Predicting Unplanned Hospital Readmissions using Patient Level Data

Mahesh Balan U
Indian Institute of Technology
Madras
maheshbalan@alumni.iitm.ac.in

Meet Gandhi
Indian Institute of Technology
Madras
meetgandhi@alumni.iitm.ac.in

Swaminathan Rammohan
Indian Institute of Technology
Madras
rammohan.swaminathan@gmail.com

Abstract

The rate of unplanned hospital readmissions in the US is likely to face a steady rise after 2020. Hence, this issue has received considerable critical attention with the policy makers. Majority of hospitals in the US pay millions of dollars as penalty for readmitting patients within 30 days due to strict norms imposed by the Hospital Readmission Reduction Program. In this study, we develop two novel models: PURE (Predicting Unplanned Readmissions using Embeddings) and Hybrid DeepR, which uses the historical medical events of patients to predict readmissions within 30 days. Both these models are hybrid sequence models that leverage both sequential events (history of events) and static features (like gender, blood pressure) of the patients to mine patterns in the data. Our results are promising, and they benchmark previous results in predicting hospital readmissions. The contributions of this study add to existing literature on healthcare analytics.

1. Introduction

The digitalization of global healthcare and availability of massive datasets like Electronic Health Records (EHR) fueled by increase in computational capabilities has redefined treatment procedures in critical care systems. The United States alone has generated 150 exabytes of healthcare data, with 48% annual growth [24]. Among the various data streams that are redefining the healthcare ecosystem, patient level data holds significant importance [14]. The use of patient level data to predict the early onset of a medical event has been demonstrated to accelerate treatment procedures for patients with cardiac failure [6], sepsis [13] and alzheimer's [29]. Generally, patient level data is sequential in nature with a chronological listing of medical events and treatment patterns associated with a patient. Moreover, most of the prior literature focusing on patient event prediction uses sequence models like

Recurrent Neural Networks (RNNs), Long Short-Term Memory Networks (LSTMs) and Convolutional Neural Networks (CNNs) with applications in different diseases [22],[6]-[7],[26],[4],[18]. Prior research leverage patient level data to predict varied outcomes like mortality, discharge time-periods and future medical events of a patient. But there has been limited focus on other potential applications of patient level data which could make a huge impact on the patients, hospitals and the government. This study examines one such application – unplanned hospital readmissions, which plays a critical role to patients, hospitals and the government towards designing and monitoring better healthcare regulations.

The availability of patient level data (PLD) has opened opportunities for innovative treatment patterns in medicine and healthcare. The primary source of data remains Electronic Health Records (EHR) and claims data. Moreover, nearly 96% of the hospitals in the US have traded paper records for online portals, and this correlates across the world. The claims data is a source of chronological listing of prescriptions (Rx), diagnoses (Dx) and procedures (Px) encountered by a patient. These details can be decoded using the International Classification of Diseases (ICD) codes created by the World Health Organization (WHO). The sequential occurrence of different types of events (Rx/Dx/Px) coupled with other records of the patient (from EHR) will act as a source of rich information that can be leveraged for treatment pathways and treatment patterns modeling. Prior research has shown evidence of identifying diseases at earlier stages in oncology [23] and dermatology [14], thus increasing the likelihood of successful treatment.

Although there has been more focus recently on modeling using patient level data, its application on a very critical use-case of predicting unplanned hospital readmission is not widely explored. Many hospitals in

HİCSS

the US have been penalized more than $270 million dollars for re-admitting patients within 30 days as a part of the Hospital Readmission Reduction Program. Also, the penalty is increased to 3% of the hospital's overall Medicare reimbursement [15]. More than 76% of the planned or unplanned hospital readmission are potentially avoidable [11]. Although current studies report numbers in the US, this remains as a fundamental problem in healthcare across the world. The traditional sequence modeling networks have shown not benchmarking performance in modeling patient level data. Given the complexity of patient level data, there comes a need to develop a more hybrid models which leverage both static (eg: age, charlson comorbidity index) and transactional (eg: historical medical events) features of the patient in order to predict the future medical pathway of the patient. In this study, we address this gap by developing two novel hybrid sequential models: Hybrid DeepR and PURE (Predicting Unplanned Readmissions using Embeddings), that leverages both static and transactional features to predict unplanned hospital readmissions. Our Hybrid DeepR model is an extended adaptation of [26], which is a convolutional neural network (CNN) that also leverages static features in addition to transactional features. The other model PURE is a bidirectional LSTM that uses both static and transactional features. We benchmark prior results by demonstrating an application of our model using patient level data in MIMIC III clinical critical care data [20].

