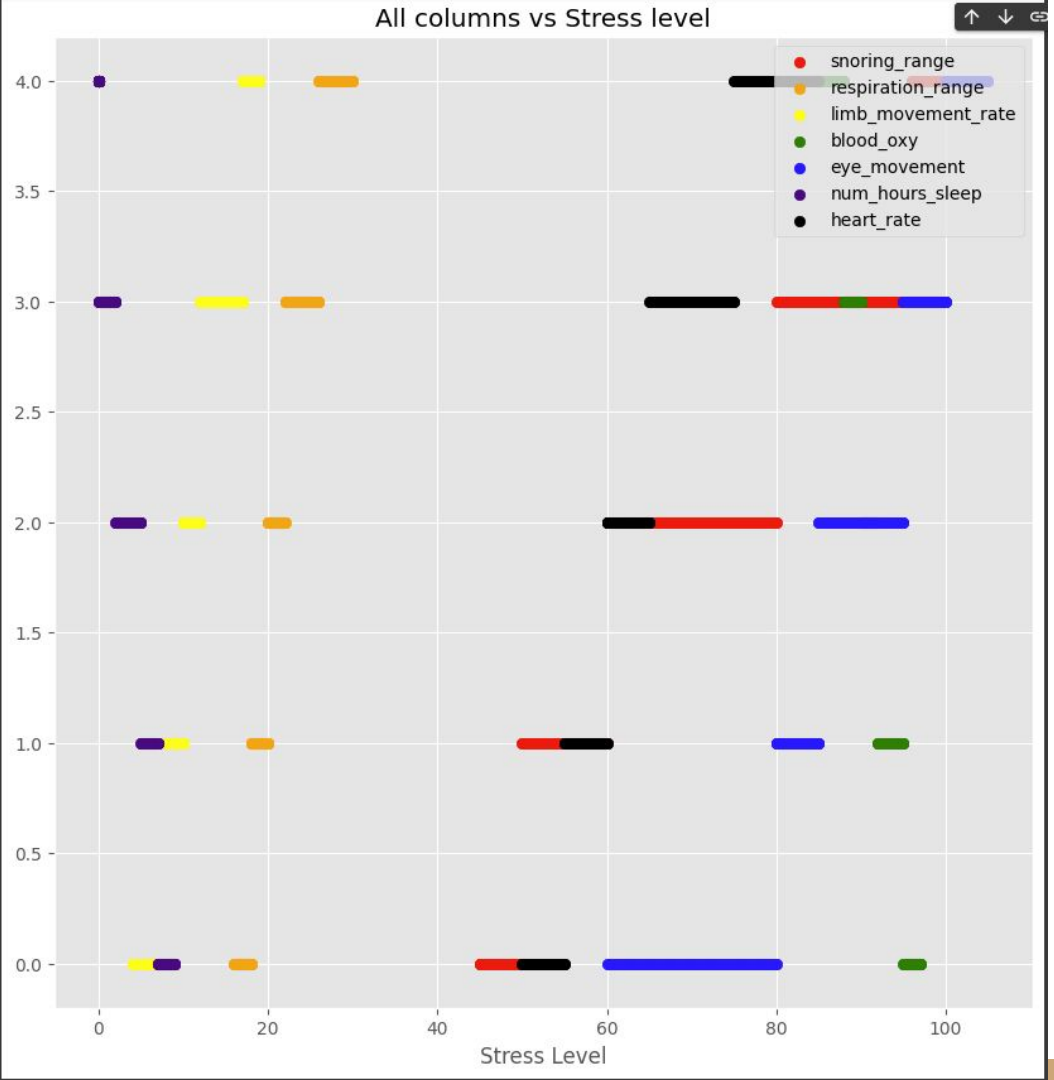
Explanation of Charts (Visualization)



Stress level versus snoring range

Stress level versus respiration rate

We used scatter plots to show a variety of different statistics with people who have different results of stress levels caused by stress, and different results received from heightened stress levels such as hours of sleep, heart rate, snoring, respiration rate, etc.



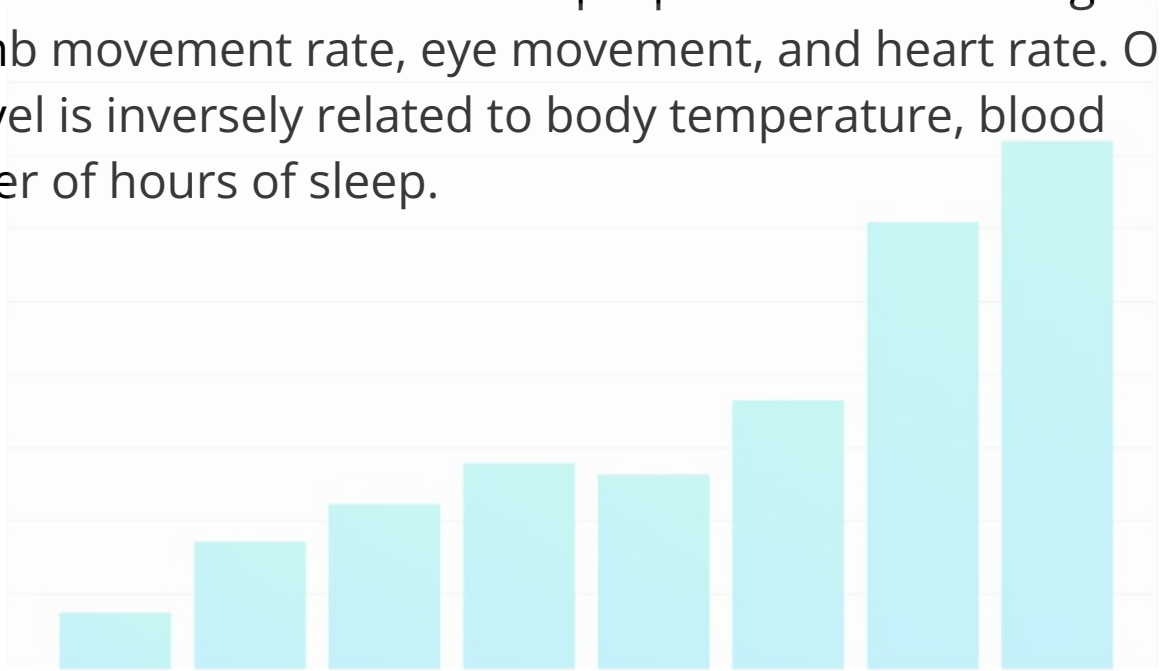
The Scatter Plot provides a visual representation of how specific factors, such as hours of sleep (Yellow) or snoring range (Red), can impact stress levels. This plot enables the reader to gain a better understanding of the relationship between these variables.

