



# به‌سوی سیستم سلامت یادگیری مبتنی بر هوش مصنوعی برای پیش بینی مرگ و میر در سطح جمعیت با استفاده از الکتروکاردیوگرام

ویجی سان<sup>۱,۶</sup>، سونیل واسو کالمادی<sup>۱,۲,۳,۴,۵</sup>، نریمان سپهروند<sup>۲,۴</sup>، امیر سلیمی ایوسف نادمی<sup>۱</sup>، کوین یابینی<sup>۲,۴</sup>، جاستین ای. ازکوویتز<sup>۲,۴</sup>، اسل گرینر<sup>۱,۳</sup>، آبرام هیندل<sup>۱</sup>، فیلی ای. مک آلیستر<sup>۲,۴</sup>، رویندر کی. ساندوهه<sup>۲,۴</sup> و پادما کائول<sup>۲,۴</sup>✉

امکان سنجی و ارزش پیوند داده های الکتروکاردیوگرام (ECG) به داده های سلامت اداری طولی در سطح جمعیت برای تسهیل توسعه یک سیستم مراقبت های بهداشتی آموزشی به طور کامل بررسی نشده است. ما مدل های یادگیری ماشین مبتنی بر ECG را برای پیش بینی خطر مرگ و میر در بین بیمارانی که به هر دلیلی به بخش اورژانس یا بیمارستان مراجعه می کنند، توسعه دادیم. با استفاده از ردیابی ها و اندازه گیری های ECG 12 لید از ECG 1605268 از 748773 اپیزود مراقبت های بهداشتی از 244077 بیمار (2007-2020) در آلبرتا، کانادا، ما مدل های یادگیری عمیق (DL) مبتنی بر ResNet و XGBoost مبتنی بر تقویت گرادیان (XGB) را برای پیش بینی مرگ و میر 30 روزه، 1 ساله و 5 ساله توسعه و اعتبارسنجی کردیم. مدل های مربوطه مرگ و میر 30 روزه، 1 ساله و 5 ساله به ترتیب روی ۱۴۶۱۷۳، ۱۴۱۰۷۲ و ۱۱۱۰۲۰ بیمار آموزش داده شدند و روی ۹۷۱۴۴، ۸۹۳۷۹ و ۵۵۶۵۰ بیمار ارزیابی شدند. در گروه ارزیابی، ۷.۶٪

۱۷.۳٪ و ۳۲.۹٪ از بیماران به ترتیب تا 30 روز، 1 سال و 5 سال فوت کردند. مدل های ResNet که تنها بر اساس داده های ECG بودند، عملکرد خوب تا عالی با مساحت زیر منحنی مشخصه عملکرد گیرنده (AUROC) (0.808-0.816) (0.۸۱۲)، CI: 0.838-0.848 (0.843 و 0.۷۹۸) (0.792-0.803) به ترتیب برای پیش بینی 30 روزه، 1 ساله و 5 ساله داشتند؛ و نسبت به مدل های XGB که بر اساس اندازه گیری های ECG بودند، با (0.780-0.788) (0.۷۸۴)، AUROC 0.782 (0.776-0.789) و ۷۴۶٪ (0.751-0.740) برتر بودند. این مطالعه اعتبار مدل های پیش بینی مرگ و میر DL مبتنی بر ECG را در سطح جمعیت نشان می دهد که می توانند برای پیش بینی در نقطه مراقبت مورد استفاده قرار گیرند.

npj: پزشکی دیجیتال (۲۰۲۳) ۳:۲۱۱ | <https://doi.org/10.1038/s41746-023-00765-3>

مرگ و میر (۱ و ۵ ساله). ما سودمندی افزوده ی گنجاندن مقادیر آزمایشگاهی (lab) موجود در زمان ECG را در عملکرد پیش بینی مدل ها بررسی کردیم و عملکرد مدل ها را در زیرگروه های جنسی و تشخیصی خاص بررسی کردیم.

## نتایج

### ویژگی ها و پیامدهای بیمار

ویژگی های گروه های بیمار مورد استفاده در مطالعه در جدول شرح داده شده است. ۱. میانگین سنی بیماران در زمان گرفتن نوار قلب (ECG) در هر دو مجموعه ی توسعه و عدم توسعه، ۶۵.۸ سال بود. با این حال، به یاد داشته باشید که ما به طور تصادفی یک نوار قلب (ECG) برای هر بیمار از مجموعه ی توسعه انتخاب کردیم تا برای ارزیابی های نهایی استفاده شود. میانگین سنی در این مجموعه ی اخیر کمی کمتر و ۶۲.۶ سال بود. این احتمالاً به این دلیل است که بیماران مسن تر نوار قلب بیشتری نسبت به بیماران جوان تر داشتند. به طور مشابه، مردان نوار قلب بیشتری داشتند، بنابراین نسبت مردان در مجموعه ی ارزیابی تصادفی کمی کمتر از مجموعه های توسعه یا عدم توسعه بود (۵۴.۷٪ در مقابل ۵۶.۷٪). این الگو برای برخی از اندازه گیری های ECG (مثلاً میانگین مدت زمان QRS: 97.9 در مقابل 101.3 میلی ثانیه؛ فاصله QT 395.1 در مقابل 399.8 میلی ثانیه)، بیماری های همراه (مثلاً نارسایی قلبی: 4.1٪ در مقابل 6.2٪؛ فیبریلاسیون دهلیزی: 9.2٪ در مقابل 15.5٪) و همچنین مقادیر آزمایشگاهی (مثلاً تریوپوئین I 2.0 در مقابل 2.2 میکروگرم در لیتر؛ کراتینین 107.5 در مقابل 116.5 میکرومول در لیتر، جدول) مشاهده شد. ۲.

مدل های مربوط به مرگ و میر 30 روزه، 1 ساله و 5 ساله به ترتیب روی ۱۴۶۱۷۳، ۱۴۱۰۷۲ و ۱۱۱۰۲۰ بیمار آموزش داده شدند و روی ۹۷۱۴۴، ۸۹۳۷۹ و ۵۵۶۵۰ بیمار ارزیابی شدند (شکل ۱). در مجموعه ارزیابی ما که شامل یک نوار قلب تصادفی به ازای هر بیمار بستری شده بود، به ترتیب ۷۳۹۹ (۷.۶٪)، ۱۵۵۰۶ (۱۷.۳٪) و ۱۸۳۰۲ (۳۲.۹٪) در 30 روز، 1 سال و 5 سال فوت کرده بودند. به طور مشابه، برای زیرمجموعه ای از نوار قلب ها برای

## مقدمه

توسعه یک سیستم سلامت یادگیرنده، سیستمی که چرخه ای از جمع آوری و تجزیه و تحلیل منظم داده های سلامت را دنبال می کند تا دانش جدیدی تولید کند که می تواند برای تصمیم گیری های بهداشتی یا بهبود سیستم کار گرفته شود، اولویت اصلی کانادا و سایر کشورها است. سیستم مراقبت های بهداشتی همگانی کانادا که توسط دولت تأمین مالی می شود، امکان پیوند قطعی داده های مراقبت های بهداشتی از مراکز مختلف مراقبت های بهداشتی (بیمارستان ها، بخش های اورژانس، کلینیک های سرپایی و مطب پزشکان) با ثبت های بیمه و آمار حیاتی را فراهم کرده است تا پیش بینی کننده های مرگ و میر را شناسایی کرده و الگوریتم های طبقه بندی ریسک را توسعه دهد. ۲-۴ این مدل ها با در دسترس بودن ادعاهای دارویی و داده های آزمایشگاهی در سطح جمعیت در برخی استان ها بهبود یافته اند. ۵-۸

الکتروکاردیوگرام (ECG) یک ابزار تشخیصی کم هزینه و در دسترس است که در اکثر بیماران در طول ویزیت مراقبت های حاد انجام می شود و حاوی اطلاعات مهمی در مورد ساختار و فعالیت الکتریکی قلب است. ۹ در سال های اخیر، پیشرفت های نمایی در منابع محاسباتی و فناوری های یادگیری ماشین، همراه با مجموعه داده های ECG دیجیتالی، فرصت هایی را برای پیش بینی های تشخیصی و پیش آگهی مبتنی بر ECG فراهم کرده است. ۱۰-۱۳ با این حال، امکان سنجی و ارزش پیوند دادن داده های ECG به داده های سلامت اداری در سطح جمعیت طولی برای کمک به پزشکان در تصمیم گیری در مورد مراقبت با هدف تکمیل چرخه کیفیت و تسهیل یک سیستم مراقبت های بهداشتی آموزشی قبلاً بررسی نشده است. ۱۴،۱

این امر ما را بر آن داشت تا از یک گروه بزرگ جمعیتی از بیماران دارای بیمه سلامت همگانی که به بخش های اورژانس یا بیمارستان ها مراجعه می کردند، برای توسعه مدل های یادگیری ماشین مبتنی بر ECG جهت پیش بینی کوتاه مدت (30 روزه) و بلندمدت استفاده کنیم.

اگرچه علوم کامپیوتر، دانشگاه آلبرتا، ادمونتون، آلبرتا، کانادا؛ مرکز VIGOUR کانادا، دانشکده پزشکی، دانشگاه آلبرتا، ادمونتون، آلبرتا، کانادا؛ موسسه هوش ماشینی آلبرتا، دانشگاه آلبرتا، ادمونتون، آلبرتا، کانادا؛ گروه پزشکی، دانشگاه آلبرتا، ادمونتون، آلبرتا، کانادا؛ موسسه قلب اسمیت، سیستم بیمارستانی مرکز پزشکی سدرز-سینای، لس آنجلس، کالیفرنیا، ایالات متحده آمریکا؛ این نویسندگان به طور مساوی مشارکت داشتند؛ ویجی سان، سونیل واسو کالمادی. ✉ [pkaul@ualberta.ca](mailto:pkaul@ualberta.ca)

جدول ۱. ویژگی های گروه های بیمار مورد استفاده در مطالعه.

| اطلاعات کامل (ن = ۱,۶۰۵,۲۶۸)               | مجموعه توسعه (ن = ۹۶۴۷۴۱) | مجموعه نگهدارنده (ن = ۶۴۰,۵۲۷) | holdout تصادفی به ازای هر بیمار در مجموعه ECG (ن = ۹۷,۶۳۱) |                 |
|--|---------------------------|--------------------------------|--|-----------------|
| سن (سال)                                   | ۱۷.۲۵ ± ۶۵.۸۰             | ۱۷.۲۲ ± ۶۵.۷۷                  | ۱۷.۲۹ ± ۶۵.۸۵  | ۱۸.۵۹ ± ۶۲.۵۷   |
| جنسیت (مرد به درصد)                        | ۵۶.۷۳                     | ۵۶.۸۱                          | ۵۶.۶   | ۵۴.۷۳           |
| اندازه گیری های ECG                        |                           |                                |  |                 |
| ضربان دهلیزی                               | ۴۶.۱۵ ± ۸۵.۶۰             | ۴۶.۱۱ ± ۸۵.۵۶                  | ۴۶.۲۰ ± ۸۵.۶۷  | ۴۰.۳۰ ± ۸۴.۰۶   |
| مدت زمان P                                 | ۱۱۶.۶۰ ± ۱۵۵.۹۲           | ۱۱۶.۳۹ ± ۱۵۶.۰۰                | ۱۱۶.۹۱ ± ۱۵۵.۷۹  | ۱۱۴.۰۶ ± ۱۶۳.۹۶ |
| فاصله RR                                   | ۲۱۳.۱۱ ± ۷۹۰.۸۱           | ۲۱۲.۶۳ ± ۷۹۰.۹۰                | ۲۱۳.۸۳ ± ۷۹۰.۶۸  | ۲۰۴.۳۵ ± ۷۹۰.۸۹ |
| شروع موج Q                                 | ۶.۵۱ ± ۵۰.۸.۸۴            | ۶.۲۹ ± ۵۰.۸.۸۲                 | ۶.۸۲ ± ۵۰.۸.۸۷   | ۶.۱۷ ± ۵۰.۹.۰۴  |
| فاصله QT اصلاح شده با ضربان قلب فریدریشیا  | ۳۸.۰۵ ± ۴۳۴.۸۶            | ۳۸.۰۴ ± ۴۳۴.۹۶                 | ۳۸.۰۷ ± ۴۳۴.۷۱   | ۳۵.۲۳ ± ۴۲۹.۵۵  |
| ضربان قلب                                  | ۲۳.۲۲ ± ۸۱.۶۴             | ۲۳.۱۸ ± ۸۱.۶۱                  | ۲۳.۲۸ ± ۸۱.۶۹  | ۲۱.۹۴ ± ۸۱.۱۳   |
| فاصله PR                                   | ۳۸.۴۶ ± ۱۶۹.۳۴            | ۳۷.۶۶ ± ۱۶۹.۴۴                 | ۳۹.۶۵ ± ۱۶۹.۱۸   | ۳۳.۶۵ ± ۱۶۵.۹۹  |
| مدت زمان QRS                               | ۲۴.۲۶ ± ۱۰۱.۳۶            | ۲۴.۲۳ ± ۱۰۱.۴۰                 | ۲۴.۳۰ ± ۱۰۱.۳۱   | ۲۱.۶۶ ± ۹۷.۸۹   |
| فاصله QT                                   | ۵۴.۸۳ ± ۳۹۹.۸۱            | ۵۴.۷۴ ± ۳۹۹.۹۴                 | ۵۴.۹۶ ± ۳۹۹.۶۳   | ۵۱.۴۱ ± ۳۹۵.۱۰  |
| فاصله QT اصلاح شده با نرخ ضربان قلب Bazett | ۴۰.۰۹ ± ۴۵۵.۰۲            | ۴۰.۰۹ ± ۴۵۵.۱۰                 | ۴۰.۰۹ ± ۴۵۴.۸۹   | ۳۷.۱۶ ± ۴۴۹.۲۳  |
| محور P جلویی                               | ۳۵.۵۲ ± ۴۴.۸۵             | ۳۵.۴۱ ± ۴۴.۸۱                  | ۳۵.۶۹ ± ۴۴.۹۱  | ۳۲.۱۹ ± ۴۵.۶۶   |
| محور QRS جلویی در ۴۰ میلی ثانیه اولیه      | ۴۶.۳۰ ± ۲۷.۵۰             | ۴۶.۳۷ ± ۲۷.۴۴                  | ۴۶.۲۰ ± ۲۷.۵۹  | ۴۲.۱۲ ± ۲۸.۴۸   |
| محور QRS جلویی در ترمینال ۴۰ میلی ثانیه    | ۸۸.۱۵ ± ۴۵.۳۶             | ۸۸.۲۸ ± ۴۵.۷۴                  | ۸۷.۹۶ ± ۴۴.۸۰  | ۸۴.۶۴ ± ۴۶.۲۴   |
| محور QRS جلویی                             | ۵۴.۳۷ ± ۱۹.۹۸             | ۵۴.۳۸ ± ۲۰.۰۴                  | ۵۴.۳۷ ± ۱۹.۸۸  | ۴۹.۸۱ ± ۲۲.۶۸   |
| محور موج ST از جلو                         | ۸۸.۲۳ ± ۹۰.۹۴             | ۸۸.۰۷ ± ۹۰.۹۸                  | ۸۸.۴۸ ± ۹۰.۸۷  | ۸۵.۱۱ ± ۷۹.۲۳   |
| محور T از جلو                              | ۶۷.۷۶ ± ۵۵.۷۰             | ۶۷.۶۰ ± ۵۵.۴۸                  | ۶۸.۰۰ ± ۵۵.۰۳  | ۵۹.۹۰ ± ۴۷.۹۷   |
| محور افقی P                                | ۴۷.۳۰ ± ۲۰.۶۹             | ۴۷.۱۴ ± ۲۰.۶۳                  | ۴۷.۵۲ ± ۲۰.۷۷  | ۴۱.۱۵ ± ۲۱.۰۱   |
| محور QRS افقی در ۴۰ میلی ثانیه اولیه       | ۴۸.۳۹ ± ۲۷.۷۹             | ۴۸.۱۷ ± ۲۷.۸۶                  | ۴۸.۷۱ ± ۲۷.۶۹  | ۳۰.۲۰ ± ۴۲.۶۸   |
| محور QRS افقی در ترمینال ۴۰ میلی ثانیه     | ۱۲۹.۵۰ ± ۳۴.۱۰            | ۱۲۹.۴۰ ± ۳۳.۹۴                 | ۱۲۹.۶۶ ± ۳۴.۳۵   | ۱۲۵.۵۸ ± ۲۶.۶۷  |
| محور QRS افقی                              | ۷۸.۱۹ ± ۰.۹۱ -            | ۷۷.۹۱ ± ۱.۱۱ -                 | ۷۸.۶۲ ± ۰.۶۱ -   | ۶۹.۲۷ ± ۴.۰۳ -  |
| محور افقی موج ST محور افقی T               | ۶۴.۹۹ ± ۹۷.۰۲             | ۶۵.۰۰ ± ۹۶.۹۸                  | ۶۴.۹۷ ± ۹۷.۰۹  | ۶۰.۲۰ ± ۹۱.۷۵   |
| بیماری های همراه                           | ۵۸.۹۸ ± ۶۴.۴۶             | ۵۸.۹۶ ± ۶۴.۲۷                  | ۵۹.۰۱ ± ۶۴.۷۳  | ۵۳.۰۵ ± ۵۹.۳۰   |
| بیماری عروق محیطی                          | (۲.۰۹٪) ۳۳۵۱۸             | (۲.۰۴٪) ۱۹,۷۱۴                 | (۲.۱۶٪) ۱۳,۸۰۴   | (۲.۲۰٪) ۲۱۴۴    |
| بیماری عروق مغزی فشار خون بالا             | (۳.۳۹٪) ۵۴,۳۴۹            | (۳.۴۴٪) ۳۳,۱۹۱                 | (۳.۳۰٪) ۲۱,۱۵۸   | (۴.۳۶٪) ۴۲۵۲    |
| زوال عقل                                   | (۲۱.۸۶٪) ۳۵۰,۸۵۹          | (۲۱.۸۰٪) ۲۱۰,۲۷۵               | (۲۱.۹۵٪) ۱۴۰,۵۸۴   | (۱۵.۷۶٪) ۱۵۳,۸۷ |
| بیماری مزمن ریوی دیابت شیرین               | (۸.۳۵٪) ۱۳۳,۹۶۳           | (۸.۳۰٪) ۸۰,۰۳۷                 | (۸.۴۲٪) ۵۳,۹۲۶   | (۹.۰۶٪) ۸۸,۴۹   |
| بیماری کلیوی                               | (۱.۹۸٪) ۱۳,۷۶۴            | (۱.۹۹٪) ۱۹,۲۱۵                 | (۱.۹۶٪) ۱۲,۵۴۹   | (۲.۱۳٪) ۲۰,۷۸   |
| بیماری کبد                                 | (۷.۴۹٪) ۱۲۰,۶۶۰           | (۷.۴۵٪) ۷۱,۸۶۰                 | (۷.۵۶٪) ۴۸,۴۰۰   | (۵.۸۲٪) ۵۶,۸۴   |
| سرطان                                      | (۱۰.۱۷٪) ۱۶۳,۲۶۲          | (۱۰.۰۵٪) ۹۶,۹۲۴                | (۱۰.۳۶٪) ۶۶,۳۳۸  | (۹.۰۱٪) ۸۸,۰۰   |
| استومی بدون انسداد قطعه ST                 | (۱.۲۶٪) ۲۰,۲۶۸            | (۱.۲۵٪) ۱۲,۰۶۲                 | (۱.۲۸٪) ۸,۲۰۶  | (۱.۱۱٪) ۱۰,۷۹   |
| STEMI                                      | (۱.۱۸٪) ۱۸,۹۰۵            | (۱.۲۱٪) ۱۱,۷۰۷                 | (۱.۱۲٪) ۷,۱۹۸  | (۱.۳۸٪) ۱۱,۳۴۶  |
| نارسایی قلبی                               | (۵.۸۵٪) ۹۳,۹۴۶            | (۵.۷۷٪) ۵۵,۶۳۲                 | (۵.۹۸٪) ۳۸,۳۱۴   | (۸.۹۱٪) ۸۶,۹۹   |
| فیبریلاسیون دهلیزی                         | (۱۰.۱۱٪) ۱۶۲,۲۷۴          | (۱۰.۰۴٪) ۹۶,۸۲۸                | (۱۰.۲۲٪) ۶۵,۴۴۶  | (۶.۶۹٪) ۶۵,۳۴   |
|  | (۶.۲۴٪) ۱۰۰,۲۰۶           | (۶.۲۶٪) ۶۰,۳۸۱                 | (۶.۲۲٪) ۳۹,۸۲۵   | (۴.۱۵٪) ۴۰,۴۹   |
|  | (۱۵.۵۳٪) ۲۴۹,۳۲۵          | (۱۵.۵۵٪) ۱۵۰,۰۵۵               | (۱۵.۵۰٪) ۹۹,۲۷۰  | (۹.۱۸٪) ۸۹,۵۸   |