2. Background

2.1. Second-order headings

Prior research using patient level data can be classified into 5 buckets – a) disease detection/classification b) sequential prediction of clinical events c) concept embedding (feature representation) d) EHR data, and e) EHR data privacy (e.g.: de-identification). However, most of them are inclined towards either disease classification or prediction of clinical events [34]. Also, these models leverage open source data that is available like Amyotrophic Lateral Sclerosis (ALS) clinical trials data [2], Parkinson's Progression Markers Initiative data [1] and MIMIC III intensive clinical care data [20]. In addition to the multimodal clinical data, researchers have also used images [10], [14], text [32] and multivariate timeseries data [32] for disease classification and predicting future clinical events. A multi-label clinical event predictor called DoctorAI using recurrent neural networks was delevloped by [6]. DoctorAI was not disease specific and it is one of the early works towards developing a generic predictive algorithm. DeepR is another such algorithm which

builds a feature representation on medical records and predict unplanned hospital readmissions [26]. As generic models might miss out on complex medical diagnostics which are challenging even for doctors to diagnose them and disease specific models have also been developed in recent studies. Many studies have demonstrated promising results in complex diagnostics spanning pediatrics [7], cardiology [8], dermatology [14], radiology [10], ophthalmology [12], and pathology [3].

There has also been more focus on predicting outcomes not related with future medical event like hospital re-admissions [16], clinical intervention [33], mortality [28] and discharge time [28]. Deep learning models have shown better results with patient level data in comparison with traditional machine learning models. Three different models were experimented– recurrent neural networks, attention-based time-aware neural networks, and a neural network with boosted time-based decision stumps [16]. They found that neural networks benchmarked good performance than all the traditional machine learning models. Similarly, an interpretable model based on convolution plus attention model architecture showed better performance in classifying diagnosis codes from clinical notes [25].

2.2. Hospital Readmissions

Researchers and practitioners in the field of medicine are increasingly focusing in reducing hospital readmissions. "An admission within 30 days after discharge from an 'index' hospitalization" is defined as hospital readmission [16]. A readmission can happen at the same venue or at a different hospital and it could be a planned or an unplanned one. Strict penalties have been enforced on hospitals through the Patient Protection and Affordable Care Act for lessening the readmission rates [30]. Therefore, hospitals have responsibility of focusing only on the high-risk patients for readmissions and carefully avert their readmission rates in an efficient way. The readmission penalties imposed globally by the Centers for Medicare & Medicaid Services (CMS) has created a necessity for identifying the patients with high risk of getting a readmission with 30 days at the time of discharge [9]. A study conducted by Medicare Payment advisory Committee revealed that more than 75% of readmission cases could have been potentially avoidable. In 2014, the CMS imposed a minimum penalty of 3% to hospitals Medicare reimbursements due to readmissions and with more stringent conditions [9].

A wide array of statistical models is available to assess patient readmission risk in the literature with a variety of existing data. The characteristics of this data include

lab test reports, patient conditions, procedures, medications and demographic conditions etc. [65]. The qualities of good prediction models must possess high predictive ability, clinical data, good performance metrics for model assessment, validate in the population to be applied, must be generalizable. Nevertheless, most of the hospital readmission risk prediction models performed poorly and hence there is a need to improve the model performance [9]. Most of the risk readmission models utilize easily interpretable models such as logistic regression. As the type of data collected with respect to volume and variety are high, there is a need for more complex models for achieving good performance metrics.

2.3. CNN for Sequence Learning

Some of the greatest successes that deep learning has had is in computer vision [31], a field of study that focuses on image and video understanding, which enables medical applications such as determining whether a patient's radiograph contains malignant tumors. Convolutional neural networks (CNNs) [21],[31], a deep-learning algorithm designed to process data that exhibits natural spatial invariance has grown to be the primary focus of this field [14]. Medical imaging has greatly benefited from the applications of CNNs spanning dermatology [13], radiology [10], ophthalmology [12] and pathology [3]. Deep-learning systems could aid physicians by offering second opinions and flagging concerning areas in images.

