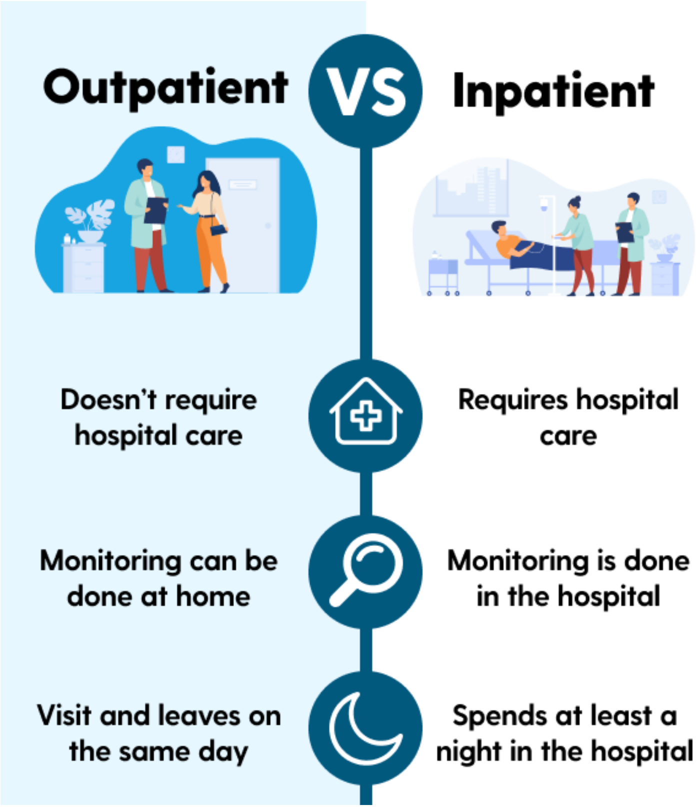
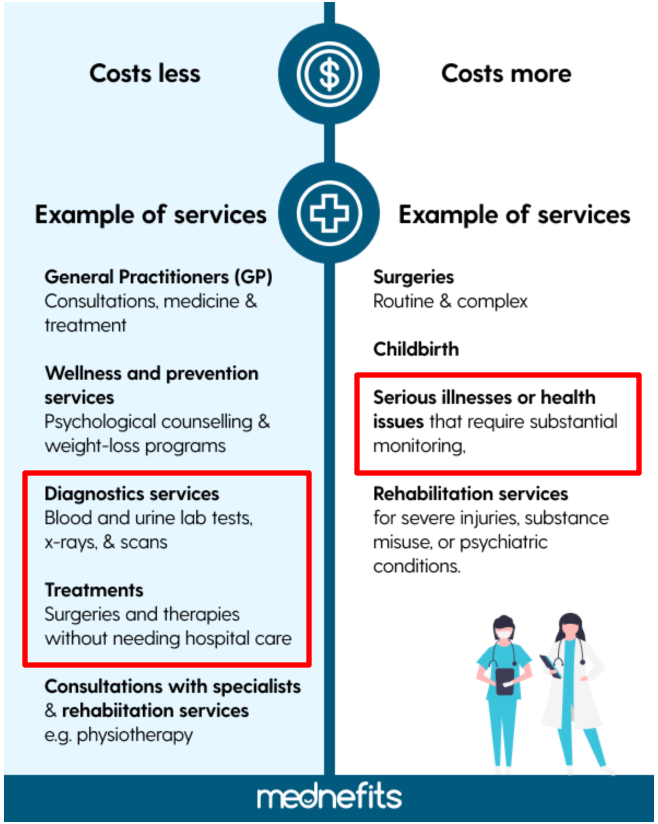
INTRODUCTION



