Defining and Monitoring Patients Clusters based on Therapy Adherence in Sleep Apnea Management

Mourya Karan Reddy Baddam
Computer Science & Engineering
University of Minnesota
Minneapolis, USA
badda004@umn.edu

Matheus Araujo
Computer Science & Engineering
University of Minnesota
Minneapolis, USA
arauj021@umn.edu

Jaideep Srivastava
Computer Science & Engineering
University of Minnesota
Minneapolis, USA
srivasta@umn.edu

***Abstract*—Obstructive Sleep Apnea (OSA) is a disorder in which breathing repeatedly stops and starts due to recurrent episodes of partial and complete airway obstruction during sleep. One of the common treatments for moderate and severe OSA cases includes the use of Continuous Positive Airway Pressure (CPAP) devices that keep the airways open. Unfortunately, about 40% of the patients using CPAP devices abandon their therapy within six months. In this work, we propose a method to cluster and monitor patients according to their therapy usage behavior aiming for a timely and appropriate intervention. In contrast to the simple rule-based methods currently employed by sleep clinics to identify non-adherent behavior, our approach uses clustering techniques to group patients based on their CPAP usage patterns. We also analyze their transition between clusters over the months using Markov Chain analysis. Our approach identifies patients who are likely to become non-adherent to therapy based on the usage patterns from previous months. This technique improves on the current methods by leveraging from data-driven frameworks the potential of predicting therapy adherence continuously as opposed to one-shot methods.**

***Index Terms*—Clustering, Obstructive Sleep Apnea, Sleep Therapy, Therapy Adherence, Markov Chain, Random Forest, SVM, XGBoost, Machine Learning**

I. Introduction

During sleep apnea events, the oxygen levels decrease, and the brain sends a reflex impulse to the muscles that support the soft tissues in the throat. That impulse causes a choking sensation, leading to interruption of sleep. The long-term effects of having multiple sleep apnea events per night are poor sleep quality, frequent awakenings, and daytime sleepiness. 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, defined by having an Apnea-Hypopnea Index (AHI)[1] less than 15 but greater than 5, the treatment includes lifestyle changes such as a healthier diet and physical exercise. However, more invasive treatment such as CPAP therapy treatment is needed for more severe cases. The USA solely has 8 million CPAP users [2], a population that grows 9% per year.

CPAP device is a mask connected to a pump that injects air into the patient's respiratory system to create positive air pressure and prevent the upper airways from collapsing. Since the therapy is not curative, patients should use the device constantly during the night for optimal effectiveness. Ensuring at least 4 hours of usage per night is essential to have significant outcomes from CPAP therapy [3], [4]. To be classified as long-term adherent, a patient should use the device for more than 4 hours for at least 70% of the days during a 30 days period (i.e. at least 21 days) [5], [6]. Owing to various social and psychological factors, the general adherence to CPAP therapy is as low as 60% of the patients [7]. The consequences of not treating sleep apnea can be devastating for an individual since its prevalence is highly associated with high blood pressure, chronic heart failure, stroke, type 2 diabetes, and depression [8], [9]. Administrating early intervention on individuals who are likely to not adhere to their therapy is critical to revert non-adherent behavior [10]. To perform a cost-effective early intervention, we need an accurate early identification of non-adherence to prioritize those individuals who need more attention.

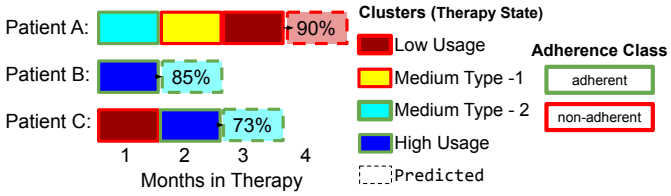

Fig. 1: Diagram showing how patients transits between clusters and how future adhrence can be predicted based on previous data.

