Healthcare Analytics 2 (2022) 100050

Contents lists available at ScienceDirect

Healthcare Analytics

journal homepage: www.elsevier.com/locate/health

Predicting emergency department utilization among children with asthma using deep learning models

Rawan AlSaad [a], Qutaibah Malluhi [a], Ibrahim Janahi [b,*], Sabri Boughorbel [c]

[a] College of Engineering, Qatar University, Doha, Qatar
[b] Department of Pediatric Pulmonology, Sidra Medicine, Doha, Qatar
[c] Qatar Computing Research Institute, Hamad Bin Khalifa University, Doha, Qatar

ARTICLE INFO

Keywords:
Deep learning
Predictive models
Electronic health record
Emergency medicine
Asthma

ABSTRACT

Pediatric asthma is a leading cause of emergency department (ED) utilization, which is expensive and often preventable. Therefore, development of ED utilization predictive models that can accurately predict patients at high-risk of frequent ED use and subsequently steering their treatment pathway towards more personalized interventions, has high clinical utility. In this paper, we investigate the extent to which deep learning models, specifically recurrent neural networks (RNNs), coupled with routinely collected electronic health record (EHR) clinical data can predict the frequency of emergency department utilization among children with asthma.

We use retrospective longitudinal EHR data of 87,413 children with asthma aged 0–18 years, who were attributed to one or more healthcare facility for at least 2 consecutive years between 2000–2013. The models were trained for the task of predicting the frequency of emergency department visits in the next 12 months. We compared prediction results of three recurrent neural network (RNN) models: bidirectional long short-term memory (BiLSTM), bidirectional gated recurrent unit (BiGRU), and reverse time attention model (RETAIN), to a baseline multinomial logistic regression model. We assessed the predictive accuracy of the models using receiver operating characteristic curve (AUC–ROC), precision–recall curve (AUC-PR), and F1-score.

The results indicated that all RNN models have similar performances reaching AUC–ROC: 0.85, AUC-PR: 0.74, and F1-score: 0.61, compared to AUC–ROC: 0.81, AUC-PR: 0.69, and F1-score: 0.56 for a baseline multinomial logistic regression.

Predictive models created from large routinely available EHR data using RNN models can accurately identify children with asthma at high-risk of repeated ED visits, without interacting with the patient or collecting information beyond the patient's EHR.

1. Introduction

Emergency departments (EDs) provide medical care for critically and acutely ill patients. Although frequent ED users only represent 1–8% of patients, they account for a substantial portion (18–28%) of all ED visits [1,2]. Therefore, patients at high-risk of repeated ED use have been extensively studied in emergency medicine literature in an effort to reduce ED crowding and costs [3–5]. However, there is currently no standardized definition for frequent ED users [6], and various cut-offs for the number of annual visits to distinguish between low and frequent ED users have been used in previous studies. These thresholds varied largely from as few as 3 to 12 or more annual visits, without a clear rationale for the choice of the selected cut-offs [1,3,6–9]. While there is a large body of literature focusing on identifying and characterizing current frequent ED users [1,7,10,11] for potential future intervention, this approach might not be efficient as a majority of frequent ED users in a given year will not continue to be frequent users in the following year [12].

The unique aspect of emergency department practices makes this field well suited to benefit from the application of machine learning techniques [13]. Patients are assessed in the ED with limited information, and physicians often find themselves balancing probabilities for risk stratification and decision-making for high-acuity patients. Therefor, the increased speed and accuracy that could be provided by machine learning techniques are especially attractive in the context of emergency medicine [14,15]. This is primarily because early detection and prediction of diseases in ED can help treat diseases more effectively and prevent unnecessary complications.

Additionally, the number of relevant studies published in the last five years shows a rapid growth of interest in machine and deep learning-based methods for emergency medicine. A recent review [16]

* Corresponding author.
E-mail addresses: 200552799@qu.edu.qa (R. AlSaad), qmalluhi@qu.edu.qa (Q. Malluhi), ijanahi@sidra.org (I. Janahi), sboughorbel@hbku.edu.qa (S. Boughorbel).

https://doi.org/10.1016/j.health.2022.100050
Received 3 October 2021; Received in revised form 28 February 2022; Accepted 2 April 2022

on artificial intelligence in emergency medicine reported that machine learning-based methods in the ED are heterogeneous in both purpose and design. For example, several studies have used supervised machine learning methods for the prediction of patient outcomes and detection of diseases at the emergency department. This includes prediction of return visits [17], prediction of hospital admission [18], prediction of pediatric asthma exacerbation [17], and stroke detection [19]. Natural language processing (NLP) models were also used in the prediction of patient disposition using triage notes [20], classification of diagnostic imaging [21], and prediction of hospital admission using EHR data [22].

Childhood asthma is one of the most common major chronic diseases, with high morbidity and an increasingly consistent burden on healthcare resource utilization [23]. It is also the primary diagnosis for 1/3 of pediatric ED visits [24,25]. However, asthma is an ambulatory care sensitive condition [26], therefore, children at high-risk for future repeated ED utilization are an important but poorly defined population who may be effectively managed by improved primary care and medications.

In this work, we investigated the use of deep learning models adapted to leverage temporal relations, such as recurrent neural network (RNN) models, together with the entire patient history of medical diagnosis from EHR, for predicting the frequency of ED utilization among children with asthma. The contribution of this work is three-fold:

- We explore the degree to which RNN models can predict the frequency of ED utilization among children with asthma, and compare the prediction accuracy of three RNN models against a baseline multinomial logistic regression.
- We demonstrate that patient's historic record of diagnosis codes routinely available in EHRs can be used to predict children with asthma at high-risk of repeated ED visits.
- We extend on previous work by including all ED visits by children with asthma, irrespective of cause, rather than only visits with asthma as the primary diagnosis.

