

توسعه و اعتبارسنجی الگوریتم های یادگیری ماشین مبتنی بر الکتروکاردیوگرام برای تشخیص بیماری های قلبی عروقی در سطح جمعیت

بررسی به روزرسانی ها

سونیل واسو کالمادی^{۱,۲,۳,۴}، امیر سلیمی^{۱,۶}، ویجی سان^۱، نریمان سپهروند^{۲,۴}، یوسف نادمی^۱، کوین باینی^{۲,۳}، جاستین ازکوویتز^{۲,۳}، آبرام هیندل^۱، فینلی مک آلیستر^{۲,۳}، راسل گرینر^۱، رویندر ساندوه^{۲,۳} و پادما کائول^{۲,۳}

الگوریتم های الکتروکاردیوگرام (ECG) مبتنی بر هوش مصنوعی، برای تشخیص زودهنگام بیماری های قلبی عروقی (CV)، از جمله مواردی که به طور سنتی با اندازه گیری های مرسوم ECG یا تفسیر تخصصی مرتبط نیستند، اهمیت پیدا می کنند. این مطالعه چنین مدل هایی را برای پیش بینی همزمان ۱۵ تشخیص رایج مختلف CV در سطح جمعیت توسعه داده و اعتبارسنجی می کند. ما یک مطالعه گذشته نگر انجام دادیم که شامل ۱,۶۰۵,۲۶۸ نوار قلب (ECG) از ۲۴۴,۰۷۷ بیمار بزرگسال مراجعه کننده به ۸۴ بخش اورژانس یا بیمارستان بود که حداقل یک نوار قلب ۱۲ لیدی از فوریه ۲۰۰۷ تا آوریل ۲۰۲۰ در آلبرتا، کانادا انجام داده بودند و ۱۵ تشخیص قلبی عروقی را که توسط کدهای طبقه بندی بین المللی بیماری ها، ویرایش دهم (ICD-10) شناسایی شده بودند، در نظر گرفتیم: فیبریلاسیون دهلیزی (AF)، تاکی کاردی فوق بطنی (SVT)، تاکی کاردی بطنی (VT)، ایست قلبی (CA)، بلوک دهلیزی-بطنی (AVB)، آنژین ناپایدار (UA)، انفارکتوس میوکارد با بالا رفتن قطعه (STEMI) (ST، غیر (NSTEMI) (STEMI، آمبولی ریوی (PE)، کاردیومیوپاتی هیپرتروفیک (HCM)، تنگی آئورت (AS)، پرولاپس دریچه میترال (MVP)، تنگی دریچه میترال (MS)، ریوی فشار خون بالا (PHTN) و نارسایی قلبی (HF). ما از یادگیری عمیق (DL) مبتنی بر ResNet با استفاده از ردیابی ECG و تقویت گرادیان شدید (XGB) با استفاده از اندازه گیری های ECG استفاده کردیم. هنگامی که اولین ECG ها در هر قسمت از ۹۷۶۳۱ بیمار مبتلا به نارسایی قلبی ارزیابی شدند، مدل های DL دارای مساحت زیر منحنی مشخصه عملکرد گیرنده (AUROC) کمتر از ۸۰٪ برای ۳ بیماری قلبی عروقی (PTE، SVT، UA)، ۸۰-۹۰٪ برای ۸ بیماری قلبی عروقی (CA، NSTEMI، VT، MVP، PHTN، AS، AF، HF) و ۹۰٪ AUROC برای ۴ تشخیص (AVB، HCM، MS، STEMI) بودند. مدل های DL با حدود ۵٪ AUROC بالاتر به طور متوسط، از مدل های XGB بهتر عمل کردند. به طور کلی، مدل های پیش بینی مبتنی بر ECG عملکرد پیش بینی خوب تا عالی را در تشخیص بیماری های شایع قلبی عروقی نشان دادند.

از فعالیت الکتریکی قلب تولید می شود. با این حال، تکنیک های استاندارد مورد استفاده پزشکان و الگوریتم های کامپیوتری برای تفسیر ECG محدود هستند، زیرا بسیاری از آنها مبتنی بر قانون هستند و تنها بخشی از کل اطلاعات موجود در ECG را در نظر می گیرند. رویکردهای دستی یا کامپیوتری و حتی روش های آماری مرسوم نمی توانند تعاملات سطح بالا بین سیگنال های ECG از چندین لید یا ... را در نظر بگیرند.

الکتروکاردیوگرام (ECG) دوازده لید، رایج ترین، کم هزینه ترین و در دسترس ترین ابزار تشخیصی برای بیماری های قلبی عروقی (CV) است. این روش تقریباً در تمام ویژگی های مراقبت های حاد و معمولاً پیش از یک بار انجام می شود. تنها در ایالات متحده، سالانه بیش از ۱۰۰ میلیون نوار قلب گرفته می شود. این مفید است زیرا ECG حاوی مقدار زیادی اطلاعات است که بینشی در مورد فیزیولوژی قلبی زمینه ای ارائه می دهد، زیرا ویژگی های مورفولوژیکی و زمانی ...

اگرچه علوم کامپیوتر، دانشگاه آلبرتا، ادمونتون، آلبرتا، کانادا؛ مرکز VIGOUR کانادا، دانشکده پزشکی، دانشگاه آلبرتا، ادمونتون، آلبرتا، کانادا؛ گروه پزشکی، دانشگاه آلبرتا، ادمونتون، آلبرتا، کانادا؛ گروه پزشکی، دانشگاه کلگری، کلگری، آلبرتا، کانادا؛ موسسه قلب اسمیت، سیستم بیمارستانی مرکز پزشکی سدرز-سینای، لس آنجلس، کالیفرنیا، ایالات متحده آمریکا؛ این نویسندگان به طور مساوی کمک کردند: سونیل واسو کالمادی، امیر سلیمی. ایمیل: pkaul@ualberta.ca



تغییرات نامحسوس، اما آموزنده ای که ممکن است نشان دهنده بیماری اولیه باشند. ظهور تحلیل های یادگیری عمیق (DL) فرصتی هیجان انگیز برای شناسایی الگوهای بالینی مرتبط اما «پنهان» در سیگنال های ECG و ارزیابی همزمان روابط تعاملی پیچیده از داده های بالینی ثبت شده معمول برای تشخیص ناهنجاری های مختلف قلبی عروقی ارائه می دهد.^{۵-۲}

تحقیقات قبلی که مدل های یادگیری ماشینی را برای پیش بینی بیماری با استفاده از ECG بررسی می کردند، عمدتاً بر بیماری های قلبی متمرکز بودند که می توانند به راحتی توسط متخصصان پزشکی بر اساس تغییرات مورفولوژیکی در الگوهای ECG تفسیر شوند - مانند آریتمی ها (فیبریلاسیون دهلیزی [AF]، تاکی کاردی بطنی [VT]، تاکی کاردی فوق بطنی [SVT])، انفارکتوس میوکارد با بالا رفتن قطعه (STEMI) (ST یا بدون) (NSTEMI) یا شرایط بلوک قلبی مانند بلوک های دهلیزی-بطنی یا بلوک های شاخه ای از جمله بلوک شاخه چپ و بلوک شاخه راست. اگرچه تعداد مدل های مبتنی بر یادگیری ماشین (ML) که از داده های ECG برای پیش بینی بیماری های قلبی عروقی فراتر از مواردی که به طور سنتی با الگوهای ECG مرتبط هستند، استفاده می کنند، در حال حاضر محدود است، اما به طور پیوسته در حال افزایش است. این مدل ها بر بیماری هایی مانند پرولاپس دریچه میترال (MVP) تمرکز دارند. ایست قلبی (CA) نارسایی قلبی (HF) ۱۲.۱۳، امبولی ریوی (PE) ۱۶.۱۵، تنگی آئورت (AS) ۱۹-۱۷، تنگی دریچه میترال (MS) ۲۲.۲۱، فشار خون ریوی (PHTN) ۲۲.۲۱ و کاردیومیوپاتی هیپرتروفیک (HCM) ۲۲.۲۱، علاوه بر این، در حالی که مطالعات موجود عمدتاً بر روی برچسب های فردی متمرکز بوده اند، هیچ تحقیق قبلی در زمینه توسعه یک سیستم پیش بینی کننده برای تشخیص همزمان این شرایط خاص انجام نشده است. فقدان مجموعه داده های پزشکی بزرگ که از نظر بالینی با مجموعه گسترده ای از برچسب های تشخیصی موجود برای یادگیری ماشینی تحت نظارت، حاشیه نویسی شده باشند، یک مشکل شناخته شده است و اعتبارسنجی های در مقیاس بزرگ در مقیاس جمعیت برای نشان دادن قابلیت اعتماد برای پذیرش موفقیت آمیز مدل های پیش بینی در عمل بالینی، که در آن شناسایی و درمان زودهنگام ممکن است به طور بالقوه بر عوارض مرتبط با بیماری، استفاده از مراقبت های بهداشتی و هزینه تأثیر بگذارد، بسیار مهم است.

براین اساس، ما از یک گروه بزرگ از بیماران در سطح جمعیت، از یک سیستم بهداشت جهانی تک پرداخت کننده، برای توسعه و اعتبارسنجی مدل های DL (مبتنی بر ردیابی های ECG دوازده گانه) و همچنین مدل های تقویت گرادیان شدید (XGB) (مبتنی بر اندازه گیری های ECG که به طور معمول جمع آوری می شوند) استفاده کردیم تا به طور همزمان ۱۵ تشخیص رایج CV را از طریق یک چارچوب پیش بینی یکپارچه پیش بینی کنیم.

نتایج

ویژگی ها و پیامدهای بیمار ویژگی های پایه گروه قبلاً شرح داده شده است.^{۲۵} به طور خلاصه، میانگین سنی بیماران ۶۵.۸ ± ۱۷.۳ سال بود و ۵۶.۷% آنها زن بودند (جدول تکمیلی ۱). مدل ها با استفاده از نوار قلب ۱۴۶۴۴۶ بیمار تحت آموزش قرار گرفتند و متعاقباً روی گروهی از افراد که از مطالعه حذف شده بودند، ارزیابی شدند. ۹۷،۶۳۱ بیمار (شکل ۱...۱)، مجموعه داده های مربوط به افراد تحت آزمایش شامل ۵۳۴۳۶ مرد و ۴۴۱۹۵ زن، که برای ارزیابی عملکرد مبتنی بر جنسیت استفاده شدند. علاوه بر این، ۹۶۱۶۴ بیمار بدون ضربان ساز قلب به طور جداگانه ارزیابی شدند تا تأثیر ضربان سازها بر عملکرد مدل بررسی شود. با پیش بینی اجرای سیستم پیش بینی ما در محل مراقبت، مدل های خود را منحصراً با استفاده از اولین نوار قلب هر بیمار بدون ضربان ساز در یک دوره خاص ارزیابی کردیم.

فراوانی و درصد ECG ها با هر یک از شرایط CV انتخاب شده به صورت کامل، توسعه یافته و تقسیم بندی های holdout و همچنین در بین اولین ECG در هر قسمت در مجموعه holdout در جدول ارائه شده است. اولین نوار قلب به ازای هر اپیزود در مجموعه ی داده های حذف شده که برای ارزیابی نهایی استفاده شد) در مقایسه با داده های کامل نوار قلب، تفاوت هایی در برچسب های تشخیصی داشت (مثلاً فراوانی برای نارسایی قلبی: ۹.۳% در مقابل ۱۵.۵% ؛ فیبریلاسیون دهلیزی: ۱۱.۵% در مقابل ۱۸.۲%).

