

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

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^{۱,۲,۳,۶}، امیر سلیمانی^۴
 کوین یانین^{۲,۴}، جاستین ازکوبیتز^۵
 روپیندر ساندو^۶ و پادما کائنول^۷

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 ۲۶۸ نوار قلب (ECG) از ۷۷ بیمار بزرگسال مراجعه کننده به ۸۴ بخش اورژانس یا بیمارستان یا پاییزدار (UA)، انجام دادیم که حداقل یک نوار قلب ۱۲ لیدی از فوریه ۲۰۰۷ تا آوریل ۲۰۲۰ در آمریکا، کانادا انجام داده بودند و ۱۵ تشخیص قلبی عروقی را که توسط کدهای طبقه بندی بین المللی بیماری ها، ویرایش DDM (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 ناپاییزدار (UA)، انجارکتوس میوکارد با بالا رفتن قطعه (STEMI)، غیر (NSTEMI)، آمبولی ریوی (PE)، کاردیومیوپاتی هیپرتروفیک (HCM)، تنگ آئورت (AS)، پرولاپس دریچه میترال (MVP)، تنگ 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 (UA)، ۸۰-۹۰٪ (PTE، SVT، UA)، ۸۰٪ برای ۸ بیماری قلبی عروقی (PTV)، ۸۰-۹۰٪ (STEMI، VT، MVP، PHTN، AS، AF، HF) و ۹۰٪ AUROC برای ۴ تشخیص (AVB، HCM، MS، STEM) 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�行یت الکتریکی قلب تولید می شود. با این حال، تکنیک های استاندارد مورد استفاده پزشکان و الگوریتم های کامپیوترا 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 کامپیوترا 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 ویژت های مراقبت های حاد و معمولاً پیش از یک بار انجام می شود. تنها در ایالات متحده، سالانه پیش 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 مورفوولوژیکی و زمانی ...

اگرچه علوم کامپیوترا، دانشگاه آمریکا، ادمونتون، آمریکا، کانادا، VIGOUR کانادا، دانشگاه پزشکی، دانشگاه آمریکا، ادمونتون، آمریکا، کانادا، گروه پزشکی، دانشگاه آمریکا، ادمونتون، آمریکا، کانادا، موسسه قلب اسمیت، سیستم بیمارستانی مرکز پزشکی سدرز-سینای، لس آنجلس، کالیفرنیا، ایالات متحده آمریکا؛ این نویسندها به طور مساوی کمک کردند: سونیل واسو کالمادی، امیر سلیمانی، ایمیل paul@ualberta.ca.

عملکردیا AUROC برابر با ۶۸.۹٪. مدل ها برای همه تشخیص ها، به جز AUROC بالای ۷۶٪ داشتند: با AUROC کمتر از ۸۰٪ برای دو تشخیص (SVT، UA)، به ترتیب افزایشی؛ AUROC در محدوده ۸۰-۹۰٪ برای هشت تشخیص (ایست قلبی [CA]، NSTEMI، VT، [MVP]، پروپلیس درجه میترال [PHTN]، فشارخون ریوی [AS]، HF، AF، به ترتیب افزایشی؛ فشارخون ریوی [PHTN]، تنگ آئورت [AS]، HF، AF، به ترتیب افزایشی؛

AUROC برای چهار تشخیص (بلوک دهلیزی-بطنی [AVB])، کاردیومیوپاتی هپرتوروفیک [HCM]، تنگ درجه میترال [MS]، STEMI، به ترتیب افزایشی). مدل مربوط به فیبریلاسیون دهلیزی (AF) بالاترین سطح زیر منحنی دقت-بادآوری (AUPRC) (با امتیاز ۵۹.۲٪ امتیاز) را داشت و بیس از آن نارسایی قلبی (HF) با ۵۶.۱٪ (امتیاز F1: ۴۶.۶٪) و STEMI (F1: ۳۹.۲٪) با ۵۴.۳٪ AUPRC (امتیاز F1: ۳۹.۲٪) قرار داشتند.

مدل DL (را دیابی ECG، سن، جنس) برای اکثر تشخیص ها، به جز AVB، که هر دو مدل عملکرد مقایسه ای داشتند، عملکرد بهتری نسبت به مدل XGB (اندازه گیری های ECG، سن، جنس) داشت. مدل های DL با بهبود متوسط ۵.۲ درصدی MVP، AUPRC، با افزایش قابل توجه ۱۱.۸ درصدی برای MS، درصدی برای ۷.۳، ۷.۱ درصدی برای NSTEMI و ۷.۱ درصدی برای STEMI، عملکرد بهتری نسبت به مدل های داشتند. مقایسه فواصل اطمینان ۹۵ درصدی از نتایج بوت استرپ نشان داد که تفاوت های معناداری بین عملکرد مدل DL با ویژگی های (سن، جنس) در مقابل بدون ویژگی های برای همه تشخیص ها به جز PTE وجود دارد، که نشان می دهد ویژگی های سن و جنس می توانند پیشفرفت های کوچک اما قابل توجهی را به پیش بینی تشخیص اضافه کنند. به طور مشابه، نتایج بوت استرپ نشان داد که مدل های DL با دیابی ECG به تنهایی، برای تشخیص های غیر از HCM، VT و AVB، عملکرد بهتری نسبت به مدل های XGB با اندازه گیری های ECG، سن و جنس دارند.

عملکرد مدل مبتنی بر جنسیت مامدل DL را که با استفاده از (دیابی ECG، سن، جنس) به طور جداگانه برای افراد فورماتیست و موئیت در مجموعه مورد نظر آموزش دیده بود، ارزیابی کردیم و به طور کلی نتایج مشابهی یافتیم (شکل ۱).^{۱۳} (تصویر بالا) مدل ها در ۱۰ مورد از ۱۵ مورد، عملکرد نسبتاً بهتری برای مردان داشتند و میانگین افزایش AUROC آنها ۱.۰٪ بود. پنجمورد از این موارد - یعنی HF، VT، STEMI، PTE، از این موارد دانش از RVT، STEMI، PTE، AS - تفاوت های معنی داری را نشان دادند. ما بیشترین تفاوت را در VT یافتیم که در آن مدل برای مردان ۶.۴٪ در مقایسه با زنان (مردان: ۸۵.۱٪، زنان: ۷۸.۷٪) و از نظر AUPRC (مردان: ۳۷.۵٪، زنان: ۲۲.۶٪) ۱۴.۸٪ در مقابل، عملکرد بهتری داشت. در مقابل، عملکرد پیش بینی برای فیبریلاسیون دهلیزی (AF) در زنان به طور قابل توجهی ۱.۲٪ AUROC بهتر از مردان بود.

عملکرد مدل مبتنی بر حضور ضربان ساز به طور مشابه، ارزیابی مدل های DL (را دیابی 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، بالافزار متوسط سپار کم AUROC به میزان ۰.۲۵٪ با ECG های غیرفعال در مقایسه با بدون آن (شکل ۱).^{۱۳} باز هم، VT بیشترین تفاوت را نشان داد، که در آن عملکرد با حذف ECG های ضربان ساز، ۳.۲٪ AUROC و AUPRC ۵.۶٪ کاهش یافت. تشخیص دیگری که تفاوت قابل توجهی را در همان جهت نشان داد، AVB بود (۱.۶٪ افت AUROC، ۵.۷٪ کاهش در AUPRC).

