



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

ویچی سان،^۱ اسونبیل واسو کالامدی،^۲ نریمان سپهروند،^۳ امیر سلیمانی،^۴ ایوسف نادمی،^۵ کوین لین،^۶ یونی،^۷ جاستین ای، ازکوویتز،^۸ اسل گرینزبرگ،^۹ آبرام هیندل،^{۱۰} فینلی ای، مک آیسترنر،^{۱۱} رویندر کی، ساندهوهو،^{۱۲} پادما کائیول،^{۱۳} دنی،^{۱۴}

امکان سنجی و ارزش پیوند 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 زیبایی ها و اندازه گیری های 12 ECG لید از 847773 ECG1605268 از 244077 2007-2020 بهداشتی از یمار 2007-2020 در آلبرتا، کانادا، ما مدل های یادگیری عمیق (DL) مبتنی بر XGBoost و ResNet مبتنی بر تقویت گردایان (XGB) را برای پیش بینی مرگ و میر 30 روزه، 1 ساله و 5 ساله توسعه و اعتبارسنجی کردیم. مدل های مربوط به مرگ و میر ۳۰ روزه، ۱ ساله و ۵ ساله به ترتیب روی ۱۴۶۷۳، ۱۴۱۷۲، ۱۴۱۷۰ و ۱۱۱۰۲ بیمار آموزش داده شدند و روی ۹۷۱۴۴، ۸۹۳۷۹ و ۵۵۶۵ بیمار ارزیابی شدند. در گروه ارزیابی، ۷٪.۶

۱۷.۳٪، ۰.۷۲۹ AUROC، ۰.۷۷۶-۰.۷۸۹ ECG میتوانند این را پیش بینی کنند. در نظریه مراقبت موردنی استفاده قرار گیرند.

پژوهشی دیجیتال npj (۲۱: ۶-۲۳)؛ <https://doi.org/10.1038/s41746-023-00765-3>

مرگ و میر (۱ و ۵ ساله). ما سودمندی افزوده‌ی گنجاندن مقادیر آزمایشگاهی موجود در زمان ECG را در عملکرد پیش‌بینی مدل‌ها بررسی کردیم و عملکردمدل‌ها را در زیرگروه‌های جنسی و تشخیصی بررسی کردیم.

نماج

ویژگی های گروه های بیمار مورد استفاده در مطالعه در جدول شرح داده است.^۱
 میانگین سنی بیماران در زمان گرفتن نوار قلب (ECG) در هر دو مجموعه تی توسعه و عدم توسعه ۸۵.۸ سال بود. با این حال، به یاد داشته باشید که ما به طور تصادفی یک نوار قلب (ECG) برای هر بیمار از مجموعه تی توسعه انتخاب کردیم تا برای ارزیابی های نهایی استفاده شود. میانگین سنی در این مجموعه ای خیر کمتر کمتر و ۶۲.۶ سال بود. این اختلافاً به این دلیل است که بیماران مسن تن نوار قلب بیشتری نسبت به بیماران جوان تراشستند. به طور مشابه، مردان نوار قلب بیشتری داشتند، بنابراین نسبت مردان در مجموعه تی ارزیابی تصادفی کمتر از مجموعه های توسعه یا عدم توسعه بود(۴۷٪ در مقابل ۵۶٪). این الگو برخی از اندازه گیری های ECG (مثلًا میانگین مدت زمان ۹۷.۹ در مقابل ۱۰۱.۳ میلی ثانیه؛ فاصله QT ۳۹۵.۱ در مقابل ۳۹۹.۸ میلی ثانیه)، بیماران های همراه (مثلًا نارسایی قلبی؛ ۴٪ در مقابل ۶٪؛ فیبریلاسیون دلهیزی ۹.۲٪ در مقابل ۱۵.۵٪) و مجمبین مقادیر آزمایشگاهی (مثلًا ترموپوئین ۲.۰ در مقابل ۲.۲ میکروگرم در لیتر؛ کراتینین ۱۰۷.۵ در مقابل ۱۱۶.۵ میکرومول در لیتر، جدول مشاهده شد.^۲)

مدل های مربوط به مرگ و میر ۳۰ روزه، ۱ ساله و ۵ ساله به ترتیب روی ۱۴۱-۷۲، ۱۴۲-۷۳ و ۱۴۳-۷۴ نموده شدند و روی ۸۹۳۷۹، ۹۷۱۴۳ و ۵۵۶۵ بیمار ارزیابی شدند (شکل ۱). ادر مجموعه ارزیابی ما که شامل یک نوار قلب تصادفی به ازای هر بیمار بستره شده بود، به ترتیب ۷۳۹۹ (٪۷.۶)، ۱۸۳۰ (٪۲.۹) و ۱۷۱۷ (٪۱۷.۳) در ۳۰ روز، ۱ سال و ۵ سال فوت کرده بودند. به طور مشابه، نیز زیرمجموعه ای، از نوار قلب های رای

توضیعه یک سیستم سلامت یادگیرنده، سیستم که چرخه ای از جمع آوری و تجزیه و تحلیل منظم داده های سلامت را دنبال می کند تا دانش جدیدی تولید کرد که می تواند برای تصمیم گیری های بهداشتی با بهبود سیستم به کار گرفته شود، اولویت اصلی کانادا و سایر کشورها است.^۱ 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 سریابی و مطب بیشکان) 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 جمیعت در برخی استان ها بهبود یافته اند.^{۵-۶}

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.^۷ در سال های اخیر، پیشرفت های نمایی در منابع محاسباتی و فناوری های یادگیری ماشین، همراه با مجموعه داده های دیجیتالی، فرصت هایی را برای پیش بینی های تشخیصی و پیش آگهی مبتنی بر ECG فراهم کرده است.^{۸-۹} با این حال، امکان سنجی و ارزش پیوند دادن داده های ECG به داده های سلامت اداری در سطح جمیعت طولی برای کمک به بیشکان در تصمیم گیری در مورد مراقبت با هدف تکمیل چرخه کیفیت و تسهیل یک سیستم مراقبت های بهداشتی آموزش قابل بررسی نشده است.^{۱۰-۱۱}

این امر ما را بر آن داشت تا از یک گروه بزرگ جمعیتی از بیماران دارای بیمه‌سلامت همگانی که به بخش‌های اورژانس یا بیمارستان‌ها مراجعه می‌کردند، برای توسعه مدل‌های یادگیری ماشین مبتنی بر ECG جهت پیش‌بینی کوتاه‌مدت (30 روزه) و بلندمدت استفاده کنیم.

اگرده علوم کامپیوتو، دانشگاه آبرتا، ادمونتون، آبرتا، کانادا. مرکز VIGOUR کانادا، دانشکده پزشکی، دانشگاه آبرتا، ادمونتون، آبرتا، کانادا. موسسه هوش مانشین آبرتا، دانشگاه آبرتا، ادمونتون، آبرتا، کانادا. اگرده پزشکی، دانشگاه آبرتا، ادمونتون، آبرتا، کانادا. موسسه قلب اسمیت، سیستم بیمارستانی مرکز پزشکی سدرز-سینای، لس آنجلس، کالیفرنیا، ایالت متحده آمریکا. گران نویسندهان به طور مساوی مشاکت داشته‌اند: درج، ساز، سمتا، واسه کلامداد. ایمیل: okau@ualberta.ca