عملکرد مدل ها و مقایسه آنها مقایسه عملکرد مدل برای مدل های DL و XGB با ردیابی های ECG (با در مقابل بدون ویژگی های سن و جنس) و اندازه گیری ها (با ویژگی های سن و جنس) برای شرایط 15 CV در شکل 1 ارائه شده است. ۲. جدول ۲ و جدول تکمیلی ۲. اعتبارسنجی مدل اصلی ما (DL: ردیابی ECG، سن، جنس) نشان داد که مدل ما برای بهترین عملکرد در آن منحنی مشخصه عملکرد گیرنده (AUROC) 95.5٪ داشت و مدل ما برای ترومبوآمبولی ریوی (PTE) بدترین عملکرد را داشت.

عملکرد با AUROC برابر با 68.9٪، مدل ها برای همه تشخیص ها، به جز AUROC، PTE، بالای 76٪ داشتند؛ با AUROC کمتر از 80٪ برای دو تشخیص (SVT، UA)، به ترتیب افزایشی؛ AUROC در محدوده 80-90٪ برای هشت تشخیص (ایست قلبی [CA]، VT، NSTEMI، پرولاپس دریچه میترال [MVP]، فشارخون ریوی [PHTN]، تنگی آئورت [AS]، HF، AF، به ترتیب افزایشی)؛ $90\% < \text{AUROC}$ برای چهار تشخیص (بلوک دهلیزی-بطنی [AVB]، کاردیومیوپاتی هیپرتروفیک [HCM]، تنگی دریچه میترال [MS]، STEMI، به ترتیب افزایشی). مدل مربوط به فیبریلاسیون دهلیزی (AF) بالاترین سطح زیر منحنی دقت-یادآوری (AUPRC) با امتیاز ۵۹.۲٪ (امتیاز $F1$: ۵۱.۶٪) را داشت و پس از آن نارسایی قلبی (HF) با (امتیاز $F1$: ۴۶.۶٪) و STEMI با ۵۴.۳٪ AUPRC (امتیاز $F1$: ۳۹.۲٪) قرار داشتند.

مدل DL با (ردیابی ECG، سن، جنس) برای اکثر تشخیص ها، به جز AVB، که هر دو مدل عملکرد قابل مقایسه ای داشتند، عملکرد بهتری نسبت به مدل XGB با (اندازه گیری های ECG، سن، جنس) داشت. مدل های DL با بهبود متوسط 5.2 درصدی در AUROC، با افزایش قابل توجه 11.8 درصدی برای MS، 8.6 درصدی برای MVP، درصدی برای NSTEMI و 7.1 درصدی برای STEMI، عملکرد بهتری نسبت به مدل های XGB داشتند. مقایسه فواصل اطمینان 95 درصدی از نتایج بوت استرپ نشان داد که تفاوت های معناداری بین عملکرد مدل DL و ویژگی های (سن، جنس) در مقابل بدون ویژگی ها برای همه تشخیص ها به جز PTE وجود دارد، که نشان می دهد ویژگی های سن و جنس می توانند پیشرفت های کوچک اما قابل توجهی را به پیش بینی تشخیصی اضافه کنند. به طور مشابه، نتایج بوت استرپ نشان داد که مدل های DL با ردیابی ECG به تنهایی، برای تشخیص هایی غیر از AVB، VT و HCM، عملکرد بهتری نسبت به مدل های XGB با اندازه گیری های ECG، سن و جنس دارند.

عملکرد مدل مبتنی بر جنسیت

ما مدل DL را که با استفاده از (ردیابی ECG، سن، جنس) به طور جداگانه برای افراد فورمالیست و مونث در مجموعه مورد نظر آموزش دیده بود، ارزیابی کردیم و به طور کلی نتایج مشابهی یافتیم (شکل ۱).^۳ (تصویر بالا) مدل ها در 10 مورد از 15 مورد، عملکرد نسبتاً بهتری برای مردان داشتند و میانگین افزایش AUROC آنها 1.0٪ بود. پنج مورد از این موارد - یعنی HF، PTE، STEMI، VT و AS - تفاوت های معنی داری را نشان دادند. ما بیشترین تفاوت را در VT یافتیم که در آن مدل برای مردان 6.4٪ در مقایسه با زنان (مردان: 85.1٪، زنان: 78.7٪) و از نظر AUPRC (مردان: 37.5٪، زنان: 22.6٪) 14.8٪ عملکرد بهتری داشت. در مقابل، عملکرد پیش بینی برای فیبریلاسیون دهلیزی (AF) در زنان به طور قابل توجهی 1.2٪ AUROC بهتر از مردان بود.

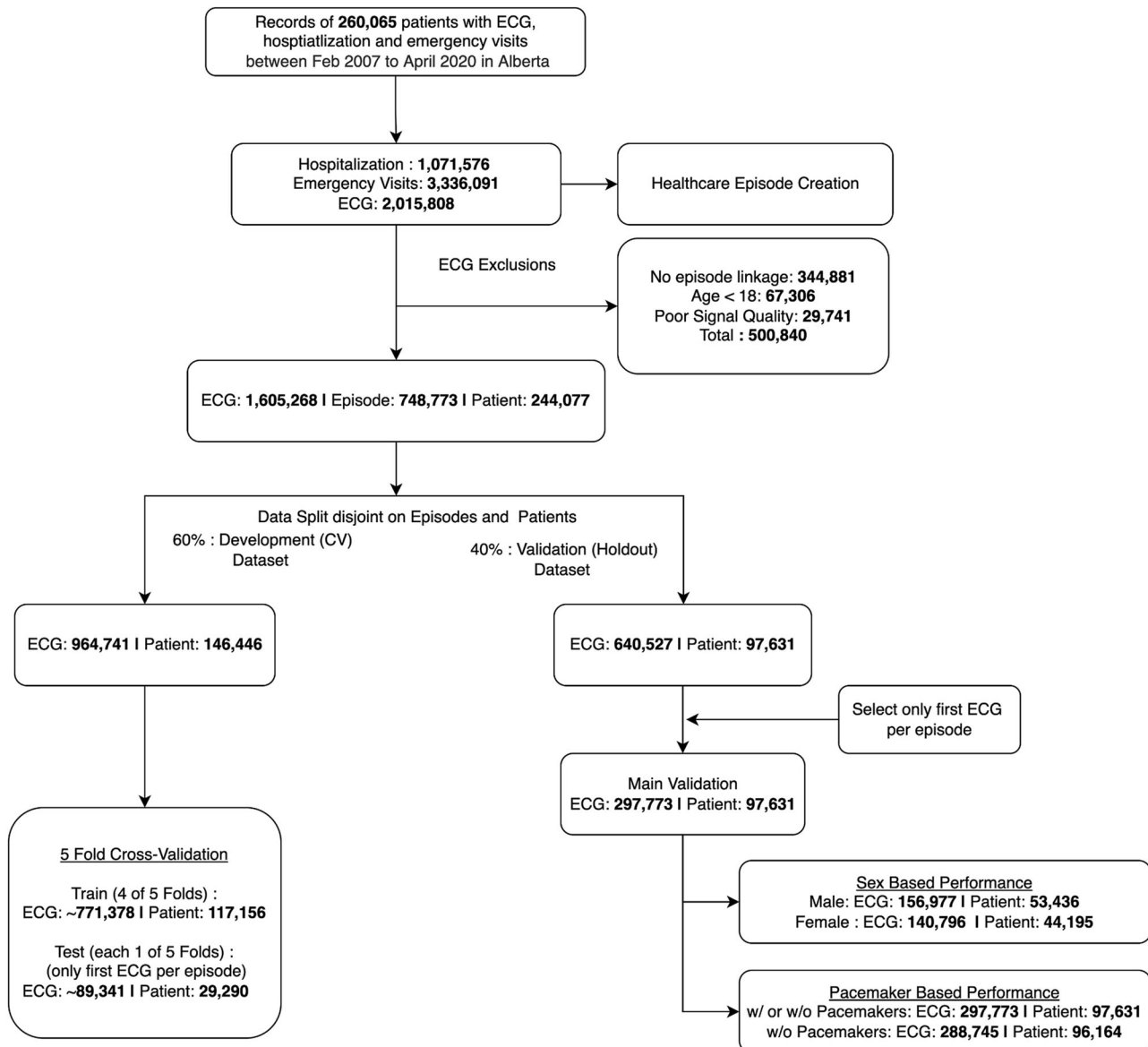
عملکرد مدل مبتنی بر حضور ضربان ساز به طور مشابه، ارزیابی مدل های DL (ردیابی ECG، سن، جنس) روی ECG های غیرفعال، پس از حذف ECG های بیماران دارای ضربان ساز و سایر دستگاه های ICD، عملکردی را نشان داد که با ارزیابی کلی قابل مقایسه است، با افزایش متوسط بسیار کم AUROC به میزان 0.25٪ با ECG های غیرفعال در مقایسه با بدون آن (شکل ۱).^۳ پل پایین). باز هم، VT بیشترین تفاوت را نشان داد، که در آن عملکرد با حذف ECG های ضربان ساز، 3.2٪ AUROC و 5.6٪ AUPRC کاهش یافت. تشخیص دیگری که تفاوت قابل توجهی را در همان جهت نشان داد، AVB بود (1.6٪ افت 5.7٪ AUROC، کاهش در AUPRC).

اعتبارسنجی خروج از بیمارستان با یک بار بستری

مطالعه ما از ECG های تهیه شده از ۱۴ بیمارستان استفاده کرد. نکته قابل توجه این است که دو مورد از این بیمارستان ها، بیمارستان های مراقبت های عالی بودند که بالاترین تعداد ECG (به ترتیب ۴۸۷،۰۴۲ و ۴۵۳،۰۸۵ ECG) را داشتند. برای هر بیمارستان عالی (H1 و H2)، ما یک اعتبارسنجی خروج از بیمارستان با آموزش بر روی ECG های سایر بیمارستان ها، به استثنای ECG اعتبارسنجی (شکل تکمیلی ۱) انجام دادیم. عملکرد مدل ECG: DL، سن، جنس با اعتبارسنجی خروج از بیمارستان با نتایج گزارش شده در مجموعه کلی موارد عدم پذیرش (جدول تکمیلی ۳) قابل مقایسه بود. در مقایسه با نتایج اعتبارسنجی اولیه، میانگین عملکرد AUROC در ۱۵ شرایط، افزایش اندک ۱.۳۸ درصدی در اعتبارسنجی H1 اما کاهش ۱.۳۴ درصدی در اعتبارسنجی H2 را نشان داد.

ارزیابی تمام holdout ECG

به عنوان یک اعتبارسنجی اضافی از مدل DL خود، ما عملکرد آن را در تمام ECG ها، به جای فقط اولین ECG، که در طول هر دوره مراقبت از بیماران در مجموعه ی مورد انتظار به دست آمده بود، ارزیابی کردیم. نتایج، همانطور که در بخش تکمیلی ذکر شده است



هرایزود ثبت شده در طول ویزیت اورژانس یا بستری شدن در بیمارستان، که نشان دهنده ی استقرار موردنظر در محل مراقبت است؛ دوم، روی تمام ECGهای از مجموعه ی مورد انتظار. علاوه بر این، ما عملکردمدل های خود را در زیرگروه های خاص بیمار که بر اساس جنسیت و وجود دستگاه های ضربان ساز قلبی یا کمکی بطنی طبقه بندی شده بودند، ارزیابی کردیم.