اعتبارسنجی خروج از بیمارستان با یک بار بسترن مطالعه‌ها از ECG های نهیه شده از ۱۴ بیمارستان استفاده کرد. نکته قابل توجه این است که دو مورد از این بیمارستان ها، بیمارستان های مراقبت های عالی بودندکه بالاترین تعداد ECG (به ترتیب ۴۲ و ۴۸٪، ۰.۸۵٪ و ۰.۵۳٪) 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 های شکل تکمیلی (۱) انجام دادیم. عملکرد مدل ECG: DL: ECG، سن، جنس با اعتبارسنجی خروج از بیمارستان با نتایج گزارش شده در مجموعه کلی موارد عدم پذیرش (جدول تکمیلی)^{۱۳} قابل مقایسه بود. در مقایسه با نتایج اعتبارسنجی اولیه، میانگین عملکرد AUROC در ۱۵ شرایط، افزایش اندک ۱.۳٪ درصدی در اعتبارسنجی H1 اما کاهش ۱.۳٪ درصدی در اعتبارسنجی H2 را نشان داد.

ارزیابی تمام ECG holdout 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 ۵۰۰ مورد انتظار به دست آمده بود، ارزیابی کردیم. نتایج، همانطور که در بخش تکمیلی ذکر شده است

تفییرات نامحسوس، اما آموزنده ای که ممکن است نشان دهنده بیماری اولیه باشد. ظهور تحیلی های بادگیری عمیق (DL) فرصتی هیجان انگیز برای شناسایی الگوهای بالینی مرتبط اما «پنهان» در سیگنال های ECG و ارزیابی همزنان پر ایجاد تعاملی پیچیده از داده های بالینی ثبت شده معمول برای تشخیص ناهنجاری های مختلف قطبی عروقی ارائه هی دهد.^{۱۴}

استفاده از ECG بررسی من کردند، عمدتاً بر بیماری های قلبی متوجه بودند که من توانده از راحتی توسعه متخصصان پژوهشی بر اساس تغییرات مورفوولوژیکی در الگوهای ECG تفسیر شوند - مانند آریتمی ها (فیبریلاسیون دهلیزی [AF])، تاکی کاردی بطنی [VT]، تاکی کاردی بطنی فوق بطنی [SVT]^{۱۵} ایفارکتوس میکاراد با بالارفتن قطعه (STEMI) ST یا بدون (STEMI) (STEMI) یا شرایط بلوک قلب مانند بلوک های دهلیزی-طنی یا بلوک های شاخه ای از جمله بلوک شاخه چپ و بلوک شاخه راست اکرچه تعداد مدل های مبتنی بر بادگیری 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 های مانند پرولپس درجه میترال (MVP) تمرکز دارند. ایست قلبی (CA)^{۱۶} نارسایی قلبی (HF)^{۱۷}، امبویلی ریوی (PE)^{۱۸}، تنگ آئورت (AS)^{۱۹} و دیگر میترال (MS)^{۲۰} و کاردیومیوپاتی هپرتوروفیک (HCM)^{۲۱} علاوه بر این، در حالی که مطالعات موجود عمدتاً بر روی برچسب های فردی متمرکز بودند، هیچ تحقیق قبلی در زمینه توسعه یک سیستم پیش بینی کننده برای تشخیص همزنان این شرایط خاص انجام نشده است. فقدان مجموعه داده های پژوهشی بزرگ که از نظر بالینی با مجموعه گستردere ای از برچسب های تشخیص موجود عمدتاً بر روی نظارت، حاسیه نویس شده باشند، یک مشکل شناخته شده است و اعتبارسنجی های در مقایسازی بزرگ در مقیاس جمعیت برای نشان دادن قابلیت اعتماد برای پذیرش موقوفیت 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.

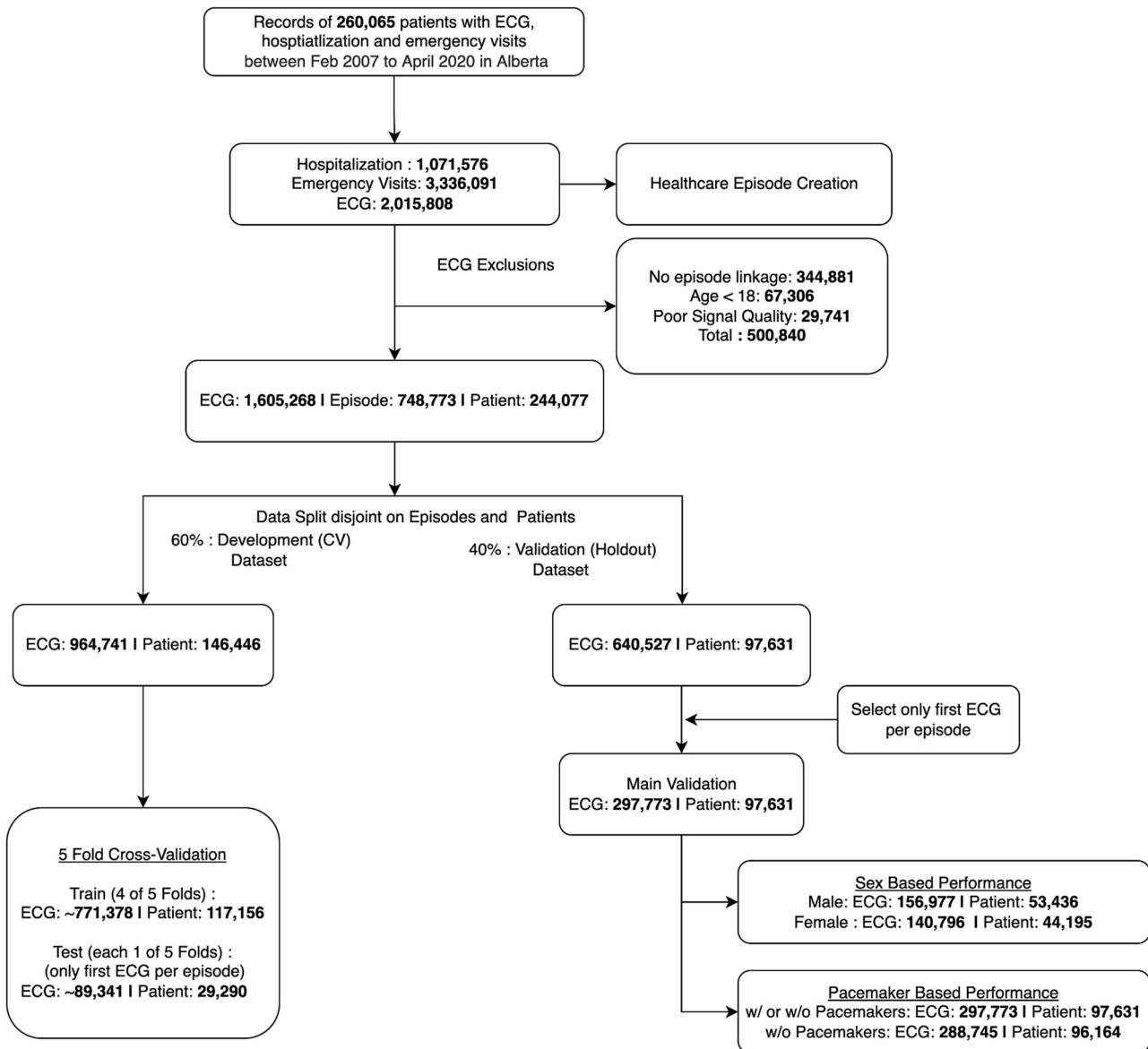
براین اساس، ما از یک گروه بزرگ از بیماران در سطح جمعیت، از یک سیستم بهداشت همانی تک پرداخت کننده، برای توسعه و اعتبارسنجی مدل های DL (مبتنی بر دیابی های دوازده گانه) و همچنین مدل های تقویت گردایان شدید (XGB) (مبتنی بر اندازه گیری های ECG که به طور معمول جمع آوری می شوند) استفاده کردیم تا به طور همزنان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

نتایج

ویژگی های پایه گروه قبلاً شرح داده شده است^{۲۰} طور خلاصه، میانگین سنی بیماران ۴۵.۸ ± ۱۷.۳ سال بود و ۷۵٪ 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. گروهی از ۹۷۶۳۱ بیمار (شکل ۱).^{۲۱} مجموعه داده های مربوط به افراد تحت آزمایش شامل ۳۴۳۶۴ مرد و ۴۱۹۵۶ 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 سازهای بعملکرد مدل ارزیابی شد. با پیش بینی احرای سیستم پیش بینی ما در محل مراقبت، مدل های خود را منحصرآ با استفاده از اولین نوار قلب هر بیمار بدون ضربان ساز در یک دوره خاص ارزیابی کردیم.