In this work, we propose a novel framework that uses clustering techniques to categorize patients into different groups based on their therapy usage behavior and then use the clusters to identify new patients who may need intervention. Our approach defines the states to which a therapy trajectory for each individual can be mapped as illustrated in Figure 1. These states depend on the level of adherence and how consistent the patients are in their therapy. After showing that patients' behavior can change over time, we use Markov Chain analysis to calculate the transition probabilities of patients from one state to another state between any two consecutive

[1]AHI is the number of apnea or hypopnea events during one night, divided by the total hours of sleep

months in the first six months of CPAP therapy. Using the transition probabilities, we then compute the steady state probabilities of adherent and non-adherent groups, which serve as a metric to quantify the overall performance of sleep therapy management. Finally, given a set of new patients, we use Machine Learning classifiers to predict if an individual will adhere or not to their therapy.

II. Related Work

Many attempts have been made to identify the key factors affecting CPAP adherence. In [11], the authors investigated device design, air humidity and psychological factors like claustrophobia. But none of them have been identified as predictive of adherence. Adherence is majorly affected by symptom amelioration, however significant symptom improvement might not occur until 6 months into therapy [12]. Alongside the adherence problem, most insurance plans require that patients are long-term adherent in the 6^{th} month of therapy, or they will no longer cover the treatment [13]. Thus, waiting till 6 months to intervene could be too late and might lead to permanent therapy abandonment by the patient.

As mentioned in [14] and [15], patients who receive real-time feedback are more likely to follow the therapy. Hence, continuous monitoring is important to effectively engage a patient in their treatment. Our framework 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. This reduces the cost of telemedicine [16] and also impacts patients' lifestyle positively [17], [18].

Early intervention requires early identification of patients who are likely to abandon therapy within 6 months. According to [19], 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 [20], 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. In a later work, [7] 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. These studies prioritize patients but only consider predictions for the 6^{th} month of therapy.

In [21], the authors performed clustering analysis on 161 CPAP patients. They found 4 clusters of patients (Great Users, Good Users, Low Users, and Slow Decliners). Also, in [23], 3 clusters of patients were found in a set of 318 patients. These works only focused on defining clustering and neither investigated how patients transit between clusters nor predicted adherence of patients to the therapy based on the clusters.

In this work, we propose an approach that can be used to predict patients' adherence to therapy for each month starting from the 2^{nd} month of the therapy. In contrast to the previous works, we quantify the transition probability of a patient who was in an adherent state to change back to a non-adherent state and vice-versa. We are the first to define therapy behavioral states and investigate how patients change their behavior over time from the first month of CPAP therapy.

III. Methodology

In this section we discuss the the steps we followed to achieve the results for the proposed problem.

A. Data

The data we used for this study was collected and shared to us by the Fairview Sleep Clinic. The dataset contains the daily time on face (number of hours the machine detects breathing) from 1815 patients who used CPAP machines with nightly granularity. Every patient considered for this work has undergone at least 6-months of therapy.

To evaluate our predictive models, we performed a retrospective study by randomly dividing patients into training(75%) and testing(25%) sets. The parameters of the classifiers used in this work were fit based on patients information from the training set. The prediction results are based on new patients from the test set.

B. Preprocessing

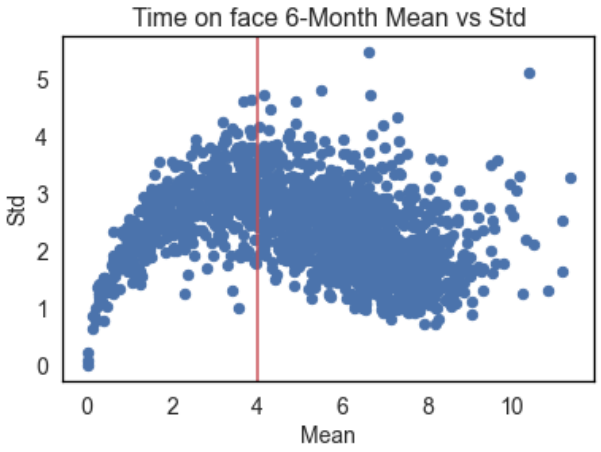

Fig. 2: Distribution of 6-month aggregate values.

We reduced the granularity of the time-series daily data to monthly by aggregating the time on face signal to the mean and standard deviation (std) for each patient. In previous studies, such as in [21], the slope of a linear regression on the hours of usage was also considered an aggregating metric. However, the slope did not incorporate significant information into the models and was disregarded. To define the clusters, the mean and std for all the 6 months of data is also computed for each patient.

