DyHealth: Making Neural Networks Dynamic for Effective Healthcare Analytics

Kaiping Zheng
National University of Singapore
kaiping@comp.nus.edu.sg

Shaofeng Cai
National University of Singapore
shaofeng@comp.nus.edu.sg

Horng Ruey Chua
National University Health System
horng_ruey_chua@nuhs.edu.sg

Melanie Herschel
Universität Stuttgart
melanie.herschel@ipvs.uni-stuttgart.de

Meihui Zhang
Beijing Institute of Technology
meihui_zhang@bit.edu.cn

Beng Chin Ooi
National University of Singapore
ooibc@comp.nus.edu.sg

ABSTRACT

In National University Hospital (NUH) in Singapore, we conduct healthcare analytics that analyzes heterogeneous electronic medical records (EMR) to support effective clinical decision-making on a daily basis. Existing work mainly focuses on *multimodality* for extracting complementary information from different modalities, and/or *interpretability* for providing interpretable prediction results. However, real-world healthcare analytics has presented another major challenge, i.e., the available data modalities evolve or change intermittently. Addressing this challenge requires deployed models to be adaptive to such dynamic modality changes.

To meet the aforementioned requirement, we develop a *modular*, *multimodal* and *interpretable* framework DyHealth to enable dynamic healthcare analytics. Specifically, different modalities are processed within their respective data modules that adhere to the interface defined by DyHealth. The extracted information from different modalities is integrated subsequently in our proposed Multimodal Fusion Module. To better handle modality changes at runtime, we further propose exponential increasing/decreasing mechanisms to support modality "hot-plug". We also devise a novel modality-based attention mechanism for providing fine-grained interpretation results on a per-input basis. We conduct a pilot evaluation of DyHealth on the patients' EMR data from the NUH division of nephrology, in which DyHealth achieves superior performance and hence, is promising to roll out for hospital-wide deployment. We also validate DyHealth in two public EMR datasets. Experimental results confirm the effectiveness, flexibility, and extensibility of DyHealth in supporting multimodal and interpretable healthcare analytics.

PVLDB Reference Format:

Kaiping Zheng, Shaofeng Cai, Horng Ruey Chua, Melanie Herschel, Meihui Zhang, and Beng Chin Ooi. DyHealth: Making Neural Networks Dynamic for Effective Healthcare Analytics.

PVLDB Artifact Availability:

The source code, data, and/or other artifacts have been made available at https://anonymous.4open.science/r/DyHealth-VLDB2022.

This work is licensed under the Creative Commons BY-NC-ND 4.0 International License. Visit https://creativecommons.org/licenses/by-nc-nd/4.0/ to view a copy of this license. For any use beyond those covered by this license, obtain permission by emailing info@vldb.org. Copyright is held by the owner/author(s). Publication rights licensed to the VLDB Endowment.

Proceedings of the VLDB Endowment, Vol. 14, No. 1 ISSN 2150-8097.
doi:XX.XX/XXX.XX

1 INTRODUCTION

National University Health System (NUHS) [4] consists of a cluster of four public hospitals, three national specialty centres, and a polyclinic in Singapore. Apart from using commercial products for managing its healthcare data for daily operations, NUHS also collaborates with universities and research institutes to create values from the data, including data acquisition, data cleaning, data integration [17], data processing [35, 38], data analytics [70] and data visualization, with the ultimate aim of improving healthcare services and outcomes.

National University Hospital (NUH) [5] is a tertiary referral hospital in Singapore and the largest hospital of NUHS, serving as the flagship hospital for the cluster. In NUH, innovative approaches are employed for healthcare analytics [1], which is critical for medical practitioners to make effective and timely decisions on patient management and resource allocation. In particular, electronic medical records (EMRs) are the main data sources for supporting healthcare analytics. EMR data are typically collected from multiple sources, and therefore it is important to address the issue of *multimodality*. In a similar vein, *interpretability* for understanding model predictions is essential, since we are making life and death decisions. We shall further elaborate on their importance below.

- **Multimodality.** A modality of EMR data represents one particular way in which the health conditions of a patient are captured. EMR data typically encompass multiple modalities: (i) structured data such as patients' demographics, diagnoses, lab tests, medications, procedures, and (ii) unstructured data such as image data (e.g., magnetic resonance imaging (MRI) scans and computerized tomography (CT) scans), and text data (e.g., doctors' notes), etc. EMR data are heterogeneous in nature, and different modalities contain complementary information for data analytics [8]. Therefore, developing models that handle multimodal data is essential for achieving effective healthcare analytics.
- **Interpretability.** Interpretability measures the extent to which predictions produced by the model can be understood by humans [42, 69]. This is important and necessary for many critical applications. Particularly, in healthcare analytics, simply reporting the prediction results of the model to clinicians without explanations is not acceptable. To mitigate this issue, it is imperative to consider interpretability in the model design so that the model can also explain "why" certain decisions are made [42]. Such interpretability is invaluable for clinicians to derive medically meaningful insights for providing better healthcare.

```
# initialize a suspected case detection
DyHealth.register_module(
    'Lab Tests',LabTestModule())

# add chest imaging as complementary criteria
DyHealth.register_module(
    'Chest Images',ImageModule1())

# upgrade module for an existing modality
DyHealth.remove_module('Chest Images')
DyHealth.register_module(
    'Chest Images', ImageModule2())
```

Figure 1: A `DyHealth` example for modality "hot-plug".

There has been a long stream of research addressing these two challenges, e.g., [19, 49, 50, 67] for multimodal modeling [8], and [15, 39, 58, 69] for prediction with interpretability [9, 41, 42]. Several dedicated healthcare frameworks [11, 26, 30, 33, 40, 59, 60, 64, 65, 68, 69] have also been developed in response to either or both challenges. Nonetheless, while working on a wide spectrum of healthcare applications, we face the challenge that the data modalities often evolve or change intermittently, which cannot be well handled by existing work. For instance, (i) the hospital may employ some scanners to support a new modality of medical images such as MRI, CT; hence, a new module needs to be integrated, (ii) the modalities required for the diagnosis of a specific disease may be updated due to medical research discovery. Therefore, after the initial installation of a healthcare analytic model in hospitals, new modules may need to be integrated into the deployed model, or certain existing modules may need to be upgraded or removed at times. Further, an existing module may also need to be upgraded to a better-performing one, e.g., via neural architecture search (NAS) [18]. Such dynamic changes of modalities require the adaptation of the deployed model at runtime. This is highly desirable in clinical practice, yet has not been investigated by existing approaches as of today.

In this work, we build upon but go beyond multimodality and interpretability, and advocate *modularity* design for a more flexible and extensible healthcare framework. Modularity [3] refers to a logical division of the framework into manageable modules for ease of implementation, upgrade, and maintenance. The modular design of healthcare analytics can be based on related functions, data flows, or other considerations. Notably, a clear division of the functional modules and their processing adhering to a standardized interface are at the core of a modular healthcare framework design. It allows different modules to be developed independently, and each module can be added, removed, or replaced in a plug-and-play manner for flexibility and extensibility.

Contributions. We present `DyHealth`, a modular, multimodal and interpretable framework for **Dy**namic **Health**care data analytics. In Figure 1, we illustrate how `DyHealth` allows to keep up with supporting COVID-19 diagnosis as diagnostic methods evolve, by simply registering a new module or removing an existing module. In the early stage of the pandemic, the confirmation of the infection relied heavily on laboratory testing [46], and `DyHealth` can initialize a suspected case detection by registering a module that processes lab tests. Later on, chest imaging started to be used in acute care of adult patients with suspected, probable or confirmed cases [47], and was included as a new criterion of World Health Organization (WHO) COVID-19 case definition [48]. In such a scenario, `DyHealth` can readily adapt to the dynamic change of modalities by plugging in a new module during runtime. Further, due to the research development of computer vision, the existing image module may need to be upgraded to a more advanced one for boosted performance [45]. In this case, `DyHealth` allows to simply replace the existing module for the modality.

In devising `DyHealth`, we first develop representative data modules for processing respective modalities in `DyHealth`. We then propose the Multimodal Fusion Module to capture the cross-modality interactions and fuse the information from different modalities. We also devise a modality-based attention mechanism for this module to dynamically capture the importance of each modality given the input sample, as the relative importance of different modalities can vary from one sample to another. After fusing the complementary multimodal information, the Prediction Module is used to perform the final prediction tasks. Therefore, `DyHealth` is designed to support multimodality. `DyHealth` achieves modularity by firstly, separating the processing of different modalities into their respective data modules and secondly, integrating exponential increasing/decreasing mechanisms to support the runtime plugging in/out of a data module. With the proposed modality-based attention mechanism and other fine-grained interpretation techniques in the data modules, `DyHealth` also provides interpretable predictions for healthcare analytics.

`DyHealth` is proposed for the next-generation EMR data analytics in NUH to support dynamic modality changes in a flexible and extensible manner during runtime. `DyHealth` is connected to the patients' EMR data from the hospital's division of nephrology for a pilot evaluation. We summarize our main contributions as follows.

- We advocate modularity in the framework design for dynamic healthcare analytics, which has not been considered before. We also propose to support multimodality and provide interpretability simultaneously.
- We devise a modular, multimodal, and interpretable framework `DyHealth` for dynamic healthcare analytics. Different modalities are first handled individually in our dedicated data modules and then fused in the novel Multimodal Fusion Module based on per-input importance for interpretability.
- We evaluate `DyHealth` on acute kidney injury (AKI) prediction in the data from the NUH division of nephrology. We also include two benchmark applications in public datasets in the evaluation. Extensive results validate the effectiveness of `DyHealth` in terms of modularity, multimodality, and interpretability.

Structure. In Section 2, we review related work. We elaborate on the overview of `DyHealth` and its detailed modular design in Section 3 and Section 4, respectively. The experimental evaluation of `DyHealth` is provided in Section 5. Finally, we conclude in Section 6.

2 RELATED WORK

Healthcare analytics utilizes patients' healthcare data, typically EMR data, to conduct analytic tasks including diagnosis [36], prognosis [44], etc. Researchers have investigated several key aspects in healthcare analytics separately, including multimodality and interpretability. There are also some related studies exploring modularity, despite not being on healthcare analytics. Further, several healthcare frameworks have also been proposed to support healthcare analytics.

2.1 Multimodality in Healthcare Analytics

Modality is generally used to characterize how things are experienced or take place, such as language, visual data, or vocal data. By fusing the information from multiple modalities (i.e., multimodality), the complementary information from different modalities can be captured, which generally contributes to more accurate and robust predictions [8].