2. Methods

2.1. Patient representation

A patient's EHR consists of a sequence of visits (encounters) a patient has made to healthcare facilities, and each visit captures the list of medical codes documented by the healthcare practitioners. We will use the following notation for representing our EHR dataset. Let $P = \{p_1, \ldots \ldots p_n\}$ be a dataset of n patients. Each patient p_j EHR is comprised of a sequence of T_j patient visits, $p_j = \{x_1, x_2, \ldots, x_{T_j}\}$, ordered by visit date $t \in \{1, T_j\}$. We express diagnosis codes as $\{d_1, d_2, \ldots, d_{|D|}\} \in D$, where D represents the entire set of unique diagnosis codes. Each patient visit x_i can be expressed as a binary vector $x_i \in \{0,1\}^{|D|}$, where the kth element is set to 1 if the ith visit contains the diagnosis code d_k, otherwise it is set to 0.

The temporal models used in this work require patient-level time-ordered data that has been collected over time. Therefore, we chose to present our EHR dataset in the form of list of lists of lists. The outermost list corresponds to patients, the intermediate list corresponds to the time-ordered visit sequence each patient made, and the innermost list corresponds to the medical codes that were documented within each visit.

A patient's visit list will be embedded before it is used as an input to the RNN models. The embedding is done as follows:

$$v_i = \sigma(W_{emb} x_i + b_x) \tag{1}$$

where m is the embedding size across D diagnosis codes, $v_i \in \mathbb{R}^m$ is the embedding of $x_i \in \mathbb{R}^D$, $W_{emb} \in \mathbb{R}^{m \times D}$ is the embedding matrix, σ is a non-linear activation function such as rectified linear unit (ReLU) or sigmoid, and b_x is the bias.

2.2. RNN models for emergency use prediction

Recurrent neural networks (RNNs) are an important and popular class of deep learning models that can utilize sequential information [27,28]. RNNs perform the same task for every element in the sequence, with the output being dependent on the previous computation. Given the power of RNNs for analyzing time-series data, we propose using three variations of RNNs for predicting the ED utilization among children with asthma: BiLSTMs, BiGRUs, and RETAIN.

2.2.1. BiLSTMs

Long-short term memory networks (LSTMs) [29] are a special architecture of RNNs, which can efficiently solve the long-term dependencies problem by introducing gating mechanism and memory cell. LSTMs also deal with the vanishing/exploding gradient problem during back propagation. Thus, they overcome both of the shortcomings that RNNs have. Bidirectional LSTM (BiLSTM) architecture [30] is used to capture both past and future information by concatenating hidden state of forward LSTM and backward LSTM. This allows the learning algorithm to better understand the context and eventually learn faster than unidirectional LSTM approach, although this might depend on the given task.

2.2.2. BiGRU

Gated recurrent unit networks (GRUs) [31] are a simplified version of LSTMs. The principal idea of a GRU is to overcome the vanishing gradient problem of a standard RNN with a gating mechanism, using two gates: an update gate and a reset gate. The update gate controls how much of previous memory to keep around, while the reset gate defines how to incorporate the previous input into the computation of the current input. Similar to BiLSTMs, bidirectional GRUs (BiGRUs), concatenate the hidden states of forward GRU and backward GRU producing output based on past and future data.

2.2.3. RETAIN

Reverse time attention model (RETAIN) is an RNN-based model which was recently introduced in [32] to help health care clinicians explaining why a model was predicting patients to be for example at risk of heart failure. RETAIN employs a factorized two-level attention mechanism to identify influential visits and significant features that contribute to the prediction. RETAIN achieves high accuracy as well as clinical interpretability. In our experiments, we compare the prediction performance of RETAIN to BiLSTMs and BiGRUs.

2.3. Multinomial logistic regression as baseline model for comparison

The main focus of this work is to explore the extent to which RNNs can predict the frequency of emergency department use, which is a time-series problem with multinomial outcomes. In addition, we try to assess if RNNs can harness the temporal information, embedded in the time-ordered EHR visits of a patient, to improve the prediction accuracy. For these two purposes, the baseline model should have two characteristics: the ability to handle multinomial outcomes and the ability to train without the need for sequential input data. We chose multinomial (multiclass) logistic regression (MLR) as our baseline model for comparison with RNNs. Multinomial logistic regression is an extension of binary logistic regression that allows for more than two categories of the outcome variable.

Fig. 1 illustrates the architecture of the multinomial logistic regression model used in our analysis. The input to the MLR model is the list of patients, and each patient is represented by a vector of aggregated features (diagnosis). For each feature, we calculated the total number of occurrences of a specific diagnosis code in any visit a patient made during his/her observation window. Clearly, this representation does not incorporate the temporal dimension of sequential events (patient visits).

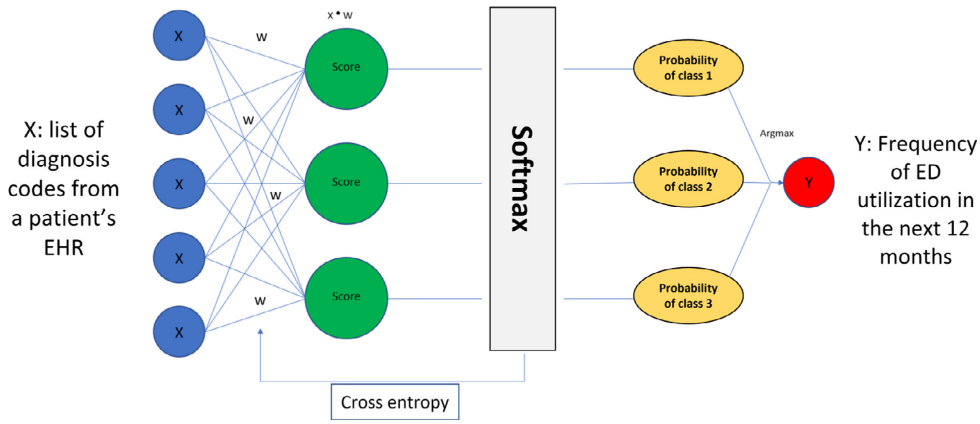

Fig. 1. Architecture of the multinomial logistic regression baseline.