شکل ۱ | فلوچارت طرح مطالعه که حجم نمونه ها را برای تقسیم بندی های مختلف آزمایش نشان می دهد. ما کل مجموعه داده های ECG را تقسیم کردیم، ۶۰٪ را برای توسعه مدل (شامل اعتبارسنجی متقابل داخلی پنج گانه برای آموزش و تنظیم دقیق) اختصاص دادیم و ۴۰٪ را به عنوان مجموعه ی پشتیبان برای اعتبارسنجی نهایی کنار گذاشتیم. برای ارزیابی، مدل های خود را با دو رویکرد ارزیابی کردیم: اول، منحصرأ روی اولین ECGها از

بنابراین، یک طرح ارزیابی جایگزین مبتنی بر رویکرد برچسب گذاری ترکیبی که قبلاً برای اهداف غربالگری به کار گرفته شده بود، بررسی شد تا بازده تشخیصی افزایش یابد. ما یک برچسب ترکیبی ایجاد کردیم به طوری که اگر هر یک از ۱۵ شرط مورد نظر ما مثبت باشد، مثبت و اگر همه شرایط منفی باشند، منفی خواهد بود.

ماتوانایی مدل DL چند برچسبی خود را برای پیش بینی مثبت بودن نوار قلب (ECG) برای برچسب ترکیبی، دوباره ارزیابی کردیم. نتایج، PPV معادل 31.64٪ با امتیاز F1 معادل 47.39٪ را نشان داد. ما همچنین یک مدل جدید تحت نظارت با برچسب ترکیبی را با استفاده از معماری مدل مشابه آموزش دادیم و عملکرد آن را در همان مجموعه داده های محدود ارزیابی کردیم. نتایج، PPV معادل 57.9٪ با امتیاز F1 معادل 63.03٪ را نشان داد. این نتایج نشان می دهد که هنگام غربالگری پیامد ترکیبی، می توان PPV بالاتری را به دست آورد.

توضیحات مدل

شکل ۴ نتایج GradCAM را نشان می دهد و حوزه هایی از ECG را که سهم و ارتباط بیشتری با پیش بینی مدل از شرایط مختلف CV دارند، برجسته می کند (برای فهرست کامل همه موارد، به شکل تکمیلی ۲ مراجعه کنید).

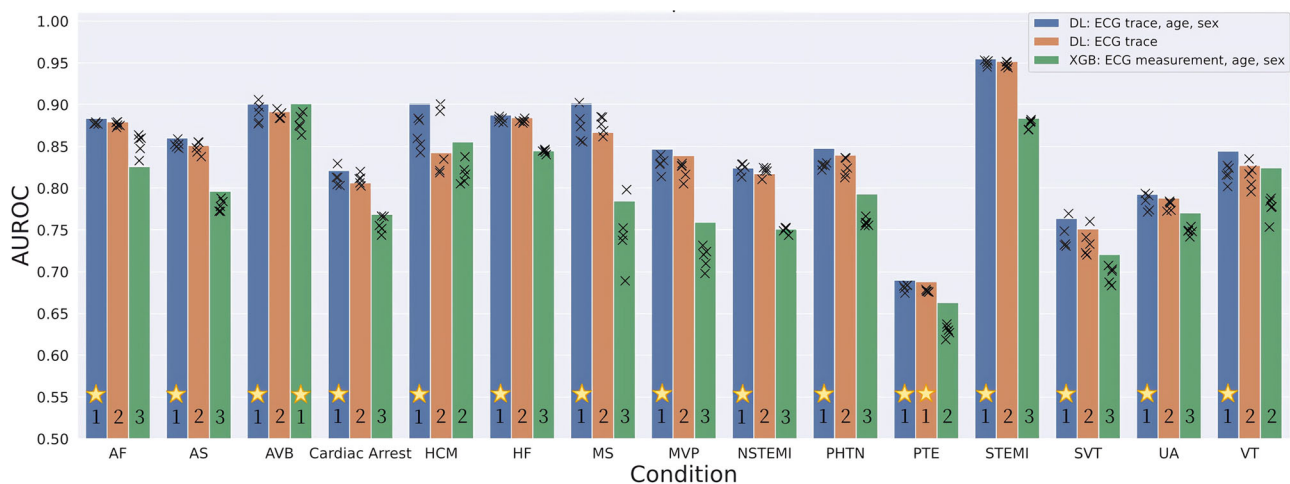
جدول ۴، عملکردی را نشان می دهد که یا برتر یا قابل مقایسه با عملکرد حاصل از استفاده صرف از ECGهای اولیه است. ارزیابی تمام ECG، نوساناتی را در نمرات AUROC و AUPRC نشان داد. کاهش کلی 2.04 درصدی در میانگین AUROC (میانگین در 15 شرایط) و افزایش همزمان 2.15 درصدی در AUPRC برای ارزیابی تمام ECG وجود داشت. نکته قابل توجه این است که نمرات F1 برای همه برچسب ها در ارزیابی تمام ECG، بهبودهایی از 1.09٪ تا 16.81٪، با افزایش متوسط 6.28٪ را نشان داد. به طور مشابه، مقادیر پیش بینی مثبت (PPV یا دقت) برای همه برچسب ها افزایشی از 0.56٪ تا 15.32٪، با بهبود متوسط 4.98٪ را نشان داد. بنابراین، می توان انتظار داشت که این الگوریتم ها در صورت اعمال بر روی ECGهای انجام شده در هر نقطه ای از دوره مراقبت، عملکرد قابل مقایسه، اگر نگوییم برتر، را نشان دهند.

ارزیابی برچسب کامپوزیت

شیوع چندین مورد از تشخیص های مورد نظر در نمونه ما پایین بود (مثلاً 0.09٪ برای MS در بین اولین ECGها)، که احتمالاً بر PPV تأثیر می گذارد. ما،

جدول ۱ | فراوانی و درصد ECG های دارای بیماری های قلبی عروقی منتخب در گروه های مورد استفاده در مطالعه

اطلاعات کامل (ن = ۱,۶۰۵,۲۶۸)	مجموعه توسعه (ن = ۹۶۴,۷۴۱)	مجموعه نگهدارنده (ن = ۶۴۰,۵۲۷)	اولین نوار قلب در هر قسمت در مجموعه holdout (ن = ۳۹۷,۷۷۳)
انفارکتوس میوکارد بدون بالا رفتن ST قطعه	۱۰.۱۱٪ (۱۶۲,۲۷۴)	۱۰.۰۴٪ (۹۶,۸۲۸)	۱۰.۱۲٪ (۴۰,۴۴۶)
انفارکتوس میوکارد با بالا رفتن ST قطعه	۶.۲۴٪ (۱۰۰,۲۰۶)	۶.۲۶٪ (۶۰,۳۸۱)	۶.۲۲٪ (۳۹,۸۲۵)
نارسایی قلبی	۱۵.۵۳٪ (۲۴۹,۳۲۵)	۱۵.۵۵٪ (۱۵۰,۰۵۵)	۱۵.۵۰٪ (۹۹,۳۷۰)
آنزیم ناپایدار	۳.۷۱٪ (۴۳,۴۶۶)	۳.۷۲٪ (۳۶,۲۳۳)	۳.۶۹٪ (۱۷,۲۴۳)
فیبریلاسیون دهلیزی	۱۸.۸۲٪ (۳۰۲,۱۴۶)	۱۸.۶۸٪ (۱۸۰,۱۹۱)	۱۹.۰۴٪ (۱۲۱,۹۵۵)
تاکی کاردی بطنی	۱.۸۵٪ (۲۹,۶۷۲)	۱.۸۰٪ (۱۷,۳۸۵)	۱.۹۲٪ (۱۲,۸۷۷)
ایست قلبی	۲.۵۲٪ (۴۰,۵۰۵)	۲.۵۰٪ (۲۴,۱۰۰)	۲.۵۶٪ (۱۶,۴۰۵)
تاکی کاردی فوق بطنی	۱.۵۰٪ (۲۴,۱۴۶)	۱.۴۸٪ (۱۴,۲۸۶)	۱.۵۴٪ (۹,۸۶۰)
بلوک دهلیزی-بطنی	۲.۴۹٪ (۴۰,۰۱۳)	۲.۴۸٪ (۲۳,۹۳۶)	۲.۵۱٪ (۱۶,۰۷۷)
آمبولی ریوی	۳.۰۲٪ (۴۸,۴۸۵)	۳.۰۲٪ (۲۸,۴۵۸)	۳.۰۳٪ (۱۳,۰۳۷)
تنگی آئورت	۱.۸۸٪ (۳۰,۱۲۰)	۱.۸۹٪ (۱۸,۲۸۱)	۱.۸۵٪ (۱۱,۸۳۹)
فشارخون ریوی	۲.۴۶٪ (۴۰,۳۳۱)	۲.۴۷٪ (۲۱,۸۶۹)	۲.۴۶٪ (۱۴,۴۶۲)
کاردیومیوپاتی هیپرتروفیک	۰.۲۸٪ (۴,۴۸۵)	۰.۳۰٪ (۲,۹۰۴)	۰.۲۵٪ (۱,۵۸۱)
افتادگی دریچه میترا	۱.۵۳٪ (۲۴,۴۸۱)	۱.۴۸٪ (۱۴,۲۷۰)	۱.۵۹٪ (۱۰,۲۱۱)
تنگی دریچه میترا	۰.۱۸٪ (۲,۹۲۵)	۰.۱۸٪ (۱,۷۴۳)	۰.۱۸٪ (۱,۱۸۲)



شکل ۲ | مقایسه عملکرد مدل AUROC برای مدل های DL و XGB با ردیابی های ECG (افزایش فراوانی و درصد ECG های دارای بیماری های قلبی عروقی منتخب در مطالعه). مدل های DL و XGB عملکرد بهتری را در تشخیص بیماری های قلبی عروقی نشان دادند. مدل های DL و XGB عملکرد بهتری را در تشخیص بیماری های قلبی عروقی نشان دادند. مدل های DL و XGB عملکرد بهتری را در تشخیص بیماری های قلبی عروقی نشان دادند.

شکل ۲ | مقایسه عملکرد مدل AUROC برای مدل های DL و XGB با ردیابی های ECG (افزایش فراوانی و درصد ECG های دارای بیماری های قلبی عروقی منتخب در مطالعه). مدل های DL و XGB عملکرد بهتری را در تشخیص بیماری های قلبی عروقی نشان دادند. مدل های DL و XGB عملکرد بهتری را در تشخیص بیماری های قلبی عروقی نشان دادند. مدل های DL و XGB عملکرد بهتری را در تشخیص بیماری های قلبی عروقی نشان دادند.

قبل از مطالعات پیش بینی مبتنی بر ECG بررسی نشده بود و نشان داد که هر دو مدل DL و XGB عملکرد پیش بینی خوب تا عالی را نشان می دهند و مدل های DL برای بیشتر شرایط CV مورد مطالعه بهتر از مدل های XGB عمل می کنند.