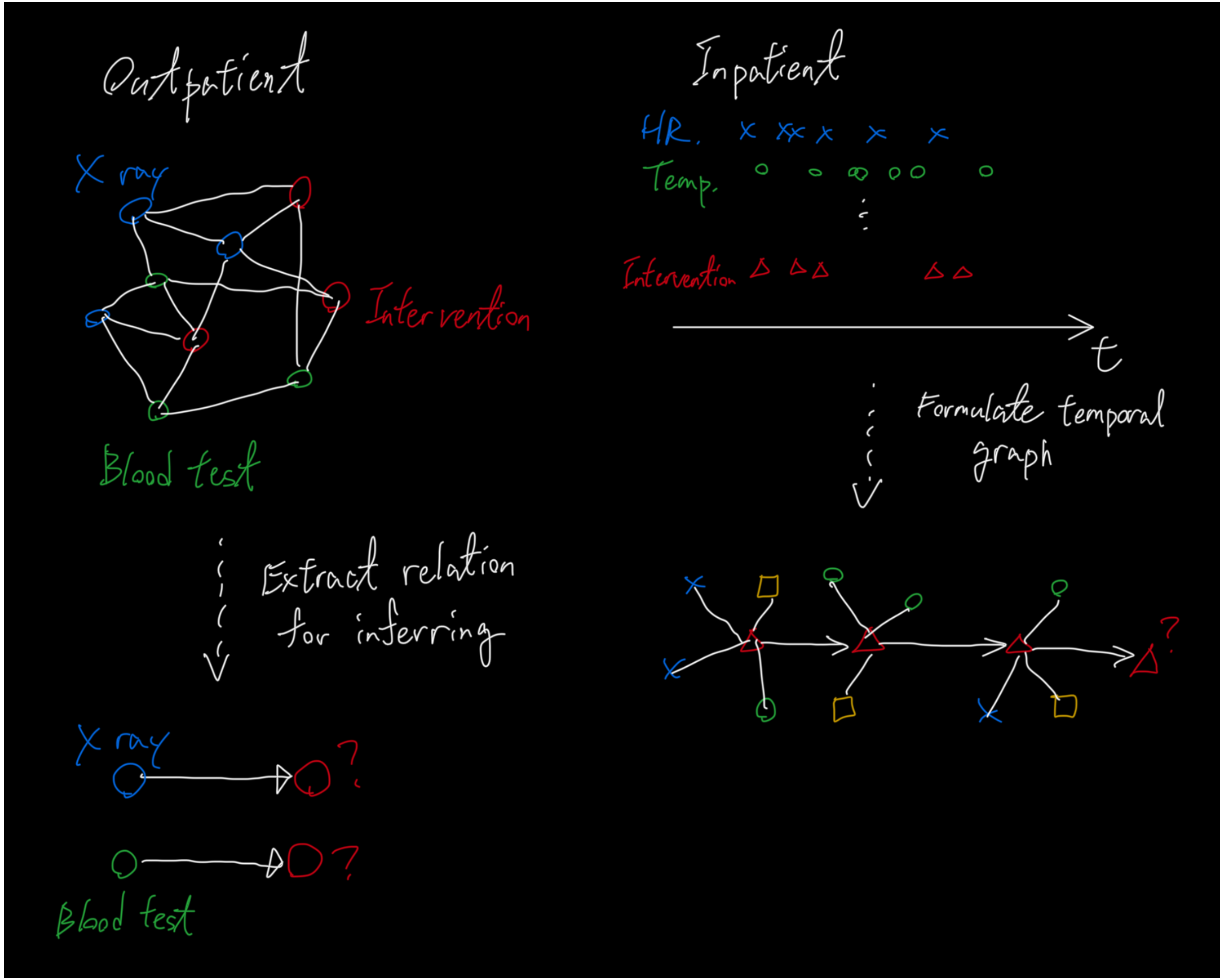


What is the difference between inpatient and outpatient services?: https://www.mednefits.com/blog/inpatient-vs-outpatient/