اندازه گیری های ECG و بیماری های همراه بر اساس تعداد نمونه های ECG بیان می شوند.

الف. مقادیر توصیفی به صورت میانگین ± انحراف معیار یا استفاده از یک تکرار نمونه برداری تصادفی ECG در مجموعه داده ی مورد نظر ارائه شده اند. نوار قلب (ECG) الکتروکاردیوگرام، ن عدد، انستومی بدون انسداد قطعه ST انفارکتوس میوکارد بدون بالا رفتن قطعه ST. ST انفارکتوس میوکارد با بالا رفتن قطعه ST.

## مقایسه مدل ها

مقایسه عملکرد مدل ها در شکل ۱ ارائه شده است. ۲ و جدول ۳ ما از ویژگی های سن و جنس به تنهایی برای ایجاد یک عملکرد مدل پایه استفاده کردیم که مساحت زیر منحنی مشخصه عملکرد گیرنده (AUROC) (میانگین و فاصله اطمینان ۹۵٪) آن برای ۳۰ روز، ۰.۷۱۶ (۰.۶۴۶-۰.۷۱۵) بود.

باتوجه به اینکه مقادیر آزمایشگاهی در دسترس بودند، مدل های مرگ و میر ۳۰ روزه، ۱ ساله و ۵ ساله به ترتیب روی ۸۴۲۳۹، ۷۸۳۴۰ و ۴۲۷۴۲ بیمار آموزش داده شدند و روی ۴۹۷۴۸، ۵۶۰۵۹ و ۲۱۷۹۶ بیمار ارزیابی شدند (شکل تکمیلی ۱). در مجموعه ارزیابی آزمایشگاهی، به ترتیب ۴۹۰۷ (۸.۸٪)، ۹۶۶۸ (۱۹.۴٪) و ۸۳۵۲ (۳۸.۳٪) در نقاط زمانی فوق الذکر فوت کرده بودند.

جدول ۲. ویژگی‌های آزمایشگاهی گروه‌های بیمار مورد استفاده در مطالعه (به صورت میانگین  $\pm$  انحراف معیار).

| اطلاعات کامل (ن = ۶۰۱,۳۰۷)                     | مجموعه توسعه (ن = ۳۶۱۵۸۵) | مجموعه نگهدارنده (ن = ۲۳۹,۷۲۲) | holdout تصادفی به ازای هر بیمار در مجموعه ECG (ن = ۳۰,۰۷۶) |
|--|---------------------------|--------------------------------|--|
| سن (سال)                                       | ۱۷.۵۷ $\pm$ ۶۴.۸۲         | ۱۷.۵۳ $\pm$ ۶۴.۹۰              | ۱۷.۵۷ $\pm$ ۶۴.۷۴  |
| جنسیت (مرد به درصد)                            | ۵۵.۸۴                     | ۵۵.۸۱                          | ۵۵.۸۸  |
| اندازه گیری‌های آزمایشگاهی                     |                           |                                |  |
| میزان فیلتراسیون گلوبولین (میلی لیتر در دقیقه) | ۲۸.۰۳ $\pm$ ۶۶.۰۱         | ۲۸.۰۴ $\pm$ ۶۵.۸۵              | ۲۷.۹۴ $\pm$ ۶۶.۶۶  |
| کراتینین (امول/لیتر)                           | ۱۱۷.۰۳ $\pm$ ۱۱۶.۴۶       | ۱۱۷.۲۰ $\pm$ ۱۱۶.۶۲            | ۱۱۷.۵۲ $\pm$ ۱۰۰.۶۳  |
| هموگلوبین (گرم در لیتر)                        | ۲۳.۲۳ $\pm$ ۱۲۶.۵۳        | ۲۳.۲۷ $\pm$ ۱۲۶.۴۶             | ۲۳.۳۲ $\pm$ ۱۲۷.۸۱   |
| پتاسیم (میلی مول در لیتر)                      | ۰.۶۴ $\pm$ ۴.۰۲           | ۰.۶۴ $\pm$ ۴.۰۲                | ۰.۶۳ $\pm$ ۳.۹۹  |
| سدیم (میلی مول در لیتر)                        | ۴.۳۵ $\pm$ ۱۳۷.۳۹         | ۴.۳۶ $\pm$ ۱۳۷.۳۸              | ۴.۳۹ $\pm$ ۱۳۷.۳۹  |
| تروپونین I (میکروگرم در لیتر)                  | ۱۰.۰۷ $\pm$ ۲.۲۴          | ۱۰.۰۹ $\pm$ ۲.۲۳               | ۹.۴۹ $\pm$ ۱.۹۹  |

نوار قلب (ECG) الکتروکاردیوگرام، میزان ترشح گلوبولین (GFR) میزان فیلتراسیون گلوبولین، ن.شماره.

گروه مقاومت. میزان مرگ و میر در بین مردان بیشتر از زنان بود (شکل تکمیلی ۴). به طور کلی، مدل‌های پیش‌آگهی در مردان کمی بهتر از زنان عمل کردند (شکل ۴.۱۰۰).

عملکرد مدل با افزودن ویژگی‌های آزمایشگاهی

ما بهبود عملکرد مدل را با افزودن ویژگی‌های آزمایشگاهی به مدل‌های ECG بررسی کردیم. ما XGB را با سن، جنس و ویژگی‌های آزمایشگاهی به عنوان مدل پایه برای مقایسه در نظر گرفتیم. مدل‌های ECG در تمام نقاط زمانی، حتی بدون مقادیر آزمایشگاهی، AUROC بالاتری نسبت به مدل‌های پایه داشتند، اما این تفاوت برای پیش‌بینی‌های با دامنه طولانی‌تر کمتر بود. افزودن ویژگی‌های آزمایشگاهی به طور قابل توجهی عملکرد مدل را در کل، در بین مدل‌ها و نقاط زمانی بهبود بخشید، با این حال، افزایش عملکرد از نظر بزرگی اندک بود (به طور متوسط 0.99٪، آزمون DeLong، همه پ > 0.001). باز هم، مدل کلی DL با ردیابی‌های ECG، سن، جنس و آزمایشگاه، بهترین عملکرد را در این مقایسه داشت (شکل ۱.۵ و جدول تکمیلی ۱).

توضیحات مدل

شکل ۶ نتایج نگاشت فعال سازی کلاس با وزن گرادیان (GradCAM) را نشان می‌دهد و نواحی ECG را که سهم و ارتباط بیشتری با عملکرد پیش‌بینی مرگ و میر مدل دارند، برجسته می‌کند. نواحی برجسته شده مختص به سرب نیستند و بر اساس داده‌های هر ۱۲ لید محاسبه شده‌اند. تجزیه و تحلیل SHAP (SHAPley Additive Explanations) از مدل XGB نشان داد که سن بالاتر، فاصله RR کمتر، محور QRS افقی (اثر شرطی)، فاصله QT اصلاح شده با نرخ Bazett بالاتر، جنس مذکر و فاصله PR کمتر بیشترین سهم را در مرگ و میر ۱ ساله داشته‌اند. افزودن ویژگی‌های آزمایشگاهی، سهم هموگلوبین پایین‌تر، میزان فیلتراسیون گلوبولین (GFR) پایین‌تر، تروپونین I پایین‌تر، کراتینین بالاتر، سدیم بسیار بالا یا پایین و پتاسیم بالا را برجسته کرد (شکل ۱.۷ برای مرگ و میر ۱ ساله و شکل تکمیلی ۵ برای سایر نقاط زمانی).

تحلیل‌های تکمیلی

ما چندین تحلیل تکمیلی برای نشان دادن استحکام مدل‌های خود انجام دادیم. ابتدا، عملکرد مدل‌های خود را در مقایسه با معماری DL طراحی شده سفارشی که توسط راگونات و همکارانش به کار گرفته شده بود، محک زدیم.<sup>۱۴</sup> آموزش مجموعه داده‌های ما با معماری مشخص شده در مطالعه آنها، منجر به AUROC برابر با 0.796 (0.792، 0.800) برای پیش‌بینی مرگ و میر ۱ ساله شد، در مقایسه با 0.826 (0.822، 0.83) با معماری ما. مدل DL مبتنی بر ResNet در مقایسه با معماری DL به کار رفته توسط راگونات و همکارانش برای هر سه نقطه زمانی، بهبود کوچک اما از نظر آماری معنی‌داری در عملکرد AUROC نشان داد (آزمون DeLong، همه پ > ۰.۰۰۱، شکل تکمیلی ۶).