مطالعات قبلی با استفاده از تشخیص ECG با قابلیت هوش مصنوعی نشان داده اند که مدل های DL و ML می توانند ریتیم ECG و ناهنجاری های مورفولوژیکی را در ECG به طور دقیق تشخیص دهند. با این حال، آنها بینشی در مورد عملکرد برای تشخیص بیماری های قلبی که به طور معمول از طریق ECG تشخیص داده نمی شوند، ارائه نکرده اند. مطالعه ما نشان می دهد که چگونه تکنیک های استاندارد یادگیری ماشین می توانند مدل هایی را یاد بگیرند که می توانند از نوار قلب ۱۲ لید ساده و آسان برای به دست آوردن، نه تنها بیماری های قلبی عروقی، بلکه اختلالاتی را که به طور معمول با استفاده از نوار قلب تشخیص داده نمی شوند، به طور دقیق پیش بینی کنند. این مدل ها ابزارهای بالقوه ای برای غربالگری زودهنگام بیماری های قلبی عروقی، به ویژه مواردی هستند که بار قابل توجهی را بر سیستم مراقبت های بهداشتی وارد می کنند و ممکن است به شناسایی دقیق تر بیماری های بالینی کمک کنند.

۱۵ تشخیص). قابل ذکر است که مناطقی که بیشترین سهم را در تشخیص داشتند عبارتند از: فواصل PR و کمپلکس های QRS در STEMI، امواج T در NSTEMI، کمپلکس های QRS در PHTN، ضربان های VT در بیماران مبتلا به VT غیر پایدار، کمپلکس های QRS در AS، امواج p در AVB و ناحیه قطعه ST در HF. شکل ۵ تحلیل های اهمیت ویژگی مدل XGB را بر اساس اندازه گیری های ECG نشان می دهد. که افزایش اطلاعات قابل توجهی را با مدت زمان P برای پیش بینی AF، ضربان قلب برای SVT، فاصله RR برای UA، مدت زمان QRS برای AVB، محور T پیشانی برای HF، محور T افقی برای NSTEMI، فاصله QT تصحیح شده با نرخ Bazett برای CA و غیره نشان می دهد.

بحث

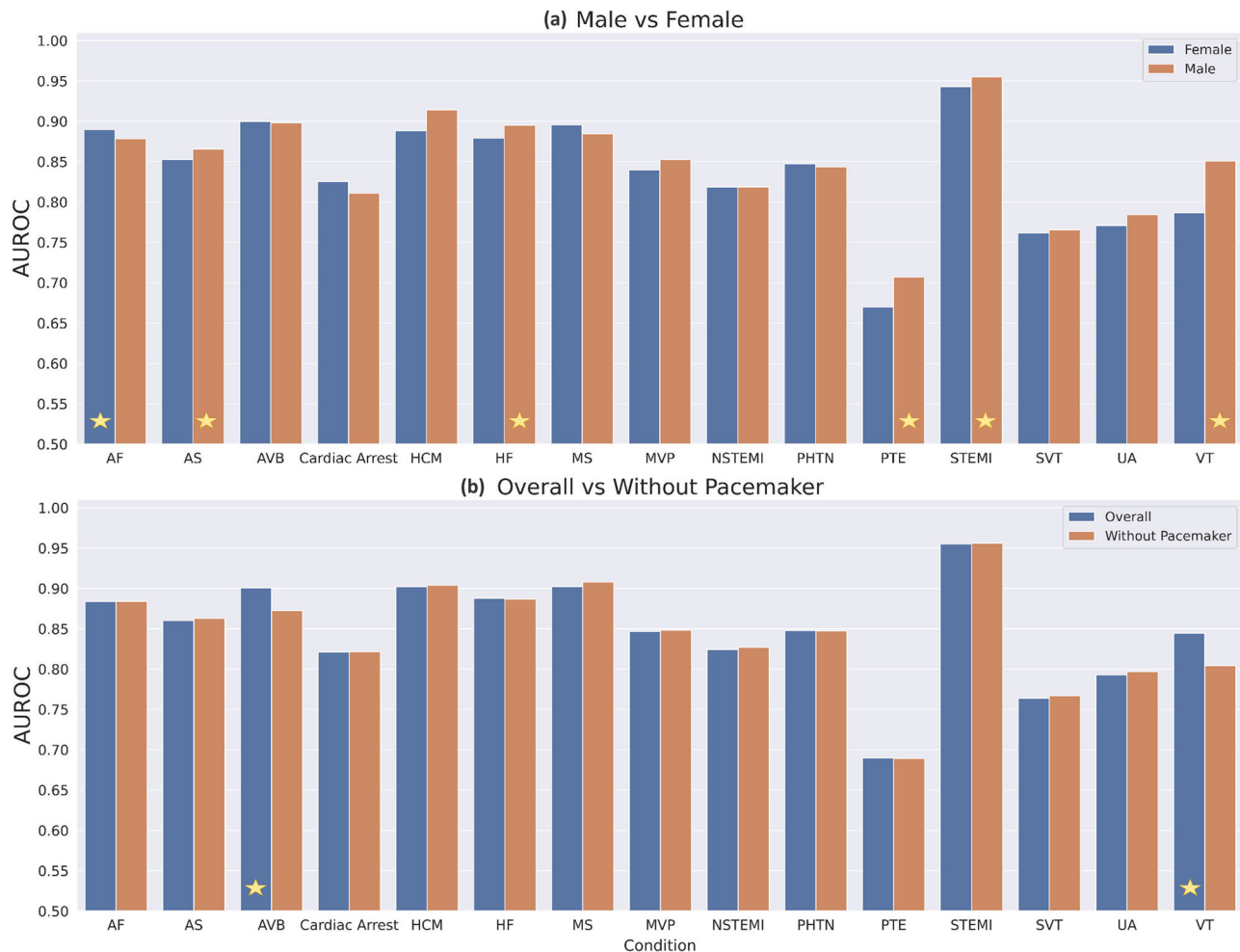
در این مطالعه بزرگ در سطح جمعیت با پرونده های سلامت اداری مرتبط شامل میلیون ها نوار قلب، ما مدل های پیش بینی مبتنی بر یادگیری ماشین را برای تشخیص بیماری های شایع قلبی عروقی از جمله موارد زیر توسعه داده و اعتبارسنجی کردیم.

شرایط مهم CV^{۳۹} مهم تر از همه، استفاده از این سیستم های طبقه بندی خودکار می تواند دقت بیشتری در مراقبت را در مناطق دورافتاده ای که دسترسی محدودی به متخصصان پزشکی و قالب و عروق واحد شرایط دارند، افزایش دهد. تحقیقات بیشتری لازم است به مشخص شود که آیا می توان از رابط های غیرالکتری خودکار مبتنی بر نوار قلب برای مدیریت زود هنگام و پیشگیری از پیشرفت بیماری و ارائه مراقبت مقرر به طرفه استفاده کرد یا خیر.

مدل های DL الگوریتم های پیچیده ای با میلیون ها پارامتر هستند و احتمالاً هنگام آموزش روی مجموعه داده های کوچک، بیش برآزش می شوند.^{۴۰-۴۲} حتی وقتی مجموعه داده های پزشکی بزرگی در دسترس هستند، معمولاً بدون برجسته یا پویا حاشیه نویسی هستند که چالش های بیشتری را برای رویکردهای یادگیری ماشینی تحت نظارت ایجاد می کند. مجموعه داده های بزرگ ECG آلبترتا ما با استفاده از استاندارد طلایی ۱۲ لید و ارتباط آن با داده های سطح جمعیت^{۳۳،۳۵} نشان دهنده یک آزمایشگاه جمعیتی طبیعی در طیف گسترده ای از متغیرهای جمعیتی، نژاد، جنس و سن است و در این روزمره ایچ اسی را برای توسعه الگوریتم های یادگیری ماشین مبتنی بر ECG برای تشخیص بیماری های قلبی عروقی رایج فراهم می کند. علاوه بر این، مطالعه حاضر بر روی اولین ECG ثبت شده در یک رخدادهای مراقبت های بهداشتی تمرکز کرد که سناریوهای واقعی از اولین تماس با پزشکی بیمار در محل مراقبت را شبیه سازی کند.

مدل ما سطح AUROC و C-index را برای ۱۲ مورد از ۱۵ بیماری، سن، بیش از ۸۰٪ برای بیماری (یعنی HCM، AF، STEMI، و AVE) بیش از ۸۰٪ نشان داد. مدل DL بر روی ECG، بیشترین دقت را در تشخیص ECGB و اندازه گیری های رویین ECG برای پیش بینی همه همای های شامل شده به جز بلوک دهلیزی-بطنی که هر دو مدل عملکرد عالی قابل مقایسه ای داشتند. ارائه داد، با بهبود ۱۸٪ در عملکرد برای درجه مبتدل و بهبود ۷۳٪ در عملکرد در تشخیص انقباض کوس میوکارد. نکته مهم این است که ما استحکام مدل را با توجه به هرگونه سوگیری بالقوه نسبت به گروه های گسسته و ECG های به دست آمده در حضور ضربان سیار قلبی یا دستگاه های بی ریتمی چ (LVAD) که می تواند تشخیص ECG را پیچیده کند، کاهش ارزیابی کردیم. ما دریافتیم که مدل DL ما همچنان قوی است و به نظر می رسد هنگام ارزیابی روی بیماران از هر دو جنس یا ضربان ساز/LVAD، به یک اندازه خوب عمل می کند. در واقع، مدل های ما حتی عملکرد بهتری را نشان دادند وقتی که ECG ها با ضربان سازها در داده های آزمایش گنجانده شدند.

