Wound and Episode Level Readmission Risk or Weeks to Readmit: Why do patients get readmitted? How long does it take for a patient to get readmitted?

Subba Reddy Oota  
Woundtech Innovative Healthcare Solutions  
soota@woundtech.net

Nafisur Rahman  
Woundtech Innovative Healthcare Solutions  
nrahman@woundtech.net

Shahid Saleem Mohammed  
Woundtech Innovative Healthcare Solutions  
shmohammed@woundtech.net

Jeffrey Galitz  
Woundtech Innovative Healthcare Solutions  
jgalitz@woundtech.net

Ming Liu  
Woundtech Innovative Healthcare Solutions  
mliu@woundtech.net

ABSTRACT

The Affordable care Act of 2010 had introduced the readmission reduction program in 2012 to reduce avoidable readmissions to control rising healthcare costs. Wound care impacts 15% [16] of Medicare beneficiaries, making it one of the major contributors of medicare health care cost. Health plans have been exploring proactive health care services that can prevent wound recurrences and readmissions from controlling wound care costs. With the rising costs of the Wound care industry, it has become of paramount importance to reduce wound recurrences & patient readmissions. What factors are responsible for a Wound to recur, which ultimately leads to hospitalization or readmission? Is there a way to identify the patients at risk of readmission before the occurrence using data-driven analysis? Patient readmission risk management has become critical for patients suffering from chronic wounds such as diabetic ulcers, pressure ulcers, and vascular ulcers. Understanding the risk & the factors that cause patient readmission can help care providers and patients avoid wound recurrences. Our work focuses on identifying patients who are at high risk of readmission and determining the time period within which a patient might get readmitted. Frequent readmissions add financial stress to the patient & Health plan and deteriorate the patient’s quality of life. Having this information can allow a provider to set up preventive measures that can delay, if not prevent, patients’ readmission.

On a combined wound & episode-level dataset of patient’s wound care information, our extended autoprognosis achieves a recall of 0.92 and a precision of 0.92 for predicting a patient’s readmission risk. For new patient class, precision and recall are as high as 0.91 and 0.98, respectively. We can also predict the amount of time (in weeks) it might take after a patient’s discharge event for a readmission event to occur through our model with a mean absolute error of 2.3 weeks.

KEYWORDS

Patient’s readmission risk, AutoPrognosis, Health care, Wound care, Chronic Wound management, Readmission prevention, Cost Control, Machine Learning

1 INTRODUCTION

Nearly 3.3 million patients were readmitted to the hospital within 30 days of being discharged in the United States as per the Agency of Healthcare Research in the year 2011 [4]. Also, over $41 billion were spent due to patient readmissions in 2011 [4]. For wound-ulcer specific readmissions, over $250 million were spent on readmissions that occurred due to diabetic wounds, and more than $11 billion were spent on pressure ulcer related readmissions 1. Patient readmissions can significantly increase cost and lead to federal fines on hospitals for poor clinical outcomes. It has become important for Wound care providers to focus on the readmission problem by determining the risk & cause of the readmission. Some of the critical questions that we try to address in our work are: (1) What factors drive wound-related readmissions?, (2) Are patients returning with new wounds or with the same wounds to the care?, (3) What is the risk of an existing wound to recur in the future?, (4) How much impact does patient’s non-compliance in matters such as wound dressing, diet, medication, compression, and exercise have on patient’s readmission risk?, (5) What is the overall readmission risk for a patient provided their wound history & non-compliance history?, and (6) In how many weeks, can we expect a patient to end up in the hospital due to wound-related problems? [2].