A number of studies have investigated multimodality [8, 23, 51], and some of them focus on healthcare analytics [19, 49, 50, 67]. Specifically, MMDL is proposed in [49] with a Feedforward Neural Network (FFN) and a Gated Recurrent Unit (GRU) [14] to handle non-temporal and temporal features respectively. Then RAIM [67] is proposed to model ICU patients' multimodal time-series data and capture the particular relationship between two modalities, i.e., the continuous monitoring data should be guided by the discrete clinical data, which renders the two modalities closely coupled and hence, makes RAIM hardly modular. Further, with a guided multichannel attention mechanism, RAIM provides interpretability for continuous monitoring data while leaving out the modality of time-series clinical data. MNN [50] considers textual clinical notes as another modality for analytics, but neither modularity nor interpretability is considered in the model design. Further, DCMN [19] captures both the intra-modal dependencies and inter-modal interactions. However, DCMN models the interactions between merely two modalities in a complex manner and is hence inherently difficult to be modular.

2.2 Interpretability in Healthcare Analytics

Interpretability measures the degree to which the decisions made can be understood by humans [9, 41]. In healthcare analytics, it is essential to take account of interpretability in the model design, as doctors need concrete explanations for deriving insights to help make certain decisions. In recent years, interpretability has attracted increasing attention.

Some existing studies are based on traditional machine learning models. Although these models can achieve certain interpretability, they are typically ineffective in modeling the longitudinal EMR data. With the attention mechanism [7] and recurrent neural networks (RNN), such as the GRU model [14], some recent studies [15, 39, 69] manage to achieve both accurate analytics and interpretability. Specifically, in [39], Dipole is proposed to capture the visit-level importance via three different attention mechanisms to provide interpretable healthcare analytics. Further, RETAIN [15] enhances the interpretability with both the visit-level attention and the variable-level attention via modeling EMR data in reverse time order. TRACER [69] is also developed to capture both time-invariant and time-variant feature importance for interpretations.

2.3 Modularity

Modularity [3] measures the degree to which a framework is divided logically into manageable modules for ease of implementation, upgrade, and maintenance. There are several domains naturally sticking to modularity, such as social networks [62], biology domain including protein networks [22], metabolic networks [52], human brain structures [12, 20], etc. However, these lines of research rely on the intrinsic modular and hierarchical organizational structures in their respective domains, which is different from our aim of modularizing the framework design for analytics.

Relatively fewer efforts have been devoted to developing a flexible and extensible framework for healthcare analytics. In practice, modularity should be one of the major concerns when designing healthcare frameworks for analyzing multimodal EMR data. For one thing, incorporating more modalities into the framework is necessary and essential for better healthcare analytics [19, 49, 50, 67], as different modalities contain complementary information. For another, the availability of each data modality changes constantly after the model deployment, which means that the framework needs to be flexible and extensible at runtime to guarantee the model's normal and smooth functioning when plugging in/out certain modality processing modules. This requires an abstraction of the modality processing and a uniform interface for fusing the information from different data modules, such that the framework can process each modality independently before the multimodal fusion. In `DyHealth`, we follow such system design principles.

2.4 Healthcare Frameworks

There are some frameworks proposed to support healthcare analytics [11, 26, 30, 33, 40, 59, 60, 64, 65, 68, 69]. Among them, some manage to provide interpretable analytic results [11, 26, 30, 33, 40, 60, 64, 68, 69]. However, they do not handle multimodal healthcare data and hence, fail to provide interpretability for multimodal healthcare analytics. Further, these frameworks do not consider modularity in the framework design.

In contrast to existing frameworks, `DyHealth` simultaneously considers all three critical aspects, namely modularity, multimodality and interpretability. It achieves modularity, and multimodality through the system design, which allows to (i) process each modality independently, (ii) integrate the complementary information from different modalities, and (iii) support modality hot-plug in a flexible and extensible way. Besides, `DyHealth` provides fine-grained interpretability for each modality of multimodal EMR data.

3 `DyHealth` FRAMEWORK

An overview of our `DyHealth` framework to equip the hospital with dynamic healthcare analytics is illustrated in Figure 2. In a nutshell, `DyHealth` handles both historical EMR data and daily generated EMR data, processes different modalities in respective data modules, then fuses these modalities in the Multimodal Fusion Module, and finally conducts predictive analytics in the Prediction Module. Through its data processing pipeline, `DyHealth` provides the following core functionalities: (i) it handles multimodal EMR data, (ii) it adapts to dynamic changes in modalities in a plug-and-play manner, and (iii) it provides interpretable prediction results for clinicians.

Data. Modern hospital databases are immutable in nature, in that new data is appended without writing over the old values (e.g., those of chronic diseases). `DyHealth` makes use of historical EMR data to train models for accurate analytics, and utilizes the newly generated EMR data, such as the data of newly admitted patients or newly

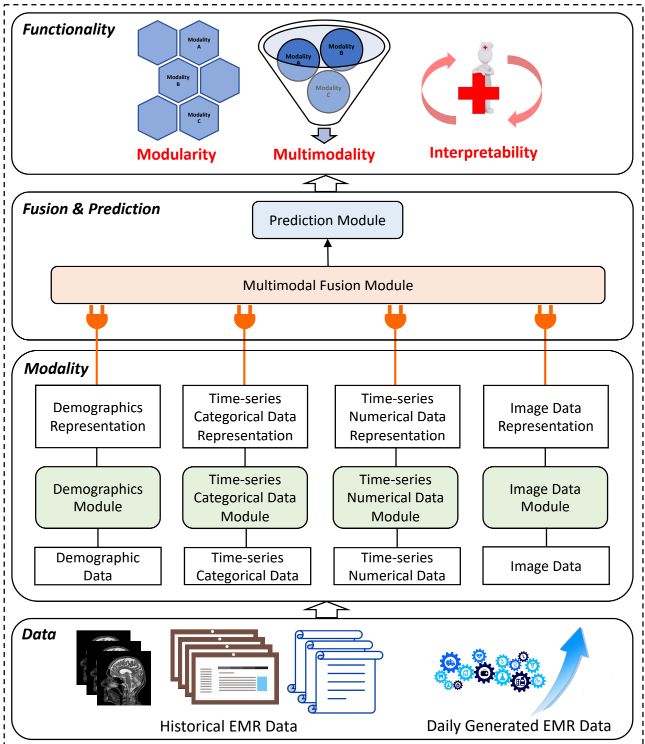

Figure 2: `DyHealth` for dynamic healthcare analytics.

prescribed medications to a hospitalized patient, for inference on a regular basis for patient monitoring and management.

Modality. `DyHealth` is designed to handle different modalities. For each modality, we design a data module for deriving the representation for the subsequent multimodal fusion and prediction. We illustrate four representative data modules in Figure 2 to showcase how `DyHealth` manages different modalities: the Demographics Module, the Time-series Categorical Data Module, the Time-series Numerical Data Module and the Image Data Module. We will elaborate on these data modules in Section 4.1. We note that for other types of modalities, users can define new customized data modules following the modular processing abstraction and the standardized interface, and then readily integrate them into `DyHealth`.

Fusion & Prediction. After obtaining the respective representations from all the data modules, `DyHealth` integrates the complementary information from different modalities in the Multimodal Fusion Module (Section 4.2). It devises a modality-based attention mechanism to dynamically assign the importance of different modalities for each sample. Further, `DyHealth` adopts exponential increasing/decreasing mechanisms for plugging in/out data modules. As a result, `DyHealth` can adapt to dynamic changes of modalities in a plug-and-play manner (Section 4.4). With the integrated representation obtained by fusing different modalities, `DyHealth` supports predictive analytics in the Prediction Module (Section 4.3).

Functionality. Based on the components above, `DyHealth` supports the following key functionalities for healthcare analytics.

- *Multimodality.* With the *Modality* component, `DyHealth` can process different modalities via respective data modules. Specifically, `DyHealth` incorporates a number of representative data modules, i.e., the Demographics Module, the Time-series Categorical Data Module, the Time-series Numerical Data Module, and the Image Data Module for demographics, time-series categorical data, time-series numerical data, and image data, respectively. In this way, the intrinsic characteristics of different modalities can be modeled effectively so that `DyHealth` can utilize the representations from these complementary information sources for more accurate and responsive analytics.
- *Modularity.* Modularity is achieved in `DyHealth` in two aspects. First, `DyHealth` provides a uniform interface for the modality processing of different data modules so that each modality is decoupled from other modalities and can be processed independently. The interface stipulates that each data module extracts a compact representation of a predefined dimension for the corresponding modality, and the representations for different modalities can hence be readily integrated into the subsequent Multimodal Fusion Module to model the cross-modality interactions. Such a standardized interface decouples the design and the implementation of each data module from other modules in the framework, thereby enabling greater flexibility and extensibility in the implementation, upgrade, and maintenance of `DyHealth`. Second, `DyHealth` supports modality hot-plug, i.e., adding or removing data modules at runtime. Specifically, these two changes of data modules at runtime are supported by the exponential increasing and decreasing mechanism respectively in `DyHealth`, which ensures that `DyHealth` adapts to the changes without affecting the performance significantly.
- *Interpretability.* `DyHealth` supports interpretable healthcare analytics at two levels. First, when fusing the information from different modalities in the Multimodal Fusion Module, a modality-based attention mechanism is integrated into `DyHealth` so that the relative importance of different modalities is dynamically modeled for each sample. In this way, `DyHealth` provides personalized interpretation results at the modality level, e.g., "for Patient A, MRI scans are the most dominant information source". Second, within each data module for each modality, more fine-grained interpretation techniques can be incorporated. For instance, in the Time-series Categorical Data Module (Section 4.1.2) and the Time-series Numerical Data Module (Section 4.1.3), the temporal attention with feature differentiation technique can provide more in-depth interpretations both at the time level and at the feature level for each sample. In summary, these interpretation results at different levels of granularity can provide valuable insights about the patients' health conditions, and therefore, assist clinicians in decision-making.

4 MODULAR DESIGN OF `DyHealth`

In this section, we first introduce several data modules for representative modalities, which are used to showcase how `DyHealth` processes different data types and therefore, demonstrate the extensibility of `DyHealth`. We then elaborate on the proposed attention-based Multimodal Fusion Module to integrate the complementary information from different modalities, which is subsequently fed to the Prediction Module. We also discuss how `DyHealth` supports the dynamic modality changes in a plug-and-play manner.

4.1 Representative Data Modules

In DyHealth, we integrate several data modules tailored to representative modalities that are commonly collected in data acquisition and highly informative to healthcare analytics. Essentially, DyHealth converts the raw data into a compact representation in each data module. In the following discussion, we denote scalars, vectors, matrices as x, $\mathbf{x}$, and $\mathbf{X}$, respectively.

