Early prediction of Migraine Episode

The central thrust of our work is the objective and unobtrusive collection of physiological, environmental, and behavioral data-enabled which would allow us to build a personalized and detailed map of factors contributing to an individual's migraine attacks. This map should translate into better migraine management and prevention strategies. The central goal is to leverage the migraine database for personalized migraine trigger identification, risk assessment, and attack prediction. Post feature extraction, we are using artificial recurrent neural networks, more specifically, long short-term memory (LSTM) networks to see if we can train our algorithms on one section of the data library to predict migraine episodes in a future section, both for the same patient and across different patients.

Data Collection:
To attain this goal, we processed our high dimensional data streams collected from Empatica E4 wearable sensor. The physiological sensor data such as heart rate and electrodermal activity (EDA) are recorded at a frequency of 4hertz whereas blood volume pulse is recorded at 64hz and XYZ raw acceleration data at 32hz. Also, the variable, quality of sleep, sleep duration, triggers, and medications taken have one reading per day – into features with clinical meaning.

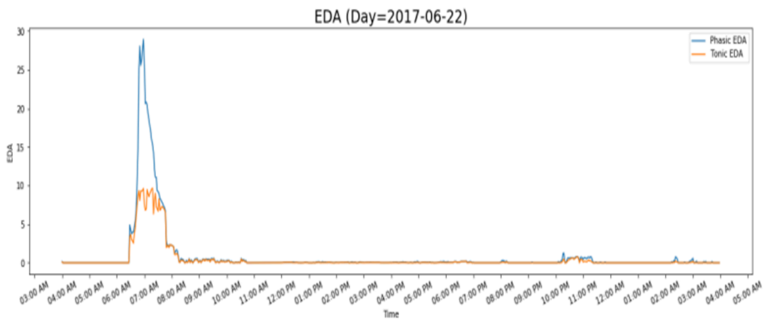
Furthermore, weather data is collected for each patient using Dark Sky API by using the geological location of the user. The data collected is Humidity, pressure, temperature, cloud cover, UV index, etc. The data is collected for 24 hours on all the days. There are no missing values.

Data Preprocessing:

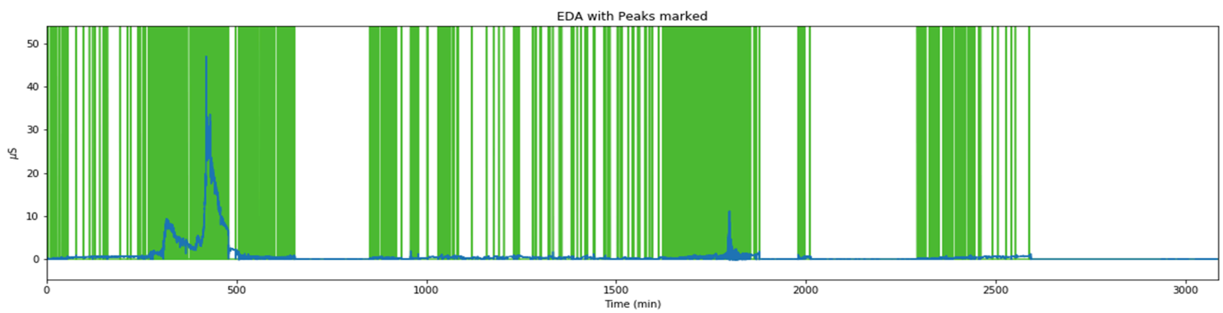
The time-series data were pre-processed by removing the noise and extracting meaningful information out of it. The table below shows the minimum, mean, maximum, and deviation of the data after normalizing the data between 0 and 1.