In addition to success in computer vision applications, CNNs also has shown promising results in sequence learning applications. Prior studies have shown that CNNs have outperformed traditional sequence learning models like RNNs and LSTMs [17], [5]. In terms of computational complexity, the computations involved in CNN are easily parallelizable in contrast to those involved in RNN which are mostly sequential, leading to a performance bottleneck. It was found that "unlike RNN, the native CNN lacks the history sensitivity required for sequence transformation; therefore, enhancing the sequential order awareness, or position-sensitivity, becomes the key to make CNN the general deep learning model". Moreover, PosNet show early evidence for successful applications of CNN for sequence learning [5]. In the same year a team from Facebook AI Research built a CNN architecture which benchmarked LSTM performance in the WMT'14 English-German and WMT'14 English-French translation at an order of magnitude faster speed, both on GPU and CPU [17]. Since then, several studies have attempted using CNNs for sequence learning applications including medical event predictions like Doctor AI [6] and DeepR [26].

2.4. Patient Level Data (PLD)

Patient level data captures several years of detailed patient-physician interactions for millions of patients. While traditional data sources provide information on physician activity, they do not allow the user to probe into the treatment regimens used for individual patients. Moreover, they offer only a limited ability to track patient behavior. Although patient identity is not provided, to ensure HIPAA compliance, a patient-specific ID allows individual patients to be tracked over time and across care settings. Anonymized patient level data (APLD), on the other hand, provides insight not only into the patient journey but also into patient behavior. This data also offers insights into the physician's prescribing behavior and treatment effectiveness. Patient-level data captures encounters of the individual patients with the healthcare system. This includes prescriptions, diagnoses, procedures, physician visits, hospitalizations, lab tests, etc.

3. Hybrid models to predict unplanned hospital readmissions

3.1. Hybrid DeepR

We build two hybrid models that use both sequential data and static data to predict future unplanned hospital readmissions at the patient level. We extended the DeepR [26] to develop Hybrid DeepR which is a CNN that learns from a sequence of events at patient level, where each event is represented by embedding vectors learnt from another skip-gram model.

The skip-gram model learns event embeddings by assessing the likelihood of any given target event c happening with a context event t, the probability $P(t|c)$ is given by:

$$P(t|c) = \frac{\exp{(\theta_t^T e_c)}}{\sum_{j=1}^{|V|} \exp{(\theta_j^T e_c)}}$$

For negative sampling, the models being trained on sets of k negative examples and 1 positive example. Given a context event c and a target event t, the prediction is expressed by:

$$P(y = 1|c, t) = \sigma(\theta_t^T e_c)$$

We adapt DeepR to handle both static and sequential data. At first a skip-gram model with negative sampling is trained to learn embeddings for every event based on the patient's history of medical events. All the unique set of events in our universe form the corpus and then an embedding vector of dimension 32 (fixed after experiments) is learnt from the corpus. These embeddings are then sequenced to create a patient level

embeddings based on the patient's past sequence of events, which is further an input to a CNN that takes static features with sigmoid activation function in the last layer to output the final predictions. The network is optimized to minimize the cross-entropy loss function.

3.2. PURE (Predicting Unplanned Readmissions using Embeddings)

Like Hybrid DeepR, the PURE model also leverages event level embeddings learnt using a skip-gram model with negative sampling. These embedding vectors (dimension: 32) are sequenced to create patient level embeddings based on the patient's past historical sequence of events. These sequences of events are input to a bidirectional LSTM which takes the static features of the patient in the last layer with sigmoid activation to finally make the predictions. Bidirectional LSTMs are very similar to LSTMs just that they train two networks with actual and reversed sequences. Following are the gates used in a typical LSTM network –

$$i_t = \sigma(w_i[h_{t-1}, x_t] + b_i)$$
$$f_t = \sigma(w_f[h_{t-1}, x_t] + b_f)$$
$$o_t = \sigma(w_o[h_{t-1}, x_t] + b_o)$$

Where,

$i_t \rightarrow represents\ input\ gate$
$f_t \rightarrow represents\ forget\ gate$
$o_t \rightarrow represents\ output\ gate$
$\sigma \rightarrow represents\ sigmoid\ function$
$w_x \rightarrow weight\ of\ respective\ gate\ (x)\ neurons$
h_{t-1}
$\rightarrow output\ of\ the\ previous\ lstm\ block\ (at\ timestamp\ t - 1)$
$x_t \rightarrow input\ at\ current\ timestamp$
$b_x \rightarrow biases\ for\ the\ respective\ gates$