Fig. 2 shows the distribution of the 6-month aggregate mean and std of usage for all the patients in our dataset. The red line represents the threshold of 4 hours, which is typically considered as the minimum hours of usage for a patient to be considered adherent. The range for the mean is almost twice the range for std.

Since most clustering methods are sensitive to different feature magnitudes due to their distance metrics, we scaled both

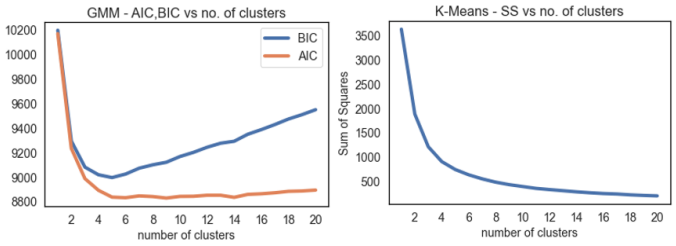

Fig. 3: Both the plots, have a sudden change of slope at K=4. Hence, 4 is the optimum number of clusters for both K-Means and the GMM models under the elbow method.

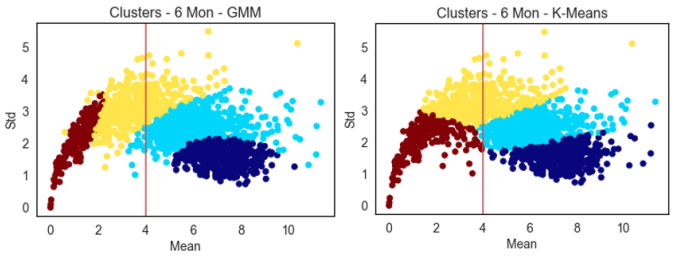

Fig. 4: Of the two algorithms, K-Means gives clusters that are more explainable and can be categorized as adherent(blue, cyan) and non-adherent(red, yellow)

std and the mean, using a standard scaler[2] [22] for monthly aggregates values. The scaler was fitted on the aggregated values of all six months. We normalized the features so that they will have a similar magnitude while clustering. Equation 1 denotes the transformation for the scaling process where z is the transformed feature, x is the original feature, and μ and σ are the mean and the standard deviation of the feature from 6-month aggregates data.

$$z = \frac{(x - \mu)}{\sigma} \tag{1}$$

C. Clustering Patients

To define the clusters of patients in the CPAP therapy, we analyzed their usage pattern derived from the pre-processing step. In this work, we use the terms clusters and therapy states interchangeably.

K-Means[3] [22] and Gaussian Mixture Models[4](GMM) [22] were both evaluated as potential clustering algorithm to be used in our work. The models were trained on the scaled 6-month aggregates data for each patient. The most important parameter in these clustering algorithms is to choose K, the number of clusters, that we expect to be generated from the clustering algorithm. In Figure 3 we show the plots of AIC (Akaike Information Criterion) & BIC (Bayesian Information Criterion) for the GMM algorithm. In the same figure we show the SS (sum of squares) for the K-Means algorithm. By using the elbow curve method in both plots, we decided that the number of clusters to be 4.

As shown in Figure 4, GMM formed clusters are more dependent on the mean value than the standard deviation. Also, comparing to the threshold of 4 hours of mean usage, none of the GMM's clusters except the blue cluster can be categorized as adherent. On the other hand, K-Means provides a better clustering. It takes into consideration the standard deviation as well as divides the adhering and non-adhering groups. Although the yellow cluster has many values before and after the threshold, given the high standard deviation of therapy usage, we can map the patients from the yellow cluster to a non-adherent group due to their undesired erratic therapy usage. We summarize the interpreted description of the clusters created by the K-Means model in Table.I.

Cluster	Colour	Description	Adherence	Distribution
0	Red	Low mean and low std dev	Non-adhering	15.53%
1	Yellow	Low mean and high std dev	Non-adhering	24.90%
2	Cyan	High mean and high std dev	Adhering	28.70%
3	Blue	High mean and low std dev	Adhering	30.85%

TABLE I: Description and categorization of the clusters found by the K-Mans method where K=4

We used the fitted K-Means model to cluster patients during 1st, 2nd, 3rd, 4th, 5th and 6th month separately using the monthly aggregate data. We did not retrain the model for each month to preserve uniformity across different time intervals while classifying patient therapy states. The distribution of patients over the mean and std axes in each month and their clusters can be visualized in the Figure 6. Even though the distributions look similar in all the months, the number and the therapy usage behaviour of the patients are different for each month.