	Minimum	Maximum	Mean	Deviation
Temp	0.00	1.00	0.82	0.16
uvIndex	0.00	1.00	0.21	0.27
tonic	0.00	1.00	0.05	0.08
HRV	0.00	1.00	0.36	0.13
IBI	0.00	1.00	0.41	0.17
motion	0.00	1.00	0.11	0.07
steps	0.00	1.00	0.05	0.10
phasic	0.00	1.00	0.01	0.06
cloudCover	0.00	1.00	0.51	0.41
pressure	0.00	1.00	0.50	0.18
visibility	0.00	1.00	0.92	0.21
temperature	0.00	1.00	0.48	0.21
humidity	0.00	1.00	0.65	0.27
windBearing	0.00	1.00	0.56	0.25
windGust	0.00	1.00	0.20	0.15
windSpeed	0.00	1.00	0.14	0.11
output	0.00	1.00	0.24	0.43

Electrodermal (EDA) is used to measure galvanic skin conductance which is related to stress. I used the tool to measure Phasic and Tonic EDA To find out the phasic and tonic components of the EDA signal they decomposed it using the cvxEDA tool [2] which uses the convex optimization technique. Skin conductance with long-term change (tonic) is preferred by researchers as they don't want to overestimate long-term changes with short-term changes (phasic) in skin conductance. Below is the resultant figure after the decomposition of the EDA signal by cvxEDA [1].

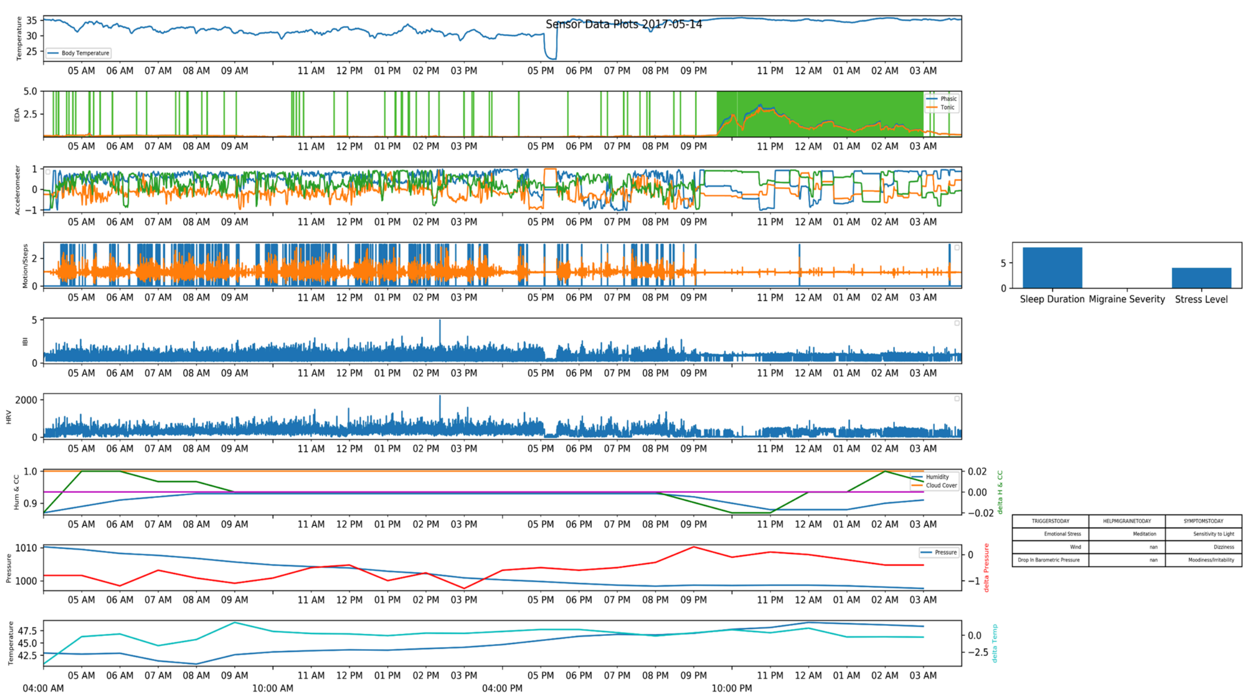


Further I detect the peaks from the signal and marked the peaks with the green region shown below.