\footnotesize Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

\footnotesize CODS COMAD 2021, January 2–4, 2021, Bangalore, India
© 2021 Association for Computing Machinery.
ACM ISBN 978-1-4503-8817-7/21/01. . . $15.00
https://doi.org/10.1145/3430984.3431005

\footnotesize 1https://www.ahrq.gov/sites/default/files/publications/files/putoolkit.pdf
To conduct our research work, we have collected a dataset of over 20,000 patients who were managed by a Wound care provider. Using data engineering techniques, we created 2 datasets from the initial dataset to represent the patient data at two different granularities. The first dataset represents a patient’s wound care episode with the wound care provider; a patient will have more than one episode if they were readmitted back to our Wound care provider’s facility. The second dataset represents all wounds that were treated by the Wound care provider for different patients; a wound can have multiple records if it had recurred in the past. Our wound care provider had implemented a process through their Electronic Health Record (EHR) system to identify a recurring wound from a new wound, which helped us determine a new wound from a recurring wound. This is an important piece of information for our work as we are trying to determine the impact of recurring wounds when a patient readmits. Almost 50% of our readmit patients returned to care due to a recurring wound making wound recurrence a major factor for readmissions. The remaining patient’s returned to care due to new wounds resulting from multiple factors such as old age, existing co-morbid conditions, and non-compliance.

After completing the data engineering task, we had to address the problem of testing the completeness of the data. To ensure that we are correctly categorizing patients as readmit or Non-readmit, we had to ensure that the patient who was discharged from our wound care provider’s service did not end up into another Wound care provider’s facility for care. In order to tackle this challenge, we limited our analysis to patients who had been discharged at least 2 months ago & had responded to the patient engagement program, a patient outreach program to check the status of the patient’s Wound after their discharge. Also, all the patients who joined the services of our Wound care provider were asked to provide Wound acquired date to keep track of Wound age & its recurrence status.

We observed that among these patients, 33% of patients had a history of readmission & were among those with the highest risk of readmission.

Our dataset mainly consisted of Patient demographics data, Patient Episode facts, Patient’s existing co-morbid conditions information, Patient’s attributes such as, Age, BMI, Location, Braden score can play a crucial role in determining a patient’s readmission risk. Apart from this, co-morbid conditions & non-compliance data can help us assess an individual’s risk of developing new wounds. Traditionally, a patient’s readmission risk has been assessed manually by the clinicians & care monitors by studying the patient’s entire clinical history. Some of the challenges of this approach are (i) Time-consuming process, (ii) Lack of consistency among assessors in their assessment process when the group of patients that require assessment is large, (iii) The reliability of these assessments for Clinical decision making is limited. Automating the patients’ risk of readmission can overcome most of these limitations and help clinicians make effective and informative decisions.

In this paper, our first goal is to build a model to predict the patient’s risk of readmission. To achieve the first goal, our proposed system mainly depends on three datasets such as, (i) would-level information, (ii) episode-level information, and (iii) a combination of wound and episode level information. Further, for Wounds that possess a high risk of recurrence, our second goal is to predict the time period during which a readmission event, i.e., weeks from the date of discharge, can be expected. We are interested in extracting information from the clinical records using machine learning models and AutoPrognosis [1]. Often machine learning models have strong predictive power, but our main focus is to build a system for automating the design of predictive modeling pipelines tailored for clinical prognosis. A model is more useful as a clinical tool if the physician understands the features underlying its predictions.

The main contributions of this paper are as follows.

Table 1: Statistics for the patient’s readmission risk attributes in our dataset. The bottom row displays common attributes across Wound and Episode level datasets.