فرآوانی 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 در جدول اولایه شده است. اولین نوار قلب به ازای هر ای پیزو در مجموعه ۵ درجه های حذف شده که برای ارزیابی های نهایی استفاده شد (در مقایسه با داده های کامل نوار قلب، تفاوت هایی در برچسب های تشخیص داشت (مثلاً فرآوانی برای نارسایی قلبی: ۷۹.۳٪ 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 دیابی های ECG مقابله دون ویژگی های سن و جنس) و اندازه گیری های سن و جنس (برای شرایط ۱۵CV در شکل ۱ ارایه شده است.^{۲۲} جدول ۲ و جدول تکمیلی ۲. اعتبارسنجی مدل اصلی ما (DL: دیابی ECG، سن، جنس) نشان داد که مدل ما برای STEMI بهترین عملکرد را با منحنی مشخصه عملکرد گیرنده ۹۵.۵٪ (AUROC) داشت و مدل ما برای ترومیوامبوی ریوی (PTE) بدترین عملکرد را داشت.



هر اپیزود ثبت شده در طول ویزیت اورژانس یا بستری شدن در بیمارستان، که نشان دهندهٔ استقرار موردنظر در محل مراقبت است؛ دوم، روی تمام ECG‌های از مجموعهٔ مورد انتظار، علاوه بر این، ما عملکرد مدل های خود را در زیرگروه‌های خاص بیمار که بر اساس جنسیت وجود دستگاه‌های ضربان سازقلبی یا کمکی بطئ طبقه‌بندی شده بودند، ارزیابی کردیم.

شکل ۱ | فلوچارت طرح مطالعه که حجم نمونه‌ها را برای تقسیم بندی‌های مختلف آزمایش نشان می‌دهد. ما کل مجموعه داده‌های ECG را تقسیم کردیم، ۶٪ را برای توسعه مدل (شامل اعتبارسنجی مقابل داخلی پنج گانه برای آموزش و تنظیم دقیق) اختصاص دادیم و ۴٪ را به عنوان مجموعهٔ پشتیبان برای اعتبارسنجی نهایی کنار گذاشتیم. برای ارزیابی، مدل‌های خود را بادو رویکرد ارزیابی کردیم؛ اول، منحصراً روی اولین ECG‌ها از

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

ماتوانایی مدل DL چند برچسبی خود را برای پیش‌بینی مثبت بودن نوار قلب (ECG) برای برچسب ترکیبی، دوباره ارزیابی کردیم. نتایج، معادل ۷۳.۶۴٪ با امتیاز F1 معمادل ۴۷.۳۹٪ را نشان داد. ما همچنین یک مدل مشابه آموزش دادیم و عملکرد آن را در همان مجموعه داده‌های محدود ارزیابی کردیم. نتایج، معادل PPV ۵۷.۹٪ با امتیاز F1 معادل ۶۳.۰۳٪ را نشان داد. این نتایج نشان می‌دهد که هنگام غربالگری پیامد ترکیبی، می‌توان PPV بالاتری را به دست آورد.

توضیحات مدل

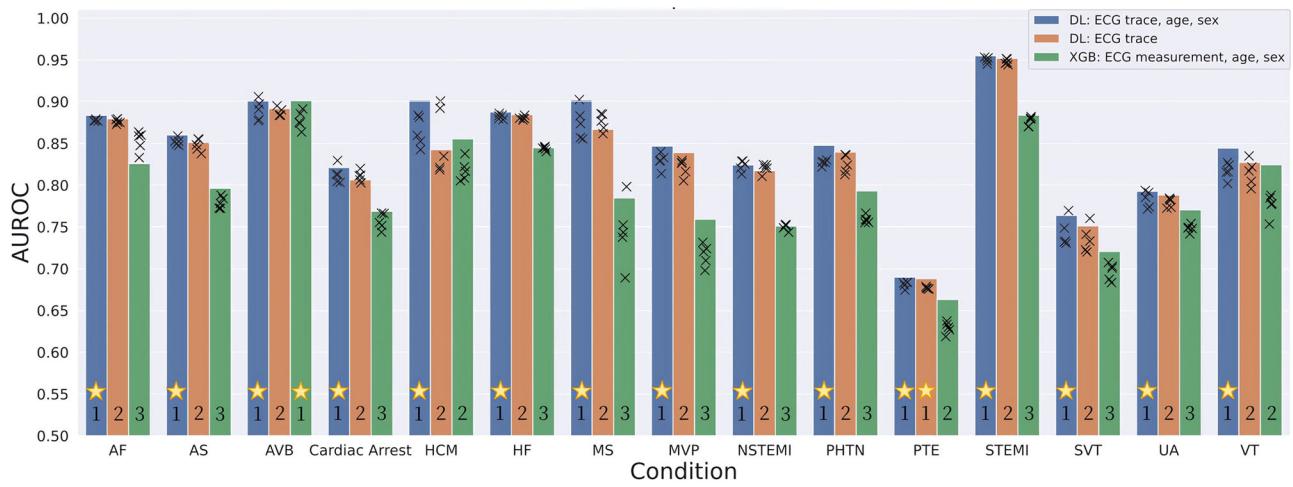
شکل ۴^۲ نتایج GradCAM را نشان می‌دهد و حوزه‌هایی از ECG را که سهم و ارتباط‌بیشتری با پیش‌بینی مدل از شرایط مختلف CV دارند، برگسته می‌کند (برای فهرست کامل همه موارد، به شکل تکمیلی ۲ مراجعه کنید).

جدول ۴، عملکردی را نشان می‌دهد که یا برتر یا قابل مقایسه با عملکرد حاصل از استفاده صرف از ECG های اولیه است. ارزیابی تمام ECG، نوساناتی را در نمرات AUROC و AUPRC نشان داد. کاهش کلی ۲.۰۴ درصدی در میانگین AUROC (میانگین در ۱۵ شرایط) و افزایش همざمان ۲.۱۵ درصدی در AUPRC برای ارزیابی تمام ECG وجود داشت. نکته قابل توجه این است که نمرات F1 برای همه برچسب‌ها در ارزیابی تمام ECG، بهبودهایی از ۱۰.۰۹٪ تا ۱۶.۸۱٪ داشت. با افزایش متوسط ۶.۲۸٪ را نشان داد. به طور مشابه، مقادیر پیش‌بینی مثبت (PPV یا دقت) برای همه برچسب‌ها افزایشی از ۰٪ تا ۵۶٪ داشت. با بهبود متوسط ۴.۹۸٪ را نشان داد. بنابراین، می‌توان انتظار داشت که این الگوریتم‌ها در صورت اعمال بر روی ECG‌های انجام شده در هر نقطه‌ای از دوره مراقبت، عملکرد قابل مقایسه‌ای نگوییم برتر، را نشان دهند.