مطالعه ما محدودیت هایی دارد که نیاز به بحث بیشتر دارد. اول همه ECG ها توسط دستگاه های از یک تولیدکننده (سیستم اطلاعات قلبی) تولید شدند که ممکن است تعمیم پذیری و ارزیابی یافته های ECG های سایر سیستم ها را محدود کند. دوم، اندازه گیری های ECG مورد استفاده در مدل های XGB بر روی دستگاه های قلبی-عروقی ارائه شدند و توسط آزمایشگاه اصلی خوانده یا توسط متخصص اطفال تنظیم نشده بودند. سوم، برجسته های ما از کدهای ICD ثبت شده در پرونده اوزان و بستری گرفته شده بودند. با توجه به ماهیت جمع آوری داده هادر پرونده های پزشکی اداری، زمان دقیق بروز بیماری در طول یک دوره مراقبت های بهداشتی را نمی توان به طور قطعی مشخص کرد. در نتیجه، در مورد نادر که یک بیماری خاص از مجموعه آوری ECG اولیه ظاهر می شود، وظیفه پیش بینی می تواند به عنوان تشخیص زود هنگام و جای پیش بینی تشخیص یک بیماری موجود تقسیم شود. این مقایسه قابل توجه است زیرا تشخیص زود هنگام در مدیریت بالینی ارزشمند است و پیش هادی در مورد عوارض احتمالی در آینده نزدیک ارائه می دهد که می تواند به همای اندازه برای مراقبت های پیشگیری باشد. چهارم، آزمایش داخلی، حتی در مقیاس قابل توجه، ممکن است در مقایسه با اعتبارسنجی های خارجی، ثانویه در نظر گرفته شود. این امر عمدتاً به این دلیل است که سوگیری های ذاتی یک سیستم سلامت و محدودیت های توانایی به دلیل شکلهای جمعیت بیمار، تجهیزات، رویه های تولید برجسته و سایر عوامل وجود دارند. ما متأسفانه، ما نتوانستیم اعتبارسنجی خارجی را برای مدل های چندبرجسته خود، ارائه دهیم زیرا هیچ مجموعه داده ECG خارجی مناسبی به ۱۵۰ برجسته تشخیصی مبتنی بر ICD-10 انتخاب شده مرتبط باشی وجود ندارد. با این حال، عملکرد اعتبارسنجی leave-one-hospital-out مدل های ECG: DL، سن، جنس، ما، استحکام مدلی های ما را در بیمارستان ها نشان می دهد. پنجم، مطالعه ما بر اساس یک گروه واقعی از بیماران مراجعه کننده به بخش های اورژانس و بیمارستان ها با میزان شیوع متفاوت تشخیصی های مورد نظر است. تغییر در میزان مثبت برجسته های مختلف می تواند توضیح دهد که چرا پیش بینی برای برخی از تشخیص های دقیق نیاز مایهتی بود. ما داده ها را تقویت با دستکاری نکردیم زیرا هدف ما این است که درجهای این مدل ها در سیستم های پرودهای پزشکی الکترونیکی مستقر کنیم. علاوه بر این، مجموعه داده های آموزشی ما که تقریباً یک میلیون ECG دارد، تعداد کافی از موارد مثبت را برای توسعه مدل های پیش بینی کننده موثر داشت. با این حال، برخی از برجسته های در مدل های ما ممکن است PPV پایین تر از حد مطلوب را نشان دهند، که مستلزم بررسی دقیق در مورد واجد شرایط بودن آنها برای ارزیابی بالینی است.



شکل ۳ | مقایسه عملکرد AUROC برای ۱۵ بیماری قلبی-عروقی برای زیرگروه‌های خاص. ارزیابی‌ها برای مردان و زنان بیماران مبتلا به سرطان ریه به طور جداگانه انجام می‌شود (الف) و همچنین نوار قلب (ECG)

بدون ضربان سازها و تمام ECGها در مجموعه‌ی ی نگه دارند (ب). ارتفاع میله‌ها نشان دهنده عملکرد در اعتبارسنجی خارجی است و مدل‌هایی که از نظر آماری عملکرد بالاتری دارند با ستاره مشخص شده‌اند.

استقرار. ششم، ما مدل‌های خود را در درجه اول با استفاده از AUROC یا C-index، یک معیار رایج در ادبیات زیست پزشکی، ارزیابی کردیم. با این حال، این معیار محدودیت‌هایی دارد و هزینه‌های طبقه‌بندی نادرست نتایج مثبت کاذب یا منفی کاذب را در نظر نمی‌گیرد. برای استقرار مدل، ارزیابی‌های سفارشی با هدف به حداقل رساندن هزینه مورد انتظار، که هزینه‌های طبقه‌بندی نادرست و سایر عوامل تخصیص منابع را در نظریه‌گیری، باید در اولویت قرار گیرند. علاوه بر این، علیرغم ماهیت جعبه سیاه برخی از رویکردهای یادگیری ماشینی، ما از تکنیک‌هایی مانند تحلیل GradCAM از مدل‌های DL (به ترتیب، تحلیل SHAP از مدل‌های XGB) برای یافتن الگوهای ECG (یا اندازه‌گیری‌های ECG) که به تشخیص بیماری‌های شایع قلبی-عروقی کمک می‌کنند، استفاده کردیم.

در نتیجه، ما با استفاده از پایگاه‌های داده جامع مرتبط اداری در سطح جمعیت‌نشان می‌دهیم که مدل‌های پیش‌بینی DL و XGB مبتنی بر ECG عملکرد پیش‌بینی خوب تا عالی را در تشخیص بیماری‌های شایع قلبی-عروقی نشان می‌دهند. مدل‌های DL ردیابی ECG دقت پیش‌بینی بهتری را در بین بیماری‌های مورد مطالعه نسبت به مدل‌های XGB مبتنی بر اندازه‌گیری‌های معمول ECG ارائه دادند. مدل‌ها به طور قابل مقایسه‌ای بین گروه‌های جنسی مختلف و در بیماران با و بدون ضربان ساز یا LVAD عمل کردند. تحقیقات آینده برای تعیین چگونگی پیاده‌سازی این مدل‌ها در عمل بالینی برای تشخیص زودهنگام و طبقه‌بندی خطر مورد نیاز است.

روش‌ها

منابع داده

این مطالعه در آلبرتا، کانادا، انجام شد که در آن یک سیستم مراقبت‌های بهداشتی تک پرداخت کننده با دسترسی جهانی و ثبت ۱۰۰٪ تمام تعاملات با سیستم مراقبت‌های بهداشتی وجود دارد.

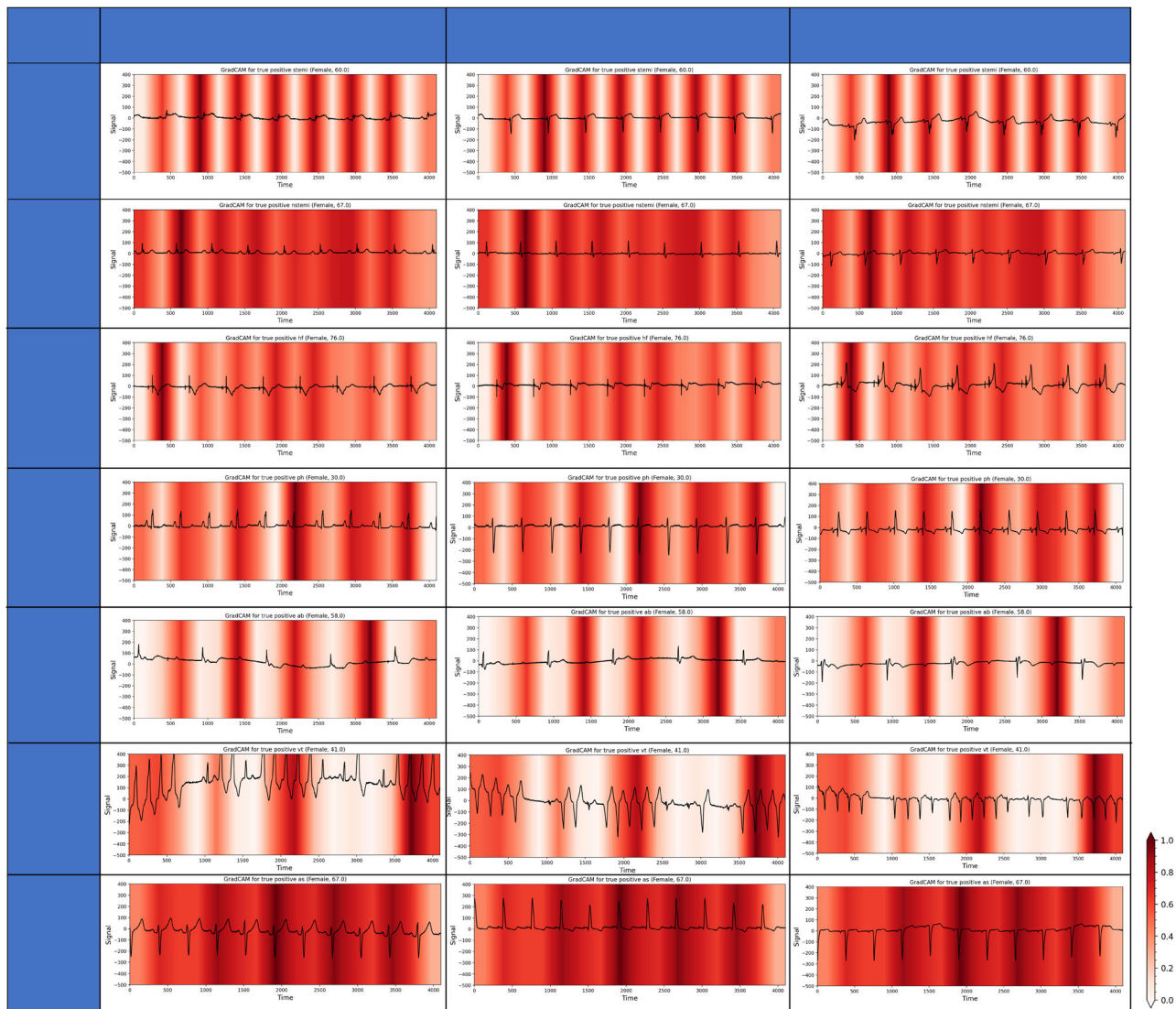
داده‌های نوار قلب (ECG) با استفاده از یک شماره سلامت بیمار منحصر به فرد به پایگاه‌های داده سلامت اداری زیر متصل شدند: (۱) پایگاه داده چکیده ترخیص (DAD) حاوی داده‌های مربوط به بستری شدن بیماران؛ (۲) پایگاه داده سیستم گزارش دهی مراقبت‌های سرپایی ملی (NACRS) شامل تمام ویزیت‌های سرپایی بیمارستانی و اورژانس؛ و (۳) ثبت طرح بیمه مراقبت‌های بهداشتی آلبرتا (AHCIP) که اطلاعات جمعیت شناختی را ارائه می‌دهد.

داده‌های نوار قلب (ECG)

ما از نمودارهای استاندارد ECG دوازده لید (سری‌های ولتاژ-زمان، نمونه برداری شده با فرکانس ۵۰۰ هرتز به مدت ۱۰ ثانیه برای هر یک از ۱۲ لید) و اندازه‌گیری‌های ECG (که به طور خودکار توسط الگوریتم داخلی سیستم Philips IntelliSpace ECG تولید می‌شوند) استفاده کردیم. اندازه‌گیری ECG شامل ضربان دهلیزی، ضربان قلب، فاصله RR، مدت موج P، محور P پیشانی، محور P افقی، فاصله PR، مدت زمان QRS، محور QRS پیشانی در ۴۰ میلی ثانیه اولیه، محور QRS افقی در ۴۰ میلی ثانیه اولیه، محور QRS پایانی، محور QRS افقی در ۴۰ میلی ثانیه پایانی، محور موج ST پیشانی (معادل انحراف ST)، محور T پیشانی، محور موج ST افقی، محور T افقی، شروع موج Q، فاصله QT اصلاح شده با نرخ Fridericia، فاصله QT، فاصله QT اصلاح شده با نرخ Bazett بود.