The data is represented as (x, y) pairs for each patient, where each x is a feature vector of length D, where D is the number of unique diagnosis codes, and the label y is an integer in $\{1, \ldots, C\}$, where C is the number of classes. Using the $Softmax$ function, we can force the output layer to be a discrete probability distribution over the C classes. The targeted loss for optimization is the multinomial loss fit across the entire probability distribution, which can be computed using cross-entropy loss. The model, including the activation function can be written as:

$$\hat{y} = Softmax(Wx + b) \tag{2}$$

where W is a matrix of size $C \times D$ and b is a vector of biases of length C.

2.4. Data source and patient cohort

Data for this research was extracted from the Cerner Health Facts EHR database (currently referred to as the Cerner Real World Data), which contains patient-level data for over 43 million patients with 240 million encounters, collected over the past two decades from approximately 500 healthcare facilities across the United States [33].

For the purpose of this study, we included children aged 0–18 years, having a primary diagnosis of asthma indicated by the International Classification of Diseases (ICD) 9 standards of 493, who were attributed to one or more healthcare facility for at least 2 consecutive years between 2000 – –2013. We excluded patients who had less than three visits (of any type) to showcase RNN model's ability to use sequential information of visits. For each patient, we extracted the complete EHR history which includes the list of all encounters (visits), and for each encounter we extracted the list of all diagnosis codes. The detailed process of cohort construction is described in Fig. 2.

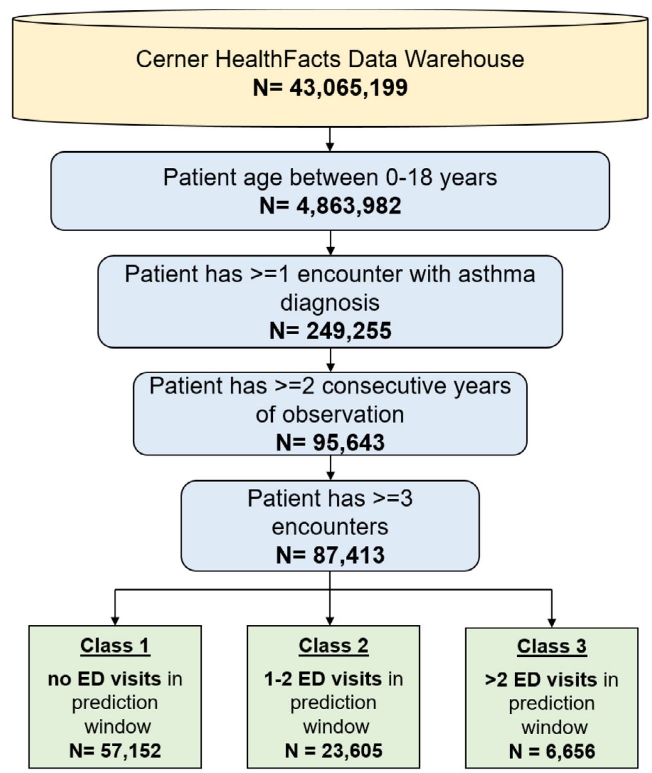

Fig. 2. Workflow of cohort construction.

2.5. Study definitions

2.5.1. Observation window and prediction window

The timeline for each patient is divided into an observation window and a prediction window based on the "index date", which is defined as 12 months prior to the last available record for each patient. The index date was also the time point at which frequency of ED utilization is to be predicted (Fig. 3). The observation window is the period before the index date (minimum requirement of $\geq$ 1 year) and the prediction window is the period after the index date (1 year for all patients). Obviously, the length of the observation window is variable for each patient depending on the available patient's history and patient age, ranging from 1 to 13 years. Data available from the observation window was used to make the prediction at the index date, while data from the prediction window were only used for defining patient labels.

2.5.2. Outcome measure

The outcome measure "ED use" was designed to capture the concept of patients who goes to the hospital for emergency medical services, during the prediction window. Therefore, we included all ED visits whether it was followed by discharge home or resulted in hospital admission. We also included all ED visits made by children with asthma, irrespective of cause, rather than only visits with asthma as the primary diagnosis. This is mainly to account for ED visits due to conditions resulting from asthma co-morbidities. The outcome variable is divided into three classes based on the frequency of ED use, which is measured by the number of visits a patient made to the emergency department during the one year prediction window. The three classes are defined as follows: class 1 refers to patients who did not make any ED visits during the prediction window, class 2 refers to patients who made one or two ED visits during the prediction window, and class 3 refers to patients who made three or more ED visits during the prediction window (frequent ED users).

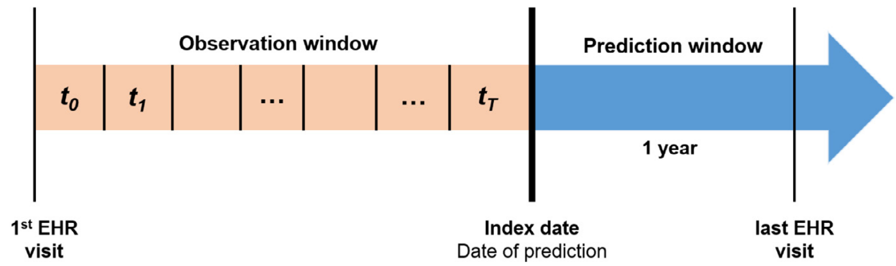

Fig. 3. Patient EHR timeline. Index date is defined as 1 year prior to the last available record for each patient. The observation window is the period before the index date and the prediction window is the period after the index date.

2.6. Implementation details

For each patient, we want to predict the class of ED utilization in the next 12 months based on previous EHR records. To validate the performance of the proposed models in this prediction task, we conducted experiments on two classes of methods: RNN-based models and baseline multinomial regression. The dataset was randomly divided into (70%) training, (10%) validation, and (20%) testing subsets. We maintained the same proportion of the three classes among the training, validation, and testing sets. The predictive accuracy of the models was evaluated using three metrics: area under the receiver operating characteristic curve (AUC–ROC), area under the precision–recall curve (AUC-PR), and F1 score. Since outcomes in our dataset are highly imbalanced, the selected metrics are well suited for this context. The RNN models were trained with Adam [34] optimizer, batch size: 16, loss: categorical cross-entropy, learning rate: 0.0001, and 1 recurrent layer with 64 hidden units. We used 10 epochs, with early stopping, which was sufficient for all models to converge. The early stopping was based on the AUC-PR estimated on the validation dataset. After an initial round of preliminary experiments, we set the embedding size of the diagnosis codes to be 200. The multinomial regression model was implemented with $L2$ regularization using newton-cg solver, and the maximum number of iterations for the solver to converge was set to 100. The RNN models were implemented using Keras 2.2.4 and the baseline model was implemented using Python Scikit-Learn 0.20.1.