دوم، ما مدل DL اولیه خود را بر روی موارد زیر ارزیابی کردیم: (الف) یک مجموعه پشتیبان که شامل ECG‌های بی‌کیفیتی بود که قبلاً وجود داشتند

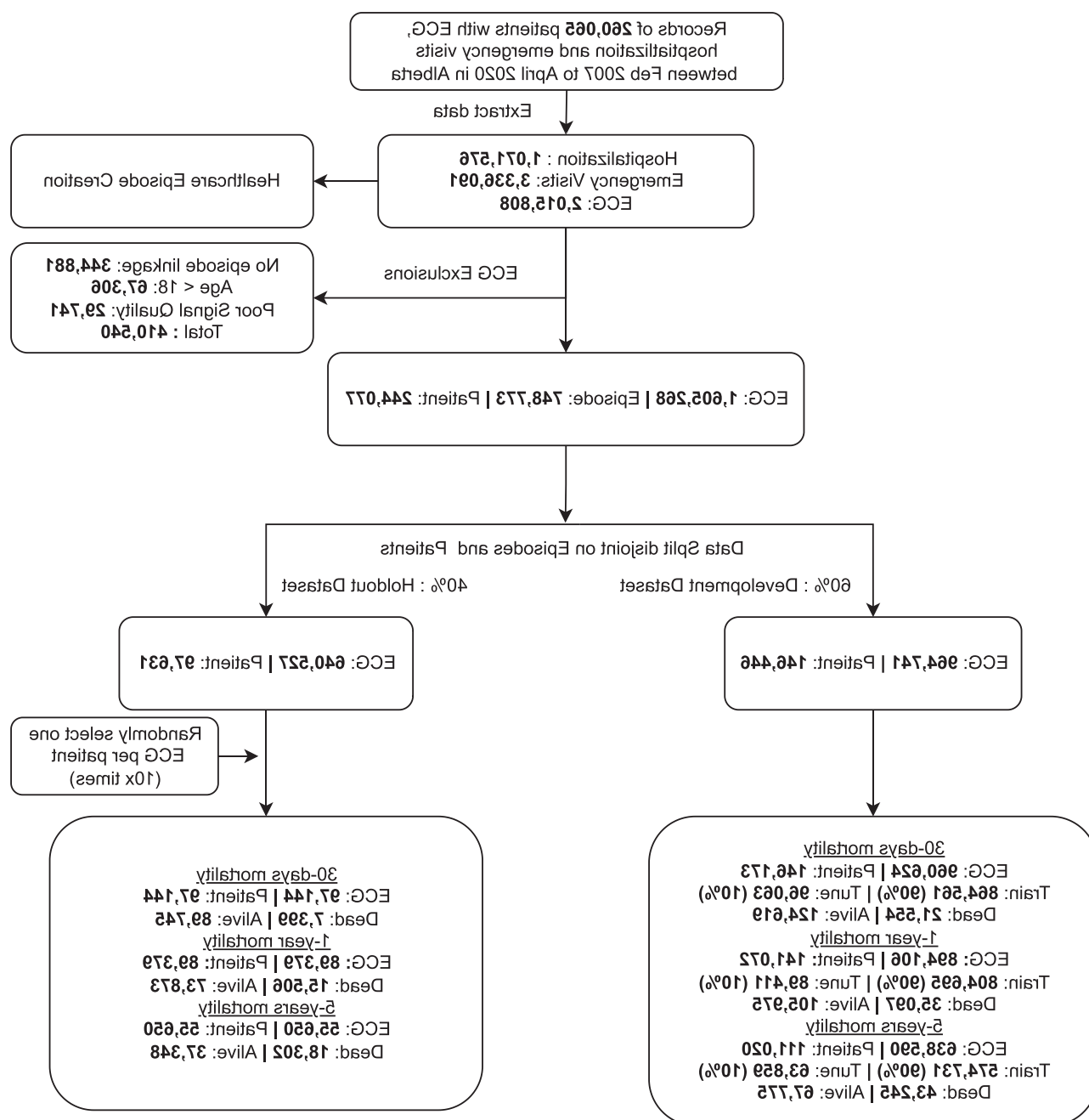
(0.704-0.723) برای مرگ و میر ۱ ساله و 0.776 (0.765-0.787) برای مرگ و میر ۵ ساله. مدل یادگیری عمیق (DL) با ردیابی‌های ECG به تنهایی، عملکرد قابل توجهی بالاتری با 0.816 (0.812، 0.838-0.848) و AUROC 0.843 و 0.798 (0.792-0.803) به ترتیب برای پیش‌بینی 30 روزه، 1 ساله و 5 ساله داشت (آزمون DeLong، همه پ > 0.001). استفاده از سن و جنس به همراه ردیابی‌های ECG، بهبودهای کوچک اما قابل توجهی را با 0.822-0.830 (0.826، 0.847-0.852) و AUROC 0.852 و 0.824 (0.822-0.824) برای سه نقطه زمانی نشان داد (آزمون DeLong، همه پ > 0.001). همه DeLong، برای هر سه نقطه زمانی نشان داد (آزمون ECG با اندازه گیری‌های XGB) XGBoost عملکرد به طور قابل توجهی بهتری نسبت به ECG با ردیابی DL پ > ۰.۰۰۱).

مدل DL با ردیابی‌های ECG، سن و جنس بهترین مدل در این مقایسه بود، با AUROC‌هایی که به طور مداوم بالاتر از 0.82 بودند. مدل XGB با اندازه گیری‌های ECG به مدت 1 سال عملکرد بهتری نسبت به مدل فقط سن و جنس نداشت (آزمون DeLong، پ = ۰.۵۷) و به مدت ۵ سال به طور قابل توجهی بدتر از سن و جنس بود (آزمون دلونگ، پ > 0.001)؛ با این حال، DL با ردیابی ECG هنوز اطلاعات مرتبطی را برای پیش‌بینی ارائه می‌دهد و در تمام نقاط زمانی به طور قابل توجهی از مدل پایه سن و جنس بهتر عمل می‌کند (آزمون DeLong، پ > ۰.۰۰۱) - که بدین ترتیب بر سودمندی پیش‌آگهی مدل‌های DL مبتنی بر ردیابی ECG نسبت به مدل‌های معمول با استفاده از اندازه گیری‌های ECG تأکید می‌کند. جدول ۳ عملکرد برتر مدل‌های DL با ردیابی‌های ECG را از نظر مساحت زیر منحنی دقت-یادآوری (AUPRC)، امتیاز F1، امتیاز Brier و سایر معیارها نشان می‌دهد.

گروه‌های خطر. ما پنج گروه خطر - گروه‌های خطر «بسیار کم»، «کم»، «متوسط»، «زیاد» و «بسیار زیاد» - را بر اساس 20 درصد نقاط برش (20-0٪، 20-40٪ و غیره) از احتمال پیش‌بینی شده مرگ از مدل‌های اصلی خود (ECG Trace، DL؛ سن، جنس) در مجموعه افراد تحت بررسی (شکل 1) استخراج کردیم.<sup>۳</sup> برای مرگ و میر ۱ ساله، شکل تکمیلی ۲ برای سایر نقاط زمانی، درصد مرگ و میر مشاهده شده در هر گروه خطر پیش‌بینی شده، کالیبراسیون خوبی را با افزایش مداوم در گروه‌های خطر نشان داد (۸.۶٪، ۳۴.۶٪، ۵۲.۳٪، ۷۰.۹٪ و ۷۸.۹٪ مرگ و میر در گروه‌های خطر «بسیار کم»، «کم»، «متوسط»، «زیاد» و «بسیار زیاد» به ترتیب).

عملکرد مدل در زیرگروه‌های تشخیصی و جنسیتی

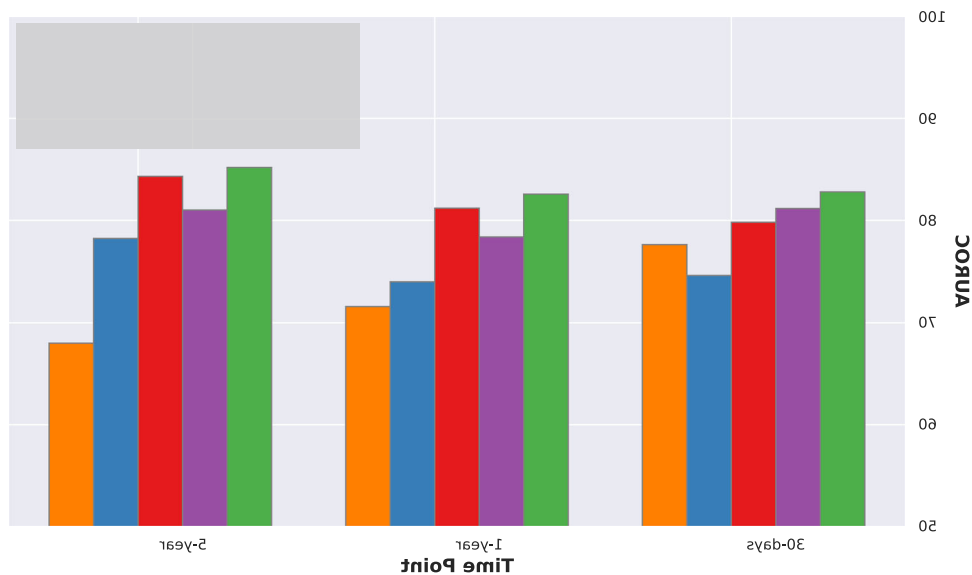
میزان مرگ و میر در گروه‌های تشخیصی مورد نظر به طور قابل توجهی متفاوت بود (شکل تکمیلی 3) و بیماران مبتلا به نارسایی قلبی در هر نقطه زمانی بالاترین میزان مرگ و میر را داشتند. شکل ۴ عملکرد مدل‌های ما را در این زیرگروه‌های تشخیصی مختلف نشان می‌دهد. مدل‌ها در بیماران مبتلا به انفارکتوس میوکارد با بالا رفتن قطعه (STEMI) و انفارکتوس میوکارد بدون بالا رفتن قطعه (NSTEMI) (به ترتیب AUROC برابر با 0.867 و 0.882 برای مرگ و میر 1 ساله) نسبت به گروه کلی بهتر عمل کردند. عملکرد مدل در زیرگروه‌های دیگر (نارسایی قلبی، دیابت و فیبریلاسیون دهلیزی) کمتر از گروه کلی بود.



شکل. اگرچه مطالعه، تقسیم بندی داده ها و طرح ارزیابی، فلوجارت طراحی مطالعه، که اندازه نمونه ها را برای مطالعه کلی، تقسیم بندی های آزمایشی و نتایج مختلف نشان می دهد. ما مجموعه داده های کلی ECG را به یک تقسیم تصادفی ۶۰٪ برای توسعه مدل (برای آموزش ۹۰٪ و زیرمجموعه ۱۰٪) و تنظیم دقیق (۱۰٪) زیرمجموعه ۶۰٪ مدل (۴۰٪ باقی مانده را به عنوان مجموعه داده های پشتیبان برای اعتبارسنجی تقسیم کردیم. انتظار می رود بیمارانی که بیماری های شدیدتری دارند، بیشتر تحت ECG قرار گیرند و در مراحل پایانی که بیماران به طور منظم تری تحت نظر هستند، ممکن است ECG های بیشتری گرفته شود. برای کاهش سوگیری بالقوه در عملکرد مدل به دلیل نمایش متفاوت فنوتیپ های بیمار، مدل های خود را با استفاده از یک ECG انتخاب شده تصادفی برای هر بیمار از چندین دوره بیماری آنها در مجموعه داده های پشتیبان ارزیابی کردیم.

سوم، برای بررسی عملکرد مدل هایمان در بیمارستان های مختلف، اعتبارسنجی خروج از بیمارستان با یک بار مراجعه برای هر یک از دو بیمارستان سطح سوم (H1 و H2) در بین 14 بیمارستان مورد مطالعه را انجام دادیم. برای اطمینان از اینکه مجموعه های آموزشی و آزمایشی ما کاملاً از هم جدا هستند، ECG های بیمارانی را که در طول دوره مطالعه در هر دو بیمارستان آموزشی و آزمایشی بستری شده بودند، از اعتبارسنجی خود حذف کردیم (شکل تکمیلی 7). ما دریافتیم که عملکرد اعتبارسنجی خروج از بیمارستان با عملکرد گزارش شده در مجموعه اعتبارسنجی کلی قابل مقایسه است (جدول تکمیلی 3). در مقایسه با نتایج اعتبارسنجی اصلی، عملکرد AUROC به میزان 1.5 بالاتر بود.

از گروه تحلیل حذف شد؛ و (ب) یک مجموعه ECG با کیفیت پایین به تنهایی (جدول تکمیلی 2). ما دریافتیم که مدل DL ما در برابر مصنوعات سیگنال ECG ومشکلات مربوط به اکتساب نسبتاً مقاوم است. افزودن ECG های با کیفیت پایین (ECG 29741) به مجموعه اصلی (ECG 640527) نتایج AUROC را تغییر نداد (کاهش تنها 0.25٪ برای مرگ و میر 1 ساله). علاوه بر این، ارزیابی بر روی ECG های با کیفیت پایین به تنهایی هنوز هم بیش از 80٪ امتیاز AUROC را برای هر سه نقطه زمانی نشان داد. به ترتیب 3.4٪، 2.6٪ و 2.5٪ کاهش در AUROC در مقایسه با ارزیابی اولیه (که ECG های با کیفیت پایین را حذف کرد) برای نقاط زمانی 30 روزه، 1 ساله و 5 ساله وجود داشت.



شکل ۲ مقایسه عملکرد مدل ها. مقایسه عملکرد مدل AUROC برای مدل های DL و XGB با اندازه گیری های ECG. میله های خطا، فواصل اطمینان ۹۵٪ را در اطراف میانگین نشان می دهند. DL با ردیاهای ECG، سن و جنسیت بهترین مدل در این مقایسه بود، با AUROC هایی که به طور مداوم بالاتر از ۰.۸۲ بودند. مساحت زیر منحنی مشخصه عملکرد گیرنده AUROC؛ یادگیری عمیق DL؛ الکتروکاردیوگرام ECG؛ XGBoost.

که می تواند برای تقویت تصمیم گیری در محل مراقبت به عنوان بخشی از یک سیستم مراقبت های بهداشتی آموزشی مورد استفاده قرار گیرد.

تا آنجا که ما می دانیم، تنها یک مطالعه دیگر، سودمندی پیش بینی کننده مدل های یادگیری ماشین مبتنی بر ECG را در سطح جمعیت بررسی کرده است. مطالعه توسط راگونات و همکاران<sup>۱۲</sup> این مطالعه بر اساس تقریباً ۱.۲ میلیون نوار قلب از کمی بیش از ۲۵۰,۰۰۰ بیمار بود که طی یک دوره ۳۴ ساله از یک سیستم مراقبت های بهداشتی بزرگ (Geisinger) در ایالات متحده جمع آوری شده بود. مدل آنها برای مرگ و میر ۱ ساله بر اساس ردیابی های ECG، سن و جنس، در یک گروه آزمایشی ۱۶۸,۹۱۴ نفری، AUROC برابر با ۰.۸۷۶ را به دست آورد، در حالی که در گروه اعتبارسنجی ما که شامل ۸۹۳۷۹ بیمار بود، AUROC برابر با ۰.۸۲۶ بود. تفاوت های متعددی در مطالعه کانادایی ما و مطالعه راگونات و همکارانش از ایالات متحده وجود دارد. گروه مامسن تر بود (میانگین سنی  $62.6 \pm 18.6$  در مقابل  $58 \pm 18$  سال) و میزان مرگ و میر ۱ ساله آنها در مقایسه با گروه آمریکایی به طور قابل توجهی بالاتر بود (۱۷.۴٪ در مقابل ۸.۴٪ میزان مرگ و میر ۱ ساله در مجموعه های مورد بررسی). میزان بالاتر مرگ و میر در کانادا با تفاوت های بین کشوری گزارش شده قبلی در جمعیت های خاص بیماران مطابقت دارد و به تفاوت در نحوه مدیریت بیماران در دو سیستم مراقبت های بهداشتی نسبت داده شده است.<sup>۱۵-۱۷</sup> راگونات و همکارانش AUROC بالاتری را مرتبط با مدل XGB خود بر اساس سن و جنس به تنهایی گزارش کردند (۰.۷۷۴)، در حالی که مدل ماکه مرگ و میر ۱ ساله را بر اساس سن و جنس بیمار پیش بینی می کرد، AUROC ۰.۷۱۶ داشت. مطالعه ما از مدل های استاندارد (ResNet DL) استفاده کرد و مبتنی بر ECG های یک تولیدکننده تجهیزات واحد (فیلیپس) بود. در حالی که مطالعه ایالات متحده از معماری DL طراحی شده سفارشی استفاده کرد و مبتنی بر ECG های تولیدکنندگان تجهیزات مختلف بود. پیاده سازی معماری DL آنها بر روی داده های ما منجر به عملکرد کوچک اما به طور قابل توجهی پایین تر در مقایسه با مدل های استاندارد DL ما شد، که نشان می دهد تغییر دامنه در سناریوهای آموزش و اعتبارسنجی ممکن است نقشی داشته باشد. با وجود این تفاوت های قابل توجه در بیماران، سیستم های بهداشتی، تجهیزات و ساختار مدل، هر دو مطالعه میزان مشابهی از بهبود در عملکرد مرتبط با افزودن ردهای ECG را یافتند. این یافته ها ارزش و قابلیت تعمیم بالقوه مدل های DL مبتنی بر ECG را برای پیش بینی مرگ و میر برجسته می کند.