گروه تحلیل

گروه مورد مطالعه قبلاً شرح داده شده است. به طور خلاصه، بیمارانی که بین فوریه ۲۰۰۷ تا آوریل ۲۰۲۰ در ۱۴ مرکز در آلبرتا، کانادا بستری شدند و شامل ۱۵۸٬۰۸۸ نوار قلب از ۳٬۳۳۶٬۰۹۱ ویزیت اورژانسی و



شکل ۴ | نمودارهای GradCAM برای مدل DL در تشخیص بیماری های قلبی عروقی مختلف. نمودارهای ECG نماینده برای یک گروه منتخب از تشخیص ها انتخاب شدند. نتایج GradCAM به کل جمعیت تعمیم داده نمی شود، اما نشان دهنده پیش بینی مدل DL برای یک مورد نماینده واحد است. نواحی تیره تر در هر نمودار در GradCAM نشان دهنده نواحی با بیشترین سهم در پیش بینی تشخیصی مدل DL هستند. فواصل PR و کمپلکس های QRS در STEMI، امواج T در

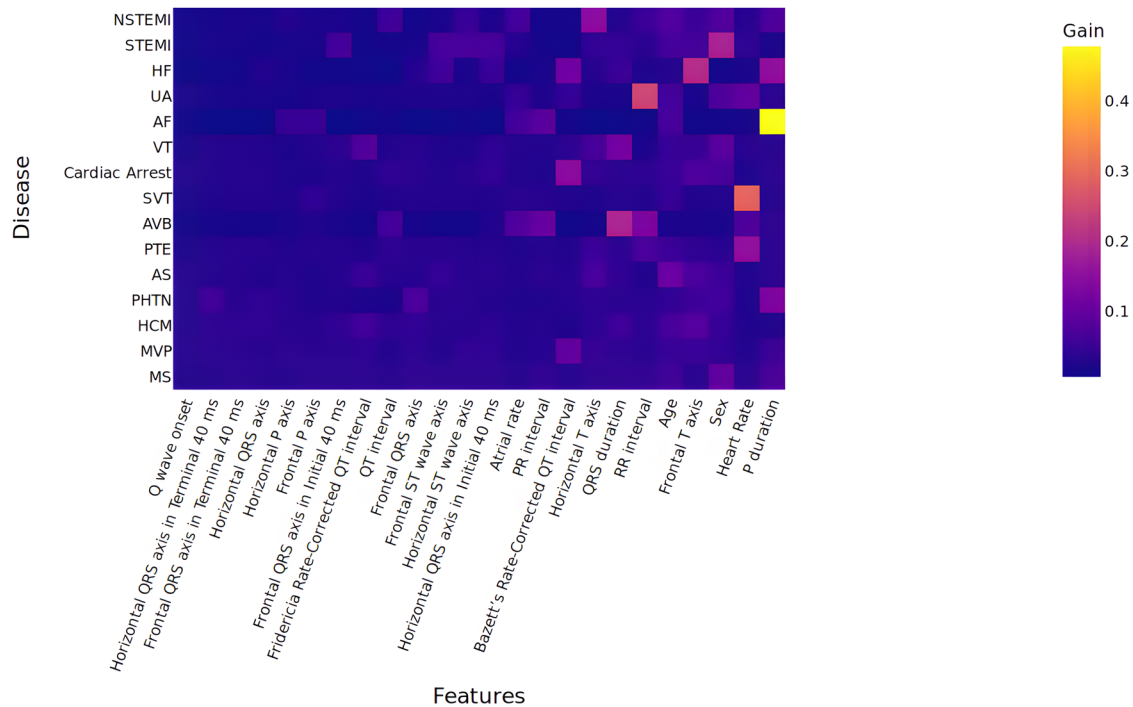
انفارکتوس میوکارد بدون قطعه ST، کمپلکس های QRS در PHTN، ضربان های VT در بیمار مبتلا به VT غیر پایدار، کمپلکس های QRS در AS، امواج p در AVB و ناحیه قطعه ST در HF بیشترین سهم را در تشخیص هر بیماری داشتند. تنگی آیورت AS، بلوک دهلیزی-بطنی AVB، یادگیری عمیق DL، نوار قلب، نارسایی قلبی HF، انفارکتوس میوکارد بدون صعود قطعه ST در NSTEMI، انفارکتوس میوکارد با صعود قطعه ST در STEMI، فشار خون ریوی PHTN، تاکی کاردی بطنی VT.

10) کدهای موجود در فیلدهای تشخیص اولیه یا هر یک از 24 فیلد تشخیص ثانویه یک رویداد مراقبت های بهداشتی مرتبط با یک نوار قلب خاص (جدول تکمیلی 5). اعتبار کدگذاری ICD در پایگاه های داده سلامت اداری قبلاً مشخص شده است. اگر نوار قلب (ECG) در طول یک دوره اورژانس یا بستری انجام می شد، برای تمام تشخیص های مورد نظر که در آن دوره ثبت شده بودند، مثبت در نظر گرفته می شد. برخی از تشخیص ها، مانند AF، SVT، VT، STEMI و AVB، که معمولاً از طریق ECG شناسایی می شوند، به عنوان کنترل مثبت در مطالعه گنجانده شدند تا اثربخشی مدل های ما در تشخیص شرایط قابل تشخیص با ECG نشان داده شود.

هدف مدل پیش بینی، خروجی دادن احتمالات کالبره شده برای هر یک از 15 وضعیت انتخاب شده بود. این مدل های آموخته شده می توانستند از ECG هایی که در هر نقطه زمانی در طول یک دوره مراقبت های بهداشتی به دست آمده بودند، استفاده کنند. توجه داشته باشید که یک ویژگی تک بیمار ممکن است شامل چندین ECG باشد. هنگام آموزش مدل، ما از تمام ECG ها (ECG های چندگانه متعلق به یک دوره شامل شدند) در مجموعه آموزش/توسعه برای به حداکثر رساندن یادگیری استفاده کردیم. با این حال، برای ارزیابی مدل های خود، ما فقط از اولین ECG در یک دوره معین در مجموعه تست/محدوده استفاده کردیم، با هدف تولید یک سیستم پیش بینی که بتواند در نقطه مراقبت، زمانی که اولین ECG بیمار به دست می آید، به کار گرفته شود.

۱۰۷۱،۵۷۶ مورد بستری در بیمارستان برای ۲۶۰،۶۵ بیمار. ملاقات های همزمان با مراقبت های بهداشتی (مراجعه به اورژانس و/یا بستری در بیمارستان) که برای یک بیمار در یک دوره ۴۸ ساعته رخ داده بود، به عنوان انتقال و بخشی از یک دوره مراقبت های بهداشتی در نظر گرفته شد. اگر تاریخ ثبت نوار قلب در بازه زمانی بین تاریخ پذیرش و تاریخ ترخیص یک دوره بود، یک رکورد نوار قلب به یک دوره مراقبت های بهداشتی مرتبط می شد. پس از حذف نوار قلب هایی که نمی توانستند هیچ دوره ای مرتبط باشند، نوار قلب بیماران زیر ۱۸ سال، و همچنین نوار قلب هایی با کیفیت سیگنال ضعیف (که از طریق پرچم های هشدار تولید شده توسط الگوریتم کیفیت داخلی سازنده دستگاه نوار قلب شناسایی می شدند)، گروه تجزیه و تحلیل ما شامل ۱،۶۰۵،۲۶۸ نوار قلب از ۷۴۸،۷۷۳ دوره در ۲۴۴،۰۷۷ بیمار بود (شکل ۱).

وظایف پیش بینی مامد های مبتنی بر ECG را برای پیش بینی احتمال تشخیص هر یک از 15 بیماری شایع قلبی عروقی خاص در یک بیمار توسعه داده و ارزیابی کردیم: VT، CA، AVB، UA، NSTEMI، STEMI، PTE، HCM، AS، MVP، MS، PHTN، AF، SVT، HF. این بیماری ها بر اساس سوابق طبقه بندی بین المللی بیماری های مربوطه، ویرایش دهم (ICD-) شناسایی شدند.



شکل 5 | نقشه حرارتی تحلیل اهمیت ویژگی ها در مدل های XGBoost با اندازه گیری های ECG، سن و جنسیت. اهمیت ویژگی مبتنی بر افزایش اطلاعات برای شرایط مختلف قلبی عروقی با مدل های XGBoost مبتنی بر اندازه گیری های ECG

افزایش اطلاعات قابل توجهی را با مدت زمان P برای پیش بینی AF، ضربان قلب برای SVT، فاصله RR برای UA و غیره نشان داد. الکتروکاردیوگرام ECG، اختصارات برای اندازه گیری ها و بیماری های ECG در جداول تکمیلی ۸ و ۹ ارائه شده است.

داده های مربوط به یک بیمار بین داده های توسعه و ارزیابی یا بین بخش های آموزش/آزمایش اعتبارسنجی متقابل داخلی به اشتراک گذاشته نشده بودند. همانطور که قبلاً ذکر شد، از آنجایی که انتظار داریم سناریوی استقرار سیستم پیش بینی ما در نقطه مراقبت باشد، مدل های خود را تنها با استفاده از اولین نوار قلب بیمار در یک اپیزود مشخص، که در طول ویزیت اورژانس یا بستری شدن در بیمارستان گرفته شده بود، ارزیابی کردیم. ارزیابی نوار قلب، اپیزودها و بیماران مورد استفاده در داده های کلی و در تقسیم بندی های تجربی در شکل ارائه شده است. جدول تکمیلی ۵. علاوه بر ارزیابی اولیه، ما آزمایش خود را گسترش می دهیم تا تمام ECG های موجود در مجموعه ی مورد نظر را نیز شامل شود، تا تطبیق پذیری مدل DL را در مدیریت ECG های ثبت شده در هر نقطه ای از یک اپیزود نشان دهیم. علاوه بر این، ما «اعتبارسنجی بدون حذف یک بیمارستان» را با استفاده از دو بیمارستان بزرگ مراقبت های عالی انجام دادیم تا استحکام مدل خود را با توجه به تفاوت های توزیعی بین محل های بیمارستان ارزیابی کنیم. برای تضمین جداسازی کامل بین مجموعه های آموزشی و آزمایشی، ECG های بیماران بستری شده در هر دو بیمارستان آموزشی و آزمایشی در طول دوره مطالعه را حذف کردیم، همانطور که در شکل تکمیلی ۱ نشان داده شده است. در نهایت، برای تأکید بر کاربرد مدل DL در سناریوهای غربالگری، ارزیابی های اضافی را با ادغام ۱۵ برجسب بیماری در یک پیش بینی ترکیبی ارائه می دهیم و در نتیجه بازده تشخیصی را افزایش می دهیم.^{۲۰}

مماساحت زیر منحنی مشخصه عملکرد گیرنده (AUROC)، معادل شاخص C) و مساحت زیر منحنی دقت-فراخوانی (AUPRC) را گزارش کردیم. همچنین، پس از باینری کردن احتمالات پیش بینی به کلاس های تشخیص/عدم تشخیص با استفاده از نقاط برش بهینه مشتق شده از مجموعه آموزشی شاخص یودن، امتیاز F1، ویژگی، فراخوان، دقت (معادل PPV) و دقت را تولید کردیم.^{۲۱} ما همچنین از معیار کالیبراسیون Brier Score استفاده کردیم.^{۲۲} (که در آن امتیاز کمتر نشان دهنده کالیبراسیون بهتر است) برای ارزیابی اینکه آیا احتمالات پیش بینی شده با نسبت های مشاهده شده مطابقت دارند یا خیر. زیرگروه های جنسیت و ضربان ساز: ما عملکرد مدل های خود را در زیرگروه های خاص بیمار، بر اساس جنسیت بیمار، بررسی کردیم. همچنین هرگونه سوگیری بالقوه را با ECG های ثبت شده در حضور ضربان ساز قلبی (از جمله ضربان ساز یا دفیبریلاتورهای قلبی قابل کاشت [ICD]) یا دستگاه های کمکی بطنی (VAD) بررسی کردیم، زیرا تفسیر ECG در این شرایط می تواند دشوار باشد، با مقایسه عملکرد مدل در ECG های بدون ضربان ساز در مجموعه نگهدارنده در مقابل مجموعه کلی نگهدارنده (شامل ECG ها با یا بدون ضربان ساز) (شکل ۱).^۱ تشخیص

در طول ویزیت اورژانس یا بستری شدن در بیمارستان (برای جزئیات بیشتر به بخش «ارزیابی» در زیر مراجعه کنید).