| Dataset            | Attributes                                                                 | #Instances | #Classes                                      |
|--------------------|---------------------------------------------------------------------------|------------|-----------------------------------------------|
| Wound-Level        | WoundStatus, PatientDischargeStatus, PalliativeCare, Wound/Ulcer Type, Wound Location, Wound Stage, DaysinTXforWounds, AvgPainLevelforWound, VisitsforWound, DaysinTXforPatients | 90328      | WoundRecurrence: {Recurring Wound (24886), New Wound (65442)}, Patient Category: { Re-AdmittedPatient (46620), NewPatient (43708)} |
| Episode-Level      | WoundsforEpisode, ChronicWoundsforEpisode, AvgDaysinTXforwounds, AvgPainLevelforEpisode, LowerExtremityWoundsforEpisode, AVGTemperature, Diabetes, Anemia, EndStageRenalDiseasewithdialysis, VenousInsufficiency, ChronicObstructivePulmonaryDisease, AtheroscleroticHeartDisease, CoronaryArteryDisease, Smoking, Edema, PeripheralArterialDisease(PVD), EndStageRenalDiseasewithdialysis(CKD), Hypertension, CongestiveHeartFailure, Obesity, WeightGain, MarkedWeightChange | 45261      | Patient Category: (Re-AdmittedPatient (21521), NewPatient (23930)) |
| EpisodeNumber, PtAge, NonComplianceWoundVisitsRate, NonComplianceDietRate,NonComplianceOffLoadRate, NonComplianceExerciseRate, NonComplianceMedicationRate, NonComplianceLimbRate, NonComplianceCompressionRate, NonComplianceDressingRate, NonComplianceSmokingRate, NonComplianceHBOVisitsRate |

To our knowledge, this is the first study to cater to the needs of practitioners, researchers, and policymakers in the field of wound care and readmission in India. The findings of this study can help in the development of interventions that can mitigate the risk of readmission.
We formulate the problem as a two-stage automation method for readmission risk and a number of weeks to readmit prediction.

In the first stage, we analyze the wound-level and episode-level datasets, feature analysis for risk of readmission, and categorize the patient as New Patient or Readmit.

In the second stage, we build a model to predict the number of weeks for a patient’s risk of readmit, which uses similar features from stage 1.

We integrate the two successful machine learning models LightGBM & CatBoost into AutoPrognosis Framework.

2 RELATED WORK

Patient readmission risk models have become effective tools in clinical decision making and provide several benefits to both health care providers and patients [8, 21]. The earlier works in the literature focused on analyzing the readmission risk from various patient datasets with different ulcer types such as diabetic [4, 19], pressure ulcers [7]. The patient readmission risk prediction models were built with the aim of identifying the patients with a high risk of hospital readmissions using direct specific interventions such as demographic details of a patient, clinical procedure-related, and diagnostic-related features for patients above 65 years of age. These models have shown to diminish the readmission effectively rates for patients after hospital discharge [14, 15, 20]. However, these studies discriminate poorly on readmissions due to the non-availability of the patient’s demographic details, medication reconciliation, and patient’s education details, etc.

Machine Learning Models for ReAdmission Risk

Motivated by the immense success of machine learning & deep learning models in AI, there has been much focus on applying machine learning & deep learning to electronic medical records. State-of-the-art AI solutions that have demonstrated high performance in diagnosing & detecting diseases, risk prediction [11], and patient sub-typing [3, 5] have been a source of motivation for our research work. All the existing works on patient’s readmission risk have used simple machine learning classifiers for risk prediction such as logistic regression, naive-bayes, and SVM models [6, 18, 21]. Moreover, the model predictions results vary from hospital to hospital, data to data. Also, these studies use traditional feature engineering methods when handling categorical and numerical variables, and each model requires separate hyper-parameter tuning. To overcome the above limitations, recently, an automated framework has been developed, known as DeepSurv (deep Cox proportional hazards neural network) [12]. Some other existing survival models that dealt with healthcare applications include deep active analysis, deep recurrent analysis, deep integrative analysis. Recently, a new framework (AutoPrognosis) [1] was developed to automate the design of predictive modeling pipelines tailored for clinical prognosis outperforms the earlier state-of-the-art models.

3 FEATURE ANALYSIS

In this section, we have offered a detailed analysis of features at wound-level and episode-level datasets. There are common attributes present in both the datasets, including admission date, patient age, episode number reported in the bottom row of Table 1.
leads to further deterioration of the wound. As stage gets to a higher level it becomes more difficult to treat a wound. Figure 3 shows that wound with full-thickness stage is having the high chances of recurring in future.