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

اطلاعات کامل (ن=۳۶۸=۱,۶۵,۳۴۸)	مجموعه توسعه (ن=۹۶۴۷۴)	مجموعه کهادرنده (ن=۵۳۷)	مجموعه ECG (ن=۹۷,۶۳۱)	holdout مجموعه ECG های بیمار در
سن (سال) ۵۶.۷۳	۵۶.۸۱	۵۶.۶	۱۷.۲۹ ± ۶۵.۸۵	۱۸.۵۹ ± ۶۲.۵۷
جنسیت (مرد به درصد) ۴۰.۷۳	۴۰.۷۳			
اندازه گیری های ECG ۵۶.۷۳				
ضریان دهلیزی P ۱۱۶.۳۹ ± ۱۶۳.۹۶	۱۱۶.۳۹ ± ۱۵۵.۷۹	۱۱۶.۹۱ ± ۱۵۵.۷۹	۱۱۶.۹۱ ± ۱۵۵.۷۹	۱۱۴.۶ ± ۱۶۳.۹۶
مدت زمان RR ۲۱۲.۶۳ ± ۷۹.۹۰	۲۱۲.۶۳ ± ۷۹.۹۰	۲۱۳.۸۳ ± ۷۹.۶۸	۲۱۳.۸۳ ± ۷۹.۶۸	۲۰۴.۰۵ ± ۷۹.۰۸۹
شروع موج Q ۶.۳۹ ± ۵.۸۰۸۲	۶.۳۹ ± ۵.۸۰۸۲	۶.۸۲ ± ۵.۸۰۸۷	۶.۸۲ ± ۵.۸۰۸۷	۶.۱۷ ± ۵.۹۰۴
فاصله QT اصلاح شده با ضربان قلب Bazett ۳۸.۰۵ ± ۴۳۴.۸۶	۳۸.۰۵ ± ۴۳۴.۸۶	۳۸.۰۷ ± ۴۳۴.۷۱	۳۸.۰۷ ± ۴۳۴.۷۱	۳۵.۲۳ ± ۴۲۹.۵۵
ضریان قلب ۲۲۳.۲۲ ± ۸۱.۶۴	۲۲۳.۲۲ ± ۸۱.۶۴	۲۳۰.۸۸ ± ۸۱.۶۱	۲۳۰.۸۸ ± ۸۱.۶۱	۲۱۰.۹۴ ± ۸۱.۱۱۳
فاصله PR ۳۸.۴۶ ± ۱۶۹.۳۴	۳۸.۴۶ ± ۱۶۹.۳۴	۳۹.۶۵ ± ۱۶۹.۱۸	۳۹.۶۵ ± ۱۶۹.۱۸	۳۳.۶۵ ± ۱۶۵.۹۹
مدت زمان QRS ۲۲۴.۲۳ ± ۱۰.۱۴	۲۲۴.۲۳ ± ۱۰.۱۴	۲۲۴.۳۰ ± ۱۰.۱۳	۲۲۴.۳۰ ± ۱۰.۱۳	۲۱۰.۶۶ ± ۹۷.۸۹
فاصله QT ۵۰.۸۳ ± ۳۹۹.۹۱	۵۰.۸۳ ± ۳۹۹.۹۱	۵۰.۹۶ ± ۳۹۹.۶۳	۵۰.۹۶ ± ۳۹۹.۶۳	۵۰.۱۱ ± ۳۹۹.۱۰
فاصله QTc Bazett ۴۰.۰۹ ± ۴۵۰.۰۲	۴۰.۰۹ ± ۴۵۰.۰۲	۴۰.۰۹ ± ۴۵۰.۱۰	۴۰.۰۹ ± ۴۵۰.۱۰	۳۷.۱۶ ± ۴۴۹.۲۳
محور P جلویی ۳۵.۵۲ ± ۴۴.۸۵	۳۵.۵۲ ± ۴۴.۸۵	۳۵.۶۹ ± ۴۴.۹۱	۳۵.۶۹ ± ۴۴.۹۱	۳۲.۱۹ ± ۴۴۵.۶۶
محور QRS جلویی در ۴۰ میلی ثانیه اولیه ۴۶.۳۷ ± ۲۷۸.۵۰	۴۶.۳۷ ± ۲۷۸.۵۰	۴۶.۲۰ ± ۲۷۸.۵۷	۴۶.۲۰ ± ۲۷۸.۵۷	۴۲.۱۲ ± ۲۸۰.۴۸
محور QRS جلویی در ترمینال ۴۰ میلی ثانیه ۸۸.۱۵ ± ۴۵۰.۳۶	۸۸.۱۵ ± ۴۵۰.۳۶	۸۷.۹۶ ± ۴۴۴.۸۰	۸۷.۹۶ ± ۴۴۴.۸۰	۸۴.۶۶ ± ۴۴۶.۲۴
محور ST از جلو ۸۸.۲۳ ± ۹۰.۹۴	۸۸.۲۳ ± ۹۰.۹۴	۸۸.۴۸ ± ۹۰.۸۷	۸۸.۴۸ ± ۹۰.۸۷	۸۵.۱۱ ± ۷۹.۲۳
محور T از جلو ۶۷.۶۰ ± ۵۵۰.۷۰	۶۷.۶۰ ± ۵۵۰.۷۰	۶۸.۰۰ ± ۵۵۰.۰۳	۶۸.۰۰ ± ۵۵۰.۰۳	۶۹.۹۰ ± ۴۵۰.۷۷
محورافقی P ۴۷.۳۰ ± ۲۰.۶۹	۴۷.۳۰ ± ۲۰.۶۹	۴۷.۵۲ ± ۲۰.۷۷	۴۷.۵۲ ± ۲۰.۷۷	۴۱.۱۵ ± ۲۱.۰۱
محورافقی QRS افقی در ۴۰ میلی ثانیه اولیه ۴۸.۱۹ ± ۲۷۸.۷۹	۴۸.۱۹ ± ۲۷۸.۷۹	۴۸.۷۱ ± ۲۷۸.۶۹	۴۸.۷۱ ± ۲۷۸.۶۹	۴۰.۲۰ ± ۴۲۶.۸۸
محورافقی QRS افقی در ۴۰ میلی ثانیه ترمینال ۱۳۹.۵۰ ± ۳۴۰.۱۰	۱۳۹.۵۰ ± ۳۴۰.۱۰	۱۳۹.۶۶ ± ۳۴۰.۳۵	۱۳۹.۶۶ ± ۳۴۰.۳۵	۱۲۵.۵۸ ± ۲۶۵.۶۷
محورافقی ST ۷۸.۱۹ ± ۹۰.۹۱	۷۸.۱۹ ± ۹۰.۹۱	۷۸.۶۲ ± ۹۰.۶۱	۷۸.۶۲ ± ۹۰.۶۱	- ۶۹.۲۷ ± ۲۰.۰۳
محورافقی T ۶۴.۹۹ ± ۹۷۸.۰۲	۶۴.۹۹ ± ۹۷۸.۰۲	۶۴.۹۷ ± ۹۷۸.۰۹	۶۴.۹۷ ± ۹۷۸.۰۹	۶۰.۲۰ ± ۹۱.۷۵
بیماری های همراه ۵۸.۹۸ ± ۶۴۰.۴۶	۵۸.۹۸ ± ۶۴۰.۴۶	۵۹.۰۱ ± ۶۴۰.۷۳	۵۹.۰۱ ± ۶۴۰.۷۳	۵۳.۰۵ ± ۵۹.۳۰
بیماری های عروق محیط بیماری ۶۳۹.۳۵۱۸	۶۳۹.۳۵۱۸	(۲۰.۴۷%) ۱۹.۷۱۴	(۲۰.۴۷%) ۱۹.۷۱۴	(۲۰.۲۰%) ۲۱۴۴
عروق مغزی فشار خون بالا ۶۳۹.۴۳۹	۶۳۹.۴۳۹	(۳۰.۳۰%) ۲۱۱۸	(۳۰.۳۰%) ۲۱۱۹۱	(۳۰.۳۶%) ۲۲۵۲
زوال عقل ۶۷.۶۰۵۶	۶۷.۶۰۵۶	(۲۱.۹۵%) ۱۴.۵۸۴	(۲۱.۸۷%) ۲۱۰۲۷۵	(۱۵.۷۶%) (۱۵۳۸۷)
بیماری مزمzon ریوی دیابت شیرین ۶۱۳۳۹۶۳	۶۱۳۳۹۶۳	(۸.۱۲%) ۵۳۰۹۲۶	(۸.۱۳%) ۸۰۰۰۳۷	(۹.۶%) (۸۸۱۴۹)
بیماری کلیوی ۶۱۳۴۲۶۲	۶۱۳۴۲۶۲	(۱.۹۶%) ۱۲۵۴۹	(۱.۹۶%) ۹۲۱۵	(۲.۱۳%) (۲۰۷۸)
بیماری کبد ۶۱۳۴۲۶۸	۶۱۳۴۲۶۸	(۷.۵۶%) ۴۸۱۴۰	(۷.۴۵%) ۷۱۱۶۰	(۵.۸۲%) (۵۶۸۴)
سرطان ۶۱۱۸۰۵	۶۱۱۸۰۵	(۱۰.۳۶%) ۴۶۳۳۸	(۱۰.۰۵%) ۹۶۹۲۴	(۹.۰۱%) (۸۸۰۰)
STEMI ۶۱۱۱۲۷۴	۶۱۱۱۲۷۴	(۱.۱۲%) ۸۰۲۶	(۱.۱۲%) ۱۲۵۲	(۱.۱۱%) (۱۷۹)
ناراساین قلبی ۶۱۰۲۷۶	۶۱۰۲۷۶	(۱.۱۲%) ۷۱۱۸	(۱.۱۲%) ۱۱۷۷	(۱.۱۳%) (۱۳۴۶)
فیبریلاسیون دهلیزی ۶۱۰۱۷۶۴	۶۱۰۱۷۶۴	(۵.۹۱%) ۳۸۱۳۱۴	(۵.۷۷%) ۵۵۶۳۲	(۸.۹۱%) (۸۶۹۹)
انتسوسیون بدون انسداد ظرفیه ST ۶۱۰۱۱۲۷۴	۶۱۰۱۱۲۷۴	(۱۰.۲۳%) ۶۵۱۴۴	(۱۰.۰۳%) ۹۶۸۲۸	(۶.۵۹%) (۶۵۳۴)
STEMI ۶۱۰۱۰۲۶	۶۱۰۱۰۲۶	(۶.۲۲%) ۳۹۱۸۲۵	(۶.۲۲%) ۴۶۳۳۸۱	(۴.۱۵%) (۴۰۴۹)
ناراساین قلبی ۶۱۰۰۵۶	۶۱۰۰۵۶	(۱۰.۵۰%) ۹۹۱۲۷	(۱۰.۵۰%) ۱۵۰۵۵	(۹.۱۸%) (۸۹۵۸)

اندازه گیری های ECG و بیماری های همراه بر اساس تعداد نمونه های ECG بیان می شوند. الگوریتم توصیفی به صورت میانگین ± انحراف معیار با استفاده از یک تکار نمونه برداشتی تصادفی ECG در مجموعه داده ای مورد نظر ارایه شده اند. نوار قلب (ECG) الکتروکاردیوگرام، عدد، استومون بدون انسداد قطعه ST انفارکتوس میوکارد بدون بالا رفتن قطعه STEMİ ST انفارکتوس میوکارد با بالا رفتن قطعه ST.

باتوجه به اینکه مقادیر آزمایشگاهی در دسترس بودند، مدل های مرگ و میر ۳. روزه، ۱ ساله و ۵ ساله به ترتیب روی ۸۴۲۳۹، ۷۸۱۳۴۰، ۳۲۷۴۲ و ۲۱۷۹۶ بیمار آموختند و روی ۵۶.۰۹٪ ۲۱۷۹۶ و ۳۰٪ بیمار ارزیابی شدند (شکل ۱). در مجموعه ارزیابی آزمایشگاهی، به ترتیب ۴۹.۷٪ (۸۸.۸٪)، ۴۹.۷٪ (۹۶.۸٪) و ۸۳.۵٪ (۸۳.۸٪) در نقاط زمانی فوق الذکر فوت کرده بودند.

مقایسه مدل ها مقایسه عملکرد مدل ها در شکل ۱ ارائه شده است. ۲ و ۳. جدول ۱ از ویژگی های سن و جنس به تهایی 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، ۰.۷۶ (۰.۶۴-۰.۸۶) بود.

جدول ۲. ویژگی های آزمایشگاهی گروه های بیمار مورد استفاده در مطالعه (به صورت میانگین ± انحراف معیار).

سن(سال) جنسیت(مرد به درصد)	اطلاعات کامل (ن=۷۰۱۳۰-۷)	مجموعه توسعه (۳۶۱۵۸۵=ن)	مجموعه نگهدارنده (ن=۳۳۹،۷۲۲)	تصادفی به ازای هر بیمار در مجموعه ECG (ن=۷۶)
آذانه گیری های آمایشگاهی میزان خپلر اسپون گلوموبول (میلی لیتر در ۴۰ دقیقه)	۱۷.۵۷ ±۶۴.۸۲	۱۷.۰۳ ±۶۴.۹.	۱۷.۶۲ ±۶۴.۷.	۱۷.۸۷ ±۶۴.۷۴
کراتینین(امول/لیتر)	۱۷.۰۳ ±۶۶.۰۱	۱۷.۰۳ ±۶۵.۸۵	۱۷.۰۳ ±۶۶.۲۵	۱۷.۹۴ ±۶۹.۶۶
هموگلوبین(گرم در لیتر)	۱۱۷.۳ ±۱۱۶.۴۶	۱۱۹.۶۲ ±۱۱۷.۲۰	۱۱۳.۰۰ ±۱۱۵.۳۵	۱۰۰.۶۳ ±۱۰۷.۵۲
پتانسیم(میلی مول در لیتر)	۰.۶۴ ±۰.۰۲	۰.۶۴ ±۰.۰۲	۰.۶۳ ±۰.۰۲	۰.۶۲ ±۰.۰۹
سدیم(میلی مول در لیتر)	۱۴.۳۵ ±۱۳۷.۳۹	۱۴.۳۶ ±۱۳۷.۳۸	۱۴.۳۳ ±۱۳۷.۳۹	۱۴.۳۹ ±۱۳۷.۳۹
تزویجین(۰-۱۰۰٪)	۱۰.۷ ±۲.۲۴	۱۰.۹ ±۲.۲۳	۱۰.۴ ±۲.۲۴	۹.۴۹ ±۱.۹۹

نوارقلب (ECG) الکتروکاردیوگرام، میزان ترشح گلومرولی (GFR) میزان فیلتراسیون گلومرولی، نشماره.

گروه مقاومت. میزان مرگ و میر در بین مردان بیشتر از زنان بود (شکل تکمیلی 4). به طور کلی، مدل های پیش آگهی در مردان کم بهتر از زنان عمل کردند (شکل ...^(۴)).

عملکرد مدل با افزودن ویژگی های آزمایشگاهی مابهیود عملکرد مدل را با افزودن ویژگی های آزمایشگاهی به مدل های ECG بررسی کردیم. ما XGB را با سن، جنس و ویژگی های آزمایشگاهی به عنوان مدل پایه برای مقایسه در نظر گرفتیم. مدل های ECG در تمام نقاط زمانی، حتی بدون مقادیر آزمایشگاهی، AUROC بالاتری نسبت به مدل های پایه داشتند، اما این تفاوت برای پیش بینی های با دامنه طولانی تر کمتر بود. افزودن ویژگی های آزمایشگاهی به طور قابل توجهی عملکرد مدل را در کل، درین مدل ها و نقاط زمانی بهبود بخشید، با این حال، افزایش عملکرد از نظربرگی اندک بود (به طور متوسط ۰.۹۹٪، آزمون DeLong، همه) ب > ۰.۰۰۱). باز هم، مدل کلی DL با ردیابی های ECG، سن، جنس و آزمایشگاه، بهترین عملکرد را در این مقایسه داشت (شکل ۱).^۵ و جدول تکمیلی ۱).