ارزیابی برچسب کامبوزیت شیوع چندین مورد از تشخیص‌های مورد نظر در نمونه‌ها پایین بود (مثلاً ۰.۰۹٪ برای MS در بین اولین ECG‌ها)، که احتمالاً بر PPV تأثیر می‌گذارد. ما

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

اطلاعات کامل (ن=۱۶۵،۲۶۸)	مجموعه نگهدارنده (ن=۴۳،۵۲۷)	مجموعه توسعه (ن=۹۶۷،۳۱)	holdout (ن=۳۹۷،۷۷۳)
انفارکتوس میوکارد بدون بالا رفتن ST	(۰.۱۱٪) (۴۲،۲۷۴)	(۰.۱۰۴٪) (۹۶،۸۲۸)	(۰.۱۰۵٪) (۴۵،۴۴۶)
نارسایی قلبی	(۰.۱۵٪) (۴۹،۳۲۵)	(۰.۱۰۲٪) (۴۵،۴۴۶)	(۰.۱۰۶٪) (۴۱،۵۵۷)
آنژین ناپایدار	(۰.۲۷٪) (۴۹،۳۲۳)	(۰.۱۸٪) (۱۰۱،۹۵۵)	(۰.۱۱۵٪) (۳۴،۱۳۹)
فیریلاسیون دهلیزی	(۰.۲۷٪) (۴۳،۴۶)	(۰.۱۸٪) (۱۰۱،۹۵۵)	(۰.۱۱۴٪) (۳۴،۱۳۹)
تاکی کاردی بطنی	(۰.۱۸٪) (۴۹،۶۷۳)	(۰.۱۰۸٪) (۱۰۱،۹۵۵)	(۰.۱۰۵٪) (۴۱،۵۵۷)
ایست قلبی	(۰.۲۰٪) (۴۰،۵۰۵)	(۰.۲۰٪) (۴۲،۲۷۴)	(۰.۰۹٪) (۳۸،۴۴۳)
تاکی کاردی فوق بطنی	(۰.۱۵٪) (۴۹،۳۲۵)	(۰.۱۰۴٪) (۹۶،۸۲۸)	(۰.۰۷٪) (۲۲،۲۲۳)
بلوک دهلیزی بطنی	(۰.۲۰٪) (۴۰،۰۱۳)	(۰.۲۰٪) (۴۳،۴۶)	(۰.۰۹٪) (۲۹،۹۴۳)
آمبولریوی	(۰.۲۰٪) (۴۰،۲۸۵)	(۰.۲۰٪) (۴۱،۵۵۷)	(۰.۱۶٪) (۴۷،۶۳)
تنگ آئورت	(۰.۱۸٪) (۴۰،۱۲۰)	(۰.۱۰۸٪) (۱۰۱،۹۵۵)	(۰.۱۰۸٪) (۳۴،۱۳۹)
فشارخون رویی	(۰.۲۶٪) (۴۶،۳۳۱)	(۰.۲۲٪) (۲۱،۸۶۹)	(۰.۱۳٪) (۴۱،۱۷)
کاردیومیوپاتی هیپرتروفیک	(۰.۲۰٪) (۴۳،۴۸۵)	(۰.۲۰٪) (۴۱،۵۵۷)	(۰.۰۱٪) (۴۰،۰۹)
افتادگی درجه میترال	(۰.۱۵٪) (۴۳،۴۸۱)	(۰.۱۰۴٪) (۱۰۱،۹۵۵)	(۰.۰۶٪) (۱۹،۱۴۹)
تنگ دریچه میترال	(۰.۱۸٪) (۴۱،۲۹۵)	(۰.۱۰۸٪) (۱۰۱،۹۵۵)	(۰.۰۹٪) (۳۷،۷۷)



شکل ۲ | مقایسه عملکرد مدل AUROC برای مدل های DL و XGB با ردیابی های ECG با بدون ویژگی های سن و جنس و اندازه گیری ها برای ۱۵ بیماری قلبی عروقی ارتفاع میله هاشان دهنده عملکرد در اعتبارسنجی خارجی و ضربدرها نشان دهنده عملکرد در هر یکاز اعتعارسنجی های مقاطعه پنج گانه است. برای هر شرایط، مدل ها بر اساس اعمالکرد دهنده عملکرد های آماری مشابه، رتبه پندی مساوی دارند و مدل که بالاترین عملکرد را دارد با یک ستاره مشخص شده است. فیریلاسیون دهلیزی AF, AS

قابل در مطالعات پیش بینی مبتنی بر ECG برسی نشده بود و نشان داد که 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 من کنند.

مطالعات قبلی با استفاده از تشخیص ECG با قابلیت هوش مصنوعی نشان داد که مدل های ML و DL می توانند ریتم ECG و ناهنجاری های مورفوولوژیک 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 لید ساده و آسان برای بدست آوردن، نه تنها بیماری های قلبی عروقی، بلکه اختلالات را که به طور معمول با استفاده از نوار قلب تشخیص داده نمی شوند، به طور دقیق پیش بینی کنند. این مدل ها ابرازهای بالقوه ای برای غربالگری روده هنگام بیماری های قلبی عروقی، به ویژه موادی هستند که با قابل توجهی را بر سیستم مراقبت های بهداشتی وارد می کنند و ممکن است به شناسایی دقیق تر بیماری های بالینی کم کنند.

۵ انتخیصی). قابل ذکر است که مناطقی که بیشترین سهم را در تشخیص داشتند عبارتند از: 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 ST در AVB و ناییه قطعه در HF، امواج p در AS، امواج QRS در SVT، امواج QT در XGB را بر اساس اندازه گیری های نشان من دهد. که افزایش اطلاعات قابل توجهی را با مدت زمان P برای پیش بینی AF، مدت زمان QRS برای RR، فاصله QT برای UA، فاصله RR برای AVB، محور T پیشانی برای HF، محور T افقی برای NSTEMI، فاصله QT تصویح شده با نرخ Bazett برای CA و غیره نشان می دهد.