3.2 Episode-level Feature Analysis

Table 1 middle row showcases the attributes used in episode-level risk of readmission. Out of 39 episode-level attributes, we majorly discuss the important features such as Total Chronic wounds treated, Days in treatment & presence of Co-morbid conditions.

Admission Date: The admission date attribute reported the date when the patient started receiving Wound care. In the current dataset, we observed that we had patients whose wound care episodes ranged from 2002-01-01 to 2020-03-20. In order to simulate the current health care trends, we decided to limit the patient dataset by including patients whose first Wound care episode was after the admission date of ≥ 2015-01-01.

Patient Age: Figure 4 displays the patient age histogram for the combined dataset where the majority of patients were in the age range of 65 to 80. The average patient age in the selected dataset is 75, the maximum age is 110, and the minimum age is 13 in our dataset. As discussed earlier, age plays a critical role in the formation of new wounds due to various medical complications that come with age.

4 APPROACH

In practice, it is an arduous task to select the right imputation method for handling missing values & it is hard to select the best classifier and fine-tune it to specific parameters. Moreover, running each model manually on a large number of samples is a cost & time-consuming task. The current existing framework AutoML [9]
does not support imputation and calibration stages are particularly important for clinical prognostic modeling. So, we employed a recent successful model AutoPrognosis useful for automated clinical prognostic models. In this paper, we integrated some of the recently acclaimed tree-based classifiers LightGBM [13] and CatBoosting [17] with AutoPrognosis framework.

4.1 AutoPrognosis Framework

The latest successful AutoPrognosis framework contains the following components, including (i) 8 imputation algorithms, (ii) 10 feature processing techniques, (iii) 20 machine learning classifiers, (iv) 3 calibration methods, and a total number of hyper-parameters (106), which is less than Auto-SKlearn framework (110) [10]. The core idea behind the AutoPrognosis framework is to configure ML pipelines automatically, where every pipeline comprises of 4 components mentioned above. The Autoprognosis pipeline configuration shown in Figure 7 is as follows. Let \( P=(A_d, A_f, A_p, A_c) \) be a pipeline with the sets of imputation algorithms \( A_d \), feature processing algorithms \( A_f \), prediction algorithms \( A_p \), and calibration algorithms \( A_c \). The space of hyper-parameter configurations for a pipeline is \( \Theta = \Theta_d \times \Theta_f \times \Theta_p \times \Theta_c \), where \( \Theta_d \) being the hyper-parameter that corresponds to imputation algorithms \( A_d \), and similarly for \( \Theta_f, \Theta_p, \) and \( \Theta_c \). Thus, the space of all possible pipeline configuration considered as \( P_\Theta \), where \( P_\Theta \in \Theta \) is a selection of algorithms \( p \in P \), and hyper-parameter settings \( \Theta \).

For a given clinical data \( D \), the main objective of the AutoPrognosis framework is to find the best pipeline configuration \( P_\Theta \in \Theta \) as follows.

\[
    p_0^* = \arg \max_{p_0 \in P_\Theta} \frac{1}{K} \sum_{i=1}^{K} \left( \text{loss}(p_0; D_{\text{train}}^i, D_{\text{valid}}^i) \right)
\]

Where \( D_{\text{train}}^i \) and \( D_{\text{valid}}^i \) are train and validation splits, \( L \) be the accuracy metric (macro avg precision, recall, etc), and \( i \) denotes the fold.

5 EXPERIMENTAL SETUP & RESULTS

To reduce clinician workload, our first goal is to predict the patient’s re-admission risk automatically from the selected attributes using three models trained separately on Wound level, Episode level, and combined Wound Episode level datasets. Using the these three datasets, our second goal is to build the weeks to readmit model. We use 70:10:20 split for the train:validation:test for all our experiments.