مطالعه ما کار راگونات و همکاران را ادامه می دهد.<sup>۱۲</sup> با توسعه مدل هایی برای پیامدهای مرگ و میر کوتاه مدت (30 روزه) و بلندمدت (5 ساله)؛ و بررسی عملکرد مدل ها در مردان و زنان به طور جداگانه. ما دریافتیم که مدل های DL ما در هر دو زمان اضافی به طور مداوم عملکرد خوبی داشتند.

۱.۲۷٪ (۸۶.۴۶-۸۵.۱۹٪) برای مدل های ۳۰ روزه، ۱.۲۵٪ (۸۲.۵۸-۸۳.۸۳٪) برای مدل های ۵ ساله و ۲.۸۹٪ (۸۲.۸-۸۵.۶۹٪) برای اعتبارسنجی H1؛ اما در اعتبارسنجی H2، برای مدل های ۳۰ روزه ۳.۶٪ (۸۱.۵۹-۸۵.۱۹٪)، برای مدل های ۵ ساله ۲.۴۱٪ (۸۰.۱۷-۸۲.۵۸٪) و برای مدل های ۵ ساله ۱.۸۴٪ (۸۰.۹۶-۸۲.۸۸٪) کمتر بود.

چهارم، ما عملکرد مدل های خود را در کل مجموعه ی «بازدارنده» (holdout)، شامل چندین نوار قلب (ECG) برای بیماران «بازدارنده» بررسی کردیم. در مجموع، 78250 (80.55٪)، 71636 (80.15٪) و 43629 (78.40٪) بیمار بیش از یک نوار قلب (ECG) برای پیش بینی های 30 روزه، 1 ساله و 5 ساله در مجموعه ی «بازدارنده» ما در دسترس داشتند. عملکرد مدل های ما پس از گنجاندن چندین نوار قلب (ECG) با عملکرد آنها در تجزیه و تحلیل اصلی بر اساس یک نوار قلب تصادفی به ازای هر بیمار قابل مقایسه بود (جدول تکمیلی 4 و شکل تکمیلی 8). علاوه بر این، ما در پیش بینی در سراسر ECG برای بیمارانی که چندین ECG داشتند، سازگاری یاقیم 85.05٪، 82.27٪ و 82.25٪ بیماران حداقل 50٪ پیش بینی های دقیق و ثابت در سراسر ECG های متعدد خود در 30 روزه، 1 ساله و 5 ساله داشتند، شکل تکمیلی 9).

## بحث

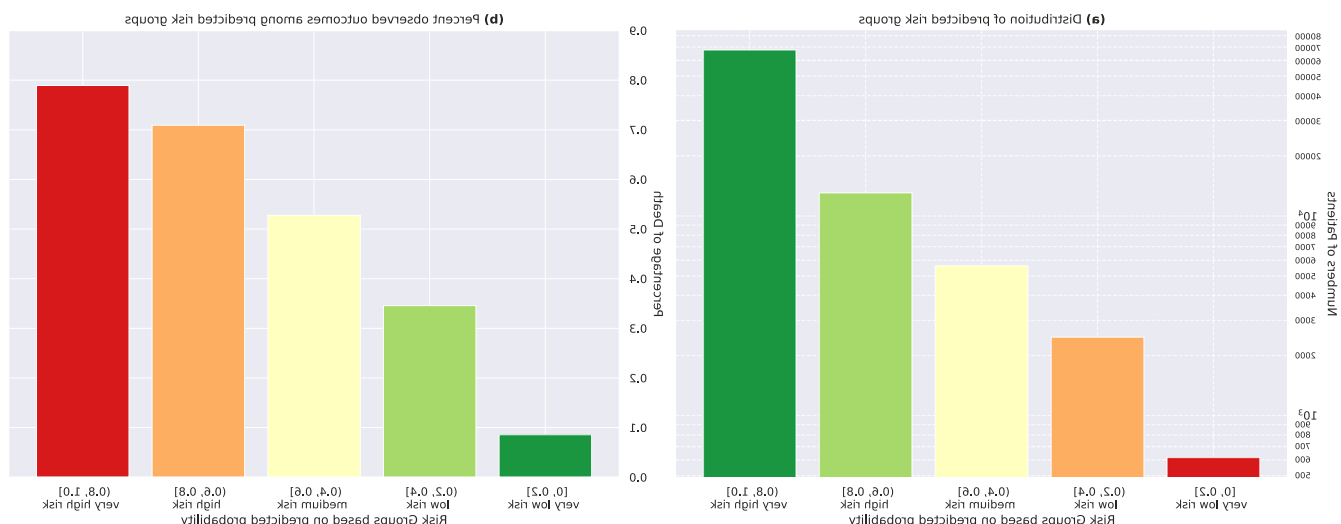
مطالعه ما، که بر اساس یک گروه بزرگ جمعیتی از بیماران با دسترسی جهانی به مراقبت های بهداشتی انجام شده است، کاربرد مدل های یادگیری ماشینی مبتنی بر داده های ECG را برای شناسایی بیماران در معرض خطر بالای مرگ و میر کوتاه مدت و بلند مدت در زمان مراجعه به اورژانس یا بیمارستان نشان می دهد. ما دریافتیم که مدل های (ResNet DL) مبتنی بر ردیابی های ECG 12 لید، در پیش بینی مرگ و میر نسبت به مدل های تقویت گرایان (XGBoost) مبتنی بر اندازه گیری های ECG گزارش شده به طور معمول، عملکرد بهتری دارند. در یک گروه اعتبارسنجی شامل تقریباً 100000 بیمار، ردیابی های ECG بیشترین اطلاعات پیش آگهی را ارائه دادند، و افزودن سن و جنس بیمار، بهبودهای افزایشی کوچکی را در عملکرد مدل ارائه داد. تجزیه و تحلیل های تکمیلی، استحکام عملکرد مدل های ما را در ECG های با کیفیت پایین، در بیمارستان های مختلف و هنگامی که چندین ECG برای هر بیمار گنجانده شده بود، نشان داد. مطالعه ما اولین مطالعه ای است که ارزش افزوده ترکیب داده های آزمایشگاهی را بررسی می کند و دریافتیم که مدل های مبتنی بر داده های ترکیبی نوار قلب، آزمایشگاه داده های جمعیت شناختی (سن و جنس بیماران) بهترین عملکرد را در پیش بینی مرگ و میر کوتاه مدت و بلندمدت دارند. این یافته ها نشان می دهد که چگونه می توان از مدل های یادگیری ماشین برای تبدیل داده های جمع آوری شده معمول در عمل بالینی به دانش استفاده کرد.



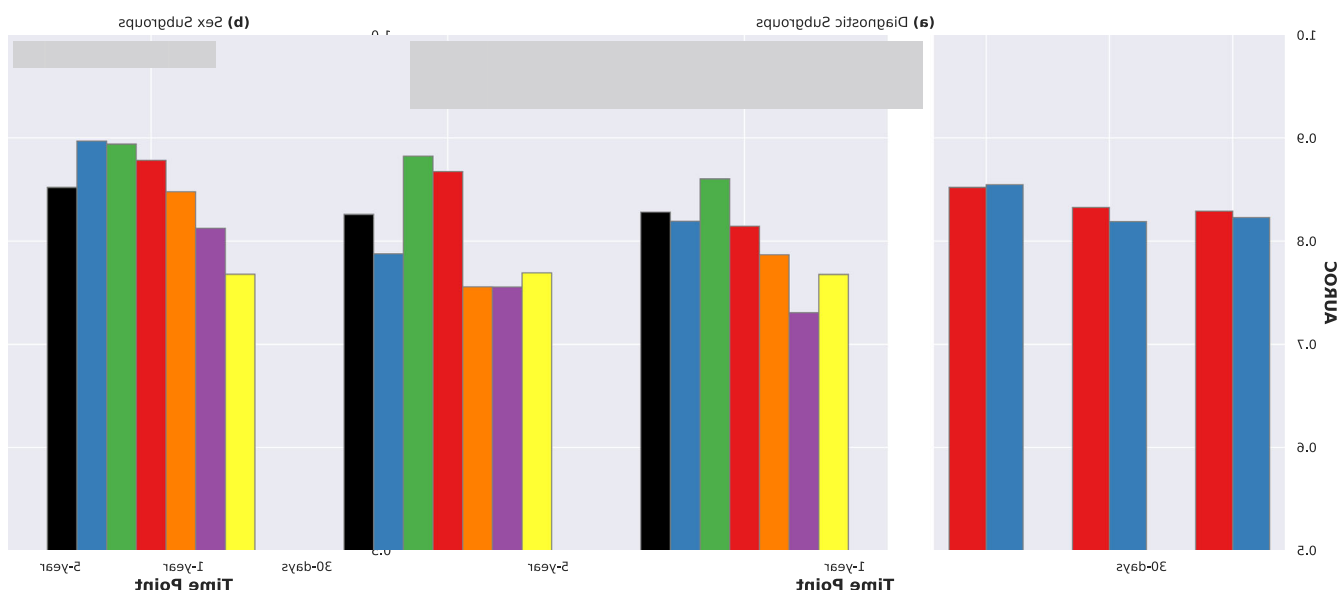
جدول ۳. ارزیابی عملکردهای مختلف مدل که به صورت میانگین (فاصله اطمینان ۹۵٪) درصد بیان شده اند.

| امتیاز زیر       | دقت                    | دقت                    | میان یاروند            | ویژگی                  | F1 امتیاز              | AUPRC                  | AUROC                  | مدل    | سن، جنسیت           | روز ۳۰ |
|------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|--------|---------------------|--------|
| (۱.۶۸-۱.۴۱)۱.۵۵  | ۶۰.۵۹<br>(۶۱.۲۱-۵۹.۹۳) | ۲۰.۳۷-۱.۷۳)۲۰.۶        | ۶۶.۳۸<br>(۷۲.۳۲-۶۰.۳۸) | ۶۰.۵۲<br>(۶۱.۱۱-۵۹.۸۶) | ۳۰.۹۹<br>(۴.۵۹-۳۰.۳۷)  | ۲۰.۹۵<br>(۴.۰۶-۲۰.۱۶)  | ۶۷.۹۹<br>(۷۱.۴۷-۶۴.۵۴) | لگسیمی | سن، جنسیت           | روز ۳۰ |
| (۶.۱۲-۵.۸۳)۵.۹۷  | ۷۹.۶۳<br>(۷۹.۹۲-۷۹.۳۳) | ۲۲.۸۵<br>(۲۳.۲۶-۲۲.۱۷) | ۷۱.۹<br>(۷۳.۰۸-۷۰.۷۶)  | ۸۰.۲۵<br>(۸۰.۵۶-۷۹.۹۶) | ۳۴.۶۷<br>(۳۵.۲۶-۳۳.۸۳) | ۳۴.۵۴<br>(۳۵.۸۱-۳۳.۲۴) | ۸۴.۳۲<br>(۸۴.۸۲-۸۳.۸۳) | مقال   | بوقلم ECG           |        |
| (۶.۶۱-۶.۳۹)۶.۴۴  | ۸۲.۰۵<br>(۸۲.۳۳-۸۱.۷۴) | ۲۱.۲۷<br>(۲۲.۰۱-۲۰.۵۳) | ۵۱.۳۸<br>(۵۲.۸۷-۴۹.۸۱) | ۸۴.۵۴<br>(۸۴.۸۲-۸۴.۲۱) | ۳۰.۰۹<br>(۳۱.۰۱-۲۹.۱۴) | ۲۲.۰۷<br>(۲۲.۹۶-۲۱.۱۴) | ۷۸.۲۴<br>(۷۸.۸۷-۷۷.۵۹) | لگسیمی |                     |        |
| (۵.۸۹-۵.۶۳)۵.۷۵  | ۷۱.۸۲<br>(۷۳.۰۷-۷۰.۶۷) | ۲۴.۲۴<br>(۲۴.۹۳-۲۳.۵۶) | ۷۱.۸۲<br>(۷۳.۰۷-۷۰.۶۷) | ۸۱.۷۵<br>(۸۲.۰۵-۸۱.۴۵) | ۳۶.۲۵<br>(۳۷.۰۹-۳۵.۳۷) | ۳۵.۰۶<br>(۳۶.۸۴-۳۴.۳۱) | ۸۵.۱۹<br>(۸۵.۶۸-۸۴.۷۱) | مقال   | نوار قلب، سن، جنسیت |        |
| (۶.۴۴-۶.۲۶)۶.۳۵  | ۸۲.۵۷<br>(۸۲.۸۳-۸۲.۲۹) | ۲۲.۱۲<br>(۲۳.۸۳-۲۲.۱۹) | ۵۶.۳۳<br>(۵۷.۸۱-۵۴.۸۱) | ۸۴.۷۱<br>(۸۴.۹۸-۸۴.۴۳) | ۳۳.۶۶<br>(۳۳.۶۷-۳۱.۶۷) | ۲۵.۷۴<br>(۲۶.۷۵-۲۴.۵۷) | ۸۱.۲۴<br>(۸۱.۶۱-۸۰.۴۵) | لگسیمی |                     | اسال   |
| (۹.۳۶-۸.۹۱)۹.۲۱  | ۶۴.۸۹<br>(۶۵.۵۵-۶۴.۲۸) | ۱۷.۸<br>(۱۸.۶۹-۱۶.۸۹)  | ۶۷.۲۳<br>(۶۹.۳۶-۶۴.۹۸) | ۶۴.۶۲<br>(۶۵.۳۳-۶۳.۹۷) | ۲۸.۱۴<br>(۲۹.۳۸-۲۶.۸۹) | ۲۲.۵۹<br>(۲۴.۱۹-۲۰.۹۳) | ۷۱.۵۹<br>(۷۲.۸۹-۷۰.۳۹) | لگسیمی | سن، جنسیت           |        |
| (۱۱.۶-۱۱.۳)۱۱.۴۴ | ۷۶.۸۸<br>(۷۷.۱۷-۷۶.۶۱) | ۴۰.۱۲<br>(۴۰.۸۲-۳۹.۴۵) | ۶۷.۴۶<br>(۶۸.۳۲-۶۶.۵۶) | ۷۸.۸۶<br>(۷۹.۱۸-۷۸.۵۴) | ۵۰.۳۱<br>(۵۱.۰۱-۴۹.۶۲) | ۴۸.۱۷<br>(۴۹.۱۱-۴۷.۱۱) | ۸۱.۲<br>(۸۱.۶۱-۸۰.۷۷)  | مقال   | بوقلم ECG           |        |
| (۱۲.۰۳-۱۲.۶۹)    | ۷۵.۶۳<br>(۷۵.۹۹-۷۵.۲۸) | ۳۵.۷۲<br>(۳۶.۴۲-۳۴.۹۸) | ۵۰.۶۲<br>(۵۱.۵۵-۴۹.۶۸) | ۸۰.۸۸<br>(۸۱.۲۲-۸۰.۵۴) | ۴۱.۸۸<br>(۴۲.۶۱-۴۱.۱۶) | ۳۵.۹۲<br>(۳۶.۸۳-۳۵.۰۲) | ۷۴.۰۲<br>(۷۴.۵۳-۷۳.۴۸) | لگسیمی |                     |        |
| (۱۱.۳۷-۱۱.۰۵)    | ۷۸.۲۶<br>(۷۸.۵۶-۷۷.۹۷) | ۴۲.۱۴<br>(۴۲.۸۳-۴۱.۴۳) | ۶۸.۰<br>(۶۸.۹۴-۶۶.۹۸)  | ۸۰.۴۱<br>(۸۰.۷۱-۸۰.۱۰) | ۵۲.۰۴<br>(۵۲.۷۵-۵۱.۲۹) | ۵۱.۲۱<br>(۵۲.۲۲-۵۰.۳۳) | ۸۲.۵۸<br>(۸۲.۹۷-۸۲.۱۸) | مقال   | نوار قلب، سن، جنسیت |        |
| (۱۲.۳۶-۱۲.۰۳)    | ۷۷.۵۲<br>(۷۷.۸۱-۷۷.۲۱) | ۳۹.۶۱<br>(۴۰.۲۹-۳۸.۹۳) | ۵۶.۳۹<br>(۵۷.۲۸-۵۵.۵۶) | ۸۱.۹۶<br>(۸۲.۲۴-۸۱.۶۶) | ۴۶.۵۳<br>(۴۷.۲۳-۴۵.۸۸) | ۴۲.۰۳<br>(۴۳.۰۵-۴۱.۰۳) | ۷۸.۳۹<br>(۷۸.۷۹-۷۷.۹۸) | لگسیمی |                     |        |
| (۱۸.۸۶-۱۸.۱۳)    | ۷۱.۱۶<br>(۷۲.۵۹-۷۰.۶۱) | ۴۸.۹۳<br>(۵۰.۸۹-۴۷.۱۹) | ۶۹.۱۹<br>(۷۰.۹۹-۶۷.۳۴) | ۷۲.۵۲<br>(۷۳.۶-۷۱.۳۵)  | ۵۷.۳۲<br>(۵۸.۹۸-۵۵.۷۸) | ۵۹.۰۳<br>(۶۱.۶۲-۵۶.۷۳) | ۷۷.۶۳<br>(۷۸.۷-۷۶.۵۳)  | لگسیمی | سن، جنسیت           | ۵ سال  |
| (۱۷.۸۳-۱۷.۲۹)    | ۷۴.۳<br>(۷۴.۷۳-۷۳.۸۳)  | ۶۰.۳<br>(۶۱.۱۷-۵۹.۶۸)  | ۶۶.۱۵<br>(۶۶.۸۸-۶۵.۲۶) | ۷۸.۳۷<br>(۷۸.۸۴-۷۷.۸۹) | ۶۳.۱۴<br>(۶۳.۷۸-۶۲.۴۵) | ۶۵.۸۸<br>(۶۶.۷۴-۶۴.۸۹) | ۷۹.۸۲<br>(۸۰.۳۹-۷۹.۱۶) | مقال   | بوقلم ECG           |        |
| (۱۹.۰۵-۱۸.۸۵)    | ۷۱.۱۴<br>(۷۱.۸۸-۷۰.۹۱) | ۵۷.۷۸<br>(۵۸.۵۷-۵۶.۸۳) | ۵۲.۰۷<br>(۵۳.۰۱-۵۱.۲۳) | ۸۱.۰۴<br>(۸۱.۵۰-۸۰.۵۶) | ۵۴.۷۸<br>(۵۵.۵۶-۵۳.۹۹) | ۵۶.۲۱<br>(۵۷.۰۸-۵۵.۱۵) | ۷۴.۶۱<br>(۷۵.۱-۷۴.۰۲)  | لگسیمی |                     |        |
| (۱۶.۵۷-۱۶.۳۳)    | ۷۶.۶۷<br>(۷۷.۱-۷۶.۲۳)  | ۶۳.۶۶<br>(۶۴.۴۸-۶۲.۸۶) | ۶۹.۶۱<br>(۷۰.۳۶-۶۸.۸۵) | ۸۰.۱۸<br>(۸۰.۶۳-۷۹.۶۹) | ۶۶.۵<br>(۶۷.۱۸-۶۵.۸۳)  | ۷۰.۰۲<br>(۷۰.۹۱-۶۹.۱۷) | ۸۲.۸<br>(۸۳.۲۴-۸۲.۳۵)  | مقال   | نوار قلب، سن، جنسیت |        |
| (۱۶.۷۲-۱۶.۳)۱۶.۵ | ۷۵.۸۴<br>(۷۶.۲۱-۷۵.۴۱) | ۶۳.۵۶<br>(۶۴.۲۷-۶۲.۸۱) | ۶۴.۱۳<br>(۶۴.۹۱-۶۳.۳۱) | ۸۱.۶۸<br>(۸۲.۱۲-۸۱.۲۴) | ۶۳.۸۴<br>(۶۴.۴۹-۶۳.۲۲) | ۶۷.۲۶<br>(۶۸.۰۵-۶۶.۳۷) | ۸۱.۱۸<br>(۸۱.۵۸-۸۰.۷۳) | لگسیمی |                     |        |