INTRODUCTION

研究構想

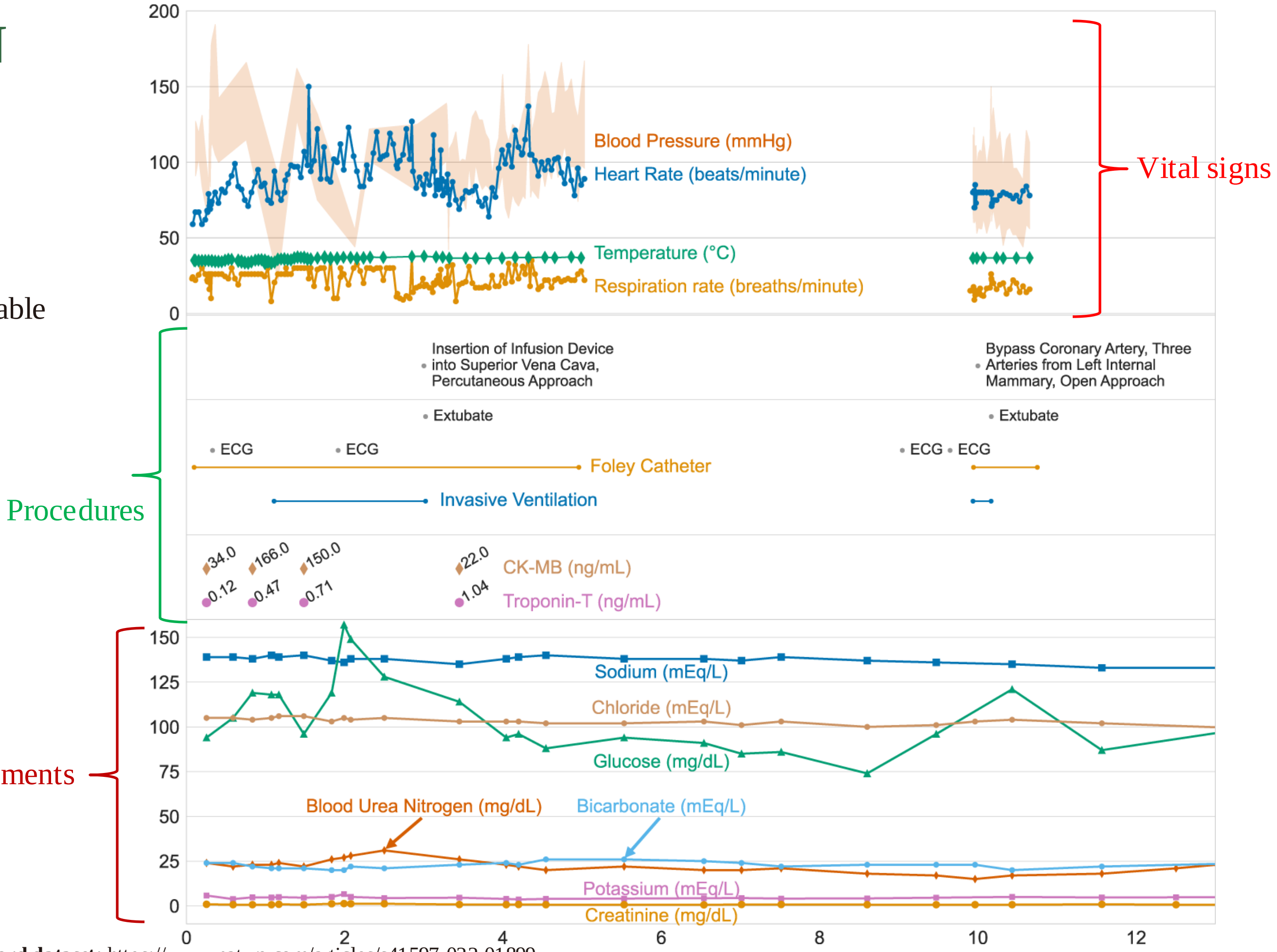
- 參考醫師思路，建構Temporal Heterogeneous Graph，時間、特徵的結構化關聯，隨著病情變化(vital sign pattern)、intervention、檢測結果(lab test)，更即時推薦用藥處置



INTRODUCTION

Data in EHR (MIMIC-IV)

- Visualization for a single patient's hospitalization
- **Frequent vital signs** are only available when the patient is in the **ICU**
- Laboratory measures are available throughout their hospitalization



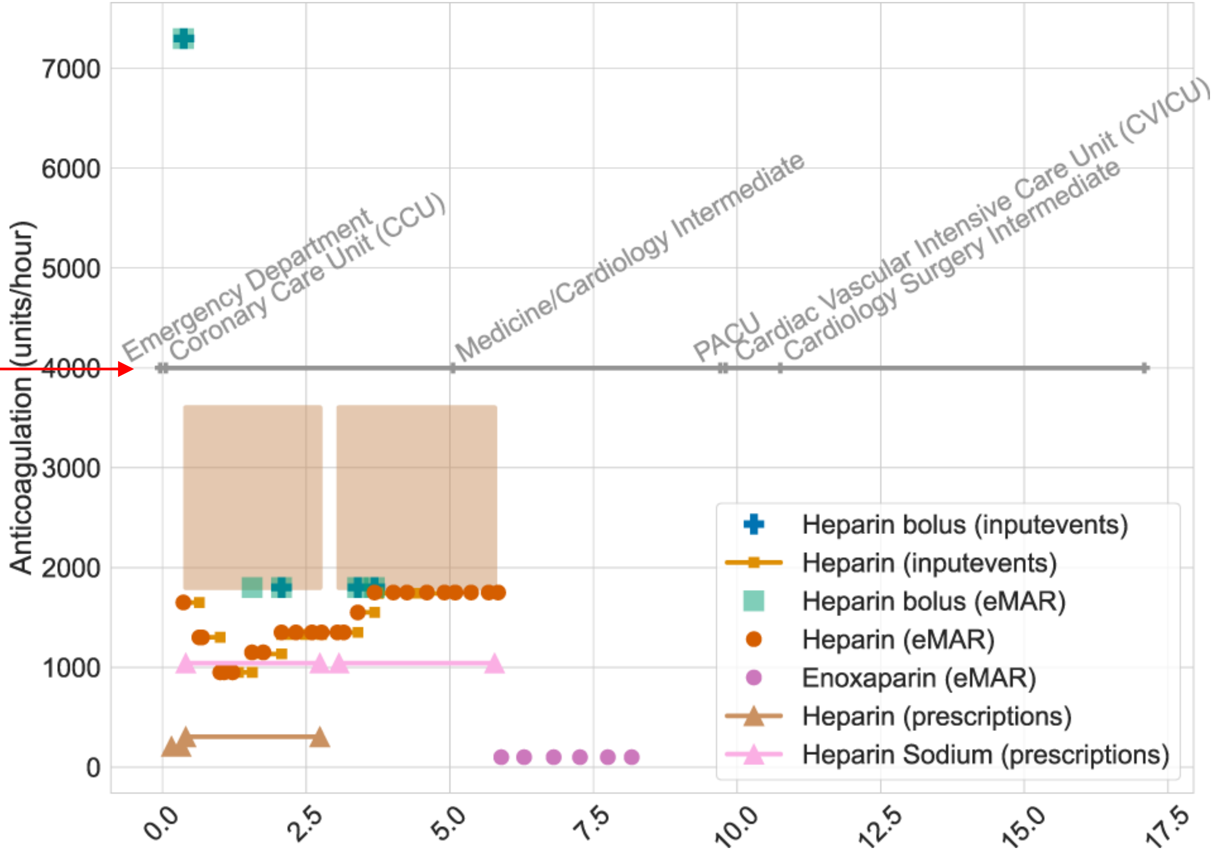
MIMIC-IV, a freely accessible electronic health record dataset: https://www.nature.com/articles/s41597-022-01899-x