**Dataset Details** Our patients re-admission risk dataset is an accumulation of 5 years (Jan 2015–Jun 2020) of patients’ wound care data captured by a Wound care organization. Data collection was carefully done by following the survival model conditions to ensure that we cover the “Patient Demographics details”, “Procedures”, “Medications”, and “Laboratory/Diagnosis of Wound condition”. Table 1 showcases the dataset details by wound-level attributes, episode-level attributes, and the common attributes used across both the datasets. Also, we have two target columns present in wound-level datasets such as, wound-recurrence, and patient category.

**Evaluation Metrics** We use classification metrics such as macro-average precision, recall, and F1-score to evaluate our methods. To understand how each class is performing, we choose macro averaging that gives each class equal weight to evaluate systems performance across both two-classes.

For the second task (weeks to readmit), we use the standard error metrics such as mean absolute error (MAE) and \( R^2 \)-score to measure the model performance.

5.1 Experiments on Wound-Level Data

As part of the first step, we use the 23 features to predict the Wound-Recurrence as well as patient category classification. Table 2, and 3 display the 5-fold cross-validation precision, recall, and F1-score results obtained using baseline logistic regression, LightGBM, AutoPrognosis, and Extended AutoPrognosis.

**Wound-Recurrence Classification** Table 2 reports the recurring wound classification results for the wound-level dataset. Of the four training methods, we achieved the best performance with extended Autoprognosis and worst with logistic regression. We can observe that the addition of two classifiers LightGBM and CatBoost models improves the recall and F1-score for recurring wound class. Further, we display the features that show a significant impact on improving the accuracy of the model in Table 7.

**Patients Category Classification** The target label recurring wound is available only on wound-level dataset, but the patient category column is available on both wound and episode datasets. To generalize the model, we predict the patient’s category (Re-Admitted or New Patient) using the four training methods mentioned above. Table 3 describes the wound-level patient’s category results, where we achieved the best performance with extended Autoprognosis and worst with logistic regression. Moreover, we can observe an increasing performance of recall and F1-score for both the classes when compared to the recurring wound model.

5.2 Experiments on Episode-Level Data

The episode-level (sequence of patients visits) dataset provides patients information at episode level rather than wound level. Using the episode-level dataset, our goal is to predict the patient’s category classification (Re-admit or New Patient) by considering the overall episodes information. Similar to the wound-level, we use all the four training methods to obtain the results on the episode-level dataset. Table 4 shows results of precision, recall, and F1-score obtained on the episode-level dataset consists of 39 features using the four methods. Observation from the Table 4 that Extended-AutoPrognosis yields the best performance compared to all the methods as well wound-level results and worst performance obtained from baseline logistic regression.
Table 2: Wound-Level Results: Wound recurrence accuracy comparison for Extended-AutoPrognosis method, baseline Logistic Regression, LightGBM, and AutoPrognosis method

| Class          | LR   | LightGBM | AutoPrognosis | AutoPrognosis-New |
|----------------|------|----------|---------------|-------------------|
|                | P    | R        | F1            | P             | R   | F1       |
| Recurring Wound| 0.66 | 0.28     | 0.39          | 0.77          | 0.61 | 0.67     |
| New Wound      | 0.78 | 0.95     | 0.86          | 0.86          | 0.94 | 0.90     |

Table 3: Wound-Level Results: Patient Re-Admit accuracy comparison for Extended-AutoPrognosis method, baseline Logistic Regression, LightGBM, and AutoPrognosis method

| Class            | LR   | LightGBM | AutoPrognosis | Extended-AutoPrognosis |
|------------------|------|----------|---------------|------------------------|
|                  | P    | R        | F1            | P             | R   | F1       | P             | R   | F1     |
| Re-Admit Patient | 0.90 | 0.72     | 0.80          | 0.96          | 0.77 | 0.86     | 0.91          | 0.79 | 0.86   | 0.94 | 0.81 | 0.88   |
| New Patient      | 0.77 | 0.92     | 0.84          | 0.81          | 0.97 | 0.88     | 0.85          | 0.96 | 0.88   | 0.86 | 0.98 | 0.90   |