توضیحات مدل

شکل ۱۰ نتایج نگاشت فعال سازی کلاس با وزن گرادیان (GradCAM) را نشان می‌دهد. نواحی ECG را که سهم و ارتباط بیسترنی را با عملکرد پیش‌بینی مرگ و میر مدل دارند، بررسی می‌کند. نواحی بررسی شده مختص به سرب نیستند و بر اساس داده‌های هر ۱۲ لید محاسبه شده اند. تجزیه و تحلیل SHAPley Additive Explanations (SHAP) 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 سهمها در مرگ و میر 1 ساله داشته اند. افزودن ویژگی های آزمایشگاهی، سهم هموگلوبین پایین تر، میزان فیلتراسیون گلومرولی (GFR) پایین تر، ترتوپوئین آ پایین تر، کراتینین بالاتر، سیدم بسیار بالا پایین و بتا سیم بالا را بر جسته کرد (شکل ۱). برای مرگ و میر ۱ ساله و شکل تکمیلی ۵ برای سایر نقاط زمانی).

تحلیل های تکمیلی

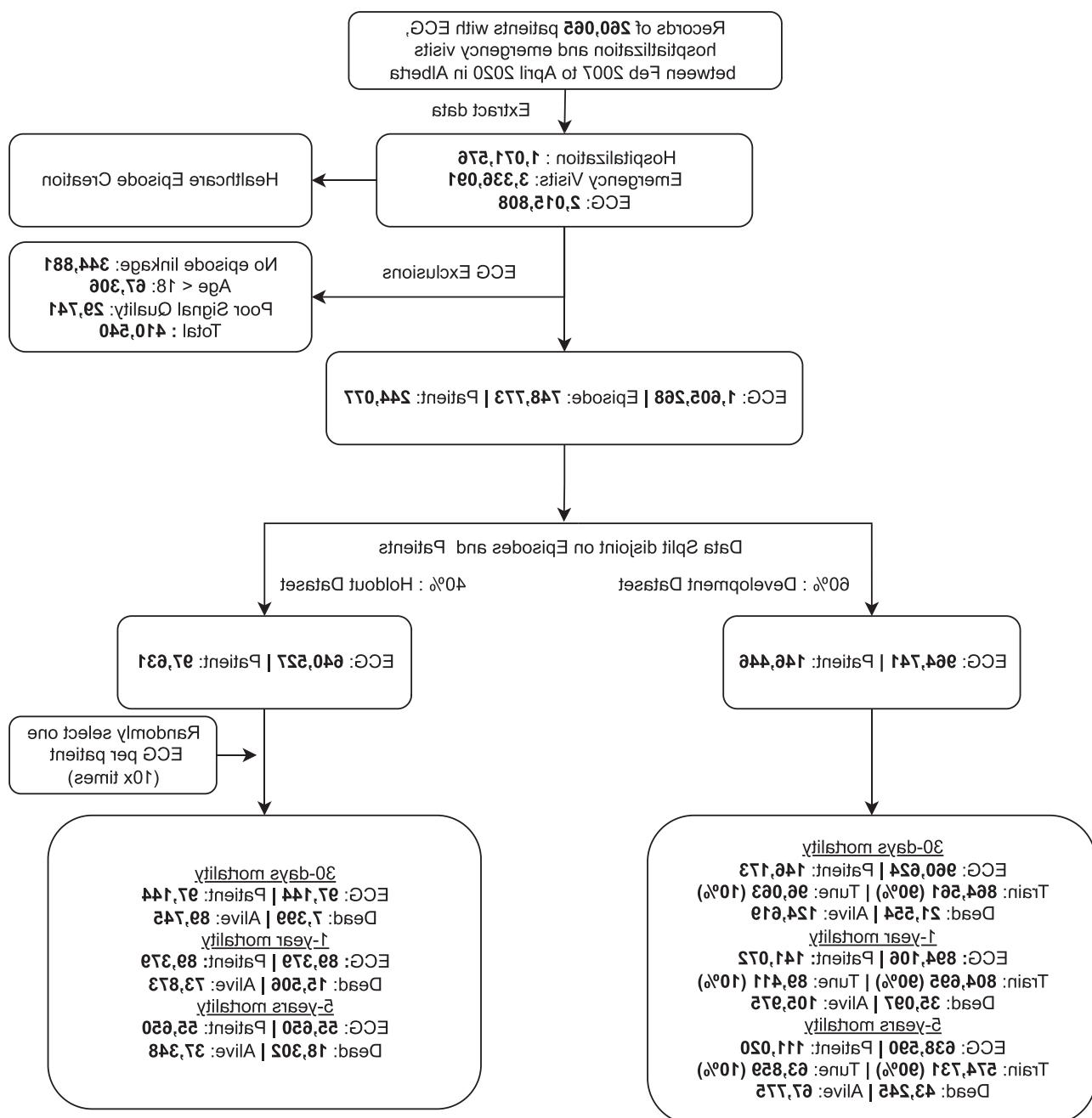
ماچندین تحلیل تکمیلی برای نشان دادن استحکام مدل های خود انجام دادیم. ابتدا، عملکرد مدل های خود را در مقایسه با معماری DL طراحی شده سفارشی که توسط راگونات و همکارانش به کار گرفته شده بود، محقق زدیم.^{۱۴} آموزش مجموعه داده های ما با معماری مشخص شده در مطالعه آنها، منجر به AUROC با مقایسه با 0.796 (برای پیش بینی مرگ و میر) 0.800 و 0.792 ساخته شد، در مقایسه با 0.826 (برای پیش بینی مرگ و میر) 0.83 (برای رفتہ توسط راگونات و 0.822 (برای معماری ما. مدل DL مبتنی بر ResNet^{۱۵} می باشد، در مقایسه با معماری DL به کار رفته توسط راگونات و همکارانش برای هر سه نقطه زمانی، بهبود کوچک اما از نظر آماری معنی داری را در عملکرد AUROC نشان داد (ازمون DeLong، همه پا...، شکا تکمیل، ...).

دوم، ما مدل DL اولیه خود را بر روی موارد زیر ارزیابی کردیم: (الف) یک مجموعه شتبانی که شامل ECG های بر کتفیتی بود که قلّاً وجود داشتند

(0.704-0.723) برای مرگ و میر 1 ساله و (0.776-0.787) برای مرگ و میر 5 ساله. مدل پادگیری عمیق (DL) با ردیابی های ECG به تنهایی، عملکرد قابل توجه بالاتری با (0.812-0.808) و (0.848-0.833) به ترتیب برای پیش بینی 30 روزه، 1 ساله و 5 ساله داشت (آزمون DeLong). همه پ < 0.001 استفاده از سن و جنس به همراه ردیابی های ECG، بهبودهای کوچک اما قابل توجه را با (0.826-0.822) و (0.857-0.847) AUROC و (0.828-0.852) AUROC ایجاد کردند. برای سه نقطه زمانی نشان داد (آزمون DeLong)، همه پ < 0.832-0.824) برای هر سه نقطه زمانی نشان داد (آزمون ECG با همه XGBoost عملکرد به طور قابل توجهی بهتری اندازه گیری های (XGB) نسبت به ECG با ردیابی DL ب (...).

مدل DL با ردیابی های ECG، سن و جنس بهترین مدل در این مقایسه بود، با AUROC اندازه گیری های ECG به طور مداوم بالاتر از 0.82 بودند. مدل XGB با اندازه گیری های ECG به مدت 1 سال عملکرد بهتری نسبت به مدل فقط سن و جنس نداشت (آزمون DeLong، $P = 0.57$) و به مدت 5 سال به طور قابل توجهی بدتر از سن و جنسیت بود (آزمون دلونگ، $P = 0.001$)؛ با این حال، DL با ردیابی ECG هنوز اطلاعات مرتبط را برای پیش بینی ارجاعیه می دهدو در تمام نقاط زمانی به طور قابل توجهی از مدل پایه سن و جنس بهتر عمل می کند (آزمون DeLong، $P < 0.001$) - که بدین ترتیب بر سودمندی پیش اگهی مدل های DL مبتنی بر ردیابی ECG نسبت به مدل های معمول باستفاده از اندازه گیری های ECG تأکید می کند. جدول ۳ عملکرد برتر مدل های DL با ردیابی های ECG را از نظر مساحت زیر منحنی دقت-یادآوری (AUPRC)، امتیاز F1، امتیاز Brier و سایر معیارها نشان می دهد.

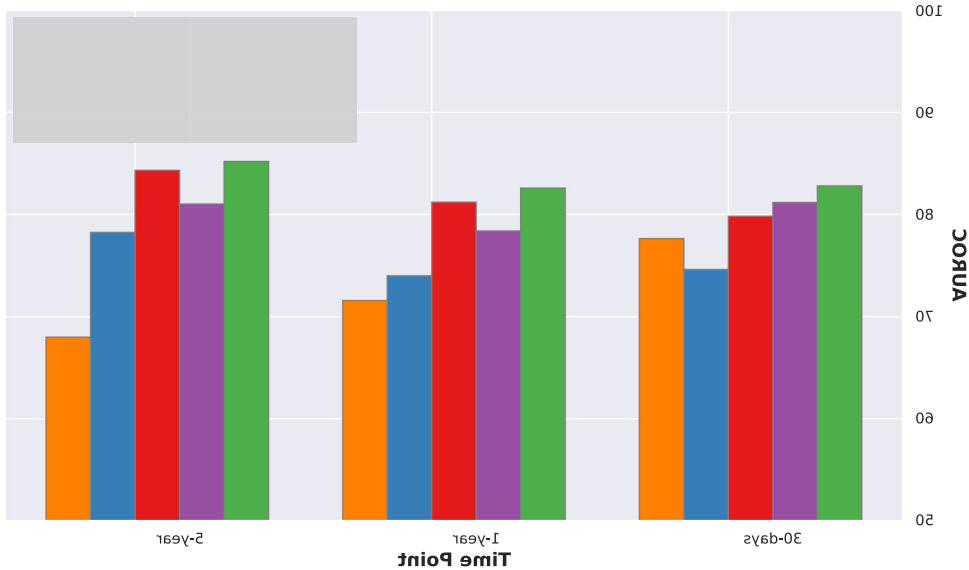
عملکرد مدل در زیرگروه های تشخیصی و جنسیتی
میزان مرگ و میر در گروه های تشخیصی مورد نظر به طور قابل توجهی متفاوت است (شکل تکمیلی 3) و بیماران مبتلا به نارسایی قلبی در هر نقطه زمانی بالاترین میزان مرگ و میر را داشتند. شکل ۴ عملکرد مدل های ما را در این زیرگروه های تشخیصی مختلف نشان می دهد. مدل ها در بیماران مبتلا به انفارکتوس میوکارد با بلا رفت رفتن قطعه (STEMI) و انفارکتوس میوکارد بدون بلا رفت رفتن قطعه (NSTMI) (به ترتیب AUROC 0.867 و 0.882 برای مرگ و میر 1 ساله) نسبت به گروه کلی بهتر عمل کردند. عملکرد مدل در زیرگروه های دیگر (نارسایی قلبی، دیابت و فیبریلاسیون دهیزی) کمتر از ۰.۸۵ کارایی داشت.