INTRODUCTION

Data in EHR (MIMIC-IV)

- Visualization of **medication information** for a single patient's hospitalization

Care units



MIMIC-IV, a freely accessible electronic health record dataset: https://www.nature.com/articles/s41597-022-01899-x

Predicting Inpatient Medication Orders From Electronic Health Record Data

Kathryn Rough, Andrew M. Dai, Kun Zhang, Yuan Xue, Laura M. Vardoulakis, Claire Cui, Atul J. Butte, Michael D. Howell and Alvin Rajkomar

2020 Google. CLINICAL PHARMACOLOGY & THERAPEUTICS

國立成功大學
National Cheng Kung University

INTRODUCTION

- **For reference**

The utility of medical predictive models is demonstrated by their ongoing use in clinical care. Routinely used examples include the Pooled Cohort Equation to estimate cardiovascular risk[1] and CHA_2DS_2-VASc to predict thromboembolism.[2] These models produce risk scores for a single outcome that assist clinicians in decision making for a specific question: Is this patient at high enough risk of cardiovascular events to require a statin? Does this patient with atrial fibrillation have sufficiently elevated thromboembolism risk to benefit from anticoagulation? Most well-known clinical predictions come from carefully designed statistical models that use a small set of patient observations and predictor variables; they fall on the simpler end of the “machine-learning spectrum.”[3]

第一段： 醫療預測模型的效用已在臨床護理中的持續使用中得到證明。常用的例子包括使用Pooled Cohort Equation來估算心血管風險，以及使用CHA_2DS_2-VASc來預測血栓栓塞。這些模型會為特定問題生成風險分數，幫助臨床醫師做出決策，例如：這名患者的心血管事件風險是否足夠高，需使用他汀類藥物？這名患有房顫的患者的血栓栓塞風險是否足夠高，從而能受益於抗凝治療？大多數眾所周知的臨床預測來自精心設計的統計模型，這些模型使用了少量的患者觀察數據和預測變量，屬於「機器學習光譜」的簡單端。

INTRODUCTION

- **For reference**

Yet, clinical care is not confined to isolated binary decisions; physicians generally choose among many treatment options, and those decisions may change as a patient's condition evolves. For instance, a physician admitting a patient with signs of severe infection may initially order intravenous fluids, an analgesic, and several broad-spectrum antibiotics. Throughout the hospitalization, additional medications will be ordered as adjunctives or to address pre-existing conditions. Eventually, a more targeted antibiotic might be ordered based on blood culture results. Developing predictive models to reliably anticipate these types of heterogeneous therapeutic actions over the course of a hospitalization could lead to increasingly useful clinical decision-support tools.

第二段： 然而，臨床護理並不僅限於單一的二元決策；醫生通常會在多種治療選擇中進行選擇，並且這些決策可能會隨著患者的病情變化而改變。例如，一名有嚴重感染跡象的患者入院時，醫生可能會首先開立靜脈輸液、止痛藥和多種廣譜抗生素。在住院期間，還會根據患者的情況添加額外的藥物來輔助治療或解決已有的病情。最終，根據血液培養結果，可能會開具更具針對性的抗生素。開發能夠可靠預測這種住院期間多樣化治療行為的預測模型，將有助於創建更加有用的臨床決策支持工具。

INTRODUCTION

- **For reference**

Medication-related prediction and machine-learning research has generally focused on selected patient populations and therapeutic areas. Examples include prediction of prescribing dynamics for sleep medications,[4] onset of vasopressor use in intensive care units,[5] progression through diabetic medications,[6] pharmaceutical treatments in pregnant women,[7] and discharge antihypertensive medications.[8] Numerous methods have also been applied to improving clinical order sets.[9–16] A more general approach was taken by Choi *et al.*[17]; using electronic health record (EHR) data, a recurrent neural network predicted the medication classes likely to be prescribed during the next outpatient encounter.