Table 1
Predictive performance of RNN models compared to baseline multinomial logistic regression.

(a) Area under receiver operating characteristic (AUC–ROC)		
Model	Micro (std. deviation)	Macro (std. deviation)
Multinomial regression	0.81 (0.002)	0.66 (0.003)
RETAIN	0.85 (0.002)	0.72 (0.002)
BiLSTM	0.85 (0.002)	0.72 (0.005)
BiGRU	0.85 (0.003)	0.72 (0.003)
(b) Area under precision–recall (AUC-PR)		
Model	Micro	Macro
Multinomial regression	0.69 (0.003)	0.43 (0.003)
RETAIN	0.74 (0.004)	0.44 (0.039)
BiLSTM	0.74 (0.002)	0.48 (0.036)
BiGRU	0.74 (0.005)	0.46 (0.002)
(c) F1 score		
Model	Micro	Weighted
Multinomial regression	0.63 (0.005)	0.56 (0.008)
RETAIN	0.66 (0.002)	0.60 (0.002)
BiLSTM	0.67 (0.002)	0.60 (0.008)
BiGRU	0.66 (0.004)	0.61 (0.005)

3. Results

3.1. Overview of emergency department users

This study used retrospective EHR data of 87,413 children with asthma. Each sample from the cohort was assigned a label from the three classes, according to our definition of the outcome variable in the Methods section. Statistics of the cohort used in this analysis is shown in Fig. 4. Class 1, which includes patients who did not visit the ED during the prediction window (12 months) accounted for 65% of children with asthma. Low ED users (class 2), which includes patients who made 1 or 2 visits to the ED, accounted for 27%. Frequent ED users (class 3), which includes patients who made 3 or more visits to the ED, accounted for 8% of all pediatric asthma patients. These figures agree with similar estimates in the literature which states that frequent ED users account for 1 to 8% of all ED patients [1,2,12].

The data also demonstrated that class 1 patients have the highest average number of visits per patient and the lowest average number of diagnosis codes per visit. This could reflect that this subgroup of patients have better access to primary care with regular follow-up visits. Such follow-up visits often help with better assessment of asthma-related risk and impairment, adjustment of treatment, and education of patients [35]. As a result, those patients are at lower risk of future ED use. Although class 3 patients also have a high rate of average number of visits per patient, however, they also have the highest average number of diagnosis codes per visit. This could reflect acute asthma which may result in more complications (diagnosis) and require patients to seek ED care more frequently.

The most frequent diagnosis codes seem to be similar among the three classes, which, in addition to asthma, includes: acute upper respiratory infections of multiple or unspecified sites (ICD-9: 465), general symptoms (ICD-9: 780), symptoms involving respiratory system and other chest symptoms (ICD-9: 786), suppurative and unspecified otitis media (ICD-9: 382), symptoms involving digestive system (ICD-9: 787), and viral and chlamydial infection in conditions classified elsewhere and of unspecified site (ICD-9: 079).

3.2. Prediction results

Table 1 presents the predictive performance of RNN and baseline models for the task of predicting the frequency of emergency department visits among pediatric asthma patients, using EHR data. We found that all the RNN models we tested showed better prediction accuracy over the baseline model of multinomial regression among all performance metrics. RNN models outperformed the baseline model by 4% AUC–ROC and 5% AUC-PR. In addition, the F1 weighted-average score for RNN models was 4%–5% higher than the baseline model. We also noticed that the BiLSTM and BiGRU models both achieved comparable performance. Hence, this confirms that GRUs may be used instead of LSTMs for similar prediction tasks, without losing significant prediction accuracy. GRUs are simpler, easier to modify, and faster to train than LSTMs [36]. Moreover, RETAIN demonstrated a predictive performance comparable to BiLSTM and BiGRU models. This supports the main idea behind the RETAIN architecture to provide comparable performance to RNN variants while offering a better interpretability.