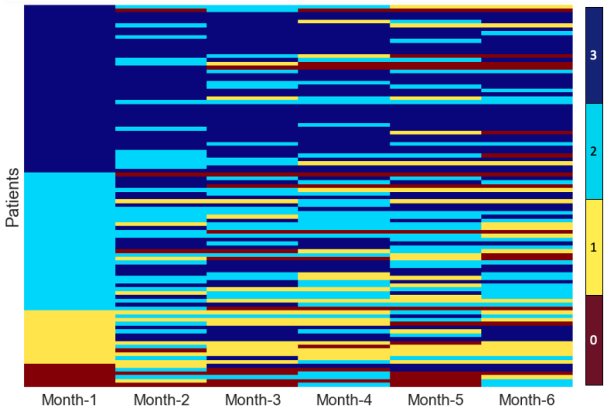

Fig. 5: Heatmap showing cluster transition over time for a random sample of 100 patients. Each thin line is one patient. As reference, the first month was ordered based on the clusters assigned.

D. Cluster Transition Probabilities

We analysed how patients change from one therapy state to another over time starting from 1st month to the 6th month of

[2]https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html

[3]https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html

[4]https://scikit-learn.org/stable/modules/mixture.html

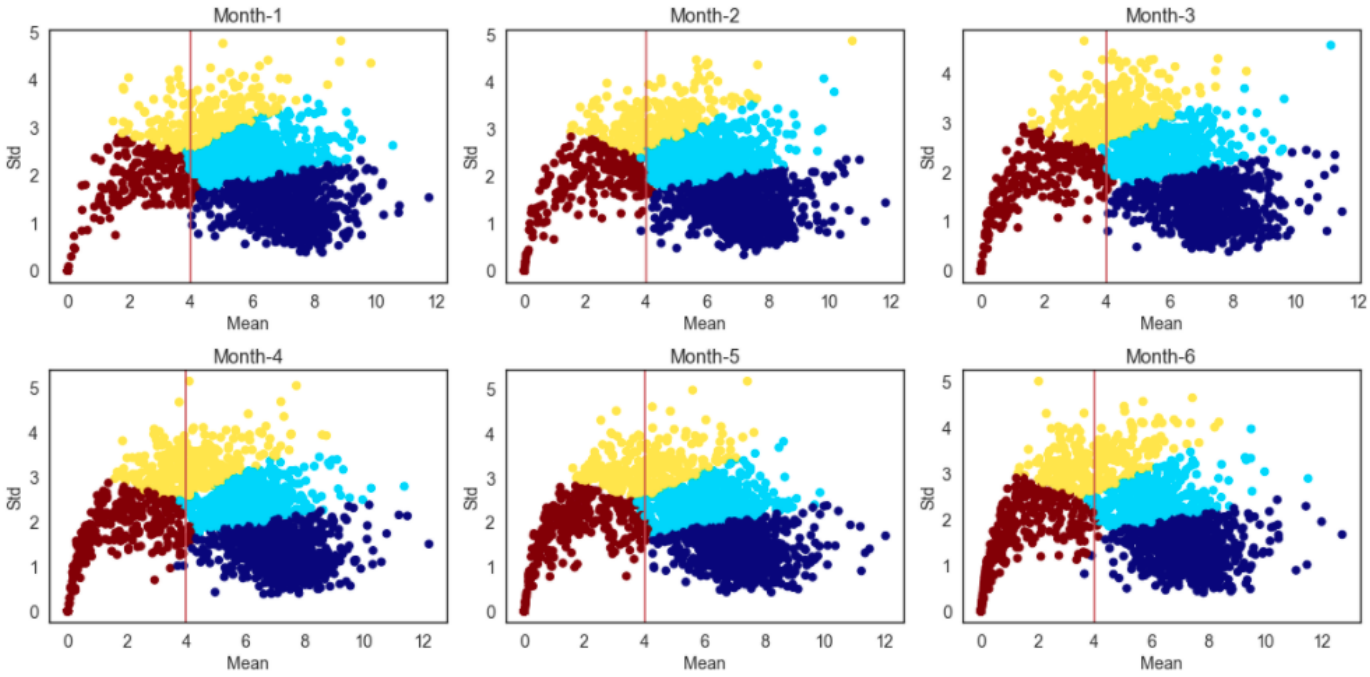

Fig. 6: We show how patients would be distributed per cluster in each month using a K-Means fitted model trained on aggregates of the 6-month period. The metrics used are the mean and standard deviation of therapy usage.