Table 4: Episode-Level Results: Patient Category accuracy comparison for Extended-AutoPrognosis method, baseline Logistic Regression, LightGBM, and AutoPrognosis method

| Class            | LR   | LightGBM | AutoPrognosis | Extended-AutoPrognosis |
|------------------|------|----------|---------------|------------------------|
|                  | P    | R        | F1            | P             | R   | F1       | P             | R   | F1   |
| Re-Admit Patient | 0.93 | 0.69     | 0.83          | 0.93          | 0.75 | 0.83     | 0.94          | 0.76 | 0.85   | 0.90 | 0.79 | 0.86   |
| New Patient      | 0.78 | 0.97     | 0.86          | 0.83          | 0.95 | 0.88     | 0.83          | 0.97 | 0.90   | 0.84 | 0.98 | 0.91   |

Table 5: WoundEpisode-Level Results: Patient Category accuracy comparison for Extended-AutoPrognosis method, baseline Logistic Regression, LightGBM, and AutoPrognosis method

| Class            | LR   | LightGBM | AutoPrognosis | Extended-AutoPrognosis |
|------------------|------|----------|---------------|------------------------|
|                  | P    | R        | F1            | P             | R   | F1       | P             | R   | F1   |
| Re-Admit Patient | 0.89 | 0.64     | 0.74          | 0.98          | 0.87 | 0.92     | 0.92          | 0.91 | 0.92   | 0.92 | 0.92 | 0.92   |
| New Patient      | 0.87 | 0.97     | 0.92          | 0.87          | 0.98 | 0.93     | 0.90          | 0.97 | 0.94   | 0.91 | 0.98 | 0.94   |

Table 6: Weeks to Re-Admit prediction comparison for LightGBM method and the baseline Linear Regression on MAE. LR=Linear Regression.

| Feature set      | LR   | LightGBM |
|------------------|------|----------|
| Wound-Level      | 3.2  | 2.6      |
| Episode-Level    | 4.6  | 3.1      |
| WoundEpisode-Level| 3.0  | 2.3      |

5.3 Experiments on Wound Episode-Level Data
Here, we combine the wound and episode level datasets by using the common patient id across both the datasets resulted in overall 66 attributes. The common attributes across two datasets are reported in Table 1. Using the wound episode-level dataset, our final goal is to predict the patient’s category classification (Re-admit or New Patient) by considering the first episode, or last episode, or overall episodes information.

Table 5 report results of precision, recall, and F1-score obtained on the wound episode-level dataset using the four methods. Observation from Table 5 that Extended-AutoPrognosis yields the best performance compared to all the methods as well wound and episode level results and worst performance obtained from baseline logistic regression. We also observe that the wound episode-based data results were better than individual datasets when we use the Extended-AutoPrognosis model, with 0.92 recall for the Re-Admitted patient class. Overall, we observe that our Extended-AutoPrognosis models provide the best results.

5.4 Weeks to ReAdmit Prediction
For the weeks to readmit prediction, our labeled data follows a power-law distribution. To perform the weeks to readmit prediction, we use the same features as used in the patient category classification model. However, here the target variable is the weeks to readmit for a particular wound-level, episode-level, and wound episode-level dataset. The results in Table 6 illustrate the performance of the LightGBM model in comparison with the baseline linear regression model. To measure the model performance, we use mean absolute error (MAE) as the metric. The minimum number of weeks to readmit is one, the maximum is 15 weeks, and the average is 6 weeks in our dataset. The LightGBM model achieves an MAE of 2.6 weeks if we consider only wound-level features, 3.1 MAE on episode-level features, and 2.3 MAE on wound episode-level features.
Table 7: Feature Importance Analysis (Task 1: Wound-level Patient Category Risk prediction; Task 2: Episode-level Patient Category Risk prediction). Features are listed in descending order of importance