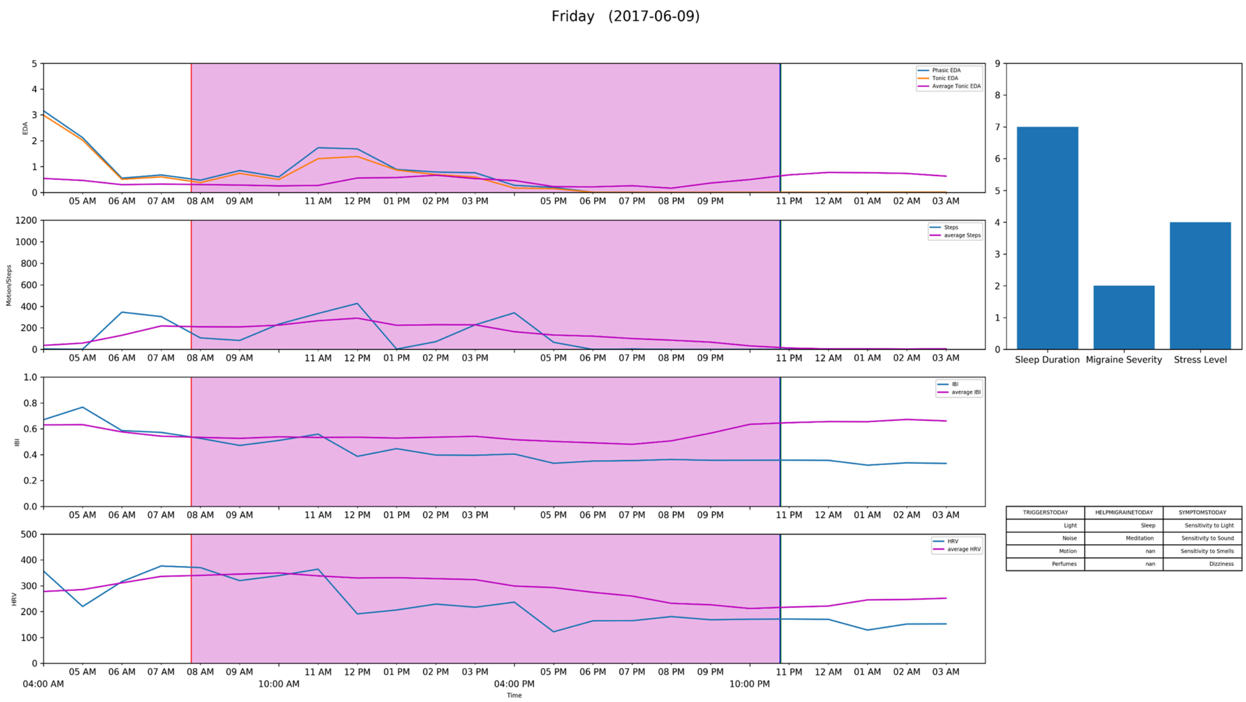


Data Visualization:

Time series data is plotted on a 24-hour time scale as shown on the x-axis and other parameters on the y-axis. On the right bar, graphs are shown that display sleep duration, migraine severity, and stress level. The table shows triggers, medications, etc on the right bottom.



The graph below shows a better meaningful visualization. It displays the average values in each subplot which helps to understand the plots better.



Data Analysis:

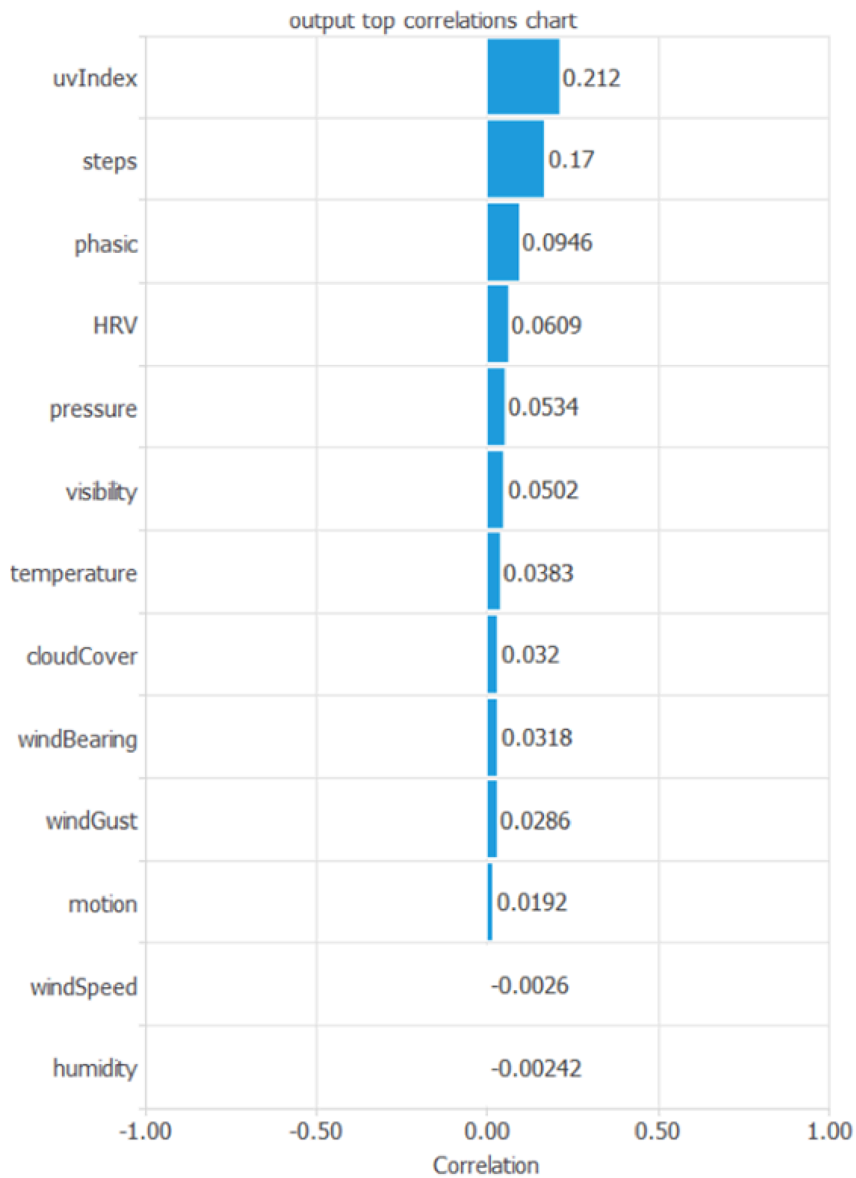
Input Correlation:

The table below shows the input correlation between different variables. The highest correlation is shown between wind gusts and windspeed.

	Temp	uvIndex	tonic	HRV	IBI	motion	steps	phasic	cloudCover	pressure	visibility	temperature	humidity	windBearing	windGust	windSpeed
Temp	1	-0.18	0.29	0.34	0.71	-0.18	-0.0094	0.082	-0.022	0.0065	0.078	0.011	0.086	0.022	-0.074	-0.066
uvIndex		1	-0.0084	0.0068	-0.29	0.046	0.12	0.14	-0.13	0.054	0.096	0.44	-0.52	0.038	0.53	0.47
tonic			1	0.017	0.27	0.019	0.098	0.74	-0.097	0.0062	0.099	0.16	-0.067	0.026	-0.064	-0.076
HRV				1	0.14	-0.13	0.21	0.049	0.056	0.073	-0.013	-0.14	0.23	-0.045	-0.12	-0.11
IBI					1	-0.12	-0.13	0.011	0.063	0.085	0.063	-0.057	0.1	-0.028	-0.15	-0.14
motion						1	0.51	0.14	0.0025	0.081	0.033	-0.034	0.00088	0.05	-0.031	-0.037
steps							1	0.22	-0.011	0.092	-0.0046	-0.075	0.089	-0.033	-0.062	-0.062
phasic								1	-0.074	0.0048	0.048	0.1	-0.062	0.041	0.0021	-0.01
cloudCover									1	-0.24	-0.42	-0.37	0.46	-0.31	0.11	0.11
pressure										1	0.23	-0.079	-0.18	-0.21	-0.31	-0.31
visibility											1	0.33	-0.45	0.39	-0.064	-0.088
temperature												1	-0.58	0.32	0.32	0.26
humidity													1	-0.28	-0.44	-0.4
windBearing														1	0.056	0.028
windGust															1	0.97
windSpeed																1