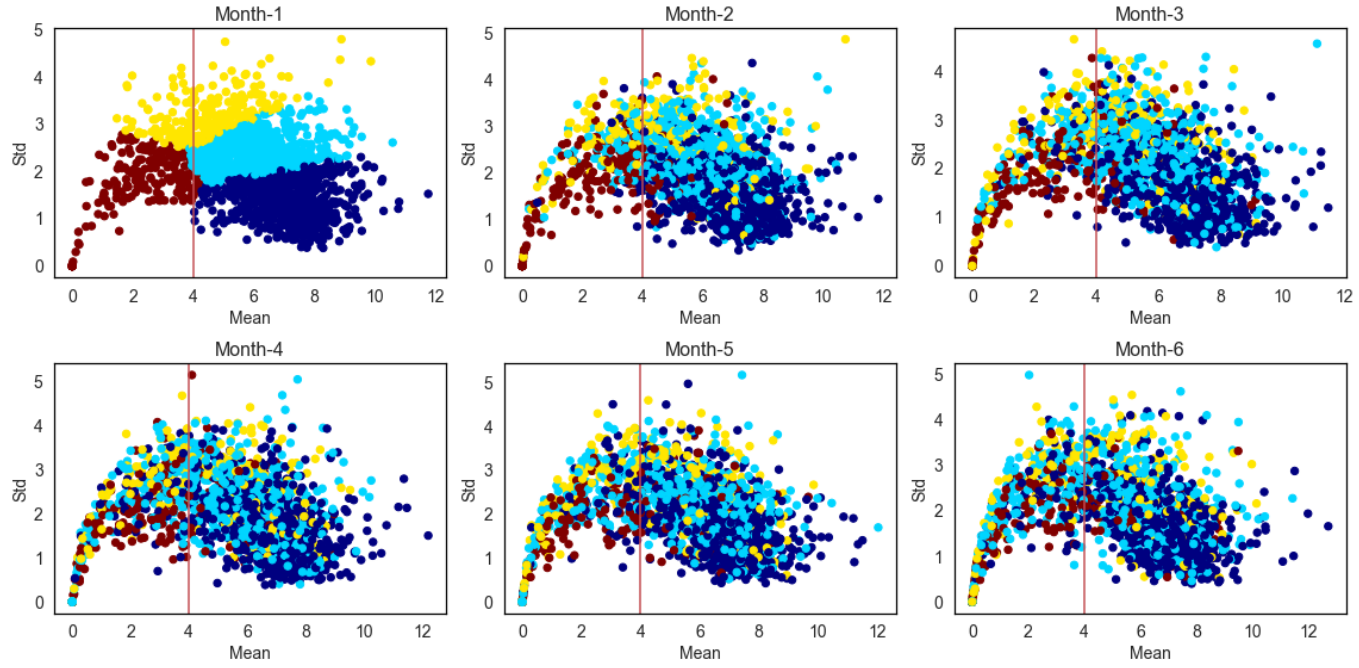

Fig. 7: This plot shows the same patients metrics as Figure 6. However, now the color from each patient's first assigned cluster was preserved. This shows how a patient can change from cluster to cluster.

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
Steady-State Probability	0.23	0.16	0.23	0.38

TABLE II: Transition probabilities matrix between the clusters and Steady State probabilities for each cluster

therapy. The heatmap in Figure 5 shows these transitions for a random stratified sample of 100 patients where each column is a month, and each thin horizontal line is a patient. We show the first month ordered by cluster from the one with lowest adherence (red) to the one with highest adherence (blue) and the subsequent columns are the clusters that patients changed to in the subsequent months. From the heatmap we observe that patients in the extreme clusters (red & blue) tend to remain in the same cluster throughout the course of the therapy. The patients in the other two clusters (yellow & cyan) keep shifting between clusters. This is also evident from Figure 7 which shows the patient distributions for all the months, but the colour of the patients represent the clusters from 1st month. The extreme clusters (Red and Blue) appears to disperse very less over time compared to the dispersion of the other two clusters (Yellow and Cyan).