ما از DL مبتنی بر ResNet برای سری های زمانی ولتاژ غنی از اطلاعات و XGB مبتنی بر افزایش گرادیان برای اندازه گیری های ECG استفاده کردیم.^{۲۳} برای تعیین اینکه آیا ویژگی های جمعیت شناختی (سن و جنس) ارزش پیش بینی کننده افزایشی به عملکرد مدل های آموزش دیده فقط بر اساس ECG اضافه می کنند، ما مدل ها را به روش زیر توسعه داده و گزارش کردیم: (الف) فقط ECG (DL: ردیابی ECG)؛ (ب) ECG + سن، جنس (DL: ردیابی ECG، سن، جنس)؛ (ج) XGB: اندازه گیری ECG، سن، جنس. این مطالعه است.)^{۲۴}؛ و (د) XGB: اندازه گیری ECG، سن، جنس.

الگوریتم های یادگیری ماشین یک روش طبقه بندی چند برجسبی با برجسب های دودویی، یعنی وجود (بله) یا عدم وجود (خیر) برای هر یک از ۱۵ تشخیص، استفاده کردیم تا احتمال ابتلای یک بیمار جدید به هر یک از این شرایط را تخمین بزنیم. از آنجایی که ورودی مدل های که از اندازه گیری های ECG استفاده می کردند، داده های جدولی ساختاریافته بود، ما گروه های درختی تقویت شده با گرادیان (XGB) را آموزش دادیم.^{۲۵} در حالی که ما از شبکه های عصبی کانولوشن عمیق برای مدل هایی با ردیابی سری های ولتاژ-زمان ECG استفاده کردیم. برای هر دو مدل XGB و DL، ما از ۹۰٪ داده های آموزشی 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 برآزش، استفاده کردیم.^{۲۶} برای یادگیری عمیق، ما یک مدل ResNet واحد را برای یک کار چند کلاسه و چند برجسبی یاد گرفتیم.^{۲۷} که هر سیگنال ECG را به ۱۵ مقدارنگاشت می کند، مربوط به احتمال حضور هر یک از ۱۵ تشخیص است. از سوی دیگر، برای تقویت گرادیان، ما ۱۵ مدل XGB دودویی متمایز را یاد گرفتیم که هر کدام سیگنال ECG را به احتمال یکی از برجسب های منفرد نگاشت می کنند. جزئیات روش شناختی پیاده سازی های مدل XGB و DL ما قبلاً شرح داده شده است.^{۲۵}

ارزیابی و تجسم طراحی ارزیابی: ما از تقسیم بندی ۶۰/۴۰ روی داده ها برای آموزش و ارزیابی استفاده کردیم. ما مجموعه داده های کلی ECG را به تقسیم بندی های تصادفی ۶۰٪ برای توسعه مدل (که از اعتبارسنجی متقابل داخلی پنج گانه برای آموزش و تنظیم دقیق مدل های نهایی استفاده کرد) و ۴۰٪ باقی مانده را به عنوان مجموعه نگه دارنده برای اعتبارسنجی خارجی نهایی تقسیم کردیم. ما اطمینان حاصل کردیم که ECG ها از

وکدهای رویه مورد استفاده برای شناسایی وجود ضربان سازها در جدول تکمیلی 7 ارائه شده است.

مقایسه مدل ها: برای هر ارزیابی، ما عملکردها را از اعتبارسنجی متقابل داخلی پنج گانه و همچنین عملکرد نهایی در مجموعه داده های پشتیبان، با استفاده از تقسیم بندی های آموزش و آزمایش یکسان برای سناریوهای مختلف مدل سازی، گزارش می کنیم. عملکردها بین مدل ها با نمونه برداری از نمونه های پشتیبان با جایگزینی به صورت جفتی مقایسه شدند تا در مجموع 10000 تکرار بوت استرپ از تفاوت های جفتی در AUROC ایجاد شود. یعنی، هر کدام بدون ضربان ساز در مقابل مدل اصلی مقایسه می شوند. گفته می شود که تفاوت در عملکرد مدل از نظر آماری معنی دار است اگر فواصل اطمینان 95٪ از میانگین تفاوت های جفتی در AUROC ها شامل مقدار صفر برای مدل های مقایسه شده نباشد.

مصورسازی ها: ما از مقادیر اهمیت ویژگی ها بر اساس اطلاعات به دست آمده برای شناسایی اندازه گیری های ECG که نقش کلیدی در پیش بینی تشخیص در مدل های XGB داشتند، استفاده کردیم. علاوه بر این، نقشه های فعال سازی گرادینان را که در پیش بینی مدل از تشخیص در مدل های DL ما نقش داشتند، با استفاده از نگاشت فعال سازی کلاس با وزن گرادینان (GradCAM) مصورسازی کردیم. در آخرین لایه کانولوشن. همچنین، ما از مقادیر اهمیت ویژگی ها بر اساس بهره اطلاعات برای شناسایی اندازه گیری های ECG که نقش کلیدی در پیش بینی تشخیص در مدل های XGB داشتند، استفاده کردیم.

دردسترس بودن 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 و فقط برای نمایش کد تولید شده است، گنجانه ایم. انتظار نمی رود که آنها سیگنال های ECG واقعی یا توزیع های برچسب 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://figshare.com/s/b593e8d7bfe7cd8500b1>

دردسترس بودن کد

کدپایه برای آموزش مدل های یادگیری عمیق مورد استفاده در این مطالعه در آدرس زیر موجود است: <https://figshare.com/s/b593e8d7bfe7cd8500b1>

دریافت: ۱۶ آگوست ۲۰۲۳؛ پذیرش: ۲۶ آوریل ۲۰۲۴؛

Published online: 18 May 2024

منابع

۱. تیسون، جی. اچ. ژانگ، جی.، دلینگ، اف. ان و دئو، آر. سی. پروفایل های خودکار و قابل تفسیر نوار قلب بیمار برای تشخیص، ردیابی و کشف بیماری. دور قلب و عروق. کیفیت نتایج ۱۲، ۲۰۱۹ (e005289)

۲. آتیا، زی و همکاران. تخمین سن و جنس با استفاده از هوش مصنوعی از نوار قلب استاندارد ۱۲ لید. سیکل. آرتمی. الکتروفیزیولوژی ۱۲، ۲۰۱۹ (e007284)

۳. آتیا، زی و همکاران. غربالگری اختلال انقباض قلب با استفاده از الکتروکاردیوگرام مجهز به هوش مصنوعی. نات. مد. ۲۵، ۷۰-۷۴ (۲۰۱۹).

۴. آتیا، زی و همکاران. یک الگوریتم ECG مبتنی بر هوش مصنوعی برای شناسایی بیماران مبتلا به فیبریلاسیون دهلیزی در طول ریتم سینوسی: یک تحلیل گذشته نگر از پیش بینی پیامد. لانست ۳۹۴، ۸۶۱-۸۶۷ (۲۰۱۹).

۵. کوون، جی. ام. و همکاران. مقایسه عملکرد هوش مصنوعی و معیارهای تشخیص مرسوم برای تشخیص هیپرتروفی بطن چپ با استفاده از الکتروکاردیوگرافی. اروپا ۲۳، ۴۱۲-۴۱۹ (۲۰۲۰).

۶. سرتیج، م.، جیرانه، ی. و حجام الحسینی، ا. یک سیستم خودکار برای تشخیص آریتمی ECG با استفاده از تکنیک های یادگیری ماشین. مجله تحقیقات پزشکی جی. کلین. ۵۴۵۰ (۲۰۲۱).

۷. گوستافسون، اس. و همکاران. توسعه و اعتبارسنجی پیش بینی سکتة قلبی مبتنی بر ECG با یادگیری عمیق در بیماران بخش اورژانس. نماینده علمی ۱۲، ۱۹۶۱۵ (۲۰۲۲).

۸. وو، ال. و همکاران. شبکه های یادگیری عمیق، انفارکتوس میوکارد با بالا رفتن قطعه ST و رگ مقصر را به طور دقیق تشخیص می دهند. قلب و عروق جلوبی. پزشکی. ۹، ۷۷۲۰۷ (۲۰۲۲).

۹. الزابتی، اس. اس و همکاران. یادگیری ماشین برای تشخیص ECG و طبقه بندی خطر انفارکتوس میوکارد انسدادی. نات. مد. ۲۹، ۱۸۰۴-۱۸۱۳ (۲۰۲۳).

۱۰. ربیرو، ای اچ و همکاران. تشخیص خودکار نوار قلب ۱۲ لیدی با استفاده از یک شبکه عصبی عمیق. نات. کمون. ۱۱، ۱۷۶۰ (۲۰۲۰).

۱۱. ایسای، آی. و همکاران. یک معماری قوی یادگیری ماشین برای تحلیل ریتم ECG قابل اعتماد در طول CPR. کنفرانس مجموعه مقالات IEEE مهندسی پزشکی زیست شناسی اجتماعی ۲۰۱۹، ۱۹۰۷-۱۹۰۹ (۲۰۱۹).

۱۲. الوالا، آ. و همکاران. شبکه های عصبی عمیق برای تشخیص نبض مبتنی بر نوار قلب در هنگام ایست قلبی خارج از بیمارستان. آنتروپی ۳۰، ۳۰۵ (۲۰۱۹).

۱۳. چوی، جی. و همکاران. یادگیری عمیق شکل موج های ECG برای تشخیص نارسایی قلبی با کسر تخلیه بطن چپ کاهش یافته. نماینده علمی ۱۲، ۱۴۳۵ (۲۰۲۲).

۱۴. راگو، آ. و همکاران. تخمین غیرتهاجمی احتقان ریوی با هدایت ECG در بیماران مبتلا به نارسایی قلبی. نماینده علمی ۱۳، ۳۹۲۳ (۲۰۲۳).

۱۵. سومانی، اس. اس و همکاران. توسعه یک مدل یادگیری ماشین با استفاده از سیگنال های الکتروکاردیوگرام برای بهبود غربالگری آمبولی ریوی حاد. یورو. قلب جی. دیجیتال هلت ۳، ۵۶-۶۶ (۲۰۲۲).

۱۶. والته سیلوا، ب.، مارکز، ج.، نوبره منزس، م.، اولیویرا، آل. و پینتو، اف. جی. تشخیص آمبولی ریوی حاد مبتنی بر هوش مصنوعی: توسعه یک مدل یادگیری ماشین با استفاده از الکتروکاردیوگرام ۱۲ لید. کشیش پورت. کاردیول. ۴۲، ۶۴۳-۶۵۱ (۲۰۲۳).

۱۷. هاتا، ای. و همکاران. طبقه بندی تنگی آئورت با استفاده از ECG به وسیله یادگیری عمیق و تحلیل آن با استفاده از grad-CAM. کنفرانس مجموعه مقالات IEEE مهندسی پزشکی زیست شناسی اجتماعی ۲۰۲۰، ۱۵۴۸-۱۵۵۱ (۲۰۲۰).

