Continuous Prediction of Adherence to Sleep Apnea Therapy Management Using Clustering

Mourya Karan Reddy Baddam[1], Matheus Araújo[1], and Jaideep Srivastava[1]

University of Minnesota, Minneapolis MN 55455, USA

Abstract. Obstructive Sleep Apnea (OSA) is a disorder in which breathing repeatedly stops and starts due to recurrent episodes of partial and complete airway obstructions during sleep.One of the common treatments for moderate and severe cases of OSA includes the use of Continuous Positive Airway Pressure (CPAP) devices that keep the airways open.But about 60% of the patients using CPAP devices abandon the therapy in middle. Long-term adherence to therapy can be improved by timely and appropriate intervention. At present, the methods followed by the sleep clinics to identify patients who could abandon the treatment is not sufficiently accurate and timely. Recently there have been attempts at identifying non-adherent patients, both in a specific period of their therapy and also in a generalized manner. We propose a novel generalized approach - Therapy Adherence Prediction using Clustering for CPAP therapy (TAPC-CPAP) - which uses clustering techniques to group the patients into clusters based on the usage pattern and identify patient groups that are likely to abandon the therapy.

Keywords: Obstructive Sleep Apnea · CPAP · Adherence · Clustering · Classification

1 Introduction

During sleep apnea attacks,the oxygen levels decrease and the brain sends a reflex impulse to a muscle in the throat. This causes a choking sensation, leading to interruption of sleep. This results in poor sleep quality, frequent awakenings, and daytime sleepiness.People with high blood pressure,chronic heart failure, stroke, type 2 diabetes, and depression are highly associated with OSA. Currently, about 22 million Americans are diagnosed with Sleep Apnea and thousands of new cases are identified every year[1]. In mild cases of sleep apnea, the treatment includes lifestyle changes such as a healthier diet and physical exercise. Fore more severe cases, aggressive therapy treatment is needed.

The continuous positive airway pressure (CPAP) device illustrated in Fig.1 is the most common and effective treatment for patients with moderate to severe OSA. CPAP device pumps air into the patient's respiratory system using an electro-mechanical air flow generator to create positive air pressure and prevent the upper airways from collapsing. Since the therapy is not curative, for optimal effectiveness patients should use the device constantly during the night. Owing to

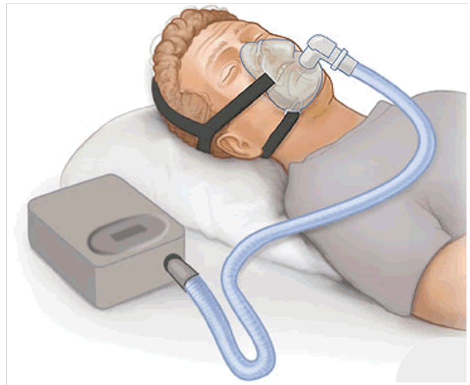

Fig. 1. Illustration of a patient using a CPAP device while sleeping

various social and psychological factors, the general adherence to CPAP therapy is as low as 60% of the patients[16]. Current approaches for adherence prediction, both one shot models and continuous prediction models[16], identify patients based on whether they will abandon therapy from 150 to 180 days after the start. But as it is possible that the behaviour of patients might not take 150 days to change, we propose a novel framework for monthly prediction of adherence to CPAP therapy.

2 Related Work

As per previous studies, the effectiveness of the therapy depends on how long the patient uses the CPAP device each night. Ensuring continuous use of the CPAP device every night by the patients is the biggest challenge in CPAP therapy[3]. According to [4] & [5], at least 4 hours of usage of CPAP device is necessary to see significant outcomes from the therapy. If the percentage of days in which the patient uses the device for less than 4 hours is less than 70% after 30 days (i.e. less than 21 days), then the patient can be categorized as non-adhering [6], [7]. Many attempts have been made to identify the key factors affecting CPAP adherence. [8] investigates the factors like device design, air humidity and psychological factors like claustrophobia. But none of them have been identified as predictive of adherence. Significant symptom improvement might not occur until 6 months into therapy [9] and adherence is majorly affected by symptom amelioration. So, waiting till 6 months to intervene could be too late and might lead to permanent therapy abandonment by the patient.