To quantify this dynamic process, we performed a Markov Chain analysis where the clusters are the potential transition states that a patient can be. We computed the state transition probabilities (p_{ij}) between the states using 2 and the steady state probabilities (π_j) for each state using 3 and 4 [24]. The results of this analysis are shown in Table II.

Transition probabilities:

$$p_{ij} = \frac{\sum_{t=1}^{t=5} Count(i \rightarrow j)}{\sum_{j=0}^{j=3} \sum_{t=1}^{t=5} Count(i \rightarrow j)} \tag{2}$$

Steady-state probabilities:

$$\pi_j = \sum_{k=0}^{k=3} \pi_k p_{kj} \tag{3}$$

$$\sum_{j=0}^{j=3} \pi_j = 1 \tag{4}$$

E. Therapy State Prediction using Machine Learning

After analyzing the state transitions of the patients from one month to another using the K-Means model, we evaluated if we can predict future month adherence for each individual. We assessed the performance of SVM [22], Random Forest[5] [22] and XGBoost [25] classifiers for adherence prediction for each month starting from 2nd to the 6th. We used one classifier of each type for each month trained on the separate training sets for different months, created from past usage data as shown in the 'Model' column of Table III. The hyper-parameters for all the classifiers were chosen by performing a grid search over a possible range of parameter values using a 5-fold cross validation on the training sets . 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 (non-adhering) and clusters 2,3 into class-1 (adhering), as shown Table I. The trained models are then tested on separate test sets and the results for this prediction are given in Table.III. We used accuracy and ROC-AUC score as the metrics to compare the predictive performance of different types of classifiers and accuracy to compare them with a baseline model. The baseline model assumes that the next month cluster is the same as the current month cluster for any patient. This is a reasonable assumption since most methods in the literature do not assume that patients might change their usage behavior. The ROC-AUC metric was chosen since it is a better indicator of performance when the dataset is not balanced[6].

IV. Results & Discussion

In Table II, we quantitatively support the observation highlighted in Figure 5 and Figure 7 in which extreme clusters (red and blue) are more stable than intermediary clusters. The first and last elements on the diagonal of this probability transition matrix are much higher than the other two diagonal elements. As the transition probabilities suggest, the patients in the red cluster (cluster 0) have the highest tendency (79% + 11%) to remain non-adherent. The yellow group of patients (cluster 1) has 3 times more chance to go to an adherence group (35% against 11%) than the red group of patients. It is also important to notice that patients on the most adherent cluster (cluster 3 - blue) have the lowest transition probability to go to cluster 0 or cluster 1 (only 5%). This reassures the need for early and effective intervention aiming at a change of habits. After the desired change is accomplished, it is likely that the patients continue with good behavior.

Hence, this approach not only predicts the adherence of patients but also gives an estimate of the severity of non-adherence. The final steady-state probabilities shown on the last line of Table II indicate that about 39% (23+16) of the patients will eventually not adhere to the therapy. Fortunately, the steady-state of the adherent groups are higher than non-adherent showing their likelihood of staying adherent in the long term. This result is also supported by [7] which mentions the adherence rate is as low as 60%. Hence, this metric can be used to monitor the performance of therapy management in the long run.

In the task of predicting future therapy state for a new set of patients given previous data, our classifiers presented interesting results. Their accuracy and AUC-ROC score increased according to the month of therapy as shown in Table III. The AUC-ROC score starts at about 0.87 for the first month and reaches 0.95 for the 6th month, meaning less false negatives and less false positive rates. The SVM classifiers have slightly

[5]https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html

[6]https://scikit-learn.org/stable/modules/generated/sklearn.metrics.roc_auc_score.html