3.3. Experimental Setup

The MIMIC-III dataset consists of data pertaining to 58,976 hospital admission stays that correspond to 46,520 patients. In the dataset, 7537 (16.8%) of the patients had more than one hospital admission. This statistic also confirms with prior literature stating about 17% of the patients are subject to hospital readmissions [16]. The average readmissions per patient was 1.26. Since MIMIC-III data specifically corresponds to critical care [19], all the readmissions are unplanned hospital readmissions. According to the Hospital

Readmission Reduction Program, hospitals are subject to penalties if they have unplanned patient readmissions within 30 days of discharge. To be consistent with reality, we consider patient readmission prediction in the immediate future of 30 days. Finally, after data processing, we had about 3.5% of the patients having readmissions within a 30-day period. We also ignored patients for whom historical medical information was not available. This led to a final data set of 45,980 patients with 1544 patients having readmissions within a 30-day period. More details provided in Table 3.

All these prior works discussed above focuses on utilizing the EHR entered components of the MIMIC-III data like hourly recording of various parameters of the patient. They use this sequential data to look for patterns leading to onset of sepsis. But we hypothesize that tracking the medical history of a patient with respect to events like prescriptions (Rx), diagnoses (Dx) and medical procedures (Px) and hence there is a better understanding to predict the future event or condition of a patient. In the past framing, there is a lag time of collecting these parameters post admission of the patient to predict the future pathway. But if we could leverage the patient level data in terms of her or his medical history, we would be having the advantage of predicting the future pathway much earlier, providing lead time for hospital planning.

The dataset was randomly split into training, validation and test. 85% of the dataset was used for training and 15% for test. Out of the 85%, 20% was used for validation purpose. The results reported in the next section are based on the performance in the test dataset. We used logistic regression as a reference model to compare, in addition to numbers reported in prior literature. The problem was framed as a binary classification, predicting whether a patient will be readmitted within the next 30 days.

Hyperparameter tuning was done for both Hybrid DeepR (CNN) and PURE (Bidirectional LSTM). For CNN, we experimented with different word2vec embedding dimension (32,64,128) given as input for medical events and the journey length of the historical medical events (50,85,120). The journey length was experimented with μ, μ+σ and μ-σ; where μ is average journey length (85) and σ is standard deviation (35). More details are provided in Tables 1 and 2.