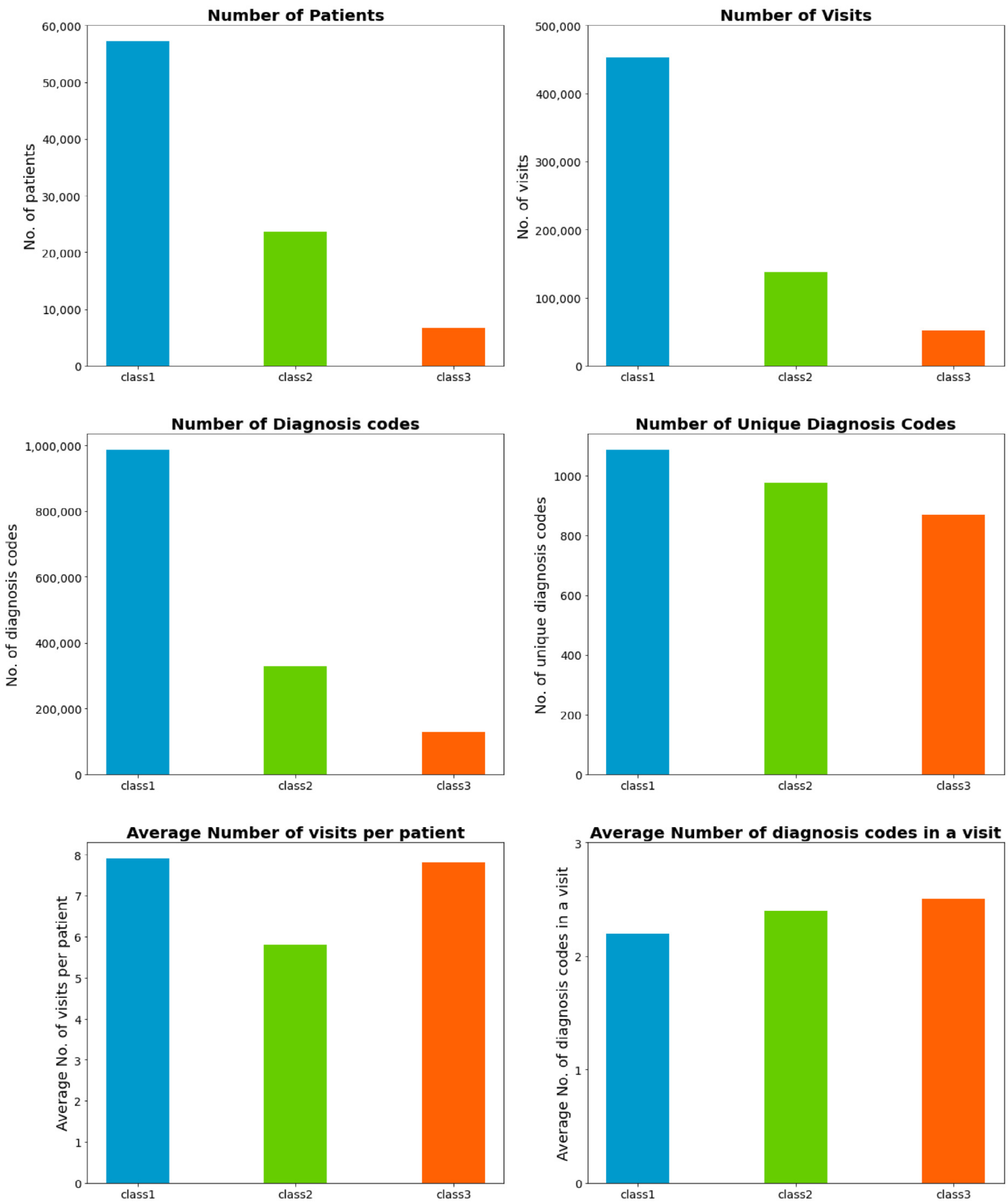

Fig. 4. Summary statistics for emergency department users cohort.

4. Discussion

In this study, we show that recurrent neural networks can effectively model the complex temporal relationships in EHR data to predict emergency department utilization among children with asthma. We also demonstrate that capturing the patient visits sequentially (as they appear in EHR) provides better predictive performance compared to traditional data representation of aggregating all features together without considering the temporal dimension. In addition, using RNNs to predict ED utilization alleviates the need for manually extracting features, because RNNs can explore the features of different variables at the same time automatically without additional computing cost.

4.1. Comparison with existing work

While there has been some progress with predicting ED utilization among children with asthma [37–40], these models are mostly trained as traditional statistical models, including logistic regression, random forests, support vector machines and decision trees. Moreover, although a few recent studies have sought to use deep learning techniques to predict the ED utilization among children with asthma [41,42], those studies could not incorporate the entire patient history accessible from the electronic health record (EHR) and only utilized a limited number of features (<100) from EHR or other sources, leaving out rich valuable information. In addition, the deep learning models used in these studies do not utilize the temporal dimension of the patients longitudinal EHR record, which could potentially improve the prediction accuracy. Furthermore, many of previous work on ED utilization has focused on predicting only asthma-related ED visits [43], ignoring other ED visits which could be caused by other co-morbidities associated with asthma [44].

Our results are consistent with a recent study [41], which utilized artificial neural networks (ANNs) to predict asthma-related ED visits or hospitalization within the next 3 months, using administrative claims data. This previous study showed that ANN slightly outperforms (with AUC = 0.845) the Lasso logistic regression (with AUC = 0.842). However, there are several key differences between our study and this previous study. First, our features vector included 1106 features representing all the ICD-9 diagnosis codes which were present in the asthma cohort. On the other hand, the previous study included only 33 features manually extracted from the claims data. These features can be broadly classified into five categories: demographics, medication, health service utilization, comorbid illnesses, and insurance gap. Manually extracting a limited set of pre-defined features may result in missing important variables which could improve the predictive performance. Second, our study was conducted on a cohort of 87,413 patients, while the previous study included a total of 28,378 patients. Using a larger population allows for better generalizability of our findings. In addition, our study utilized the temporal relationships in the data using RNN architectures, while in the previous study the temporal relationships in the claims data were neither represented in the input data nor incorporated in the ANN models architecture. Moreover, the prediction task in the previous study is limited to asthma-related visits and over a period of 3 months only, while our study extends the prediction to include all ED visits by children with asthma within the next 12 months, irrespective of cause, rather than only visits with asthma as the primary diagnosis. This has the advantage of considering ED visits caused by comorbidities which are directly or indirectly associated with asthma, rather than only ED visits with asthma as the primary diagnosis.

4.2. Models interpretability

Prediction of clinical outcomes with RNNs can have high accuracy but are unfortunately difficult to interpret as a result of feature-interactions, temporal-interactions, and non-linear transformations. Several important model-agnostic interpretability techniques exist, and while none of them are perfect, they can be used to interpret the results of simple and complex machine learning models [45]. Among these methods, we describe the Shapley Additive explanations (SHAP) technique [46]. The concept of the SHAP method is inspired by the game-theory and is based on computing the contribution score for each feature for individual predictions. A prediction can be explained by assuming that each feature value of the instance is a "player" in a game and the contribution of each player is computed by including and excluding the player from all subsets of the rest of the players. In SHAP, the authors first describe the class of additive feature attribution methods, which unifies six current methods, including LIME [47], Layer-Wise Relevance Propagation [48] and DeepLIFT [49], which all use the same explanation model. Then, they suggest SHAP values as a unified measure of feature importance that maintains three necessary properties: local accuracy, missingness, and consistency. Finally, they describe several different methods for estimating SHAP values, as well as experiments demonstrating not only the improved performance of these values in terms of distinguishing between different output classes, but also in terms of better aligning with human intuition, when compared to many other existing interpretability techniques. In our work, SHAP method can be used to interpret the predictions of the RNN models and provide insight into the relationships they have learned, which is necessary for validate and trusting the predictions produced by the RNN models used in our work. Such interpretability mechanisms are likely to yield increased transparency in the steps by which predictions are generated by the RNN models and in a traceable component to learn from their possible deficiencies.