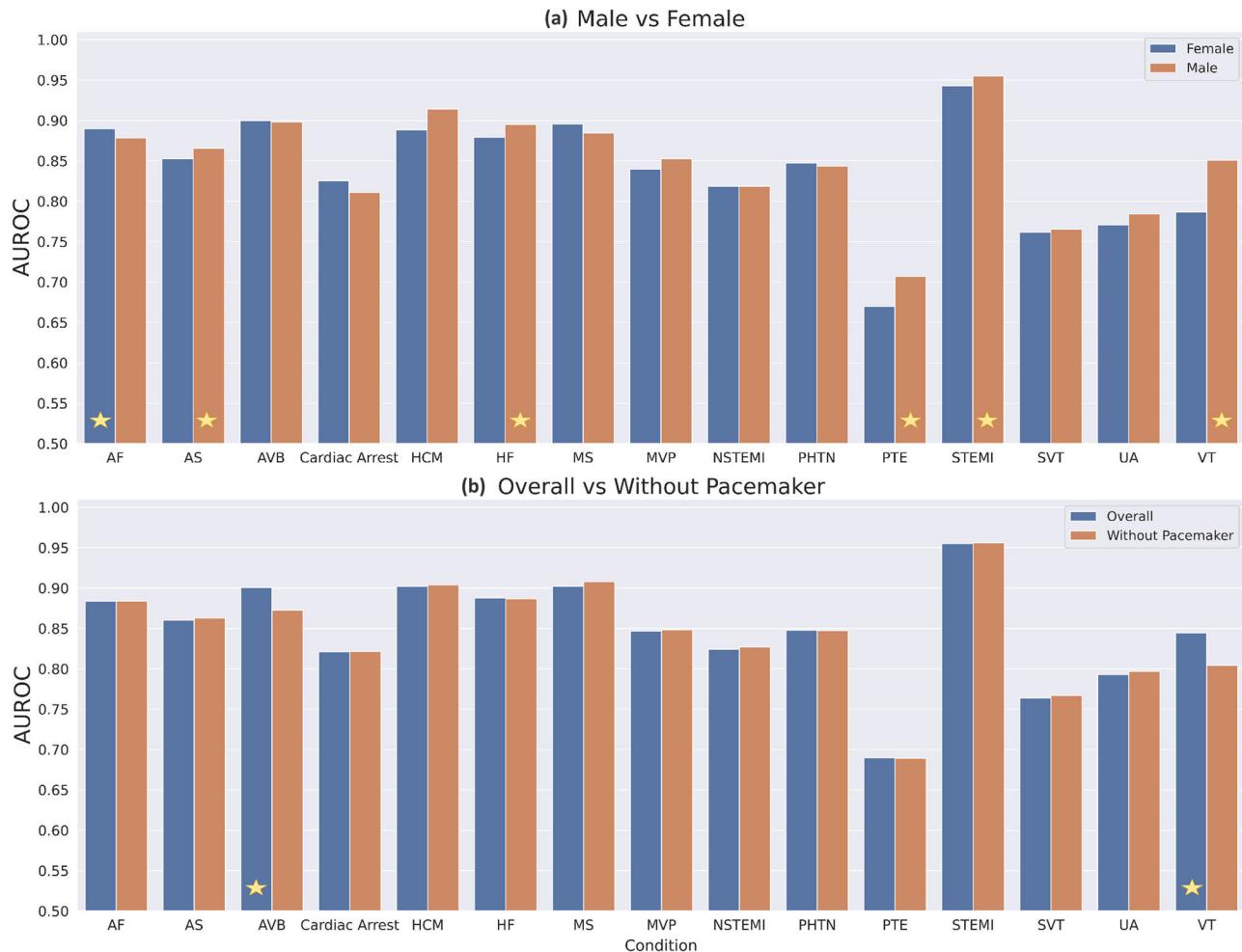
بحث

در این مطالعه بزرگ در سطح جمعیت با پرونده های سلامت اداری مرتب شامل میلیون هانوار 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.

شرايط فهم ۷۶) هم فرم از همه، استفاده از این سیستم های طبقه بندی خودگار من تواند دسترسی به مراقبت را در هنرمندانه ای از افرادهایی که نسبت پس از حدودی که مخصوصاً این افراد را محدود نمایند، ممکن نماید. حقیقات ۷- میشتری از زمان است مخصوص شود که آنچه از توافق این افراد باشد که خودگار مبتنی بتواری قلوب بخوبی مدیریت زده نگاه و پیشگیری از پیشرفت بیماری و ارائه مراقبت مقرر و به طرفه استفاده کرد یا خیر.

مطالعه‌ها محدودیت‌های دارد که نیاز به بحث پیشتر پردازد. اول همه GECG‌ها نوسودستگاه هیلیس یا لیک تپیدکننده (سیستم اطلاعاتی هیلیس) کاربرد نیافرند و ممکن است عودهمین پیری و ویژگی‌های باقیه های ECG سایر سیستم‌ها را محدود کنند. همچومن اندیشه هایی که در اینجا در خصوص درجه دل های XGB از طریق دستگاه های خارجی می‌باشند و توسعه آنها یک‌نگاه اصلی خوانده یا توسعه نیافرند. مخصوصاً ایشان تنظیم‌کننده بودند. سوم، بیجسی‌هایها از کدهای ICD شیفت‌شده در پرونده اوزرانش و بسترهای گرفته شده بودند. با توجه به ماهیت جمع آوری داده هار پرونده های پرسک اداری، زمان دقیق بروز بیماری در طول یک دوره کامل مراقبت های پیش از اینم تعلق به طبقه قطعی مشخص نیست. بیان توجه در مورد اداری که یک تیغه‌واری ایجاد می‌کند از جمله این ECG های خطا به مری شهد و بخطفه پیش بینی می‌نماید که اینها ممکن است مبتلایان بیماری را تشخیص دهند. جایی بیشتر تشخیص بیکاری موجو در تفسیر شود. این تهاوارها غالباً توجه ایستاد زنگ تشخیص زودهنگام ردمیریت پالپیت ارتشید ایست و پیشتر هایی که ممکن هستند اینها را ایمنی در آینده زدیک را ایجاد نمی‌کنند که تواند به همراه انتزاع برای مراقبت های بشتر می‌گذرد. چهارم، آزمایش داخلی، حتی در مقایسه قابل توجه، ممکن است در مقایسه با عبارت‌سنجه های خارجی پیش از اینم نظر نظر گرفته شون اینها امکانات ایمنی دلیل است.

عبدالرسنجی های خارجی توانده به نظر کوچک شود این مجموعه به این دلیل است که سوکری هایی که از آنچه یک سیستم سلامتی وارد چشم می شوند باید لیکن شناخت های جمعیت بیمار، تجهیزات روشی های تولید جسمی و تهییر عوامل اجتماعی و اقتصادی باشد. متأسفانه متوسلیم انتقالی استخراجی برای مغول های چینی جنس خود را رای دهیم زیرا هیچ مجموعه داشته ECG خارجی همانی که پس از این جنگ ب عملکرد اعتبر انسنجی ICD-10 انتخاب شده مرتبط باشند وجود ندارد. بالین حال، سن، جنس ماست که انتخاب مدل های leave-one-hospital-out از انتخاب مدل های DL: ECG بر اساس یک کوچک اتفاقی ایستگران مرتعه کنندگان باید بخوبی های فورانی شوند. یعنی در میزان بیمارستان های اچیزیتی شووند و منقوص شوند و تشوییض های هورمونی ایستگیری در میزان مثبت بررسی های پنهانی از توجهی درجه دهنده که جراحتیش بین های براخ از شخصیت هایی متفاوت باشند. همچنانکه در اینجا مثبت بررسی های براخی در میزان انتخاب مدل های ایستگرانی که در آنها مبتدا از این دلیل هایی دستگیری شوند هدف ما این است که در این مجموعه داده های دیسپلیشن های بروندگانشکی لکترونیکی مستقر کیم. علاوه بر این، مجموعه داده های اموزش ما که تقریباً یک میلیون ECG دارد، تعداد کافی از موارد مثبت را برای توسعه مدل های پیش بینی گندم مؤثر داشته است. با این حال، برای این برجسته های قدر که این ممکن است PPV پایین تر از حد مطلوب را نشان دهد، که مستلزم بررسی دقیق در مورد واحد شرایط بودن آنها برای ارزیابی بالینی است.



بدون ضربان سازها و تمام ECG ها در مجموعه‌ی نگه دارنده (ب). ارتفاع میله نشان دهنده ممکن‌کرد در اعتیارستنجی خارجی است و مدل هایی که از نظر آماری عملکرد بالاتری دارند با ستاره مشخص شده‌اند.

شکل ۳ | مقایسه عملکرد AUROC برای مدل ECG، سن، جنس برای ۱۵ بیماری کلی 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)

داده‌های نوار قلب (ECG) با استفاده از یک شماره سلامت بیمار منحصر به‌فرد به پایگاه‌های داده سلامت اداری زیر متصل شدند: (1) پایگاه داده چکیده ترخیص (DAD) حاوی داده‌های مربوط به بسترهای شدن بیماران؛ (2) پایگاه داده سیستم‌گزارش دهی مراقبت‌های سرپایی ملی (NACRS) شامل تمام ویزیت‌های سرپایی بیمارستانی و اورژانس؛ و (3) ثبت طرح بیمه مراقبت‌های بهداشتی آبرتا (AHCIP) که اطلاعات جمعیت شناختی را ارائه می‌دهد.