4.1.1 Demographics Module. Demographic features in EMR data refer to socio-economic information about patients (such as age, gender, race, etc) and are acknowledged to play a crucial role in healthcare analytics. Given D input demographic features, the i-th feature can be categorized and represented with a one-hot encoded vector $\mathbf{d_i}$. Next, by concatenating these vectors, we have the encoded input vector $\mathbf{d} = [\mathbf{d_1}, \ldots, \mathbf{d_i}, \ldots, \mathbf{d_D}]$ as the raw representation of demographic features. We then transform the input vector $\mathbf{d}$ into an embedding space:

$$\mathbf{O_d} = \phi_d(\mathbf{W_d d}) \tag{1}$$

where $\mathbf{W_d}$ is the learnable transformation matrix, and ϕ_d applies the element-wise rectified linear unit (ReLU) activation function. The embedded representation $\mathbf{O_d}$ is further transformed by:

$$\mathbf{s_d} = \mathbf{W_{sd} O_d} + \mathbf{b_{sd}} \tag{2}$$

where $\mathbf{W_{sd}}$ and $\mathbf{b_{sd}}$ are the learnable weight matrix and bias vector. $\mathbf{s_d}$ is the representation of this data module and will be subsequently fused with other modalities in the Multimodal Fusion Module. The dimensionality of $\mathbf{s_d}$ is predefined by the modality processing interface of DyHealth, which is the same for all the data modules.

4.1.2 Time-series Categorical Data Module. There exist different types of time-series categorical modalities in EMR data, such as diagnoses, procedures, etc, serving as key indicators of patients' health conditions. In DyHealth, we devise a Time-series Categorical Data Module to model such modalities as in Figure 3.

Formally, the time-series categorical data are collected over T_c time steps, and each step obtains F binary features. The features collected at each time step t can be denoted as $\mathbf{c_t} = [c_{t,1}, \ldots, c_{t,f}, \ldots, c_{t,F}]$, where $c_{t,f} = 1$ indicates the presence of the f-th binary feature at time t, e.g., whether a particular diagnosis is made. We then transform $\mathbf{c_t}$ into an embedding space:

$$\mathbf{q_t} = \phi_c(\mathbf{W_c c_t}) \tag{3}$$

where ϕ_c is the ReLU activation function, and $\mathbf{W_c}$ is the learnable embedding matrix for time-series categorical data.

Bidirectional RNN. We then feed all the T_c embedded vectors into a bidirectional RNN model [56] to capture the temporal dynamics of the time-series data:

$$(\mathbf{h_1^{(c)}}, \ldots, \mathbf{h_t^{(c)}}, \ldots, \mathbf{h_{T_c}^{(c)}}) = BIRNN(\mathbf{q_1}, \ldots, \mathbf{q_t}, \ldots, \mathbf{q_{T_c}}) \tag{4}$$

where $BIRNN(\cdot)$ is a bidirectional GRU model [56]. Specifically, the hidden representation $\mathbf{h_t^{(c)}}$ is the concatenation of $\overrightarrow{\mathbf{h_t^{(c)}}}$ and $\overleftarrow{\mathbf{h_t^{(c)}}}$, where $\overrightarrow{\mathbf{h_t^{(c)}}}$ is calculated from $\mathbf{q_1}$ to $\mathbf{q_t}$ through a forward GRU model, and $\overleftarrow{\mathbf{h_t^{(c)}}}$ from $\mathbf{q_{T_c}}$ to $\mathbf{q_t}$ through a backward GRU model. The bidirectional RNN model is adopted in this data module in order to obtain a more comprehensive representation of a patient's time-series EMR data. This is achieved due to the capability of bidirectional RNN in modeling time-series dynamics from both directions, which imitates how clinicians examine a patient's EMR data to analyze the patient's health conditions.

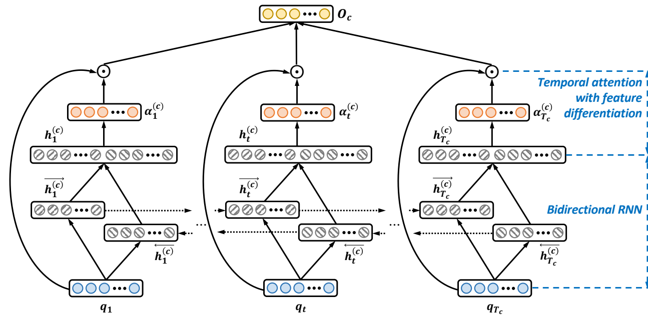

Figure 3: Time-series Categorical Data Module of DyHealth.

Temporal attention with feature differentiation. Further, in order to differentiate the varying influence of medical features at different time steps, we devise a self-attention mechanism that has been shown to support similar tasks well [13, 66]. In particular, our attention mechanism can give the importance of each time step, and meanwhile, differentiate the influence of different features at each time step. We calculate the temporal attention vector $\boldsymbol{\alpha}_\mathbf{t}^{(\mathbf{c})}$:

$$\boldsymbol{\alpha}_\mathbf{t}^{(\mathbf{c})} = \tanh(\mathbf{W}_{\boldsymbol{\alpha}}^{(\mathbf{c})} \mathbf{h_t^{(c)}} + \mathbf{b}_{\boldsymbol{\alpha}}^{(\mathbf{c})}) \tag{5}$$

where $\boldsymbol{\alpha}_\mathbf{t}^{(\mathbf{c})}$ denotes the influence of each medical feature at each time step, which thus supports both temporal and feature-wise interpretability. With $\boldsymbol{\alpha}_\mathbf{t}^{(\mathbf{c})}$, we can aggregate the outputs $\mathbf{q_t}$ from bidirectional RNN at each time step to calculate a compact representation $\mathbf{O_c}$ of the time-series categorical modality:

$$\mathbf{O_c} = \sum_{t=1}^{T_c} \boldsymbol{\alpha}_\mathbf{t}^{(\mathbf{c})} \odot \mathbf{q_t} \tag{6}$$

where $\odot$ denotes Hadamard product, and $\mathbf{O_c}$ summarizes the information from all the T_c time steps. We then transform $\mathbf{O_c}$ to the final representation of the predefined dimension:

$$\mathbf{s_c} = \mathbf{W_{sc} O_c} + \mathbf{b_{sc}} \tag{7}$$

where $\mathbf{W_{sc}}$ and $\mathbf{b_{sc}}$ are the learnable weight matrix and bias vector. $\mathbf{s_c}$ is the representation of this data module and will be forwarded to the Multimodal Fusion Module.

4.1.3 Time-series Numerical Data Module. Time-series numerical modalities constitute an essential part of EMR data, including lab tests, prescribed medications with the recommended dosages, etc. In DyHealth, we design a Time-series Numerical Data Module to incorporate the information of such modalities.

Given G numerical features collected over T_n time steps, the numerical features at each time step can be represented as $\mathbf{n_t} = [n_{t,1}, \ldots, n_{t,g}, \ldots, n_{t,G}]$, where $n_{t,g}$ records the g-th feature value at time t. Then, we can obtain the time-series numerical data $\mathbf{n_t}$ for $t = 1, 2, \ldots, T_n$. In a way similar to the Time-series Categorical Data Module in Figure 3, the time-series representation $\mathbf{n_t}$ is first fed into

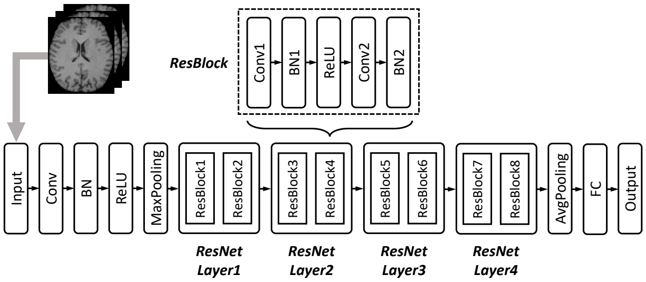

Figure 4: Image Data Module of DyHealth. Conv: convolutional layer, BN: batch normalization, MaxPooling: max pooling layer, ResBlock: residual block, AvgPooling: average pooling layer, FC: fully connected layer.

a bidirectional RNN model for modeling the temporal dynamics:

$$(\mathbf{h}_1^{(\mathbf{n})}, \dots, \mathbf{h}_\mathbf{t}^{(\mathbf{n})}, \dots, \mathbf{h}_{\mathbf{T_n}}^{(\mathbf{n})}) = BIRNN(\mathbf{n_1}, \dots, \mathbf{n_t}, \dots, \mathbf{n_{T_n}}) \tag{8}$$

where $BIRNN(\cdot)$ is a bidirectional GRU, and $\mathbf{h}_\mathbf{t}^{(\mathbf{n})} = [\overrightarrow{\mathbf{h}_\mathbf{t}^{(\mathbf{n})}}; \overleftarrow{\mathbf{h}_\mathbf{t}^{(\mathbf{n})}}]$ is the concatenation of the calculated representations from both directions likewise.

With the representation $\mathbf{h}_\mathbf{t}^{(\mathbf{n})}$, we further differentiate the importance of different numerical features at different time steps, via the temporal attention to obtain fine-grained interpretations. Then we calculate the representation of this time-series numerical modality in a similar way as in the Time-series Categorical Data Module:

$$\boldsymbol{\alpha}_\mathbf{t}^{(\mathbf{n})} = \tanh(\mathbf{W}_{\boldsymbol{\alpha}}^{(\mathbf{n})} \mathbf{h}_\mathbf{t}^{(\mathbf{n})} + \mathbf{b}_{\boldsymbol{\alpha}}^{(\mathbf{n})}) \tag{9}$$

$$\mathbf{O_n} = \sum_{t=1}^{T_n} \boldsymbol{\alpha}_\mathbf{t}^{(\mathbf{n})} \odot \mathbf{n_t} \tag{10}$$

where $\boldsymbol{\alpha}_\mathbf{t}^{(\mathbf{n})}$ indicates each feature's influence at each time step and contributes to fine-grained interpretability in DyHealth. We further transform the derived representation $\mathbf{O_n}$ into a final representation for this modality (with the weight matrix $\mathbf{W_{sn}}$ and the bias vector $\mathbf{b_{sn}}$) as follows:

$$\mathbf{s_n} = \mathbf{W_{sn}} \mathbf{O_n} + \mathbf{b_{sn}} \tag{11}$$

The representation $\mathbf{s_n}$ follows the dimension predefined by the modality processing interface of DyHealth and will be integrated with other modalities in the Multimodal Fusion Module.

4.1.4 Image Data Module. EMR data contain medical image data such as MRI scans, CT scans that are generally 3D scans. Such image modalities are useful for medical diagnosis in that they can help detect organ abnormality of patients and hence, facilitate diagnostic medicine and biomedical research.