Hence, it is important to identify such patients who are likely to abandon therapy within 6 months for intervention. According to [10], the adherence prediction for 30 days can be achieved using the first 3 days data of number of hours the patients are using the CPAP device along with their age and race. In [2], the authors propose two one-shot models: one for 13^{th} day and the other for 30^{th} day, which predict the adherence of the patient 5 months later. [16] proposes a continuous adherence prediction framework (CTAP-CPAP), in which, 150 models were trained - one of each day until the 150^{th} day - and all of them predict the adherence of the patient between 150^{th} to 180^{th} day. This is better than the previous approach of only two models, but is still lacking in the sense that the prediction is only during the last 30 days of the 6 month therapy as shown in Fig.2. Therefore, we propose a novel approach which can be used to predict the adherence of patients to therapy for each month starting from the 2^{nd} month of the therapy.

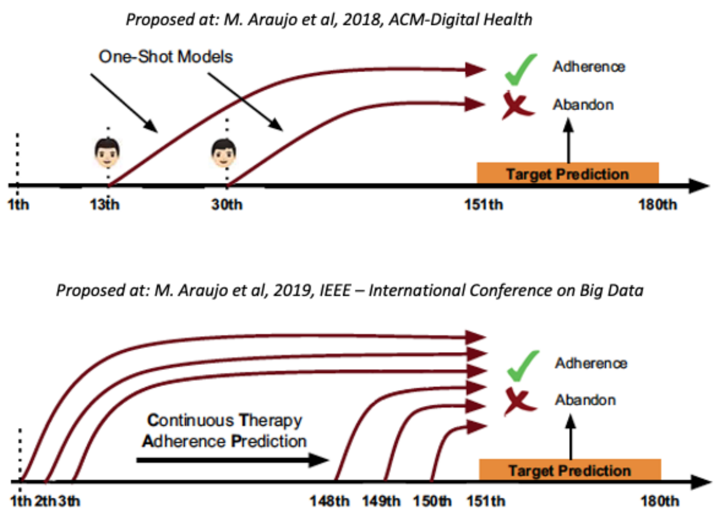

Fig. 2. Illustration of one shot models [2] and CTAP-CPAP [16] models. Each arrow represents one model trained for that day

2.1 Need for Continuous Adherence Prediction

The current models, both one-shot and CTAP, are able to predict adherence of patients to therapy well, but are only designed to do so for the 6th month of the therapy as shown in Fig.2. The goal of this work is to establish a framework that can monitor adherence of patients to therapy in each month following the first month of the therapy as shown in Fig.3.As mentioned in [11] and [12], patients who receive real-time feedback are more likely to follow the therapy. Hence, continuous monitoring is important in effectively treating a patient. TAPC-CPAP tries to identify patients who are likely to drop the therapy much before the 6^{th} month threshold and thereby helps in real-time intervention in such cases.

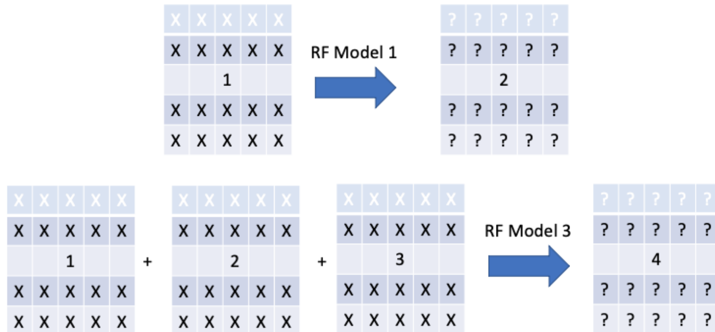

Fig. 3. TAPC-CPAP model uses extracted features from previous months to predict the adherence in the coming month of the therapy, thereby providing continuous prediction of adherence

This reduces the cost of telemedicine [13] and also impacts patients' lifestyle positively [14], [15].

3 Methodology

In this section we discuss the methodology and the steps we followed to achieve our best solution for the proposed problem.