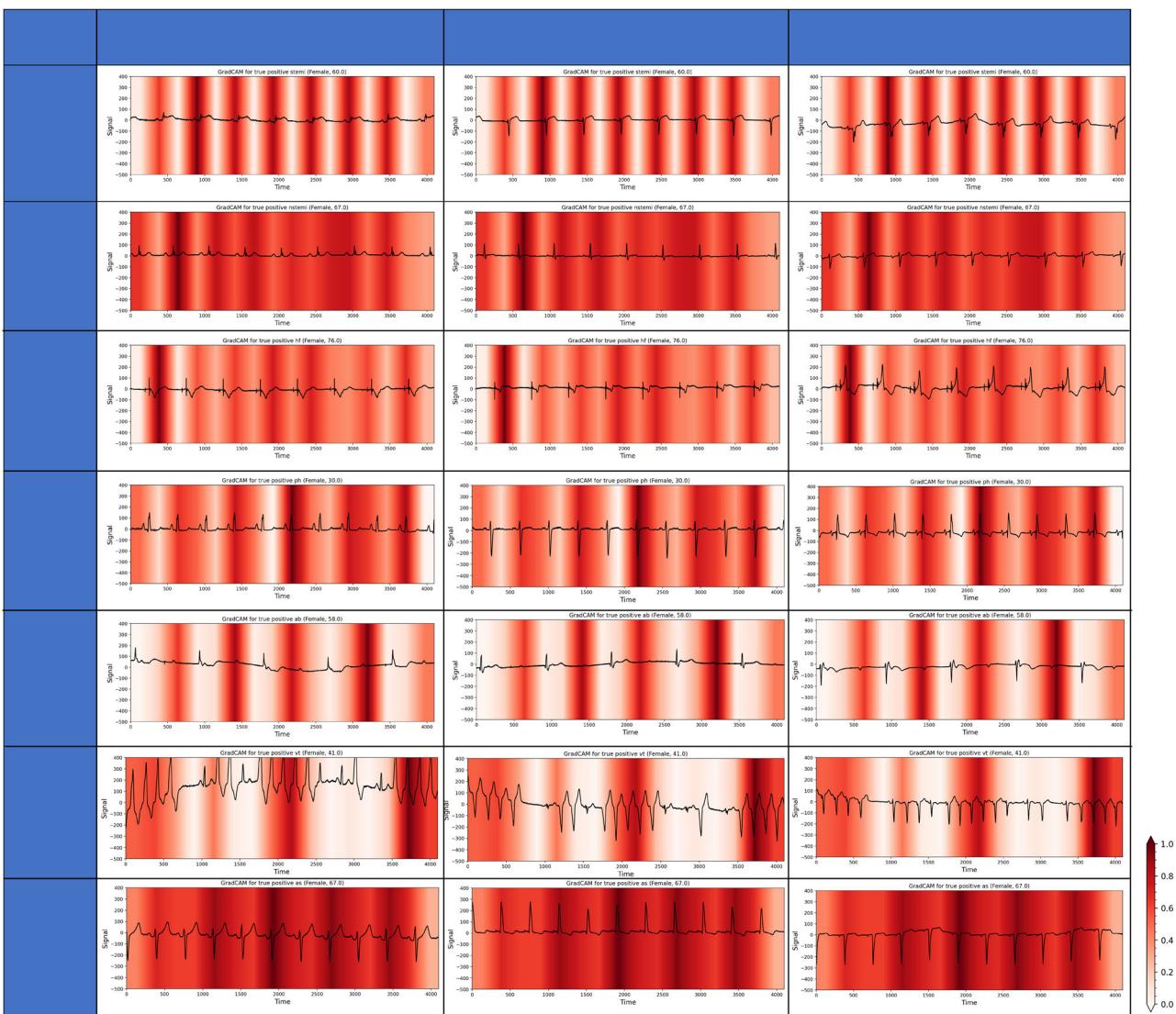
استقرار. ششم، ما مدل‌های خود را در درجه اول با استفاده از AUROC با C-index معیار را در ادبیات زست پژوهی، ارزیابی کردیم. با این حال، این معیار محدودیت‌هایی دارد و هزینه‌های طبقه‌بندی نادرست نتایج مثبت کاذب یا منفی کاذب را در نظر نمی‌گیرد. برای استقرار مدل، ارزیابی‌های سفارشی با هدف به حداقل رساندن هزینه‌های مورد انتظار، که هزینه‌های طبقه‌بندی نادرست و سایر عوامل تخصیص منابع را در نظر می‌گیرند، باید در اولویت قرار گیرند.^{۲۰}علاوه بر این، علیرغم ماهیت جعبه‌سیاه برخی از رویکردهای یارگیری 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.

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

روش ها
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.

نوار قلب (ECG) موردنظر مطالعه قبل از شرح داده شده است.^{۲۱} به طور خلاصه، بیمارانی که بین ۲۰۰-۷۰۰ تا آوریل ۲۰۲۰ در ۱۴ مرکز در آلبرتا، کانادا بسترهای شدن و شامل معادل انحراف (ST)، محور T پیشانی، محور موج ST افقی، محور T افقی، شروع موج Q، فاصله QT اصلاح شده با نرخ Fridericia، QT، فاصله QT، اصلاح شده با نرخ Bazett بود.

گروههای موردنظر مطالعه قبلاً شرح داده شده است.^{۲۲} به طور خلاصه، بیمارانی که بین ۲۰۰-۷۰۰ تا آوریل ۲۰۲۰ در ۱۴ مرکز در آلبرتا، کانادا بسترهای شدن و شامل



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

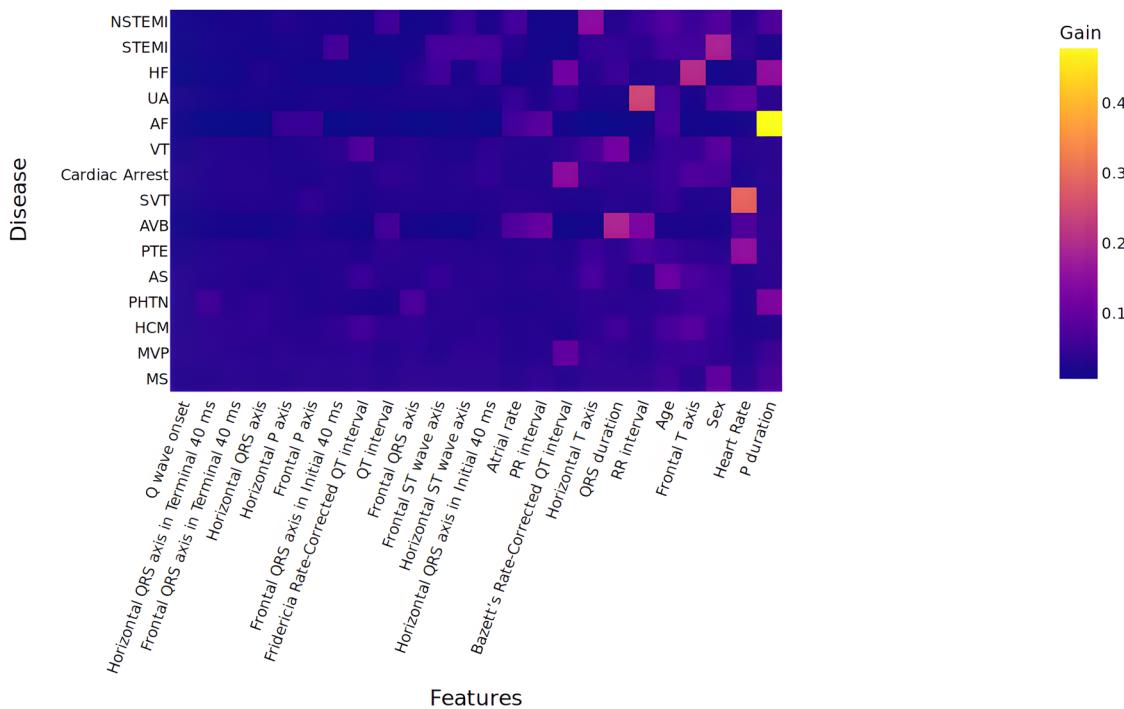
۷۶-۵۷،۰۷۱ امورد بسترنی در بیمارستان برای ۶۵،۰۶۴ بیمار. ملاقات های همزمان با مرابت های بهداشتی (مراجه به اورژانس و یا بسترنی در بیمارستان) که برای یک بیمارادر یک دوره ۴۸ ساعته رخ داده بود، به عنوان انتقال و بخش از یک دوره مرابت های بهداشتی در نظر گرفته شد. اگر تاریخ ثبت نوار قلب در بازه زمانی بین تاریخ پذیرش و تاریخ ترجیص یک دوره بود، یک رکورد نوار قلب به یک دوره مرابت های بهداشتی مرتبط می شد. پس از حذف نوار قلب هایی که نمی توانستند به 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 من شدند)، گروه تجزیه و تحلیل ما شامل ۳۶۵،۰۶۱ نوار قلب از ۷۷۳،۸۷۴ دوره در ۷۷ بیماران بود (شکا...).