۱۸. گوتو، اس. و همکاران. رویکرد یادگیری فدرال چندملیتی برای آموزش مدل های ECG و اکوکاردیوگرام برای تشخیص کاردیومیوپاتی هیپرتروفیک. گردش خون ۱۴۶، ۷۵۵-۷۶۹ (۲۰۲۲).

۱۹. کوهن-شلی، م. و همکاران. غربالگری الکتروکاردیوگرام برای تنگی دریچه آئورت با استفاده از هوش مصنوعی. یورو. هارت جی. ۴۲، ۲۸۸۵-۲۸۹۶ (۲۰۲۱).

۲۰. اولوا-سرنای، ای ای و همکاران. rECHOmend: یک رویکرد یادگیری ماشینی مبتنی بر نوار قلب برای شناسایی بیماران در معرض خطر بالای بیماری ساختاری قلب تشخیص داده نشده که توسط ... قابل تشخیص است.

اکوکاردیوگرافی گردش خون ۱۴۶، ۳۶-۴۷ (۲۰۲۲).

۲۱. آراس، ام. ای و همکاران. تشخیص فشار خون ریوی با استفاده از الکتروکاردیوگرام با استفاده از یادگیری عمیق. جی. کارت. شکست. ۲۹، ۱۰۱۷-۱۰۲۸ (۲۰۲۳).

۲۲. لیو، سی. ام. و همکاران. الکتروکاردیوگرام مجهز به هوش مصنوعی، تشخیص و پیش بینی مرگ و میر در بیماران مبتلا به فشار خون ریوی را بهبود می بخشد. آسیا ACC ۲۷، ۲۵۸-۲۷۰ (۲۰۲۲).

۲۳. چن، ال. فو، جی. و جیانگ، سی. پیش بینی ژنوتیپ مبتنی بر الکتروکاردیوگرام ۱۲ لید مشتق شده از یادگیری عمیق برای کاردیومیوپاتی هیپرتروفیک: یک مطالعه آزمایشی. آن. مدی. ۵۵، ۲۲۳۵۵۶۴ (۲۰۲۳).

۲۴. کو، دبلیو. وای. و همکاران. تشخیص کاردیومیوپاتی هیپرتروفیک با استفاده از الکتروکاردیوگرام مجهز به شبکه عصبی کانولوشن. مجله آمریکایی کاردیولوژی. ۷۵، ۷۲۲-۷۳۳ (۲۰۲۰).

۲۵. سان، دبلیو و همکاران. به سوی سیستم سلامت مبتنی بر یادگیری هوش مصنوعی برای پیش بینی مرگ و میر در سطح جمعیت با استفاده از الکتروکاردیوگرام ها. مجله پزشکی NPJ Digit ۲۱، ۶ (۲۰۲۳).

۲۶. لیو، ایکس.، وانگ، اچ.، لی، زد. و کین، ال. یادگیری عمیق در تشخیص نوار قلب: یک بررسی. سیستم مبتنی بر دانش ۲۷، ۱۷۸۷-۱۷۸۷ (۲۰۲۱).

۲۷. ماننت، جی. و همکاران. دقت تشخیص فیبریلاسیون دهلیزی بر روی نوار قلب توسط پزشکان مراقبت های اولیه و نرم افزار تشخیصی تفسیری: تجزیه و تحلیل داده های حاصل از آزمایش غربالگری فیبریلاسیون دهلیزی در سالمندان (SAFE). بی ام جی ۳۸۰، ۳۳۵ (۲۰۰۷).

۲۸. ورونس، جی. و همکاران. دقت پزشک اورژانس در تفسیر الکتروکاردیوگرام ها با احتمال انفارکتوس میوکارد با بالا رفتن قطعه ST: آیا کافی است؟ کارت حاد. مراقبت ۱۸، ۷-۱۰ (۲۰۱۶).

۲۹. تران، دی تی و همکاران. بار مالی فعلی و آینده پذیرش بیمارستانی به دلیل نارسایی قلبی در کانادا: یک تحلیل هزینه باز (FCMA). ۱۲، ۱۰۶۱۰ (E365-E370)

۳۰. سومانی، س. و همکاران. یادگیری عمیق و الکتروکاردیوگرام: مروری بر جدیدترین فناوری های موجود. اروپا ۲۳، ۱۱۷۹-۱۱۹۱ (۲۰۲۱).

۳۱. کلیفورد، جی. دی و همکاران. طبقه بندی AF از یک ثبت کوتاه تک لیدی ECG: چالش PhysioNet/computing در کاردیولوژی ۲۰۱۷.

ولزوما نمایانگر دیدگاه های خدمات درمانی آلبرتا یا هیچ یک از تأمین کنندگان مالی آن نیستند.

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 NS و AS، SVK، PK، تجزیه و تحلیل ها را انجام دادند؛ YN و AS، WS، SVK، این مطالعه را آغاز کردند، بودجه آن را تأمین کردند و مسئول کل مطالعه بودند؛ SVK و PK

منافع رقابتی

نویسندگان هیچ گونه تضاد منافی را اعلام نمی کنند.

اخلاق

این مطالعه توسط هیئت اخلاق پژوهشی دانشگاه آلبرتا (Pro00120852) تأیید شده است، از جمله اینکه نیازی به رضایت آگاهانه بیمار ندارد. این گزارش بر اساس دستورالعمل های «گزارش شفاف از یک مدل پیش بینی چند متغیره برای پیش آگهی یا تشخیص فردی مبتنی بر هوش مصنوعی» (TRIPOD-AI) ساختار یافته است.^{۲۳}

اطلاعات تکمیلی

اطلاعات تکمیلی نسخه آنلاین شامل مطالب تکمیلی است که در آدرس زیر موجود است.

<https://doi.org/10.1038/s41746-024-01130-8>

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

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

یادداشت ناشر اسپرینگر نیچر در مورد ادعاهای مربوط به صلاحیت قضایی در نقشه های منتشر شده و وابستگی های سازمانی بی طرف باقی می ماند.

دسترسی آزاد این مقاله تحت مجوز بین المللی Creative Commons Attribution 4.0 منتشر شده است که استفاده، اشتراک گذاری، اقتباس، توزیع و تکثیر در هر رسانه یا قالبی را مجاز می داند، مادامی که به نویسنده (گان) اصلی و منبع، اعتبار کافی داده شود، پیوندی به مجوز Creative Commons ارائه شود و در صورت ایجاد تغییرات، مشخص شود که آیا تغییراتی ایجاد شده است یا خیر. تصاویر یا 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 Creative Commons مقاله گنجانده نشده باشد و استفاده مورد نظر شما طبق مقررات قانونی مجاز نباشد یا از مجاز 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/>.

© نویسنده (گان) ۲۰۲۴

محاسبه کار دیول. ۴۴. <https://doi.org/10.22489/CinC.2017.065-469> (۲۰۱۷)

۳۲. هانون، ای وای و همکاران. تشخیص و طبقه بندی آریتمی در سطح متخصص قلب در الکتروکاردیوگرام های سریایی با استفاده از یک شبکه عصبی عمیق. نات. مد. ۲۵، ۶۵-۶۹ (۲۰۱۹).

۳۳. سان، دلیو و همکاران. بهبود تشخیص کووید-۱۹ مبتنی بر نوار قلب و پیش بینی مرگ و میر با استفاده از سوابق پزشکی قبل از همه گیری در مقیاس جمعیتی. در: سری های زمانی برای سلامت در NeurIPS. <https://doi.org/10.48550/arXiv.2211.10431> (۲۰۲۲).

۳۴. 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. در: تصویربرداری پزشکی با NeurIPS ترکیب می شود. <https://doi.org/10.48550/arXiv.2210.06291> (۲۰۲۲).

۳۵. دراموند، سی. و هولت، منحنی های هزینه RC: روشی بهبود یافته برای تجسم عملکرد طبقه بندی کننده. مآخ. یاد بگیر. ۹۵، ۶۵-۱۳۰ (۲۰۰۶).

۳۶. کوان، اچ. و همکاران. ارزیابی اعتبار داده های اداری ICD-9-CM و ICD-10 در ثبت شرایط بالینی در یک پایگاه داده منحصر به فرد با کدگذاری دوگانه. خدمات درمانی. پژوهش. ۱۴۴۱-۱۴۴۴، ۴۳ (۲۰۰۸).

۳۷. کوان، اچ. و همکاران. الگوریتم های کدگذاری برای تعریف بیماری های همراه در داده های اداری ICD-9-CM and ICD-10. مراقبت پزشکی. ۱۱۳۰، ۴۳-۱۱۳۹ (۲۰۰۵).

۳۸. چن، تی. و گسترین، سی. XGBoost: یک سیستم تقویت درختی مقیاس پذیر. در: مجموعه مقالات بیست و دومین کنفرانس بین المللی ACM SIGKDD در مورد کشف دانش و داده کاوی، 785-794 (انجمن ماشین های محاسباتی، 2016).

۳۹. پرشلت، ل. توقف زودهنگام - اما چه زمانی؟ در: شبکه های عصبی: ترفندهای حرفه ای؛ ویرایش دوم (ویراستاران: مونتاون، جی.، اور، جی.بی. و مولر، کی. -). ۵۳-۶۷ (اشپرینگر برلین هایدلبرگ، ۲۰۱۲).

۴۰. یودن، شاخص WJ برای رتبه بندی آزمون های تشخیصی. سرطان. ۳۲، ۳۵-۱۹۵۰ (۱۹۵۰).

۴۱. بریر، جی. دلیو. تأیید پیش بینی های بیان شده بر حسب احتمال. دوشنبه. هواشناسی. نسخه ۱.۷۸-۳ (۱۹۵۰).

۴۲. سلواراجو، آر. و همکاران. گراد-کم: توضیحات بصری از شبکه های عمیق از طریق محلی سازی مبتنی بر گرادین. در: کنفرانس بین المللی IEEE در مورد بینایی کامپیوتر (ICCV) 2017، شماره های 618 تا 626 (2017).

۴۳. مونز، KGM و همکاران. گزارش شفاف یک مدل پیش بینی چند متغیره برای پیش آگهی یا تشخیص فردی (TRIPOD): توضیح و شرح. آن. کارآموز. پزشکی. ۱۶۲، W1-W73 (2015).

تقدیرنامه ها

این مطالعه با کمک هزینه عملیاتی از موسسه تحقیقات سلامت کانادا (کمک هزینه شماره PJT-178158) تأمین مالی شده است. دکتر کائول دارای کرسی علوم جنسی و جنسیتی در موسسه تحقیقات سلامت کانادا و کرسی قلب و سگته مغزی در تحقیقات قلبی عروقی است. داده ها از انبار داده های سازمانی خدمات بهداشتی آلبرتا با پشتیبانی پلتفرم خدمات داده و تحقیقاتی AbSPORU استخراج شده است که توسط Alberta Innovates، CIHR، بنیاد بیمارستان دانشگاه، دانشگاه آلبرتا، دانشگاه کلگری و خدمات بهداشتی آلبرتا تأمین مالی می شود. تفسیر و نتیجه گیری های موجود در اینجا متعلق به محققان است.