3.1 Data

Tha data we used for this study is part of the Patient Characteristics and CPAP usage data actively tracked by Fairview Sleep Clinic and shared to us. The dataset contains the recorded signals for 1815 patients with nightly granularity. It contains the following parameters:

- **PRESSURE** value is the measurement of upper limit of pressure delivered by the CPAP device
- **AHI** value gives the average number of Sleep apnea events per hour
- **FACE TIME** value represents the number of hours the machine detects breathing (indicates usage of device)
- **LEAKAGE** value indicates the leakage of pressure in L/min (indicates the quality of mask fit)

The patients are randomly divided into training(80%) and testing(20%) sets. The classification models are trained using the training set data and the metrics are evaluated on the test sets.

3.2 Preprocessing

The granularity of the data is reduced to monthly by aggregating the fields to find the mean and standard deviation(std) of each of the parameters for each patient for each month. The aggregated values are then standardized using the 6-month mean and std of the parameters. This is done to normalize the parameters so that all of them will have a similar influence while clustering.

3.3 Clustering Patients

The main idea of this study is to cluster the patients into different groups based on the usage pattern and identify the groups that are not adhering to the therapy. Another salient feature of this study is to analyze the behaviour of the patients and find the transition patterns between different clusters, by using Markov Chain Principle and calculating the transition probabilities between the clusters.

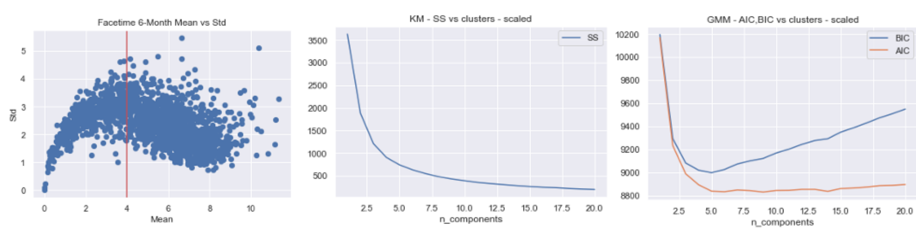

Fig. 4. For the distribution of 6-month aggregate values as shown in the first figure, it is clear that 4 is the apt number of clusters for both K-Means and the GMM models based on the elbow curves shown above.

As adherence is majorly defined based on the usage time of the device, we used the aggregated values of **FACE TIME** parameter (mean and std) to define the clusters. We used the K-means model and Gaussian Mixture Models(GMM) for clustering and compared the clusters created using both the models, with and without standardization of the parameters. The number of clusters were decided as 4 based on the elbow curve for both the models using sum of scores as the parameter for K-Means and AIC, BIC for GMM as seen in Fig.4 .

The models were trained on the 6-month mean and std and the same models were used to cluster the monthly aggregate values to preserve uniformity across different time periods, as the distribution of aggregates is similar across all time periods as shown in Fig.6.As shown in Fig.5 out of 4 combinations for K-Means and GMM using standardized and non-standardized data, K-Means with standardized data is the option that provides the best possible clustering that can be explained well and divided into adhering and non-adhering groups as explained below in Table.1. Using this trained K-means model, all the standardized monthly aggregates are now clustered into 4 clusters a shown in Fig.6.