In DyHealth, we integrate 3D-ResNet [24], a state-of-the-art model for 3-dimensional data, into our Image Data Module for processing 3D medical images. The detailed architecture of this module is shown in Figure 4. The input 3D image is first transformed by a convolutional layer, which is followed by batch normalization, ReLU, and a max-pooling layer sequentially. Then, the transformed representation is further processed by four consecutive ResNet layers. Each ResNet layer has two residual blocks with shortcut connections that help signals flow to the next layer for residual learning and ease of training. Next, after an average pooling layer,

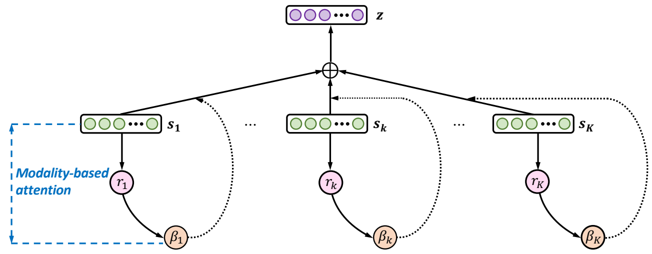

Figure 5: Multimodal Fusion Module of DyHealth.

the original 3D medical image is transformed into a compact representation $\mathbf{O_v}$. We then feed $\mathbf{O_v}$ into a fully connected layer for an affine transformation to obtain the final representation $\mathbf{s_v}$. We note that $\mathbf{s_v}$ is of the same dimension predefined by the data module interface of DyHealth, and this $\mathbf{s_v}$ will interact with other modalities in the Multimodal Fusion Module.

4.2 Multimodal Fusion Module

As discussed in Section 4.1, we can obtain a compact representation of each modality from respective data modules, i.e., $\mathbf{s_d}$, $\mathbf{s_c}$, $\mathbf{s_n}$ and $\mathbf{s_v}$ that are of the same dimension defined by the data module interface of DyHealth. We note that the standardization of the data module dimension is necessary to achieve flexibility and extensibility in DyHealth. These modalities provide complementary information and hence, can be integrated for a more comprehensive view of the available EMR data. Further, for a given sample, the modalities are not equally important, and the relative importance of different modalities varies from one sample to another. Therefore, when fusing the information from different modalities, we need a modality-based attention mechanism to model the importance of each modality in a per-instance manner. To this end, we design the Multimodal Fusion Module for DyHealth as illustrated in Figure 5 and its detailed design is introduced as follows.

Modality-based attention. Suppose there are K modalities to fuse, we denote their representations derived from the corresponding data modules as $\{\mathbf{s_1}, \dots, \mathbf{s_k}, \dots, \mathbf{s_K}\}$. Then, the proposed multimodal fusion with modality-based attention computes the integrated data representation $\mathbf{z}$:

$$\mathbf{z} = \tanh(\sum_{k=1}^{K} \beta_k \mathbf{s_k} + b) \tag{12}$$

where $\mathbf{s_k}$ is the compact representation of the k-th modality, and the modality-based attention weights $\boldsymbol{\beta}$ (i.e., $[\beta_1, \dots, \beta_k, \dots, \beta_K]$) are computed as follows:

$$[\beta_1, \dots, \beta_k, \dots, \beta_K] = \text{softmax}([r_1, \dots, r_k, \dots, r_K]) \tag{13}$$

where r_k is obtained by:

$$r_k = \mathbf{w}_{\boldsymbol{\beta}}^\top \mathbf{s_k} + b_\beta \tag{14}$$

Based on such a modality fusion mechanism, the advantages of the Multimodal Fusion Module are threefold. First, the cross-modality interactions only need to be handled within this module instead of preceding data modules. Such a functionality division renders DyHealth modular, flexible and extensible during runtime, which will be further discussed in Section 4.4. Second, the Multimodal Fusion Module integrates the information from different modalities, and meanwhile learns the varying importance of different modalities for each sample. Hence, this module provides a more comprehensive representation taking account of all the collected modalities of EMR data, which contributes to more effective analytics. Third, with the modality-based attention, the Multimodal Fusion Module is able to support fine-grained interpretability in a per-instance manner, e.g., showing which modality is more prominent for the given instance. Such interpretation results are valuable for the reference of clinicians on patient management, etc.

4.3 Prediction Module

After fusing the information from multimodal EMR data, we feed the integrated representation into the Prediction Module for predictive analytics. Specifically, we feed the integrated data $\mathbf{z}$ obtained in the Multimodal Fusion Module into a multilayer perceptron (MLP) model: $\hat{\mathbf{y}} = \Psi_{MLP}(\mathbf{z})$, where the prediction labels $\hat{\mathbf{y}}$ is a m-dimension vector for an m-class classification task and a scalar for a regression task. In the case of binary classification, the loss function is the cross-entropy loss (CEL), while in the case of regression, the loss function is the mean squared error (MSE).

With the loss function specified, we can train the model parameters of `DyHealth` in an end-to-end manner, including the parameters of all the data modules, the Multimodal Fusion Module, and the Prediction Module, via gradient-based optimizers.

4.4 `DyHealth`'s Support for Modality Hot-plug

We further elaborate on how `DyHealth` supports modularity. As shown in Figure 2, in a practical application scenario that involves multiple modalities, we design a uniform modality processing interface, where each modality is processed independently within its own data module for extracting intra-modality representation. After the processing of each data module, the extracted and standardized representations for all modalities can be readily integrated into the Multimodal Fusion Module to model the inter-modality correlation. As a result of such a processing pipeline, the detailed processing of each modality can be decoupled from the framework, which thus provides modularity for `DyHealth`.

In real-world healthcare delivery, the availability of different modalities changes constantly, and the deployed models need to provide uninterrupted service for life-and-death medical decisions, e.g., to predict the health conditions of a hospitalized patient for taking timely interventions. As a consequence, it is necessary and imperative to support modality hot-plug in the framework, i.e., plugging in a new modality or plugging out an existing modality. Such flexibility and extensibility at runtime enable the framework to adapt to dynamic modality changes without a drastic decrease in performance. `DyHealth` achieves this by adopting an *exponential increasing mechanism* for plugging in a new modality, and an *exponential decreasing mechanism* for plugging out an existing modality.

Exponential increasing mechanism to plug in a modality. Suppose a new modality $\mathbf{s_{K+1}}$ arrives, for instance, the hospital employs some MRI scanners to collect patients' MRI images for better analytics. We design an exponential increasing mechanism for such a scenario. Specifically, we introduce an increasing multiplier $\tau(\cdot)$ to gradually increase the importance of $\mathbf{s_{K+1}}$'s representation when fusing different modalities in Equation 12. The mechanism ensures that $\mathbf{s_{K+1}}$ can be integrated smoothly without causing a sudden and drastic change in `DyHealth`'s performance. $\tau(\cdot)$ is calculated as:

$$\tau(H) = \frac{\exp(\frac{H}{H_{max}}) - 1}{e - 1} \tag{15}$$

where H is the number of iterations. This means that the new modality $\mathbf{s_{K+1}}$ is not entirely plugged into `DyHealth` until after H_{max} iterations. The integration of $\mathbf{s_{K+1}}$ starts from $H = 0$ ($\tau(H) = 0$) to $H = H_{max}$ ($\tau(H) = 1$).

Exponential decreasing mechanism to plug out a modality. Similarly, when an existing modality $\mathbf{s_k}$ is going to be plugged out of `DyHealth`, we multiply the decreasing multiplier $\lambda(\cdot)$ to the representation of $\mathbf{s_k}$ in Equation 12. The $\lambda(\cdot)$ gradually decreases the influence of the modality to be removed for iterations $H = \{0, 1, \ldots, H_{max}\}$:

$$\lambda(H) = \frac{\exp(\frac{H_{max}-H}{H_{max}}) - 1}{e - 1} \tag{16}$$

Therefore, the modality is not entirely plugged out until after H_{max} iterations. With such a gradual decreasing mechanism, `DyHealth` can still function properly with only a small decrease in performance when plugging out certain modalities at runtime.

We note that in these two modality change scenarios, `DyHealth` can readily support plugging in a new modality $\mathbf{s_{K+1}}$ or plugging out an existing modality $\mathbf{s_k}$ in the Multimodal Fusion Module (Equation 13) without any further changes. Therefore, with the proposed exponential increasing/decreasing mechanisms, `DyHealth` can readily adapt to dynamic changes in modalities during runtime.

5 EXPERIMENTS

We demonstrate the experimental results of `DyHealth` in the pilot evaluation, which is conducted in the NUH division of nephrology, specifically on the hospital-acquired AKI prediction application. In this evaluation, clinicians validate and verify `DyHealth`'s interpretation results, and in turn, `DyHealth` assists clinicians in understanding why a certain patient develops AKI. As the NUH patients' EMR data (i.e., **NUH dataset**) are highly private, we further adopt two other public and popular benchmark EMR datasets, **MIMIC-III dataset** [28] and **IXI dataset** [2] for evaluation, with the former focusing on evaluating modularity due to its larger number of modalities than the other datasets, and the latter focusing on evaluating `DyHealth`'s performance on medical image modalities.

5.1 Hospital-acquired AKI Prediction

AKI, short for acute kidney injury, develops in 4% of patients who are admitted to NUH and involves over 3000 patients annually [37]. AKI generally indicates a poor prognosis of patients in terms of prolonged hospitalization, sustained kidney function deterioration, and even a significant risk of kidney failure and mortality in the long run [16, 25]. For patients with a high risk of developing AKI, existing management measures may reduce the incidence of the disease or its downstream complications, even if AKI still develops. Nonetheless, such strategies must be promptly implemented, which requires AKI to be diagnosed in the subclinical phase, i.e., way

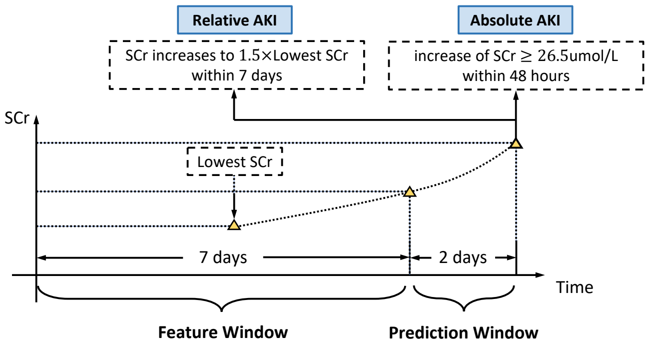

Figure 6: Detection criteria of hospital-acquired AKI.

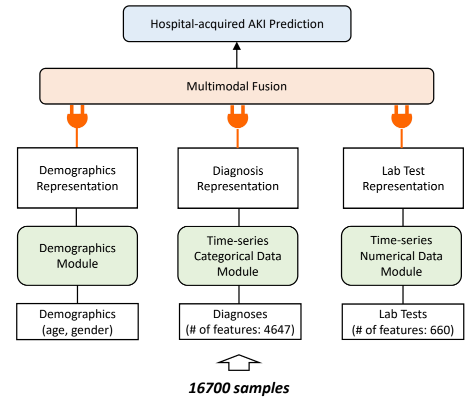

Figure 7: `DyHealth` for hospital-acquired AKI prediction.

before the onset of AKI. As a consequence, AKI prediction is significantly motivating and valuable for the NUH division of nephrology to take preemptive measures for optimized patient management.

