Migraine Prediction Engine

In-Lab Study

Abstract

Migraines affect more than a billion people globally. Two-thirds of migraineurs are women at the peak productive years of their lives. The average migraineur has 4.2 migraines per month, and migraines render the patient unable to function for anywhere between 4 to 72 hours. Treatments for preventing and treating migraines are not very effective at present. Migraineurs report that the unpredictability that comes from not knowing when the next episode will happen is as distressing as the headache and associated symptoms that come with an attack.

Migraineurs have long tried to predict migraine by keeping track of triggers and helpers via paper diaries, but data shows that digital diaries are better at helping patients predict migraines. We combine digital biomarkers from wearables, diary data from a smartphone app, and weather data to build a migraine model for participants. We analyzed 1,252 hours of wearable data, 733 days of diary data, and approximately two years of weather data. We used LSTM-based models to predict the next migraine 60% to 73% of the time with two to eight hours of lead time. Different combinations of factors from our multi-stream data are important for different patients implying that each patient has an individual migraine map that does not generalize to other patients.

Keywords: migraine, migraine prediction, digital biomarkers, LSTM, digital diary, weather, time series data.

Introduction

Migraine is a primary neurological episodic disorder which affects more than a billion people around the globe [1]. It is the second leading cause of years lived with disability (YLDs) worldwide [2] and was responsible for 41.1 million YLDs in 2019, accounting for 4.8% of the total YLDs [3].

One out of six adults in the US (15.3%) has migraine or severe headache, with prevalence rates of 20.7% among women and 9.7% among men [4]. Women account for two-thirds of all migraine patients and prevalence is highest from the ages 18 to 55, their peak productive years [5]. The economic burden of migraine in the US exceeds $28B annually [6][7][8].

A migraine episode lasts anywhere between 4 to 72 hours [9], and 90% of migraineurs are unable to function normally during the episode [10]. Migraine is characterized by intense, throbbing pain typically located on one side of the head. Associated symptoms include nausea, vomiting, and sensitivity to light and sound often, which accompany the pain [11]. Patients report the unpredictability of migraine attacks to be as distressing as the pain and disability associated with an attack. The lack of control over keeping appointments causes anxiety that in turn augments the risk for a migraine attack, perpetuating a vicious cycle [12].

Patients use both prophylactic (preventative) and abortive medications to treat migraine. Prophylactic medications reduce the frequency, severity, and duration of migraine attacks [13]. Roughly 38% of individuals with episodic migraines could find relief through preventive therapy, but less than 13% actually opt for prophylactic medications [14]. Beta-blockers, the most common of the prophylactic medications, have been reported to decrease the frequency of migraines in 46% of the users. However, Beta-blockers are unsuitable for asthmatic patients and can cause lethargy, dizziness, sleep issues, and gastrointestinal problems as possible side effects. [15].

Abortive medications aim to alleviate the symptoms of a migraine episode [16]. They include triptans, nonsteroidal anti-inflammatory drugs (NSAIDs), and antiemetics [17].

NSAIDs are less effective than triptans. A study showed that triptans were effective for 60% of non-responders to NSAIDs [18]. However, Triptans need to be taken soon after the onset of symptoms for optimal effectiveness. Standard dose triptans provided sustained headache relief at 24 hours in 29 to 50% of patients and sustained freedom from pain in 18 to 33% of patients [19].

Both NSAIDs and Triptans usage needs to be limited to avoid overuse headaches [18][19]. Furthermore, these treatments are not effective in addressing other symptoms, such as nausea, photophobia (sensitivity to light) and fatigue, which are significant aspects of the suffering related to migraine.

Therefore, both prophylactics and abortives are inadequate in giving relief from migraine. Prophylactic medications are under-utilized, and abortive treatments do not provide adequate relief and have undesirable side effects.

A complementary approach to managing migraine focuses on identifying potential triggers, such as stress, lack of sleep, or certain foods [20]. Migraine triggers tend to accumulate in the form of a stack, which leads to the onset of an attack [21]. Gago-Veiga et al. found that during a 2-month period, 67.6% of the participating migraineurs were able to anticipate at least one migraine attack, but only 35.3% demonstrated proficiency as good predictors by accurately forecasting more than 50% of their attacks [22].

Historically, migraineurs have kept migraine diaries to track their episodes and identify potential triggers and helpers in an effort to prevent and/or predict their next attack [23]. Personal trigger identification has long relied on patients identifying and associating environmental, physiological, psychological, and behavioral factors with a migraine attack, retrospectively logging these in paper-based migraine diaries, and then attempting “N=1 experiments” of trigger avoidance [24]. This approach is prone to recall bias and requires significant cooperation from the patient. [25][26][27].