第三段： 與藥物相關的預測和機器學習研究通常集中於特定的患者群體和治療領域。例子包括：睡眠藥物開具動態的預測，重症監護病房中血管加壓劑使用的預測，糖尿病藥物治療進展的預測，孕婦的藥物治療，和抗高血壓藥物出院開具的預測。還有許多方法應用於改善臨床訂單的設置。一個更為普遍的方法是由**Choi 等人**提出的：利用電子健康記錄（EHR）數據，一個遞歸神經網路預測了患者下一次門診就診時可能會開具的藥物類別。

INTRODUCTION

- **For reference**

The clinical applicability of previous work on medication prediction has been limited by narrow patient populations, infrequent timing of predictions (i.e., at the encounter or day level), or by aggregating medications into broad categories. In this paper, we build on previous work by training models to predict which specific medication compounds physicians will order across a general inpatient population throughout their hospitalization, without restricting to particular patient cohorts or therapeutic areas. We evaluate the models' capacity to produce predictions whenever orders are placed due to patient needs and clinical workflow.

第四段： 過去關於藥物預測的研究臨床應用受到了限制，原因包括：患者群體狹窄，預測時間點不頻繁（例如，僅在就診或每日層面上），或是將藥物分類過於廣泛。在本文中，我們基於以往的研究，訓練模型來預測住院期間，臨床醫生會開具哪些具體的藥物，而不限制於特定患者群體或治療領域。我們評估了這些模型是否能夠在患者住院期間隨著病情和臨床工作流程變化的情況下，準確預測下達的訂單。

INTRODUCTION

Key points

- 醫生通常會在**多種治療選擇**中進行選擇，並且這些決策可能會**隨著患者的病情變化而改變**。
    - 例如，一名有嚴重**感染**跡象的患者入院時，醫生可能會首先開立靜脈輸液、止痛藥和多種廣譜抗生素。在住院期間，還會根據**患者的情況**添加額外的藥物來**輔助治療或解決已有的病情**。最終，根據**血液培養結果**，可能會開具更具**針對性的抗生素**。開發能夠可靠預測這種住院期間多樣化治療行為的預測模型，將有助於創建更加有用的臨床決策支持工具。
- 過去關於**藥物預測的研究臨床應用**受到了限制，原因包括：**患者群體狹窄**，**預測時間點不頻繁**（例如，僅在**就診**或**每日層面**上），或是將**藥物分類過於廣泛**。
- 在本文中，我們基於以往的研究，訓練模型來**預測住院期間，臨床醫生會開具哪些具體的藥物**，而不限制於特定患者群體或治療領域。我們評估了這些模型是否能夠在患者住院期間隨著**病情**和**臨床工作流程變化**的情況下，準確預測下達的醫囑(用藥)。

INTRODUCTION

Q1:醫師在做用藥決策時，會參考病人哪些資料? 有什麼考慮?

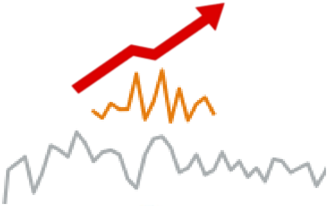
1. 病人用藥紀錄
2. 嘗試性用藥處置
   - 例如: 插管→感染→等待實驗室檢驗結果釐清確切感染類型(革蘭氏陰/陽性, …)，需要一段時間，先嘗試可能的用藥，查看病人後續病況反映，無效則做用藥調整 [1]
3. 同疾病類型(如: 敗血症)病人，用藥處置較相近
4. 醫生用藥習慣(同一種症狀，可有不同可行的用藥)
5. DDI (藥物交互作用)