شكل اگروده مطالعه، تقسیم بندی داده ها و طرح ارزیابی-فلوچارت 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 ۶٪ برای توسعه مدل (برای آموختن ۹٪) و تنظیم دقیق ۱۰٪ زیرمجموعه ۶٪ مد (و ۴٪ باقی مانده را به عنوان مجموعه داده های پشتیبان برای اعتبارسنجی تقسیم کردیم. انتظار می رود بیمارانی که بیماری های شدیدتری دارند، پیشتر تحت ECG قرار گیرند و در مراحل پایانی که بیماران به طور منظم نری تحت نظر هستند، ممکن است ECG های پیشتری گرفته شود. برای کاهش سوکبری بالقوه در عملکرد مدل به دلیل نمایش متفاوت فتوتیپ 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 مراجعت 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 آموختنی و آزمایشی ما کامل از هم جدا هستند، ECG های بیمارانی را که در طول دوره مطالعه در هر دو بیمارستان آموختنی و آزمایشی بستره 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 تحیل حذف شد؛ و (ب) یک مجموعه ECG با کیفیت پایین به تنها (ECG جدول تکمیلی 2) ما دریافتیم که مدل DL ما در برای مصنوعات سینگال ECG و مشکلات 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 ۰.۲۵٪ برای مرگ و میر ۱ ساله)، علاوه بر این، ارزیابی بر روی ECG های با کیفیت پایین به تنها هنوز هم بیش از ۸۰٪ امتیاز AUROC را برای هر سه نقطه زمانی نشان داد. به ترتیب ۳.۴٪، ۲.۶٪ و ۲.۵٪ 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 ردیاها و اندازه گیری های ECG. میله های خط، فواصل اطمینان ۹۵٪ را در اطراف میانگین نشان می دهند. DL با ردیا های 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؛

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

تا آنکه 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 همکاران¹⁴ این مطالعه بر اساس تقریباً ۱.۳ میلیون نوار قلب از کمی پیش از ۲۵٪ 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 جنسی، در یک گروه آزمایشی ۹۱٪ AUROC، برابر با ۸۷٪. را به دست آورد، در حالی که در گروه اعتبارسنجی ما که شامل ۸۹۳۷۹ بیمار بود، AUROC، برابر با ۸۲٪ بود. تفاوت های متعددی در مطالعه کانادایی ما و مطالعه راگونات و همکارانش از ایالات متحده وجود دارد. گروه مامسن تر بود (میانگین سنی ۶۲.۶ ± ۱۸.۶ در مقابل ۵۸ ± ۱۸ سال) و میزان مرگ و میر ۱ ساله آنها در مقایسه با گروه آمریکایی به طور قابل توجهی بالاتر بود (۴٪ در مقابل ۸٪ میزان مرگ و میر ۱ ساله در مجموعه های موردن 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 درنحوه مدیریت بیماران در دو سیستم مراقبت های بهداشتی نسبت داده شده است.¹⁵ راگونات و همکارانش AUROC بالاتر از مرتبط با مدل XGB است¹⁶. مکانیسم اساس سن و جنس به تهایی گزارش کردن (0.774)، در حالی که مدل ماکه مرگ و میر ۱ ساله را بر اساس سن و جنس بیمار پیش بینی کرد، AUROC 0.716 داشت. مطالعه ما از مدل های استاندارد (ResNet) DL AUROC 0.716 استفاده کرد و مبتنی بر ECG های یک تولیدکننده تجهیزات واحد (فیلیپس) بود. در حالی که مطالعه ایالات متحده از عماری DL طراحی شده سفارشی استفاده کرد و مبتنی بر ECG های تولیدکننده ای داده های مخالف بود. پیاده سازی عماری DL آنها بر روی داده های ما منجر به عملکرد کوچک اما به طور قابل توجهی دارند. با توجه به این تفاوت های قابل توجه در بیماران، سیستم های بهداشتی، تجهیزات و ساختار مدل، هر دو مطالعه میزان مشابه از بیهود در عملکرد مرتبط با افزودن ردیابی ECG را یافتند. این یافته ها از قابلیت تعیین 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.¹⁶ را توسعه مدل های برای پیامدهای مرگ و میر کوتاه مدت (۳۰ روزه) و بلندمدت (۵ ساله)؛ و بررسی عملکرد مدل های 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.

۰.۱۳٪ (۸۴.۴۶-۸۵.۸۳٪) برای مدل های ۳۰ روزه، ۰.۱۲٪ (۸۲.۵۸-۸۳.۸۳٪) برای مدل های ۱۲ ماه و ۰.۱۲٪ (۸۲.۸-۸۵.۵۹٪) برای مدل های ۵ ساله در اعتبارسنجی H1؛ اما در اعتبارسنجی H2، برای مدل های ۳۰ روزه ۰.۱۳٪ (۸۱.۵۹-۸۵.۱۹٪)، برای مدل های ۱۲ ماه ۰.۱۷٪ (۸۰.۵۸٪) و برای مدل های ۵ ساله ۰.۱۸٪ (۸۰.۹۶-۸۲.۸٪) کمتر بود.

چهارم، ما عملکرد مدل های خود را در کل مجموعه ای «بازاراندنه» (holdout)، شامل چندین نوار قلب (ECG) برای بیماران «بازاراندنه» بررسی کردیم. در مجموع، ۷۸۲۵۰ (۸۰.۱۵٪) ۷۱۶۳۶ (۷۸.۴۰٪) بیمار پیش از یک نوار قلب (ECG) برای پیش بینی های ۳۰ روزه، ۱ ساله و ۵ ساله در مجموعه ای «بازاراندنه» ما در دسترس داشتند. عملکرد مدل های ما پس از گجاندن 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 ۴ و شکل تکمیلی ۸). علاوه بر این، ما در پیش بینی در سراسر ECG برای بیماران حداقل ECG داشتند، سازگاری یافته (۸۵.۰۵٪)، سراسر ECG های متعدد خود در ۳۰ روزه، ۱ ساله و ۵ ساله داشتند، شکل تکمیلی (9).

بحث

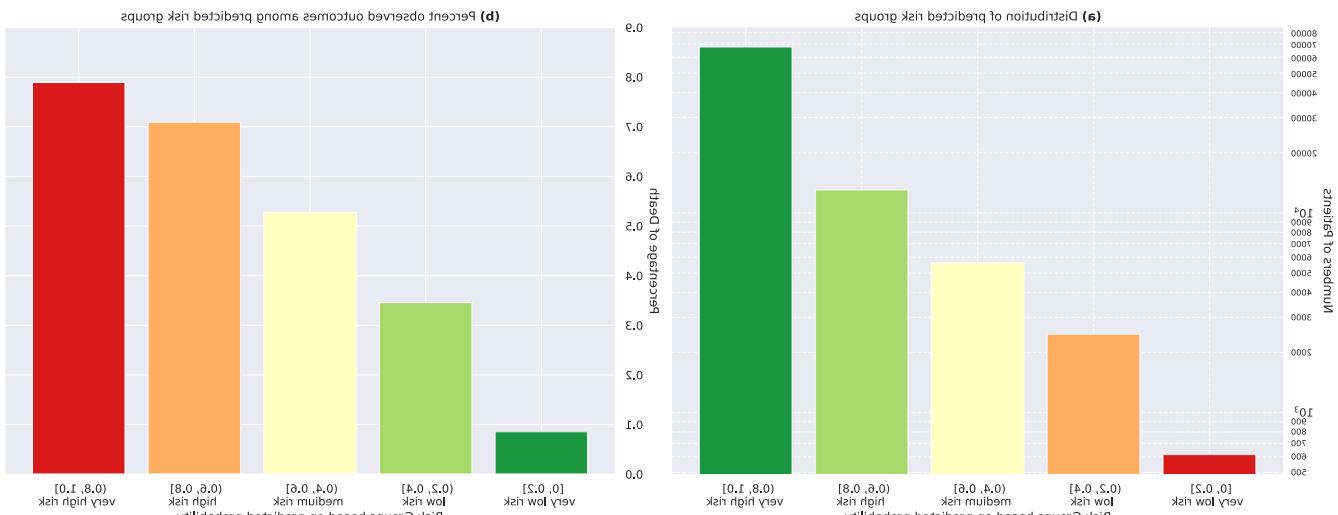
مطالعه ما، که بر اساس یک گروه بزرگ جمعیت از بیماران با دسترسی جهانی به مراقبت های بهداشتی انجام شده است، کاربرد مدل های یادگیری ماشینی مبتنی بر داده های ECG را برای شناسایی بیماران در معرض خطر بالای مرگ و میر کوتاه مدت و بلند مدت در زمان مراجعه به اورژانس یا بیمارستان نشان می دهد. ما دریافتیم که مدل های DL (ResNet) مبتنی بر داده های ECG ۱۲ یا ۱۰۰۰۰۰ بیمار، ردیابی های ECG مبتنی بر اندازه گیری های ECG گزارش شده به تقویت گردید (XGBoost)؛ این تقویت با کمتری بهتری دارند. در یک گروه اعتبارسنجی شامل تقریباً ۰.۱۳٪ (۸۴.۴۶-۸۵.۸۳٪) برای مدل های ۳۰ روزه، ۰.۱۲٪ (۸۲.۵۸-۸۳.۸۳٪) برای مدل های ۱۲ ماه و ۰.۱۲٪ (۸۲.۸-۸۵.۵۹٪) برای مدل های ۵ ساله در اعتبارسنجی H1 است که ارزش افزوده داده های ای تکمیلی، استحکام عملکرد مدل های ما را دریافتیم که مدل های مبتنی بر داده های تقویتی با کمیت 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 داده های آزمایشی کوچک را در عملکرد ECG های با کمیت پایین، در بیمارستان های تقویتی با کمیت پایین، در بیمارستان های مختلط شناختی (سن و جنس بیماران) بهترین عملکرد را در پیش بینی مرگ و میر کوتاه مدت و بلند مدت دارند. این یافته ها نشان می دهد که چکونه 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 عملکرد های مختلف مدل که به صورت مبادگن (فاصله اطمینان ۹۵٪) درصد بیان شده اند.