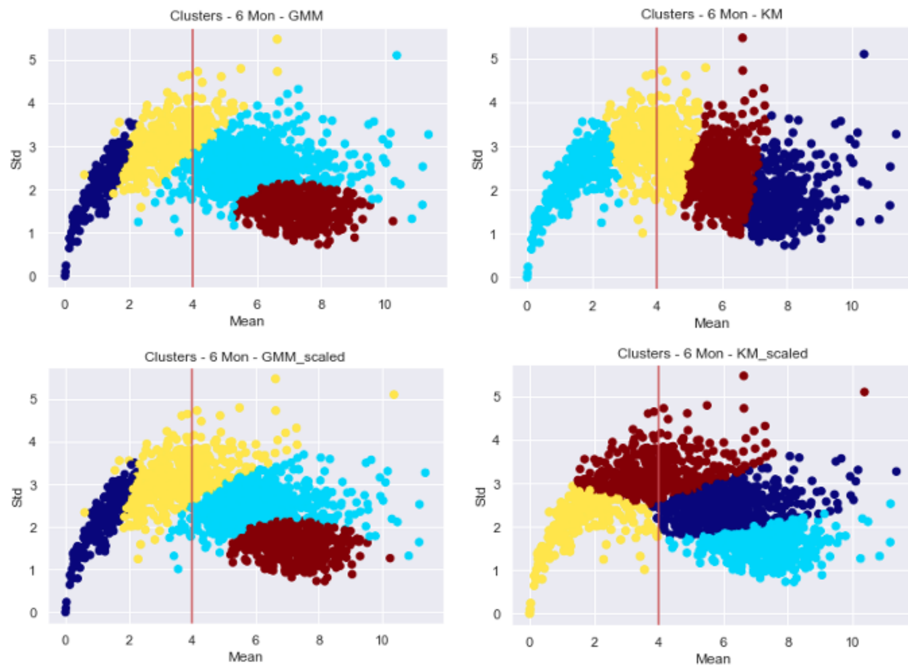

Fig. 5. Out of all the clustering combinations, K-Means with standardized data (d) is the one that explains the clusters well

Table 1. Description and categorization of the clusters

Cluster	Colour	Description	adherence
Cluster - 0	Yellow	Low mean usage and low standard deviation	Non-adhering
Cluster - 1	Red	Below threshold mean usage and high standard deviation	Non-adhering
Cluster - 2	Dark Blue	Above threshold mean usage and high standard deviation	Adhering
Cluster - 3	Sky Blue	High mean usage and low standard deviation	Adhering

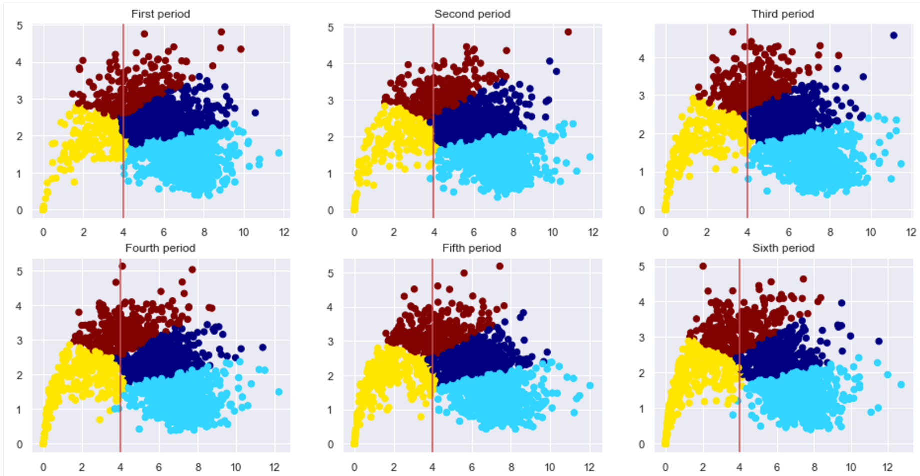

Fig. 6. Standardized monthly aggregates clustered using K-Means model

3.4 Behaviour Analysis

The other important feature of this study is the analysis of the transition behaviour of the patients from one month to the next month. It is clear from the sample heatmap shown in Fig.7 that, most of the patients in the extreme clusters (yellow & sky blue) tend to remain in the same cluster throughout the course of the therapy. But the patients in the other two clusters (dark blue & maroon) keep shifting between the clusters. Hence, considering the process as Markov chain where the clusters are the states, we computed the state transition probabilities between the states which are shown in Table.2.

Table 2. Transition probabilities between the clusters

Clusters	0	1	2	3
0	0.79	0.11	0.07	0.04
1	0.17	0.48	0.26	0.09
2	0.07	0.19	0.44	0.30
3	0.02	0.03	0.19	0.76

The final steady state probabilities shown in Table.3 indicate that about 39% (23+16) of the patients will eventually not adhere to the therapy and the rest will continue to more or less adhere to the therapy in the long term. These statistics can be used to monitor the performance of the therapy management.