Output Correlation:

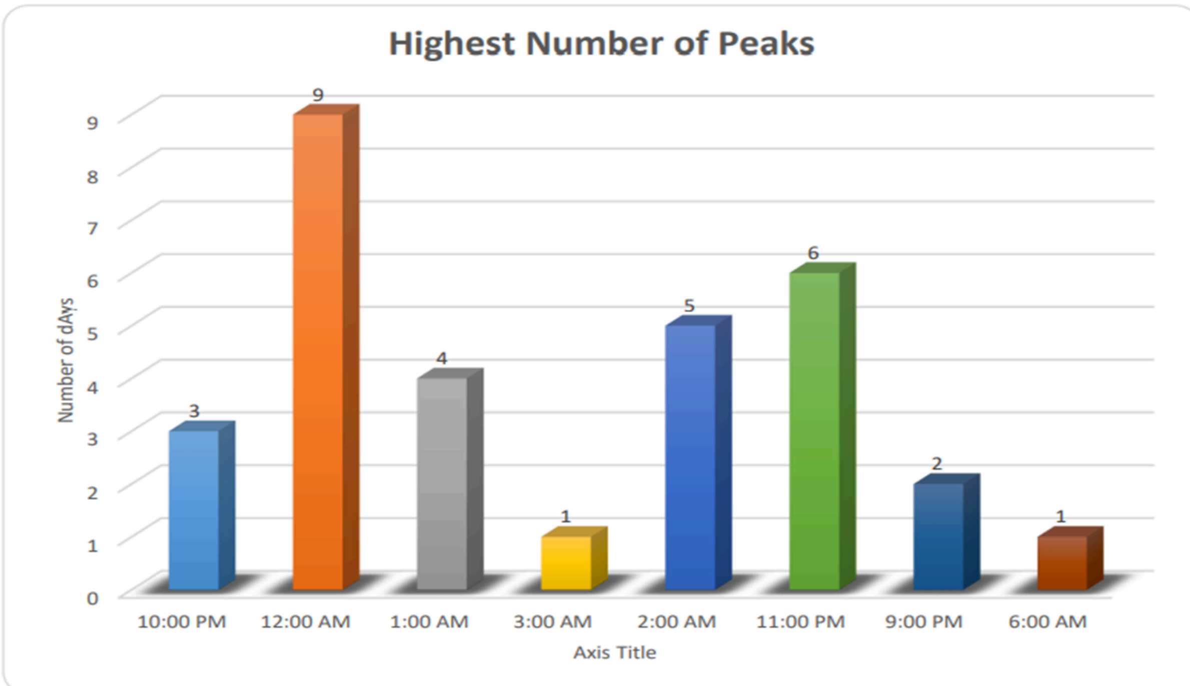
The next chart illustrates the dependency of the target output with the input variables. The following table shows the value of the correlations between all input and target variables. The maximum correlation (0.212283) is the yield between the input variable UV index and the target variable output