We use the **NUH dataset** that records over 100,000 patients' EMR data including diagnoses, lab tests, etc, to predict if a hospitalized patient will develop AKI in the current admission, i.e., the hospital-acquired AKI prediction. The medical definition of AKI is given by the KDIGO clinical practice guideline [31]. Specifically, AKI is defined based on a lab test serum creatinine (SCr). When there exists an increase in SCr of $26.5umol/L$ or more within 48 hours, absolute AKI is detected; when there is an increase in SCr by 1.5 times baseline (i.e., the lowest SCr value) or higher within the last 7 days, relative AKI is detected. The detection criteria of both absolute AKI and relative AKI are illustrated in Figure 6. For each hospitalized admission, we check both criteria to derive the corresponding label, and satisfying either criterion corresponds to a positive label.

Further, if a hospitalized admission has a positive label, i.e., AKI is developed in this admission, we record the corresponding AKI detection time, trace two days back as *"Prediction Window"* (not used as input) and trace seven more days back as *"Feature Window"* (used as input). Otherwise, if an admission is negative, we take the time when the latest medical feature appears in the corresponding patient's EMR data, and set it as the end of Prediction Window to derive Feature Window accordingly. The relationship between Feature Window and Prediction Window is shown in Figure 6. With such settings, we can predict hospital-acquired AKI by a two-day lead time, with `DyHealth` serving as an AKI surveillance framework that is deployable in real time. Moreover, such a lead time is necessary for AKI preventive strategies to make a meaningful impact in clinical outcomes when implemented in a timely manner, which contributes to reducing the AKI duration and hospital days in affected patients [32, 57].

5.1.1 Experimental Set-up.

Data Preprocessing. We utilize the patients' demographics (age and gender), time-series diagnoses, and time-series lab tests in the NUH dataset as input. For non-time series features of demographics, we first conduct binning on the two features, convert the value of each feature to a one-hot encoding representation, concatenate the representations and then conduct an embedding for this representation. As illustrated in Figure 7, such demographics information is processed in `DyHealth` via a Demographics Module.

For time-series medical features, we divide Feature Window with a length of seven days into seven *"Time Windows"*, aggregate the time-series features within each Time Window as input, and predict if the patient will develop AKI in this admission in two days. The time-series diagnoses and lab tests are handled by `DyHealth` through a Time-series Categorical Data Module, and a Time-series Numerical Data Module respectively as in Figure 7. We note that for the lab test with a numerical value x, we conduct a min-max normalization, i.e., $x' = (x - min)/(max - min)$ and then use the normalized feature values for further analytics.

In the NUH dataset, we have 16700 samples with all three modalities available. The feature number of each modality and the architecture of `DyHealth` for the hospital-acquired AKI prediction are illustrated in Figure 7.

Baseline Methods. We first compare `DyHealth` with several state-of-the-art methods in time-series healthcare analytics.

- **RETAIN** [15] proposes a two-level attention mechanism based on a reverse time attention model structure, to achieve improved interpretability for healthcare analytics.
- **Dipole** [39] devises an attention mechanism based on a bidirectional RNN model to achieve the visit-level interpretability. Specifically, there are three attention mechanisms as follows.
- **Dipole**$_{loc}$ uses a location-based attention mechanism to capture the attention weights solely based on the current hidden state.
- **Dipole**$_{gen}$ uses a general attention mechanism to capture the relationship between the current hidden state and each previous hidden state, via learning a weight matrix.
- **Dipole**$_{con}$ uses a concatenation-based attention mechanism to capture the relationship between the current hidden state and each previous hidden state, via learning the attention weights based on their concatenation.
- **TRACER** [69] employs a feature-wise transformation subnetwork and a self-attention subnetwork to capture the time-invariant and the time-variant feature importance respectively for interpretable healthcare analytics.

In these baseline methods, we concatenate the medical features of different modalities as input in each Time Window. Apart from these baselines, we also compare `DyHealth` with the following modality fusion methods to validate its effectiveness in integrating the complementary information from different modalities.

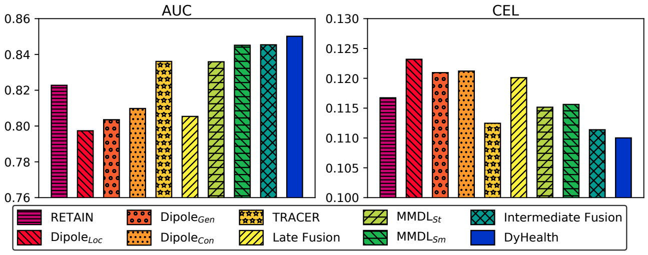

Figure 8: Comparison results for AKI prediction.

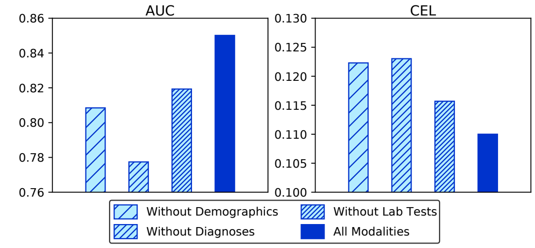

Figure 9: Ablation study results for AKI prediction.

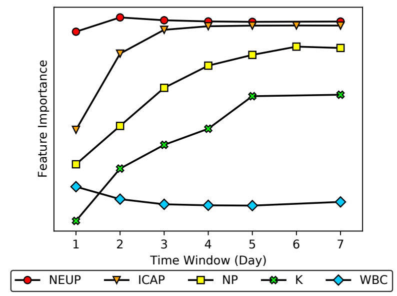

Figure 10: Interpretation results of an example patient who developed AKI in the NUH dataset.

- **Late Fusion** first calculates the final prediction results of each modality separately, i.e., the probability distribution for classification applications, or the value of the predicted target for regression applications. It then averages the prediction results of different modalities to fuse the complementary information.
- **MMDL** [49], short for multimodal deep learning model, uses an ensemble of feedforward network (FFN) and GRU to process non-temporal features and temporal features separately and then learns a shared representation from different modalities. Two variants of MMDL are discussed as follows.
- **MMDL$_{St}$** first handles the non-temporal features in an FFN model. It then concatenates the temporal features from different time-series modalities and further learns the temporal dynamic behavior with a GRU model. Finally, it combines the learned representation from the FFN model and that from the GRU model into a shared latent representation for prediction.
- **MMDL$_{Sm}$** processes the non-temporal features in an FFN model, similar to MMDL$_{St}$. However, MMDL$_{Sm}$ uses different GRU models for different temporal modalities separately. Next, it fuses the learned representations from the FFN model and different GRU models to derive the shared representation.
- **Intermediate Fusion** integrates different modalities in an unweighted manner, i.e., in Equation 12 when fusing the extracted representations of different modalities, the modality-based attention weights $\boldsymbol{\beta}$ are not taken into account.

Evaluation Metrics. The hospital-acquired AKI prediction is formulated as a binary classification problem. We evaluate the effectiveness of `DyHealth` in terms of the area under the ROC curve (AUC) and the average CEL per sample. A better-performing classifier should achieve a higher AUC value and a lower CEL value. In the experiments, we partition the dataset into the following parts: 80% for training, 10% for validation, and 10% for testing. Based on the best performance achieved on the validation data, we choose the hyperparameter setting, apply the model to testing data and report the average AUC value and CEL value on the testing data of three different repeats.

5.1.2 *Experimental Evaluation.*

Comparison with Baseline Methods. We demonstrate the experimental results of `DyHealth` compared with the baselines in Figure 8. `DyHealth` achieves the best performance with the highest AUC and the lowest CEL, which confirms the effectiveness of `DyHealth` in making use of multimodal EMR data.

Specifically, compared with RETAIN, the three variants of Dipole (among which Dipole$_{con}$ outperforms the other two) and TRACER, `DyHealth` can provide more accurate analytics. The superior performance of `DyHealth` can be attributed to its modular design for processing each modality and the integration of the complementary information from different modalities.

Analyzing the performance of different fusion methods, we have the following observations. First, MMDL and Intermediate Fusion outperform Late Fusion, since Late Fusion only fuses the final predictions, and thus lacks the capability to exploit the cross-modality interactions effectively. Further, for the two variants of MMDL, MMDL$_{Sm}$ achieves a higher AUC than MMDL$_{St}$, and performs similarly to Intermediate Fusion in terms of AUC. This indicates that processing different time-series modalities separately as in MMDL$_{Sm}$ can improve the analytic performance. Finally, compared with Intermediate Fusion, `DyHealth` is more accurate, which validates the effectiveness of `DyHealth` in modeling the varying importance of modalities on a per-input basis.

Ablation Study. We next conduct an ablation study of `DyHealth` in terms of multimodality modeling. The experimental results are shown in Figure 9. We find that among the three modalities, the diagnosis modality is the most informative one, the removal of which leads to the largest decrease in AUC. This confirms the crucial importance of clinicians' judgment on patients' health conditions. Besides diagnoses, the demographics modality is also critical to the performance, which means that the patients' intrinsic characteristics play an essential role in the prediction. Finally, the importance of the lab test modality is relatively lower than the other two. This indicates that the underlying information of lab tests may be partially carried by diagnoses which therefore can be considered a surrogate data source for lab tests.

5.1.3 *Interpretability.* With clinicians validating and verifying the interpretation results provided by `DyHealth`, we demonstrate the interpretability of `DyHealth`. Specifically, we show `DyHealth`'s interpretations for a patient who developed AKI after two days in

Figure 10. The x-axis denotes each Time Window in Feature Window as input, and the y-axis "Feature Importance" corresponds to $\boldsymbol{\alpha}_{\mathbf{t}}^{(\mathbf{n})}$ in Time-Series Numerical Data Module for processing the lab tests (in Equation 9). The features involved in Figure 10 are: "Neutrophils %" (NEUP), "Ionised CA, POCT" (ICAP), "Sodium, POCT" (NP), "Serum Potassium" (K), and "White Blood Cell" (WBC).

As in Figure 10, we find that NEUP shows the highest Feature Importance among all features, and WBC shows relatively stable importance across time. These suggest infection or inflammation that directly contributes to AKI causation. Further, when we investigate K, NP and ICAP, we find these three kinds of ionised electrolytes show similarly increasing changing patterns in terms of their importance over time. K and NP are important electrolytes in the human body, which are vitally important to cellular metabolism especially K, while NP levels reflect the free water balance, and these are regulated by the kidneys [10]. Besides, ICAP is related to hypocalcemia, which is in evolving kidney dysfunction and may be related to disordered mineral metabolism in kidney disease and hyperphosphatemia [34]. Hence, the increasing patterns of these features indicate the presence of the electrolyte and water imbalance.

In a nutshell, based on the interpretation results, we presume this patient has worsening inflammation or infection that contributes to a high risk of developing AKI; evolving kidney injury is accompanied by worsening serum electrolyte and water imbalance, which therefore explains the predictive performance.

