Big Data & Smart City Group,
Peking University

ConCare: Personalized Clinical Feature Embedding via Capturing the Healthcare Context

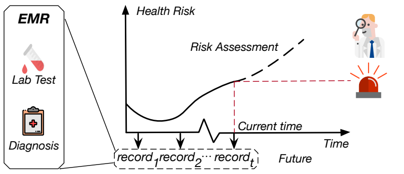
Liantao Ma, Chaohe Zhang, Yasha Wang, Wenjie Ruan, Jiangtao Wang, Wen Tang, Xinyu Ma, Xin Gao, Junyi Gao

Association for the Advancement of Artificial Intelligence

SCAN ME!

I. Background

- **By evaluating patient's health status comprehensively, physicians can:**
  - ✓ Select follow-up personalized treatments
  - ✓ Prevent adverse outcomes
  - ✓ Assign medical resources effectively
  - ✓ Reduce the medical cost
- **Practical application:**
  - ✓ Prognosis of End Stage Renal Disease (ESRD)

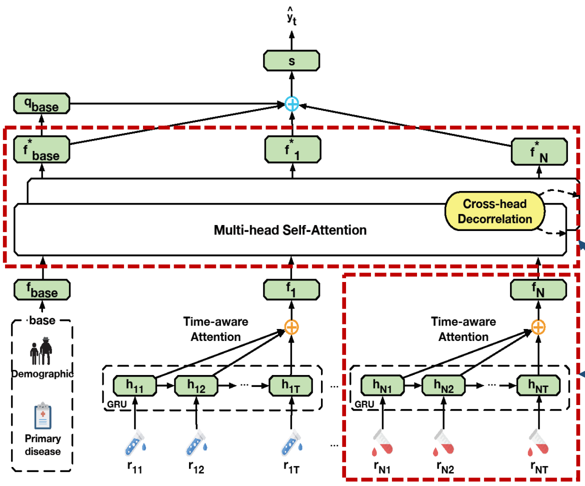


II. EMR Data

- A type of multivariate time series data that records patients' dynamic visits and static baseline information.
- **Static Information:**
  - ✓ Demographics (Age/Gender/…)
  - ✓ Primary diseases (Diabetes/…)
- **Dynamic Information:**
  - ✓ Lab test results (Blood glucose/…)

	Demographic				Primary disease				…
Patient ID	Gender	Age	Height	…	CGN	Diabetes	CHD	…	…
98	M	53.92	160.54		1	0	0		
118	F	42.51	171.41		0	1	0		

		Blood lab test		Physical lab test		New-onset complications			…
Patient ID	Date	Albumin	…	Weight	…	URTI	Cancer	…	…
98	19/12/2015	35.0		66.3		0	0		
98	22/01/2016	38.2		65.2		1	0		
118	09/07/2015	32.8		49.8		0	1		

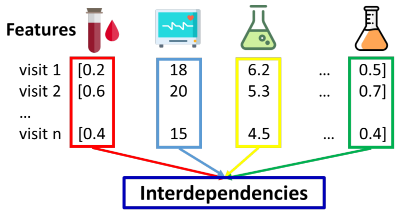


III. Framework of ConCare

We propose a multi-channel healthcare predictive model, which can learn the representation of health status and perform the health prediction by more deeply considering the personal health context.

- **Take the embedding of all features as healthcare context.**
  - ✓ The feature encoder is adopted to combine all the static information and dynamic records based on **self-attention, and learn interdependency among features**.
- **Capture the impact of time interval for different features.**
  - ✓ The multi-channel time series embedding module with **time-aware attention** is developed to separately learn the representation of each dynamic feature.

Insight 1: Extract Personal Health Context (Feature Interdependency)

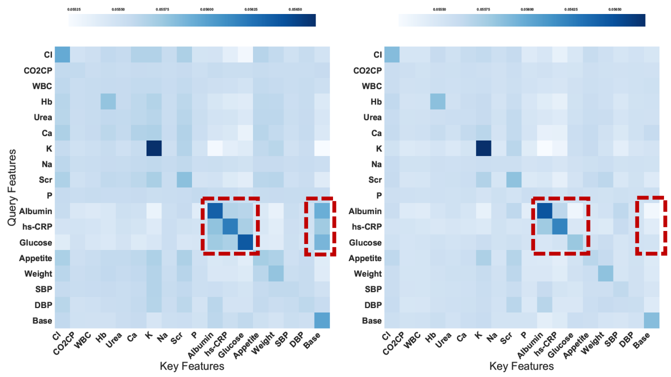


- A certain value of a clinical feature (e.g., blood glucose) may imply different meanings to patients with diverse static baselines (e.g., diagnosis of diabetes as a primary disease).
- We explicitly extract inter-dependencies between clinical features to learn the personal health context and regenerate the feature embedding under the context, by a **multi-head self-attention mechanism**.

Case study on experiment:

- Most of the clinical features are more likely to respond to themselves, which is denoted by the diagonal of two matrices.
- By comparing the two figures, in the box of **Glucose-Glucose** position, the model pays much more attention to the glucose in **patients with diabetes**.
- Besides, there are relatively high interdependencies between **blood albumin, hyper-sensitive C-reactive protein (hs-CRP), glucose and the static baseline information (including the diagnosis of diabetes)** for patients with diabetes.

Cross-Feature Interdependency:
Patients with (Left) / without (Right) Diabetes



This is consistent with the medical research and clinical experience.

Insight 2: Capturing the Impact of Time Interval for Different Features

- It is usually assumed that the more recent clinical records weight more than previous records.
  - ✓ The historical record of serum potassium several years ago has little influence on current healthcare prediction.
- However, historical records also contain valuable clinical information, which may not be revealed in the latest record.
- We learn the representation of different feature sequences via separate GRUs, and adaptively captures the effect of time intervals between records of each feature **by time-aware attention**.

Case study on experiment: Time-decay rates for different features

- ConCare attends more on the **short-term of serum creatinine (Scr), potassium (K), white blood cell count (WBC), calcium (Ca), carbon dioxide combining power (CO2CP), chlorine (Cl), hemoglobin (Hb).** They are relatively fast changing indicators, reflecting patients' infection status or dialysis adequacy, etc.
- ConCare attends more on the **long-term of Body Weight, Albumin, Na and systolic pressure (SBP).** They are usually related to nutrition intake and reflect the patient's condition over a period of time.