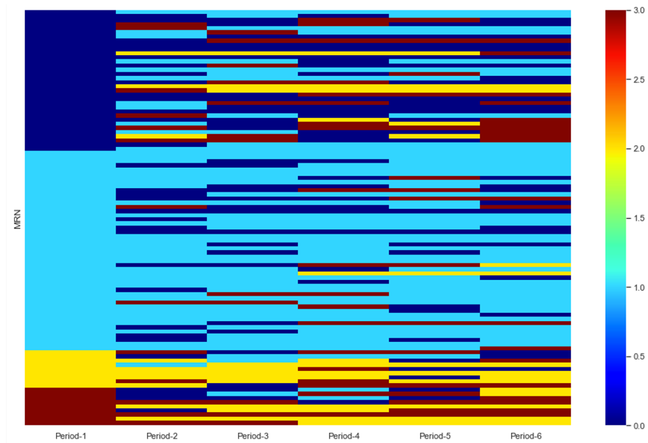

Fig. 7. Transition Heatmap for a random sample of 100 patients

Table 3. Steady State probabilities for the cluster transitions

Clusters	0	1	2	3
Probability	0.23	0.16	0.23	0.38

3.5 Adherence prediction

After clustering the patients using the K-Means model, 5 Random Forrest models, one for each month from 2 to 6, are trained to predict the clusters for the patients, using the data of the previous months as depicted in Fig.3. The data is split into training set(80%) and test set(20%) and the models are trained using training data. We followed a binary classification approach for this step as we only need to classify the patients as adhering and non-adhering. For this, we combined the clusters 0,1 into class-0 and 2,3 into class-1. The trained models are then tested on the test set and the results for this process are given in Table.4. We used accuracy as the metric to compare it with the baseline model of classifying all patients as adhering.

It is noticeable that although the class imbalance decreases the performance of the models improve. Hence, it is evident that TAPC-CPAP approach of categorizing patients into clusters and using them for adherence prediction is an efficient way for continuous prediction of adherence in Sleep Apnea therapy.

4 Conclusion and Future work

Though the existing models like one-shot models and CTAP-CPAP models perform a good job at prediction of adherence to CPAP therapy for sleep apnea, they

Table 4. Classification accuracy and adhering class proportions

Model	Accuracy	Class ratio (1/1+0)
RF 1: 1→ 2	0.818	0.738
RF 2: 1+2→ 3	0.813	0.697
RF 3: 1+2+3→ 4	0.847	0.653
RF 4: 1+2+3+4→ 5	0.852	0.645
RF 5: 1+2+3+4+5→ 6	0.882	0.636

do so only for a specific time period in the treatment (6$^{\text{th}}$ month). The proposed TAPC-CPAP framework provides a novel approach for continuous prediction of adherence to CPAP therapy right from the 2$^{\text{nd}}$ month of the therapy and thereby helps in identifying the patients who are likely to drop from the therapy much earlier than 6 months. The performance of the proposed models is significantly better than the baseline model of classifying everyone as adherent, even though there was inherent class imbalance present. This performance can further be improved by extracting time series features from the CPAP signal data and using them for clustering and classification as they provide much richer information compared to aggregation. Further, other sophisticated classification models can also be used to improvise on the performance of TAPC-CPAP framework.