4.3. Limitations

A potential limitation of our study is that we chose to restrict patient history to medical diagnosis while leaving out other patient data such as medications, lab tests, procedures, and patients demographics. We anticipate that expanding the sources of historical data may improve model performance. Another limitation is that EHR databases often lack important variables which are vital for identifying and assessing asthma patients. Such variables include: lung function, socioeconomic status, environmental factors, and adherence to treatment and lifestyle [50]. These challenges are, however not insurmountable. Furthermore, since our outcome classes are highly imbalanced, it may be required to use customized performance metrics that are designed to handle data imbalance. Matthews Correlation Coefficient (MCC) [51] is widely used in Bioinformatics as a performance metric, and its recent extension [52] to imbalanced data could be used as a performance metric for optimization to overcome the issue of data imbalance in our dataset. Finally, since this is an observational study with diagnoses or conditions that vary over time, standard approaches for adjustment of confounding are biased when there exist time-dependent confounders that are also affected by previous conditions. A possible approach to mitigate this limitation is to use marginal structural models, which allows for adjustment of confounding in those situations [53–55].

5. Conclusion

Given the power of RNNs for modeling the temporal relationships embedded in time-series data, we proposed to use variants of RNNs (BiLSTM, BiGRU, and RETAIN) to predict emergency department utilization among pediatric asthma patients, using EHR data. RNN models showed better predictive accuracy over baseline multinomial regression models for multi-class prediction. However, to address the primary cause of frequent ED use and appropriately decide on interventions, it is necessary to further understand what subset of features have the largest contribution to the prediction. Therefore, future work may focus on utilizing attention-based mechanisms, that has been shown effective for identifying important features, to better understand the significant variables contributing to the prediction score.

CRediT authorship contribution statement

Rawan AlSaad: Developed the idea, Implemented the methods, Conducted the experiments, Drafted the manuscript. **Qutaibah Malluhi:** Supervised every step of the work, Provided critical revision of the manuscript. **Ibrahim Janahi:** Provided clinical input to the cohort design. **Sabri Boughorbel:** Supervised every step of the work, Provided critical revision of the manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Availability of data and materials

The Cerner Health Facts Database (currently referred to as the Cerner Real World Data) is not publicly available. It is available to research affiliates at contributing hospitals, upon a request made directly to Cerner Corporation.

Acknowledgments

The authors acknowledge Sidra Medicine for the support provided to access the data, through a Data Use Agreement with Cerner. All authors read and approved the final version.

Funding information

Not applicable

Ethics approval and consent to participate

This research was approved by the Institutional Review Board at Sidra Medicine (protocol number 1804023464). Informed consent was exempted because of the retrospective nature of this research. Patient data were anonymized and de-identified by the data provider.

References

[1] K.A. Hunt, E.J. Weber, J.A. Showstack, D.C. Colby, M.L. Callaham, Characteristics of frequent users of emergency departments, Anna. Emerg. Med. 48 (1) (2006) 1–8, http://dx.doi.org/10.1016/j.annemergmed.2005.12.030.

[2] K.K. Fuda, R. Immekus, Frequent users of massachusetts emergency departments: A statewide analysis, Anna. Emerg. Med. 48 (1) (2006) 16.e1–16.e8, http://dx.doi.org/10.1016/j.annemergmed.2006.03.001.

[3] E.A. Hooker, P.J. Mallow, M.M. Oglesby, Characteristics and trends of emergency department visits in the united states (2010–2014), J. Emerg. Med. 56 (3) (2019) 344–351, http://dx.doi.org/10.1016/j.jemermed.2018.12.025.

[4] P.B. Patel, M.A. Combs, D.R. Vinson, Reduction of admit wait times: The effect of a leadership-based program, in: S.M. Schneider (Ed.), Acad. Emerg. Med. 21 (3) (2014) 266–273, http://dx.doi.org/10.1111/acem.12327.

[5] J. Crilly, N. Bost, L. Thalib, J. Timms, H. Gleeson, Patients who present to the emergency department and leave without being seen, Eur. J. Emerg. Med. 20 (4) (2013) 248–255, http://dx.doi.org/10.1097/mej.0b013e328356fa0e.

[6] D.J. Wu, E. Hipolito, A. Bilderback, S.O. Okelo, A. Garro, Predicting future emergency department visits and hospitalizations for asthma using the pediatric asthma control and communication instrument – emergency department version (PACCI-ED), J. Asthma 53 (4) (2016) 387–391, http://dx.doi.org/10.3109/02770903.2015.1115520.

[7] P. Milbrett, M. Halm, Characteristics and predictors of frequent utilization of emergency services, J. Emerg. Nurs. 35 (3) (2009) 191–198, http://dx.doi.org/10.1016/j.jen.2008.04.032.

[8] S. Zuckerman, Y.-C. Shen, Characteristics of occasional and frequent emergency department users, Med. Care 42 (2) (2004) 176–182, http://dx.doi.org/10.1097/01.mlr.0000108747.51198.41.

[9] F.S. Blank, H. Li, P.L. Henneman, H.A. Smithline, J.S. Santoro, D. Provost, A.M. Maynard, A descriptive study of heavy emergency department users at an academic emergency department reveals heavy ED users have better access to care than average users, J. Emerg. Nurs. 31 (2) (2005) 139–144, http://dx.doi.org/10.1016/j.jen.2005.02.008.

[10] M.B. Doupe, W. Palatnick, S. Day, D. Chateau, R.-A. Soodeen, C. Burchill, S. Derksen, Frequent users of emergency departments: Developing standard definitions and defining prominent risk factors, Anna. Emerg. Med. 60 (1) (2012) 24–32, http://dx.doi.org/10.1016/j.annemergmed.2011.11.036.