AUPRC مساحت زیر منحنی دقت-فراخوانی، AUROC مساحت زیر منحنی عملکرد گیرنده، دی ال یادگیری عمیق، نوار قلب (ECG) الکتروکاردیوگرام، ایکس جی بی ایکس جی بوست



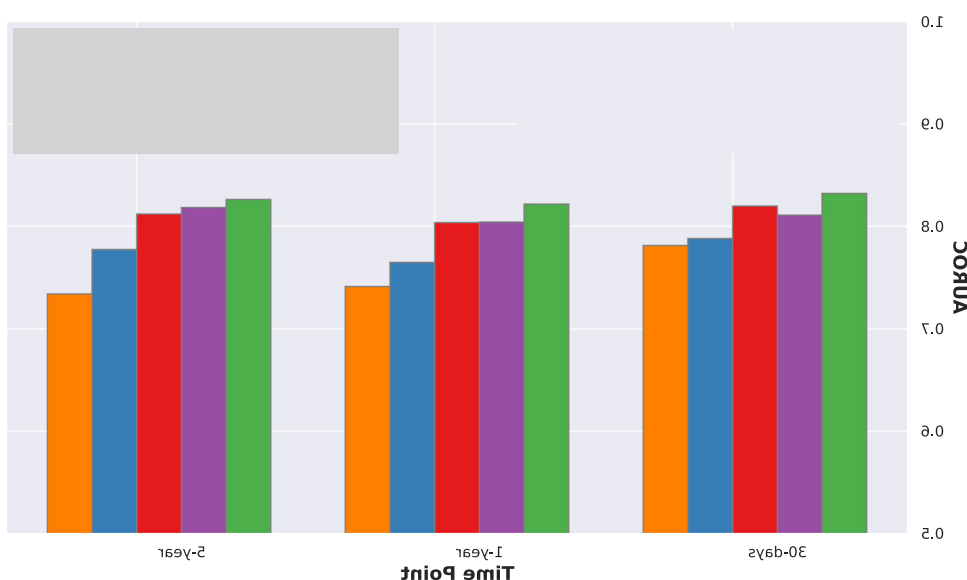
شکل ۳: گروه های خطر پیش بینی شده برای مرگ و میر ۱ ساله. الف: پنج گروه خطر - گروه های خطر «بسیار کم»، «کم»، «متوسط»، «زیاد» و «بسیار زیاد» را بر اساس 20 درصد نقاط برش (0 تا 20 درصد، 20 تا 40 درصد و غیره) از احتمال پیش بینی شده مرگ از مدل های اصلی خود (DL: نوار قلب، سن، جنس) در مجموعه افراد تحت بررسی، استخراج کردیم. ب: درصد مرگ های مشاهده شده در هر گروه خطر پیش بینی شده، کالیبراسیون خوبی را با افزایش مداوم در گروه های خطر نشان داد (به ترتیب ۷۸.۹٪، ۷۰.۹٪، ۵۲.۳٪، ۳۴.۶٪، ۱۸.۶٪). مرگ در گروه های خطر «بسیار کم»، «کم»، «متوسط»، «زیاد» و «بسیار زیاد». یادگیری عمیق DL، نوار قلب.



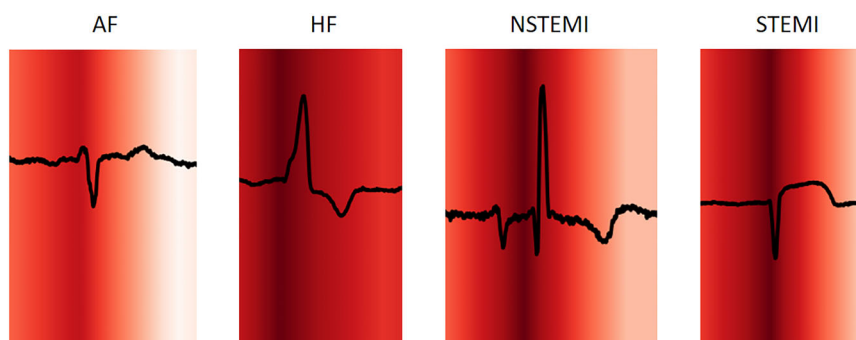
شکل ۴: عملکرد مدل در زیرجمعیت های تشخیصی و مبتنی بر جنسیت. الف: عملکرد مدل های DL: ردیابی های ECG، سن، جنسیت در زیرگروه های مختلف تشخیص اولیه. این مدل ها در بیماران مبتلا به STEMI و NSTEMI (AUROC 0.867 و 0.882 برای مرگ و میر 1 ساله، به ترتیب) نسبت به گروه کلی بهتر عمل کردند. عملکرد مدل در زیرگروه های دیگر (نارسایی قلبی، دیابت و فیبریلاسیون دهلیزی) کمتر از گروه کلی افراد تحت مطالعه بود. ب: مدل های پیش آگهی در مردان کمی بهتر از زنان عمل کردند. AUROC سطح زیر منحنی مشخصه عملکرد گیرنده، یادگیری عمیق DL، الکتروکاردیوگرام ECG، انفارکتوس میوکارد بدون بالا رفتن قطعه ST در مرحله NSTEMI، انفارکتوس میوکارد با بالا رفتن قطعه ST در مرحله STEMI.

مما معتقدیم مطالعه ما اولین مطالعه ای است که ارزش پیش آگهی افزایشی حاصل از گنجاندن داده های مربوط به آزمایش های آزمایشگاهی منتخب را نشان می دهد. ما مدل های خود را به صورت متوالی ساختیم، که فقط با سن و جنس شروع شد و سپس ردیابها یا اندازه گیری های ECG و متعاقباً داده های آزمایشگاهی را اضافه کردیم. AUROC برای مدل مرگ و میر ۱ ساله از ۰.۸۱ برای مدل مبتنی بر سن، جنس و ردیابهای ECG به ۰.۸۳ برای مدل مبتنی بر سن، جنس، ردیابهای ECG و آزمایش افزایش یافت. داده های آزمایشگاهی ممکن است اطلاعات پیش آگهی بیشتری را در جمعیت های خاص بیمار (مثلاً بیماران مبتلا به سندرم های حاد کرونری یا بیماری کلیوی و غیره) ارائه دهند و افزودن سایر معیارهای آزمایشگاهی مانند ALT، AST، و HbA1c ممکن است عملکرد مدل های ما را بهبود بخشد. این بررسی ها به عنوان بخشی از مطالعات تحقیقاتی آینده برنامه ریزی شده اند.

امتیاز (AUROC) برابر با 0.85 در 30 روز و 0.83 در 5 سال) و به طور مشابه در هر دو جنس. با این حال، مشاهده کردیم که عملکرد مدل پایه سن + جنس برای پیش بینی های بلندمدت، بالاتر و به مدل های DL فقط با ECG نزدیک ترمی شود. این نشان می دهد که اگرچه ممکن است مزایای واضحی در رابطه با کاربرد پیش بینی ECG در هدایت کوتاه مدت تا میان مدت درمان ها وجود داشته باشد، استفاده از ECG به تنهایی بدون ویژگی های سن و جنس ممکن است ارزش پیش بینی کننده کافی برای نتایج مرگ و میر 5 ساله نداشته باشد. همچنین، ما عملکرد مدل متفاوتی را در زیرگروه های تشخیصی مشاهده کردیم، به طوری که مدل ها برای بیماران مبتلا به انفارکتوس میوکارد نسبت به سایر گروه های بیماری عملکرد بهتری داشتند.



شکل ۵ مقایسه عملکرد مدل با و بدون ویژگی‌های آزمایشگاهی، میله‌های خطا، فواصل اطمینان ۹۵٪ را در اطراف میانگین نشان می‌دهند. مدل‌های ECG در تمام نقاط زمانی، حتی بدون مقادیر آزمایشگاهی، AUROC بالاتری نسبت به مدل پایه (XGB: سن، جنس، آزمایشگاه) داشتند، اما این تفاوت برای پیش‌بینی‌های با بازه زمانی طولانی‌تر کمتر بود. افزودن ویژگی‌های آزمایشگاهی، عملکرد مدل را در کل، در مدل‌ها و نقاط زمانی مختلف به طور قابل توجهی بهبود بخشید، با این حال، افزایش عملکرد از نظر بزرگی اندک بود (به طور متوسط ۰.۹۹٪، آزمون DeLong، همه  $p > 0.001$ ). به طور کلی، مدل DL با ردیابی‌های ECG، سن، جنس و آزمایشگاه، بهترین مدل از نظر عملکرد در مقایسه بود. AUROC سطح زیر منحنی مشخصه عملکرد گیرنده، یادگیری عمیق DL، الکتروکاردیوگرام XGBoost، ECG.



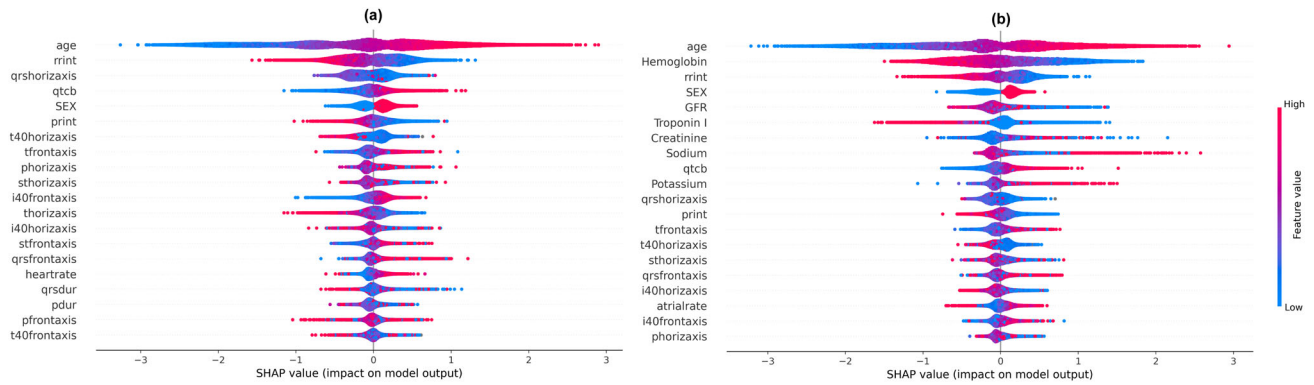
شکل ۶ نقشه حرارتی GradCAM برای مرگ و میر ۱ ساله، ردیابی‌های ECG نماینده با تشخیص‌های اولیه AF، HF، STEMI و NSTEMI انتخاب شدند. نواحی تیره‌تر در هر ردیاب GradCAM نشان دهنده نواحی با بیشترین سهم در پیش‌بینی مرگ و میر ۱ ساله مدل ResNet DL: ECG هستند. مناطق برجسته شده مختص به سرب نیستند و بر اساس داده‌های هر ۱۲ لید محاسبه شده‌اند. فیبریلاسیون دهلیزی AF، یادگیری عمیق DL، الکتروکاردیوگرام ECG، نارسایی قلبی STEMI، HF، انفارکتوس میوکارد با بالا رفتن قطعه NSTEMI، ST، انفارکتوس میوکارد بدون بالا رفتن قطعه ST.