Cell phone based electronic migraine diaries do not overcome these limitations. Curelator, for example, asks customers to log their migraine triggers and helpers for 90 days and then provides them with association maps [28][29]. Thus, patients have to know their triggers and helpers a priori, which limits the usefulness of the approach.

Single data stream models that utilize patient-reported stress as the only input can predict migraine attacks in individual patients with 65% accuracy [30]. We know from our own patient surveys that less than 25% of migraineurs would use a model with 65% predictive accuracy. Weather variables can also exacerbate migraine symptoms for a subset of migraineurs. However, previous pooled analysis of all patients does not reveal any association between migraine attacks and weather variables [31].
However, tracking weather variables' trends continuously for an individual migraineur using their location data from their smartphone may capture the role of weather in their migraine frequency.

The advent of wearable digital devices allows individuals to continuously monitor physical activity, heart rate, sleep patterns, and Electrodermal activity (EDA) [32]. About 30% of the adults in the US use healthcare wearables, of which 47.3% use them on a daily basis. Moreover, 82.38% of the regular users showed a willingness to share their data with healthcare providers [32].
EDA, temperature, Heart rate and accelerometer from wrist wearables can predict stress [30]. A study on triggers revealed that stress served as a trigger for 64% of its participants [33]. Ordas et al. reported a higher body temperature during headache days and a positive correlation between body temperature and pain intensity [34]. Heart Rate Variability (HRV) calculated from Inter-Beat-Interval (IBI) collected from an individual's wearable, is shown to decrease in people with episodic migraine, particularly during the migraine ictal period [35].
Given the preceding evidence, a personal data library that has continuous streams of digital biomarkers from a wearable device, a digital diary from a smartphone and weather variables for an individual patient could predict their next episode with enough precision to allow the patient to undertake preventive measures and/or cancel their commitments if a migraine episode were inevitable. Our survey of 106 migraine patients shows that migraineurs prefer a predictive tool that over-predicts rather than misses migraine episodes, and has at least 85% accuracy with a lead time of 24 hours. Our study attempts to build this data library for four of our patients.

Research questions:

- How to leverage machine learning techniques to predict a migraine episode in advance using biomarker data (from wearables) and patient diary data
- How do different data streams contribute to the predictive model? Essentially what factors, and changes within them point towards an impending episode.
- What are common triggers and helpers in our pool?

Research Objective:

- Apply machine learning techniques over longitudinal wearable and diary data to predict migraine 24 hours in advance
- Feature selection - find what features are important
- Reach an educated guess or a rough plan as to how this study can be carried out at a larger scope

Methodology

We recruited patients with a diagnosis of migraine from a primary care practice in Massachusetts, USA. Participants received information regarding the study and provided their consent. The Institutional Review Board (IRB) at MIT granted approval for the study protocol, ensuring adherence to ethical guidelines.
We used Everion and Empatica wearables to collect physiological data from our participants. We encouraged the participants to wear the device continuously throughout the study period and have 1,252 hours of wearable data from our study participants. The wearables captured streams for electrodermal activity (EDA), heart rate (HR), blood volume pulse (BVP), the time between an individual's heartbeat (IBI), and accelerometer data. We collected weather data using the Weatherbit API, which spanned the entire duration of the study.

We designed a digital diary app with feedback from our patients, incorporating state-of-the-art design principles into its interfaces and functions. Participants maintained a digital diary, recording the onset and end of migraine events, their severity as well as, the presence of lurking migraines, associated symptoms, helpers, triggers, stress levels, sleep time, and quality of sleep on a scale of 1 to 5.

We included a total of five participants, comprising four females and one male. Among the female participants, three were menstruating, while one was menopausal. Their demographics are outlined in the table below.

Data collection spanned a total of 1071 days. Throughout this period, we collected a total of 733 participant entries covering 656 days. We recorded 156 instances of lurking migraine and 201 instances of migraine. Common triggers that emerged included emotional stress, poor sleep, and temperature changes.

	Participant				
	P_1 (carol)	P_2 (dagmar)	P_3 (katherine)	P_4 (krisztina)	P_5 (suili)
Sex	Female	Female	Female	Female	Male
Menstruating	No	Yes	Yes	Yes	No
Age	56	46	36	36	34
Days covered by sensor data	251	241	79	40	29
Days covered by diary data	259	241	82	43	31
Lurking Migraine	82	24	18	0	2

Had Migraine		87	99	9	6	0
Common triggers		Neck Pain, Emotional Stress, Cold Temperature, Drop in Barometric Pressure, Wind	Lack of Sleep, Citrus/Bananas, Emotional Stress, Chocolate, Dehydration	Emotional Stress, cough, period, Lack of Sleep, Hunger	N/A	Lack of Sleep, Alcohol, Emotional Stress
Common helpers		Caffeine, Massage, Heat On Neck, Sunglasses, Sleep	Hydration, Medications, Migrelief Preventative, Avoid Wheat, Sleep	Exercise, Medications, Sleep	Drinking Water, Sleep	Exercise, Sleep, Medications
Average stress levels	With Migraine	2.172	1.277	2.889	1.027	0.0
	Without Migraine	2.448	1.855	2.013	1.0	1.821
Average sleep quality	With Migraine	3.207	2.109	3.222	1.784	0.0
	Without Migraine	3.524	2.87	3.32	1.583	2.231
Migraine episodes	Before 12 PM	7	35	1	3	0
	After 12 PM	80	64	8	3	0

We applied supervised and unsupervised learning models to the three rich data streams for the patients in our pool: 1) diary data, 2) wearable sensor data, and 3) weather data.