[11] R. Drewek, L. Mirea, A. Rao, P. Touresian, P.D. Adelson, Asthma treatment and outcomes for children in the emergency department and hospital, J. Asthma 55 (6) (2017) 603–608, http://dx.doi.org/10.1080/02770903.2017.1355381.

[12] E. LaCalle, E. Rabin, Frequent users of emergency departments: The myths, the data, and the policy implications, Anna. Emerg. Med. 56 (1) (2010) 42–48, http://dx.doi.org/10.1016/j.annemergmed.2010.01.032.

[13] J. Stewart, P. Sprivulis, G. Dwivedi, Artificial intelligence and machine learning in emergency medicine, Emerg. Med. Australas. 30 (6) (2018) 870–874, http://dx.doi.org/10.1111/1742-6723.13145.

[14] H. Ehrlich, M. McKenney, A. Elkbuli, The niche of artificial intelligence in trauma and emergency medicine, Am. J. Emerg. Med. 45 (2021) 669–670, http://dx.doi.org/10.1016/j.ajem.2020.10.050.

[15] K. Grant, A. McParland, S. Mehta, A.D. Ackery, Artificial intelligence in emergency medicine: Surmountable barriers with revolutionary potential, Anna. Emerg. Med. 75 (6) (2020) 721–726, http://dx.doi.org/10.1016/j.annemergmed.2019.12.024.

[16] A. Kirubarajan, A. Taher, S. Khan, S. Masood, Artificial intelligence in emergency medicine: A scoping review, J. Am. Coll. Emerg. Physicians Open 1 (6) (2020) 1691–1702, http://dx.doi.org/10.1002/emp2.12277.

[17] S.J. Patel, D.B. Chamberlain, J.M. Chamberlain, A machine learning approach to predicting need for hospitalization for pediatric asthma exacerbation at the time of emergency department triage, in: R. Cloutier (Ed.), Acad. Emerg. Med. 25 (12) (2018) 1463–1470, http://dx.doi.org/10.1111/acem.13655.

[18] A.D. Hond, W. Raven, L. Schinkelshoek, M. Gaakeer, E.T. Avest, O. Sir, H. Lameijer, R.A. Hessels, R. Reijnen, E.D. Jonge, E. Steyerberg, C.H. Nickel, B.D. Groot, Machine learning for developing a prediction model of hospital admission of emergency department patients: Hype or hope? Int. J. Med. Inf. 152 (2021) 104496, http://dx.doi.org/10.1016/j.ijmedinf.2021.104496.

[19] V. Abedi, N. Goyal, G. Tsivgoulis, N. Hosseinichimeh, R. Hontecillas, J. Bassaganya-Riera, L. Elijovich, J.E. Metter, A.W. Alexandrov, D.S. Liebeskind, A.V. Alexandrov, R. Zand, Novel screening tool for stroke using artificial neural network, Stroke 48 (6) (2017) 1678–1681, http://dx.doi.org/10.1161/strokeaha.117.017033.

[20] N.W. Sterling, R.E. Patzer, M. Di, J.D. Schrager, Prediction of emergency department patient disposition based on natural language processing of triage notes, Int. J. Med. Inf. 129 (2019) 184–188, http://dx.doi.org/10.1016/j.ijmedinf.2019.06.008.

[21] K. Yadav, E. Sarioglu, H.-A. Choi, W.B. Cartwright, P.S. Hinds, J.M. Chamberlain, Automated outcome classification of computed tomography imaging reports for pediatric traumatic brain injury, in: M. Hauswald (Ed.), Acad. Emerg. Med. 23 (2) (2016) 171–178, http://dx.doi.org/10.1111/acem.12859.

[22] X. Zhang, J. Kim, R.E. Patzer, S.R. Pitts, A. Patzer, J.D. Schrager, Prediction of emergency department hospital admission based on natural language processing and neural networks, Methods Inf. Med. 56 (05) (2017) 377–389, http://dx.doi.org/10.3414/me17-01-0024.

[23] G. Ferrante, S.L. Grutta, The burden of pediatric asthma, Front. Pediatr. 6 (2018) http://dx.doi.org/10.3389/fped.2018.00186.

[24] J.S. Weissman, Rates of avoidable hospitalization by insurance status in massachusetts and maryland, JAMA: J. Am. Med. Assoc. 268 (17) (1992) 2388, http://dx.doi.org/10.1001/jama.1992.03490170060026.

[25] T. Wang, T. Srebotnjak, J. Brownell, R.Y. Hsia, Emergency department charges for asthma-related outpatient visits by insurance status, J. Health Care Poor Underserved 25 (1) (2014) 396–405, http://dx.doi.org/10.1353/hpu.2014.0051.

[26] A.J. Trachtenberg, N. Dik, D. Chateau, A. Katz, Inequities in ambulatory care and the relationship between socioeconomic status and respiratory hospitalizations: A population-based study of a Canadian city, Annal. Family Med. 12 (5) (2014) 402–407, http://dx.doi.org/10.1370/afm.1683.

[27] J. van der Westhuizen, J. Lasenby, The unreasonable effectiveness of the forget gate, 2018, CoRR abs/1804.04849, arXiv:1804.04849.

[28] H. Salehinejad, J. Baarbe, S. Sankar, J. Barfett, E. Colak, S. Valaee, Recent advances in recurrent neural networks, 2018, CoRR abs/1801.01078, arXiv:1801.01078.

[29] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (1997) 1735–1780, http://dx.doi.org/10.1162/neco.1997.9.8.1735.

[30] F.A. Gers, J. Schmidhuber, F. Cummins, Learning to forget: Continual prediction with LSTM, Neural Comput. 12 (10) (2000) 2451–2471, http://dx.doi.org/10.1162/089976600300015015.

[31] K. Cho, B. van Merrienboer, Ç. Gülçehre, F. Bougares, H. Schwenk, Y. Bengio, Learning phrase representations using RNN encoder-decoder for statistical machine translation, 2014, CoRR abs/1406.1078, arXiv:1406.1078.