اینکه یافته‌های ما تا چه حد قابل تعمیم به ECG‌های سایر تولیدکنندگان تجهیزات است، باید مشخص شود. دوم، داده‌های آزمایشگاهی فقط از سال ۲۰۱۲ به بعد در دسترس بودند و نه برای همه بیماران. میزان مرگ و میر یک ساله در بین بیمارانی که داده‌های آزمایشگاهی داشتند (۱۹.۴٪) کمی بیشتر از بیمارانی بود که داده‌های آزمایشگاهی نداشتند (۱۷.۳٪). در حالی که افزودن ویژگی‌های آزمایشگاهی منجر به بهبودهای کوچک اما قابل توجهی در عملکرد مدل شد، ارزیابی اینکه آیا این موضوع با میزان بالاتر پیامدهای نامطلوب در این زیرگروه مرتبط است یا خیر، دشوار است. سوم، همانطور که در بالا ذکر شد، فهرست آزمایش‌های آزمایشگاهی موجود در مدل‌های ما جامع نیست. چهارم، ECG تصادفی ما به ازای هر گروه بیمار، در مقایسه با کل مجموعه بیماران تحت درمان، کمی جوان‌تر بود و بیماران مرد کمتری داشت و بیماری‌های همراه کمتری نیز داشت که به ECG‌های بیشتری که در بیماران مسن‌تر، مرد و دارای عوارض بالینی با بیماری‌های همراه بیشتر انجام می‌شد، نسبت داده می‌شود. با این حال، ما در پیش‌بینی بین ECG‌ها در بین بیمارانی که چندین ECG در مجموعه بیماران تحت درمان داشتند، سازگاری قابل توجهی یافتیم. و در آخر، اندازه‌گیری‌های ECG مورد استفاده در مدل‌های XGB از طریق دستگاه‌های فیلیپس ارائه شده‌اند و توسط آزمایشگاه اصلی یا متخصصان انسانی بررسی نشده‌اند.

مدل‌های DL با شبکه‌های عصبی کانولوشنی، در شناسایی و تفسیر الگوهای مورد استفاده مدل برای پیش‌بینی، به عنوان جعبه‌های سیاه در نظر گرفته می‌شوند. ما تکنیک‌هایی مانند ایجاد نقشه‌های حرارتی GradCAM را برای این منظور امتحان کرده ایم که نشان می‌دهد فواصل PR، کمپلکس‌های QRS و تغییرات ST-T، به ویژه بخش اولیه کمپلکس QRS، بیشترین سهم را در پیش‌بینی مرگ و میر دارند. این امر عمدتاً در شرایط مختلف بیماری ثابت بود، با این حال، تقریباً تمام بخش‌های ECG در پیش‌بینی مدل DL در بیماران مبتلا به نارسایی قلبی نقش داشتند. لازم به ذکر است که این تکنیک‌های تجسم، حوزه‌ای از تحقیقات فعال هستند و استخراج تفاسیر بالینی معنادار از سیگنال‌های ECG با ضربان‌های قلبی متعدد، چالش برانگیز است. به عنوان مکمل، ما از تجزیه و تحلیل SHAP برای مدل‌های XGB استفاده کردیم که چند پارامتر ECG (مانند فاصله RR پایین‌تر، محور QRS افقی پایین‌تر، فاصله QT بالاتر و فاصله PR پایین‌تر) را که بیشترین سهم را در پیش‌بینی مرگ و میر در دوره‌های مختلف پیگیری داشتند، برجسته کرد.

مطالعه‌ها محدودیت‌هایی دارد. اولاً تمام ECG‌های مورد مطالعه ما از یک تولیدکننده (سیستم اطلاعاتی فیلیپس) بودند.





شکل ۱ اهمیت ویژگی مبتنی بر SHAP برای مرگ و میر ۱ ساله. الف تحلیل SHAP مدل های XGBoost با اندازه گیری های ECG، سن و جنس نشان داد که سن بالاتر، فاصله RR پایین تر، محور QRS افقی (اثر شرطی)، فاصله QT اصلاح شده با نرخ Bazett بالاتر، جنس مذکر و فاصله PR پایین تر بیشترین سهم را در مرگ و میر ۱ ساله داشته اند. ب افزودن ویژگی های آزمایشگاهی، سهم هموگلوبین پایین تر، میزان فیلتراسیون گلوبمرولی (GFR) پایین تر، تروپونین I پایین تر، کراتینین بالاتر، سدیم بسیار بالا یا پایین و پتاسیم بالا را برجسته کرد. شرح اندازه گیری های ECG در جدول تکمیلی ۵ ارائه شده است.

شروع موج، فاصله QT تصحیح شده با سرعت Fridericia، ضربان قلب، فاصله PR، مدت زمان QRS، فاصله QT، فاصله QT تصحیح شده با سرعت Bazett، محور P پیشانی، محور QRS پیشانی در ۴۰ میلی ثانیه اولیه، محور QRS پیشانی در ۴۰ میلی ثانیه پایانی، محور QRS پیشانی، محور موج ST پیشانی (معادل انحراف ST)، محور T پیشانی، محور P افقی، محور QRS افقی در ۴۰ میلی ثانیه اولیه، محور QRS افقی در ۴۰ میلی ثانیه پایانی، محور QRS افقی، محور موج ST افقی و محور T افقی.

#### داده های آزمایشگاهی

داده های آزمایشگاهی متمرکز در سطح جمعیت از سال ۲۰۱۲ به بعد در دسترس هستند. داده های مربوط به مجموعه ای منتخب از آزمایش ها شامل کراتینین، میزان فیلتراسیون گلوبمرولی (GFR) که از کراتینین، هموگلوبین، پتاسیم، سدیم و تروپونین I محاسبه شده است، در صورت وقوع در همان روز، با داده های ECG مرتبط شدند. آزمایش ها بر اساس ارتباط آنها با پیامدهای نامطلوب، استفاده معمول در عمل و در صورت در دسترس بودن برای بخش قابل توجهی از بیماران انتخاب شدند.

#### گروه تحلیل

گروه مورد مطالعه شامل بیمارانی بود که بین فوریه ۲۰۰۷ و آوریل ۲۰۲۰ در آلبرتا، کانادا، در ۱۴ مرکز بستری شده بودند. داده های ما شامل ۲۰۱۵،۸۰۸ نوار قلب، ۳،۳۳۶،۹۱۰ مراجعه به اورژانس، ۱۰۷،۵۷۶ بستری در بیمارستان و ۱۲۶،۰۶۵ بیمار بود. مواجهه های همزمان با مراقبت های بهداشتی برای یک بیمار (مراجعه به اورژانس و/یا بستری شدن در بیمارستان) که در یک بازه زمانی کوتاه رخ داده بود، به عنوان انتقال (به عنوان مثال، از بخش اورژانس به پذیرش بیمارستان یا از بیمارستان عمومی به بیمارستان سطح سوم) در نظر گرفته شد و به اپیزودها گروه بندی شدند. نمودار جریان درخت تصمیم گیری مورد استفاده برای تعریف اپیزود در شکل تکمیلی ۱۰ نشان داده شده است.

اگر تاریخ ثبت نوار قلب (ECG) در بازه زمانی بین تاریخ پذیرش و تاریخ ترخیص یک دوره بود، یک رکورد ECG به یک دوره مراقبت های بهداشتی مرتبط می شد (شکل تکمیلی ۱۱). ECG های با کیفیت پایین از طریق پرچم های هشدار دهنده تولید شده توسط الگوریتم کیفیت داخلی سازنده دستگاه ECG برای وجود آرتیفکت عضلانی، نویز AC، انحراف خط پایه، برش QRS و لیدهای جدا شده شناسایی شدند. پس از حذف ECG هایی که نمی توانستند هیچ دوره ای مرتبط باشند، ECG های بیماران زیر ۱۸ سال و همچنین ECG های با کیفیت سیگنال پایین، گروه تجزیه و تحلیل ما شامل ۷۴۸۷۷۳ از ۷۴۸۷۷۳ دوره در ۲۴۴۰۷۷ بیمار بود. به شکل ۱ مراجعه کنید. ابرای نمودار جریان طراحی مطالعه، که اندازه نمونه ها را برای مطالعه کلی، تقسیم بندی های تجربی و پیامدهای مختلف نشان می دهد. در تجزیه و تحلیل های تکمیلی، عملکرد مدل های خود را در بیمارستان های مختلف و روی ECG های بی کیفیت که از مطالعه اصلی حذف شده بودند، ارزیابی کردیم.

در نتیجه، مطالعه ما نشان می دهد که مدل های DL مبتنی بر ECG می توانند برای شناسایی بیمارانی که در معرض خطر بالای مرگ و میر کوتاه مدت یا بلندمدت هستند، مورد استفاده قرار گیرند. این مدل ها در مردان و زنان به طور یکسان عملکرد خوبی دارند و می توانند با گنجاندن داده های مربوط به آزمایش های آزمایشگاهی که به طور معمول انجام می شوند، تقویت شوند. مطالعات آینده برای ارزیابی سودمندی ارائه ارزیابی ریسک بر اساس داده های ECG در عمل بالینی در حال برنامه ریزی هستند.

#### روش ها

مجموعه داده ها

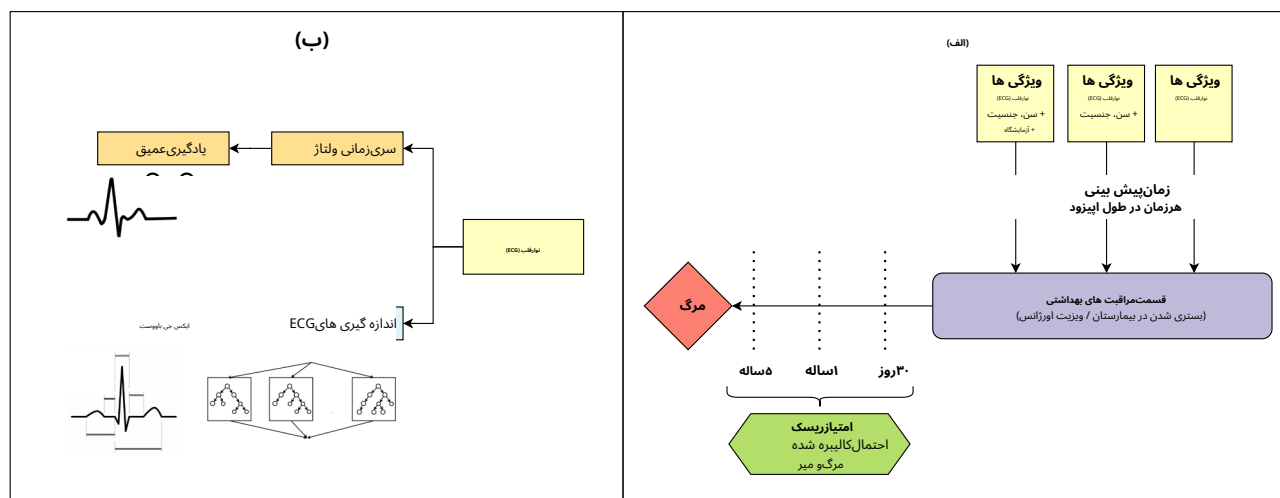
استان آلبرتا، کانادا، دارای یک سیستم مراقبت های بهداشتی تک پرداخت کننده (وزارت بهداشت: آلبرتا هلت) و تک ارائه دهنده (خدمات بهداشتی آلبرتا) است. حدود ۴.۴ میلیون نفر از ساکنان این استان به خدمات بیمارستانی، سرپایی، آزمایشگاهی و پزشکی دسترسی جهانی دارند.

برای این مطالعه، داده های نوار قلب (ECG) با استفاده از یک شماره سلامت بیمار منحصر به فرد به پایگاه های داده سلامت اداری زیر متصل شدند: (۱) پایگاه داده چکیده ترخیص (DAD) که شامل داده های مربوط به بستری ها از جمله تاریخ پذیرش، تاریخ ترخیص، تشخیص اصلی، تا ۲۴ تشخیص دیگر و وضعیت ترخیص (انتقال، ترخیص به منزل، فوت) است. (۲) پایگاه داده سیستم گزارش دهی مراقبت های سرپایی ملی (NACRS) از تمام ویزیت های کلینیک سرپایی بیمارستان (از جمله بخش اورژانس) که شامل تاریخ پذیرش، تشخیص اصلی، تا ۹ تشخیص دیگر و وضعیت ترخیص است. (۳) رجیستری طرح بیمه مراقبت های بهداشتی آلبرتا (AHCIP) که اطلاعات جمعیت شناختی (سن، جنس) و تاریخ فوت را ارائه می دهد. (۴) داده های آزمایشگاهی متمرکز و (۵) رجیستری مرگ و میر وضعیت حیاتی. در صورت وجود وضعیت یا تاریخ های متناقض مرگ و میر (۱.۱٪ از بیماران)، رجیستری وضعیت حیاتی نسبت به سوابق رجیستری NACRS، DAD، AHCIP و اولویت داده شد.

این مطالعه توسط هیئت اخلاق پژوهشی دانشگاه آلبرتا (Pro00120852) تأیید شد. هیئت اخلاق تشخیص داد که این تحقیق یک بررسی گذشته نگر از پایگاه داده است که برای آن رضایت آزمودنی برای دسترسی به اطلاعات سلامت قابل شناسایی شخصی، منطقی، امکان پذیر یا عملی نخواهد بود.

داده های نوار قلب (ECG)

این مطالعه از ردیاب های استاندارد ECG دوازده لید و اندازه گیری های ECG از سیستم ECG Philips IntelliSpace استفاده کرد. ردیاب های ECG سری های زمانی ولتاژ بودند که با فرکانس ۵۰۰ هرتز به مدت ۱۰ ثانیه برای هر یک از ۱۲ لید (۵۰۰) نمونه برداری شدند. ۱۲۰۰۰ اندازه گیری ولتاژ برای هر ECG، اندازه گیری های ECG به طور خودکار توسط الگوریتم داخلی سازنده دستگاه ECG تولید می شوند (جدول تکمیلی ۵). این اندازه گیری های اخیر شامل ضربان دهلیزی، مدت زمان P، فاصله Q، RR، می شود.