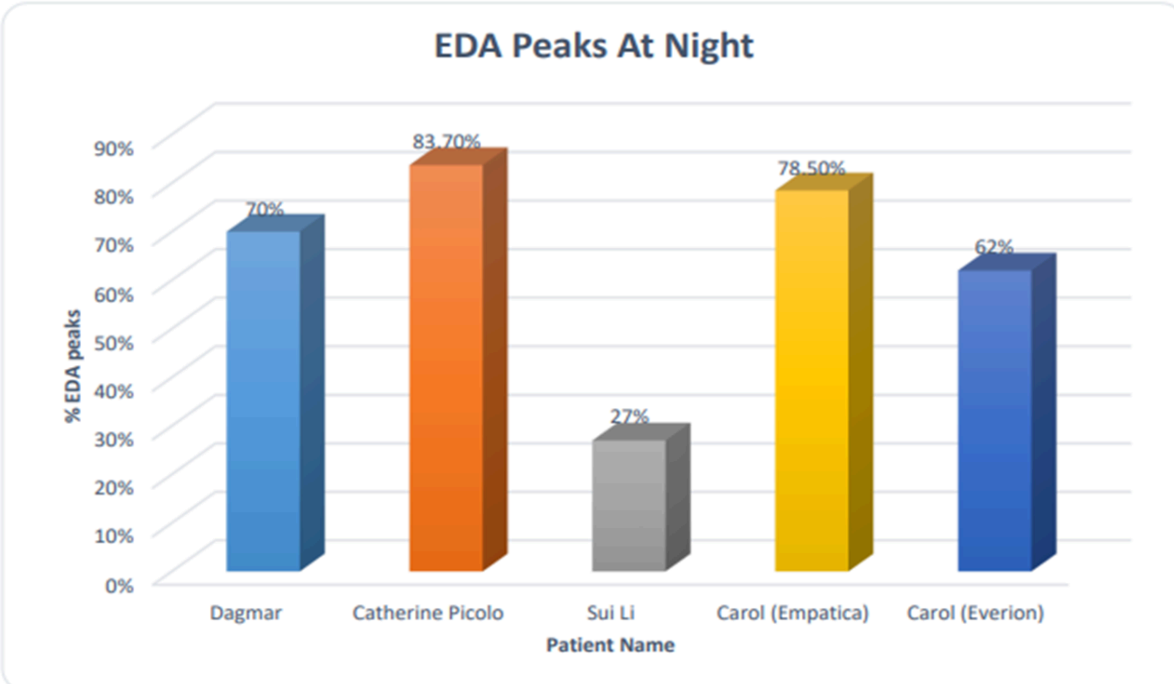


Statistical Analysis:

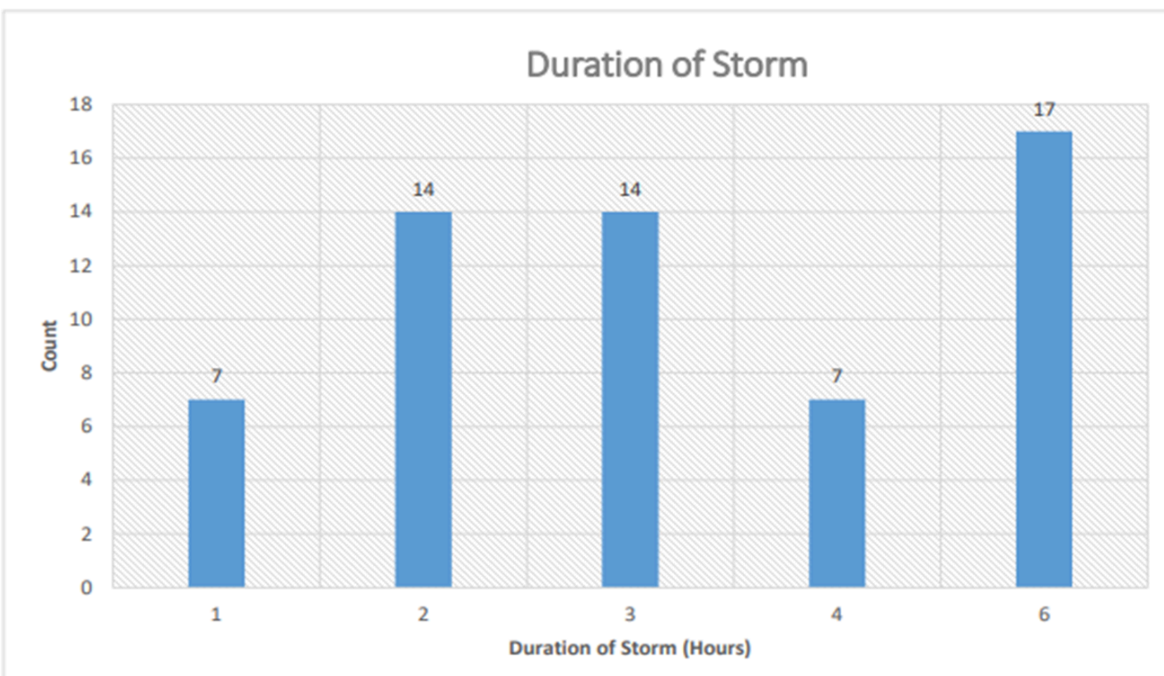
The bar plot shown below displays the highest number of peaks recorded during nighttime as the patient feels the most stress during the stages of sleep at night.