Figure 1 Experimental Setup
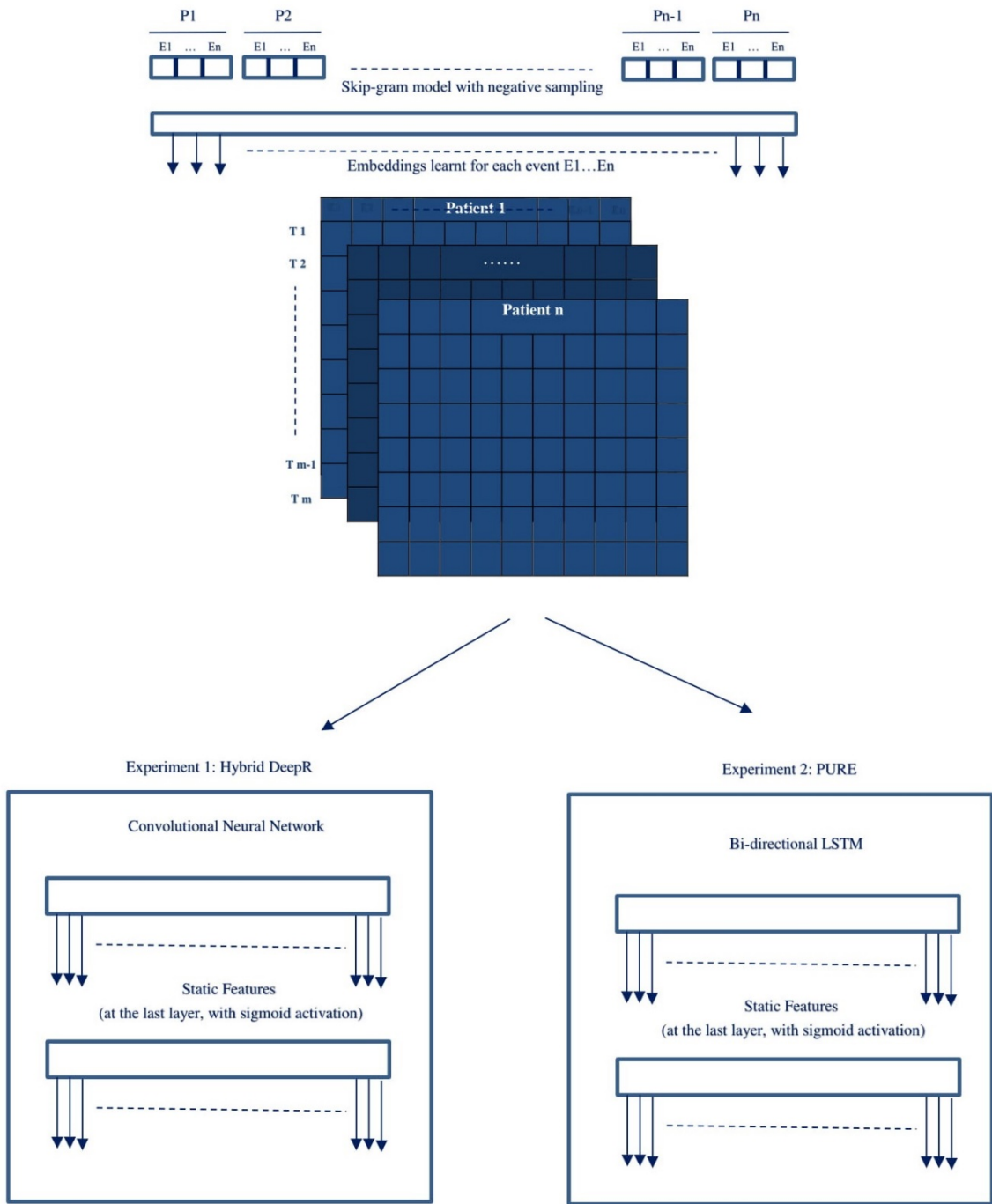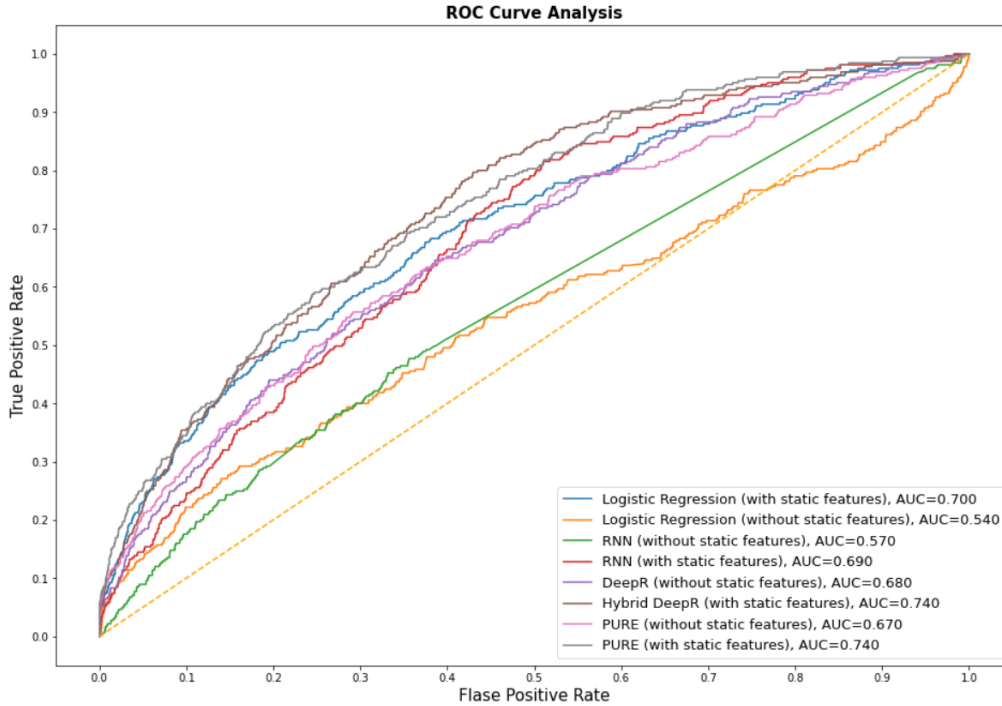
Figure 2 AUROC of all the models