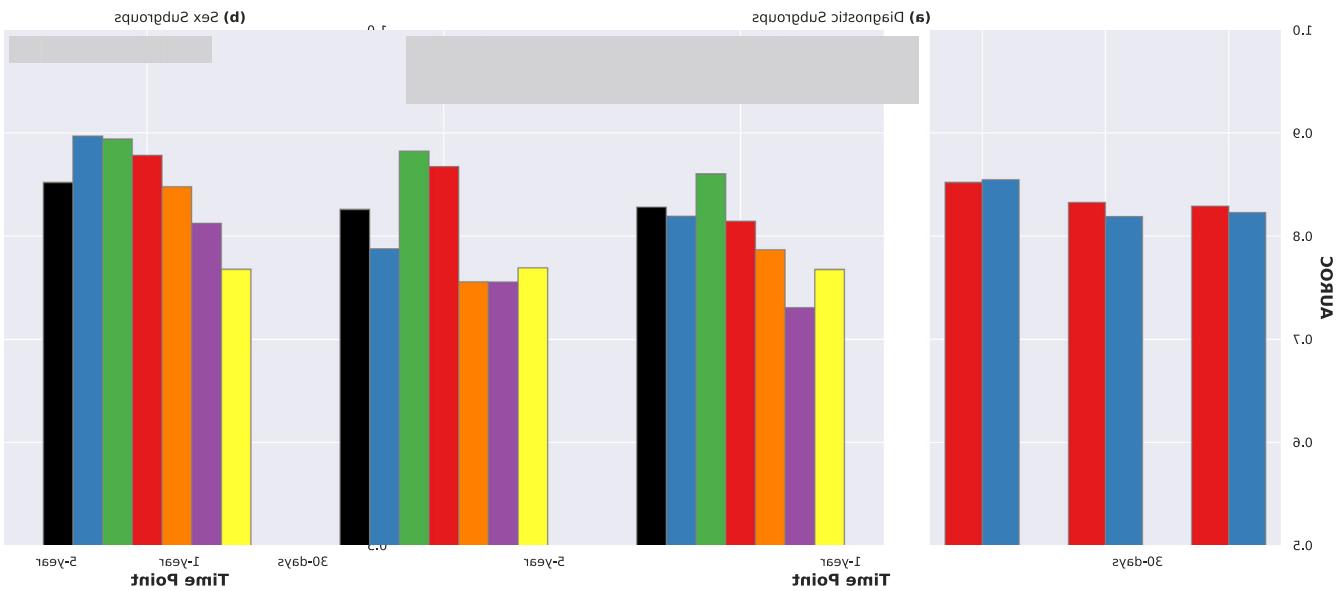
پردازشگاه	مدل	AUROC	AUPRC	F1 امتیاز
جداول ایرانی	مکارهای مختلف مدل که به صورت می‌آگین (اصاله اطهربنام ۷۵٪) درصد بیان شده‌اند.			
ویژگی های ناقله زمانی				

منتشرشده با همکاری بیمارستان بوندانگ دانشگاه ملی سئول

npj Digital Medicine (2023) 21



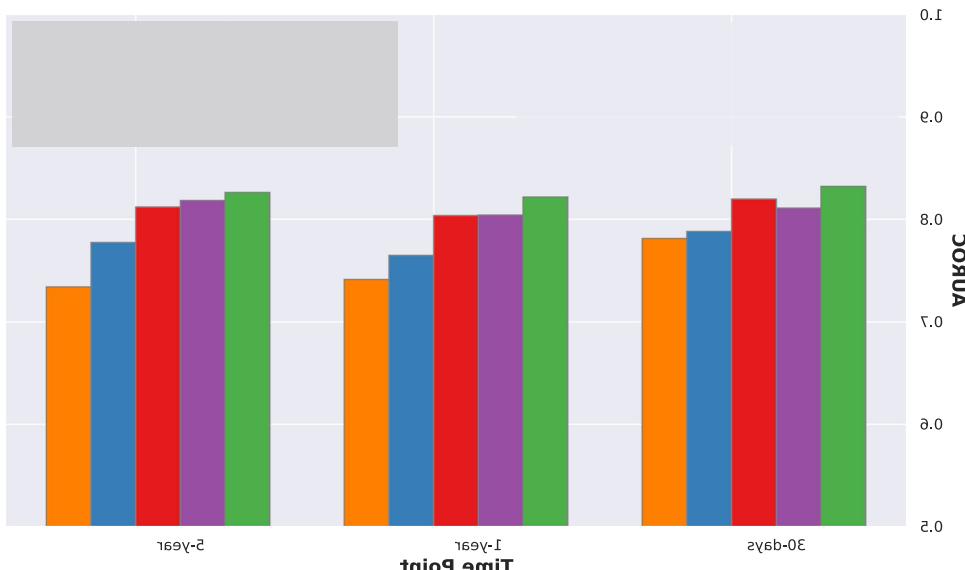
شکل ۳: گروه های خطر پیش بینی شده برای مرگ و میر ا ساله، الفمیا نیچه گروه خطر - گروه های خطر «بسیار کم»، «کم»، «متوسط»، «زیاد» و «بسیار زیاد» را بر اساس درصد نقاط برش ۰ تا ۲۰ درصد، ۲۰ تا ۴۰ درصد و غیره از احتمال پیش بینی شده مرگ از مدل های اصلی خود (DL): نوار قلب، سن، جنس در مجموعه افراد تحت بررسی استخراج 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٪، ۵۲.۳٪، ۳۴.۶٪ و ۲۸.۶٪ مرگ در گروه های خطر «بسیار کم»، «کم»، «متوسط»، «زیاد» و «بسیار زیاد»). یادگیری عمیق DL، نوار قلب.



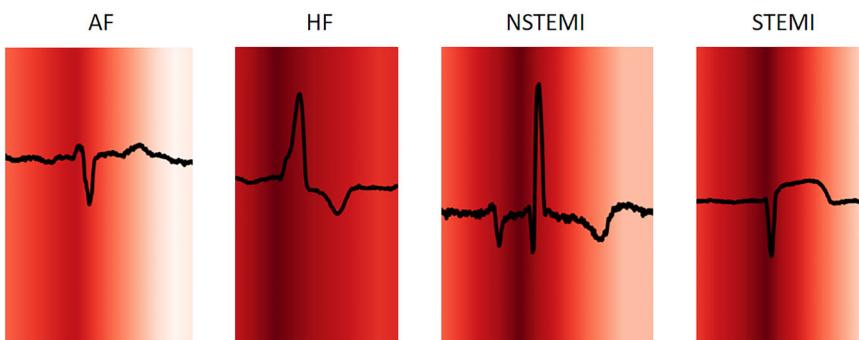
شکل ۴: عملکرد مدل در زیرمجموعیت های تشخیصی و مبتنی بر جنسیت. الف عملکرد مدل های ECG، سن، جنسیت در زیرگروه های مختلف تشخیص اوایله، این مدل ها در بیماران مبتلا به STEMİ و NSTEMİ AUROC ۰.۸۶۷ و ۰.۸۸۲ برای مرگ و میر ۱ ساله، به ترتیب) نسبت به گروه کلی بهتر عمل کردند. عملکرد مدل در زیرگروه های دیگر (نارسایی قلبی، دیابت و فیبرولاسیون 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 STEMİ در مرحله ST در مرحله STEMİ.

مامعتقدمیم 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 متواالی ساختیم، که فقط با سن و جنس شروع شد و سپس ردپاها با اندازه گیری های ECG و متعاقباً داده های آزمایشگاهی را اضافه کردیم. AUROC برای مدل مرگ و میر ۱ ساله از ۰.۸۱ به مدل مبتنی بر سن، جنس و ردپاهای ECG به ۰.۸۳. برای مدل مبتنی بر سن، جنس، ردپاهای ECG و آزمایش افزایش یافت. داده های آزمایشگاهی ممکن است اطلاعات پیش آگهی بیشتری را در جمعیت های خاص بیمار (مثلًاً بیماران مبتلا به سندروم های حاد کرونی یا بیماری کلیوی و غیره) ارائه دهند و افزون سایر معیارهای آزمایشگاهی مانند ALT، AST، HbA1c و ممکن است عملکرد مدل های ما را بهبود بخشد. این بررسی ها به عنوان بخشی از مطالعات تحقیقاتی آینده برنامه ریزی شده اند.

امتیازAUROC برابر با ۰.۸۵ در ۳۰ روز و ۰.۸۳ در ۵ 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 نزدیک ترین شود. این نشان می دهد که اگرچه ممکن است مزایای واضح در رابطه با کاربرد پیش بینی ECG در هدایت کوتاه مدت تا میان مدت درمان ها وجود داشته باشد، استفاده از ECG به تنها بدون ویژگی های سن و جنس ممکن است ارزش پیش بینی کننده کافی برای نتایج مرگ و میر ۵ ساله نداشته باشد. همچنین، ما عملکرد مدل متفاوتی را در زیرگروه های تشخیص 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 ویژگی های آزمایشگاهی. میله های خطای اطمینان ۹۵٪ را در اطراف میانگین نشان می دهند. مدل های ECG در تمام نقاط زمانی، حتی بدون مقادیر آزمایشگاهی AUROC بالاتری نسبت به مدل پایه (XGB: سن، جنس، آزمایشگاه) داشتند، اما این تفاوت برای بیشین های با بازه زمانی طولانی ترکمتر بود. افزودن ویژگی های آزمایشگاهی، عملکرد مدل را در کل، در مدل ها 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 ردیابی های ECG، سن، جنس و آزمایشگاه، بهترین مدل از نظر عملکرد در مقایسه بود. AUROC سطح زیر منحنی مشخصه عملکرد گیرنده، یادگیری عمیق DL، الکتروکاردیوگرام ECG، XGBoost

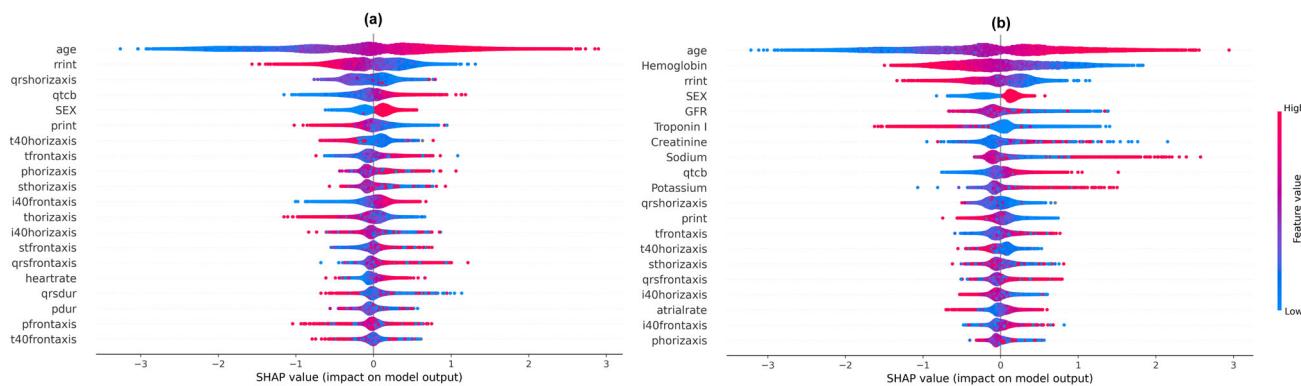


شکل ۶ نقشه حرارتی GradCAM برای مرگ و میر ۱ ساله. ردپاهای ECG نماینده با تشخیص های اولیه AF، HF، STEMI و NSTEMI انتخاب شدند. نواحی تیره تر در هر ردپادر GradCAM نشان دهنده نواحی با بیشترین سهم در پیش بینی مرگ و میر ۱ ساله مدل ResNet DL: ECG هستند. مناطق برجسته شده مختص به سرب نیستند و بر اساس داده های هر ۱۲ لید محاسبه شده اند. فیریلاسیون دهلیزی AF، یادگیری عمیق DL، الکتروکاردیوگرام ECG، نارسایی قلبی HF، STEMI، انفارکتوس میوکارد با بالارفتن قطعه ST، NSTEMI، STEMI، انفارکتوس میوکارد بدون بالا رفتن قطعه ST.

اینکه یافته های ما تا چه حد قابل تعمیم به ECG های سایر تولیدکنندگان تجهیزات است، باید مشخص شود. دوم، داده های 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 داده های آزمایشگاهی داشتند (۱۹٪) کمی بیشتر از بیمارانی بود که داده های آزمایشگاهی نداشتند (۱۷٪). در حال که افزودن ویژگی های 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 آزمایش های آزمایشگاهی موجود در مدل های ما جامع نیست. چهارم، تصادفی ما به ازای هر گروه بیمار، در مقایسه با کل مجموعه بیماران تحت درمان، کم جوان تر بود و بیماران مرد کمتری داشت و بیماری های همراه کمتری نیز داشت که به ECG های بیشتری که در بیماران مسن تر، مرد و دارای عوارض بالینی با بیماری های همراه بیشتر انجام نشود، نسبت داده می شود. با این حال، ما در پیش بینی بین این دو گروه از ECG در مجموعه بیماران تحت درمان داشتند، بین بیمارانی که چندین سازگاری قابل توجهی یافیم، و در آخر، اندازه گیری های ECG مورد استفاده در مدل های XGB از طریق دستگاه های فیلیپس ارائه شده اند و توسط آزمایشگاه اصلی یا متخصصان انسانی بررسی نشده اند.