5.1.4 Summary. Based on this pilot evaluation in the NUH division of nephrology, DyHealth achieves an AUC exceeding 0.85 on AKI prediction with different modalities modeled in a modular manner, which is validated to be medically effective in clinical practice. With such accurate predictive analytics, DyHealth enables the risk stratification of hospitalized patients for detailed biomarker or clinical assessment on the true AKI risk. As for interpretability, DyHealth unveils valuable reference information on a patient's underlying problems and insightful patient-specific trends via its interpretations and hence, aids the evaluation of AKI etiology and guides clinicians in NUH to take timely interventions accordingly. In the near future, we expect the promising roll-out of DyHealth for hospital-wide real-time deployment to highly facilitate our upstream endeavors to prevent AKI and/or its complications in hospitalized patients.

5.2 In-hospital Mortality Prediction

In this application, we make use of the **MIMIC-III dataset** [28], a public EMR dataset that records the data of more than 40,000 patients in critical care units. In this dataset, we investigate the in-hospital mortality prediction, i.e., to predict if a patient will pass away in the current admission. Specifically, each admission corresponds to one visit of a patient and the admissions lasting over 48 hours are selected as samples.

5.2.1 Experimental Set-up.

Data Preprocessing. We take into account the non-time series patients' demographics information and the time-series lab tests, output events, input events, and prescriptions, which are all in the structured form. The demographic features include age, gender and as illustrated in Figure 11, they are processed by DyHealth via a Demographics Module.

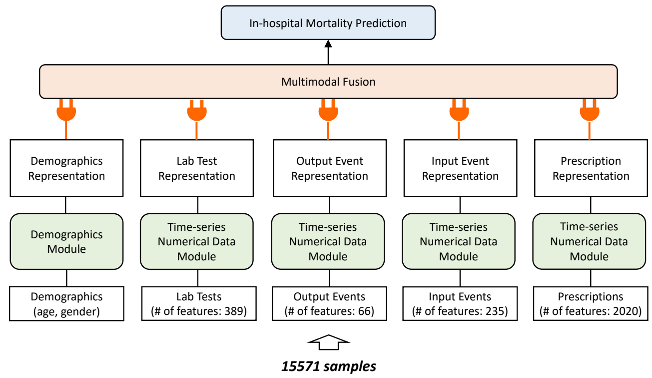

Figure 11: DyHealth for in-hospital mortality prediction.

For the other modalities, we adopt a 48-hour Feature Window with a 2-hour Time Window. We then aggregate the time-series features of each modality within each Time Window as input and predict if the patient will pass away in the hospital.

Specifically, lab tests refer to the lab events in the MIMIC-III dataset, recording laboratory-based measurements. As lab tests indicate the patients' health conditions, we average the values of the same lab test appearing in each Time Window and then, conduct a min-max normalization of the values. We then configure a Time-series Numerical Data Module to process lab tests.

Input events and output events form a crucial information source for analyzing the patients in ICU. Input events correspond to the fluids administered to the patients including oral/tube feedings, intravenous solutions with medications, etc, while output events refer to the fluids excreted from the patients such as urine output, etc. Both modalities represent an accumulated feature of the patients' human body; therefore, the values of the same feature in each Time Window are then aggregated by sum. Then these two modalities are both handled by DyHealth through a Time-series Numerical Data Module as shown in Figure 11.

Prescriptions record medications prescribed to the patients. As an accumulated type of features given over a time period, we also sum the feature values within each Time Window, and configure a Time-series Numerical Data Module for prescriptions.

The architecture of DyHealth for in-hospital mortality is illustrated in Figure 11. We extract 15571 samples with all five modalities available. In this application, we compare DyHealth with the same set of baseline methods as in Section 5.1, as both applications are predictive analytics based mainly on time-series EMR data. Besides, because this application is formalized as a binary classification problem, we also adopt AUC and CEL as the evaluation metrics.

5.2.2 Experimental Evaluation.

Comparison with Baseline Methods. As illustrated in Figure 12, we find that DyHealth achieves the best-performing analytics among all methods in terms of both AUC and CEL. To be specific, DyHealth outperforms RETAIN, Dipole, and TRACER due to the ability to process each individual modality in a modular and effective manner. Further, DyHealth is more accurate than Late Fusion, MMDL

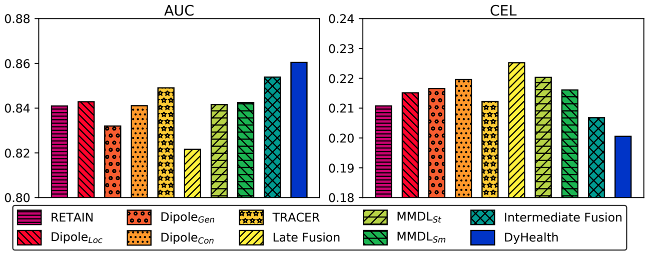

Figure 12: Comparison results for mortality prediction.

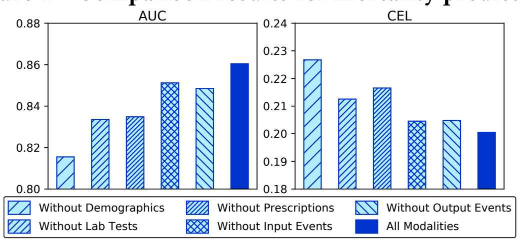

Figure 13: Ablation study results for mortality prediction.

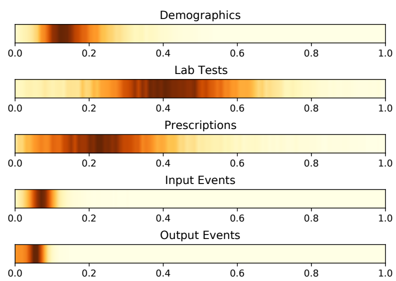

Figure 14: The distribution of $\boldsymbol{\beta}$ for different modalities over all samples for in-hospital mortality prediction.

with two variants and Intermediate Fusion, as a result of its Multimodal Fusion Module for integrating various modalities in an attention-based manner.

Ablation Study. We remove one modality from input each time in order to evaluate its influence on the overall performance. The experimental results are illustrated in Figure 13.

We find that the demographics modality has the most significant influence on the mortality prediction, which indicates that the mortality of a hospitalized patient is highly related to his/her personal characteristics. Next, both lab tests and prescriptions are more important as compared with input events and output events. We conjecture that lab tests and prescriptions directly represent patients' health conditions and reflect clinicians' judgment; hence, these two modalities tend to be more informative than other implicit modalities, e.g., input events and output events.

5.2.3 Modularity. Since the MIMIC-III dataset has a larger number of modalities than the other two datasets, we evaluate the modularity of DyHealth in this dataset.

To start with, we illustrate the distribution of $\boldsymbol{\beta}$ (calculated in Equation 13) for the five involved modalities in the MIMIC-III dataset over all samples in Figure 14. The experimental results demonstrate that our proposed modality-based attention mechanism can reflect the relative importance of various modalities on a per-sample basis. Further, when all modalities are incorporated

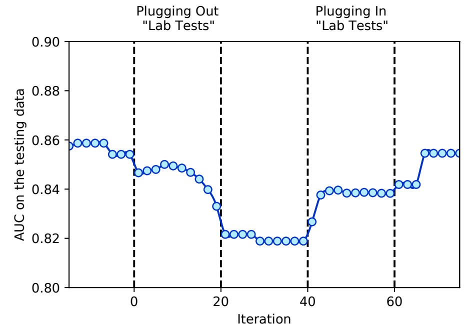

Figure 15: Modularity experimental results w.r.t. lab tests in the MIMIC-III dataset.

in DyHealth for analytics, the relative importance from higher to lower tends to be: lab tests, prescriptions, demographics, input events, and output events.

We also design an experiment to demonstrate how DyHealth supports modality hot-plug by first plugging out the lab test data module and then plugging this data module back in. The experimental results are illustrated in Figure 15, in which the x-axis denotes the iterations, and the y-axis is the corresponding AUC achieved on the testing data. Specifically, we start with DyHealth incorporating all modalities with an AUC of around 0.86. To keep the deployed framework up to date, we regularly train the parameters of DyHealth on the training data available at each iteration. At Iteration 0, we start plugging out the lab test data module via the proposed exponential decreasing mechanism with $H_{max} = 20$. At Iteration 20, the lab test data module is totally plugged out with AUC decreased to around 0.82. Next, DyHealth operates with the remaining four modalities for another 20 iterations. At Iteration 40, we start plugging in the lab test data module back to DyHealth via the proposed exponential increasing mechanism. The plug-in process is completed at Iteration 60 (with $H_{max} = 20$), after which DyHealth continues to be trained with all the five modalities in the training dataset. As shown in Figure 15, we can observe that during the plug-out process (Iterations $0 - 20$), DyHealth still functions properly and its performance decreases smoothly, and during the plug-in process (Iterations $40 - 60$), DyHealth does not undergo sudden changes. These findings confirm the efficacy of DyHealth in supporting modality hot-plug with our proposed exponential increasing/decreasing mechanisms and further validate the modularity of DyHealth.

5.3 Brain Age Prediction

Aging tends to have a significant influence on the brain [29] that is generally visible in MRI scans. In recent years, it is demonstrated that MRI scans can be used to predict the chronological age accurately [21], which is considered as an estimation of the biological brain age. Brain age prediction is to predict a person's chronological age based on his/her brain data. An accurate brain age prediction model is of high medical significance in that it provides a way of estimating biological brain age and therefore, can be used to discover possible diseases and genetic factors related to abnormal brain aging.

In this brain age prediction application, we use the Information eXtraction from Images (IXI), i.e., **IXI dataset** [2]. The IXI dataset

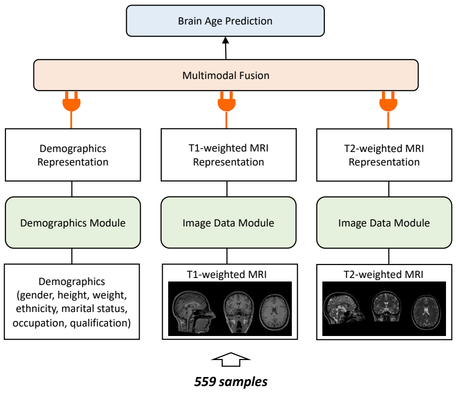

Figure 16: DyHealth for brain age prediction. For both T1-weighted MRI and T2-weighted MRI that are 3D medical images, three views of the MRI scan are shown (from left to right): sagittal, coronal and axial.

collects about 600 magnetic resonance images of healthy and normal subjects, recording their different types of MRI data, such as T1-weighted MRI and T2-weighted MRI, demographics, etc.

5.3.1 Experimental Set-up.

Data Preprocessing. We incorporate the demographics, T1-weighted MRI, and T2-weighted MRI of the subjects as input, and aim to predict the chronological age of the subjects.