Table 1 Hyperparameter tuning for Hybrid DeepR

w2v_dim	journey_len	Test AUC for different window_size (ws)		
		ws=3	ws=5	ws=6
32	50	0.661	0.653	0.657
32	85	0.656	0.669	0.657
32	120	0.652	0.638	0.658
64	50	0.648	0.639	0.634
64	85	0.649	0.664	0.657
64	120	0.660	0.649	0.669
128	50	0.643	0.632	0.645
128	85	0.668	0.656	0.657
128	120	0.668	0.651	0.677

Table 2 Hyperparameter tuning for PURE

w2v_dim	journey_len	Test AUC
32	50	0.645
32	85	0.656
32	120	0.655
64	50	0.653
64	85	0.657
64	120	0.674
128	50	0.623
128	85	0.650
128	120	0.653

Table 3 Patient data details (static features)

Static Features	Description
Gender	Gender of the patient (Male/Female)
Age	Calculated from Date of Birth (DOB)
Duration of Stay	The duration of stay in number of days
Admission Type	4 levels: Elective, Urgent, Newborn or Emergency
Admission Location	Previous location of patient before admission
Discharge Location	Type of hospital where discharge happened
Insurance	Insurance Type: Public, Private, Self-pay, Medicare
Language	Native language
Religion	Religion of the patient
Marital Status	Marital status of the patient
Ethnicity	Ethnic group

Table 4 Results of different models on predicting readmission

Models	AUROC		
	Train (#31266)	Val (#5518)	Test (#9196)
Logistic Regression (without static features)	0.61	0.57	0.55
Logistic Regression (with static features)	0.79	0.64	0.70
RNN (without static features)	58.3	53.8	56.5
RNN (with static features)	76.2	68.9	68.9
DeepR (without static features)	0.78	0.65	0.68
Hybrid DeepR	0.93	0.72	0.74
PURE (without static features)	0.82	0.66	0.67
PURE (with static features)	0.87	0.74	0.74

3.4. Experimental Results

The results from both the models Hybrid DeepR and PURE benchmarked the AUROC numbers reported by prior literature and that of logistic regression that we used in our dataset. The numbers used for benchmarking are based on the predictions on test dataset. The train and validation numbers are reported in Table 1. The Hybrid DeepR (with static features) reported an AUROCC of 0.727 and PURE (with static features) reported an AUROC of 0.733. These numbers benchmarked results of logistic regression (AUROC: 0.54) and those reported in prior literature (AUROC: 0.68; [16]). We also conducted more experiments to evaluate the lift in model performance by removing static features. Modeling details are visually represented in Figure 1 and consolidated results are reported in Figure 2 and Table 1.

In addition to better performance, our models also have the following advantages –

i.   The medical events history that includes Rx, Dx and Px can also be used to predict future outcomes, along with readmission. This will enable several decision support systems in hospital planning.
ii. Since the models leverage prior historical sequence of events, there is considerable lead time in

predicting much in advance as these details are available in plenty during the stay of the patient in the hospital. Other studies use EHR data which includes various parameters being recorded at hourly intervals during the stay in the hospital planning.

4. Discussion and Implications

We find that PURE and Hybrid DeepR have a significantly high AUROC scores (0.74) for predicting unplanned hospital readmissions using MIMIC-III dataset. The results benchmark numbers from Logistic Regression, prior research including DeepR which doesn't include static features [26], and the work of [16]. Moreover, the lead time with which PatImg-2D predicts hospital readmission is very high compared to other studies. Following are the key advantages of PURE and Hybrid DeepR in addition to improved AUROC –

a.   The flexibility to stack any type of medical events makes these hybrid models more flexible to even limited information available for a patient.
b.   The static features are available for most of the patients and these act as covariates in predicting the outcome, which makes the predictions more stable and foolproof. As seen from Table 1, the static features provide a huge lift in model performance compared in experiments.
c.   The model is embedding based and is very generic. Hence, this could be extended for predicting any future patient pathway modeling.