مدل های DL با شبکه های عصبی کانولوشنی، در شناسایی و تفسیر الگوهای مورد استفاده مدل برای پیش بینی، به عنوان جعبه های سیاه در نظر گرفته می شوند. ما تکنیک هایی مانند اجبار نقشه های حرارتی GradCAM برای این منظور امتحان کرده ایم که نشان می دهد فواصل PR، QRS و تغییرات ST-T، به ویژه بخش اولیه کمپلکس QRS، بیشترین سهم را در پیش بینی مرگ و میر دارند. این امر عمدتاً در شرایط مختلف بیماری ثابت بود، با این حال، تقریباً تمام بخش های ECG در پیش بینی مدل DL در بیماران مبتلا به نارسایی قلبی نقش داشتند. لازم به ذکر است که این تکنیک های تجسم، حوزه ای از تحقیقات فعلی هستند و استخراج تفاصیر بالینی معنادار از سیگنال های قلبی ECG با ضربان های قلبی متعدد، پالش برانگیز است. به عنوان مکمل، ما از تجزیه و تحلیل SHAP برای مدل های XGB استفاده کردیم که چند پارامتر ECG (مانند فاصله RR پایین تر، محور QRS افقی پایین تر، فاصله QT بالاتر و فاصله PR پایین تر) را که بیشترین سهم را در پیش بینی مرگ و میر در دوره های مختلف پیگیری داشتند، برجسته کرد.

مطالعه ما محدودیت هایی دارد. اولاً، تمام ECG های مورد مطالعه ما از یک تولیدکننده (سیستم اطلاعاتی فیلیپس) بودند.



شکل ۷. 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 QT پایین تر، محور QRS افقی (اثر شرط)، فاصله QT اصلاح شده با نخر Bazett بالاتر، جنس مذکور و فاصله PR پایین تر پیشترین سهم را در مرگ و میر ۱ ساله داشته اند. ب. افزون و پیچیدگاهی از آزمایشگاهی، سهم هموگلوبین پایین تر، تروپونین I پایین تر، کراتینین بالاتر، سدیم بسیار بالا پایین و پتاسیم بالا بر جسته 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 اوایله، محور QRS پیشانی در ۴۰ میلی ثانیه پایانی، محور ST پیشانی، محور موج QRS پیشانی (اعمال انحراف ST)، محور T پیشانی، محور P افقی، محور QRS افقی در ۴۰ میلی ثانیه اوایله، محور T 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 گلومرولی (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 پامدهای 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 آمریکا، کانادا، در ۱۴ مرکز بستره شده بودند. داده های ما شامل ۲۰۱۵، ۲۰۸۸ نوار قلب، ۳۳۶، ۵۷۶، ۱، بستره در بیمارستان و ۶۵، ۲۶۰، ۱۴۰ بیمار بود. مواجهه های همزنان با مراقبت های بهداشتی برای یک بیمار (مراجعة به اورژانس و/یا بستره 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 اینزوده گروه بندی شدند. نمودار جریان درخت نصیم گیری مورد استفاده برای تعریف اپیزود در شکل تکمیلی ۱ نشان داده شده است.

اگر تاریخ ثبت نوار قلب (ECG) در بازه زمانی بین تاریخ پذیرش و تاریخ ترجیخ یک دوره بود، یک رکورد ECG به یک دوره مراقبت های بهداشتی مرتبط می شد (شکل تکمیلی ۱۱). ECG های با کیفیت پایین از طرق 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 نمى توانستند به هیچ دوره ای مرتبط باشند، ECG های بیماران زیر ۱۸ سال و همچنین ECG های با کیفیت سیگال پایین، گروه تجزیه و تحلیل ما شامل ۷48773 ECG از ۱605268 دوره در ۷44077 بیمار بود. به شکل ۱ مراجعة 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 بنی کیفیت که از مطالعه اصلی حذف شده بودند، ارزیابی کردیم.

درنتیجه، مطالعه ما نشان می دهد که مدل های DL مبتنی بر ECG می تواند 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

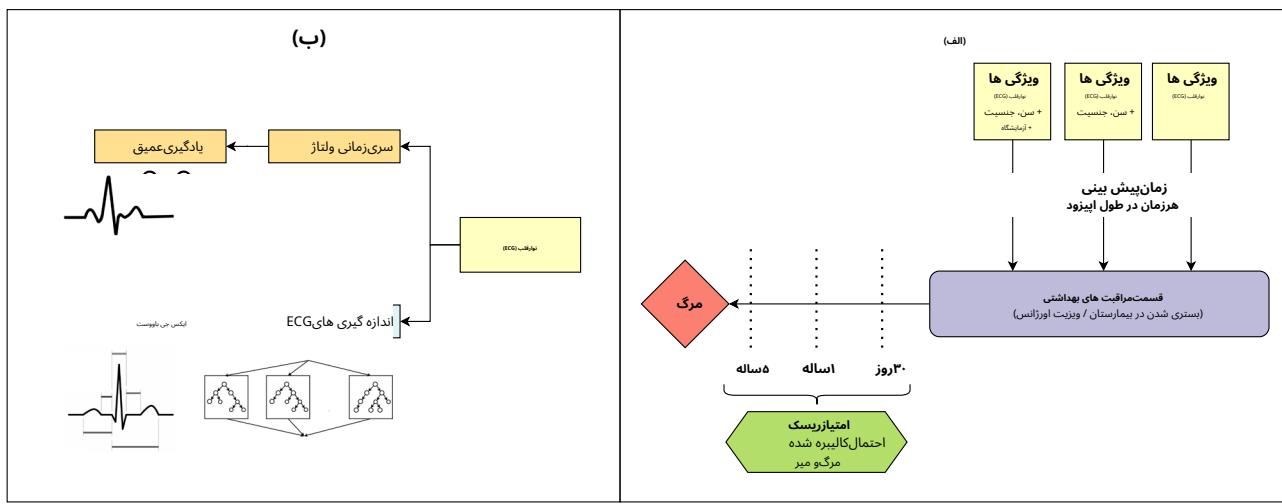
استان البرتا، کانادا، دارای یک سیستم مراقبت های بهداشتی تک پرداخت کننده (وزارت بهداشت: البرتا هلت) و تک ارائه دهنده خدمات بهداشتی آبرتا) است. حدود ۴۰۰۰ میلیون نفر از ساکنان این استان به خدمات بیمارستانی، سرپایی، آزمایشگاهی و پزشکی دسترسی جهانی دارند.

برای این مطالعه، داده های نوار قلب (ECG) با استفاده از یک شماره سلامت بیمار منحصر به فرد به پایگاه داده اداری زیر متصل شدند: (۱) پایگاه داده چکیده ترجیح (DAD) که شامل داده های مربوط به ستری ها (از جمله تاریخ پذیرش، تاریخ ترجیح، تشخیص اصلی، تا ۲۴) تشخیص دیگر و وضعیت ترجیح (انتقال، ترجیص به منزل، فوت) است. (۲) پایگاه داده سیستم گزارش دهنده مراقبت های سرپایی ملی (NACRS) از تمام ویزیت های کلینیک سرپایی بیمارستان (از جمله بخش اورژانس) که شامل تاریخ پذیرش، تشخیص اصلی، تا ۲۴ تشخیص دیگر و وضعیت ترجیح است. (۳) رجیستری طرح بیمه مراقبت های بهداشتی آبرتا (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 حیاتی نسبت به سوابق رجیستری مرگ و میر (۱٪ از بیماران)، رجیستری وضعیت داده شد.

این مطالعه توسط هیئت اخلاق پژوهشی دانشگاه آبرتا (Pro00120852) تأیید شد. هیئت اخلاق ترجیح 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 عمل نخواهد بود.

داده های نوار قلب

این مطالعه از ردیاب های استاندارد ECG دوازده لید و اندازه گیری های ECG از سیستم Philips IntelliSpace ECG استفاده کرد. ردیاب های ECG سری های زمانی ولتاژ بودند که با فرکانس ۵۰ هرتز به مدت ۱۰ ثانیه برای هر یک از ۱۲ لید (۵۰۰ نمونه برداشت شدند. ۱۲۰۱۰ 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, Q منشأ شده با همکاری بیمارستان بوندانگ دانشگاه ملی سئول



شکل ۱: شماتیک وظایف پیش‌بینی، انواع ویزگی‌ها و انتخاب الگوریتم‌های پادگیری. الفا: احتمال مرگ بیمار در عرض ۳۰ روز، ۱ سال و ۵ سال، از روز اخذ ECG، متمرکز بود. ECG های مورد استفاده در این مدل ها می‌توانستند در هر نقطه زمانی در طول یک دوره مراقبت‌های بهداشتی اخذ شده باشند. مدل ها شامل ویزگی‌هایی با من فقط نوار قلب (ECG)، دوم نوار قلب + سن، جنس و آناتور قلب + سن، جنس + آزمایش‌های آزمایشگاهی. هدف مدل‌های پیش‌بینی از احتمال کالبیره شده از مرگ و میر است که می‌تواند در ارزیابی ریسک بیمار مورد استفاده قرار گیرد. بد داده‌های نوار قلب بیمار معمولاً توسط مراکز درمانی به یکی از دو شکل زیر باگانی می‌شوند: من به عنوان یک گزارش بالینی از اندازه گیری‌های خلاصه شده ECG مانند QT، مدت زمان QRS و غیره بادوام‌کمتر راچی، به عنوان سری زمانی ولتاژ خام ردیابی سیگنال ECG. به منظور تسهیل کاربرد گستردۀ تر، ما از الگوریتم‌های خلاصه شده ECG مانند مناسب برای فرمت‌های داده استفاده کردیم، یعنی پادگیری عمیق مبتنی بر ResNet برای سری زمانی ولتاژ چند کاناله غنی از اطلاعات و XGBoost مبتنی بر تقویت گردید. ECG گردانی برای اندازه گیری‌های پادگیری از زیراگهای ECG.

+ بهتر ترتیب زیر: (الف) فقط نوار قلب، (ب) نوار قلب + سن، جنس و (ج) نوار قلب
 سن، جنس + آزمایشگاه، که در آن داده‌های نوار قلب می‌توانند با نمودارهای ولتاژ-زمان یا اندازه گیری‌ها باشند.
 نوار قلب بیمارانی که قبل از نقاط زمانی ۳۰ روزه، ۱ ساله و ۵ ساله سانسور شده بودند، به دلیل نامشخص بودن وضعیت مرگشان از تجزیه و تحلیل حذف شدند و آنها که پس از نقاط زمانی سانسور شده بودند، به عنوان «زنده» در نظر گرفته شدند. برای آمزوش، نوار قلب بیمارانی که دچار حادثه شده بودند، صرف نظر از کامل بودن پیگیری، برای به حداقل رساندن پادگیری حفظ شد (به عنوان مثال، برای مرگ و میر ۵ ساله، نوار قلب بیمارانی را که پس از ۳۱-۰۳-۲۰۱۵ وارد گروه شدند اما قبل از ۳۱-۰۳-۲۰۲۰ فوت کردند، حفظ کردیم)، با این حال، برای ارزیابی، تمام نوار قلب‌هایی که پیگیری کامل نداشتند، صرف نظر از وضعیت مرگ یا سانسور آنها، حذف شدند (عنی، برای مرگ و میر ۵ ساله، تمام نوار قلب‌های پس از ۳۱-۰۳-۲۰۱۵ را در مجموعه ارزیابی حذف کردیم، زیرا آنها پنج سال پیگیری لازم را نداشتند). تعداد ECG‌ها، ابیزووها و بیماران مورد استفاده برای مدل سازی‌های نقطه زمانی در داده‌های کلی و در تقسیم بندی‌های تجربی در شکل ۱ ارائه شده است.^۱ و جدول تكميلي ۶.