[32] E. Choi, M.T. Bahadori, A. Schuetz, W.F. Stewart, J. Sun, RETAIN: interpretable predictive model in healthcare using reverse time attention mechanism, 2016, CoRR abs/1608.05745, arXiv:1608.05745.

[33] J.P. DeShazo, M.A. Hoffman, A comparison of a multistate inpatient EHR database to the HCUP nationwide inpatient sample, BMC Health Serv. Res. 15 (1) (2015) http://dx.doi.org/10.1186/s12913-015-1025-7.

[34] D.P. Kingma, J. Ba, Adam: A method for stochastic optimization, 2014, pp. 1–15, arXiv:1412.6980.

[35] K.A. Nelson, J.M. Garbutt, M.J. Wallendorf, K.M. Trinkaus, R.C. Strunk, Primary care visits for asthma monitoring over time and association with acute asthma visits for urban medicaid-insured children, J. Asthma 51 (9) (2014) 907–912, http://dx.doi.org/10.3109/02770903.2014.927483.

[36] J. Chung, C. Gulcehre, K. Cho, Y. Bengio, Empirical evaluation of gated recurrent neural networks on sequence modeling, 2014, pp. 1–9, arXiv:1412.3555.

[37] L.T. Das, E.L. Abramson, A.E. Stone, J.E. Kondrich, L.M. Kern, Z.M. Grinspan, Predicting frequent emergency department visits among children with asthma using EHR data, Pediatr. Pulmonol. 52 (7) (2017) 880–890, http://dx.doi.org/10.1002/ppul.23735.

[38] J. Wu, S.J. Grannis, H. Xu, J.T. Finnell, A practical method for predicting frequent use of emergency department care using routinely available electronic registration data, BMC Emerg. Med. 16 (1) (2016) http://dx.doi.org/10.1186/s12873-016-0076-3.

[39] S. Giangioppo, V. Bijelic, N. Barrowman, D. Radhakrishnan, Emergency department visit count: a practical tool to predict asthma hospitalization in children, J. Asthma (2019) 1–10, http://dx.doi.org/10.1080/02770903.2019.1635151.

[40] T. Lieu, C. Quesenberry, M. Sorel, G. Mendoza, A. Leong, Computer-based models to identify high-risk children with asthma, Am. J. Respir. Crit. Care Med. 157 (4) (1998) 1173–1180, http://dx.doi.org/10.1164/ajrccm.157.4.9708124.

[41] X. Wang, Z. Wang, Y.M. Pengetnze, B.S. Lachman, V. Chowdhry, Deep learning models to predict pediatric asthma emergency department visits, 2019, pp. 1–7, arXiv:1907.11195.

[42] T. Goto, C.A. Camargo, M.K. Faridi, R.J. Freishtat, K. Hasegawa, Machine learning–based prediction of clinical outcomes for children during emergency department triage, JAMA Netw. Open 2 (1) (2019) e186937, http://dx.doi.org/10.1001/jamanetworkopen.2018.6937.

[43] S. Ram, W. Zhang, M. Williams, Y. Pengetnze, Predicting asthma-related emergency department visits using big data, IEEE J. Biomed. Health Inf. 19 (4) (2015) 1216–1223, http://dx.doi.org/10.1109/jbhi.2015.2404829.

[44] N. Ullmann, V. Mirra, A.D. Marco, M. Pavone, F. Porcaro, V. Negro, A. Onofri, R. Cutrera, Asthma: Differential diagnosis and comorbidities, Front. Pediatr. 6 (2018) http://dx.doi.org/10.3389/fped.2018.00276.

[45] P. Linardatos, V. Papastefanopoulos, S. Kotsiantis, Explainable AI: A review of machine learning interpretability methods, Entropy 23 (1) (2020) 18, http://dx.doi.org/10.3390/e23010018.

[46] S.M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, in: NIPS17, Curran Associates Inc., Red Hook, NY, USA, 2017, pp. 4768–4777.

[47] M.T. Ribeiro, S. Singh, C. Guestrin, Why should I trust you?: Explaining the predictions of any classifier, 2016, arXiv:1602.04938.

[48] A. Binder, G. Montavon, S. Lapuschkin, K.-R. Muller, W. Samek, Layer-wise relevance propagation for neural networks with local renormalization layers, in: ICANN, 2016.

[49] A. Shrikumar, P. Greenside, A. Kundaje, Learning important features through propagating activation differences, in: Proceedings of the 34th International Conference on Machine Learning - Volume 70, in: ICML17, JMLR.org, 2017, pp. 3145–3153.

[50] R.E. Greenblatt, E.J. Zhao, S.E. Henrickson, A.J. Apter, R.A. Hubbard, B.E. Himes, Factors associated with exacerbations among adults with asthma according to electronic health record data, Asthma Res. Prac. 5 (1) (2019) http://dx.doi.org/10.1186/s40733-019-0048-y.

[51] B. Matthews, Comparison of the predicted and observed secondary structure of T4 phage lysozyme, Biochimica Et Biophys. Acta (BBA) - Protein Struct. 405 (2) (1975) 442–451, http://dx.doi.org/10.1016/0005-2795(75)90109-9.

[52] S. Boughorbel, F. Jarray, M. El-Anbari, Optimal classifier for imbalanced data using Matthews correlation coefficient metric, in: Q. Zou (Ed.), PLOS ONE 12 (6) (2017) e0177678, http://dx.doi.org/10.1371/journal.pone.0177678.

[53] J.M. Robins, M.A. Hernan, B. Brumback, Marginal structural models and causal inference in epidemiology, Epidemiology 11 (5) (2000) 550–560, http://dx.doi.org/10.1097/00001648-200009000-00011.

[54] B. Lim, A. Alaa, M.v.d. Schaar, Forecasting treatment responses over time using recurrent marginal structural networks, in: Proceedings of the 32nd International Conference on Neural Information Processing Systems, in: NIPS18, Curran Associates Inc., Red Hook, NY, USA, 2018, pp. 7494–7504.

[55] I. Bica, A.M. Alaa, J. Jordon, M. van der Schaar, Estimating counterfactual treatment outcomes over time through adversarially balanced representations, 2020, arXiv:2002.04083.