شکل ۸: شمای کلی وظایف پیش بینی، انواع ویژگی ها و انتخاب الگوریتم های یادگیری. الف: این مطالعه بر توسعه و ارزیابی مدل های مرگ و میر مبتنی بر ECG برای پیش بینی احتمال مرگ بیمار در عرض 30 روز، 1 سال و 5 سال، از روز اخذ ECG، متمرکز بود. ECG های مورد استفاده در این مدل ها می توانند در هر نقطه زمانی در طول یک دوره مراقبت های بهداشتی اخذ شده باشند. مدل ها شامل ویژگی هایی با منقش نوار قلب (ECG)، دوقنوار قلب + سن، جنس و آانوار قلب + سن، جنس + آزمایش های آزمایشگاهی، هدف مدل های پیش بینی، ارائه یک احتمال کالبره شده از مرگ و میر است که می تواند در ارزیابی ریسک بیمار مورد استفاده قرار گیرد. داده های نوار قلب بیمار معمولاً توسط مراکز درمانی به یکی از دو شکل زیر بایگانی می شوند: من به عنوان یک گزارش بالینی از اندازه گیری های خلاصه شده ECG مانند فاصله QT، مدت زمان QRS و غیره یادوم کمتر رایج، به عنوان سری زمانی ولتاژ خام ردیابی سیگنال ECG. به منظور تسهیل کاربرد گسترده تر، ما از الگوریتم های یادگیری مناسب برای فرمت های داده استفاده کردیم، یعنی یادگیری عمیق مبتنی بر ResNet برای سری زمانی ولتاژ چند کاناله غنی از اطلاعات و XGBoost مبتنی بر تقویت گرادیان برای اندازه گیری های ECG.

زیرگروه تجزیه و تحلیل آزمایشگاهی ما شامل ۶۰۱،۳۰۷ نوار قلب از ۳۳۰،۶۳۷ مورد از ۱۴۱،۰۱۷ بیمار که داده های مربوط به هر شش آزمایشگاه مورد نظر برای آنها در دسترس بود. به طور متوسط، ۴۱.۹٪ از نوار قلب ها را می توان برای هر سال مالی از ۲۰۱۲ به ۲۰۱۹ به آزمایش های آزمایشگاهی مرتبط کرد. برای نمودار جریان طراحی مطالعه برای مدلی که شامل مقادیر آزمایشگاهی است و اندازه نمونه ها را برای زیرمطالعه کلی، تقسیم بندی های تجربی و پیامدهای مختلف نشان می دهد، به شکل تکمیلی ۱ مراجعه کنید.

#### وظایف پیش بینی

این مطالعه بر توسعه و ارزیابی مدل های مرگ و میر مبتنی بر ECG برای پیش بینی احتمال مرگ بیمار در عرض 30 روز، 1 سال و 5 سال، از روز اخذ ECG، متمرکز بود. ECG های مورد استفاده در این مدل ها می توانند در هر نقطه زمانی در طول یک دوره مراقبت های بهداشتی اخذ شده باشند (شکل ۸، پانل سمت چپ). هدف مدل های پیش بینی، ارائه یک احتمال کالبره شده از مرگ و میر است که می تواند در ارزیابی ریسک بیمار مورد استفاده قرار گیرد.

داده های ECG بیمار معمولاً توسط مراکز مراقبت های بهداشتی به یکی از دو فرمت زیر بایگانی می شوند: یا (الف) به عنوان یک گزارش بالینی از اندازه گیری های خلاصه شده ECG مانند فاصله QT، مدت زمان QRS و غیره یا (ب) که کمتر رایج است، به عنوان سری زمانی ولتاژ خام از ردیابی سیگنال ECG. به منظور تسهیل کاربرد گسترده تر، ما مدل هایی را برای تطبیق با هر یک از این فرمت های ECG توسعه دادیم. ما از الگوریتم های یادگیری مناسب برای فرمت های داده، یعنی DL مبتنی بر ResNet برای سری زمانی ولتاژ غنی از اطلاعات و XGBoost مبتنی بر تقویت گرادیان (XGB) برای اندازه گیری های ECG استفاده کردیم (شکل ۱ را ببینید). پانل سمت راست «الگوریتم های یادگیری» در زیر. XGB به دلیل عملکرد پیشرفته اش با داده های جدولی ساختاریافته، زمان آموزش سریع، پشتیبانی از مقادیر گمشده و توابع قابل توضیح انتخاب شد. به همین ترتیب، ما از معماری ResNet برای یادگیری عمیق (DL) بر اساس عملکرد موفق آن در مطالعات قبلی با مجموعه داده های قابل مقایسه از ردیابی های ECG استفاده کردیم. در تحلیل تکمیلی، ما عملکرد مدل های استاندارد DL خود را در مقایسه با معماری DL طراحی شده سفارشی که توسط یک مطالعه قبلی استفاده شده بود، محک زدیم. این مطالعه از شبکه های عصبی کانولوشن عمیق (DNN) با استفاده از پنج شاخه برای تطبیق با مدت زمان های مختلف اخذ ECG در لیدهای مختلف استفاده می کرد.<sup>۱۲</sup>

به منظور بررسی ارزش پیش بینی افزایشی که داده های جمعیت شناختی (سن، جنس) و آزمایشگاهی به عملکرد مدل های آموزش دیده فقط بر اساس ECG اضافه می کنند، ما مدل هایی را در ... توسعه دادیم.

به ترتیب زیر: (الف) فقط نوار قلب، (ب) نوار قلب + سن، جنس و (ج) نوار قلب + سن، جنس + آزمایشگاه، که در آن داده های نوار قلب می توانند یا نمودارهای ولتاژ-زمان یا اندازه گیری ها باشند.

نوار قلب بیمارانی که قبل از نقاط زمانی 30 روزه، 1 ساله و 5 ساله سانسور شده بودند، به دلیل نامشخص بودن وضعیت مرگشان از تجزیه و تحلیل حذف شدند و آنهایی که پس از نقاط زمانی سانسور شده بودند، به عنوان «زنده» در نظر گرفته شدند. برای آموزش، نوار قلب بیمارانی که دچار حادثه شده بودند، صرف نظر از کامل بودن پیگیری، برای به حداکثر رساندن یادگیری حفظ شد (به عنوان مثال، برای مرگ و میر 5 ساله، نوار قلب بیمارانی که پس از 31-03-2015 وارد گروه شدند اما قبل از 31-03-2020 فوت کردند، حفظ کردیم). با این حال، برای ارزیابی، تمام نوار قلب هایی که پیگیری کامل نداشتند، صرف نظر از وضعیت مرگ یا سانسور آنها، حذف شدند (یعنی، برای مرگ و میر 5 ساله، تمام نوار قلب های پس از 31-03-2015 را در مجموعه ارزیابی حذف کردیم، زیرا آنها پنج سال پیگیری لازم را نداشتند). تعداد ECG ها، اپیزودها و بیماران مورد استفاده برای مدل سازی هر نقطه زمانی در داده های کلی و در تقسیم بندی های تجربی در شکل ۱ ارائه شده است. جدول تکمیلی ۶.

#### الگوریتم های یادگیری

ماز یک روش طبقه بندی با برچسب های دودویی، یعنی مرده یا زنده به ترتیب ظرف 30 روز، 365 روز (1 سال) یا 1825 روز (5 سال) از تاریخ اخذ ECG، برای تخمین احتمال زنده ماندن یک بیمار جدید حداقل 30 روز، 365 روز، 1825 روز پس از اخذ ECG استفاده کردیم. از آنجایی که ورودی مدل هایی که از اندازه گیری های ECG استفاده می کردند، داده های جدولی ساختاریافته بود، ما مدل های گروه درختی تقویت شده با گرادیان (XGB) را آموزش دادیم، در حالی که برای مدل هایی با ردیابی سری زمانی ولتاژ-زمان ECG از شبکه های عصبی کانولوشن عمیق استفاده کردیم.<sup>۱۲-۱۹</sup> برای هر دو مدل XGB و DL، ما از ۹۰٪ داده های توسعه برای آموزش مدل استفاده کردیم و ۱۰٪ باقی مانده را به عنوان یک مجموعه تنظیم برای ردیابی افت عملکرد و «توقف زودهنگام» فرآیند آموزش، برای کاهش احتمال بیش برازش (متفاوت از داده های ماندگار) استفاده کردیم.<sup>۲۰</sup>

مدل XGB از تابع هدف لگاریتم درستمایی استفاده کرد. پارامترهایی مانند حداکثر عمق درخت، حداقل وزن فرزند و وزن مثبت مقیاس بر اساس اعتبارسنجی متقابل داخلی جستجوی شبکه ای ۵-لایه در مجموعه های آموزشی تنظیم شدند. مدل ها حداکثر برای ۲۰۰ دوره آموزش داده شدند و اگر افت عملکرد در مجموعه تنظیم برای ۱۰ دوره متوالی کاهش نمی یافت، فرآیند آموزش متوقف می شد.

زیرگروه های جنسی و تشخیصی. ما عملکرد مدل های خود را در زیرگروه های خاص بیمار، بر اساس تشخیص های مورد نظر و جنسیت بیمار، در مجموعه بیمارانی تحت بررسی خود بررسی کردیم. زیرگروه های تشخیصی بر اساس مسئول ترین تشخیص تعیین شده در زمان ترخیص بودند و شامل موارد زیر بودند: انفارکتوس میوکارد بدون بالا رفتن قطعه (NSTEMI) ST، انفارکتوس میوکارد با بالا رفتن قطعه (STEMI) ST، نارسایی قلبی، فیبریلاسیون دهلیزی، دیابت شیرین و فشار خون بالا. کدهای ICD مورد استفاده برای شناسایی این شرایط در جدول تکمیلی 7 ارائه شده است.

مقایسه مدل ها. برای هر ارزیابی، ما از 10 تکرار انتخاب تصادفی یک ECG واحد برای هر بیمار استفاده کردیم تا ثبات در عملکرد مدل را نشان دهیم. در طول آموزش این مدل ها، ما از تمام ECG های موجود در مجموعه آموزشی به همراه پرچسب های مرگ و میر مربوطه استفاده کردیم. تقسیم بندی های آموزش و آزمایش یکسان (از جمله انتخاب های تصادفی) برای سناریوهای مختلف مدل سازی استفاده شد، به طوری که عملکرد می توانست به طور مستقیم مقایسه شود. نمرات عملکرد بین مدل ها با بوت استرپ کردن 100 نمونه با نمونه گیری با جایگزینی از هر یک از 10 تکرار انتخاب تصادفی ECG که در بالا ذکر شد، مقایسه شدند تا در مجموع 1000 تکرار بوت استرپ ایجاد شود (شکل تکمیلی 13). تفاوت در عملکرد مدل ها بر اساس همپوشانی فواصل اطمینان 95٪ از میانگین نمرات AUROC مدل های مقایسه شده ارزیابی شد. ما همچنین مقدار p از آزمون DeLong را گزارش کرده ایم. برای نشان دادن اینکه آیا AUROC های دو مدل از نظر آماری تفاوت معنی داری دارند یا خیر.

تجسم هاما از نگاشت فعال سازی کلاس با وزن گرادیان (GradCAM) برای تجسم نقشه های فعال سازی گرادیان که در پیش بینی مدل از مرگ و میر در مدل های یادگیری عمیق ما نقش داشتند، استفاده کردیم. برای دستیابی به این هدف، آخرین لایه کانولوشن که حاوی اطلاعات سطح بالای مدل عمیق CNN و ردپاهای نماینده از مجموعه ارزیابی است، انتخاب شد. همچنین، ما SHAP (توضیحات افزایشی SHapley) برای شناسایی اندازه گیری های ECG که نقش کلیدی در پیش بینی میانگین مرگ و میر در مدل های XGB داشتند، استفاده کردیم. ۲۸.

مطالعه ما بر اساس دستورالعمل های گزارش دهی شفاف یک مدل پیش بینی چند متغیره برای پیش آگهی یا تشخیص فردی مبتنی بر هوش مصنوعی (TRIPOD-AI) گزارش شده است. ۲۹.

نقش منبع تأمین مالی: این مطالعه توسط کمک هزینه عملیاتی از موسسه تحقیقات سلامت کانادا (کمک هزینه شماره PJT-178158) تأمین مالی شده است. سازمان تأمین مالی هیچ نقشی در طراحی مطالعه، تجزیه و تحلیل داده ها، تفسیر نتایج یا تهیه نسخه خطی نداشته است.

## دسترسی به داده ها

داده های زیربنایی این مقاله توسط خدمات بهداشتی آلبرتا تحت شرایط یک توافق نامه تحقیقاتی ارائه شده است. سوالات مربوط به دسترسی به داده ها را می توان مستقیماً از آنها پرسید. ما یک مجموعه داده ECG را که به صورت مصنوعی با استفاده از رمزگذاری های خودکار متغیر و صرفاً به منظور نمایش کد تولید شده است، گنجانه ایم. انتظار نمی رود که آنها سیگنال های ECG واقعی را به طور دقیق نشان دهند. مجموعه داده آزمایشی به صورت آزاد در دسترس است و می توان آن را از اینجا دانلود کرد. <https://doi.org/10.6084/m9.figshare.21612786.v1>

## در دسترس بودن کد

کدپایه برای آموزش مدل های یادگیری عمیق مورد استفاده در این مطالعه در آدرس زیر موجود است: <https://doi.org/10.6084/m9.figshare.21612786.v1>

دریافت: ۲۶ اگوست ۲۰۲۲؛ پذیرفته شده: ۲۱ ژانویه ۲۰۲۳  
Published online: 06 February 2023

## منابع

۱. مک گینیس، جی. ام، فاینبرگ، اچ. وی. و دزائو، وی. جی. پیشبرد نظام سلامت آموزشی.

مجله پزشکی انگلستان ۳۷۵(۱۰۱۳۵): ۱-۲ (۲۰۲۱).

۲. ازکوویتز، جی. ای و همکاران. آیا در زنده ماندن از سندرم حاد کرونری با ایلا به نارسایی قلبی پس از آن، شکاف جنسیتی وجود دارد؟ گردش خون ۲۲۳۱-۲۲۳۹ (۲۰۲۰).

برای مدل DL، ما یک شبکه عصبی کانولوشن (CNN) مبتنی بر معماری شبکه عصبی باقیمانده پیاده سازی کردیم. ۲۲. متشکل از یک لایه کانولوشن، ۴ بلوک باقیمانده با ۲ لایه کانولوشن در هر بلوک، و به دنبال آن یک لایه مترامک (در مجموع ۶،۴۰۰،۴۳۳ پارامتر مدل). ما از نرمال سازی دسته ای، ReLU و حذف بعد از هر لایه کانولوشن استفاده کردیم. معماری ما بر اساس یک مطالعه در مقیاس بزرگ که قبلاً منتشر شده بود، برای شناسایی ناهنجاری ها در ECG های ۱۲ لید با برخی اصلاحات برای تطبیق ورودی داده های جدولی و خروجی مرگ و میر بود (شکل تکمیلی ۱۲). هر نمونه ECG به صورت یک ماتریس عددی ۱۲ × 4096 بارگذاری شد. اگر از ویژگی های اضافی مانند سن، جنس یا ویژگی های آزمایشگاهی استفاده می شد، آنها به صورت ویژگی دودویی (جنسیت؛ 1 ویژگی) یا مقادیر پیوسته (سن، ویژگی های آزمایشگاهی؛ 1 + 6 ویژگی) وارد می شدند. سپس به یک لایه 5N واحدی پنهان (که در آن N تعداد ویژگی های جدولی است) منتقل می شدند. سپس به لایه مترامک الحاق می شدند و در نهایت برای تولید خروجی به یک تابع سیگموئید منتقل می شدند. از اتروپی متقاطع دودویی به عنوان تابع زیان با نرخ یادگیری اولیه 1 × 10<sup>-۴</sup> استفاده شد. ۲۳. بهینه ساز آدام، تابع فعال سازی ReLU، اندازه هسته ۱۶، اندازه دسته ۵۱۲، و نرخ خروج ۰.۲. سایر پارامترهای پیش فرض، مدل ها برای حداکثر ۵۰ دوره (epoch) یادگیری شدند. نرخ یادگیری به ۱ × ۱۰<sup>-۴</sup> کاهش یافت. اگر هیچ بهبودی در تلفات تنظیم برای هفت دوره متوالی وجود نداشت، و اگر تلفات در مجموعه تنظیم برای ۲ دوره کاهش نیافت، فرآیند یادگیری متوقف شد. مدل ها با استفاده از 2.2 TensorFlow و XGBoost 1.5.1 را پایتون 3.8 پیاده سازی شدند. ما تمام مدل های خود را روی درایور NVIDIA نسخه 418.88 با 8 پردازنده گرافیکی Tesla V100-SXM2 و 32 گیگابایت رم برای هر پردازنده گرافیکی آموزش دادیم. آموزش هر مدل DL تقریباً 30 دقیقه در هر دوره طول کشید.