The highest number of peaks was 9 at midnight and the lowest was 1 at 3 am and 6 am.





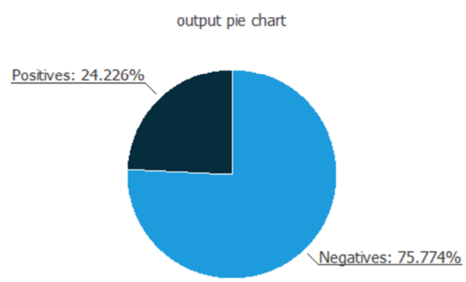
The above bar plot shows the number of peaks recorded for each patient during nighttime. The lowest number of peaks were recorded for patient 3 and the highest peaks were recorded for patient 2.



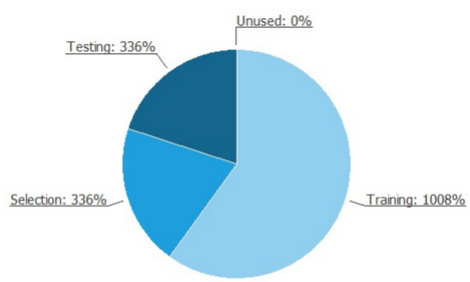
The above bar plot shows the number of storms recorded for each duration of the hour. The highest duration of storm recorded was 14 for 6 consecutive hours and the lowest was recorded 7 times for 1 hour and 4 hours.

Classification:

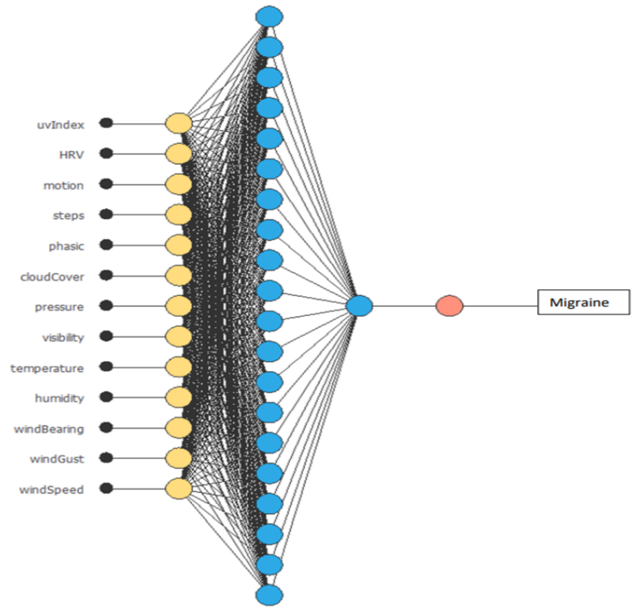
The data shows that the patient has Migraine for 24.226% of the time out of the whole data set as shown in the pie chart below.