[1]https://www.mmh.org.tw/child/know_health_view.php?docid=98

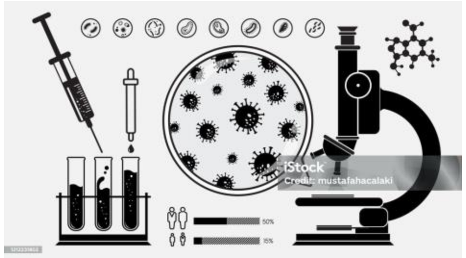
INTRODUCTION

Q2:醫生思路

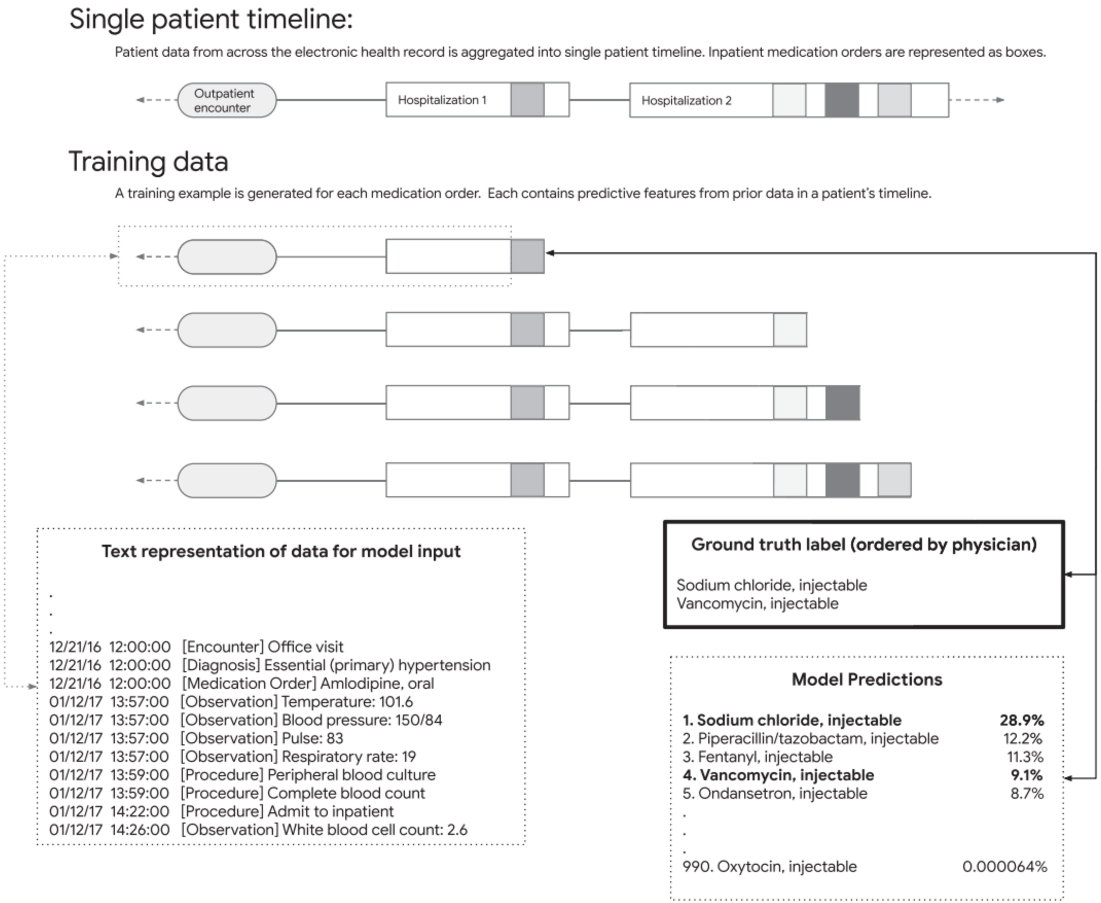
1. Vital sign pattern出現異狀 (看變化趨勢判斷，如: 體溫上升(發燒), …)
2. 觀察病人疾病類型、年齡、性別等資訊綜合考量，揣測可能感染類型(細菌、黴菌、病毒)
3. Intervention code處置
   - Ex: provider order entry(實驗室檢驗感染源), medication code(緊急嘗試性用藥處置)





INTRODUCTION



Predictive Modeling with Temporal Graphical Representation on Electronic Health Records

Kathryn Rough, Andrew M. Dai, Kun Zhang, Yuan Xue, Laura M. Vardoulakis, Claire Cui, Atul J. Butte, Michael D. Howell and Alvin Rajkomar