وظایف پیش بینی مامدل 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 مریبوطه، و در این دهم (ICD) شناساری، شدند.

بیماران متبل به ST ایمید پایدار، تمیلیس های QRS در AS، امواج ST در AVB و ناحیه قطعه HF بیشترین سرورنگی را در تشخیص هر بیماری داشتند. تنگ آئورت AS، بلوك دهیکار-بطی، AVB، یادگیری عیقیم DL، نوار قلب، نارسایی قلبی HF، انفارکتوس میوکار-دینون، NSTEMI، انفارکتوس میوکارد با صعود قطعه ST در STEMI، فشار خون رویی PHTN، تاکی کاردی بطي 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 مخصوص شده است.^{۲۰} اگر نوار قلب (ECG) در طول یک دوره اورژانس یا بسترنی انجام شده است، 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 مطالعه‌گنجانده شدند تا اثربخشی مدل های ما در تشخیص شرایط قابل تشخیص باشد.

هدف مدل پیش بینی، خروجی دادن احتمالات کالایبره شده برای هر کی از ۱۵ 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 وزیرت تک بیمار ممکن است شامل چندین ECG باشد. هنگام آموختن مدل، ما از تمام ECG های چندگاهه متعلق به یک دوره شامل شدند) در مجموعه آموختن/توسعه برای به حداقل 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 سمار به دست می آمد، به کار گرفته شود.



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

شکل ۵ | 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

داده های مربوط به یک بیمار بین داده های توسعه و ارزیابی یا بین بخش های آموخته از مایباش اعتبارسنجی مقابله داشتند. این بخش از داده های اشتراک گذاشته شدند. همانطور که قبل از داده های آنچنان که انتظار داریم سنتاریوی 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 بستره شدن در بیمارستان گرفته شده بود، ارزیابی کردیم. تعداد نوار قلب، اپیزودها و بیماران مورد استفاده در داده های کلی و در تقسیم بندی های تحریجی در شکل ارائه شده است.^{۱۰} حدول تکمیلی ۵. علاوه بر ارزیابی اولیه، ما آزمایش خود را گسترش می دهیم تا تمام ECG های موجود در مجموعه دی مورد نظر را نیز شامل شود، تا تطبیق بذیری 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 آموخته و آزمایش، ECG های بیماران بستره شده در هر دو بیمارستان آموخته و آزمایش 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 اغام ۱۵ برچسب بیماری در یک پیش بینی ترکیبی ارائه می دهیم و در نتیجه بازده تشخیصی را افزایش می دهیم.^{۱۱}

مامساحت زیر منحنی مشخصه عملکرد گیرنده AUROC، معادل شاخص C) و مامساحت زیر منحنی دقت-فراخوان (AUPRC) را گزارش کردیم. همچنین پس از بازبینی کردن احتمالات پیش بینی به کلاس های تشخیص/عدم تشخیص با استفاده از نقاط برتر بهینه مشتق شده از مجموعه آموخته شاخص بودن، امتیاز F1، ویژگی، فراخوان، دقت (معادل PPV) و دقت را تولید کردیم.^{۱۲} ما همچنین از معیار 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 دیفیریلاتورهای قلبی که ICD-[II]) 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).^{۱۳} تشخیص

در طول ویزیت اورژانس یا بستره شدن در بیمارستان (برای جزئیات بیشتر به بخش «ازبیابی» در زیر مراجعه کنید). ماز 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.^{۱۴} برای تعیین اینکه آیا ویژگی های جمعیت شناختی (سن و جنس) ارزش پیش بینی کننده افزایشی به عملکرد مدل های آموخته دیده فقط بر اساس ECG اضافه می شوند، ما مدل ها را به روش زیر توسعه داده و گزارش کردیم: (الف) فقط ECG (DL: ریدیابی ECG؛ سن: سن)؛ (ب) ECG + سن، جنس (DL: ریدیابی ECG، سن، جنس [که مدل اصلی ارائه شده در این مطالعه است])؛ و (ج) XGB (ECG؛ اندازه گیری ECG، سن، جنس).

الگوریتم های یادگیری
ماز یک روش طبقه بندی چند برچسب 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 جدول ساختاریافته بود، ما گروه های درخت تقویت شده با گرادیان (XGB) را آموخت دادیم.^{۱۵} در حالی که ما از شبکه های عصبی کانولوشن عمیق برای مدل هایی با ریدیابی سری های ولتاژ-زمان ECG استفاده کردیم، برای هر دو مدل XGB و DL، ما از ۹۰٪ داده های آموخت شی برای آموخت مدل استفاده کردیم. برای این مدل ابتدا «توافق زودهنگام» فرآیند آموخت، برای کاهش احتمال بیش برآش، استفاده کردیم.^{۱۶} برای یادگیری عمیق، ما یک مدل ResNet واحد را برای یک کار چند کلاسه و چند برچسبی یاد گرفتیم. اکه هر سیگنال ECG را به ۱۵ مقدارگاشتن می کنند، مربوط به احتمال حضور هر یک از ۱۵ تشخیص است. از سوی دیگر، برای تقویت گرادیان، ما ۱۵ مدل XGB دودویی گرادیان، ما ۱۵ مدل سیگنال ECG را به احتمال یکی از برچسب های منفرد نگاشتند. جزئیات روش شناختی پیاده سازی های مدل XGB و DL و ما قبل از شرح داده شده است.^{۱۷}

ازبیابی و تجسم
طراجی ارزیابی: ما از تقسیم بندی ۶/۰-۴/۰ روی داده ها برای آموخت و ارزیابی استفاده کردیم. ما مجموعه داده های کلی ECG را به تقسیم بندی های تصادفی ۶٪ برای تقویت مدل (که از اعتبارسنجی مقابله داخل پنج گانه برای آموخت و تنظیم دقیق مدل های نهایی استفاده کرد) و ۴٪ باقی مانده را به عنوان مجموعه نگه دارنده برای اعتبارسنجی خارجی نهایی تقسیم کردیم. ما اطمینان حاصل کردیم که اینها از

۹. الزایتی، اس. اس و همکاران. یادگیری ماشین برای تشخیص ECG و طبقه بندی خطرناک توسعه میوکارد انسدادی. نات. مد. ۲۹، ۱۸۰۴-۱۸۱۳ (۱۴۳-۱۱۱).

۱۰. ریبرو، ای ج و همکاران. تشخیص خودکار نوار قلب ۱۲ لیدی با استفاده از یک شبکه عصبی عمیق. نات. کمون. ۱۷۶، ۱۷۰ (۲۰۲).

۱۱. ایاسن، آی. و همکاران. یک معماری قوی یادگیری ماشین برای تحلیل ریتم ECG قابل اعتماد بر طول CPR. کنفرانس مجموعه مقاالت IEEE مهندسی پزشکی زیست شناسی اجتماعی ۱۹، ۱۹-۱۹ (۲۰۱۹).

۱۲. الولا، آ. و همکاران. شبکه های عصبی عمیق برای تشخیص نبض میتنی بر نوار قلب در هنگام ایست قلبی خارج از بیمارستان. آنترورپی ۳۰، ۵-۲ (۲۰۱۹).