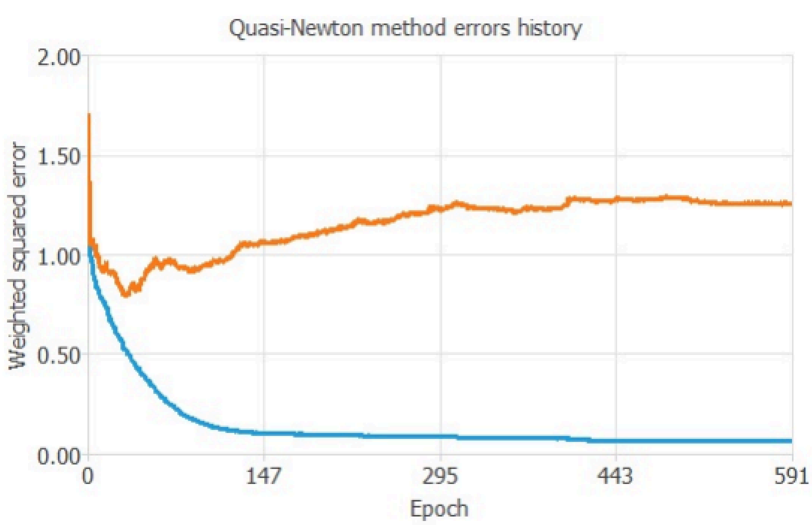
The total number of instances is 1680. The number of training instances is 1008 (60%), the number of selection instances is 336 (20%), the number of testing instances is 336 (20%), and the number of unused instances is 0 (0%).



A graphical representation of the network architecture is depicted next. It contains a scaling layer, a neural network, and a probabilistic layer. The yellow circles represent scaling neurons, the blue circle's perceptron neurons, and the red circle's probabilistic neurons. The number of inputs is 13, and the number of outputs is 1. The complexity, represented by the number of hidden neurons, is 20.



Optimization Algorithm
The quasi-Newton method is used here for training. It is based on Newton's method but does not require the calculation of second derivatives. Instead, the quasi-Newton method computes an approximation of the inverse Hessian at each iteration of the algorithm, by only using gradient information.
Quasi-Newton method errors history
The following plot shows the training and selection errors in each iteration. The blue line represents the training error and the orange line represents the selection error. The inial value of the training error is 1.42045, and the final value a=er 591 epochs is 0.065082. The inial value of the selection error is 1.70254, and the final value a=er 591 epochs is 1.25714.



Confusion table

The number of correctly classified instances is 272 (81%), and the number of misclassified instances is 64 (19%).

	Predicted positive	Predicted negative
Real positive	49 (14.6%)	26 (7.74%)
Real negative	38 (11.3%)	223 (66.4%)

Binary Classification Results

The table below shows the classification accuracy of 80.95% with an error rate of 19.0476%.

	Description	Value
Classification accuracy	Ratio of instances correctly classified	0.809524
Error rate	Ratio of instances misclassified	0.190476
Sensitivity	Portion of real positive which are predicted positive	0.653333
Specificity	Portion of real negative predicted negative	0.854406
Precision	Portion of predicted positive which are real positive	0.563218
Positive likelihood	Likelihood that a predicted positive is a real positive	4.48737
Negative likelihood	Likelihood that a predicted negative is a real negative	2.46463
F1 score	Harmonic mean of precision and sensitivity	0.604938
False positive rate	Portion of real negative which are predicted positive	0.145594
False discovery rate	Portion of predicted positives which are real negatives	0.436782
False negative rate	Portion of real positive which are predicted negatives	0.346667
Negative predictive value	Portion of predicted negative which are real negative	0.895582
Matthews correlation	Correlation between the targets and the outputs. It takes a value between -1 and +1	0.48265
Youdens index	Probability that the prediction method will make a correct decision as opposed to guessing	0.507739
Markedness	Probability of predicting the classifier labels from the real classes.	0.417625

Clustering Algorithm

I tried implementing the hierarchical clustering algorithm with a time of the day as shown on the x-axis and y-axis of the whole day. Interesting information can be extracted after comparing the graphs with patients having higher stress and patients with low stress. I will try to do further analysis by plotting the heat map of patients with days having migraine and days with no migraine.