Predicting unplanned hospital readmissions with greater accuracy has potential clinical impact. First, it helps in better patient queue management. Second, hospitals avoid paying penalties (in millions of dollars) for unplanned patient readmission which could be avoided, particularly in countries like the United States which has strict patient readmission policies. Third, it also helps in effective patient treatment planning by taking a deep dive into the patient trajectories (historical medical events) for patients with higher probability of future readmissions.

5. Discussion and Implications

In this study, we hypothesized that hybrid models (static + transactional features) could enable better predictive power when modeling with patient level data. We present PURE and Hybrid DeepR which leverage patient level data along with event embeddings and

static features to predict hospital readmissions within 30 days using the MIMIC-III dataset. Our results are promising compared to reference models and prior research (AUROC: 0.74). This supports our hypothesis that hybrid models perform better when modeling with patient level data. The hospital readmission problem is a very important issue faced by major hospital across the world and this study has huge implications towards better regulation by hospital and government towards patient demand planning. Our paper contributes to existing literature on advancements of deep learning for healthcare and medicine.

However, this paper does have its limitations and leaves huge scope for future research. The paper builds on the sequence numbers of the event as proxy for the exact time-step. This is because of a constraint of the MIMIC-III dataset which doesn't have exact time-step of the occurrence of every medical event (Rx, Dx, Px). Another future direction for research is to extend these embeddings to multiple dimensions like taking weighted product of embeddings and beyond. Although this comes at huge computational cost, we believe it could demonstrate even better predictive performance. We hope our findings open up huge scope for future research in this area.

6. References

[1]   I. M. Baytas, C. Xiao, X. Zhang, F. Wang, A. K. Jain, and J. Zhou, "Patient subtyping via time-aware LSTM networks," in *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*, 2017, pp. 65-74.

[2]   B. K. Beaulieu-Jones and C. S. Greene, "Semi-supervised learning of the electronic health record for phenotype stratification," *Journal of biomedical informatics,* vol. 64, pp. 168-178, 2016.

[3]   P. Charoentong *et al.*, "Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade," *Cell reports,* vol. 18, no. 1, pp. 248-262, 2017.

[4]   Z. Che, S. Purushotham, K. Cho, D. Sontag, and Y. Liu, "Recurrent neural networks for multivariate time series with missing values," *Scientific reports,* vol. 8, no. 1, pp. 1-12, 2018.

[5]   Q. Chen and R. Wu, "CNN is all you need," *arXiv preprint arXiv:1712.09662,* 2017.

[6]   E. Choi, M. T. Bahadori, A. Schuetz, W. F. Stewart, and J. Sun, "Doctor ai: Predicting clinical events via recurrent neural networks," in *Machine Learning for Healthcare Conference*, 2016, pp. 301-318.

[7]   E. Choi *et al.*, "Multi-layer representation learning for medical concepts," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 1495-1504.

[8]   E. Choi, A. Schuetz, W. F. Stewart, and J. Sun, "Using recurrent neural network models for early detection of heart failure onset," *Journal of the American Medical Informatics Association,* vol. 24, no. 2, pp. 361-370, 2017.

[9]   S. A. Choudhry, J. Li, D. Davis, C. Erdmann, R. Sikka, and B. Sutariya, "A public-private partnership develops and externally validates a 30-day hospital readmission risk prediction model," *Online journal of public health informatics,* vol. 5, no. 2, p. 219, 2013.

[10]   M. Cicero *et al.*, "Training and validating a deep convolutional neural network for computer-aided detection and classification of abnormalities on frontal chest radiographs," *Investigative radiology,* vol. 52, no. 5, pp. 281-287, 2017.

[11]   M. P. A. Commission, "Report to the Congress: Promoting Greater

Efficiency in Medicare," 2007, doi: http://www.medpac.gov/docs/default-source/reports/Jun07_EntireReport.pdf.

[12]   J. De Fauw *et al.*, "Clinically applicable deep learning for diagnosis and referral in retinal disease," *Nature medicine,* vol. 24, no. 9, pp. 1342-1350, 2018.

[13]   T. Desautels *et al.*, "Prediction of sepsis in the intensive care unit with minimal electronic health record data: a machine learning approach," *JMIR medical informatics,* vol. 4, no. 3, p. e28, 2016.

[14]   A. Esteva *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature,* vol. 542, no. 7639, pp. 115-118, 2017.