۱۳. چوی، ج. و همکاران. یادگیری عمیق شکل موج های ECG برای تشخیص نارسایی قلبی با کسر تخلیه بطن چپ کاهش یافته نماینده علمی ۱۱، ۱۴۲۳۵ (۰۲۲).

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. بورو. قلب جی. دیجیت هلت ۶۶-۵۶، ۳ (۲۰۲۲).

۱۶. والته سیلو، ب.. مارکز، ج. نوبره منس، م. اولیویرا، آ. و پینتو، اف. جی. تشخیص آمبولی ریوی حاد میتنی بر هوش مصنوعی: توسعه یک مدل یادگیری ماشین با استفاده از الکتروکاردیوگرام ۱۲ لید. کشیش پورت. کاردیول. ۶۵-۶۴۳، ۴۲ (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 مقاالت IEEE مهندسی پزشکی زیست شناسی اجتماعی ۱۵۵۱-۱۵۴۸، ۲۲ (۲۰۲۰).

۱۸. گوتون، اس. و همکاران. رویکرد یادگیری فدرال چندملیتی برای آموزش مدل های ECG و اکوا کاردیوگرام برای تشخیص کاردیومیوپاتی هیپرتروفیک. گردش خون ۷۸۹-۷۵۵، ۱۴۶ (۲۰۲۲).

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

۲۰. اولو-سرنا، ای و همکاران. ECHOMend: 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. اکوا کاردیوگرافی گردش خون ۴۷-۳۶، ۱۴۶ (۲۰۲۳).

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

۲۲. لیو، س.-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. آسیاACC ۲۷-۲۵۸، ۲۷ (۲۰۲۲).

۲۳. چن، ال.. فو، جی. و جیانگ، سی. پیش بینی ژنتیک مبتنی بر الکتروکاردیوگرام الاید متشتق شده از یادگیری عمیق برای کاردیومیوپاتی هیپرتروفیک: یک مطالعه آزمایشی. آن. مدی. ۲۳۳۵۵۶۴، ۵۵ (۲۰۲۳).

۲۴. کو، دبلیو.-اوی. و همکاران. تشخیص کاردیومیوپاتی هیپرتروفیک با استفاده از الکتروکاردیوگرام مجهز به شبکه عصبی کانولوشن. مجله آمریکاین کاردیولوژی. ۷۷-۷۲، ۷۲ (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 الکتروکاردیوگرامها. مجله پزشکی آسیاACC ۲۱-۲۰، NPDigit (۲۰۲۳).

۲۶. لیو، ایکس.. وانگ، اچ. لی، زد. و کین، ال. یادگیری عمیق در تشخیص نوار قلب: یک بررسی. سیستم مبتنی بر دانش ۱-۱۷۸۷، ۲۲۷ (۲۰۲۱).

۲۷. مانت، جی. و همکاران. دقت تشخیص فیریلایسیون دهیزی بر روی نوار قلب توسط پزشکان مراقبت های اولیه و نرم افزار تشخیصی تقسیری: تجزیه و تحلیل داده های حاصل از آزمایش غربالگری فیریلایسیون دهیزی در سالمدان (SAFE). بی ام جی ۳۳۵-۳۸۰ (۲۰۰۷).

۲۸. ورونس، جی. و همکاران. دقت پزشک اورژانس در تقسیر الکتروکاردیوگرام ها با اختلال اනفارکتوس میوکارد با بالا رفتن قطعه ST: آیا کافی است؟ کارت حاد. مراقبت ۱۷، ۱-۷ (۲۰۱۶).

۲۹. تران، دی تی و همکاران. بار مالی فعلی و آبیده بذیرش بیمارستانی به دلیل نارسایی قلبی در کانادا: یک تحلیل هزینه. باز CMAJ ۴۰، E365-E370 (۲۰۱۶).

۳۰. سومانی، س. و همکاران. یادگیری عمیق و الکتروکاردیوگرام: مروری بر جدیدترین فناوری های موجود. اردوی ۲۳۳-۱۱۹۱ (۲۰۲۱).

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

وکدهای رویه مورد استفاده برای شناسایی وجود ضربان سازها در جدول تکمیلی ۷ ارایه شده است.

مقایسه مدل ها: برای هر ارزیابی، ما عملکردها را از اعتبار سنجی مقابل داخلی پنج گانه و همچنین عملکرد نهایی در مجموعه داده های پشتیبانی، 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 پشتیبانی با جایگزینی به صورت جفتی مقایسه شدند تا در مجموع 10000 نکرار بوت استرپ از تفاوت های جفتی در AUROC ایجاد شود. یعنی، هر کدام بدون ضربان ساز در مقابل مدل اصلی مقایسه می شوند. گفته ای شود که تفاوت در عملکرد مدل از نظر آماری معنی دار است اگر فوایل اطمینان 95٪ از میانگین تفاوت های جفتی در AUROC ها شامل مقدار صفر برای مدل های مقایسه شده نباشد.

تصویرسازی 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 گرادیان را که در پیش بینی مدل از تشخیص در مدل های DL ما نقش داشتند، با استفاده از نگاشت فعال سازی کلاس با و گرادیان (GradCAM) تصویرسازی کردیم.²² 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.

در دسترس 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، گنجانده ایم. انتظار نمی رود که آنها سیگنال های ECG واقعی با توزیع های برچسب را به طور دقیق نشان دهند. مجموعه داده از مایش 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 ارج
است: share.com/s/b593e8d7bfe7cd8500b1
دریافت: ۱۶ آگوست ۲۰۲۳ | پذیرش: ۲۶ آوریل
Published online: 18 May 2024

منابع

۱. ایتیسون، جی اچ، زانگ، جی، دلینگ، اف ان و دیو، آر سی. پروفایل های خودکار و قابل تفسیر نوار قلب بیمار برای تشخیص، ردابی و کشف بیماری دور قلب و عروق. *کیفیت نتایج* ۱۲، (۰۵) ۲۰۱۹.

۲. آتیا، زی و همکاران. تخمین سن و جنس با استفاده از هوش مصنوعی از نوار قلب استاندارد ۱۲ لید. *سیکل. آریتمی، الکتروفیزیولوژی*. ۱۲، (۰۷) ۲۰۱۹.

۳. آتیا، زی و همکاران. غربالگری اختلال انقباض قلب با استفاده از الکتروکاردیوگرام مجهر به هوش مصنوعی. *نات. مد. ۲۵*، ۷۴-۷۰. (۰۶) 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، ۸۶۱-۸۷۶. (۰۶) ۲۰۱۹.

۵. کوون، جی-ام، و همکاران. مقایسه عملکرد هوش مصنوعی و معیارهای تشخیص مرسموپرای تشخیص هیپوتروفی بطن چپ با استفاده از الکتروکاردیوگرافی. *اروپا* ۱۲، ۴۱۹-۴۲۰. (۰۲) ۲۰۲۰.

۶. سریتیح، م.، جباره، ی.، و حجام الحسنی، ا. یک سیستم خودکار برای تشخیص آریتمی ECG با استفاده از تکنیک های یادگیری ماشین. *مجله تحقیقات پزشکی جن. کلینی*. ۱، (۰۲) ۴۵۴-۴۱۱. (۰۲) ۲۰۲۱.

۷. گوستافسون، اس. و همکاران. توسعه و اعتبارسنجی پیش بینی سکته قلبی مبتنی بر ECG با یادگیری عمیق در بیماران بخش اورژانس. *نماینده علمی* ۱۲، ۱۹۶۱۵. (۰۲) ۲۰۲۲.

۸. او، ال. و همکاران. شبکه های یادگیری عمیق، انفارکتوس میوکارد با بالا رفتگ قطعه ST و RG مقصیر را به طور دقیق تشخیص می دهند. *قلب و عروق جلویی. پزشکی*. ۹، (۰۲) ۷۹۷۲۰-۷۹۷۲۱.