2024 International Joint Conference on Artificial Intelligence (IJCAI)

國立成功大學
National Cheng Kung University

Methodology

Overview

論文提出了一個名為Temporal Heterogeneous Graph (THG)的異質圖模型，來表示病人的電子病歷(EHR)。

為了充分**捕捉時間依賴性和結構信息**，模型利用各種**圖卷積**與**Transformer**技術。

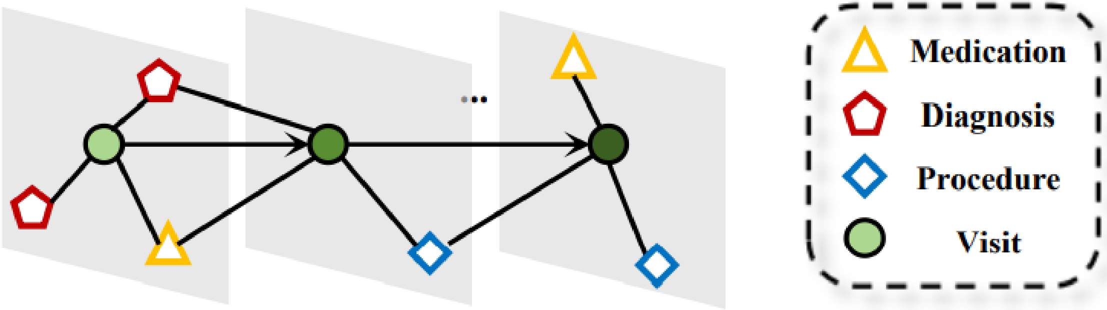
整個方法論分為以下幾個主要步驟：

1. Graph Construction（圖結構構建）
2. Temporal Heterogeneous Message Passing（時序異質消息傳遞）
3. Spatial Encoding（空間編碼）
4. Patient Explainer（病人解釋器）

Methodology

Graph Construction（圖結構構建）

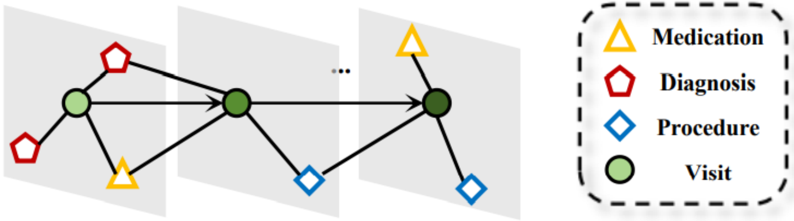
- 將每個病人的EHR構建成一個異質圖G=(N,E)
  - N (Nodes, 節點): 節點包含四種類型
    - N_v：表示每次病患就診的"時間點節點"（Time-aware visit nodes）
    - N_e：表示每次病患就診內的 "醫療事件節點"（Medical event nodes），包括：
      - 診斷（Diagnosis）、治療程序（Procedure）、藥物使用（Medication）



(c) Temporal Graphical Representation

Methodology

Graph Construction（圖結構構建）



(c) Temporal Graphical Representation

- 將每個病人的EHR構建成一個異質圖$G=(N,E)$
  - E (Edges, 邊): 節點之間的連結：
    - E_{ev}：表示在一次就診中，某個醫療事件（如診斷、藥物）與該次就診的聯繫。

      $$E_{ev}=\{(u_1,u_2):\forall u_1\in N_e,\forall u_2\in N_v\}$$

      其中，u_1為醫療事件節點，u_2為就診節點。
    - E_{vv}：表示歷史就診之間的時間順序聯繫（時間順序邊）。

      $$E_{vv}=\{(u_t,u_{t+1}):\forall u\in N_v, 0\leq t<T\}$$

      其中，u_t和u_{t+1}分別表示不同時間點的就診節點。
  - 目的：
    - 捕捉**病患健康狀況**在**不同就診時間點**之間的**變化**。
    - 建立**時間**與**醫療事件**之間的**依賴關係**。

Methodology

Temporal Heterogeneous Message Passing (時序異質消息傳遞)

3.1 Temporal Embedding (時間嵌入)

在 TRANS 中，時間嵌入分為兩部分：

- **(1) Time2Vec 嵌入（適用於就診節點）** 將每次就診的具體時間點進行時間向量化，並通過 Time2Vec 方法進行編碼。

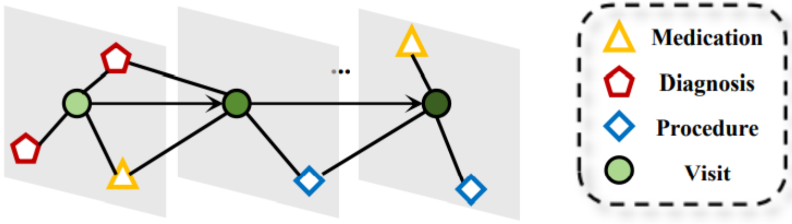
$$T(t) = \sqrt{\frac{1}{d}}\left[\cos(\omega_1 t), \sin(\omega_1 t), \ldots, \cos(\omega_d t), \sin(\omega_d t)\right]$$