[15]   A. Esteva *et al.*, "A guide to deep learning in healthcare," *Nature medicine,* vol. 25, no. 1, pp. 24-29, 2019.

[16]   J. Futoma, J. Morris, and J. Lucas, "A comparison of models for predicting early hospital readmissions," *Journal of biomedical informatics,* vol. 56, pp. 229-238, 2015.

[17]   J. Gehring, M. Auli, D. Grangier, D. Yarats, and Y. N. Dauphin, "Convolutional sequence to sequence learning," in *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, 2017: JMLR. org, pp. 1243-1252.

[18]   S. Ghanvatkar and V. Rajan, "Deep Recurrent Neural Networks for Mortality Prediction in Intensive Care using Clinical Time Series at Multiple Resolutions," 2019.

[19]   A. E. Johnson *et al.*, "A comparative analysis of sepsis identification methods in an electronic database," *Critical care medicine,* vol. 46, no. 4, p. 494, 2018.

[20]   A. E. Johnson *et al.*, "MIMIC-III, a freely accessible critical care database," *Scientific data,* vol. 3, p. 160035, 2016.

[21]   Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature,* vol. 521, no. 7553, pp. 436-444, 2015.

[22]   V. Liu, P. Kipnis, M. K. Gould, and G. J. Escobar, "Length of stay predictions: improvements through the use of automated laboratory and comorbidity variables," *Medical care,* pp. 739-744, 2010.

[23]   A. S. Marr, C. Zhang, and A. K. Ganti, "Resected small cell lung cancer—time for more?," *Journal of Thoracic Disease,* vol. 8, no. 8, pp. E755-E757, 2016. [Online].   Available: http://jtd.amegroups.com/article/view/8584.

[24]   S. Medicine, "Harnessing the Power of Data in Health,"   2017,   doi: http://med.stanford.edu/content/dam/sm/sm-news/documents/StanfordMedicineHealthTrendsWhite Paper2017.pdf.

[25]   J. Mullenbach, S. Wiegreffe, J. Duke, J. Sun, and J. Eisenstein, "Explainable prediction of medical codes from clinical text," *arXiv preprint arXiv:1802.05695,* 2018.

[26]   P. Nguyen, T. Tran, N. Wickramasinghe, and S. Venkatesh, "$\mathtt {Deepr} $: a convolutional net for medical records," *IEEE journal of biomedical and health informatics,* vol. 21, no. 1, pp. 22-30, 2016.

[27]   S. A. Pendergrass and D. C. Crawford, "Using electronic health records to generate phenotypes for research," *Current protocols in human genetics,* vol. 100, no. 1, p. e80, 2019.

[28]   S. Purushotham, C. Meng, Z. Che, and Y. Liu, "Benchmark of deep learning models on large healthcare mimic datasets," *arXiv preprint arXiv:1710.08531,* 2017.

[29]   L. Qiu, V. Rajan, and B. Tan, "Battling Alzheimer's Disease through Early Detection: A Deep Multimodal Learning Approach," 2019.

[30]   A. Rajkomar *et al.*, "Scalable and accurate deep learning with electronic health records," *NPJ Digital Medicine,* vol. 1, no. 1, p. 18, 2018.

[31]   O. Russakovsky *et al.*, "Imagenet large scale visual recognition challenge," *International journal of computer vision,* vol. 115, no. 3, pp. 211-252, 2015.

[32]   B. Shickel, P. J. Tighe, A. Bihorac, and P. Rashidi, "Deep EHR: a survey of recent advances in deep learning techniques for electronic health record (EHR) analysis," *IEEE journal of biomedical and health informatics,* vol. 22, no. 5, pp. 1589-1604, 2017.

[33]   H. Suresh, N. Hunt, A. Johnson, L. A. Celi, P. Szolovits, and M. Ghassemi, "Clinical intervention prediction and understanding with deep neural networks," in *Machine Learning for Healthcare Conference*, 2017, pp. 322-337.

[34]   C. Xiao, E. Choi, and J. Sun, "Opportunities and challenges in developing deep learning models using electronic health records data: a systematic review," *Journal of the American Medical Informatics Association,* vol. 25, no. 10, pp. 1419-1428, 2018.

[35]   F. Yu and H. H. Ip, "Semantic content analysis and annotation of histological images," *Computers in Biology and Medicine,* vol. 38, no. 6, pp. 635-649, 2008.