## ارزیابی و تجسم

معیارهای ارزیابی. ما مجموعه داده های کلی ECG را به یک تقسیم تصادفی ۶۰٪ برای توسعه مدل (برای آموزش ۹۰٪ زیرمجموعه از ۶۰٪) و تنظیم دقیق (۱۰٪ زیرمجموعه از ۶۰٪) (مدل) و ۴۰٪ باقی مانده را به عنوان مجموعه داده های پشتیبان برای اعتبارسنجی تقسیم کردیم (شکل ۱). او جدول تکمیلی ۶ برای مدل اصلی بدون آزمایشگاه، و شکل تکمیلی ۱ برای مدل ثانویه با آزمایشگاه. ما اطمینان حاصل کردیم که ECG های یک بیمار بین مجموعه داده های توسعه و ارزیابی به اشتراک گذاشته نشده باشند. ما معیارهای عملکرد زیر را در مجموعه داده های پشتیبان گزارش کردیم - مساحت زیر منحنی مشخصه عملکرد گیرنده (AUROC)، که با نام شاخص C نیز شناخته می شود و مساحت زیر منحنی دقت-یادآوری (AUPRC). همچنین، ما احتمالات پیش بینی را با استفاده از نقاط برش بهینه مشتق شده از مجموعه آموزشی، به کلاس های مرده/زنده تقسیم کردیم. ۲۴. امتیاز F1، ویژگی، فراوانی، دقت و صحت را تولید کردیم. علاوه بر این، کالیبراسیون مدل های خود را ارزیابی کردیم تا ببینیم آیا احتمالات پیش بینی شده با نسبت های مشاهده شده با استفاده از امتیاز بریر مطابقت دارند یا خیر (مقدار پایه ۲۵٪ است؛ امتیاز کمتر نشان دهنده کالیبراسیون بهتر است). ۲۵.

نمونه گیری ارزیابی. انتظار می رود بیمارانی که بیماری های شدیدتری دارند، بیشتر تحت نوار قلب قرار گیرند و در مراحل پایانی بیماری، وقتی بیماران به طور منظم تری تحت نظر باشند، ممکن است تعداد بیشتری نوار قلب گرفته شود. این تنوع در زمان و تعداد نوار قلب در بین بیماران می تواند منجر به سوگیری بالقوه در عملکرد مدل به دلیل نمایش متفاوت فنوتیپ های بیمار شود. برای کاهش چنین سوگیری، ما مدل های خود را با استفاده از یک نوار قلب تصادفی انتخاب شده برای هر بیمار از چندین دوره بیماری در مجموعه بیماران مبتلا، ارزیابی کردیم (شکل 1). او جدول تکمیلی ۶ برای مدل اصلی بدون آزمایشگاه، و شکل تکمیلی ۱ برای مدل ثانویه با آزمایشگاه. این استراتژی نمونه برداری می تواند به جای استفاده از جدیدترین ECG یا ECG که نزدیک ترین زمان به مرگ بیمار گرفته شده است، بیشتر نمایانگر استقرار مدل در یک سناریوی دنیای واقعی بر روی ECG جدید از یک بیمار نادیده گرفته شده در نظر گرفته شود (برای جزئیات به بخش «مقایسه مدل ها» مراجعه کنید). در تحلیل های تکمیلی، عملکرد مدل های خود را در کل مجموعه داده های ذخیره شده، شامل چندین ECG به ازای هر بیمار، و همچنین سازگاری پیش بینی در ECG ها برای بیماران با چندین بررسی کردیم.

۳. لی، دی اس و همکاران. پیش بینی کننده های مرگ و میر در میان ساکنان مراقبت های طولانی مدت مبتلا به عفونت COVID-19. *SARS*. مجله پزشکی سالمندان. جامعه شناسی. ۳۳۷۷:۶۹-۳۳۸۸ (۲۰۲۱).
۴. استیپلر، جی-ای و همکاران. اعتبارسنجی خارجی نمرات اصلاح شده LACE، LACE+ و LACE+ برای پیش بینی بستری مجدد یا مرگ پس از ترخیص از بیمارستان. جی. ایول. کلینیک. پرکتیس. ۱۳۹۰-۱۳۹۰ (۲۰۲۱).
۵. ساراک، ب. و همکاران. آزمایش لیپید، درمان اصلاح کننده لیپید و واجد شرایط بودن مهارکننده PCSK9 (پروپروتئین کانونرتاز سوبوتیلیسین-کسین نوع 9) در 27979 بیمار مبتلا به سندرم حاد کرونری. دور قلب و عروق. کیفیت نتایج. 2021(۰۱۲). e006646
۶. اورلاندی، م. و همکاران. معرفی داروهای ضدانعقاد خوراکی مستقیم، شکاف های درمانی را برای بیماران ضعیف مبتلا به فیبریلاسیون دهلیزی غیردریچه ای برطرف نکرده است. می توان. مجله کاردیول. ۸۴-۷۷ (۲۰۲۲).
۷. مک آلیستر، اف-ای و همکاران. استاتین ها و عفونت SARS-CoV-2: نتایج یک مطالعه کوهورت آینده نگر مبتنی بر جمعیت بر روی 469749 بزرگسال از 2 استان کانادا. انجمن قلب آمریکا، 2021(۰۱۲). e022330
۸. ساندو، آر کی و همکاران. استفاده همزمان از داروهای P-گلیکوپروتئین یا سیتوکروم 3A4 و داروهای ضد انعقاد خوراکی آنتاگونیست غیر ویتامین K در فیبریلاسیون دهلیزی غیر دریچه ای. یورو. قلب. مجله کیفیت. مراقبت. بالینی. نتایج. ۲۰۱-۱۹۵ (۲۰۲۲).
۹. دیانتی ملکی، ن. احتشامی افشار، ا. و آزمسترانگ، پی. دلیو. استفاده از شاخص های الکتروکاردیوگرام ایسکی میوکارد برای طبقه بندی خطر و تصمیم گیری در مورد استراتژی های خون رسانی مجدد. جی. الکتروکاردیول. ۵۲۰۴۷-۵۲۲ (۲۰۱۴).
۱۰. آتیا، زی. آ.، هارمون، دی. ام، بهر، ای. آر. و فریدمن، پی. ای. کاربرد هوش مصنوعی در الکتروکاردیوگرام. یورو. هارت جی. ۴۷۱۷.۴۲-۴۷۳۰ (۲۰۲۱).
۱۱. مینچوله، آ.، کمیس، ج.، لیون، آ. و رودریگز، ب. یادگیری ماشین در الکتروکاردیوگرام. جی. الکتروکاردیول. ۵۷۵، فصل های ۶۱ تا ۶۴ (۲۰۱۹).
۱۲. سان، دلیو. و همکاران. نوار قلب برای غربالگری با توان عملیاتی بالا برای بیماری های متعدد: اثبات مفهوم با استفاده از یادگیری عمیق چند تشخیصی از مجموعه داده های مبتنی بر جمعیت. تصویربرداری پزشکی با NeurIPS ترکیب می شود. <https://doi.org/10.48550/arXiv.2210.06291> (۲۰۲۱).
۱۳. کالیف، آر. ام. مزایای انتقال کیفیت به سطح ملی. ام. هارت جی. ۱۵۶، ۱۰۱۹-۱۰۲۲ (۲۰۰۸).
۱۴. راگونات، س. و همکاران. پیش بینی مرگ و میر از داده های ولتاژ الکتروکاردیوگرام ۱۲ لید با استفاده از یک شبکه عصبی عمیق. نات. مد. ۸۸۶.۲۶-۸۹۱ (۲۰۲۰).
۱۵. دکتر سامسکی و همکاران. روند بستری مجدد و مدت اقامت بیماران بستری شده به دلیل نارسایی قلبی در کانادا و ایالات متحده. جما. کاردیول. ۴۴۴-۴۵۳ (۲۰۱۹).
۱۶. کائپول، پی. و همکاران. تفاوت در درمان، پیامدها و کیفیت زندگی در میان بیماران مبتلا به نارسایی قلبی در کانادا و ایالات متحده. نارسایی قلبی ACC، ۵۲۳-۵۳۰ (۲۰۱۳).
۱۷. کائپول، پی. و همکاران. مرگ و میر طولانی مدت بیماران مبتلا به انفارکتوس حاد میوکارد در ایالات متحده و کانادا: مقایسه بیمارانی که در استفاده جهانی از استریتوکیناز و t-PA برای شریان های کرونری مسدود شده (GUSTO) ثبت نام کرده اند. گردش خون. ۱۷۶-۱۷۵ (۲۰۰۴).
۱۸. مراقبت های بهداشتی حرفه ای فیلیپس. IntelliSpace ECG. سیستم مدیریت ECG. به صورت آنلاین در آدرس زیر موجود است: <https://www.usa.philips.com/healthcare/product/HC860426/intellispace-ecg-ecg-management-system> (۲۰۲۲).
۱۹. جن، تی. و گسترین، سی. ایکس جی بوس: یک سیستم تقویت درختی مقیاس پذیر. در مجموعه مقالات بیست و دومین کنفرانس بین المللی ACM SIGKDD در مورد کشف دانش و داده کاوی (صفحات ۷۸۵-۷۹۴). نیویورک، نیویورک، ایالات متحده آمریکا: (2016). ACM.
۲۰. ریبریو، ای اچ و همکاران. تشخیص خودکار نوار قلب ۱۲ لیدی با استفاده از یک شبکه عصبی عمیق. نات. کمون. ۱۱. ۱۷۶۰ (۲۰۲۰).
۲۱. پرشلت، ل. توقف زودهنگام - اما چه زمانی؟ در شبکه های عصبی: ترفندهای تجارت (ویراستاران. اور، جی. بی. و مولر، کی. آر). 55-69 (1998).
۲۲. هی، ک.، ژانگ، ایکس، رن، اس. و سان، جی. یادگیری عمیق باقیمانده برای تشخیص تصویر. در مجموعه مقالات کنفرانس IEEE در مورد بینایی کامپیوتر و تشخیص الگو (CVPR) (2016) (778-770).
۲۳. کینگما، دی پی و با، جی. آدام: روشی برای بهینه سازی تصادفی. به عنوان مقاله کنفرانسی در سومین کنفرانس بین المللی بازنمایی های یادگیری منتشر شد، سن دیگو. <https://doi.org/10.48550/arXiv.1412.6980> (۲۰۱۵).
۲۴. یودن، شاخص WJ برای رتبه بندی آزمون های تشخیصی. سرطان. ۳۲.۳-۳۵ (۱۹۵۰).
۲۵. بریر، جی. دلیو. تأیید پیش بینی های بیان شده بر حسب احتمال. آب و هوای دوشنبه. ۱.۷۸-۱ (۱۹۵۰).

© نویسنده (گان) ۲۰۲۳

۲۶. سان، ایکس. و زو، دلیو. پیاده سازی سریع الگوریتم دلونگ برای مقایسه مساحت های زیر منحنی های مشخصه عملیاتی گیرنده همبسته. فرآیند سیگنال IEEE. نامه. ۱۳۸۹.۲۱-۱۳۹۳ (۲۰۱۴).
۲۷. سلواراجو، آر. آر. و همکاران. گراد-کم: توضیحات بصری از شبکه های عمیق از طریق محلی سازی مبتنی بر گرادین. کنفرانس بین المللی IEEE در زمینه بینایی کامپیوتر (ICCV) 2017 <https://doi.org/10.1109/iccv.2017.74> (۲۰۱۷).
۲۸. لوندبرگ، اس. ام و همکاران. از توضیحات محلی تا درک جهانی با هوش مصنوعی قابل توضیح برای درختان. نات. ماخ. هوش. ۵۶.۲-۶۷ (۲۰۲۰).
۲۹. مونز، KGM و همکاران. گزارش شفاف از یک مدل پیش بینی چند متغیره برای پیش آگهی یا تشخیص فردی (TRIPOD): توضیح و شرح آن. کارآموز. پزشکی. ۱۶۲ (2015) W1-W73

## تقدیرنامه ها

داده هاز انبار داده های سازمانی خدمات بهداشتی آلبرتا با پشتیبانی پلتفرم خدمات داده و تحقیقات AbsPORU استخراج شده اند که توسط SVK، PK، WS و NS تجزیه و تحلیل ها را انجام دادند؛ WS، SVK، AS، YN بررسی و اظهار نظر کردند SVK، PK، WS و NS تجزیه و تحلیل ها را انجام دادند؛ WS، SVK، AS، YN این مطالعه را آغاز کردند، بودجه آن را تأمین کردند و مسئول کل مطالعه بودند؛ PK و SVK بهداشتی آلبرتا یا هیچ یک از تأمین کنندگان مالی نیست.

## مشارکت های نویسنده

پیش نویس مقاله را تهیه کردند؛ سایر نویسندگان، تجزیه و تحلیل ها و مقاله را به طور انتقادی بررسی و اظهار نظر کردند WS، SVK، PK، WS و NS تجزیه و تحلیل ها را انجام دادند؛ WS، SVK، AS، YN این مطالعه را آغاز کردند، بودجه آن را تأمین کردند و مسئول کل مطالعه بودند؛ PK و SVK

## منافع متضاد

نویسندگان هیچ گونه تضاد منافی را اعلام نمی کنند.

## اطلاعات تکمیلی

اطلاعات تکمیلی نسخه آنلاین شامل مطالب تکمیلی است که در آدرس زیر موجود است. <https://doi.org/10.1038/s41746-023-00765-3>

مکاتبات و درخواست های مربوط به مطالب باید به پادما کائول ارسال شود.

اطلاعات مربوط به چاپ مجدد و مجوزها موجود است در تجدید چاپ <http://www.nature.com/>

یادداشت ناشران: این مجله در مورد ادعاهای مربوط به صلاحیت قضایی در نقشه های منتشر شده وابستگی های سازمانی بی طرف باقی می ماند.

دسترسی آزاد این مقاله تحت مجوز بین المللی Commons Attribution 4.0 Creative منتشر شده است که استفاده، اشتراک گذاری، اقتباس، توزیع و تکثیر در هر رسانه یا قالبی، مادامی که به نویسنده (گان) اصلی و منبع، اعتبار کافی داده شود، پیوندی به مجوز Creative Commons ارائه شود و در صورت ایجاد تغییرات، ذکر شود. تصاویر سایر مطالب شخص ثالث در این مقاله، در مجوز Creative Commons مقاله گنجانده شده اند. مگر اینکه در خط اعتباری مطلب، خلاف آن ذکر شده باشد. اگر مطلبی در مجوز Commons Creative مقاله گنجانده نشده باشد و استفاده مورد نظر شما طبق مقررات قانونی مجاز نباشد یا از حدمجاز تجاوز کند، باید مستقیماً از دارنده حق چاپ اجازه بگیرید. برای مشاهده نسخه ای از این مجوز، به <http://creativecommons.org/licenses/by/4.0/>.