Model	SVM Acc.	SVM AUC	RF Acc.	RF AUC	XGB Acc.	XGB AUC	Baseline Acc.
1→2	0.85	0.88	0.84	0.88	0.82	0.87	0.82
1+2→3	0.81	0.83	0.80	0.89	0.81	0.89	0.80
1+2+3→4	0.82	0.89	0.81	0.89	0.82	0.90	0.84
1+2+3+4→5	0.87	0.92	0.86	0.93	0.85	0.93	0.86
1+2+3+4+5→6	0.89	0.94	0.88	0.95	0.87	0.95	0.87

TABLE III: Non-adherence classification accuracy & AUC-ROC score for our random forest models. We also present the baseline accuracy as reference.

better accuracy over Random Forest and XGBoost classifiers. But both Random Forest and XGBoost classifiers have higher ROC-AUC scores compared to the SVM classifiers, which indicates they have higher confidence in prediction. The continuous improvement over time is expected since more data is available for each subsequent month. When comparing our classifiers performance with the baseline, we observe a consistent better prediction accuracy. Thus, machine learning is justified in our context. These results indicate that the approach of categorizing patients into clusters and using it for adherence prediction is an efficient way for continuous prediction of adherence to the CPAP therapy.

V. Conclusion

In this, we established a framework that can monitor adherence of patients to therapy in each month following the first month of CPAP therapy. This approach contrasts with previous works that tried to model adherence only for a specific timeframe. Here we used a cluster analysis technique to define patients' clusters according to their CPAP usage. We showed how patients can change from one cluster to another cluster over the first 6 months of therapy. After observing these dynamics, we applied a Markov Chain analysis to calculate the likelihood of a patient transitioning between clusters, giving the transition probability matrix of the process. We also leveraged machine learning potential by using SVM, Random Forest and XGBoost classifiers to predict the next state for each patient. The performance of our classifiers is outstanding, reaching 0.95 for the AUC-ROC sore and 88.5% of accuracy. We left the investigation of other time-series features for the clustering model as well as the use of deep learning approaches for classifying next therapy states as future work.

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] D'Ambrosio, D. (2019, March 19). Healthcare Startup Grows 2,805 Percent In 7 Years With CPAP Sanitizer. Forbes. https://www.forbes.com/sites/danieldambrosio/2018/12/21/healthcare-startup-grows-2805-percent-in-7-years-with-cpap-sanitizer/.

[3] 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.

[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] 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.

[6] 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.

[7] 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.

[8] Botros, N., Concato, J., Mohsenin, V., Selim, B., Doctor, K., & Yaggi, H. K. (2009). Obstructive sleep apnea as a risk factor for type 2 diabetes. The American journal of medicine, 122(12), 1122-1127.

[9] Ejaz, S. M., Khawaja, I. S., Bhatia, S., & Hurwitz, T. D. (2011). Obstructive sleep apnea and depression: a review. Innovations in clinical neuroscience, 8(8), 17.

[10] Jonikas, M. A., & Mandl, K. D. (2012). Surveillance of medication use: early identification of poor adherence. Journal of the American Medical Informatics Association, 19(4), 649-654.

[11] C. M. DiNapoli, "Improving continuous positive airway pressure adherence among adults," Journal of Nursing Education and Practice, vol. 5, no. 2, p. 110, 2014.

[12] 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.

[13] Centers for Medicare and Medicaid Services. (2016). Positive airway pressure (PAP) devices: complying with documentation & coverage requirements. US Centers for Medicare & Medicaid Services. Baltimore, MD: Medicare Learning Network.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] 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.

[21] Babbin, S. F., Velicer, W. F., Aloia, M. S., & Kushida, C. A. (2015). Identifying longitudinal patterns for individuals and subgroups: an example with adherence to treatment for obstructive sleep apnea. Multivariate behavioral research, 50(1), 91-108.

[22] Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011

[23] D. Ferreira-Santos and P. Pereira Rodrigues, ”Finding Groups in Obstructive Sleep Apnea Patients: A Categorical Cluster Analysis,” 2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS), Karlstad, Sweden, 2018, pp. 387-392, doi: 10.1109/CBMS.2018.00074.

[24] Gubner, J. A. (2006). Probability and random processes for electrical and computer engineers. Cambridge University Press.

[25] Tianqi Chen and Carlos Guestrin. 2016. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD ’16). Association for Computing Machinery, New York, NY, USA, 785–794. DOI:https://doi.org/10.1145/2939672.2939785