We kept the datasets for each patient separate because our literature review suggested that triggers, helpers, and symptoms related to migraine were individual to the patient, and previous attempts at predicting migraine episodes by pooling data had not been successful (). For example, excessive screen time may trigger for one migraineur but not for another. Similarly different migraineurs have a different threshold for tolerating lack of sleep. To convert this qualitative data in a form that is ready for model training, we used one-hot vectors for each unique trigger, symptom and helper. For missing data, such as in sleep duration, sleep quality and stress level, we used a 4 period moving average with 2 periods on each side to fill in the missing values.

Supervised learning methods like K-Nearest Neighbors (KNN), Neural Networks (Multilayer Perceptron or MLP) and Random Forests have strong predictive capabilities and are applied on each patient's data. For each of these models, we primarily used Python's scikit-learn machine learning library and experimented with PyTorch and Tensorflow for neural networks to cross-validate the results across libraries. We built the KNN architecture for 3, 5, 7, and 9 neighbors and used euclidean distance for distance calculation. The neural network architecture had 4 hiddens layers with 10, 8, 4 and 2 neurons respectively and 2500 maximum iterations at a random state of 1. The random forest classifier had 100 trees with a max depth of five. A five-fold cross validation was performed on each model for robustness in results. To better account for the time-series nature of the data, we trained a Long-Short Term Memory network (LSTM) with class weights of 2:1 for migraine against no migraine to correct for the class imbalance since the migraine class was a minority class. The architecture for LSTM is described below with sensor data.

Our literature review suggested that weather variables like temperature, humidity, atmospheric pressure, and cloud cover can play a vital role in triggering migraines. We used Weatherbit API to get hourly and daily weather data from June 2016 till November 2019 and used it in isolation as well as in conjunction with the diary data to predict migraines by training the LSTM model described above, but with a slightly modified architecture.

Voting Algorithm:

We transformed our LSTM model's minute-by-minute predictions into daily predictions by using predictions as a voting mechanism. We calculated the average minutes a patient was in a migraine state on the days that a patient self-reported a migraine. We treated that average as the minimum threshold for minutes required for a day to be classified as a migraine day in the test data.

Harmonic Mean of the two predictions

The probability for a day having a migraine is calculated as the ratio of the number of minutes predicted to belong to the migraine class and the average number of minutes a person reported experiencing a migraine in the training data day.

This gave us the probabilities for each day having a migraine as shown below in table 1.3 under the column of 'sensor probability'.

We had two sets of probabilities for a day having a migraine: one derived from the sensor data using the approach above and another from the LSTM's prediction using diary data. We then calculated the harmonic mean of the two probabilities to reach a final probability for a day having a migraine for the patient. We chose the harmonic mean as it punishes a decrease in value for either input data, which reflects greater strictness we imposed on predicting a day to have a migraine. This results in the probabilities shown in the following table:

Date	Sensor Probability	Diary Probability	True Label	Harmonic Mean
2017-06-21	0.633663	0.487105	1	0.523424
2017-06-22	0.316832	0.461213	0	0.405743
2017-06-23	1.000000	0.485736	0	0.574346
2017-06-24	1.000000	0.465585	1	0.554482
2017-06-25	1.000000	0.494134	0	0.582540
2017-06-26	0.560548	0.536648	1	0.543601
2017-06-27	0.656512	0.515092	1	0.550678
2017-06-28	0.000000	0.482489	1	0.000000
2017-06-29	0.801219	0.000000	1	0.000000
2017-06-30	0.513328	0.475376	1	0.486159
2017-07-01	0.863671	0.000000	0	0.000000
2017-07-02	0.373191	0.495004	1	0.450855
2017-07-03	0.258949	0.517487	0	0.398213
2017-07-04	1.000000	0.522507	1	0.609870
2017-07-05	0.773800	0.530312	0	0.585591

Results:

We find that if we set a prediction threshold for migraine probability to be a minimum of 0.5, our approach predicts 8 out of 9 migraine attacks either through the sensor, diary data or a combination of both through their harmonic mean.
Diary Data:

After hyperparameter tuning, KNN predicted a migraine two days before it occurs with an accuracy of 65% for Patient 1, where the F1 score for ‘no migraine’ is 71% and for ‘migraine’ is 55% and the precision is 62% when predicting a migraine. The neural network performed with a comparable accuracy of 64%, with F1 score for ‘no migraine’ being 68% and for ‘migraine’, 57%. The random forest was the best out of all the models in predicting a migraine, with an F1 score of 60% for migraine class. The results for models trained and tested on the digital diary data are mentioned in the table 1.1 below:

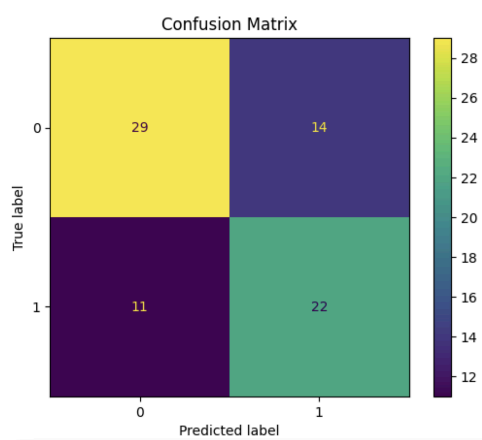

Figure: the confusion matrix for LSTM on carol → taking 3 day's context and predicting 1 day in advanced

Patient	Model	Accuracy (%)	F1 Score for Migraine (%)	F1 Score for No Migraine (%)	ROC AUC	PR AUC
Patient 1	KNN	65	55	71	0.51	0.47
	Neural Network	64	57	68	0.60	0.55
	Random Forest	58	60	56	0.56	0.54
	LSTM	67	64	70	0.67	0.55
Patient 2	KNN	65	38	76	0.58	0.43
	Neural Network	66	39	77	0.56	0.40
	Random Forest	59	58	69	0.58	0.36
	LSTM	55	48	60	0.43	0.38
Patient 3	KNN	95	0	98	0.55	0.1
	Neural Network	95	0	98	0.50	0.09
	Random Forest	91	0	95	0.45	0.08
	LSTM	4	8	0	0.13	0.16
Patient 4	KNN	80	0	89	0.42	0.19
	Neural Network	80	0	89	0.50	0.19
	Random Forest	80	0	89	0.51	0.46
	LSTM	35	42	27	0.33	0.19

Weather Data:

The LSTM model did not gain any significant contextual information from the weather features and was less successful in predicting migraines, with F1 score of 30%. Despite using class weights to give more importance to the minority class of migraines, this model was not able to catch any patterns in migraines by looking at a context of 3-5 days and predicting migraine 1 day in advance.

Discussion:

The objective of our study was to develop a predictive model capable of forecasting migraine episodes with a lead time of 24 hours, with an 85% accuracy rate. While we did not achieve the desired 24-hour lead time, our model correctly predicted 8 out of 9 migraine episodes in our test set with a lead time of 2 hours combined with a 2-hour context, yielding an accuracy rate of 88%. It is important to acknowledge that our study looked at only five patients and needs to be replicated for a larger patient cohort. However, we believe that our approach of tracking multiple data streams over time is sound and once applied to a larger participant pool, will enhance both the lead time and the predictability of the model.

Our literature review of migraine prediction data suggested that sensor, digital diary, and weather datasets were good candidates to be part of a robust predictive model. While the sensor data, and history of triggers and helpers from the digital diaries were helpful in training the LSTM, we found that weather data added no discernible predictive power to the model. One explanation for this can be that our study participants were not as sensitive to variations in weather over the course of a single season. One of our participants had xyz % more headaches in the winter months in New England than in the summer months. However migraine prediction with 1-10 days consideration of past weather data was not helpful in predicting migraine episodes for any of our patients.

Patients report correlation between fatigue brought on from vigorous physical activity and migraine episodes[36]The steps data from the accelerometer sensor did not yield a significant contribution to our model either.

An interesting observation was that the diary logs were adequate in the prediction of migraine episodes during periods where there was no accompanying sensor data. This means that migraineurs who only maintain digital diary logs can benefit from predictive models.

Our survey of 106 migraineurs had told us that patients will use a model with greater than 85% accuracy and that they preferred a model that overpredicts but does not miss any migraines, i.e., gives more false positives than false negatives. Our model only missed one migraine and predicted xyz…

This is a proof-of-concept study that provides the framework for a larger pool of participants who wear a wearable device for sensor data and log their migraine(s) via a digital diary for a year. Our data suggests that each patient has a unique migraine map in which even if the triggers and helpers are the same, their effect size is different. With a larger participant pool, migraineurs can potentially be categorized into several archetypes. If that proves to be the case, the training process for an individual's migraine predictive model would shorten based on the early recognition of their particular archetype.