| Nr | Task 1 | Task 2 |
|----|--------|--------|
| 1  | Days in TX for Patients | Avg BMI |
| 2  | Patient Age | Avg Days in TX for Wounds |
| 3  | Days Prior to TX | Patient Age |
| 4  | Days in TX for Wounds | Patient Discharge Status |
| 5  | Wound Location | Avg Pain Level for Episode |
| 6  | Patient Discharge Status | NonCompliance Wound Visits Rate |
| 7  | Wound Type | NonCompliance Dressing Rate |
| 8  | Visits for Wound | NonCompliance OffLoad Rate |
| 9  | Avg Pain Level for Wound | NonCompliance Limb Rate |
| 10 | Wound Stage | Lower Extremity Wounds for Episode |
| 11 | NonCompliance Dressing Rate | Chronic Wounds for Episode |
| 12 | NonCompliance Wound Visits Rate | NonCompliance Diet Rate |
| 13 | NonCompliance OffLoad Rate | Wounds for Episode |
| 14 | NonCompliance Limb Rate | Edema |

5.5 Feature Importance Analysis

Table 7 display the significant impact of features across wound and episode level datasets. We observed that age and BMI are the most important patient attributes. Wound Location, Type, Stage were the best predictor for the ReAdmit/New Patient classifier, which is expected since the wound attributes intuitively correlates with the seriousness of the wound. Similarly, Non-Compliance attributes play a major role in the readmit risk prediction.

6 CONCLUSION

In this paper, we present the patient’s risk of readmission model using the extended-autoprosigosis method. To achieve the first goal, our proposed system mainly depends on three datasets such as, (i) wound-level information, (ii) episode-level information, and (iii) combination of wound and episode level information. Our second goal is to predict the time period during which a readmission event, i.e., weeks from the date of discharge, can be expected. We are the first to perform extensive experiments on a large dataset for patients’ risk of readmission on wound care data. Our experiments show that extended-autoprosigosis offers accurate readmission predictions, and therefore can be practically be deployed. This model can serve as a useful tool for care managers to have better insight into a patient’s readmission and preemptively prevent avoidable readmissions.

REFERENCES

[1] Ahmed Alaa and Mihaela Schaar. 2018. AutoPrognosis: Automated Clinical Prognostic Modeling via Bayesian Optimization with Structured Kernel Learning. In International Conference on Machine Learning. 139–148.

[2] Charles A Basile, Christine VanZandbergen, Gordon Tait, Asaf Hanish, Brian Leas, Benjamin French, C William Hanson, Maryam Behta, and Craig A Umscheid. 2013. The readmission risk flag: Using the electronic health record to automatically identify patients at risk for 30-day readmission. Journal of hospital medicine 8, 12 (2013), 689–695.

[3] Ineci M Baytas, Cao Xiao, Xi Zhang, Fei Wang, Andi K Jain, and Jiayu Zhou. 2017. Patient subtyping via time-aware lstm networks. In Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining. 65–74.

[4] TD Briefing. 2014. Ahrq: The conditions that cause the most readmissions. The Daily Briefing (2014).

[5] Chao Che, Cao Xiao, Jian Liang, Bo Jin, Jiayu Zho, and Fei Wang. 2017. An rnn architecture with dynamic temporal matching for personalized predictions of Parkinson’s disease. In Proceedings of the 2017 SIAM International Conference on Data Mining. SIAM. 198–206.

[6] Avishik Choudhury and Christopher M Greene. 2018. Evaluating patient readmission risk: a predictive analytics approach. arXiv preprint arXiv:1812.11028 (2018).

[7] William W Dzwierzynski, Kimberly Spitz, Arthur Hartz, Clare Guse, and David L Larson. 1998. Improvement in resource utilization after development of a clinical pathway for patients with pressure ulcers. Plastic and reconstructive surgery 102, 6 (1998), 2006–2011.