The demographics include each subject's gender, height, weight, ethnicity, marital status, occupation, and qualification. For height and weight, we first conduct binning on the values and then transform each feature into the corresponding one-hot encoding vector. For other demographic features that are categorical, we convert them into one-hot representation as well. We next concatenate the representations of all the seven features above and configure a Demographics Module in DyHealth for processing (as in Figure 16).

For MRI images, we conduct the following standard preprocessing workflow for medical brain image data: (i) skull stripping based on a pre-trained model of UNET [54], (ii) N4 bias field correction [63], (iii) template registration [27], and (iv) voxel intensity normalization via a Gaussian mixture model. After such standard preprocessing of MRI images, we configure two Image Data Modules for T1-weighted MRI and T2-weighted MRI respectively.

The architecture of DyHealth for brain age prediction is shown in Figure 16. We have 559 samples with all the modalities available.

Evaluation Metrics. In brain age prediction, we compare DyHealth with Intermediate Fusion and Late Fusion. The other baselines are not compared as they are not applicable to processing image data. Since the goal is to predict the chronological age of the subjects, the problem is formulated as a regression task, with the loss function of MSE. In the experimental evaluation, we adopt mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R^2) as evaluation metrics, and report their values on the testing data. A better-performing regression model has a lower MAE value and a lower RMSE value, while a higher R^2 value.

5.3.2 Experimental Evaluation.

Comparison with Baseline Methods. Results in Figure 17 show that DyHealth outperforms the baselines consistently in terms of MAE, RMSE, and R^2. Such superior performance of DyHealth can

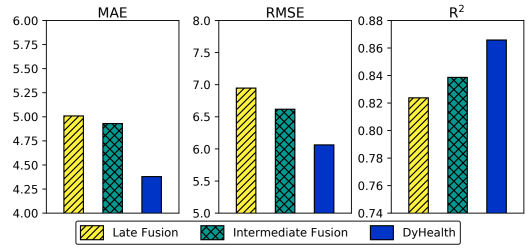

Figure 17: Comparison results for brain age prediction.

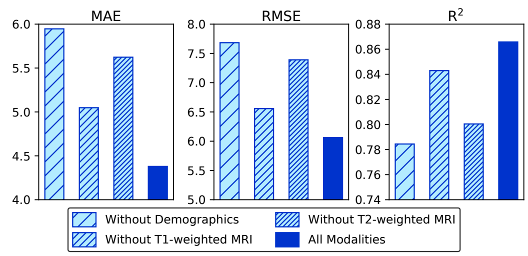

Figure 18: Ablation study results for brain age prediction.

be attributed to the effectiveness of our proposed modality-based attention mechanism in the Multimodal Fusion Module for integrating the complementary information from different modalities.

Ablation Study. We conduct an ablation study to evaluate the influence of each modality. According to the experimental results in Figure 18, we find that demographics are the most influential modality on the performance of brain age prediction, followed by the T2-weighted MRI modality and the T1-weighted MRI modality. These findings confirm that among the factors that are related to the aging of the human brain, demographics are of vital importance [6], including gender [53, 55], education [43, 61], etc.

6 CONCLUSIONS

Based on our observations on real-world healthcare analytics, dynamic modality changes are prevalent. However, to the best of our knowledge, no existing studies take modularity into consideration and support modularity, multimodality, and interpretability simultaneously and satisfactorily. To achieve this, we design DyHealth, a modular, multimodal and interpretable framework for dynamic healthcare analytics. The key idea is to distill the information from different modalities in respective data modules that adhere to the same interface defined by DyHealth, and then integrate the derived complementary information. With such a modular design and the devised mechanisms for handling dynamic modality changes, DyHealth supports modality hot-plug. Furthermore, to fuse information from different modalities, we propose a modality-based attention mechanism that contributes to the fine-grained interpretability on a per-input basis. Through extensive experiments on AKI prediction in the data from the NUH division of nephrology and two other benchmark applications in public datasets, we demonstrate the effectiveness, flexibility, and extensibility of DyHealth in supporting multimodal health data analytics through modularity. Experimental evaluations also showcase how DyHealth supports interpretability and dynamic changes in modalities.

REFERENCES

[1] 2020. Healthcare Analytics. https://en.wikipedia.org/wiki/Health_care_analytics.

[2] 2020. IXI dataset. https://brain-development.org/ixi-dataset/.

[3] 2020. Modularity. https://en.wikipedia.org/wiki/Modularity.

[4] 2022. National University Health System. https://en.wikipedia.org/wiki/National_University_Health_System.

[5] 2022. National University Hospital. https://en.wikipedia.org/wiki/National_University_Hospital.

[6] Habtamu M Aycheh, Joon-Kyung Seong, Jeong-Hyeon Shin, Duk L Na, Byungkon Kang, Sang W Seo, and Kyung-Ah Sohn. 2018. Biological brain age prediction using cortical thickness data: A large scale cohort study. *Frontiers in aging neuroscience* 10 (2018), 252.

[7] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2015. Neural machine translation by jointly learning to align and translate. In *3rd International Conference on Learning Representations.*

[8] Tadas Baltrušaitis, Chaitanya Ahuja, and Louis-Philippe Morency. 2018. Multimodal machine learning: A survey and taxonomy. *IEEE transactions on pattern analysis and machine intelligence* 41, 2 (2018), 423–443.

[9] Or Biran and Courtenay Cotton. 2017. Explanation and justification in machine learning: A survey. In *IJCAI-17 workshop on explainable AI (XAI)*, Vol. 8. 8–13.

[10] D. A. Black. 1969. The biochemistry of renal failure. *Br J Anaesth* 41 (1969), 264–268.

[11] Zhengping Che, David Kale, Wenzhe Li, Mohammad Taha Bahadori, and Yan Liu. 2015. Deep computational phenotyping. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* 507–516.

[12] Guangyu Chen, Hong-Ying Zhang, Chunming Xie, Gang Chen, Zhi-Jun Zhang, Gao-Jun Teng, and Shi-Jiang Li. 2013. Modular reorganization of brain resting state networks and its independent validation in Alzheimer's disease patients. *Frontiers in human neuroscience* 7 (2013), 456.

[13] Jianpeng Cheng, Li Dong, and Mirella Lapata. 2016. Long Short-Term Memory-Networks for Machine Reading. In *EMNLP*. The Association for Computational Linguistics, 551–561.

[14] Kyunghyun Cho, Bart van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. 2014. Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP).* 1724–1734.

[15] Edward Choi, Mohammad Taha Bahadori, Jimeng Sun, Joshua Kulas, Andy Schuetz, and Walter Stewart. 2016. Retain: An interpretable predictive model for healthcare using reverse time attention mechanism. In *Advances in Neural Information Processing Systems.* 3504–3512.

[16] Horng-Ruey Chua, Weng-Kin Wong, Venetia Huiling Ong, Dipika Agrawal, Anantharaman Vathsala, Hui-Ming Tay, and Amartya Mukhopadhyay. 2020. Extended mortality and chronic kidney disease after septic acute kidney injury. *Journal of Intensive Care Medicine* 35, 6 (2020), 527–535.

[17] Jian Dai, Meihui Zhang, Gang Chen, Ju Fan, Kee Yuan Ngiam, and Beng Chin Ooi. 2018. Fine-grained Concept Linking using Neural Networks in Healthcare. In *SIGMOD Conference.* ACM, 51–66.

[18] Thomas Elsken, Jan Hendrik Metzen, and Frank Hutter. 2019. Neural Architecture Search: A Survey. *J. Mach. Learn. Res.* 20 (2019), 55:1–55:21.

[19] Yujuan Feng, Zhenxing Xu, Lin Gan, Ning Chen, Bin Yu, Ting Chen, and Fei Wang. 2019. DCMN: Double Core Memory Network for Patient Outcome Prediction with Multimodal Data. In *2019 IEEE International Conference on Data Mining (ICDM).* IEEE, 200–209.

[20] Luca Ferrarini, Ilya M Veer, Evelinda Baerends, Marie-José van Tol, Remco J Renken, Nic JA van der Wee, Dirk J Veltman, André Aleman, Frans G Zitman, Brenda WJH Penninx, et al. 2009. Hierarchical functional modularity in the resting-state human brain. *Human brain mapping* 30, 7 (2009), 2220–2231.

[21] Katja Franke, Gabriel Ziegler, Stefan Klöppel, Christian Gaser, Alzheimer's Disease Neuroimaging Initiative, et al. 2010. Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: exploring the influence of various parameters. *Neuroimage* 50, 3 (2010), 883–892.

[22] Hunter B Fraser. 2005. Modularity and evolutionary constraint on proteins. *Nature genetics* 37, 4 (2005), 351–352.

[23] Wenzhong Guo, Jianwen Wang, and Shiping Wang. 2019. Deep multimodal representation learning: A survey. *IEEE Access* 7 (2019), 63373–63394.

[24] Kensho Hara, Hirokatsu Kataoka, and Yutaka Satoh. 2018. Can Spatiotemporal 3D CNNs Retrace the History of 2D CNNs and ImageNet?. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* 6546–6555.

[25] Wei-Zhen Hong, Sabrina Haroon, Titus Lau, Soh-Heng Hui, Valerie Ma, Jing-Xian Goh, and Horng-Ruey Chua. 2019. Transitional care program to facilitate recovery following severe acute kidney injury. *Journal of Nephrology* 32, 4 (2019), 605–613.

[26] Jianying Hu, Fei Wang, Jimeng Sun, Robert Sorrentino, and Shahram Ebadollahi. 2012. A healthcare utilization analysis framework for hot spotting and contextual anomaly detection. In *AMIA annual symposium proceedings*, Vol. 2012. American Medical Informatics Association, 360.

[27] Mark Jenkinson, Peter Bannister, Michael Brady, and Stephen Smith. 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17, 2 (2002), 825–841.

[28] Alistair EW Johnson, Tom J Pollard, Lu Shen, Li-wei H Lehman, Mengling Feng, Mohammad Ghassemi, Benjamin Moody, Peter Szolovits, Leo Anthony Celi, and Roger G Mark. 2016. MIMIC-III, a freely accessible critical care database. *Scientific data* 3 (2016).

[29] Benedikt Atli Jónsson, Gyda Bjornsdottir, TE Thorgeirsson, Lotta María Ellingsen, G Bragi Walters, DF Gudbjartsson, Hreinn Stefansson, Kari Stefansson, and MO Ulfarsson. 2019. Brain age prediction using deep learning uncovers associated sequence variants. *Nature communications* 10, 1 (2019), 1–10.

[30] David C Kale, Zhengping Che, Mohammad Taha Bahadori, Wenzhe Li, Yan Liu, and Randall Wetzel. 2015. Causal phenotype discovery via deep networks. In *AMIA Annual Symposium Proceedings*, Vol. 2015. American Medical Informatics Association, 677.