الگوریتم‌های پادگیری

مازای یک روش طبقه‌بندی با برچسب‌های دودویی، یعنی مرده یا زنده به ترتیب طرف ۳۰ روز، ۳۶۵ روز (۱ سال) یا ۱۸۲۵ روز (۵ سال) از تاریخ اخذ ECG، برای تخمین احتمال زنده مادرن یک بیمار جدید حداقل ۳۰ روز، ۳۶۵ روز، ۱۸۲۵ روز پس از اخذ ECG استفاده کردیم. از آنجایی که ورودی مدل‌هایی که از اندازه گیری‌های ECG استفاده می‌کردند، داده‌های جدولی ساختاریافته بود، ما مدل‌هایی که گروههای XGB را آموزش دادیم، در حالی که برای مدل‌هایی با ردیابی سری‌های ولتاژ-زمان ECG از شبکه‌های عصبی کانولوشن عمیق استفاده کردیم.^{۱۹} برای هر دو مدل XGB و DL، ما از ۹۰٪ داده‌های توسعه برای آموزش مدل استفاده کردیم و ۱٪ باقی مانده را به عنوان یک مجموعه تنظیم برای ردیابی افت عملکرد و «توقف‌زودهنگام» فرآیند آموزش، برای کاهش احتمال پیش برازش (متفاوت از داده‌های ماندگار) استفاده کردیم.^{۲۰}

مدل XGB از تابع هدف لگاریتم درستنمایی استفاده کرد. پارامترهای مانند حداقل درخت، حداقل وزن فرزند و وزن مثبت مقایسه بر اساس اعتبارسنجی متقابل داخلی شبکه ای-لایه در مجموعه‌های آموزش تنظیم شدند. مدل‌ها حداکثر برای ۳۰۰ دوره متوازن داده شدند و اگر افت عملکرد در مجموعه تنظیم برای ۱۰ دوره متوازن کاهش نمی‌یافت، فرآیند آموزش متوقف می‌شد.

زیراگره‌های تجزیه و تحلیل آزمایشگاهی ما شامل ۴۱۳۷ نوار قلب از ۶۳۷ مورد از ۱۴۱ بیمار که داده‌های مربوط به هر شش آزمایشگاه موردنظر برای آنها در دسترس بود. به طور متوسط، ۴۱.۹٪ از نوار قلب‌ها می‌توان برای هر سال مالی از ۲۰۱۹ تا ۲۰۲۰ به آزمایشگاهی مرتبط کرد. برای نمودار جریان طراحی مطالعه برای مدل که شامل مقادیر آزمایشگاهی است و اندازه نمونه‌ها را برای زیرمطالعه کلی، تقسیم بندی‌های تجربی و پیامدهای مختلف نشان می‌دهد، به شکل تکمیلی ۱ مراجعه کنید.

وظایف پیش‌بینی

این مطالعه بر تسویه و ارزیابی مدل‌های مرگ و میر مبتنی بر ECG برای پیش‌بینی احتمال مرگ بیمار در عرض ۳۰ روز، ۱ سال و ۵ سال، از روز اخذ ECG، متمرکز بود. ECG های مورد استفاده در این مدل ها می‌توانستند در هر شرکه مراقبت‌های بهداشتی اخذ شده باشند (شکل ۱).^۱ هدف مدل‌های پیش‌بینی ارائه یک احتمال کالبیره شده از مرگ و میر است که می‌تواند در ارزیابی ریسک بیمار مورد استفاده قرار گیرد.

داده‌های ECG بیمار معمولاً توسط مراکز مراقبت‌های بهداشتی به یکی از دو فرمت زیر باگانی می‌شوند: (الف) به عنوان یک گزارش بالینی از اندازه گیری‌های خلاصه شده ECG مانند QT، مدت زمان QRS و غیره یا (ب) کمتر رایج است، به عنوان سری زمانی ولتاژ خام از ردیابی سیگنال ECG. به منظور تسهیل کاربرد گستردۀ تر، ما مدل‌های را برای تطبیق با هر یک‌جا این فرمت‌های ECG توسعه دادیم. ما از الگوریتم‌های پادگیری مبنای ResNet برای سری زمانی ولتاژ‌گذشتگی از اطلاعات و XGBoost استفاده کردیم (شکل ۱ را بینید).^۱ بنابراین، داده‌های ECG برای سری زمانی DL مبتنی بر تقویت گرادیان (XGB) برای اندازه گیری‌های ساختاریافته اش با «الگوریتم‌های پادگیری» در زیر. XGB به دلیل عملکرد پیشرفته اش با داده‌های جدولی ساختاریافته، زمان آموزش سریع، پشتیبانی از مقادیر گمشده و توابع قابل توضیح انتخاب شد.^{۲۱} به همین ترتیب، ما از معماری ResNet برای پادگیری عمیق (DL) بر اساس عملکرد موفق آن در مطالعات قبلی با مجموعه داده‌های قابل مقایسه از ردیابی‌های ECG استفاده کردیم.^{۲۲} در تحلیل تکمیلی، ما عملکرد مدل‌های استاندارد DL را در مقایسه با عماری DL طراحی شده سفارشی که توسط یک مطالعه قبلی (DNN) با بود، محک زدیم. این مطالعه از شبکه‌های عصبی کانولوشن عمیق (DNN) با استفاده از پنج شاخه برای تطبیق با مدت زمان‌های مختلف اخذ ECG در این مطالعه از لیدهای مختلف استفاده می‌کرد.^{۲۳}

به منظور بررسی ارزش پیش‌بینی افزایشی که داده‌های جمعیت شناختی (سن، جنس) و آزمایشگاهی به عملکرد مدل‌های آموزش دیده فقط بر اساس ECG اضافه می‌کنند، ما مدل‌های را در ... توسعه دادیم.

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 بیماران تحت بررسی خود بررسی کردیم. زیرگروه های تشخیصی بر اساس مسیوول ترین تشخیص تعیین شده در زمان ترخیص بوند و شامل مواد زیر بودند: افوارکتوس میوکارد بدون بالا رفتن قطعه (STEMI)، افوارکتوس میوکارد بالا رفتن قطعه (STEMII)، ST، نارسایی قلبی، فیبریلاسیون دهلیزی، دیابت شیرین و فشار خون بالا. کدهای ICD مورد استفاده برای شناسایی این شرایط در جدول تکمیلی ۷ ارائه شده است.

مقایسه مدل ها برای هر ارزیابی، ما از ۱۰ تکرار انتخاب تصادفی یک ECG واحد برای هر بیمار استفاده کردیم تا ثبات در عملکرد مدل را نشان دهیم. در طول آموختن این مدل ها، ما از تمام ECG های موجود در مجموعه آموختن یه همراه برچسب های مرگ و میر مربوطه استفاده کردیم. تقسیم بندی های آموختن و آزمایش یکسان (از جمله انتخاب های تصادفی) برای سازی های مختلف مدل سازی استفاده شد، به طوری که عملکرد من توانست به طور مستقیم مقایسه شود. نمرات عملکرد بین مدل ها با بوت استرپ کردن ۱۰۰ نمونه با نمونه گیری با جایگزینی از هر ۱۰ تکرار انتخاب تصادفی ECG که در بالا ذکر شد، مقایسه شدند تا در مجموع ۱۰۰۰ تکرار بوت استرپ ایجاد شود (شکل تکمیلی ۱۳). تفاوت در عملکرد مدل ها بر اساس همپوشانی فواصل اطمینان ۹۵٪ از میانگین نمرات AUROC مدل های مقایسه شده ارزیابی شد. ما همچنین مقدار p آزمون DeLong را گزارش کرده ایم.^{۲۴} برای نشان دادن اینکه آیا AUROC های دو مدل از نظر آماری تفاوت معنی داری دارندیما خیر.

تجسم هاما از نگاشت فعال سازی کلاس با وزن گرادیان (GradCAM) برای مدل های یادگیری عمیق ما نقش داشتند، استفاده کردیم.^{۲۵} برای دستیابی به این هدف، آخرين لایه کانولوشن که حاوی اطلاعات سطح بالای مدل عمیق CNN و ریاهای نماینده از مجموعه ارزیابی است، انتخاب شد. همچنین، ما از SHAP (توضیحات افزایشی) (SHAPley Additive exPlanations) برای شناسایی اندازه گیری های ECG که نقش کلیدی در پیش بینی میانگین مرگ و میر در مدل های XGB داشتند، استفاده کردیم.^{۲۶}

مطالعه ما بر اساس دستورالعمل های گزارش دهنده شفاف یک مدل پیش بینی چند متغیره برای پیش آگهی یا تشخیص فردی مبتنی بر هوش مصنوعی (AI-TRIPOD) گزارش شده است.^{۲۷}

نقش منبع تأمین مالی: این مطالعه توسط کمک هزینه عملیاتی از موسسه تحقیقات سلامت کانادا (کمک هزینه شماره ۱۷۸۱۵۸ PJT) 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 بهداشت آبرتا تحت شرایط یک توافق نامه تحقیقات ارایه شده است. سوالات مربوط به دسترسی به داده ها را می توان اسنادیاً آنها رسید. ما یک مجموعه داده ECG را که بصورت مجموعه ای از مذکورهای خودکار متغیر و صرفاً به منظور نمایش که تولید شده است. گنجانده ایم، انتظار نمی رو که آنها ECG و سیگنال های ECG و صرفاً به منظور نمایش که تولید شده 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>

دردسترس بودن کد

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

درباره: ۲۶ آگوست ۲۰۲۲؛ پذیرفته شده: ۲۱ ژانویه ۲۰۲۳
Published online: 06 February 2023

منابع

۱. مک گینیس، جن. ام.، فابنبرگ، اج. وی. و دزاویو، وی. جن.، پیشبرد نظام سلامت آموختن. مجله پژوهش ایسلند (N. Engl. J. Med.) ۳۷۸(۵-۶): ۳۸۵-۳۹۰.
۲. ازکوویتز، جن. ای. و همکاران. آیا در زنده ماندن از سندروم حاد کرونیک یا ابتلاء به نارسایی قلبی پس از آن، شکاف جنسیت وجود دارد؟ کردش خون ۱۳۹-۲۲۳۱، ۱۳۹-۲۲۳۹. (۲۰۲۰).