其中，$\omega_1, \ldots, \omega_d$ 是可學習的參數。編碼結果將與就診節點初始特徵 h_v 進行拼接，得到時間編碼向量。

- **(2) 邊的時間編碼（適用於連接不同就診的邊）** 針對時間索引的邊進行時間編碼，具體公式如下：

$$\alpha_t = L(T(t))$$

其中，α_t 是經線性層 $L(\cdot)$ 映射後的時間因子，表示在消息傳遞過程中對時間的權重控制。



(c) Temporal Graphical Representation

Methodology

Temporal Heterogeneous Message Passing（時序異質消息傳遞）

3.2 Heterogeneous Message Passing (異質消息傳遞)

在進行異質圖的消息傳遞時，模型使用了一種基於注意力機制的異質消息傳遞方式。主要分為以下幾個步驟：

1. **節點特徵提取與鄰居信息聚合：**

$$q_v = L(h_v^{(l)}), \quad k_{e_t} = L(h_{e_t}^{(l)})$$

- $h_v^{(l)}$ 表示當前就診節點的特徵，$h_{e_t}^{(l)}$ 表示鄰居醫療事件節點的特徵。

2. **注意力計算：**

$$Att_i(e_t, v_t) = \frac{\alpha_t \cdot (q_v W^{\mu(e)} k_{e_t}^\top)}{\sqrt{d}}$$

- $W^{\mu(e)}$ 是根據醫療事件類型（診斷、藥物、程序）來決定的權重矩陣，α_t 是時間因子，最終通過 Softmax 得到歸一化權重。

3. **消息聚合：**

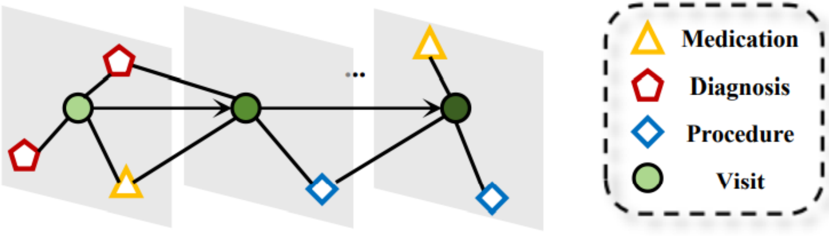
$$\tilde{h}_v^{(l+1)} = \text{Softmax}_{\forall e_t \in N(v_t)}(Att_i(e_t, v_t)) \cdot L(h_{e_t}^{(l)})$$

其中，$\tilde{h}_v^{(l+1)}$ 為多頭注意力機制中各頭的最終聚合特徵。

4. **特徵更新：**

$$h_v^{(l+1)} = \gamma \cdot L(\sigma(\tilde{h}_v^{(l+1)})) + (1-\gamma) \cdot h_v^{(l)}$$

- 使用跳躍連接（skip connection）進行融合，γ 是平衡係數。



(c) Temporal Graphical Representation

Methodology

4. 空間編碼 (Spatial Encoding)

1. **全局位置編碼 (Global Positional Encoding)：**
   - 使用拉普拉斯矩陣 (Laplacian Matrix) 的特徵值來獲取節點在圖上的全局位置：

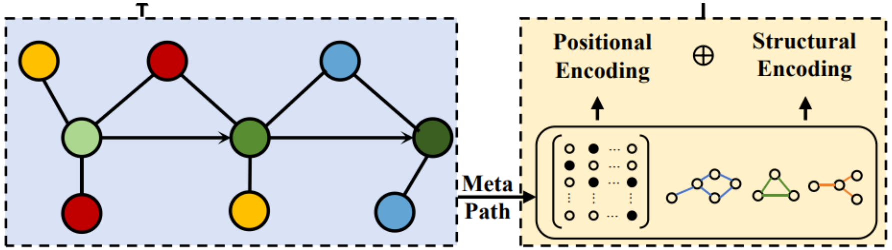
$$P(e) = [v_e^{(1)}, v_e^{(2)}, \ldots, v_e^{(k)}]$$

     - `P(e)` 為節點 `e` 的位置編碼。
     - `v_e^{(i)}` 為第 `i` 個特徵向量的元素。

2. **局部結構編碼 (Local Structural Encoding)：**
   - 使用隨機遊走 (Random Walk) 來捕捉節點在圖中的局部結構資訊：

$$S(e) = [d_e^{(1)}, d_e^{(2)}, \ldots, d_e^{(k)}]$$

     - `S(e)` 為節點 `e` 的結構特徵。
     - `d_e^{(i)}` 為隨機遊走步驟中與自己連接的機率。



Patient Graph Spatial Encoding

Methodology

5. 異質時間圖卷積模型：TRANS

1. 模型架構：
   - 由三個主要模組組成：
     1. 異質時間訊息傳遞模組 (Heterogeneous Temporal Message Passing Module)：
        - 聚合節點間的時間特徵。
     2. 空間編碼模組 (Spatial Encoder Module)：
        - 整合結構和位置特徵到節點表示中。
     3. 病患解釋器 (Patient Graph Explainer)：
        - 用於辨識哪些診斷、治療或用藥節點對模型預測最重要。
2. 預測模組 (Predictor)：
   - 使用最終圖卷積得到的節點特徵進行預測。