What We Learned

The most important thing learned here is the skill of collaboration. We obtained experience in working as a team for a computer science project specialised towards Artificial intelligence. We learned about supervised learning and learned Linear Regression by finding the best fit linear line. We gained further knowledge in cleaning data.

What we can learn from the chart

Based on the charts, we can see that stress level is proportional to snoring level, respiratory rate, limb movement rate, eye movement, and heart rate. On the other hand, stress level is inversely related to body temperature, blood oxygen levels, and number of hours of sleep.



Why this matters

As noted before, we were interested to see how certain physical conditions (snoring range, respiration rate, etc.) connected with stress in order to deduce potential reasonings or indicators of stress in sleep. This is essential because stress can lead to many health complications, as well as unease in our daily lives. Most of these seem to make sense with how we see stress.

Something to note is that, while working with our data, we noticed that the points of the graph were too “perfect.” Unfortunately, we later found out that many of the points are computer generated. We did not expect this, however it does bring up a couple of interesting considerations on the impacts of AI.

Although AI can help us with processing data and analyzing outcomes, too much reliance on it can skew the results. This dataset relied too much on AI. It makes it so that, rather than gaining meaningful information from the data points, people are growing a bit skeptical of its results.

AI should be used to help with extrapolating data and organizing data, but it should not be used to “generate” data. Even though this data seems insightful, it probably does not accurately represent the general population.

Next Steps

As explained before, we can definitely conclude linear relationships between some factors from our graphs (positive relationship between eye movement and stress level, etc). Knowing this, we can perhaps dive deeper into why certain symptoms or actions are happening, and learn to mitigate or encourage certain qualities to decrease stress levels.

We could probably find some data or continue researching online into the potential factors we have found that indicates greater stress levels, and start thinking about what we can do to improve these qualities (what causes greater eye movement and can we minimize them? How can we healthily regulate our heart rate/decrease heart rate?). Then, we can see/fact check if this dataset holds.

This data provided a good starting point for us to continue finding relationships between stress level and other variables. Perhaps we can find another dataset concerning this topic and re-evaluate our findings.

What we would’ve done if we had more time

- A stretch goal for us is to conduct research personas with people at different ages, with different jobs to see the examples of causes of stress. We would do user testing to see what type of category the user is facing stress from. And if different type of routines can lead to more stress than others.

- Since there are so many different ways to identify stress such as heart rate, sleep, muscle movement, food, eye movement, etc there can be a lot of devices with hardware that can be built with haptics and sensors for this type of detection.

- Another way to expand this project is to make a chatbot that asks questions to the users to see any symptoms they do have. And for the chatbot to provide suggestions to manage stress.

Link to Code

https://colab.research.google.com/drive/1Y0xh0ZpC1nm50Kg5LWeb_qSxM_BdT7J?authuser=1

References

1. L. Rachakonda, A. K. Bapatla, S. P. Mohanty, and E. Kougianos, “SaYoPillow: Blockchain-Integrated Privacy-Assured IoMT Framework for Stress Management Considering Sleeping Habits”, IEEE Transactions on Consumer Electronics (TCE), Vol. 67, No. 1, Feb 2021, pp. 20-29.
2. L. Rachakonda, S. P. Mohanty, E. Kougianos, K. Karunakaran, and M. Ganapathiraju, “Smart-Pillow: An IoT based Device for Stress Detection Considering Sleeping Habits”, in Proceedings of the 4th IEEE International Symposium on Smart Electronic Systems (iSES), 2018, pp. 161--166.

Link to Dataset:

https://www.kaggle.com/datasets/laavanya/human-stress-detection-in-and-through-sleep

References Cont.

Han, Kuem Sun, et al. “Stress and Sleep Disorder.” *Experimental Neurobiology*, U.S. National Library of Medicine, Dec. 2012, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3538178/.

“Manage Stress.” *Manage Stress - MyHealthfinder*, Office of Disease Prevention and Health Promotion, 2 Apr. 2023, https://health.gov/myhealthfinder/health-conditions/heart-health/manage-stress#:~:text=Preventing%20and%20managing%20long%2Dterm,which%20tasks%20to%20do%20first