برای مدل DL، ما یک شبکه عصبی کانولوشن (CNN) مبتنی بر معماری شبکه عصبی با قیمانده پیاده سازی کردیم.^{۲۸} متشکل از یک لایه کانولوشن، ۴ بلوک با قیمانده با ۲ لایه کانولوشن در هر بلوک، و به دنبال آن یک لایه متردام (در مجموع 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 مقدماتی استفاده شده بود، برای شناسایی ناهنجاری ها در ECG های ۱۲ لید با برخی اصلاحات برای تطبیق ورویدی داده های جدولی و خروجی مرگ و میر بود (شکل تکمیلی ۱۲). هر نمونه ECG به صورت یک ماتریس عددی 12×4096 بارگذاری شد. اگر از ویژگی های آزمایشگاهی استفاده می شد، آنها به صورت ویژگی دودویی (جنسيت: ۱ و ویژگی: ۰) یا مقادیر پیوسته (سن، ویژگی های آزمایشگاهی: ۱ و ویژگی: ۰) وارد می شدند، سپس به یک لایه ۵N واحد پنهان (که در آن N تعداد ویژگی های جدولی است) منتقل می شدند، سپس بالایه متردام الحق می شدند و در نهایت برای تولید خروجی به یک تابع سیگموئید منتقل می شدند. از آنتروپی مقاطعه دودویی به عنوان تابع زبان با نرخ یادگیری اولیه 10×10^{-6} استفاده شد. بهینه ساز آدم، تابع فعال سازیReLU، اندازه هسته ۱۶، اندازه دسته ۵۱۲، و نرخ خروج ۰.۲ با سایر پارامترهای پیش فرض مدل ها برای حداکثر ۵ دوره (epoch) یادگیری شدند. نرخ یادگیری به ۱۰ کاهش یافت.^{۲۹} اگر هیچ بهبودی در تلفات تنظیم برای دوره دو همتوالی وجود نداشت، و اگر تلفات در مجموعه تنظیم برای دوره دو کاهش نیافت، فرآیند یادگیری متوقف شد. مدل ها با استفاده از ۲.۲ XGBoost ۱.۵.۱ و Tensorflow ۴۱۸.۸۸ نسخه NVIDIA با ۳۲ گیگابایت رم برای هر پردازنده گرافیکی Tesla V100-SXM2 دادیم. آموختن هر مدل DL تقریباً ۳۰ دقیقه در هر دوره طول کشید.

ارزیابی و تجسم معیارهای ارزیابی. ما مجموعه داده های کلی ECG را به یک تقسیم تصادفی ۶۴٪ برای توسعه مدل (برای آموختن ۹۰٪ زیرمجموعه از ۶۰٪) و تنظیم دقیق (۱۰٪ زیرمجموعه از ۶۰٪) مدل (۴۰٪ باقی مانده را به عنوان مجموعه داده های پشتیبان برای اعتبارسنجی تقسیم کردیم (شکل ۱) و جدول تکمیلی ۶ برای مدل هایی بدون آزمایشگاه، و شکل تکمیلی ۱ برای مدل ثانویه با آزمایشگاه. مطالعه این حاصل کردیم که ECG های یک بیمار بین مجموعه داده های توسعه ارزیابی به اشتراک گذاشته نشده باشند. ما معیارهایی داریم که با نام شاخص C نیز شناخته می شود) و مساحت زیر منحنی دقت-باداوزی (AUROC)، که با نام شاخص C مساحت زیر منحنی مشخصه پیش بینی را با استفاده از نقاط برگزینه شده از مجموعه داده های پیش بینی بهینه متنطبق شده از مجموعه آموختن، به کلاس های مرده/زنده تقسیم کردیم.^{۲۰} امتحانات F1، ویژگی، فراخوانی، و صحت را تولید کردیم، علاوه بر این، کالیبراسیون مدل های خود را ارزیابی کردیم تا بینیمیم آیا احتمالات پیش بینی شده با نسبت های مشاهده شده با استفاده از امتیاز اختیاری مطابقت دارند یا خیر (دقیقای ۰.۲۵٪ است؛ امتیاز کمترین شان دهنده کالیبراسیون بهتر است).^{۲۰}

نمونه گیری ارزیابی. انتظار می روی بیمارانی که بیماری های شدیدتری دارند، بیشتر تحت نارق قلب قرار گیرند و در مراحل پایانی بیماری، وقتی بیماران به طور منظم تری تحت نظر باشند، ممکن است تعداد بیشتر نارق قلب گرفته شود. این تنوع در زمان و تعداد نارق قلب در بین بیماران می تواند منجر به سوگیری بالقوه در عملکرد مدل به دلیل نمایش متفاوت فوتوبی های بیمارشون. برای کاهش چنین سوگیری، ما مدل های خود را با استفاده از یک نارق قلب تصادفی انتخاب شده برای هر بیمار از چندین دوره بیماری در مجموعه بیماران مبتلا، ارزیابی کردیم (شکل ۱). 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، بینش 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 ECG کردیم.

۳. ل.، دی اس و همکاران. پیش بینی کنده های مرگ و میر در میان ساکنان مراقبت های طولانی مدت مبتلا به عفونت-2-CoV-19. مجله پژوهشی سالموندان. جامعه شناسی. ۹۶، ۳۳۷۷-۳۳۸۸. (۰۱-۰۲).
۴. استپیلر، ج.ا. و همکاران. اعتبارسنجی خارجی نمرات اصلاح شده LACE+ و LACE-. برای پیش بینی مستری مجدد با مرگ پس از ترجیح از بیمارستان. ج.ا. ایول. کلینیک. پرکسیس. (۰۱-۰۲). ۱۳۹۷-۱۳۹۰.
۵. ساراک، ب. و همکاران. آزمایش لبپید، درمان اصلاح کننده لبپید و واحد شرایط بودن مهارتمند PCSK9 (پروپروتئین کاتورتاز سوتیلیسین-کسین نوع ۹) در 27979 بیمار مبتلا به سندروم کرونری دور قلب و عروق. کیفیت نتایج. (۰۱-۰۲). ۲۰۲۱. e006646.
۶. اوپلاندی، م. و همکاران. معرفی داروهای ضد اینقاد خوارکی مستقیم، شکاف های درمانی را برای بیماران ضعیف مبتلا به فیبریلاسیون دهلیزی غیر دریچه ای برطرف نکرده است. می توان. مجله کاردیولوژی. (۰۱-۰۲). ۸۷-۸۶.
۷. مک آیسترن، اف.ا. و همکاران. استانی ها و عقوبات SARS-CoV-2: نتایج یک مطالعه کوهوورت آینده نگرفته بر روی 469749 بزرگسال از ۲ استان کانادا. انجمن قلب آمریکا. (۰۱-۰۲). ۰۰۲۳۰.
۸. ساندوو، ار.ک. و همکاران. استفاده همزمان از داروهای P-گلکوپروتئین با سیتوکروم ۳A4 و داروهای ضد اینقاد خوارکی آنتاکونوپیست غیر ویتابین K در فیبریلاسیون دهلیزی غیر دریچه ای. پورول. قلب. مجله کیفیت. مراقبت. پالین. نتایج ۱۹۵-۲۰۱. (۰۱-۰۲).
۹. دیاتش ملکی، ن.، احتشامی افتخاری، ا. و ارمستانگ، پ. پرکسیس. استفاده از شاخص های الکتروکاردیوگرام اسکمی میوکارد برای طبقه بندی خطر و تضمیم گیری در مورد استراتژی های خون رسانی مجدد. ج.ا. الکتروکاردیوگرام. پورول. کاردیوگرام. ج.ا. (۰۱-۰۲). ۵۰-۴۵.
۱۰. آقیان، ز.، هارمون، د.، ام.، یهود، ا.، و فیدمن، پ.، ای. کاربرد هوش مصنوعی در الکتروکاردیوگرام. پورول. هارت. ج.ا. ۴۷۸۷-۴۷۸۶. (۰۱-۰۲).
۱۱. مینچوله، آ.، کمپس، ج.، لیون، آ.، و رودریگز، ب.، یادگیری ماشین در الکتروکاردیوگرام. ج.ا. الکتروکاردیوگرام. (۰۱-۰۲). ۶۱-۶۲.
۱۲. سان، دبلیو. و همکاران. نوار قلب برای غربالگری با توان عملیاتی بالا برای بیماری های متعدد: اثبات مفهومیها استفاده از یادگیری عمیق چند تنسیضی از مجموعه داده های مبتلی به جمعیت. تصویربرداری پرکسیس کیا با شود. (۰۱-۰۲). <https://doi.org/10.48550/arXiv.2210.06291>.
۱۳. کالیف، آ.، ام.، مزایای انتقال کیفیت به سطح مل. ام.، هارت. ج.ا. ۱۵۶-۱۴۹. (۰۱-۰۲).
۱۴. راکونات، س.، و همکاران. پیش بینی مرگ و میر از داده های ولتاز الکتروکاردیوگرام ۱۷ لید با استفاده از یک شبکه عصبی عمیق. نات. مد. ۱۶۶-۱۶۵. (۰۱-۰۲).
۱۵. دکتر سامسکی و همکاران. روند بسترهای مجدد و مدت اقامت بیماران بسترهای شده به دلیل نارسایی قلبی در کانادا و ایالات متحده. جما. کاردیولوژی. (۰۱-۰۲). ۴۳-۴۴.
۱۶. کائیول، ب.، و همکاران. تفاوت در درمان، پیامدها و کیفیت زندگی در میان بیماران مبتلا به نارسایی قلبی در کانادا و ایالات متحده. نارسایی قلبی جما. (۰۱-۰۲). ۵۳-۵۲.
۱۷. کائیول، ب.، 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 استریتیوکنیاز و t-PA برای شریان های کرونری مسدود شده (GUSTO) I- ثبت نام کرده اند. گردش خون. (۰۱-۰۲). ۱۷۵۴-۱۷۵۶.
۱۸. مراقبت های بهداشتی حرfe ای فیلیپس. ECG. به صورت آنلاین. IntelliSpace ECG. در آردرس زیر موجود است: <https://www.usa.philips.com/healthcare/product/HC860426/intellispace-ecg-ecg-management-system>. (۰۱-۰۲).
۱۹. جن، ت.، و گسترن، س.، ایکس جی بوسیت: یک سیستم تقویت درختن مقیاس بذر. در مجموعه مقالات بیست و دوین کنفرانس بین المللی ACM SIGKDD در مورد کشف داشش و داده کاوی (صفحات ۷۸۵-۷۸۷). پیوپورک، نیوپورک، ایالات متحده آمریکا. (۰۱-۰۲).
۲۰. ریبیکو، ای.ج. و همکاران. تشخیص خودکار نوار قلب ۱۲ لیدی با استفاده از یک شبکه عصبی عمیق. نات. کمون. (۰۱-۰۲). ۱۷۶-۱۷۵.
۲۱. پرشلت، ل.، توقف زودهنگام - اما چه زمانی؟ در شبکه های عصبی: ترقندهای تجارت (ویراستاران. اور. ج.ن.، و مولر. ج.ا.). (۰۱-۰۲). ۵۵-۶۹.
۲۲. ه.، ک.، ژانک، ایکس.، رن، اس.، و سان، جن. یادگیری عمیق باقیمانده برای تشخیص تصویر. در مجموعه مقالات کنفرانس IEEE در مورد بینایی کامپیوتر و تشخیص الگو (CVPR) (۰۱-۰۲). ۷۷۸.
۲۳. کینگما، دی بی و با، جن. آدام: روشن برای بهینه سازی تصادفی. به عنوان مقاله کنفرانسی در سومین کنفرانس بین المللی بازنمایی های یادگیری مجدد شد. سن دیگو. (۰۱-۰۲). <https://doi.org/10.48550/arXiv.1412.6980>.
۲۴. بیودن، شاخص W7 برای رتبه بندی آزمون های تشخیصی. سلطان. ۳۵-۳۲. (۰۱-۰۲).
۲۵. بیری، ج.، دبلیو.، تأیید پیش بینی های بیان شده بر حسب احتمال. آب و هواي دوشنبه. (۰۱-۰۲). ۱۷۸-۱۷۳.

© نویسنده(گان) ۲۰۲۳