[8] Elizabeth Eby, Christine Hardwick, Maria Yu, Steve Gelwicks, Ketra Deschamps, Jen Xie, and Tom George. 2015. Predictors of 30 day hospital readmission in patients with type 2 diabetes: a retrospective, case–control, database study. Current medical research and opinion 31, 1 (2015), 107–114.

[9] Matthias Feurer, Aaron Klein, Katharina Eggensperger, Jost Springenberg, Manuel Blum, and Frank Hutter. 2015. Efficient and robust automated machine learning. In Advances in neural information processing systems. 2962–2970.

[10] Matthias Feurer, Aaron Klein, Katharina Eggensperger, Jost Tobias Springenberg, Manuel Blum, and Frank Hutter. 2019. Auto-sklearn: efficient and robust automated machine learning. In Automated Machine Learning. Springer, Cham, 113–134.

[11] Joseph Futoma, Jonathan Morris, and Joseph Lucas. 2015. A comparison of models for predicting early hospital readmissions. Journal of biomedical informatics 56 (2015), 229–238.

[12] Jared L Katzman, Uri Shaham, Alexander Cloninger, Jonathan Bates, Tingting Jiang, and Yuval Kluger. 2018. DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. BMC medical research methodology 18, 1 (2018), 24.

[13] Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. 2017. Lightgbm: A highly efficient gradient boosting decision tree. In Advances in neural information processing systems. 3146–3154.

[14] Qijun Kripalani, Cecelia N Theohal, Beth Anctil, and Eduard E Vanlambhikelis. 2014. Reducing hospital readmission rates: current strategies and future directions. Annual review of medicine 65 (2014), 471–485.

[15] Aaron L Leppin, Michael R Gionfriddo, Maya Kessler, Juan Pablo Brito, Frances S Mair, Katie Gallacher, Zhou Wang, Patricia J Erwin, Tanya Sylvestre, Kasey Boehner, et al. 2014. Preventing 30-day hospital readmissions: a systematic review and meta-analysis of randomized trials. JAMA internal medicine 174, 7 (2014), 1095–1107.

[16] Samuel R Nussbaum, Marissa J Carter, Caroline E Fife, Joaquin DaVanzo, Randall Haught, Marcia Nusgart, and Donna Cartwright. 2018. An economic evaluation of the impact, cost, and medicare policy implications of chronic nonhealing wounds. Value in Health 21, 1 (2018), 27–32.

[17] Liudmila Prokhorenkova, Gleb Gusev, Aleksandr Vorobev, Anna Veronika Dorogus, and Andrey Gulin. 2018. CatBoost: unbiased boosting with categorical features. In Advances in neural information processing systems. 6638–6648.

[18] Joseph S Ross, Gregory K Mulvey, Brett Staufer, Vishnu Patlolla, Susannah M Berrheim, Patricia S Keenan, and Harlan M Krumholz. 2008. Statistical models and patient predictors of readmission for heart failure: a systematic review. Archives of internal medicine 168, 13 (2008), 1371–1386.

[19] Marc D Silverstein, Huanying Qin, S Quay Mercer, Jaclyn Fong, and Ziad Haydar. 2015. Prognostic Modeling via Bayesian Optimization with Structured Kernel Learning. In Advances in Neural Information Processing Systems. 6585–6594.

[20] Daniela Urma and Chi-Cheng Huang. 2017. Interventions and strategies to reduce 30-day readmission rates. Hospital Medicine Clinics 6, 2 (2017), 216–228.

[21] Binchen Zheng, Jinghe Zhang, Song Won Yoon, Sarah S Lam, Mohammad Khamseh, and Srikantam Poranika. 2015. Predictive modeling of hospital readmissions using metaheuristics and data mining. Expert Systems with Applications 42, 20 (2015), 7110–7120.