[31] John A Kellum, Norbert Lameire, Peter Aspelin, Rashad S Barsoum, Emmanuel A Burdmann, Stuart L Goldstein, Charles A Herzog, Michael Joannidis, Andreas Kribben, Andrew S Levey, et al. 2012. Kidney disease: improving global outcomes (KDIGO) acute kidney injury work group. KDIGO clinical practice guideline for acute kidney injury. *Kidney international supplements* 2, 1 (2012), 1–138.

[32] Nitin V Kolhe, Timothy Reilly, Janson Leung, Richard J Fluck, Kirsty E Swinscoe, Nicholas M Selby, and Maarten W Taal. 2016. A simple care bundle for use in acute kidney injury: a propensity score-matched cohort study. *Nephrology Dialysis Transplantation* 31, 11 (2016), 1846–1854.

[33] Bum Chul Kwon, Min-Je Choi, Joanne Taery Kim, Edward Choi, Young Bin Kim, Soonwook Kwon, Jimeng Sun, and Jaegul Choo. 2018. Retainvis: Visual analytics with interpretable and interactive recurrent neural networks on electronic medical records. *IEEE transactions on visualization and computer graphics* 25, 1 (2018), 299–309.

[34] David E Leaf and Marta Christov. 2019. Dysregulated mineral metabolism in AKI. In *Seminars in Nephrology*, Vol. 39. Elsevier, 41–56.

[35] Zheng Jye Ling, Quoc Trung Tran, Ju Fan, Gerald C. H. Koh, Thi Nguyen, Chuen Seng Tan, James Wei Luen Yip, and Meihui Zhang. 2014. GEMINI: An Integrative Healthcare Analytics System. *Proc. VLDB Endow.* 7, 13 (2014), 1766–1771.

[36] Zachary C Lipton, David C Kale, Charles Elkan, and Randall Wetzel. 2016. Learning to diagnose with LSTM recurrent neural networks. In *International Conference on Learning Representations.*

[37] Sanmay Low, Anantharaman Vathsala, Tanusya Murali Murali, Long Pang, Graeme MacLaren, Wan-Ying Ng, Sabrina Haroon, Amartya Mukhopadhyay, Shir-Lynn Lim, Bee-Hong Tan, et al. 2019. Electronic health records accurately predict renal replacement therapy in acute kidney injury. *BMC nephrology* 20, 1 (2019), 1–11.

[38] Zhaojing Luo, Sai Ho Yeung, Meihui Zhang, Kaiping Zheng, Lei Zhu, Gang Chen, Feiyi Fan, Qian Lin, Kee Yuan Ngiam, and Beng Chin Ooi. 2021. MLCask: Efficient Management of Component Evolution in Collaborative Data Analytics Pipelines. In *ICDE.* IEEE, 1655–1666.

[39] Fenglong Ma, Radha Chitta, Jing Zhou, Quanzeng You, Tong Sun, and Jing Gao. 2017. Dipole: Diagnosis prediction in healthcare via attention-based bidirectional recurrent neural networks. In *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining.* 1903–1911.

[40] Fenglong Ma, Yaqing Wang, Houping Xiao, Ye Yuan, Radha Chitta, Jing Zhou, and Jing Gao. 2018. A general framework for diagnosis prediction via incorporating medical code descriptions. In *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM).* IEEE, 1070–1075.

[41] Tim Miller. 2019. Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence* 267 (2019), 1–38.

[42] Christoph Molnar et al. 2018. Interpretable machine learning: A guide for making black box models explainable. *E-book at< https://christophm.github.io/interpretable-ml-book/>, version dated* 10 (2018).

[43] Moyra E Mortby, Richard Burns, Andrew L Janke, Perminder S Sachdev, Kaarin J Anstey, and Nicolas Cherbuin. 2014. Relating education, brain structure, and cognition: the role of cardiovascular disease risk factors. *BioMed research international* 2014 (2014).

[44] DR Mould. 2012. Models for disease progression: new approaches and uses. *Clinical Pharmacology & Therapeutics* 92, 1 (2012), 125–131.

[45] Mizuho Nishio, Shunjiro Noguchi, Hidetoshi Matsuo, and Takamichi Murakami. 2020. Automatic classification between COVID-19 pneumonia, non-COVID-19 pneumonia, and the healthy on chest X-ray image: combination of data augmentation methods. *Scientific reports* 10, 1 (2020), 1–6.

[46] World Health Organization et al. 2020. *Laboratory testing for 2019 novel coronavirus (2019-nCoV) in suspected human cases: interim guidance, 14 January 2020.* Technical Report. World Health Organization.

[47] World Health Organization et al. 2020. *Use of chest imaging in COVID-19: a rapid advice guide, 11 June 2020.* Technical Report. World Health Organization.

[48] World Health Organization et al. 2020. *WHO COVID-19: case definitions: updated in public health surveillance for COVID-19, published 16 December 2020.* Technical

Report. World Health Organization.

[49] Sanjay Purushotham, Chuizheng Meng, Zhengping Che, and Yan Liu. 2018. Benchmarking deep learning models on large healthcare datasets. *Journal of biomedical informatics* 83 (2018), 112–134.

[50] Zhi Qiao, Xian Wu, Shen Ge, and Wei Fan. 2019. MNN: Multimodal Attentional Neural Networks for Diagnosis Prediction. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019, Macao, China, August 10-16, 2019*, Sarit Kraus (Ed.). ijcai.org, 5937–5943.

[51] Dhanesh Ramachandram and Graham W Taylor. 2017. Deep multimodal learning: A survey on recent advances and trends. *IEEE Signal Processing Magazine* 34, 6 (2017), 96–108.

[52] Erzsébet Ravasz, Anna Lisa Somera, Dale A Mongru, Zoltán N Oltvai, and A-L Barabási. 2002. Hierarchical organization of modularity in metabolic networks. *science* 297, 5586 (2002), 1551–1555.

[53] Stuart J Ritchie, Simon R Cox, Xueyi Shen, Michael V Lombardo, Lianne M Reus, Clara Alloza, Mathew A Harris, Helen L Alderson, Stuart Hunter, Emma Neilson, et al. 2018. Sex differences in the adult human brain: evidence from 5216 UK Biobank participants. *Cerebral Cortex* 28, 8 (2018), 2959–2975.

[54] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*. Springer, 234–241.

[55] Amber NV Ruigrok, Gholamreza Salimi-Khorshidi, Meng-Chuan Lai, Simon Baron-Cohen, Michael V Lombardo, Roger J Tait, and John Suckling. 2014. A meta-analysis of sex differences in human brain structure. *Neuroscience & Biobehavioral Reviews* 39 (2014), 34–50.

[56] Mike Schuster and Kuldip K. Paliwal. 1997. Bidirectional recurrent neural networks. *IEEE Trans. Signal Processing* 45, 11 (1997), 2673–2681.

[57] Nicholas M Selby, Anna Casula, Laura Lamming, John Stoves, Yohan Samarasinghe, Andrew J Lewington, Russell Roberts, Nikunj Shah, Melanie Johnson, Natalie Jackson, et al. 2019. An organizational-level program of intervention for AKI: a pragmatic stepped wedge cluster randomized trial. *Journal of the American Society of Nephrology* 30, 3 (2019), 505–515.

[58] Ying Sha and May D Wang. 2017. Interpretable predictions of clinical outcomes with an attention-based recurrent neural network. In *Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*. 233–240.

[59] Parikshit Sondhi, Jimeng Sun, Hanghang Tong, and ChengXiang Zhai. 2012. SympGraph: a framework for mining clinical notes through symptom relation graphs. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. 1167–1175.

[60] Qiuling Suo, Fenglong Ma, Giovanni Canino, Jing Gao, Aidong Zhang, Pierangelo Veltri, and Gnasso Agostino. 2017. A multi-task framework for monitoring health conditions via attention-based recurrent neural networks. In *AMIA annual symposium proceedings*, Vol. 2017. American Medical Informatics Association, 1665.

[61] Megan E Thow, Mathew J Summers, Nichole L Saunders, Jeffery J Summers, Karen Ritchie, and James C Vickers. 2018. Further education improves cognitive reserve and triggers improvement in selective cognitive functions in older adults: The Tasmanian Healthy Brain Project. *Alzheimer's & Dementia: diagnosis, assessment & disease monitoring* 10 (2018), 22–30.

[62] Amanda L Traud, Peter J Mucha, and Mason A Porter. 2012. Social structure of facebook networks. *Physica A: Statistical Mechanics and its Applications* 391, 16 (2012), 4165–4180.

[63] Nicholas J Tustison, Brian B Avants, Philip A Cook, Yuanjie Zheng, Alexander Egan, Paul A Yushkevich, and James C Gee. 2010. N4ITK: improved N3 bias correction. *IEEE transactions on medical imaging* 29, 6 (2010), 1310–1320.

[64] Fei Wang, Noah Lee, Jianying Hu, Jimeng Sun, Shahram Ebadollahi, and Andrew F Laine. 2012. A framework for mining signatures from event sequences and its applications in healthcare data. *IEEE transactions on pattern analysis and machine intelligence* 35, 2 (2012), 272–285.

[65] Fei Wang and Jimeng Sun. 2015. PSF: a unified patient similarity evaluation framework through metric learning with weak supervision. *IEEE journal of biomedical and health informatics* 19, 3 (2015), 1053–1060.

[66] Kelvin Xu, Jimmy Ba, Ryan Kiros, Kyunghyun Cho, Aaron C. Courville, Ruslan Salakhutdinov, Richard S. Zemel, and Yoshua Bengio. 2015. Show, Attend and Tell: Neural Image Caption Generation with Visual Attention. In *ICML (JMLR Workshop and Conference Proceedings, Vol. 37)*. JMLR.org, 2048–2057.

[67] Yanbo Xu, Siddharth Biswal, Shriprasad R Deshpande, Kevin O Maher, and Jimeng Sun. 2018. Raim: Recurrent attentive and intensive model of multimodal patient monitoring data. In *Proceedings of the 24th ACM SIGKDD international conference on Knowledge Discovery & Data Mining*. 2565–2573.

[68] Ye Yuan, Guangxu Xun, Fenglong Ma, Qiuling Suo, Hongfei Xue, Kebin Jia, and Aidong Zhang. 2018. A novel channel-aware attention framework for multi-channel eeg seizure detection via multi-view deep learning. In *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*. IEEE, 206–209.

[69] Kaiping Zheng, Shaofeng Cai, Horng Ruey Chua, Wei Wang, Kee Yuan Ngiam, and Beng Chin Ooi. 2020. TRACER: A Framework for Facilitating Accurate and Interpretable Analytics for High Stakes Applications. In *Proceedings of the 2020 International Conference on Management of Data, SIGMOD*. ACM.

[70] Kaiping Zheng, Jinyang Gao, Kee Yuan Ngiam, Beng Chin Ooi, and James Wei Luen Yip. 2017. Resolving the Bias in Electronic Medical Records. In *KDD*. ACM, 2171–2180.