References

1. N. A. Kerner and S. P. Roose, "Obstructive sleep apnea is linked to depression and cognitive impairment: evidence and potential mechanisms," The American Journal of Geriatric Psychiatry, vol. 24, no. 6, pp. 496– 508, 2016.
2. M. Araujo, R. Bhojwani, J. Srivastava, L. Kazaglis, and C. Iber, "Ml approach for early detection of sleep apnea treatment abandonment: A case study," in Proceedings of the 2018 International Conference on Digital Health. ACM, 2018, pp. 75–79.
3. H. M. Engleman, N. Asgari-Jirhandeh, A. L. McLeod, C. F. Ramsay, I. J. Deary, and N. J. Douglas, "Self reported use of cpap and benefits of cpap therapy a patient survey," Chest, vol. 109, no. 6, pp. 1470–1476, 1996.
4. T. E. Weaver and R. R. Grunstein, "Adherence to continuous positive airway pressure therapy: the challenge to effective treatment," Proceedings of the American Thoracic Society, vol. 5, no. 2, pp. 173–178, 2008.
5. H. M. Engleman, S. E. Martin, and N. J. Douglas, "Compliance with cpap therapy in patients with the sleep apnoea hypopnoea syndrome." Thorax, vol. 49, no. 3, pp. 263–266, 1994.
6. N. B. Kribbs, A. I. Pack, L. R. Kline, P. L. Smith, A. R. Schwartz, N. M. Schubert, S. Redline, J. N. Henry, J. E. Getsy, and D. F. Dinges, "Objective measurement of patterns of nasal cpap use by patients with obstructive sleep apnea," American Review of Respiratory Disease, vol. 147, no. 4, pp. 887–895, 1993.
7. T. E. Weaver and A. M. Sawyer, "Adherence to continuous positive airway pressure treatment for obstructive sleep apnea: implications for future interventions," The Indian journal of medical research, vol. 131, p. 245, 2010.

8. C. M. DiNapoli, “Improving continuous positive airway pressure adherence among adults,” Journal of Nursing Education and Practice, vol. 5, no. 2, p. 110, 2014.
9. C. A. Kushida, D. A. Nichols, T. H. Holmes, S. F. Quan, J. K. Walsh, D. J. Gottlieb, R. D. Simon Jr, C. Guilleminault, D. P. White, J. L. Goodwin et al., “Effects of continuous positive airway pressure on neurocognitive function in obstructive sleep apnea patients: the apnea positive pressure long-term efficacy study (apples),” Sleep, vol. 35, no. 12, pp. 1593–1602, 2012.
10. R. Budhiraja, S. Parthasarathy, C. L. Drake, T. Roth, I. Sharief, P. Budhiraja, V. Saunders, and D. W. Hudgel, “Early cpap use identifies subsequent adherence to cpap therapy,” Sleep, vol. 30, no. 3, 2007.
11. A. Malhotra, M. E. Crocker, L. Willes, C. Kelly, S. Lynch, and A. V. Benjafield, “Patient engagement using new technology to improve adherence to positive airway pressure therapy: a retrospective analysis,” Chest, vol. 153, no. 4, pp. 843–850, 2018.
12. X. Rafael-Palou, E. Vargiu, C. Turino, A. Steblin, M. S. de-la Torre, and F. Barbe, “Towards an intelligent monitoring system for patients with obstrusive sleep apnea,” EAI Endorsed Transactions on Ambient Systems, vol. 4, no. 16, 12 2017.
13. N. Fox, A. Hirsch-Allen, E. Goodfellow, J. Wenner, J. Fleetham, C. F. Ryan, M. Kwiatkowska, and N. T. Ayas, “The impact of a telemedicine monitoring system on positive airway pressure adherence in patients with obstructive sleep apnea: a randomized controlled trial,” Sleep, vol. 35, no. 4, pp. 477–481, 2012.
14. S. Kandel, I. Shalom, O. Kitzen, Q. Ahmed, and M. Weinstein, “Early intervention with single session” mask-fitting” improves cpap adherence in obstructive sleep apnea syndrome patients,” in A38. UNDERSTANDING CARE DELIVERY AND OUTCOMES IN OBSTRUCTIVE SLEEP APNEA. American Thoracic Society, 2018, pp. A1495–A1495.
15. C. Iber, L. Kazaglis, H. Steffens, E. Silbernick, and M. John, “Cpap adherence in sleep apnea managed with virtual care and ehr integration,” in SLEEP, vol. 40. OXFORD UNIV PRESS INC JOURNALS DEPT, 2001 EVANS RD, CARY, NC 27513 USA, 2017, pp. A195–A195.
16. M. Araujo, L. Kazaglis, C. Iber and J. Srivastava, ”A Data-Driven Approach for Continuous Adherence Predictions in Sleep Apnea Therapy Management,” 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 2019, pp. 2716-2725, doi: 10.1109/BigData47090.2019.9006476.